Simulations of Polymer Translocation

Design of the cover: Jelle Henze
ISBN: 978-90-393-4852-9

Simulations of Polymer Translocation

Simulaties van de Translocatie van Polymeren

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. J. C. Stoof, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 9 juli 2008 des middags te 2.30 uur

door

Henk Vocks

geboren op 1 april 1979, te Venlo

Promotor: Prof. dr. G. T. Barkema

Contents

Chapter 1

Polymer physics and introduction to polymer translocation

Abstract

This chapter will acquaint the reader with several properties of long, flexible polymers. A short introduction to polymer physics is necessary to appreciate many of the results presented in this thesis. No experimental techniques will be discussed. Well-known introductory texts on polymer physics are "Introduction to Polymer Physics" by Masao Doi [1] and "Scaling Concepts in Polymer Physics" by Pierre-Gilles de Gennes [2], and this introduction closely follows presentations found in both. The latter part of this chapter provides an introduction to polymer translocation.

As in De Gennes' book, all instances of the Boltzmann constant k_{B} are removed. From the onset we will explicitly work in three dimensions only.

1.1 Static properties of single chains

1.1.1 A single ideal chain

Perhaps the simplest (and crudest) model for a flexible polymer is depicted in figure 1.1. The polymer lives on a periodic lattice, with lattice distance λ. On a random site we put the "head"-monomer, with number 0. Consecutive

monomers are randomly chosen to be one of the z nearest-neighbors. Hence this model allows multiple occupancy of sites. The N^{th} monomer, the "tail", completes the random walk. A more advanced lattice polymer model, detailed in the next chapter, forms the basis for all numerical results presented in this thesis.

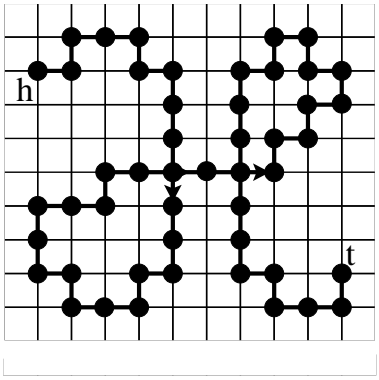

Figure 1.1: Square lattice with a random walk. The zeroth monomer is denoted "h", the N^{th} "t". The lattice coordination number z is 4.

The total number of distinct polymers with degree of polymerization $(N+1)$, all starting at the origin, is easily shown to be

$$Z_{\text{ideal}} = z^N. \tag{1.1}$$

Since the end-to-end vector

$$\mathbf{r} = \mathbf{r}_{\text{h}} - \mathbf{r}_{\text{t}} = \sum_{n=0}^{N-1} (\mathbf{r}_{n+1} - \mathbf{r}_n) = \sum_{n=0}^{N-1} \mathbf{v}_n \tag{1.2}$$

is just a sum of randomly chosen bond vectors, we can immediately write down for the mean squared end-to-end distance

$$\langle \mathbf{r}^2 \rangle = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \langle \mathbf{v}_n \cdot \mathbf{v}_m \rangle = \sum_{n=0}^{N-1} \langle \mathbf{v}_n^2 \rangle = N\lambda^2 \equiv R_0^2. \tag{1.3}$$

So the mean size of the polymer is proportional to $N^{1/2}$.

We can also calculate the probability distribution function of $\mathbf{r}$. $P(\mathbf{r}, N)$ is the probability to find monomer N at position $\mathbf{r}$, given that its head is located at the origin. From figure 1.1 it is easy to see that

$$P(\mathbf{r}, N) = \frac{1}{z} \sum_{i=1}^{z} P(\mathbf{r} - \mathbf{b}_i, N - 1), \tag{1.4}$$

with $\mathbf{b}_i$ the z different bond vectors of length λ. If the polymer is very long, $N \gg 1$, $|\mathbf{r}| \gg \lambda$, then the previous equation reduces to

$$\frac{\partial P}{\partial N} = \frac{\lambda^2}{6} \frac{\partial^2 P}{\partial \mathbf{r}^2}, \tag{1.5}$$

which may be solved to yield

$$P(\mathbf{r}, N) = \left(\frac{3}{2\pi N \lambda^2} \right)^{3/2} \exp\left(-\frac{3\mathbf{r}^2}{2N\lambda^2} \right). \tag{1.6}$$

Eq. (1.6) is very convenient, since from it follows immediately an expression for the free energy of a chain at fixed elongation

$$F(\mathbf{r}) = -T \log\left(P(\mathbf{r}, N)\right) = F(0) + \frac{3T\mathbf{r}^2}{2R_0^2}, \tag{1.7}$$

where T is the temperature and $F(0)$ a constant independent of r. Note that there is no enthalpic contribution to this free energy. Without getting ahead of matters too much, later it will be shown that the "entropic spring constant", that one obtains by taking the second derivative of Eq. (1.7) with respect to r, is the foremost 'ingredient' to understanding polymer translocation.

Many models of polymers are known. The random walk on a square lattice of figure 1.1 is perhaps the simplest model. A more detailed description may include interactions along the backbone of the polymer, such that $\langle \mathbf{v}_n \cdot \mathbf{v}_m \rangle \neq 0$ for $m \neq n$. If these interactions extend only to bonds up to a finite distance along the chain, then $\langle \mathbf{v}_n \cdot \mathbf{v}_m \rangle$ will decay exponentially for large $|n - m|$. For such systems the mean squared end-to-end distance is always proportional to N for large N, and the distribution of $\mathbf{r}$ is Gaussian. Models with such short-ranged interactions are in a sense equivalent to the random walk model and are called ideal.

1.1.2 A single nonideal chain

Many polymer models are ideal, despite short-ranged interactions along the backbone. But with short-ranged interactions only, nothing is preventing those monomers that are far apart along the chain from occupying the same region in space. In reality, no chain can ever intersect itself. This "excluded-volume effect" is incorporated in a so-called "self-avoiding walk" (SAW), simply defined as a path that never intersects itself.

The total number of SAWs of N steps is asymptotically of the form

$$Z_{\mathrm{SAW}} = \alpha \hat{z}^N N^{\gamma-1}, \tag{1.8}$$

with α a constant, and $\hat{z}^N$ serving the same purpose as z^N in Eq. (1.1), but now with $\hat{z}$ numerically somewhat smaller than the coordination number. The last factor, $N^{\gamma-1}$ is called the enhancement factor, with exponent γ, which in three dimensions is numerically known to be $\gamma = 1.1575 \pm 0.0006$ [3].

An ideal chain, with mean size $R_0 \sim N^{1/2}$, will on average be smaller than an excluded-volume chain, for which $R_{\mathrm{F}} \sim N^{\nu}$, with $\nu > 0.5$. This swelling effect was predicted by Flory [4] first, who used simple arguments to obtain a remarkably accurate estimate, $\nu \approx 0.6$. Today's best estimates are around $\nu = 0.5877 \pm 0.0006$ [5].

Flory's argument proceeds as follows. A chain of length N, with a (yet to be determined) radius R will have an internal monomer concentration

$$c_{\mathrm{int}} \simeq \frac{N}{R^3}. \tag{1.9}$$

For the monomer-monomer repulsive interaction we write, per unit volume,

$$F_{\mathrm{rep}} = \frac{1}{2} T v c^2, \tag{1.10}$$

with v in units of volume, and c the local concentration of monomers. Next, one makes a mean-field approximation, which reads

$$\langle c^2 \rangle \longrightarrow \langle c \rangle^2 \sim c_{\mathrm{int}}^2. \tag{1.11}$$

Since all correlations between monomers are ignored, we write for the total repulsive energy in volume R^3

$$F_{\mathrm{rep}|tot} \simeq \frac{1}{2} T v c_{\mathrm{int}}^2 R^3 = \frac{1}{2} T v \frac{N^2}{R^3}. \tag{1.12}$$

The repulsive term causes the polymer to swell, but certainly not without limits. As already observed in Eq. (1.7), as the polymer increases its size, its entropy is reduced. Therefore Flory added an elastic energy contribution

$$\frac{F}{T} \simeq \frac{1}{2} v \frac{N^2}{R^3} + \frac{3}{2} \frac{R^2}{N\lambda^2}. \tag{1.13}$$

The size of the excluded-volume chain is obtained by minimizing the free energy, to find

$$R_{\mathrm{F}} \sim N^{\nu}, \tag{1.14}$$

with $\nu = 3/5$ the Flory exponent.

1.2 Dynamic properties of single chains

1.2.1 The Rouse model

A polymer suspended in a dilute solution is in perpetual movement. In 1953 P. E. Rouse presented a simple model with which a single chain's relaxation modes can be understood. The model has penetrated every aspect of polymer physics, which justifies a complete derivation here.

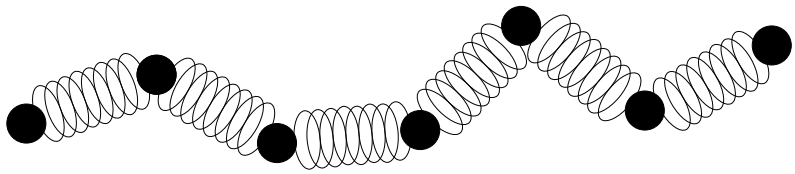

Figure 1.2: Bead-spring model with 7 beads and 6 springs.

The starting point is a bead-spring model of a polymer. In figure 1.2 beads are connected by springs. Physically, springs should be thought of as stretches of the polymer long enough to contain sufficient numbers of monomers for Gaussian statistics to apply. In Eq. (1.7) we derived an expression for the elastic energy of the spring connecting beads n and $n+1$,

$$F_{n,n+1} = \frac{3T}{2} \frac{(\mathbf{r}_{n+1} - \mathbf{r}_n)^2}{\langle (\mathbf{r}_{n+1} - \mathbf{r}_n)^2 \rangle} = \frac{3T}{2} \frac{(\mathbf{r}_{n+1} - \mathbf{r}_n)^2}{\lambda^2}, \tag{1.15}$$

where λ is the average extension of a spring. The polymer's total elastic energy simply reads

$$F_{\mathrm{el|tot}} = \sum_{n=0}^{N-1} F_{n,n+1}, \tag{1.16}$$

with N the number of beads.

Since the polymer is floating in a solution, we will assume the velocity of the beads to be proportional to the force exerted on them. To also incorporate thermal fluctuations, a Langevin equation of motion for the beads seems appropriate:

$$\frac{\mathrm{d}\mathbf{r}_n}{\mathrm{d}t} = -\frac{1}{\xi}\frac{\partial F_{\mathrm{el|tot}}}{\partial \mathbf{r}_n} + \mathbf{g}_n, \tag{1.17}$$

with $\mathbf{g}_n$ a random force with zero mean and variance given by

$$\langle g_{n\alpha}(t) g_{m\beta}(t')\rangle = 2\delta_{nm}\delta_{\alpha\beta}\frac{T}{\xi}\delta(t-t'), \tag{1.18}$$

with ξ the friction coefficient of a bead. Together with Eqs. (1.15) and (1.16) the equations governing the dynamics of the polymer become, for $n = 0, 1, \cdots, N$,

$$\frac{\mathrm{d}\mathbf{r}_n}{\mathrm{d}t} = \frac{3T}{\xi\lambda^2}(\mathbf{r}_{n+1} + \mathbf{r}_{n-1} - 2\mathbf{r}_n) + \mathbf{g}_n, \tag{1.19}$$

together with $\mathbf{r}_{-1} \equiv \mathbf{r}_0$ and $\mathbf{r}_{N+1} \equiv \mathbf{r}_N$.

To solve this Rouse model, the conventional approach is to let n be a continuous variable: $\mathbf{r}_n(t)$ becomes $\mathbf{r}(n,t)$, and Eq. (1.19) turns into

$$\frac{\partial \mathbf{r}}{\partial n} = \frac{3T}{\xi\lambda^2}\frac{\partial^2 \mathbf{r}}{\partial n^2} + \mathbf{g}(n,t), \tag{1.20}$$

together with the boundary conditions $\partial\mathbf{r}/\partial n = 0$ at $n = 0$ and $n = N$. Normalized coordinates

$$\mathbf{x}_p(t) = \frac{1}{N}\int_0^N \mathrm{d}n \cos\left(\frac{p\pi n}{N}\right)\mathbf{r}(n,t), \tag{1.21}$$

with $p = 0, 1, 2, \cdots$, decompose the polymer's motion into separate modes

$$\frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t} = -\frac{k_p}{\xi_p}\mathbf{x}_p + \mathbf{g}_p. \tag{1.22}$$

Here

$$\xi_0 = N\xi \, , \, \xi_p = 2N\xi \, , \, k_p = \frac{6T\pi^2}{N\lambda^2} p^2. \tag{1.23}$$

The random force $\mathbf{g}_p(t)$ has a variance given by

$$\langle g_{p\alpha}(t) g_{q\beta(t')} \rangle = 2\delta_{pq}\delta_{\alpha\beta}\frac{T}{\xi_p}\delta(t-t'), \tag{1.24}$$

while the normalized coordinates can be shown to obey

$$\langle (\mathbf{x}_0(t) - \mathbf{x}_0(0))_\alpha (\mathbf{x}_0(t) - \mathbf{x}_0(0))_\beta \rangle = \delta_{\alpha\beta}\frac{2T}{\xi_0}t, \tag{1.25}$$

$$\langle \mathbf{x}_{p\alpha}(t)\mathbf{x}_{q\beta}(0)\rangle = \delta_{pq}\delta_{\alpha\beta}\frac{T}{k_p}\exp\left(-t/\tau_p\right), \tag{1.26}$$

with

$$\tau_p = \frac{\xi_p}{k_p} = \frac{\tau_1}{p^2} = \frac{1}{p^2}\frac{\xi N^2\lambda^2}{3\pi^2 T}. \tag{1.27}$$

Observe that in the Rouse model the longest relaxation time $\tau_1 \equiv \tau_{\mathrm{R}}$ scales as N^2.

With the solution of the Rouse model at hand, let us investigate some of its consequences. Firstly, the position of the center of mass is given by the first normal coordinate $\mathbf{x}_0(t)$. Therefore, we find for the average squared displacement of the center of mass

$$\langle (\mathbf{r}_{\mathrm{CM}}(t) - \mathbf{r}_{\mathrm{CM}}(0))^2 \rangle = 3\frac{2T}{\xi_0}t = \frac{6T}{N\xi}t = 6D_{\mathrm{R}}t. \tag{1.28}$$

The polymer's diffusion constant D_{R} is inversely proportional to the degree of polymerization N.

Next we investigate the average squared displacement of the nth bead as a function of time

$$\phi(n,t) \equiv \langle (\mathbf{r}(n,t) - \mathbf{r}(n,0))^2 \rangle. \tag{1.29}$$

By expressing $\mathbf{r}(n,t)$ in terms of the normalized coordinates $\mathbf{x}_p(t)$, one easily obtains

$$\phi(n,t) = 6D_{\mathrm{R}}t + \frac{4N\lambda^2}{\pi^2}\sum_{p=1}^{\infty}\cos^2(\frac{p\pi n}{N})\frac{1}{p^2}(1-\exp\left(-tp^2/\tau_{\mathrm{R}}\right)). \tag{1.30}$$

For $t \gg \tau_{\mathrm{R}}$, the averaged squared displacement is $6D_{\mathrm{R}}t$ + constant$\cdot N$. On the other hand, for $t \ll \tau_{\mathrm{R}}$,

$$\phi(n,t) \approx \frac{2}{\pi^{3/2}} N\lambda^2 (\frac{t}{\tau_{\mathrm{R}}})^{1/2}, \tag{1.31}$$

i.e., at short times, the averaged squared displacement increases as $t^{1/2}$.

1.2.2 Excluded-volume effects and the Zimm model

As we have seen above, the average size of an ideal chain scales as $N^{1/2}$. But real polymers are not ideal, with excluded-volume effects which cause the chain to swell to N^{ν}. The Rouse model was derived with an ideal chain in mind. Consequently, $\tau_{\mathrm{R}} \sim N^2$. If excluded-volume effects are taken into account, the longest relaxation time can be shown to increase to $\tau_{\mathrm{R}} \sim N^{1+2\nu}$. In short, as the diffusion of a polymer is proportional to N^{-1}, the time needed to diffuse over its own length is proportional to $N^{2\nu}/N^{-1} = N^{1+2\nu}$.

Experimentally, in contrast to Rouse's result (1.28), the diffusion constant is proportional to $N^{-\nu}$. It is well-known that the discrepancy is due to Rouse's neglect of hydrodynamic interaction effects: whenever a force $\mathbf{f}_n$ is applied to a bead, the bead's resulting motion will introduce a velocity field in the fluid which decreases in strength with distance as $(\mathbf{r}-\mathbf{r}_n)^{-1}$. Hydrodynamic effects were added to the equations of motion of the bead spring model by Zimm, who confirmed the measured diffusion constant. Moreover, the chain's longest relaxation time turns out to be $\tau_{\mathrm{d}} \sim N^{3\nu}$.

Computer simulation of polymers with hydrodynamic interactions are becoming more common nowadays [6, 7, 8, 9, 10]. Still the additional computational overhead is prohibitive for most purposes. As we refrain from their inclusion, the longest relaxation time of single unperturbed chains, in our simulations, is proportional to $\tau_{\mathrm{R}} \sim N^{1+2\nu}$.

1.3 Polymer melts

1.3.1 Molten chains are ideal

In very dilute polymer solutions, in which the separate chains are not in contact with each other, we have seen that the excluded-volume interactions cause the chains to swell. Their average size increases with N as N^{ν}. In a polymer melt or in a highly concentrated solution of polymers we are looking at the

opposite limit: each and every chain is in contact with many others. It was Flory who came to understand first that these highly interacting polymers are ideal: the average size increases as $R_0 \sim N^{1/2}$ and the chains obey Gaussian statistics. To instill the reader with some qualitative understanding, without a long derivation (for which we refer the reader to the excellent derivation in Doi and Edwards [11]), we repeat De Gennes' self-consistent-field arguments.

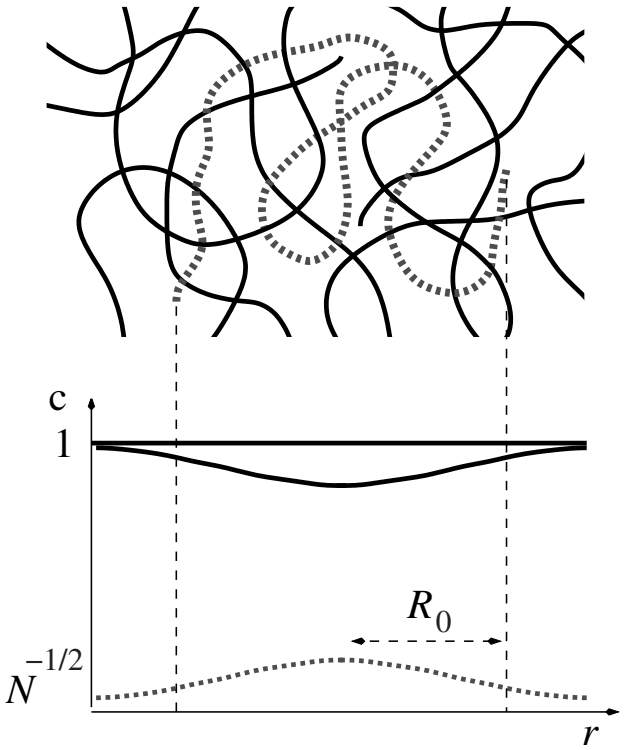

Figure 1.3: Top: representation of a polymer melt. Bottom: concentration of light and dark monomers.

In figure 1.3 a single dashed "light chain" is located in a dense melt of equivalent "dark chains". The total monomer concentration c is 1. A light monomer experiences a repulsive potential U_{tot} essentially proportional to the local monomer density c, to which there are two parts. Firstly, the concen-

tration of light monomers is peaked around the center of gravity of the light chain. The excluded-volume interactions of the dashed monomers create an effective potential $U_{\rm l}$, with outward pointing force $-\frac{\partial U_{\rm l}}{\partial r}$. In the single-chain problem, it is this force that causes the polymer to swell and show nonideal behavior.

On the other hand it is known that the density fluctuations in a melt are very weak. Therefore, around the location of the light chain, the concentration of dark monomers must be lowered. The resulting density gradient leads to an inward pointing force $-\frac{\partial U_{\rm d}}{\partial r}$.

In equilibrium the total force on any monomer is zero. We conclude $\frac{\partial U_{\rm tot}}{\partial r} = 0$: the force responsible for the swelling of the single polymer is canceled and the chain remains ideal.

Note that the local concentration of light monomers is very low, scaling as $N/R_0^3 \sim N^{-1/2}$, with R_0 the mean size of an ideal chain (from Eq. (1.3)), but still much higher than the concentration $c^* \approx N/R_{\rm F}^3 \sim N^{-0.8}$, with $R_{\rm F}$ the mean size of an excluded volume chain (from Eq. (1.14)), needed for polymers in dilute solution to start to overlap.

1.3.2 Reptation theory

The viscosity and self-diffusion of concentrated polymer systems depend very strongly on chain length. Experimentally, for chains long enough to be entangled, it is well-known that

$$\eta \sim N^{m_\eta}, \tag{1.32}$$

where m_η is of order 3.4 ± 0.2. For the diffusion constant it is reported that $D_{\rm d} \sim N^{-\delta}$, with δ at least 2, although the exact value of δ is shrouded in controversy.

A chain in a dense polymeric system moves in a network composed of the surrounding polymers. For chains long enough ($N \gtrsim 10^2$), entanglement effects come into play. These effects can be thought of as the topological restrictions on the movement of a polymer in a matrix of other polymers. Theoretically, It was only with the introduction of the reptation concept by De Gennes in 1971, that a foothold was gained [12]. As a first step, De Gennes considered the Brownian motion of a polymer in a fixed network, as in figure 1.4. The dots, which represent the network, are impenetrable. It is possible to use the Rouse model to describe the motion of the polymer in the network, and to express the viscosity and self-diffusion in terms of the parameters that enter

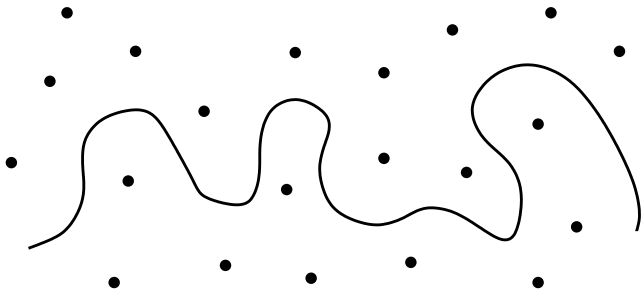

Figure 1.4: Chain in a static network of dots.

the Rouse model, i.e., degree of polymerization N, friction coefficient ξ and average monomer-monomer distance λ, augmented with the average separation of the dots a. We already know from Eq. (1.31), that at very short times the average squared displacement of a monomer increases proportionately to $t^{1/2}$, since the polymer has not been given enough time to notice the surrounding polymers. On very long time scales, when the polymer has diffused over its own length and any correlations with its conformation at $t = 0$ have vanished, ordinary diffusion applies, i.e., $\phi(n, t) \sim t$. De Gennes' reptation theory deals with the intermediate times.

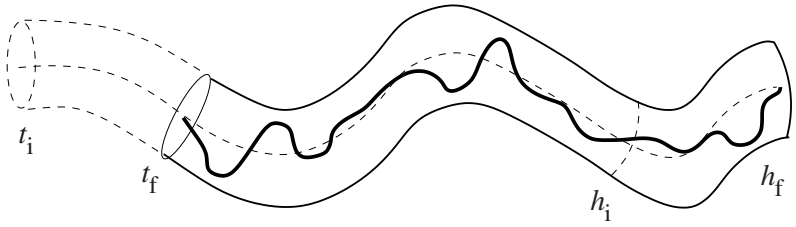

Figure 1.5: The motion of the tube. The dotted line through the center represents the primitive chain, while the thick line represents the polymer chain.

De Gennes used the tube concept, introduced earlier by Edwards, to argue that the dots effectively cause the polymer to move in a tube with radius a. The polymer is significantly longer than the tube, and the excess length along the primitive path allows the chain to diffuse away from its initial constraints. Due to this motion the tube itself changes with time: in figure 1.5, the vacated

part t_i to t_f disappears, while h_i to h_f appears in a newly created section. This snake-like motion bears the name reptation.

In actual concentrated solutions and melts the confining dots are polymers themselves, which move in a similar fashion as the test chain. Consequently, the tube is not a static object. The mechanism of Constraint Release (CR) describes the deterioration of the tube walls with time. Other interactions between the chains are ignored as well. Nevertheless, for highly entangled systems, it is widely accepted that the reptating motion of the chain is the dominant contribution to the dynamics.

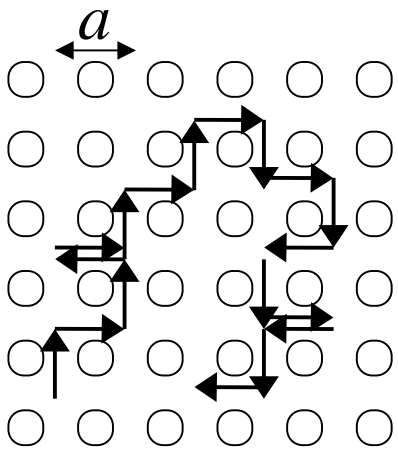

Figure 1.6: Doi's lattice model of reptation [1].

We will follow Doi's lattice model of reptation [1], not only because of its clarity, but also because it connects rather straightforwardly to Rubinstein's historically important Repton model [13], as well as to our Slithering Snake algorithm with which we have generated large equilibrated densely packed polymer samples (chapter 2.5). As in figure 1.6, a chain with Z bonds and $Z+1$ monomers is placed on a lattice with lattice spacing λ and lattice coordination z. The bond vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_Z$ are randomly chosen from the z unit vectors. Every time interval Δt we let the whole chain reptate along one lattice spacing, in the direction of either the head or the tail. All bond vectors $\mathbf{u}_n$ are renamed $\mathbf{u}_{n\pm1}$, with $n < Z$ respectively $n > 1$, and $\mathbf{u}_1$ resp. $\mathbf{u}_Z$ is

taken in a random direction. In short

$$\mathbf{u}_n(t+\Delta t) = \mathbf{u}_{n+\sigma(t)}(t), \tag{1.33}$$

with $\sigma(t)$ a random variable $+1$ or -1, and $\mathbf{u}_0$ or $\mathbf{u}_{Z+1}$ randomly chosen. Therefore, the equations describing the coordinates of the monomers $\mathbf{r}_n$ ($n = 0, 1, \cdots, Z$) are

$$\mathbf{r}_n(t+\Delta t) = \mathbf{r}_{n+\sigma(t)}(t), \tag{1.34}$$

with $\mathbf{r}_{-1}(t+\Delta t) = \mathbf{r}_0(t) - a\boldsymbol{\rho}(t)$, $\mathbf{r}_{Z+1}(t+\Delta t) = \mathbf{r}_Z(t) + a\boldsymbol{\rho}(t)$, and random unit vector $\boldsymbol{\rho}(t)$.

Δt and Z can be expressed in the parameters that appear in the Rouse model. The Z bonds form an ideal chain with a mean squared end-to-end distance Za^2. For a Rouse chain, the end-to-end distance is $N\lambda^2$, which tells us that

$$Z = N\lambda^2/a^2. \tag{1.35}$$

Δt is the time required for the chain to diffuse a distance a along its own contour. For a Rouse chain the diffusion constant, $D_{\mathrm{R}} = T/N\xi$, expresses the displacement of the center of mass. Since the polymer can only go forward or backward along its primitive path, we obtain for this "curvilinear displacement" $\langle s^2(t)\rangle = 2D_{\mathrm{R}}t$, where the factor of 2 indicates one-dimensional motion (cf. Eq. (1.28)). In reptation theory it is assumed that interactions with the tube wall may be included by using an effective friction constant ξ'. Within the time t the reptating chain makes $t/\Delta t$ jumps, each over a distance a. Hence $\langle s^2(t)\rangle = (t/\Delta t)a^2$, and

$$\Delta t = \frac{a^2}{2D_{\mathrm{R}}} = \frac{a^2 N\xi'}{2T}. \tag{1.36}$$

Let us first find the self-diffusion constant of the chain. For the center of mass

$$\mathbf{r}_{\mathrm{CM}} = \frac{1}{Z+1}\sum_{n=0}^{Z}\mathbf{r}_n(t) \tag{1.37}$$

simple substitution suffices to show that

$$\mathbf{r}_{\mathrm{CM}}(t+\Delta t) = \mathbf{r}_{\mathrm{CM}}(t) + \sigma(t)\mathbf{f}(t), \tag{1.38}$$

with

$$\mathbf{f}(t) = \frac{\mathbf{P}(t) + a\boldsymbol{\rho}(t)}{Z+1}, \tag{1.39}$$

and the end-to-end vector $\mathbf{P}(t) = \mathbf{r}_Z(t) - \mathbf{r}_0(t)$. Thus

$$\langle(\mathbf{r}_{\rm CM}(t) - \mathbf{r}_{\rm CM}(0))^2\rangle = \frac{t}{\Delta t}\langle f(t)^2\rangle. \tag{1.40}$$

But in equilibrium $\langle \mathbf{P}^2(t)\rangle = Za^2$, so that we find, for $Z \gg 1$,

$$\langle(\mathbf{r}_{\rm CM}(t) - \mathbf{r}_{\rm CM})^2\rangle = \frac{t}{\Delta t}\frac{a^2}{Z} = \frac{2D_{\rm R}}{Z}t. \tag{1.41}$$

Hence we find for the self-diffusion constant of this reptation model

$$D_{\rm d} = \frac{D_{\rm R}}{3Z} = \frac{T}{3N^2\xi'}\left(\frac{a^2}{\lambda^2}\right). \tag{1.42}$$

For this model of reptation, the diffusion constant of the center of mass is proportional to N^{-2}.

Similarly, the correlation function $\langle \mathbf{P}(t)\cdot\mathbf{P}(0)\rangle$ of the end-to-end vector $\mathbf{P}(t)$ can be expressed as

$$\langle \mathbf{P}(t)\cdot\mathbf{P}(0)\rangle = a^2\sum_{n=1}^{Z}\sum_{m=1}^{Z}\psi_{n,m}(t), \tag{1.43}$$

with $\psi_{n,m}(t) \equiv \langle \mathbf{u}_n(t)\cdot\mathbf{u}_m(0)\rangle$. After a time Δt, $\mathbf{u}_n(t)$ changes randomly to $\mathbf{u}_{n+1}(t)$ or $\mathbf{u}_{n-1}(t)$. Therefore, $\psi_{n,m}(t)$ satisfies the equation

$$\psi_{n,m}(t+\Delta t) = \frac{1}{2}\left[\psi_{n+1,m}(t) + \psi_{n-1,m}(t)\right], \tag{1.44}$$

with boundary conditions $\psi_{0,m}(t) = 0$, $\psi_{Z+1,m}(t) = 0$, and at $t = 0$ we have $\psi_{n,m}(0) = \delta_{nm}$. To solve the equation, in the limit $Z \gg 1$, the difference equation may be rewritten as

$$\frac{\partial\psi_{n,m}}{\partial t} = \frac{D_{\rm R}}{a^2}\frac{\partial^2\psi_{n,m}}{\partial n^2}, \tag{1.45}$$

with appropriate boundary and initial conditions $\psi_{0,m}(t) = \psi_{Z,m}(t) = 0$ and $\psi_{n,m}(0) = \delta(n-m)$. The solution is

$$\psi_{n,m}(t) = \frac{2}{Z}\sum_{p=1}^{\infty}\sin\left(\frac{np\pi}{Z}\right)\sin\left(\frac{mp\pi}{Z}\right)\exp\left(-\frac{tp^2}{\tau_{\rm d}}\right). \tag{1.46}$$

Here we have defined the reptation time $\tau_{\rm d} = \frac{Z^2 a^2}{\pi^2 D_{\rm R}} = \frac{1}{\pi^2}\frac{\xi' N^3 \lambda^4}{T a^2}$. We finally find

$$\langle \mathbf{P}(t) \cdot \mathbf{P}(0) \rangle = Z a^2 \psi(t), \tag{1.47}$$

with

$$\psi(t) = \sum_{p=\text{odd}} \frac{8}{p^2 \pi^2} \exp\left(-\frac{t p^2}{\tau_{\rm d}}\right). \tag{1.48}$$

It is easily seen that $\psi(0) = 1$, and it decays with "rotational" relaxation time proportional to N^3. Physically, $\psi(t)$ can be interpreted as the fraction of the polymer that is still in the tube after time t, while $\tau_{\rm d}$, also often referred to as the disentanglement time, is a measure of the time needed for the polymer to vacate the initial tube.

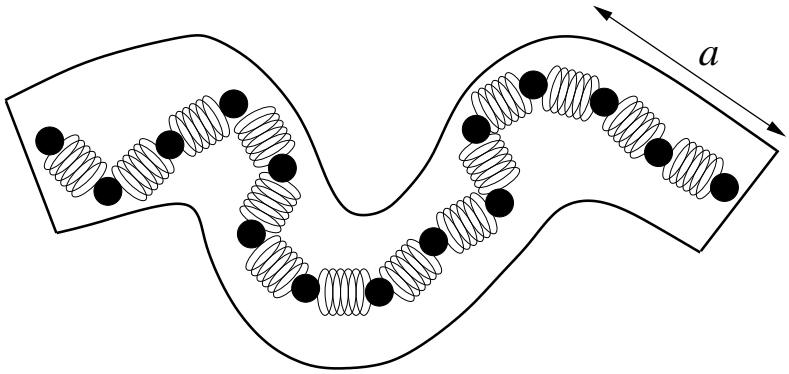

Figure 1.7: A bead-spring chain in a tube.

A different approach to reptation theory is by starting from a Rouse chain trapped in a tube, as in figure 1.7. Our aim here is to study the mean squared displacement of a single bead n. As noted at the beginning of the discussion, $\phi(n,t) \sim t^{1/2}$ for times that are too short for the polymer to become aware of the topological restrictions that constitute the confining tube. It is important to realize that the tube represents topological restrictions, more than just excluded-volume interactions. The tube is not empty, but contains many monomers from many different chains.

We define the entanglement time $\tau_{\rm e}$ as the time needed for the polymer to experience the restrictions posed by the tube wall. For $t > \tau_{\rm e}$, the diffusion will be reduced to curvilinear motion along the tube's primitive path. We

have seen that the mean size of a polymer in the melt is proportional to $N^{1/2}$. Therefore, in real-space coordinates, the mean squared displacement will cross over to $\phi(n,t) \sim (t^{1/2})^{1/2} = t^{1/4}$, for $t > \tau_e$.

For the single-chain Rouse model, the mean squared displacement of the single bead n crosses over from $t^{1/2}$ to t at $t \approx \tau_R$. Here, let us assume that also in the tube the polymer's longest relaxation time is given by τ_R, albeit with a modified friction coefficient ξ' to account for the presence of the surrounding monomers. For times $t > \tau_R$ the motion is still restricted to the curvilinear path. Therefore, we expect a crossover to $\phi(n,t) \sim t^{1/2}$ at $t \approx \tau_R$.

Finally, as is illustrated in figure 1.8, we can estimate the disentanglement time τ_d from retrieving the time needed for a monomer to diffuse over the length of a molten polymer coil $R_0 \sim N^{1/2}$. Hence $\phi(n,t) \sim t$, for $t > \tau_d$.

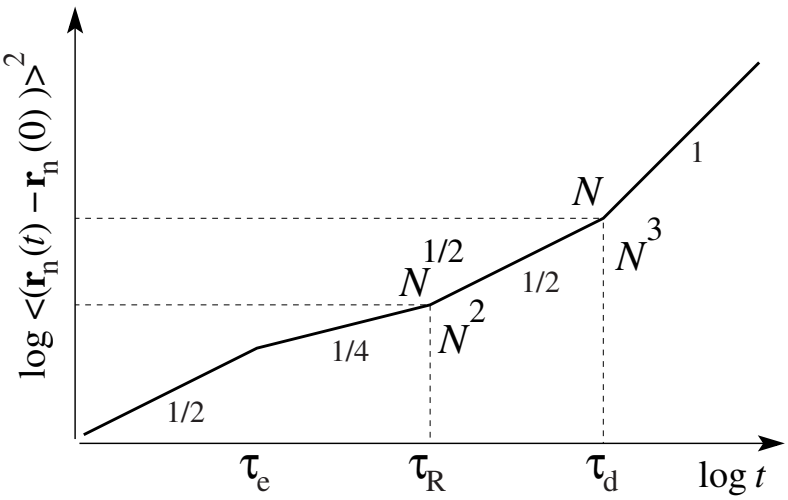

Figure 1.8: Mean squared displacement of monomer n plotted against time, on a double-logarithmic plot, for well-entangled polymers.

Conventional reptation theory concludes $\eta \sim N^{m_\eta}$, where $m_\eta = 3$. It also predicts a diffusion constant $D_d \sim N^{-2}$. Experimentally, for strongly entangled solutions and melts, $m_\eta = 3.4 \pm 0.2$ and $D_d \sim N^{-\delta}$ with δ at least 2. A compilation of data by Lodge [14] suggests $\delta = 2.28 \pm 0.05$, although many authors measure $\delta = 2$. Many authors have tried to bridge the gaps. According to Doi [15], standard reptation theory wrongly assumes the length

of the polymer along the tube to be fixed. By including these Contour Length Fluctuations (CLF) Doi obtained for the disentanglement time

$$\tau_{\mathrm{d}} \sim Z^3 \left(1 - \frac{C}{\sqrt{Z}}\right)^2, \tag{1.49}$$

with C a numerical constant of approximately 1.5 [16, 17]. For $10 < Z < 100$, $\tau_{\mathrm{d}} \sim Z^{3.4}$. Many authors also consider Constraint Release to be a mechanism that contributes to the disagreement [18, 19]: the surrounding chains that constitute the entanglements are free to diffuse and relax stress themselves. Hence the tube erodes with a finite rate anywhere along its length. But, to quote De Gennes [2]: "tube renewal is not observable" for longer chains in dense systems, although Wang [20] concludes that CR may explain a rapid crossover from $\delta = 2.4$ to $\delta = 2$. Several authors report measurements that support $\tau_{\mathrm{d}} \sim N^3$ and $D_{\mathrm{d}} \sim N^{-2}$, for $N \to \infty$ [21, 20]. Alternative explanations of the 3.4 power-law have been put forward as well, but none so far are considered satisfactory.

Recently, Liu *et al.* [22] suggested that multi-chain effects, i.e., CR, are of similar importance as the single chain effects, i.e., CLF. Promoting the multi-chain effects, not just CR, is not at odds with new simulation results by Barkema and Panja [23]. They have presented highly accurate numerical simulations for dense polymeric systems with next chapter's polymer model. The accompanying theoretical view also has significant consequences for polymer translocation in a dense environment, which will be the topic of chapter 4. Without getting ahead of matters too much, the interchain interactions leads to the introduction of a new time scale $\sim N^{\beta}$, with $\beta = 2.56 \pm 0.10$, that replaces the Rouse time scale in figure 1.8. Hence Barkema and Panja find[1] $\tau_{\mathrm{d}} \sim N^{3.28 \pm 0.05}$.

[1]The scaling with N of the longest relaxation time scale $\tau_\beta \sim N^\beta$ implies a rescaling of a (limited) number of modes with index p to $\tau_p \sim p^\beta$ (cf. Eq. (1.27)). Consequently, the mean squared displacement of the n^{th} bead calculated in Eq. (1.30) for the Rouse model, may have to be replaced by $\phi(n, t \ll \tau_\beta) \sim \int_0^\infty \mathrm{d}p(1 - \exp(-tp^\beta/\tau_\beta))/p^2 \approx (t/\tau_\beta)^{1/\beta}$ in a dense environment. Therefore, following Doi and Edwards [11], chapter 6.4.4, the mean squared displacement of a bead in regimes i and ii obeys $\sim t^{1/\beta}$ resp. $\sim t^{1/2\beta}$, leading to a scaling with N of the disentanglement time $\tau_{\mathrm{d}} \sim N^{1+\beta}$, and thus exceeding the estimate $\tau_{\mathrm{d}} \sim N^{3.28 \pm 0.05}$ of [23].

1.4 Introduction to polymer translocation through a narrow pore

To facilitate the study of the translocation of a polymer through a narrow pore in a membrane, we introduce several time scales, such as the translocation time, dwell time and unthreading time, as well as their mutual relations. The left-hand side of figure 1.9 shows two cells, each of volume V, connected through a pore in the separating membrane. The translocation time τ_{t} is the mean time required for a polymer to either leave cell A or B completely. The dwell time τ_{d} is the typical time scale the polymer spends in the pore, while the unthreading time τ_{u} is the average time needed to vacate the pore starting from an equilibrated state in which the polymer is threaded halfway through the pore.

We will restrict ourselves to cases in which the pore is just wide enough to allow the passage of a single monomer at a time. Hence the monomer number in the pore s separates a threaded polymer in two segments of length $s-1 \approx s$ and $N-s$ on either side of the membrane. For an ideal chain on a lattice we derived the partition function, Eq. (1.1), $Z_{\mathrm{ideal}} \sim z^N$. Similarly, for the self-avoiding walk, Eq. (1.8), $Z_{\mathrm{SAW}} \sim \hat{z}^N N^{\gamma-1}$. However, to compose the partition sum for chains tethered to an impenetrable wall, we must necessarily exclude contributions from chains crossing the wall. For an ideal tethered chain the partition function is easily shown to be of form [30]

$$Z_{\mathrm{ideal}}^{\mathrm{tethered}} \sim z^N N^{-1/2}, \tag{1.50}$$

while for a self-avoiding walk it is known that

$$Z_{\mathrm{SAW}}^{\mathrm{tethered}} \sim \hat{z}^N N^{\gamma_1-1}, \tag{1.51}$$

with an adjusted exponent $\gamma_1 = 0.68$ [31].

Let us also assume that each monomer carries a charge q and that a voltage $2V$ is applied across the pore, as in the left part of figure 1.9. Then, we have for the partition sum of a threaded polymer

$$Z^{\mathrm{threaded}}(N,s) = Z^{\mathrm{teth.A}}(s)\, Z^{\mathrm{teth.B}}(N-s) \exp\left(+2qVs/T\right), \tag{1.52}$$

with s the monomer in the pore. From it we obtain the free energy of a threaded self-avoiding walk

$$\begin{aligned} F^{\mathrm{threaded}}(N,s) &= -T \log Z^{\mathrm{threaded}}(N,s) \\ &\sim T(1-\gamma_1)\log\left[s(N-s)\right] - 2qVs - T\left(N \log \hat{z}\right). \end{aligned} \tag{1.53}$$

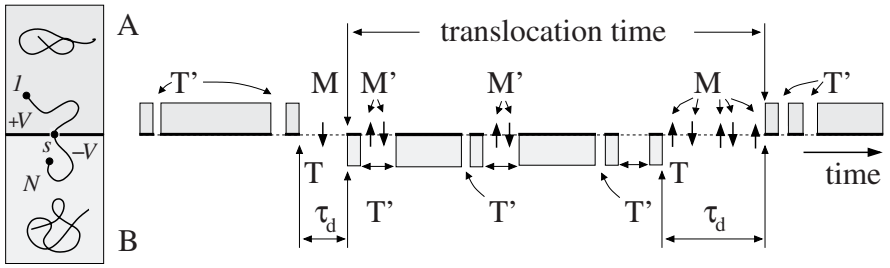

Figure 1.9: Left: two cells A and B, each of volume V, that are connected by a pore in a membrane. An applied voltage $2V$ may be present across the pore. A polymer of length N is located in the box, which can either be in cell A or B, or threaded in the pore. Right: a typical unbiased polymer translocation process, i.e., $2V = 0$.

Figure 1.10 shows the entropic barrier comprised of the first two terms of Eq. (1.54) for a polymer of length $N = 1000$ and for several values of $2qV$. Early theories of polymer translocation adopted a mean-field type approach to explain the scaling of the time scales of translocation with degree of polymerization, vesicle volume and other relevant parameters [24, 25, 26, 27, 28, 29, 30]. In these theories polymer translocation is described by a Fokker-Planck equation for first passage over the entropic barrier in terms of the reaction coordinate s. For unbiased polymer translocation, for ideal polymers, the mean-field theories predict a mean dwell time scaling as N^3 for Rouse dynamics, and as $N^{2.5}$ for Zimm dynamics [30]. However, the use of the (equilibrium) free energy to determine the transition rates from s to $s \pm 1$ implicitly assumes that at a fixed reaction coordinate s the polymer equilibrates faster than the typical time for the reaction coordinate to change its value by ± 1. This is not necessarily the case for longer polymers, or polymers in higher spatial dimensions. The authors of Ref. [32, 33, 34] found that for a self-avoiding polymer, with Rouse dynamics, the dwell time for unbiased translocation scales as the Rouse time τ_{R}. They concluded that the dynamics of a translocating polymer is anomalous. Their scaling of the dwell time was later confirmed in Ref. [35, 36]. Klein Wolterink *et al.* [37] found, with simulations on the extended repton model (chapter 2), that $\tau_{\mathrm{d}} \sim N^{2.40 \pm 0.05}$, while Dubbeldam *et al.* [38, 39] modeled unbiased translocation with a fractional Fokker-Planck equation, leading to a scaling $\tau_{\mathrm{d}} \sim N^{2.52 \pm 0.04}$ after some corrections [40]. Only recently, Panja *et al.* [41] have characterized the anomalous dynamics of translocation and its

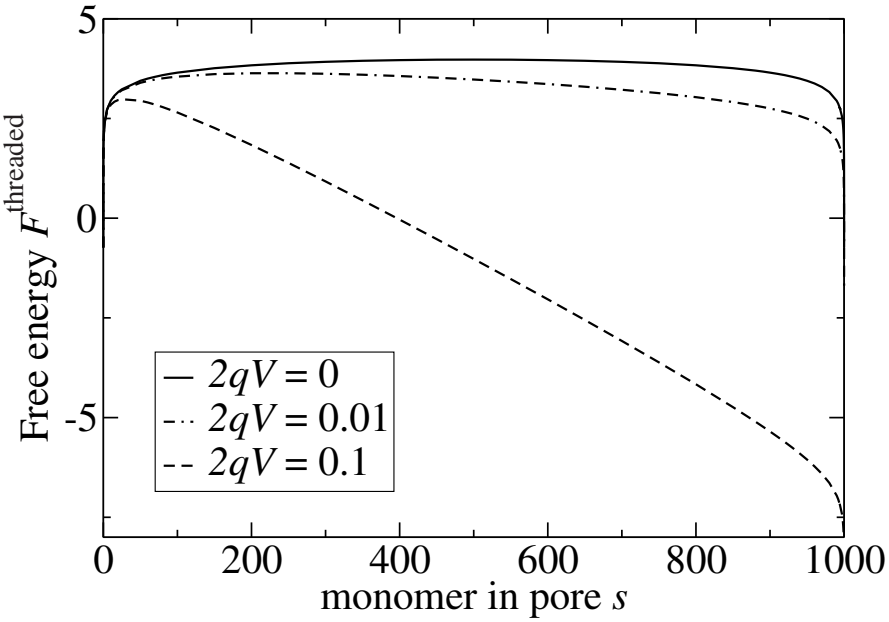

Figure 1.10: Free energy $F^{\text{threaded}}(N = 1000, s)$ for $2qV = 0$, 0.01, and 0.1.

underlying physics to derive the scaling $\tau_{\text{d}} \sim N^{2+\nu}$.

The strong influence of $2qV$ on the free energy barrier, cf. figure 1.10, warrants the study of field-driven polymer translocation in chapter 3. The simulation results will support the theoretical foundation laid down by Panja *et al.* [41]. Moreover, the electric field will play a pivotal role in chapter 5, where we investigate in how far translocation can be used to determine RNA secondary structures. In chapter 4 the topic changes to unbiased polymer translocation in a crowded environment of other polymers.

To accommodate the latter study, it is helpful to spell out the differences between the various time scales. We will follow the derivation by Klein Wolterink *et al.* [37]. To define the quantities precisely, we introduce the following states of the polymer. In state A (B), the entire polymer is located in cell A (B). States M, M' are defined as the states in which the middle monomer is located halfway between both cells. In state T and T' the polymer is threaded, but the monomer is not in the middle of the pore. Finally, the distinction between states M and T respectively M' and T', is that for the

former the polymer is en route from state A to B, or from state B to A, while for the latter it returns to where it came from, i.e., A (B) returns to A (B) without reaching B (A) first. With these definitions the translocation process can be characterized as a sequence of states, as shown in the right-hand side of figure 1.9. In this picture, the dwell time $\tau_{\rm d}$ is the mean time that the polymer spends in states M and T, while the translocation time $\tau_{\rm t}$ is the mean time starting at the first instant that the polymer reaches state B (A) after leaving state A (B), until it reaches state A (B) again. The unthreading time $\tau_{\rm u}$ is merely introduced for computational purposes. During the dwelling process, a polymer must pass through a state M at least once. Hence the unthreading time is defined as the mean time needed for the polymer to unthread starting from a state M. It is expected, and observed from simulations, that the scaling of the unthreading time is the same as the scaling of the dwell time. Clearly, computationally, to unthread from a state M is preferred, since no entropic barrier needs to be traversed.

Lastly, a relation can be found between $\tau_{\rm d}$ and $\tau_{\rm t}$. Let us define $p_{\rm M} + p_{\rm T}$ as the fraction of the translocation time $\tau_{\rm t}$ spent in states M and T. Also, $p_{\rm M'} + p_{\rm T'} \equiv 1 - \tau_{\rm d}/\tau_{\rm t}$. The probability that the polymer is threaded exactly halfway is an equilibrium property, i.e., the sum of the probabilities $p_{\rm M}$ and $p_{\rm M'}$ is found from the contribution of these states to the partition sum $Z^{\rm threaded}$ (Eq. (1.52) with $s = N/2$) to be

$$p_{\rm M} + p_{\rm M'} \sim Z^{\rm threaded}/Z_{\rm SAW} \sim N^{-\gamma+2\gamma_1-1}/V, \tag{1.54}$$

where a linear scaling with V of $Z_{\rm SAW}$ has been used. (Observe that for an ideal chain $p_{\rm M} + p'_{\rm M} \sim 1/V$.) The ratios $f_{\rm M} \equiv p_{\rm M}/p_{\rm M'}$ and $f_{\rm T} \equiv p_{\rm M}/p_{\rm T}$ are non-equilibrium properties which can be obtained from targeted simulations. With these quantities, the translocation time is found from the dwell time, using

$$\tau_{\rm t} = \tau_{\rm d} \frac{f_{\rm T}\,(1 + f_{\rm M})}{(p_{\rm M} + p_{\rm M'})\, f_{\rm M}\,(1 + f_{\rm T})}. \tag{1.55}$$

For unbiased polymer translocation of a single chain it was confirmed by direct simulations that

$$f_{\rm M}\,(N) \sim c, \tag{1.56}$$

with c a constant of order 1. We confirm the relation to hold also for unbiased translocation in a crowded environment. From these simulations we have also learned that

$$f_{\rm T}\,(N) \sim N^{-1}, \tag{1.57}$$

which is known to hold for single self-avoiding polymers too.

From these latter observations, together with Eq. (1.54), we find for the scaling of the translocation time for self-avoiding single chains[2]

$$\tau_{\rm t} \sim V\tau_{\rm d} N^{\gamma-2\gamma_1}, \tag{1.58}$$

which for ideal chains reduces to $\tau_{\rm t} \sim V\tau_{\rm d}$.

[2]Both Refs. [37, 40] wrongly state $\tau_{\rm t} \sim V\tau_{\rm d} N^{1+\gamma-2\gamma_1}$ instead of Eq. (1.58).

Chapter 2

Computational model

Abstract

All results presented in this thesis are tested against a highly efficient Monte Carlo algorithm for simulating polymers on a lattice. This extended repton model by Van Heukelum and Barkema [42, 43, 44, 45] finds its historic roots in Rubinstein's repton model [13], that was specifically designed to test de Gennes' reptation model [2]. The extended repton model was created to study dense melts of polydisperse chains of moderate length. For polymer translocation, under various external fields, in different environments, and for monomers with specific affinities, several further extensions are introduced. The resulting algorithm is well suited to the study of long polymers. An efficient slithering snake algorithm for preparing well-equilibrated melts of very long polymers is studied.

2.1 Lattice polymer models

In the previous chapter we introduced Doi's lattice model of reptation to investigate the immediate consequences of reptation theory. In this model, the longest relaxation time is proportional to the third power of the polymer length. Experimentally, however, $\tau_{\mathrm{d}} \sim \eta \sim N^{3.4\pm0.2}$ is consistently measured. Similarly, although off-lattice, the bead-spring model's longest relaxation time is proportional to N^2, and not $N^{1+2\nu}$. Obviously, the use of too simplistic models may lead to erroneous conclusions. On the other hand, too sophisticated models, besides missing the clarity of their simple counterparts, often are very expensive computationally. Given the presence of time scales of order

$N^{3.4}$, with N easily exceeding a thousand, it is evident that successful application of lattice polymer models hinges on finding the simplest representation of both the polymer configurations and dynamics. The extended repton model by Van Heukelum and Barkema strikes a good compromise between both requirements. Before its introduction in section 2.3, let us first briefly mention several other historically important models.

Already in 1962, Verdier and Stockmayer [46] simulated single self-avoiding walks with unit bond lengths between monomers on square and cubic lattices. Colloquially speaking, models do not get any simpler than this. Equilibrium distributions of the end-to-end distance of the polymer, as well as relaxation properties of initially stretched chains were investigated. Later, by filling the simulation box with multiple chains, semi-dilute and dense solutions were studied [47, 48] to confirm the crossover from $\nu = 0.588$ to $\nu = 0.5$. In a related model, with slightly adjusted dynamics, dense systems with monomer density $c = 0.5$ with up to 40 chains of length $N = 800$ were investigated [49, 50].

Another well-known model is the bond-fluctuation model by Carmesin and Kremer [51, 52]. The original version is restricted to two dimensions. Each monomer occupies several lattice sites, while bonds can only be of length 2, $\sqrt{5}$, $2\sqrt{2}$, 3, $\sqrt{10}$, or $\sqrt{13}$. The dynamics is restricted to single-monomer moves over a single lattice spacing. In later incarnations the model is extended to three dimensions [53]. Among its applications here too we find the study of dense systems. In the simulations by Kreer *et al.* [54] a system with monomer density $c = 0.5$ contains chains of up to length $N = 500$, while the dynamics progressed for ten million time steps.

The previous two models were both used to study the transition from Rouse to reptation dynamics, which was achieved by simply placing a lot of polymers in the simulation box. Rubinstein's repton model [13], to be discussed shortly, but also the cage model by Evans and Edwards [55, 56, 57] take an entirely different approach by only simulating a single chain. They both cleverly restrict the sideways moves of the monomers, so as to effectively incorporate De Gennes' reptation dynamics.

2.2 The repton model

Rubinstein introduced the repton model in 1987 [13]. It was specifically created to study the dynamics of reptation. It is a logical extension of the lattice

model of reptation (chapter 1.3.2), extended to also incorporate Doi's contour length fluctuations [15]. It deserves a short introduction, for it is the direct predecessor of the polymer model from which all our numerical results stem. The version described below is defined with dynamics and unit of time slightly different from Rubinstein's repton model [13].

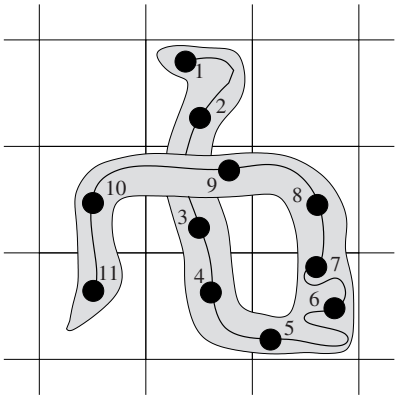

Figure 2.1: The repton model on a square lattice. The bond length between monomers is either zero or one. Therefore multiple monomers can occupy the same site. The polymer backbone obeys ideal statistics.

In reptation theory the contour length of the polymer exceeds the primitive path of the tube. The excess stored length diffuses along the primitive path. Rubinstein proposed the "repton model" as a discretized version of the reptation model. Its main features are illustrated in figure 2.1. The polymer is represented as a random walk, in which the consecutive monomers are also allowed to reside on the same site, i.e., bonds can be of length one or zero. A unit of "stored length" corresponds to zero bond length. The dynamics of the repton model consists of single-monomer hops along the chain, in addition to moves of the ends.

An inner monomer can be in only one of three states. If both or neither of its nearest neighbors are on the same site, then no move is possible. However, if only one of them is on the same site, then the stored length enables the monomer to make a jump along the chain. In figure 2.1, internal monomers 2, 5, and 7 can hop.

An end-monomer can be in only one of two states: either it resides on the

same site as its neighbor, and if so it may hop to any of the z nearest-neighbor lattice sites, or its lack of stored length leaves no alternative for change other than to jump to the site of its neighbor (for the polymer to retract). In figure 2.1 monomer 1 can move to any nearest neighbor site, while monomer 11 can only go to 10.

In every elementary move a randomly chosen monomer is moved if allowed by the rules defined above. Otherwise nothing is done. Time is incremented by one unit after N elementary moves, with N the degree of polymerization.

The average number of stored lengths is set by the lattice coordination number z. Let us show this by example. In figure 2.1 the polymer with curvilinear length 8 contains 3 units of stored length. We will name this configuration **A**. Monomer 11 in figure 2.1 can only go to neighbor 10, and in doing so it increases the number of stored lengths along the chain to 4. This new configuration is named **B**.

The next time monomer 11 is selected, it can go to any of the z surrounding sites, and hence the probability for it to go back to where it came from, is only $1/z$. From the requirement of detailed balance [58], for the move from **A** to **B** we have

$$P_{\mathbf{A}} p_{\mathbf{AB}} = P_{\mathbf{B}} p_{\mathbf{BA}}, \tag{2.1}$$

with $P_{\mathbf{A(B)}}$ the weight of configuration $\mathbf{A}(\mathbf{B})$ in the partition sum Z, and $p_{\mathbf{AB}}$ the probability, given that monomer 11 is selected, to go from configuration **A** to **B**. For the dynamics defined above, $p_{\mathbf{AB}} = 1$ and $p_{\mathbf{BA}} = 1/z$. Therefore, $P_{\mathbf{B}}/P_{\mathbf{A}} = z$.

The repton model's success is largely based on its ability to project the higher dimensional reptating polymer on the one-dimensional repton model. The lattice coordination number z is the only parameter to define the lattice. For example, $z = 4$ may represent a simple square lattice, $z = 6$ is appropriate for a simple cubic lattice, and $z = 12$ conforms to a Face Centered Cubic (FCC) lattice. The simple dynamics allowed Barkema *et al.* [59] to implement the model with multispin-coding techniques, which radically increased its numerical efficiency. They have been able to show that to leading order $N^2 D = 1/3$, with D the polymer's diffusion constant; this was later proven analytically [60, 61].

2.3 Extended repton model

The extended repton model by Van Heukelum and Barkema [42, 43, 44, 45] was specifically designed to study the dynamical properties of dense polymer solutions and melts, with a strong emphasis on computational efficiency. For polymer translocation through a narrow pore we are mainly interested in the dynamics of single chains. Thus, numerical efficiency becomes less of an issue. Moreover, with the introduction of various external fields and monomer affinities, a significant increase of the execution time per elementary move is unavoidable. Therefore, a detailed description of the multi-spin coding techniques, responsible for much of the code's efficiency, is omitted, but we refer the interested reader to Ref. [62]. It must be noted that the current implementation of the extended repton model uses a slightly modified dynamics, including a redefinition of the unit of time.

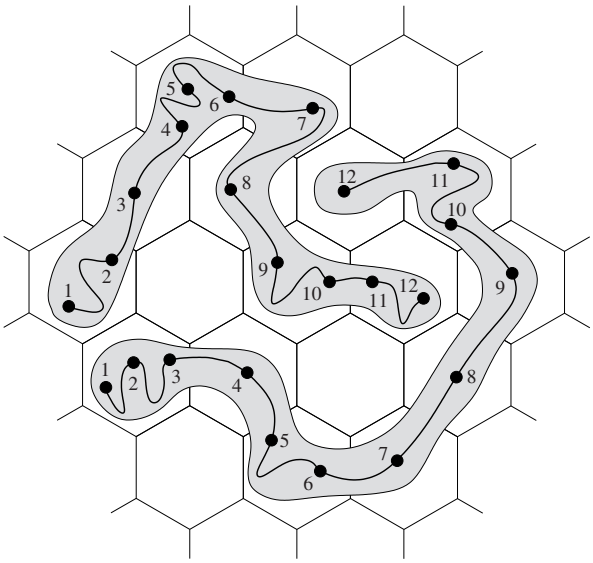

Figure 2.2: The extended repton model on a triangular lattice.

Rubinstein's repton model [13] simulates the dynamics of a single chain obeying random-walk statistics. The dynamics is limited to the diffusion of stored length along the chain. The primary extensions to the repton model comprise both excluded-volume constraints on the chains and sideways motion

of the interior monomers.

Figure 2.2 depicts two model polymers on a triangular lattice. As in the repton model, stored length may accumulate on a single lattice site. Therefore, the contour lengths of these polymers with degree of polymerization 12 are only 7 and 8. However, the chains respect excluded-volume constraints, i.e., primitive paths cannot cross or return, and only consecutive monomers can reside on the same lattice site.

The dynamics is reminiscent of the repton model, in that stored length can diffuse along the chain. This "reptation dynamics" randomly chooses a monomer, be it an interior or an end-monomer, and moves it in accordance with repton dynamics. If an interior monomer is selected, and only one of its neighbors is on the same site, then it hops to the site of the other neighbor. In figure 2.2 monomers 2, 4, 6, 9, 10, and 11 of the upper polymer, and 3, 5, 6, 10, and 11 of the lower polymer can reptate. If an end-monomer is picked, and it is not on the same lattice site as its neighbor, then stored length can be created. Monomer 12 of the lower polymer can join monomer 11. All the other exterior monomers can use the stored length to hop to any of the remaining 5 nearest neighbor lattice sites, provided no excluded-volume constraints are violated.

The sideways motion of monomers, which we call "Rouse dynamics", is new to the repton model. A random interior monomer is selected. If only one of its neighbors is on the same site, then the stored length permits a local increase of the polymer's contour length. For example, monomers 2, 4, 6, 9, 10, and 11 of the upper polymer and monomers 3, 5, 6, 10, and 11 of the lower polymer can hop to either of the two neighboring sites, provided these sites are free. Of course the opposite moves are allowed as well. For example, monomer 7 can either join monomer 6 or monomer 8 to shorten the upper polymer's contour length. These single-monomer moves are easily implemented on lattice structures which contain loops of three sites, such as the triangular lattice and the face-centered-cubic (FCC) lattice. Without these three-site loops the Rouse dynamics would only be possible by allowing multiple bond lengths between adjacent monomers.

In the repton model the average number of stored lengths only depends on the lattice coordination number z. From the dynamics we deduced the relative weight of chains with L and K stored lengths to be $P_L/P_K = z^{L-K}$. For the extended repton model, we will reason the other way round. For some positive constant W, we will impose $P_L/P_K = W^{L-K}$. An appropriate dynamics can

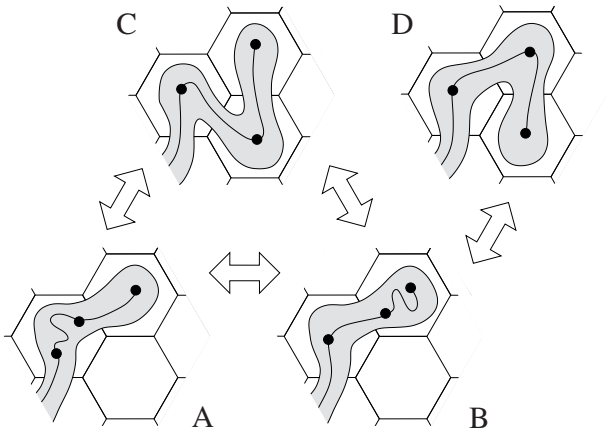

Figure 2.3: Detailed investigation of the Monte Carlo moves in the extended repton model.

then be constructed from the principle of detailed balance. In figure 2.3 the last three monomers of four nearly identical chains are shown. According to the reptation moves defined above, the middle monomer of chain **A** can hop to become chain **B**. Similarly, from the Rouse moves, **C** can become **A**, or **B**. The end-monomer in **D** can only hop to become **B**.

Since chains **A** and **B** contain equal numbers of stored length, the detailed balance reads

$$\frac{p_{\mathbf{AB}}}{p_{\mathbf{BA}}} = \frac{P_{\mathbf{B}}}{P_{\mathbf{A}}} = 1. \tag{2.2}$$

Hence, reptation dynamics for the interior monomers simply proceeds by selecting a random monomer to attempt a jump along the chain's contour.

A move from **C** to either **A** or **B** results in an additional stored length. Thus,

$$\frac{p_{\mathbf{AC}}}{p_{\mathbf{CA}}} = \frac{P_{\mathbf{C}}}{P_{\mathbf{A}}} = \frac{1}{W}, \tag{2.3}$$

and similarly for $\mathbf{C} \leftrightarrow \mathbf{B}$. All our simulation results are obtained on an FCC lattice, with $W = 4$. While for the triangular lattice only one other configuration equivalent to **C** can be reached, it is important to recognize that for the FCC lattice four other such configurations exist. Therefore, let's write

$p_{\mathbf{AC}} = Q/4$, with Q a constant still to be determined from the detailed balance equation and the requirement $p_{\mathbf{CA}} = 1/2$. From here,

$$\frac{1}{2} = p_{\mathbf{CA}} = W p_{\mathbf{AC}} = W\frac{Q}{4}, \tag{2.4}$$

to conclude $p_{\mathbf{AC}} = 1/2W$, with W at least 2. For Rouse dynamics a random interior monomer is picked, and if both its neighbors are on different lattice sites, cf. **C**, a flip of the coin decides to which nearest neighbor the monomer is moved. For the reversible moves, cf. $\mathbf{A} \to \mathbf{C}$, one of four equivalent configurations is picked, and the move proceeds with probability $1/2W$.

For the end-monomers in **C** and **D**,

$$\frac{p_{\mathbf{BD}}}{p_{\mathbf{DB}}} = \frac{P_{\mathbf{D}}}{P_{\mathbf{B}}} = \frac{1}{W}. \tag{2.5}$$

Also, since $p_{\mathbf{BD}} = 1/z$, we have $p_{\mathbf{DB}} = W/12$.

Note that the Monte Carlo algorithm based on reptation dynamics, together with moves of the end-monomers, is ergodic, even without the Rouse dynamics. Every unit of time N such monomer moves are attempted per chain, be they interior or not.

The computational time required for an elementary Rouse move is much higher than for a reptation move, since the former requires the lattice coordinates of the chosen monomer, while the latter only needs bond vectors, allowing for a very fast multi-spin coding implementation. Therefore, reptation moves are attempted R times as frequently as Rouse moves. The study of field-driven polymer translocation (chapter 3), and the translocation of RNA (chapter 5), have been performed with $R = 1$, while for translocation in a crowded environment (chapter 4) we have picked $R = 10$, which corresponds to a roughly equivalent amount of computational effort being invested in either kinds of moves. Recent work on a related lattice polymer model [63] provides evidence that the overall characteristics of the dynamics of a single polymer in bulk solution is controlled by the parameter $NR^{-1/2}$; provided the polymers are sizable, one can boost reptation dynamics over Rouse dynamics quite a bit before the polymer dynamics changes qualitatively.

2.4 External fields and affinity of monomers

An obvious advantage of lattice models is the easy inclusion of excluded-volume interactions: lattice sites are either empty or not. For our studies

of polymer translocation through a narrow pore, it is equally simple to just toggle the status of the wall's lattice sites to be occupied, while leaving open a pore. Another advantage of on-lattice polymers is the straightforward inclusion of several external fields, or the ability to easily bond monomers together. Several aspects of our implementation choices are discussed below.

2.4.1 Electric fields across the pore

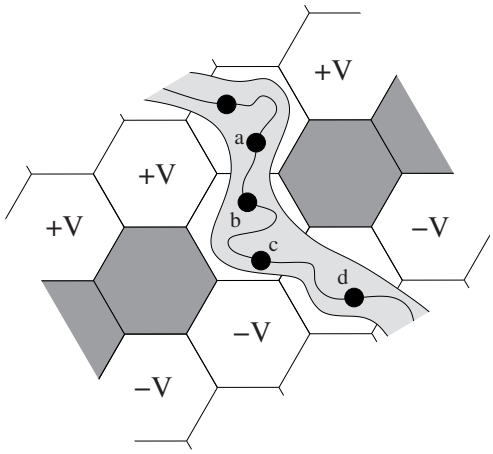

Figure 2.4: The voltage difference across the pore drives the translocation of the polymer.

In chapter 3 we investigate field-driven polymer translocation, i.e., a potential difference applied across the pore drives the translocation of the charged monomers. In figure 2.4 a voltage difference $2V$ is applied across the pore; if all monomers carry a (positive) charge q, then the relative Boltzmann weight of monomers **a** and **b** is simply $P_{\mathbf{a}}/P_{\mathbf{b}} = \exp(-qV/T)$, while an equivalent relation must be satisfied by monomers **c** and **d**. Without any electric field the extended repton model attempts hops from **a** to **b** as often as it tries to move **b** to **a**. Hence, we incorporate the electric field by attempting $\exp(qV) - 1$ additional moves from just left to within the pore, cf., **a** to **b**, as well as an

equal number of moves from within to just right of the pore, cf., **c** to **d**. Note that a move is defined by one hop of the reptation type and $1/R$ hops of the Rouse type.

2.4.2 Pulling of a molecule with optical tweezers

Several experimental groups have shown the ability to manipulate individual RNA and DNA molecules. A common approach is to bind one of its ends to a polystyrene bead, which can be trapped using infrared lasers, allowing control of its position [64, 65, 66, 67, 67, 69, 70]. These optical tweezers are characterized by a trap stiffness k_{tw} that relates the distance of the bead from the trap center, $dz = z_{\mathrm{tw}} - z_{\mathrm{b}}$, to the restoring optical force $F_{\mathrm{tw}} = -k_{\mathrm{tw}} dz$.

The polystyrene bead typically has a diameter of roughly half a micron [65]. On the other hand, the persistence length of a poly(U) RNA molecule is of the order of a nanometer. In our study of the translocation of RNA we relate the latter length scale to the lattice spacing λ. The incommensurate bead diameter forces upon us a severe approximation: only the chain's first monomer experiences the effect of the harmonic potential

$$V_{\mathrm{tw}} = \frac{1}{2} k_{\mathrm{tw}} \left(\mathbf{r}_{\mathrm{tw}} - \mathbf{r}_{\mathrm{b}} \right)^2 , \tag{2.6}$$

with $\mathbf{r}_{\mathrm{tw}}$ the center of the optical trap, and $\mathbf{r}_{\mathrm{b}}$ the location of the first monomer, representing the polystyrene bead.

Contrary to the monomers, the center of the optical trap is not restricted to the lattice. For clarity we refer to figure 2.5, which depicts two stacked FCC unit-cells. It is helpful to note that the set of points

$$\mathbf{r} = t\hat{\mathbf{t}} + u\hat{\mathbf{u}} + v\hat{\mathbf{v}} + w\hat{\mathbf{w}}, \tag{2.7}$$

$$\begin{aligned}
\hat{\mathbf{t}} &= \frac{1}{\sqrt{2}} \left(-\hat{\mathbf{x}} + \hat{\mathbf{y}} \right) , \\
\hat{\mathbf{u}} &= \frac{1}{\sqrt{2}} \left(-\hat{\mathbf{y}} + \hat{\mathbf{z}} \right) , \\
\hat{\mathbf{v}} &= \frac{1}{\sqrt{2}} \left(+\hat{\mathbf{x}} + \hat{\mathbf{y}} \right) , \\
\hat{\mathbf{w}} &= \frac{1}{\sqrt{2}} \left(-\hat{\mathbf{y}} - \hat{\mathbf{z}} \right) ,
\end{aligned} \tag{2.8}$$

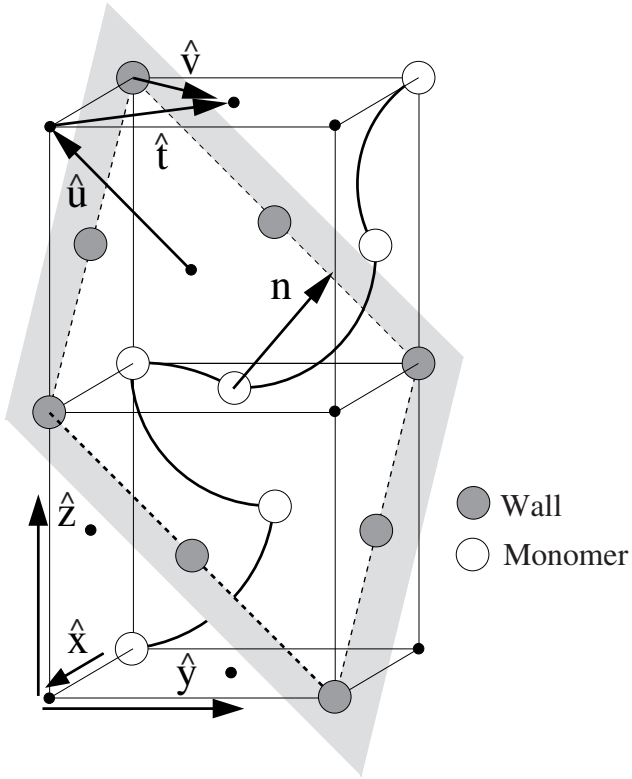

Figure 2.5: Two stacked FCC unit-cells. The open spheres represent a typical configuration of a polymer in transit through the pore.

with t, u, v, and w restricted to integers, and constrained to $t + u + v + w = 0$, forms an FCC-lattice. Therefore, in our implementation of the extended repton model, the lattice sites are conveniently identified by vectors $t\hat{\mathbf{t}} + u\hat{\mathbf{u}} + v\hat{\mathbf{v}}$. In all our polymer translocation simulations the wall is defined as a plane with constant v-coordinate, i.e., a plane spanned by vectors $\hat{\mathbf{t}}$ and $\hat{\mathbf{u}}$ with normal $\hat{\mathbf{n}} = 1/\sqrt{3}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})$. The eight solid spheres in figure 2.5 form the circumference of the pore, while the open spheres represent the lattice sites occupied by the polymer.

We take the center of the optical trap to move along the normal to the wall, i.e., $\mathbf{r}_{\mathrm{tw}} = v_{\mathrm{tw}} t \hat{\mathbf{n}}$, with v_{tw} the trap's speed and time t.

Inversely, for any point in space we can write

$$\mathbf{r} = \frac{1}{\sqrt{2}}(v-t)\hat{\mathbf{x}} + \frac{1}{\sqrt{2}}(v+t-u)\hat{\mathbf{y}} + \frac{1}{\sqrt{2}}u\hat{\mathbf{z}}, \tag{2.9}$$

with which it is straightforward to express Eq. (2.6) in the (t, u, v)-coordi- nate system as

$$V_{\text{tw}} = \frac{1}{2}k_{\text{tw}}\left[dt^2 + dv^2 + du^2 - dt\ du - dv\ du\right]. \tag{2.10}$$

Since $v_{\text{tw}} \ll \lambda/\Delta t$ the position of the trap center is approximately stationary during a time step, and a simple Metropolis accept/reject

$$p_{\text{acc}}(t, t+\Delta t) = \min\left[1, \exp\left(\{V_{\text{tw}}(t) - V_{\text{tw}}(t+\Delta t)\}/T\right)\right] \tag{2.11}$$

suffices to update the first monomer's position.

2.4.3 RNA, or the affinity of base pairs

An RNA molecule is a linear polymer in which the nucleotides are linked together by means of phosphodiester bonds. These bonds link the 3' carbon in the ribose of one nucleotide to the 5' carbon in the ribose of the adjacent nucleotide. Besides the ribose sugar and phosphate, each nucleotide also contains an organic (nitrogenous) base: either Adenine (A) or Guanine (G), which are Purines, or, Cytosine (C) or Uracil (U), which are Pyrimidines. In the canonical Watson-Crick base pairing, adenine forms a base pair with uracil, as does guanine with cytosine. The strength of the interaction between C and G is stronger than between A and U, because the former pair has three hydrogen bonds joining them while the latter has only two.

We extend the lattice polymer to resemble an RNA molecule by labeling monomers to be either A, G, C, U, or none of these. Each such labeled nucleotide carries a list of complementary nucleotides that it can bond with. Let us impose that nucleotides with a separation of less than 3 units along the backbone are not able to form a bond. Then, in figure 2.6, nucleotide $\mathbf{C}_1$ finds on its list nucleotides $\mathbf{G}_4$ and $\mathbf{G}_7$. The lists of both $\mathbf{G}_4$ and $\mathbf{G}_7$ only comprise $\mathbf{C}_1$. $\mathbf{A}_6$ can bond with $\mathbf{U}_2$ as well as with $\mathbf{U}_3$, while the list of $\mathbf{A}_5$ only contains $\mathbf{U}_2$. Similarly, $\mathbf{U}_3$ can bond with $\mathbf{A}_6$, while $\mathbf{U}_2$ allows both $\mathbf{A}_5$ and $\mathbf{A}_6$. Spatially, however, bonded nucleotides need to be nearest neighbors, cf. nucleotides $\mathbf{C}_1$ and $\mathbf{G}_7$. Hence, we dismiss monomer hops that require the breaking of a bond.

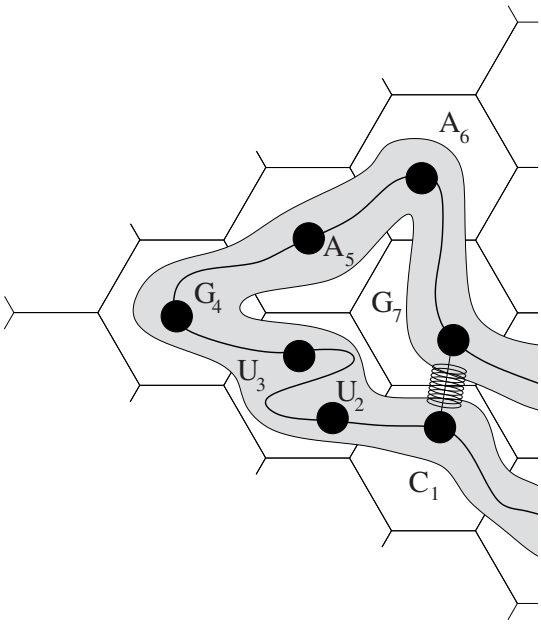

Figure 2.6: An RNA molecule with a single CG bond with affinity E_{CG}.

To handle the base pairing itself, we add new Monte Carlo steps. Let us refer to figure 2.6 as configuration **A**. The same polymer without the $\mathbf{C}_1\mathbf{G}_7$-bond is **B**. The new elementary steps satisfy the detailed balance equation

$$p_{\mathbf{BA}}/p_{\mathbf{AB}} = P_{\mathbf{A}}/P_{\mathbf{B}} = \exp\left(E_{CG}/T\right), \tag{2.12}$$

with E_{CG} the affinity of the base pair. Every step a random monomer is chosen, and if it is without a bond, then a member on its list is picked randomly. Only if this other nucleotide is free and if it is located on a nearest neighbor lattice site, a new base pair will be formed.

In every elementary step the formation of the $\mathbf{C}_1\mathbf{G}_7$-bond is attempted, on average,

$$p_{\mathbf{BA}} = \frac{1}{N}\left(\frac{1}{N_{\mathbf{C}_1}} + \frac{1}{N_{\mathbf{G}_7}}\right) = \frac{1}{N}\frac{3}{2} \tag{2.13}$$

times. In the opposite direction, from Eq. (2.12),

$$p_{\mathbf{AB}} = \frac{1}{N}\left(\frac{1}{N_{\mathbf{C}_1}} + \frac{1}{N_{\mathbf{G}_7}}\right)\exp\left(-E_{CG}/T\right). \tag{2.14}$$

In configuration **A**, the probability to pick either of the two bonded monomers is $2/N$. Thus, to satisfy Eq. (2.14), we have to break the bond with probability

$$p_{\text{break}} = p_{\mathbf{AB}}N/2 = \frac{1}{2}\left(\frac{1}{N_{\mathbf{C}_1}} + \frac{1}{N_{\mathbf{G}_7}}\right)\exp\left(-E_{CG}/T\right). \tag{2.15}$$

2.5 Equilibration of melts

To investigate the translocation of a long polymer in a dense crowded environment of others (chapter 4), we need to be able to generate well-equilibrated polymeric systems. In the previous chapter we have seen that in a melt De Gennes' standard reptation theory predicts a longest relaxation time $\tau_{\text{d}} \sim N^3$, while experimentally the situation is worse, with $\tau_{\text{d}} \sim N^{3.4\pm0.2}$. The dynamics of the extended repton model adheres to the same scaling laws. Hence, it is clearly undesirable to only rely on the dynamics of the extended repton model to prepare equilibrated samples of long chains. Here we describe an implementation of the slithering snake algorithm which proves extremely effective to equilibrate systems with hundreds of polymers, each with thousands of monomers. For these systems, equilibration times are reduced from quite literally decades, to mere days.

2.5.1 Internal distances, and the near-ideality of chains

It is far from trivial to assess whether a dense polymeric system is near equilibrium, since the question is intimately related to the rather intangible concept of entanglements [71].

Auhl *et al.* [72] describe several methods to equilibrate a coarse-grained model of the bead-spring type. They characterize the chain conformations in the melt mainly by relying on the mean squared internal distance $\langle R^2\rangle(|i-j|, N)$ averaged over all segments of size $n = |i-j|$ along the chains, where $i < j \in [1, N]$ are monomer indices. The authors measure the efficiency of their methods from the observation that $\langle R^2\rangle(n)/n$ approaches a "target function" as the system approaches equilibrium.

In figure 2.7 we show these target functions that we have obtained from samples that are at unit monomer-density, all in a periodic simulation box with $L \times U \times V = 104^3$ sites. The degree of polymerization varies from 169 to 10816. In the inset, for $N = 10816$, we have plotted n versus $\langle R^2 \rangle(n)$ on a logarithmic scale. Clearly, $\langle R^2 \rangle(n) \sim n^{1/2}$, for n not too small, to show that for not too short internal distances the chain statistics are near to ideal. However, excluded-volume effects are observed for monomers that are connected through a small number of bonds.

While it is certainly true that a unique target function exists in equilibrium, this is not to say that it allows an easy identification of this equilibrium. Significant time-averaging is required to overcome the fluctuations of $\langle R^2 \rangle(n)$, which are strongly present in these systems far from the thermodynamic limit.

Consequently, after we have presented our implementation of the slithering snake algorithm, we will propose an alternative measure that permits a more accurate determination of the equilibration times, well suited to lattice polymers with stored length.

2.5.2 Slithering snake algorithm

The Slithering Snake Algorithm (SSA) was invented independently by Kron [73, 74] and Wall and Mandel [75, 76]. The elementary moves of this algorithm are slithering, reptation-like, motions of a single chain: a monomer is appended at one end of the chain, and one monomer is simultaneously deleted from the other end.

Three variants of the algorithm are proposed by the same authors. In the first version by Wall and Mandel [75], the "head" and "tail" of the chain are interchanged only when an attempted move is rejected. This algorithm violates the detailed-balance condition. In Wall and Mandel's second version, for each step the head and tail are chosen randomly with fifty percent probability. This variety satisfies detailed balance. Note that Doi's lattice model of reptation (chapter 1.3.2) is a slithering snake algorithm of this type for ideal chains. Kron's algorithm is much more complicated than either of these, but it does not satisfy detailed balance.

It must be noted that for self-avoiding walks these SSAs are nonergodic [75]. Frozen configurations occur when both ends of the chain are trapped.

From Doi's lattice model of reptation we have learned that the autocorrelation time scales as $\tau \sim N^2$ (observe that τ_{d} in Eq. (1.47) carries an extra N to connect simulation time steps to the Rouse modes). A simple heuristic

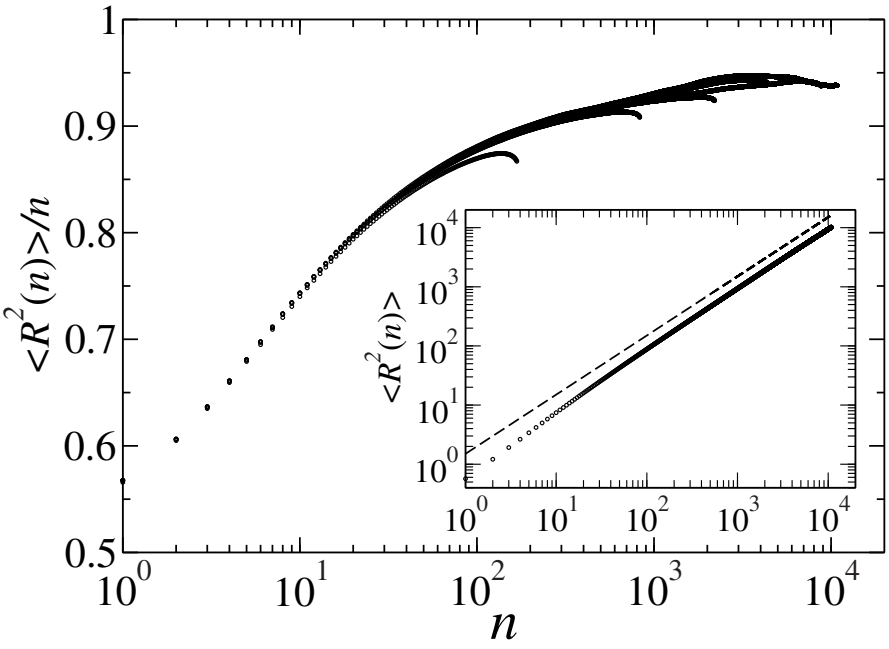

Figure 2.7: Mean squared distance $\langle R^2(n)\rangle$ as a function of curvilinear distance n, obtained from simulations with slithering snake moves. Each "target function" is averaged over 8 independent simulations, and time-averaged over at least 100 configurations obtained from the slithering snake algorithm after equilibration from $\tau_{\rm eq} \lesssim t \lesssim 2\tau_{\rm eq}$. All simulations were conducted at unit monomer-density, in a simulation box of size $L \times U \times V = 104^3$ lattice sites with periodic boundary conditions. For the degree of polymerization we have used $N = 169$, 832, 2197, 4394, 8788, and 10816, which amounts to at least 104 polymers in the simulation box. Inset: double-logarithmic plot of the mean squared internal distance for $N = 10816$ only. The dashed line has a slope of 1/2, to show the near-ideality of a polymer in the melt.

argument suggests that also for the self-avoiding walks in an ergodic class, the slithering snake respects $\tau \sim N^2$ [77].

The intrinsic connection with reptation theory has attracted several researchers to use the SSA to test the concepts of reptation theory in dense polymeric systems. in 1985 Deutsch put seven self-avoiding chains of length 300 in a cubic simulation box [78]. Every time step head and tail were randomly chosen. His simulation results suggest that the viscosity of concentrated polymer systems is proportional to $N^{3.25\pm0.1}$, i.e., the autocorrelation time for the SSA exceeds N^2 for dense systems. More recently, Mattioni *et al.* [79] have used an SSA for the bond-fluctuation model to present results that are in qualitative agreement with the activated-reptation hypothesis by Deutsch [78, 80], and later extended by Semenov and Rubinstein [81, 82, 83]. It is beyond the scope of the section to discuss whether the activated-reptation hypothesis leads to a viable explanation of the experimentally measured viscosity $N^{3.4\pm0.2}$. Irrespective of the exact mechanism, clearly the time required to bring dense polymeric samples to equilibrium with the slithering snake algorithm exceeds the lower bound N^2.

But as long as the autocorrelation time does not exceed $N^{3.4\pm0.2}$, it ought to be advantageous to employ an SSA to reach equilibrium, instead of a more 'realistic' dynamics. Computationally, another benefit of the SSA is that in every time step, irrespective of chain length, only moves of the end-monomers are attempted, i.e., the cost of a single time step only scales with the number of chains, whereas for ordinary local dynamics the cost is proportional to the number of monomers in the simulation box.

Our implementation is an algorithm of the first type: the head and tail are only interchanged when an attempted move is rejected. (It should be noted that this choice is arbitrary, and we assume it to be without consequences.) To ensure an appropriate density of stored length along the chains, as required by the extended repton model, we use the following algorithm:

1. with probability $z/(W+z)$, resp. $W/(z+W)$ we attempt to add a new head monomer and to shrink the tail, by randomly selecting one of the head's z nearest neighbor lattice sites, or the site of the head itself.

2. if the selected site is not forbidden, then the move is accepted. Otherwise, head and tail are interchanged.

2.5.3 Results and conclusions

The mean squared internal distance $\langle R^2 \rangle(n)$ converges to a target function as the system progresses towards equilibrium. Unfortunately, extensive time-averaging is required to reach a similar degree of smoothness as witnessed in figure 2.7. Luckily, for lattice polymers with stored lengths, other measures can be devised.

At $t = 0$ compressed polymers, i.e., polymers with $N - 1$ units of stored length, are placed randomly in the simulation box. As the polymers unfold with the passing of time, the stored length density along their contours has to approach its equilibrium value. Computationally, at a monomer density of $c = 1$, we have measured an average of $\rho_{\mathrm{sl}}(t \to \infty) = 0.43$ units of stored lengths per lattice site of the box, with a near-negligible influence of polymer length. The excess stored length density can only have reached the equilibrium value after its head and tail have moved along a curvilinear distance of at least qN and $(1 - q)N$, with $0 < q < 1$, which, in equilibrium, amounts to complete tube renewal. Therefore, we presume that the onset of equilibrium is recorded at the earliest time for which $\rho_{\mathrm{sl}}(t) - \rho_{\mathrm{sl}}(\infty) < \epsilon$, with ϵ sufficiently small. Unfortunately, the assumed monotonic decrease of $\rho_{\mathrm{sl}}(t < \tau_{\mathrm{eq}})$ is often violated for $t \lesssim \tau_{\mathrm{eq}}$, i.e., $\rho_{\mathrm{sl}}(t)$ often overshoots its equilibrium value, before returning to it. Moreover, the non-negligible fluctuations of $\rho_{\mathrm{sl}}(t)$ worsen the clear identification of τ_{eq}.

A related, but better behaved criterion can be constructed easily: we define τ_{eq} as the time after which the standard deviation σ_{sl} of the distribution of units of stored lengths per chain σ_{sl} reaches a plateau, indicating the onset of equilibrium. The inset of figure 2.8 displays $\sigma_{\mathrm{sl}}(t)$, measured from one simulation each, for polymers with N ranging from 169 to 10816. To be precise, we define τ_{eq} to be the earliest time for which

$$\sigma_{\mathrm{sl}}(t) - (1 + \epsilon) \left\langle \sigma_{\mathrm{sl}} \left(t > \tau_{\mathrm{eq}} \right) \right\rangle < 0, \tag{2.16}$$

with ϵ roughly of the order of the relative magnitude of the fluctuations of $\sigma_{\mathrm{sl}} \left(t > \tau_{\mathrm{eq}} \right)$.

Figure 2.8 depicts τ_{eq} for various polymer lengths. A fair description of the effective scaling of τ_{eq} is given by $\tau_{\mathrm{eq}} \sim N^2 \exp\left(0.26 N^{1/3}\right)$. Mattioni *et al.* [79] have analyzed their bond-fluctuating SSA in terms of a similar expression, that finds its origin in the activated-relaxation hypothesis. Alternatively, for longer chains, $\tau_{\mathrm{eq}} \approx 3.3 \cdot 10^{-4} N^{3.4}$. The tiny prefactor, together with the extremely low computational cost of a slithering snake move, still makes for a

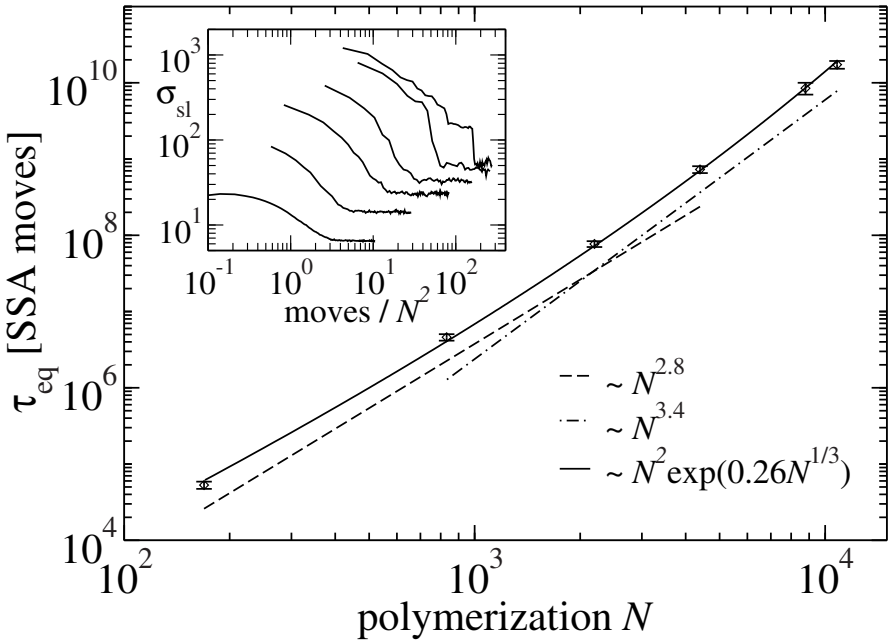

Figure 2.8: Equilibration times τ_{eq} measured with Eq. (2.16). The simulation box with $L \times U \times V = 104^3$ lattice sites and periodic boundary conditions, is filled to unit monomer-density with polymers with $N = 169$, 832, 2197, 4394, 8788, and 10816 monomers. The error bars are obtained from 8 independent simulations. Inset: σ_{sl} from one of eight independent simulations.

very efficient technique. We acknowledge that finite-size effects may influence the quantitative results.

Chapter 3

Field-driven polymer translocation

This chapter has been published as *Pore-blockade Times for Field-Driven Polymer Translocation*, H. Vocks, D. Panja, G. T. Barkema, and R. C. Ball, J. Phys.: Condens. Matter **20**, 095224 (2008), apart from minor modifications.

Abstract

We study pore blockade times for a translocating polymer of length N, driven by a field E across the pore in three dimensions. The polymer performs Rouse dynamics, i.e., we consider polymer dynamics in the absence of hydrodynamic interactions. We find that the typical time the pore remains blocked during a translocation event scales as $\sim N^{(1+2\nu)/(1+\nu)}/E$, where $\nu \simeq 0.588$ is the Flory exponent for the polymer. We show that this scaling behavior stems from the polymer dynamics at the immediate vicinity of the pore — in particular, the memory effects in the polymer chain tension imbalance across the pore. This result, along with the numerical results by several other groups, violates the lower bound $\sim N^{1+\nu}/E$ suggested earlier in the literature. We discuss why this lower bound is incorrect and show, based on conservation of energy, that the correct lower bound for the pore-blockade time for field-driven translocation is given by $\eta N^{2\nu}/E$, where η is the viscosity of the medium surrounding the polymer.

3.1 Introduction

Molecular transport through cell membranes is an essential mechanism in living organisms. Often, the molecules are too long, and the pores in the membranes too narrow, to allow the molecules to pass through as a single unit. In such circumstances, the molecules have to deform themselves in order to squeeze — i.e., translocate — themselves through the pores. DNA, RNA and proteins are such naturally occurring long molecules [84, 85, 86, 87, 88] in a variety of biological processes. Translocation is also used in gene therapy [89, 90], and in delivery of drug molecules to their activation sites [91]. Consequently, the study of translocation is an active field of research: as a cornerstone of many biological processes, and also due to its relevance for practical applications.

Recently, translocation has found itself at the forefront of single-molecule detection experiments [92, 93, 94, 95, 96, 97, 98, 99], as new developments in the design and fabrication of nanometer-sized pores and etching methods may lead to cheaper and faster technology for the analysis and detection of single macromolecules. In these experiments, charged polymeric molecules, suspended in an electrolyte solution, are initially located on one side of a membrane. The membrane is impenetrable to the molecule except for a nanometer-sized pore. Between the two different sides of the membrane, a DC voltage difference is then applied, which drives the molecule through the pore. When the molecule enters the pore, it affects the electrical resistivity of the circuit, leading to a dip in the electric current supplied by the voltage source. The magnitude and the duration of these dips have proved to be very effective in determining the size and the length of the molecule. The usage of protein pores (modified α-haemolysin, mitochondrial ion channel, nucleic acid binding/channel protein etc.) and the etching of specific DNA sequences inside the pores [89, 100, 101, 102] have opened up promising new avenues of fast, simple and cheap technology for single macromolecule detection, analysis and characterization, perhaps even allowing DNA sequencing at the nucleotide level.

The subject of this chapter is (charged) polymer translocation in three dimensions through a narrow pore in an otherwise impenetrable membrane placed at $z = 0$, as the polymer is driven by a DC voltage across the pore. Our interest is in the scaling behavior for the typical pore-blockade time during a translocation event with polymer length N. In practice, the electric field due to the applied voltage decays rapidly with increasing distance from the

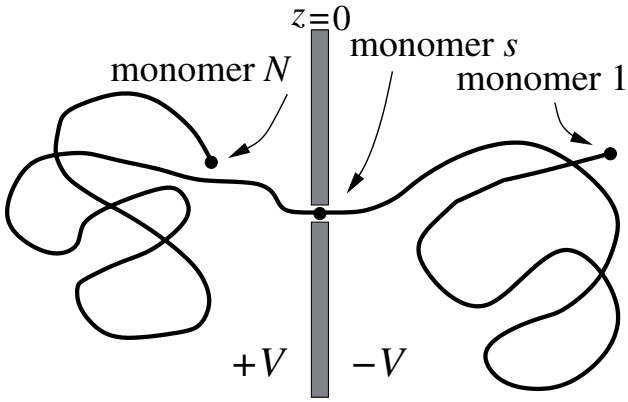

Figure 3.1: Snapshot of a translocating polymer in a two-dimensional projection of our three-dimensional system. Across the pore of size unity a voltage difference $2V$ is applied. The monomer located within the pore is labeled s.

pore, and for simplicity it is often assumed that only those polymer segments residing within the pore feel the driving force due to the field. For our theory and simulations too, we consider a polymer which only experiences a force acting on its monomers that reside in the pore, as illustrated in Fig. 3.1.

To substantiate our theoretical analysis we use extensive Monte Carlo simulations with a three-dimensional self-avoiding lattice polymer model. For the voltage profile across the pore we choose

$$V(z) = \begin{cases} +V & (z \leq -1) \\ 0 & (z = 0) \\ -V & (z \geq 1) \end{cases} . \tag{3.1}$$

Thus, during translocation through the pore, the energy gained by each monomer carrying a charge q, in dimensionless units, is given by $\Delta U = 2qV/k_BT$. From now on, favoring notational simplicity, we choose both q and k_BT to be unity. Since we also choose the lattice spacing to be unity in our simulations, the strength of the electric field acting on each monomer within the pore is given by $E = V$.

Details of the lattice polymer model used in this chapter can be found in chapter 2: the polymer moves through a sequence of random single-monomer hops to neighboring lattice sites. These hops can either be “reptation moves”,

along the contour of the polymer, or "Rouse moves", in which the monomer jumps "sideways" and changes the contour. The definition of time used throughout this chapter is such that every monomer attempts a reptation move as well as a sideways move with rate unity. There is no explicit solvent in our analysis, i.e., the polymer performs Rouse dynamics.

Our conventions to study this problem, throughout this chapter, are the following. We place the membrane at $z = 0$. We fix the middle monomer (monomer number $N/2$) of a polymer of total length N at the pore, apply the voltage as in Eq. (3.1) and thermalize the polymer. At $t = 0$ we release the polymer and let translocation commence. We define the typical time when the polymer leaves the pore as the dwell time $\tau_{\rm d}$: it scales with N in the same way as the pore-blockade time in a full (field-driven) translocation event.

This problem has recently been studied in Refs. [32, 33], in which a lower bound $\propto N^{1+\nu}/E$ has been argued for $\tau_{\rm d}$. This lower bound was derived in the limit of unimpeded polymer movement, i.e., for an infinite pore, or equivalently, in the absence of the membrane. In Ref. [32] the authors also suggested that the dynamics of translocation is anomalous (see also Ref. [33] in this context).

In the recent past, Panja *et al.* have been investigating the microscopic origin of the anomalous dynamics of translocation. They have set up a theoretical formalism, *based on the microscopic dynamics of the polymer*, and showed that the anomalous dynamics of translocation stem from the polymer's memory effects, in the following manner. Translocation proceeds via the exchange of monomers through the pore: imagine a situation when a monomer from the left of the membrane translocates to the right. This process increases the monomer density in the right neighborhood of the pore, and simultaneously reduces the monomer density in the left neighborhood of the pore. The local enhancement in the monomer density on the right of the pore *takes a finite time to dissipate away from the membrane along the backbone of the polymer* (similarly for replenishing monomer density on the left neighborhood of the pore). The imbalance in the monomer densities between the two local neighborhoods of the pore during this time causes an enhanced chance of the translocated monomer to return to the left of the membrane, thereby giving rise to *memory effects*. The ensuing analysis enabled Panja *et al.* to provide a proper microscopic theoretical basis for the anomalous dynamics. Further theoretical analysis then led us to the conclusion that in the case of unbiased translocation, i.e., when the polymer is not subjected to an external force, the

dwell time scales with length as $\tau_{\rm d} \sim N^{2+\nu}$ [40, 41, 103], both in two and three dimensions. The approach based on the polymer's memory effects also works beautifully for pulled translocation, during which a force F is applied at the head of the polymer: Panja and Barkema have shown that if FN^{ν} is sufficiently large, then the dwell time scales as $\tau_{\rm d}/N^{2+\nu} \sim (FN^{\nu})^{-1}$ [104]. In this work, we push ahead with the same formalism to demonstrate that it reveals the physics of field-driven translocation too, thus providing a *unified underlying theoretical basis* for translocation, based on the theory of polymer dynamics.

authors	two dimensions	three dimensions
Kantor *et al.* [32]	1.53 ± 0.01	–
Luo *et al.* [35, 105, 36]	1.69 ± 0.04	1.42 ± 0.01
Cacciuto *et al.* [106]	1.55 ± 0.04	–
Wei *et al.* [107]	–	1.27
Milchev *et al.* [108]	–	1.65 ± 0.08
Dubbeldam *et al.* [39]	–	1.5

Table 3.1: Existing numerical results on the exponent for the scaling of $\tau_{\rm d}$ with N for field-driven translocation. Note that the proposed lower bound $1 + \nu$ of Ref. [32] is 1.75 and 1.59 in two and three dimensions respectively.

Returning to the lower bound for the scaling of the dwell time with polymer length N for field-driven translocation as proposed in Ref. [32], we note that subsequent numerical studies, including the one by the authors of Ref. [32] themselves, did not immediately settle the scaling for $\tau_{\rm d}$ with N. In Table 3.1 we present a summary of the existing numerical results on the exponent for the scaling of $\tau_{\rm d}$ with N for field-driven translocation. All results quoted are for self-avoiding polymers in the absence of hydrodynamic interactions in the scaling limit.

More recently, this lack of consensus prompted Panja *et al.* to investigate the issue of field-driven translocation in two dimensions, via a related problem, viz., polymer translocation in three dimensions out of strong planar confinements [103]. They showed that the actual lower bound for $\tau_{\rm d}$ for field-driven translocation is given by $\eta N^{2\nu}/E$, where η is the viscosity of the surrounding medium. This inequality is derived from the principle of conservation of energy: it was shown in Ref. [103] that although the presence of the

memory effects suggests that the scaling of $\tau_{\rm d}$ could behave as $N^{(1+2\nu)/(1+\nu)}$, since $(1+2\nu)/(1+\nu) < 2\nu$ in two dimensions, conservation of energy overrides the memory effects in the polymer — high precision simulation data suggested, in accordance with those of Refs. [32, 106] that the actual scaling of $\tau_{\rm d}$ for field-driven translocation in two dimensions is given by $\tau_{\rm d} \sim N^{2\nu}$. In three dimensions $2\nu < (1+2\nu)/(1+\nu)$, implying that in three dimensions $\tau_{\rm d} \sim N^{(1+2\nu)/(1+\nu)}$, which is the central result of this chapter.

This chapter is organized in the following manner. In Sec. 3.2 we derive the lower bound $N^{2\nu}$ for $\tau_{\rm d}$ for field-driven translocation. In Sec. 3.3.1 we discuss a method to measure the polymer's chain tension at the pore. In Sec. 3.3.2 we analyze the memory effects in the imbalance of the polymer's chain tension at the pore. In Sec. 3.4 we discuss the consequence of these memory effects on the translocation velocity $v(t)$, and obtain the scaling relation of $\tau_{\rm d}$ with the polymer length N. We end this chapter with a discussion in Sec. 3.5.

3.2 Lower bound for $\tau_{\rm d}$ for field-driven translocation

As noted in Sec. 3.1, a lower bound for the dwell time $\tau_{\rm d} \sim N^{1+\nu}/E$ has been proposed in Ref. [32]. The underlying assumption behind this result is that, with or without an applied field, the mobility of a polymer translocating through a narrow pore in a membrane will not exceed that of a polymer in bulk (i.e., in the absence of the membrane). This mobility is then obtained under two more assumptions for the behavior of a polymer under a driving field:

(i) To mimic the field acting on a translocating polymer, the field on the polymer in bulk has to act on a monomer whose position along the backbone of the polymer changes continuously in time. As a result, there is no incentive for the polymer to change its shape from its bulk equilibrium shape, i.e., the polymer can still be described by a blob with radius of gyration $\sim N^{\nu}$ in the appropriate dimension.

(ii) The polymer's velocity is proportional to DE, where E is the applied field, and D is the diffusion coefficient scaling as $D \sim 1/N$ for a Rouse polymer.

Of these two assumptions, note that (ii) is obtained as the steady state solution of the equation of motion of a Rouse polymer, in bulk, with uniform velocity and vanishing internal forces, see Eq. (1.28). We have already witnessed

in many occasions [32, 33, 34, 109, 38, 39, 40, 41, 104] that the dynamics of translocation through a narrow pore is anomalous (subdiffusive), as a consequence of the strong memory effects discussed in the previous section, and also that these memory effects are so strong that the velocity of translocation is not constant in time [104, 103]. The anomalous dynamics and the memory effects are crucial ingredients that question the validity of the lower bound $N^{1+\nu}$ for $\tau_{\rm d}$ for field-driven translocation.

It is however possible to derive a lower bound for $\tau_{\rm d}$ for field-driven translocation, based on the principle of conservation of energy. Consider a translocating polymer under an applied field E which acts only at the pore. By definition, the N monomers of the polymer translocate through the pore in a time $\tau_{\rm d}$. The total work done by the field in this time $\tau_{\rm d}$ is then given by EN. During translocation, each monomer travels over a distance of order $\sim R_g$, leading to an *average* monomer velocity $v_m \sim R_g/\tau_{\rm d}$. The rate of loss of energy due to the viscosity η of the surrounding medium per monomer is given by ηv_m^2. For a Rouse polymer, the frictional force on the entire polymer is a sum of frictional forces on individual monomers, leading to the total energy loss due to the viscosity of the surrounding medium during the entire translocation event scaling as $\sim N\tau_{\rm d}\eta v_m^2 = N\eta R_g^2/\tau_{\rm d}$. This loss of energy must be less than or equal to the total work EN done by the field, which yields us the inequality[1] $\tau_{\rm d} \geq \eta R_g^2/E = \eta N^{2\nu}/E$.

3.3 Memory effects in the chain tension perpendicular to the membrane

A translocating polymer can be thought of as two segments of polymers tethered at the pore, while the segments are able to exchange monomers between them through the pore. In Ref. [41] Panja *et al.* developed a theoretical method to relate the dynamics of translocation to the imbalance of chain tension between these two segments across the pore. The key idea behind this method is that the exchange of monomers across the pore responds to $\phi(t)$, this imbalance of chain tension; in its turn, $\phi(t)$ adjusts to $v(t)$, the transport velocity of monomers across the pore. Here, $v(t) = \dot{s}(t)$ is the rate of exchange of monomers from one side to the other.

[1] For unbiased translocation, EN in this argument is to be replaced by the difference in free energy (or entropy) of a threaded polymer, corresponding to $s = N/2$, and the translocated polymer, corresponding to $s = N$. This leads to the inequality $\tau_{\rm d} \geq \eta N^{1+2\nu}$.

The memory effects discussed in Sec. 3.1 in terms of relaxation of excess monomers (or the lack of monomers) in the immediate vicinity of the pore translates immediately to that of the imbalance of the chain tension across the pore — local accumulation of excess monomers reduces the chain tension, while local lack of monomers enhances it. Quantitatively speaking, in the presence of memory effects, the chain tension imbalance across the pore $\phi(t)$ and the velocity of translocation $v(t)$ are related by

$$\phi(t) = \phi_{t=0} + \int_0^t dt' \mu(t-t') v(t') \tag{3.2}$$

via the (field-dependent) memory kernel $\mu(t)$, which could be thought of as time-dependent 'impedance' of the system. Using the Laplace transform, this relation could be inverted to obtain $v(t) = \int_0^t dt' a(t-t')[\phi_{t=0} - \phi(t')]$, where $a(t)$ can be thought of as the 'admittance' of the system. In the Laplace transform language, these are related to each other as $\mu(k) = a^{-1}(k)$, where k is the Laplace variable representing inverse time [40, 41, 104, 103].

3.3.1 Chain tension perpendicular to the membrane

Measuring chain tension directly is difficult. We therefore use a method developed earlier [103, 104] to monitor the chain tension near the pore.

By definition, the chain tension imbalance $\phi(t)$ is the difference of the chain tensions on the right and the left side of the pore: $\phi(t) = \Phi_{\rm R}(E,t) - \Phi_{\rm L}(E,t)$. Both $\Phi_{\rm R}(E,t)$ and $\Phi_{\rm L}(E,t)$ are functions of the applied electric field E across the pore. Note, from the applied potential (3.1), that the field E acts on the monomers at site $z=-1$ towards the pore, while it acts on those at site $z=1$ away from the pore. Using the convention that $E<0$ (resp. $E>0$) implies a field acting towards (resp. away from) the membrane, we have

$$\Phi(E, t=0) = \begin{cases} \Phi_{\rm L}(t=0) & (E<0) \\ \Phi_{\rm R}(t=0) & (E>0) \end{cases} . \tag{3.3}$$

Now consider a different problem, where one end of a polymer is tethered to a fixed membrane, yet monomers are allowed to spontaneously enter or leave the tethered end, under the effect of an electric field E. Then, following the methodology described in Refs. [103, 104], we have

$$\Phi(E, t=0) = k_B T \, \ln \frac{P_+}{P_-} , \tag{3.4}$$

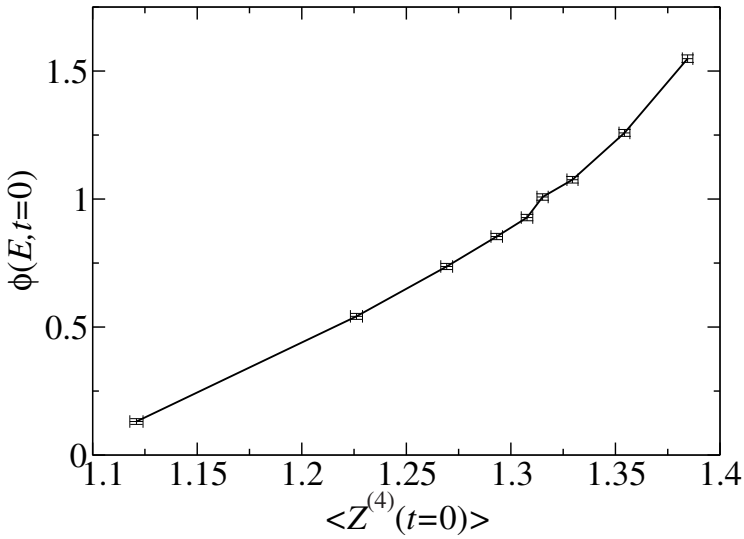

Figure 3.2: $\langle Z^{(4)}(t=0)\rangle$ vs. $\Phi(E,t=0)$, for $N/2=200$ and electric field values $E=-0.5,-0.25,-0.1,-0.05,0,0.05,0.1,0.25$, and 0.5 respectively. The angular brackets for $\langle Z^{(4)}(t=0)\rangle$ indicate an average over $32,000$ polymer realizations, which are also used to obtain $\Phi(E,t=0)$.

where P_- (resp. P_+) is the probability that the left (or the right) polymer segment has one monomer less (resp. one extra monomer).

Note that even for $E=0$, as already stressed in Ref. [104], there is nonzero chain tension Φ_0 at the pore, due to the presence of the membrane. A polymer's free energy close to a membrane is higher than its free energy in bulk. In other words, the membrane repels the polymer, and as a result, for a polymer with one end tethered to a membrane, the monomers close to the membrane are more stretched than they would be in the bulk.

For a translocating polymer Eq. (3.4) cannot be used, so to compute $\Phi_{\rm R}(t)$ and $\Phi_{\rm L}(t)$ one needs a suitable proxy. In the cases of unbiased translocation [40, 41], translocation with a pulling force [104] and translocation out of planar confinements [103], it has been shown that the center of mass distance of the first few, say 4 to 5 monomers from the membrane provides an excellent proxy for Φ. In this chapter we follow the same line. The average distance $\langle Z^{(4)}(t=0)\rangle$ is plotted as a function of the chain tension $\Phi(E,t=0)$ for various values of E in Fig. 3.2. This figure shows that under an applied field,

$\Phi(E, t = 0)$ is a reasonably linear function well proxied by $Z^{(4)}$. The positive curvature seen in Fig. 3.2, i.e., the deviation from linearity, is seen only for $E > 0$. We believe that this is partly due to the saturation of $Z^{(4)}$. [By definition, in our lattice model the distance of the center of mass of the first 4 monomers from the membrane cannot exceed $(1 + 2 + 3 + 4)/4 = 2.5$.]

3.3.2 Memory effects in the chain tension

From Eq. (3.2), the behavior of the memory kernel $\mu_{\rm R}(t)$ for the polymer segment on the right side of the membrane can be obtained with a sudden introduction of p extra monomers at the pore, corresponding to an instantaneous current $v(t) = p\delta(t)$. Physically, $v(t) = p\delta(t)$ with $p > 0$ (resp. $p < 0$) means that we tether a polymer of length N halfway through the pore at $t \to -\infty$, let it thermalize till $t = 0$, and then introduce p extra monomers at the tethered end of the right (resp. left) segment at $t = 0$. We then ask for the time-evolution of the mean response $\langle \delta\Phi_{\rm R}(t) \rangle$, where $\delta\Phi_{\rm R}(t)$ is the shift in chemical potential for the right segment of the polymer at the pore. This means that for the translocation problem (with both right and left segments), we would have $\phi(t) = \delta\Phi_{\rm R}(t) - \delta\Phi_{\rm L}(t)$, where $\delta\Phi_{\rm L}(t)$ is the shift in chemical potential for the left segment at the pore due to an opposite input current to it.

In earlier works [40, 41], using $v(t) = p\delta(t)$ for a polymer of length N tethered halfway at the pore as described in the above paragraph, Panja *et al.* showed that for unbiased polymer translocation, i.e., for $E = 0$, this mean response, and hence $\mu(t)$ takes the form $\mu(t) \sim t^{-\alpha} \exp[-t/\tau_{\rm R}(N/2)]$, with $\alpha = (1 + \nu)/(1 + 2\nu)$ and $\tau_{\rm R}(N/2) \sim (N/2)^{1+2\nu}$ [note that for $E = 0$ there is a trivial symmetry between the right and the left segment of the polymer, hence $\mu_{\rm R}(t) = \mu_{\rm L}(t) \equiv \mu(t)$]. For the benefit of smooth reading, we repeat the argument below. The terminal exponential decay $\exp[-t/\tau_{\rm R}(N/2)]$, with $\tau_{\rm R}(N/2) \sim (N/2)^{1+2\nu}$ is expected from the relaxation dynamics of the entire right segment (of length $N/2$) of the polymer [40]. To understand the physics behind the exponent α, we use the well-established result for the relaxation time t_n for n self-avoiding Rouse monomers scaling as $t_n \sim n^{1+2\nu}$. Based on the expression of t_n, we anticipate that by time t the extra monomers will be well equilibrated across the inner part of the chain up to $n_t \sim t^{1/(1+2\nu)}$ monomers from the pore, but not significantly further. This internally equilibrated section of $n_t + |p|$ monomers extends only $r(n_t) \sim n_t^{\nu}$, less than its equilibrated value $(n_t + |p|)^{\nu}$, because the larger-scale conformation has yet to

adjust: the corresponding compressive force from these $n_t + |p|$ monomers is expected by standard polymer scaling [2] to follow $f/(k_BT) \sim \delta r(n_t)/r^2(n_t) \sim \nu|p|/[n_t r(n_t)] \sim t^{-(1+\nu)/(1+2\nu)}$, for $|p| \ll n_t$. This force f must be transmitted to the membrane, through a combination of decreased tension at the pore and increased incidence of other membrane contacts. The fraction borne by reducing chain tension at the pore leads us to the inequality $\alpha \geq (1+\nu)/(1+2\nu)$. It seems unlikely that the adjustment at the membrane should be disproportionately distributed between the chain tension at the pore and other membrane contacts, leading to the expectation that the inequality above is actually an equality.

When the electric field is applied at the pore, and the same monomer injection method is used to probe the memory kernels $\mu_{\rm R}(t)$ and $\mu_{\rm L}(t)$, we expect the above arguments to hold again: since the field is applied very locally at the base of the tethered polymer segments, it does not destroy the broader structure of the polymer. However, we do expect to see deviations from the $t^{-\alpha}\exp[-t/\tau_{\rm R}(N/2)]$ at short times. Indeed, we have confirmed this picture — for various field strengths we tracked $\langle\delta\Phi_{\rm R}(t)\rangle$ and $\langle\delta\Phi_{\rm L}(t)\rangle$ by measuring the distance of the average center of mass of the first 4 monomers from the membrane, $\langle Z^{(4)}(t)\rangle$, in response to the injection of extra monomers near the pore at $t = 0$. Specifically we consider the equilibrated right and left segments of the polymer, each of length $N/2 = 200$ (with the middle monomer threaded at the pore), adding 5 extra monomers at the tethered end of the right and the left segment each at $t = 0$, corresponding to $|p| = 5$, bringing the length of each segment up to $N/2 + |p|$. Using the proxy $\langle Z^{(4)}(t)\rangle$ for both segments we then track $\langle\delta\Phi_{\rm R}(t)\rangle$ and $\langle\delta\Phi_{\rm L}(t)\rangle$, denoting them by values $E > 0$ and $E < 0$ respectively in Fig. 3.3. The deviations from the expected power-law $t^{-(1+\nu)/(1+2\nu)}$ at short times and the $\exp[-t/\tau_{\rm R}(N/2)]$ at long times makes the precise identification of the power-law $t^{-(1+\nu)/(1+2\nu)}$ difficult. Nevertheless, there is an extended regime where this power-law can be identified reasonably clearly, yielding us $\mu_{\rm R}(t) = \mu_{\rm L}(t) \equiv \mu(t) = t^{-(1+\nu)/(1+2\nu)}\exp[-t/\tau_{\rm R}(N/2)]$.

3.4 Scaling behavior of $\tau_{\rm d}$ with N

The memory kernel we obtained in Sec. 3.3 can be termed as the "static memory kernel", as it is obtained under the condition that before the injection of the extra monomers both segments were thermalized. When the applied field is not too strong, we can expect the static memory kernel to yield the

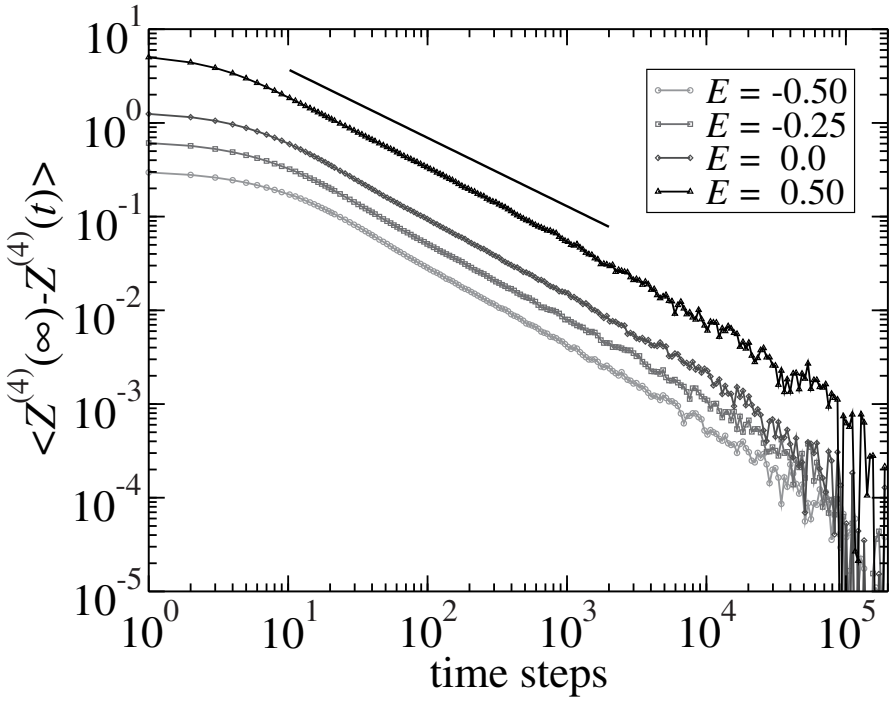

Figure 3.3: Probing the memory kernels by $\langle Z^{(4)}(\infty) - Z^{(4)}(t)\rangle$ following monomer injection at the pore corresponding to $v(t) = p\delta(t)$, with $|p| = 5$. Physically, $v(t) = p\delta(t)$ with $p > 0$ (resp. $p < 0$) means that we tether a polymer of length N halfway at the pore at $t \to -\infty$, let it thermalize till $t = 0$, and then introduce $|p|$ extra monomers to the right (resp. left) segment at the tether point at $t = 0$. Following our notation in Eq. (3.3), the $E < 0$ data (resp. $E > 0$ data) correspond to $\mu_{\rm L}(t)$ [resp. $\mu_{\rm R}(t)$]. The data presented correspond to an average over $500,000$ polymer realizations, with $N/2 = 200$. The steeper drop at longer times corresponds to the exponential decay $\exp[-t/\tau_{\rm R}(N/2)]$. For clarity, the data have been multiplied by 16^E. The solid line corresponds to the power-law $t^{-\frac{1+\nu}{1+2\nu}} \approx t^{-0.73}$.

scaling of translocation velocity with time, in the following manner.

An inverse Laplace transform of Eq. (3.2) yields us

$$v(k) = \frac{\phi_{t=0}}{k\mu(k)} \; - \; \frac{\phi(k)}{\mu(k)}\,, \tag{3.5}$$

where k is the Laplace variable representing inverse time. Thereafter, using the power-law part of $\mu(t) \sim t^{-(1+\nu)/(1+2\nu)}$, i.e., $\mu(k) \sim k^{(1+\nu)/(1+2\nu)-1}$, and Laplace-inverting Eq. (3.5), we get

$$v(t) = \int_0^t \mathrm{d}t' \, (t-t')^{-\frac{1+3\nu}{1+2\nu}} \, \left[\phi_{t=0} - \phi(t')\right] \, . \tag{3.6}$$

If $\phi(t)$ goes to a constant[2] $\neq \phi_{t=0}$, then Eq. (3.6) reduces to

$$v(t) \sim t^{-\frac{\nu}{1+2\nu}}, \text{ i.e., } s(t) = N/2 + \int_0^t \mathrm{d}t' \, v(t') \sim t^{\frac{1+\nu}{1+2\nu}}, \tag{3.7}$$

where $[s(t) - N/2]$ is the distance unthreaded after time t; the $N/2$ appears in Eq. (3.7) as $s(0) = N/2$.

In Fig. 3.4 we show the behavior of $[\phi_{t=0} - \phi(t)]$ by means of the proxy variable $\langle z^{(4)}(0) - z^{(4)}(t)\rangle$ for $E = V = 0.05$, 0.15, and 0.25 respectively, where $z^{(4)}$ is the difference between the $Z^{(4)}$ values of the right and left segment of the polymer, i.e., $z^{(4)}(t) = Z^{(4)}_{\mathrm{R}}(t) - Z^{(4)}_{\mathrm{L}}(t)$. Indeed the quantity $[\phi_{t=0} - \phi(t)]$ approaches a constant rather quickly. We also note that the relation between this constant and the applied field E is almost linear.

For strong fields, there is no *a priori* reason that the dynamics can still be described by the static memory kernel instead of a suitably replacing "dynamic memory kernel", but we find that the scaling $s(t) \sim t^{(1+\nu)/(1+2\nu)}$ is obeyed for fairly strong fields as well: in Fig. 3.5 we plot the average time $\langle t\rangle$ to unthread a distance s to show this scaling. Note the strong finite-size effects for the scaling behavior as shown by the deviation from the $t^{(1+2\nu)/(1+\nu)}$ for larger values of s. The presence of such strong finite-size effects indicates that without the aid of $s(t)$ vs. t curves, determining the scaling of τ_{d} with N will almost certainly lead to erroneous identification of the scaling laws —

[2]Note that with $\phi(t)$ a constant, strictly speaking, the integral (3.6) does not converge. The divergence stems from the assumption that $\mu(t) \sim t^{-(1+\nu)/(1+2\nu)}$ holds all the way to $t \to 0$. This is clearly not true as can be seen from Fig. 3.3, which provides the required cutoff for the convergence of the integral (3.6).

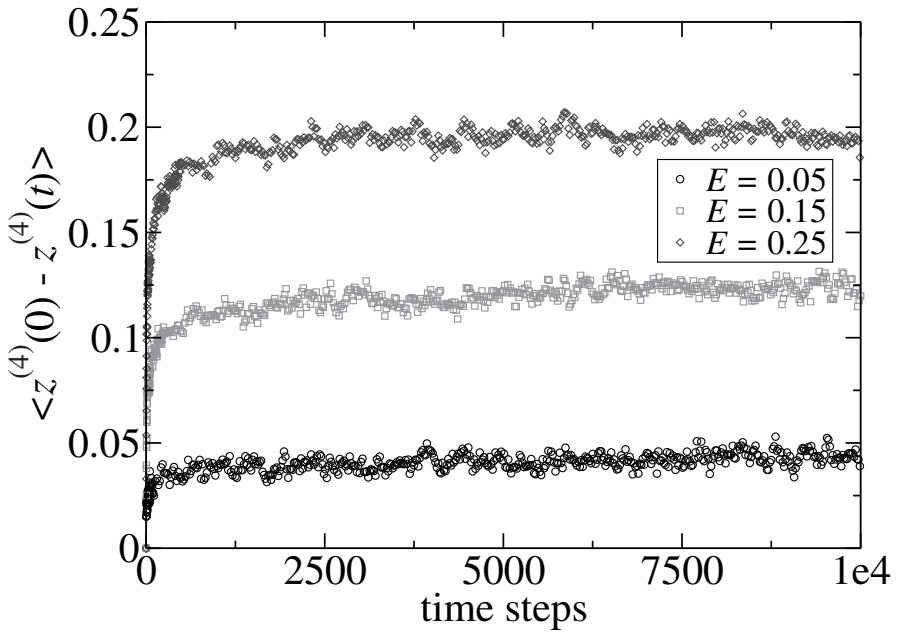

Figure 3.4: Behavior of $[\phi_{t=0} - \phi(t)]$ for $N = 200$ as a function of t, shown by means of the proxy variable $< z^{(4)}(0) - z^{(4)}(t) >$. demonstrating that $[\phi_{t=0} - \phi(t)]$ reduces to a constant very quickly: $E = 0.05$ (circles), 0.15 (squares), and 0.25 (triangles). To generate these averages 16,000 individual polymers were unthreaded for each value of E.

we believe that these finite-size effects are responsible for the wide range of existing numerical scaling results, as summarized in Table 3.1. Nevertheless, Fig. 3.5 shows that these finite-size effects do not increase linearly with N, leading us to the scaling for $\tau_{\rm d}$ as

$$\tau_{\rm d} \sim N^{(1+2\nu)/(1+\nu)}/E\,, \tag{3.8}$$

which is obtained from the condition that $s(\tau_{\rm d}) = N$. For the above analysis to hold, the dwell time must be less than $\tau_{\rm R}$, which Eq. (3.8) confirms. Note that the E-dependence of Eq. (3.8) is only numerically obtained from Fig. 3.5. Note also that the curves in Fig. 3.5 for $E = 0.05$ tend to 'sag' a bit. We attribute this to our numerically inspired definition of $s(\langle t\rangle)$, as the mean time

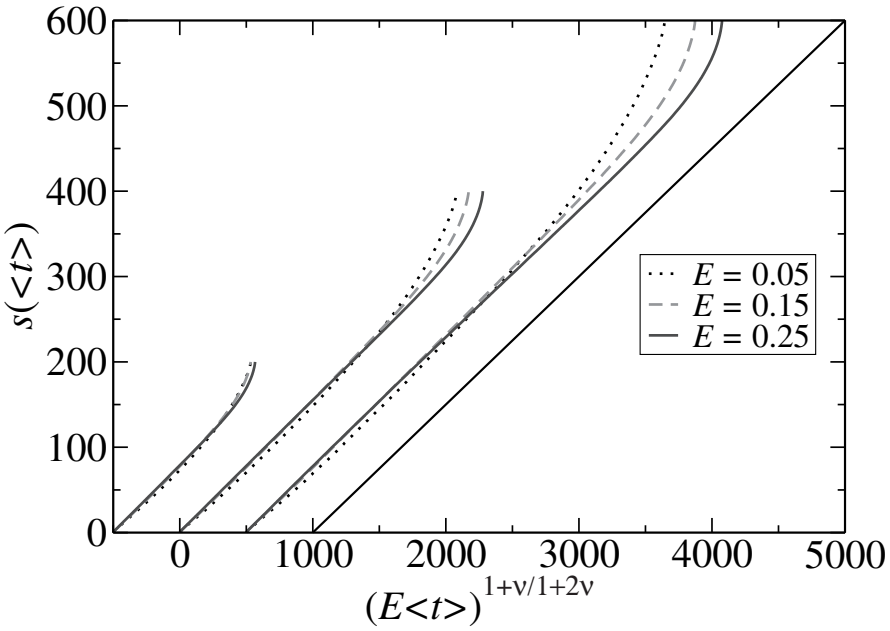

Figure 3.5: The average time $\langle t \rangle$ to unthread distance s for three different field strengths, for $N = 400$ (average over 16,000 polymer realizations for each field), $N = 800$ (average over 16,000 polymer realizations for each field), and $N = 1,200$ (5,000 polymer realizations for $E = 0.05$, and 7,500 polymer realizations each for $E = 0.15$ and $E = 0.25$). The data for $N = 800$ correspond to real-time values, while the data for $N = 400$ and $N = 1,200$ have been shifted by ∓ 500 units along the x-axis for clarity. The solid line has been added as a guide to the eye.

to unthread a distance s, as opposed to, e.g., the numerically less favorable measure of distance $\langle s(t) \rangle$, i.e.,the monomer which is most likely to reside in the pore at time t. At small fields, the polymer has ample time for fluctuations, pushing the time of first arrival up. Numerically, for $E = 0.15$ and 0.25, the exponent $\partial[\log s\,(t)]/\partial(\log t)$ is found to be 0.73 ± 0.02, in agreement with the theoretical value $(1+\nu)/(1+2\nu)$. The sagging and finite-size effects discussed above cause the apparent exponent to be slightly larger, ranging from 0.74 to 0.79, for $E = 0.05$.

With decreasing field strength, especially in the range where the thermal fluctuations are comparable to the work done by the field to translocate the

entire polymer, given by $EN \simeq k_BT = 1$, one should obtain a crossover from the above scaling (3.8) to $\tau_{\rm d} \sim N^{2+\nu}$ for unbiased translocation [40, 41]. This suggests that if $\tau_{\rm d}/N^{2+\nu}$ is plotted as a function of EN, then one should obtain a scaling collapse; i.e., there exists a scaling function f such that $\tau_{\rm d} = N^{2+\nu} f(EN)$. However, EN as a scaling variable is simply numerically inconsistent with Fig. 3.5 and Eq. (3.8). Instead $\tau_{\rm d} = N^{2+\nu} f(E, N)$ is the proper description of the situation, with $f(E, N)$ approaching a constant for $E \to 0, N \to \infty$ and $f(E, N)$ behaving as $E^{-1} N^{-\nu-1/(1+\nu)}$ for $E \sim O(1)$ and $N \to \infty$. Note that E in this paragraph should be interpreted as the dimensionless quantity $qV/(k_BT)$.

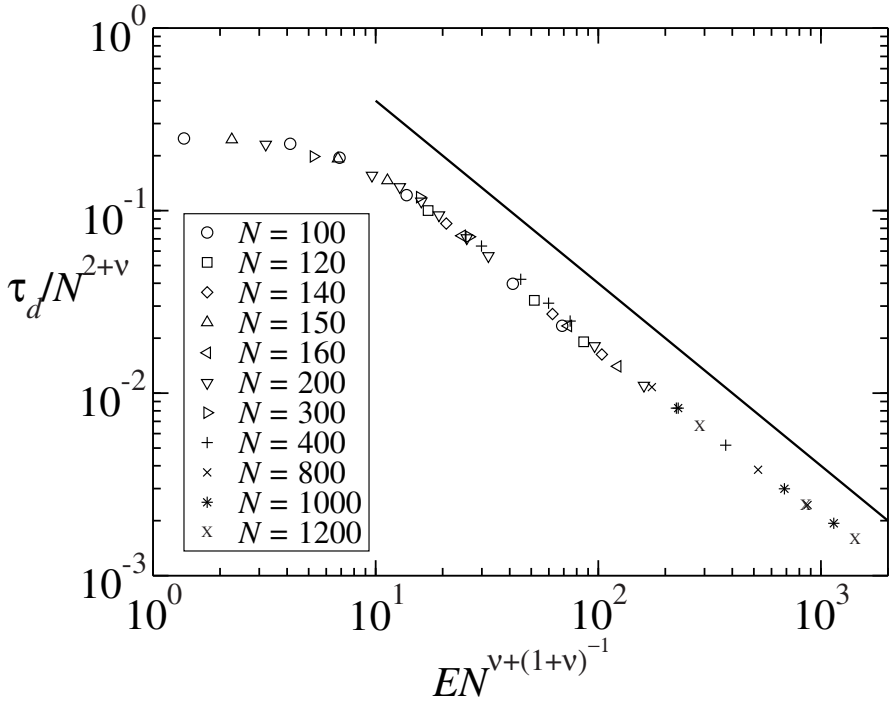

Figure 3.6: Data collapse in terms of $x = EN^{\nu+1/(1+\nu)}$ and $y = \tau_{\rm d}/N^{2+\nu}$ for various fields. The mean unthreading times $\tau_{\rm d}$ have been obtained with at least $1,000$ polymers (up to $32,000$ for smaller N values) for each N and field strength E: the statistical error bars are smaller than the size of the symbols. The solid line $y \sim 1/x$ for moderate field strengths in support of Eq. (3.8) has been added as a guide to the eye.

To demonstrate the scaling behavior of Eq. (3.8) for $E \simeq O(1)$, we plot $\tau_{\rm d}/N^{2+\nu}$ as a function of $EN^{\nu+1/(1+\nu)}$ in Fig. 3.6. Keeping in mind that this way of plotting the data does not necessarily yield a data collapse at small but nonzero E, as discussed above, we also plot several data points for small E, in order to demonstrate that for $E \to 0$ our results in this chapter are consistent with that of unbiased translocation [40, 41, 103].

3.5 Discussion

We have studied polymer translocation in three dimensions through a narrow pore in an otherwise impenetrable membrane, as the polymer is driven by a field E across the pore. The polymer performs Rouse dynamics, i.e., we considered polymer dynamics in the absence of hydrodynamic interactions. We found that the typical time the pore remains blocked during a translocation event, for moderate field strengths scales as $\sim N^{(1+2\nu)/(1+\nu)}/E$, where $\nu \simeq 0.588$ is the Flory exponent for the polymer. In line with Refs. [40, 41, 104, 103], we showed that this scaling behavior stems from the polymer dynamics at the immediate vicinity of the pore — in particular, the memory effects in the polymer chain tension imbalance across the pore. We also showed that our results in this chapter are consistent with that of unbiased translocation [40, 41, 103] in the limit $E \to 0$.

The above results for finite E, along with the numerical results by several other groups, violate the lower bound $\sim N^{1+\nu}/E$ suggested earlier in the literature [32]. We also discussed why this lower bound is incorrect and showed, based on conservation of energy, that a correct lower bound for the pore-blockade time for field-driven translocation is given by $\eta N^{2\nu}/E$, where η is the viscosity of the medium surrounding the polymer. Our theoretical analysis has been supported by high precision computer simulation data, generated with a three-dimensional self-avoiding lattice polymer model.

Having worked out the physics of field-driven polymer translocation in the absence of hydrodynamic interactions, it is worthwhile to reflect on the scaling of pore-blockade times as a function of the polymer length N in the presence of hydrodynamic interactions. Hydrodynamical interactions will modify the memory kernel $\mu(t)$ — changing it from $t^{-(1+\nu)/(1+2\nu)} \exp(-t/\tau_{\rm R})$ to $t^{-(1+\nu)/(3\nu)} \exp(-t/\tau_{\rm Zimm})$ [40, 41], where $\tau_{\rm Zimm}$ is the Zimm relaxation time, scaling as $N^{3\nu}$ in good solvent for a polymer of length N. This implies that the pore-blockade time will behave as $N^{3\nu/(1+\nu)}$ under the influence of hydro-

dynamic interactions. In this context we note that the scaling of the pore-blockade time has been experimentally measured to scale as $N^{1.26\pm0.07}$ [99]. In the scaling limit $3\nu/(1+\nu) \simeq 1.11$. The value for ν suggested in Ref. [99] is 0.611 ± 0.016, for which $3\nu/(1+\nu) \simeq 1.14\pm0.02$, a bit closer to 1.26 ± 0.07. For a physical explanation of the scaling of the pore-blockade times with polymer length, the authors of Ref. [99] arrived at an answer 2ν using a macroscopic view of the translocating polymer, assuming that the translational velocity of the center of mass of the untranslocated part is constant in time, and (implicitly) that the memory kernel is a δ-function. Our analysis in this chapter, as well as in Refs. [40, 41, 104, 103] based on memory effects, therefore, casts serious doubts on the physical interpretation of Ref. [99]: it has been repeatedly shown that the velocity of translocation is not uniform in time, and the same part of the polymer visits the pore a large number of times. Although so far our work has not incorporated hydrodynamic interactions explicitly, it is difficult to imagine that introducing hydrodynamic interactions will mysteriously wipe out the entire memory effects in the polymer.

Chapter 4

Polymer translocation in crowded environments

Abstract

We study unbiased translocation of a polymer with length N, surrounded by equally long polymers, through a narrow pore in a static membrane. We show that in dense polymeric systems a relaxation time exists that scales as $\tau_{\beta_t} \sim N^{\beta_t} = N^{2.65\pm0.2}$, much longer than the Rouse time $\tau_R \sim N^2$. If the polymers are well entangled, we find that the mean dwell time scales as $N^{3.33\pm0.1}$, while for shorter, less entangled polymers, we measure dwell times scaling as $N^{2.7\pm0.1}$. These conclusions are supported by a detailed investigation of the average squared displacement of the reaction coordinate $s(t)$, which is the number of the monomer in the pore at time t.

4.1 Introduction

Molecular transport across cell membranes is an essential mechanism in biological systems. The molecules are often too long, compared with the narrow pores in the membranes, to allow them to pass through as a single unit. In such circumstances, the molecules have to deform themselves in order to translocate through the pores. DNA, RNA and proteins are such naturally occurring long molecules [84, 85, 86, 87, 88] in a variety of biological processes. Technological applications include gene therapy [89, 90], the delivery of drug molecules to their activation sites [91], and as a fast, simple, cheap

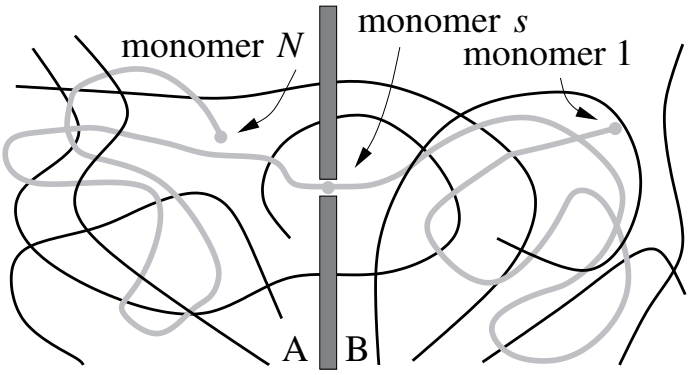

Figure 4.1: Snapshot of a translocating polymer in a two-dimensional projection of our three-dimensional system. The number of the monomer of this translocating polymer that is located within the pore is labeled s. The surrounding polymers do not feel the membrane.

technology for single macromolecule detection, analysis and characterization [93, 94, 95, 96, 97, 98, 92].

As discussed in chapters 1.4 and 3.1, a considerable amount of theoretical work has focused on many aspects of polymer translocation. However, a notable exception is the translocation of polymers in crowded environments. We are only aware of one recent study by Gopinanthan and Woon Kim [110]. In their theory polymer translocation is described by a Fokker-Planck equation for first passage over the entropic barrier: they explicitly treat the process quasistatically with the polymer segments on both sides being in equilibrium at all times.

Crowding is of significant importance, not only from a theoretical point of view. In the cellular cytoplasm macromolecular aggregates and other inclusions can occupy a volume fraction as high as 50% [111]. Crowding is known to have a significant influence on reaction rates, protein folding rates, and equilibria *in vivo* [112, 113].

The subject of this chapter is a detailed investigation of unbiased polymer translocation through a narrow pore in an otherwise impenetrable membrane in a crowded environment. The surrounding environment is composed of polymers that are indistinguishable from the translocating polymer, except they can cross the membrane at any point. For the majority of the experiments

we impose a monomer density c equal to one monomer per lattice site in our periodic simulation box. Figure 4.1 provides a schematic two-dimensional projection of our three-dimensional system.

Our interest is in the scaling behavior of the typical pore-blockade time $\tau_{\rm d}$ during a translocation event, as a function of polymer length N. From extensive simulations with the extended repton model of chapter 2.3, we characterize the anomalous dynamics of translocation. From it, we infer that for long well-entangled polymers the dwell time scales as $\tau_{\rm d} \sim N^{2+\beta_{\rm t}/2} = N^{3.33\pm0.1}$, with $\beta_{\rm t} = 2.65 \pm 0.2$; this is an exponent that indicates the presence of a long relaxation time scale proportional to $N^{\beta_{\rm t}}$, exceeding the Rouse time $\tau_{\rm R} \sim N^2$ of a single polymer with ideal statistics. A very similar relaxation time scale has recently been proposed by Barkema and Panja [23]. However, for shorter, and therefore less entangled polymers, we measure $\tau_{\rm d} \sim N^{2.7\pm0.1}$. We propose several relaxation mechanisms that may contribute to the latter result.

4.2 Average squared displacement

The important time scales for translocation can be obtained from the time evolution of the reaction coordinate $s(t)$, which is the number of the monomer s in the pore at time t. In our study of field-driven polymer translocation, chapter 3, we used the numerically inspired definition $s(\langle t \rangle)$, as the mean time to unthread over a distance s. However, for unbiased polymer translocation $\langle s(t) \rangle$ is zero. To explore the temporal behavior of the reaction coordinate, we determine $P_N(s_1, s_2, t)$, the probability distribution that at time t' a polymer of length N is in a configuration for which monomer s_1 is threaded in the pore, and the polymer evolves at time $t' + t$ into a configuration in which monomer s_2 is threaded in the pore.

First, as is detailed in [40], we recall some of the properties of these probability distributions for various values of s_1, s_2, t, and N. As long as neither s_1 nor s_2 are too close to the end of the polymer, $P_N(s_1, s_2, t)$ depends only on $s_2 - s_1$. It was shown that $P_{400}(s_1, s_2, t)$ for the unbiased translocation of the single chain was near indistinguishable for $s_1 = N/4$, $N/2$ and $3N/4$ at $t = 100$ time units, with the ratio of reptation to Rouse moves set to $R = 10$ (see chapter 2.3), a choice which is respected in this chapter too. Also, it was shown that the distributions are well fitted with Gaussians.

Given these two observations we construct distributions $P_N(s_2 - s_1, i\Delta t)$ from the conservative requirement $17N/40 < s_2 < 24N/40$ (while for $N =$

2197, for reasons of computational efficiency, $N/4 < s_2 < 3N/4$). Further memory constraints limit the number of distributions $P_N(s_2 - s_1, i\Delta t)$ to fixed time intervals Δt, with Δt ranging from tens to thousands of time steps, and $1 < i \leq 5000$. We define $\langle s_N^2(t)\rangle$ to be the variance of the distribution $P_N(s,t)$, which is straightforwardly found from a fit to a Gaussian. Henceforth, we omit the subscript N, unless it is in the interest of clarity.

From studies of unbiased polymer translocation of a single chain [40, 41], it is observed, and theoretically understood, that

$$\langle s^2(t)\rangle \sim t^{\frac{1+\nu}{1+2\nu}}, \tag{4.1}$$

with ν the Flory exponent. (Similarly, for field-driven polymer translocation we derived Eq. (3.7).) We have argued in chapter 1, and observed in figure 2.7, that polymers in a dense environment of other polymers show ideal scaling, i.e, $\nu = 1/2$. Hence, at least at short time scales, but long enough to observe ideal scaling, we expect for polymer translocation in a crowded environment[1]

$$\langle s^2(t)\rangle \sim t^{\frac{3}{4}}. \tag{4.2}$$

In figure 4.2, we plot the deviations from scaling Eq. (4.2), as a function of time, for various polymer lengths. The polymers forming the environment do not interact with the wall. At short times, the squared displacement does not depend on polymer length: the curves for different polymer lengths fall on top of each other. However, (starting with the shortest polymer length) there is a time $\tau_{\beta_t}(N)$ after which the average squared displacement increases much faster. We attribute this upward trend to the loss of memory beyond this time $\tau_{\beta_t}(N)$: if the squared displacement is measured over a time much longer than $\tau_{\beta_t}(N)$, the dynamics consists of many uncorrelated events, and thus will result in normal diffusive behavior.

Another feature shown in figure 4.2 is that the curves for the longer polymers, $N = 512$ or more, show a downward trend well before they show the upward trend discussed above. We associate this crossover to a slower anomalous diffusion with the onset of entanglement. Once the polymers are sufficiently long, sideways motion is hindered and longitudinal motion (reptation) remains. As argued by Doi and Edwards [11], the dynamics of entangled polymers is characterized by $\langle s^2(t)\rangle \sim t^{1/2}$. Indeed, a function $g(t)$ satisfying an

[1] Note that monomers that are near each other along the backbone show nonideal scaling, as is witnessed in figure 2.7. Therefore it cannot be ruled out that initially, e.g., the first few hundreds of time steps, $\langle s^2(t)\rangle \sim t^{0.73}$.

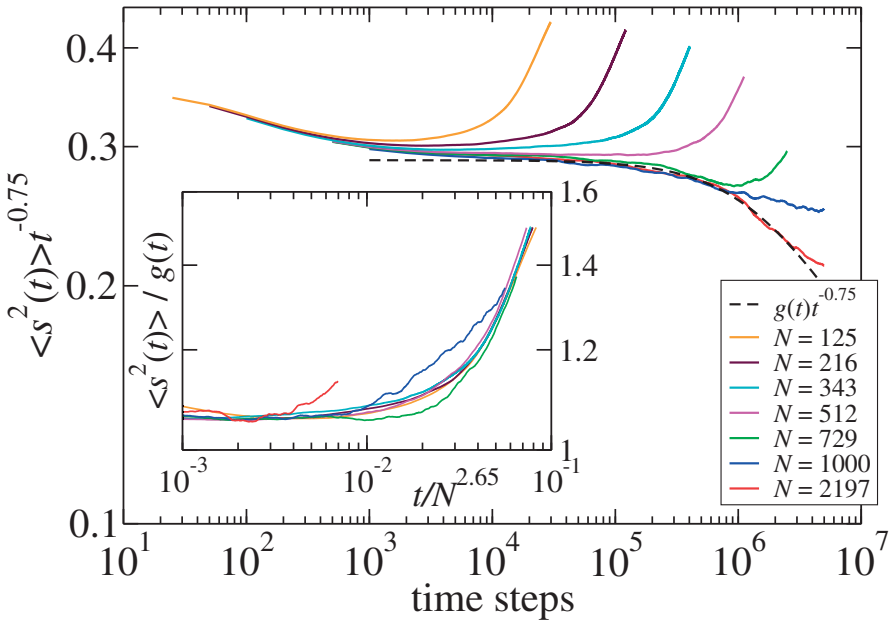

Figure 4.2: Double-logarithmic plot of the mean squared displacement of the reaction coordinate $\langle s^2(t)\rangle$ divided by $t^{0.75}$, as a function of time t, for polymer lengths $N =$ 125, 216, 343, 512, 729, 1000, and 2197, all at monomer density $c = 1$, and with 512 polymers in the periodic box, of which 16 are threaded in 16 equidistant pores in the wall. The surrounding polymers do not feel the presence of the wall. The curves for $N \geq 512$ required around 10,000 processor hours. The dashed line is a fit of $\langle s^2_{2197}(t)\rangle$ with $t < 3 \cdot 10^6$ to Eq. (4.3) from which $a = 2.92 \cdot 10^2$, $b = 2.91 \cdot 10^{-5}$, and $q = 4.56$. Inset: t/N^{β_t} versus $c_N\langle s^2(t)\rangle/g(t)$, with $\beta_t = 2.65$. The constant c_N introduces a small arbitrary vertical shift of the curves to improve the data collapse; it ranges from 1 for the shortest, to 1.07 for the longest polymers.

equation of the form

$$g(t)^{-q} = a \cdot \left(t^{3/4}\right)^{-q} + b \cdot \left(t^{1/2}\right)^{-q} \tag{4.3}$$

fits the data quite well. This function features a crossover from power-law behavior with exponent 3/4 at short times, to power-law behavior with an exponent of 1/2 at long times; the fitting parameter q tunes the "sharpness" of the crossover.

To quantify the scaling of the time τ_{β_t} after which the mean squared displacement increases linearly with time, we show $\langle s^2(t')\rangle / g(t')$ in the inset of figure 4.2, in which $t' \equiv t/N^{\beta_t}$ is rescaled time. As shown in this inset, the curves for various polymer lengths seem to deviate from $g(t')$ more or less at the same time, provided that time is rescaled by a factor of N^{β_t} with $\beta_t = 2.65 \pm 0.2$.

According to standard reptation theory (see chapter 1.3.2) the longest relaxation time scale in the tube is given by $\tau_R \sim N^2$, and our value for β_t is thus at odds with this finding. In a recent work, Barkema and Panja [23] measure, using the extended repton-model, the dynamical response of the surrounding polymers in a dense environment, also at monomer density $c = 1$. They observe a longest relaxation time scale "in the tube" $\tau_\beta \sim N^\beta$, with $\beta = 2.56 \pm 0.1$. The numerical closeness of β and β_t suggests a strong physical connection. We expect, in the limit $N \to \infty$, that β_t and β become identical.

4.3 Dwell times

In the simulations used to compute the average squared displacement, we also monitor each event in which the polymer is completely located on side A or B, except for the first or last monomer, both of which are not allowed to leave the pore. If at some time t_A the polymer is located on side A, and at a later time t_B is it located on side B, while between these times it has segments on both sides, then $t_B - t_A$ is an estimate for the dwell time, cf. figure 1.9. The resulting estimates for the dwell time, for various polymer lengths, are reported in Table 4.1.

The dwell time τ_d can also be obtained from the average squared displacement, using

$$\langle s^2(\tau_d)\rangle \sim N^2. \tag{4.4}$$

In the inset of figure 4.2 it was shown that the average squared displacement is well described by the function $g(t)$ in Eq. (4.3), until $\tau_t \approx f \cdot N^{\beta_t}$ with

N	$\tau_{\mathrm{d}}(c=1)$	$\tau_{\mathrm{d}}(c=1/8)$	single chains
50	$8.94(1)\cdot10^3$		
64	$1.80(1)\cdot10^4$		
125	$1.14(1)\cdot10^5$	$1.94(1)\cdot10^5$	
216	$5.00(2)\cdot10^5$	$8.21(5)\cdot10^5$	$8.62(5)\cdot10^5$
343	$1.77(1)\cdot10^6$	$2.72(2)\cdot10^6$	$2.87(2)\cdot10^6$
512	$5.46(6)\cdot10^6$	$7.57(8)\cdot10^6$	$8.06(9)\cdot10^6$
729	$1.57(4)\cdot10^7$	$1.92(3)\cdot10^7$	$1.94(2)\cdot10^7$
1000	$4.1(2)\cdot10^7$		$4.3(1)\cdot10^7$

Table 4.1: Average dwell times, for various polymer lengths, and monomer densities $c = 1$, $c = 1/8$, and in an box void of other polymers. The uncertainty is given by the mean standard error of the dwell times. To these estimates of the dwell times a correction of at most 10% is added to compensate for truncation effects.

$\beta_{\mathrm{t}} = 2.65$, and $f = 0.056$, after which it is expected to increase linearly in time. Hence, we expect

$$\langle s^2(t)\rangle = \frac{t}{\tau_{\beta_{\mathrm{t}}}} \cdot \langle s^2(\tau_{\beta_{\mathrm{t}}})\rangle = \frac{t}{\tau_{\beta_{\mathrm{t}}}} g(\tau_{\beta_{\mathrm{t}}}). \tag{4.5}$$

Combining Eqs. (4.4) and (4.5), an estimate for the dwell times is given by

$$\tilde{\tau}_{\mathrm{d}} \sim N^2 \frac{\tau_{\beta_{\mathrm{t}}}}{g(\tau_{\beta_{\mathrm{t}}})} \sim \frac{N^{2+\beta_{\mathrm{t}}}}{g(fN^{\beta_{\mathrm{t}}})}. \tag{4.6}$$

Figure 4.3 compares this estimated dwell time with the measurements in Table 4.1 at monomer density $c = 1$. The agreement, for polymers with $N = 216$ or more, is quite good, indicating that our assumptions, specifically the linear extrapolation of the average squared displacement beyond $\tau_{\beta_{\mathrm{t}}}$, is valid.

Based on Eq. (4.5), the dwell time shows a crossover behavior as a function of polymer length. For short polymers, the first term in Eq. (4.3) dominates, and $g(\tau_{\beta_{\mathrm{t}}}) \sim N^{\beta_{\mathrm{t}}\cdot 3/4} = N^{1.99\pm0.15}$. In this regime, the dwell time scales with polymer length as $\tau_{\mathrm{d}} \sim N^{2+(1-3/4)\beta_{\mathrm{t}}} = N^{2.66\pm0.05}$. Once the polymers are sufficiently long, the second term in Eq. (4.3) dominates, and $g(\tau_{\mathrm{d}}) \sim N^{\beta_{\mathrm{t}}\cdot 1/2} = N^{1.33\pm0.1}$. In this regime, we predict the dwell time to scale with

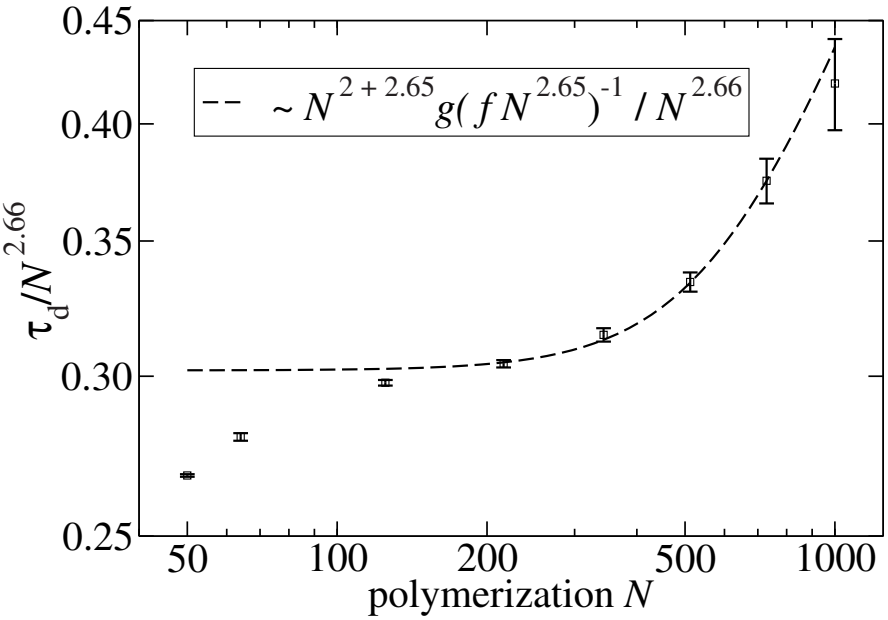

Figure 4.3: The dwell times at monomer density $c = 1$ rescaled to show the deviations from the (no-entanglement) scaling $N^{2.66}$. The dashed line is a fit of the average dwell times to Eq. (4.6), from which it is found that $f \approx 0.056$, a value consistent with the inset of figure 4.2.

polymer length as[2] $\tau_d \sim N^{2+(1-1/2)\beta_t} = N^{3.33\pm0.1}$. For the extended repton model, at unit monomer density, a direct numerical verification of this result is computationally out of reach.

4.4 Memory effects in the z-component of the chain tension at the pore

The quantitative formulation of the anomalous dynamics of unbiased polymer translocation by Panja *et al.* in Refs. [40, 41] was based upon the key observation that a translocating polymer is comprised of two polymer seg-

[2]Following footnote 1, chapter 1, the dwell time for a polymer in a strongly entangled crowded environment may scale with N as $\tau_d \sim N^{1+\beta_t}$. This exceeds the scaling $\sim N^{2+\beta_t/2}$ proposed above.

ments, tethered at the pore, that are able to exchange monomers between them through the pore. The velocity of translocation $v(t) = \dot{s}(t)$, representing monomer current, responds to $\phi(t)$, the imbalance in the monomer chemical potential across the pore acting as "voltage". Simultaneously, $\phi(t)$ also adjusts in response to $v(t)$. In the presence of memory effects, they are related to each other by $\phi(t) = \phi_{t=0} + \int_0^t \mathrm{d}t' \mu(t-t')v(t')$ via the memory kernel $\mu(t)$, which can be thought of as the (time-dependent) "impedance" of the system. Supposing a zero-current equilibrium condition at time 0, this relation can be inverted to obtain $v(t) = \int_0^t \mathrm{d}t' a(t-t')\phi(t')$, where $a(t)$ can be thought of as the "admittance". In the Laplace transform language, $\tilde{\mu}(k) = \tilde{a}^{-1}(k)$, where k is the Laplace variable representing inverse time. Via the fluctuation-dissipation theorem, they are related to the respective autocorrelation functions as $\mu(t-t') = \langle \phi(t)\phi(t') \rangle_{v=0}$ and $a(t-t') = \langle v(t)v(t') \rangle_{\phi=0}$. Note that with

$$\langle s^2(t) \rangle = 2 \int_0^t a(t')(t-t')\mathrm{d}t', \tag{4.7}$$

the translocation problem is mapped onto a dynamical property of anchored chains. In particular, if $\langle s^2(t) \rangle \sim t^\alpha$, then for $\alpha < 1$ the impedance response which we can measure from anchored chains must scale as $\mu(t) \sim t^{-\alpha}$.

In this section, following Refs. [40, 41] we determine the memory kernel $\mu(t)$ to describe the dynamics of translocation in a crowded environment (at unit monomer density). The procedure that we follow is that at time $t = 0$ we impose a delta-function in $v(t)$, equivalent to a step-function in $s(t)$. In practice, this is realized by injecting $p = 5$ monomers at the base of an equilibrated tethered chain at time $t = 0$. Next, we want to keep track of the chain tension, as it relaxes from its out-of-equilibrium value right after this injection, to its new equilibrium value. Since the chain tension is numerically a noisy observable, we measure instead its proxy $Z_4(t)$, the center of mass of the first four monomers, following Refs. [40, 41] and chapter 3.

The result, which is obtained by averaging the relaxation of roughly 9 million tethered polymers of length $N = 256$, is plotted in figure 4.4 and its inset. The power-law behavior of $d\langle Z_4(t) \rangle / dt$ at long times with exponent $-(1+\alpha_\mathrm{t})$ shows that $\phi(t) - \phi(\infty) \sim Z_4(t) - Z_4(\infty) \sim t^{-\alpha_\mathrm{t}}$, with $\alpha_\mathrm{t} = 0.78 \pm 0.02$. The short-time behavior is probably sensitive to details of the model and the numerical experiment. From similar relaxation experiments (albeit with significantly reduced accuracy (not shown)) of tethered chains with respectively 125/2 and 216/2 monomers, in a crowded environment of polymers with twice

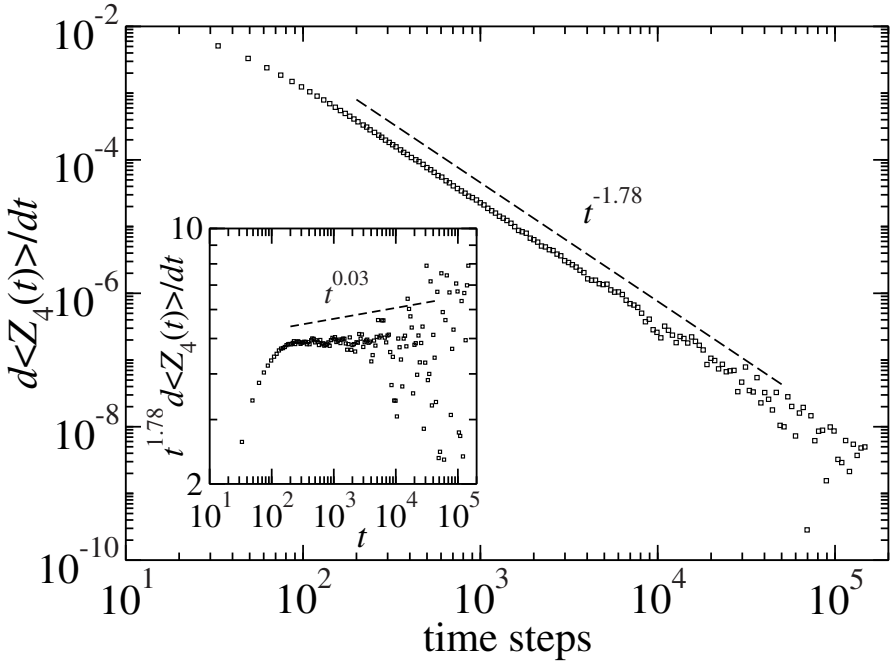

Figure 4.4: Time derivative of $\langle Z_4(t)\rangle$, the average distance between the tethered end and the center of mass of the first four monomers as a function of time, after an injection of monomers at the tethered end at $t = 0$. Before the injection of $p = 5$ monomers, the tethered polymers have a length of 256 monomers, and are surrounded by other polymers with 512 monomers. The membrane is only impenetrable for the tethered polymers. The data is obtained by averaging the relaxation of over 9 million tethered polymers. Each simulation simultaneously relaxes 32 polymers, 16 on either side of the membrane, while 496 polymers form the crowded environment at unit monomer density. These simulations required around $100,000$ processor hours. Inset: same graph, multiplied with $t^{1.78}$ to show the deviation from $t^{-1.75}$.

as many monomers, we find that both $\alpha_t(N = 125/2)$ and $\alpha_t(N = 216/2)$ exceed 0.78, which indicates that α_t is N dependent, at least for short chains.

From Eq. (4.7), together with $\alpha_t = 0.78 \pm 0.02$, we expect to measure an average squared displacement $\langle s^2(t)\rangle \sim t^{\alpha_t} = t^{0.78\pm0.02}$. However, figure 4.2 shows, for $N = 512$ and $3,000 < t < 10,000$, that, consistently with Eq. (4.2), $\langle s^2(t)\rangle \sim t^{0.75}$, while at later times the effective exponent is slightly reduced due to entanglement effects.

Although α_t and the prediction 3/4 seem inconsistent, it must be noted that the difference is small enough to be explained by systematic errors beyond our control.

Alternatively, the inconsistency may indicate the presence of another relaxation mechanism, corresponding to $\mu(t) \sim t^{-\alpha_t}$. It is possible that $\langle s^2(t)\rangle \sim t^{\alpha_t}$ is simply obscured by entanglement effects that force the anomalous diffusion to cross over to $t^{1/2}$. These entanglement effects are not visible in the $\langle Z_4(t)\rangle$ experiments, since the displacements of the polymer are simply too small.

A likely candidate for such a relaxation mechanism is known from recent work by Barkema and Panja [23]. At time $t = 0$ they inject into a tagged polymer, surrounded by a dense polymeric system, p extra monomers. Barkema and Panja convincingly show that in a dense polymer system the longest relaxation time scale in the polymer scales as $\tau_\beta \sim N^\beta$, with $\beta = 2.56 \pm 0.1$. Thus, those monomers of the tagged polymer that are within a distance along the backbone $n_t \sim t^{1/\beta}$ from n^* can be considered equilibrated to this new situation. However, after a time t, this internally equilibrated part stretches in space only to $r(n_t) \sim \sqrt{(n_t)}$, and not to the equilibrated distance $r(n_t) \sim \sqrt{(n_t + p)}$. From a simple scaling argument it is estimated that the free energy increases by an amount

$$\delta F \sim T \cdot \frac{\delta r^2\,(n_t)}{r^2\,(n_t)}. \tag{4.8}$$

Therefore these p monomers encounter a "chemical potential"

$$\mu\,(t) = \frac{1}{p}\frac{\partial F}{\partial n_t} = \frac{1}{p}\frac{\partial F}{\partial r\,(n_t)}\frac{\partial r\,(n_t)}{\partial n_t} \sim t^{-2/\beta}. \tag{4.9}$$

Consequently, it is expected that the excess stored length density $\delta\rho_{n^*}(t)$ at monomer n^* decays $\sim t^{-2/\beta}\exp(-t/\tau_\beta)$. The relation between local strain relaxation and local density of stored lengths implies $\alpha = 2/\beta$. Barkema and Panja continue by measuring α explicitly, to find 0.78 ± 0.03 for $N = 200$.

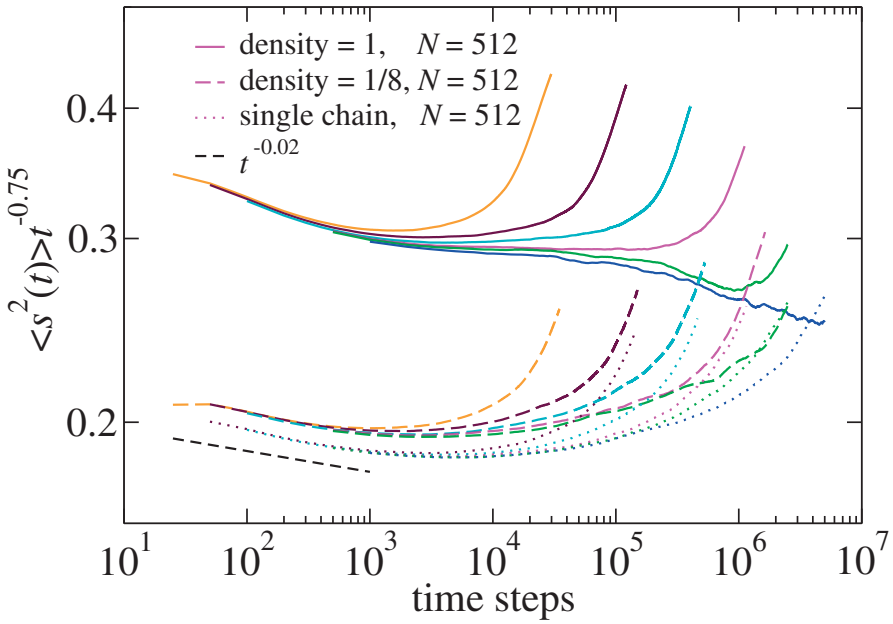

Figure 4.5: Double-logarithmic plot of the average squared displacement divided by $t^{0.75}$ at monomer density $c = 1$ (solid curves, same as in figure 4.2), $c = 1/8$ (dashed curves), and for single chains (dotted curves) for various polymer lengths (the same color coding as in figure 4.2). The size of the simulation box is the same for $c = 1$ and $c = 1/8$. The dashed line $\sim t^{-0.02}$ is added as a guide to the eye, corresponding to the (short time) average squared displacement $\sim t^{1+\nu/1+2\nu}$ of a single chain [40, 41].

A related puzzle is the following. Comparing the translocation of a single polymer with the polymer in a crowded environment, the decay of the force is expected to scale with time as $n(t)^{-(1+\nu)}$, in which ν switches from 0.588 to 1/2. At the same time, the relation between the typical distance of relaxation along the polymer and time alters from $t \sim n(t)^{1+2\nu}$ to $n(t)^{\beta_t}$. While in the single polymer the combination of these two scalings yields a relaxation mechanism proportional to $f(t) \sim t^{-(1+\nu)/(1+2\nu)}$, we would now expect a power-law $f(t) \sim t^{-3/(2\beta_t)} \sim t^{-0.57\pm0.04}$; we do not observe such an exponent.

4.5 Influence of monomer density

In the preceding paragraphs we studied polymer translocation in an environment of polymers of equal length, at a monomer density $c = 1$. Here, we will only touch upon the influence of monomer density. Figure 4.5 shows the average squared displacement at monomer densities $c = 1$, $c = 1/8$, as well as in a simulation box with no other than the translocating polymer. The polymer lengths are varied from $N = 125$ to $N = 1000$ (color coding equivalent to figure 4.2). For clarity $\langle s^2(t)\rangle$ is divided by $t^{0.75}$.

The most prominent difference between low and high density curves is the apparent increase of the prefactor of the mean squared displacement with monomer density c. Table 4.1 contains corresponding dwell times, as obtained from the same simulations. We do not offer a full explanation. However, we have measured that the radius of gyration of the surrounding chains at $c = 1/8$ exceeds the radius of gyration at $c = 1$ by a factor 1.36. The latter chains show ideal scaling, while the former show signs of swelling ($\nu > 0.5$). Correspondingly, we have measured that the stored length density per monomer decreases from roughly 0.43 at $c = 1$, to 0.29 at $c = 1/8$, and finally 0.28 for the chains in the void. (Both quantities show (for $N \geq 512$) no length dependence within the stated accuracy.) For the extended Repton model, the acceptance rate of the elementary moves depends on the density of stored length. Therefore, at least part of this discrepancy can be attributed to properties specific for the lattice polymer model used. It would be very interesting to see whether the strong density-dependence of the overall prefactor also persists for different polymer models. Similarly ill-understood is the influence of density on the behavior at very short times.

Also worth noting is that the curves in figure 4.5 for monomer density $c = 1/8$ closely resemble the single chain curves, contrary to the $c = 1$ curves. It has been shown in Refs. [40, 41] that the longest relaxation time scale in the absence of crowding corresponds to the Rouse time $\tau_{\mathrm{R}} \sim N^{1+2\nu}$. Therefore, the longest relaxation time scale at monomer density $c = 1$, $\tau_{\beta_{\mathrm{t}}} \sim N^{\beta_{\mathrm{t}}}$, is not mirrored at comparatively lower monomer densities: the presence of a relaxation time scaling with N as $\tau_{\beta_{\mathrm{t}}} \sim N^{\beta_{\mathrm{t}}}$ seems to indicate (very) dense crowded environments.

4.6 Conclusions

We have shown that unbiased polymer translocation in a (very) dense crowded environment of identical polymers is sub-diffusive up to the relaxation time scale $\tau_{\beta_t} \sim N^{\beta_t}$, after which it has a further Fickian regime before the longer dwell time is exhausted. In dense systems of long chains, the mean squared displacement along the chain at times shorter than the relaxation time τ_{β_t} is well described by a function $g(t)$, Eq. (4.3): it features a crossover from power-law behavior with exponent 3/4 at short times, to power-law behavior with an exponent of 1/2 at long times. The exponent 1/2 characterizes the dynamics of entangled polymers. The exponent 3/4 is taken from the mean squared displacement $\langle s^2(t)\rangle \sim t^{(1+\nu)/(1+2\nu)}$ observed for unbiased polymer translocation through a narrow pore of a free, single chain [40, 41]; in the crowded environment $\nu = 0.5$. In our high-resolution numerical simulation data, we observe an exponent of 0.78 ± 0.02, deviating from the expected exponent of 3/4; this deviation cannot be explained by statistical fluctuations. Perhaps, there is a link between the exponent $\alpha = 0.78$ observed by Barkema and Panja for the decay of the restoring force in response to polymer enfolding [23].

In summary, we predict a mean dwell time scaling with N as $\tau_d \sim N^{3.33\pm0.1}$ for well-entangled polymers, while in the opposite limit of no entanglements $\tau_d \sim N^{2.66\pm0.05}$.

Chapter 5

Translocation of RNA

This chapter, apart from minor modifications, is taken from e-print arXiv: 0801.3947, *Can Translocation Be Used to Determine RNA Secondary Structures?*, H. Vocks, D. Panja, and G. T. Barkema.

Abstract

Using a combination of theory and high-precision simulations, we study the translocation of an RNA molecule, pulled through a nanopore by an optical tweezer, as a method to determine its secondary structure. The resolution with which the elements of the secondary structure can be determined is limited by thermal fluctuations, ruling out single-nucleotide resolution under normal experimental conditions. Two possible ways to improve this resolution are strong stretching of the RNA with a back-pulling voltage across the membrane, and stiffening of the translocated part of the RNA by biochemical means.

5.1 Introduction

New developments in design and fabrication of nanometer-sized pores and etching methods, in recent times, have put translocation at the forefront of single-molecule experiments [92, 93, 94, 95, 96, 97, 98, 99, 70], with the hope that translocation may lead to cheaper and faster technology for the analysis of biomolecules. The underlying principle is that of a Coulter counter [114]: molecules suspended in an electrolyte solution pass through a narrow pore in a membrane. The electrical impedance of the pore increases with the entrance

of a molecule as it displaces its own volume of solution. By applying a voltage across the pore, the passing molecules are detected as current dips. For nanometer-sized pores (slightly larger than the molecule's cross-section) the magnitude and the duration of these dips have proved to be rather effective in determining the size and length of these molecules [115].

With a membrane placed at $z = 0$, a translocation experiment for determining the secondary structure of an RNA molecule [66, 67, 68, 116, 117, 118, 119] composed of N nucleotides proceeds as follows (see Fig. 5.1). One end of the folded RNA, which is almost completely located on the left side ($z < 0$) of the membrane, is pulled through a nanopore to the right side ($z > 0$), and a latex bead is attached to nucleotide 1. An optical tweezer captures the bead, and pulls the RNA through the pore at a constant speed $v_{\rm tw}$. A potential difference $2V$ is applied across the membrane to increase the tension in the translocated part of the RNA so that no secondary structure can form between the tweezer and the pore. During this process, the force exerted by the tweezer on the RNA is monitored. Since the pore is narrow, for translocation to proceed, the bonds between the base pairs forming the secondary structures must be broken at the pore. The breaking of the base pair-bonds is detected as increased force on the tweezer. The force on the tweezer as a function of time can then be translated into the binding energies of the base pairs as a function of distance along the RNA, yielding a wealth of information on the secondary structure of the RNA.

Given that one is dealing with nanometer-scale distances in this experiment, thermal fluctuations are expected to blur the correlation between the force exerted by the tweezer and the nucleotide number located in the pore. The central question addressed in this chapter, therefore, is the level of resolution (in units of a nucleotide) that can be achieved by this experiment. We address this question with a combination of theory and high-precision computer simulations.

The approach we follow in this chapter is that we consider a poly(U) RNA with a sequence composition $U_{30}\,(U_{60}G_{32}U_6C_{32})_2\,U_{60}$, represented as a polymer of length $N = 350$ nucleotides, wherein *each nucleotide is represented as a monomer*. In our simulation all nucleotides reside on a FCC lattice. Multiple occupation of the same site is forbidden, i.e., the polymer is described by a self-avoiding-walk. For practical purposes, this restriction is lifted for consecutive nucleotides along the chain. Two neighboring C and G-nucleotides can form a bond with an affinity $E_{CG} = 2.5\,k_{\rm B}T$. As a result, the secondary

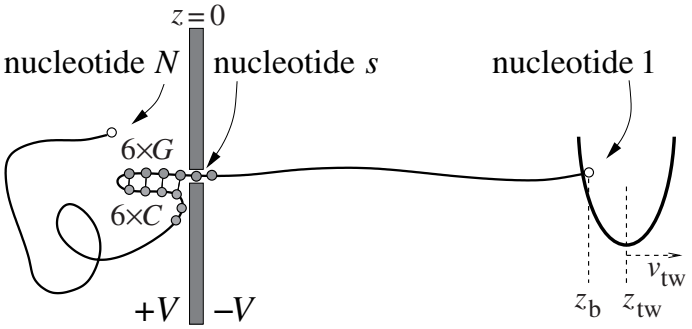

Figure 5.1: The experiment in schematics, illustrated by only 6 CG-bonds for clarity. A RNA molecule composed of N nucleotides is pulled through a narrow pore in a membrane (placed at $z = 0$) towards the right ($z > 0$) using an optical tweezer, represented by a parabolic potential, at a constant speed $v_{\rm tw}$. The bottom of the potential is located at $z_{\rm tw}$; and the latex bead, attached to nucleotide 1 is located at $z_{\rm b}$. The nucleotide located in the pore is labeled s. A potential difference $2V$ is also applied across the membrane.

structure of this polymer consists of two hairpins, each with 32 CG-bonds. The latex bead, i.e., nucleotide 1 feels a spherically symmetric harmonic trap with spring constant $k_{\rm tw}$, centered around the location of the optical tweezer at a distance $z_{\rm tw}$ from the membrane. The dynamics of the polymer consists of single-nucleotide hops to nearest-neighbor sites, attempted at random with rate unity, and accepted with Metropolis probabilities. This model, as detailed in chapter 2, describes both reptation and Rouse moves, but does not include explicit hydrodynamics. This model has been used before successfully to simulate polymer translocation under various circumstances [40, 41, 103, 104, 120]. Since the model is a variant of a freely-jointed-chain, we expect it to reproduce poly(U) RNA behavior reasonably well [65].

To correspond to experimental parameters we use lattice spacing $\lambda = 0.5\,{\rm nm}$, comparable to the persistence length (as well as the typical inter-nucleotide distance) for poly(U) [64]. The resulting forces measured at the tweezer are $\lesssim 60\,{\rm pN}$, similar to experimental values [64]. Based on the diffusion coefficient $2\text{-}5\times10^{-6}\ {\rm cm}^2/{\rm s}$ for U_6 [121], one unit of time in our simulations corresponds roughly to 100 ps. The tweezer velocity is λ per 300,000 units of time, or $\sim 20\,\mu{\rm m/s}$, comparable to typical experimental velocities. The time scale $\lambda/v_{\rm tw}$ is larger than that of the longest time-scale of the translocated part

of the RNA, implying that *the translocated part of the RNA can be treated as properly thermalized at all times.* If we assume nucleotides to carry an effective charge $q = 0.6\,e$, then our V-values satisfy $4\,\mathrm{mV} < V < 30\,\mathrm{mV}$.

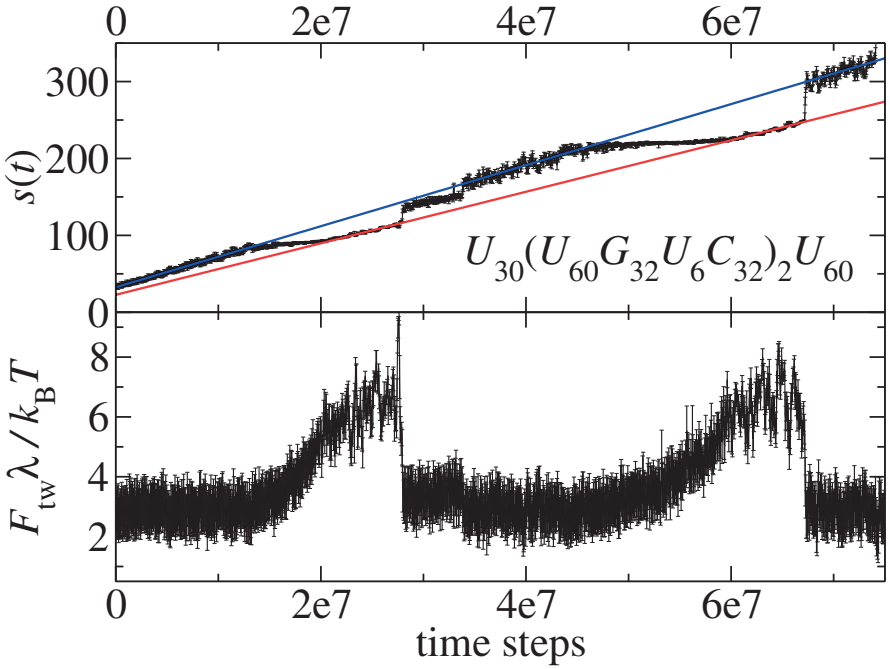

Figure 5.2: Upper panel: nucleotide in the pore s as a function of time for a poly(U) RNA of composition $U_{30}\,(U_{60}G_{32}U_6C_{32})_2\,U_{60}$, pulled with constant velocity $v_{\mathrm{tw}} = \lambda$ per 300,000 time steps. The binding energy of each bond is set to $E_{CG} = 2.5\,k_{\mathrm{B}}T$, with $2qV = 1.5\,k_{\mathrm{B}}T$. Every data point is an average of 1,500 consecutive measurements each 100 time steps apart, with the standard deviation represented by the error bars. The two straight lines are guides to the eye. Lower panel: the corresponding chain tension measured by means of the optical tweezer, with $k_{\mathrm{tw}} = 1\,k_{\mathrm{B}}T/\lambda$.

A typical simulation output is presented in Fig. 5.2. It consists of the force $F_{\mathrm{tw}}(t)$ exerted by the optical tweezer as a function of time: *this information is readily accessible in real experiments.* We also monitor the number $s(t)$ as a function of time: *this information is typically not accessible in real experiments.* In the simulation of figure 5.2, the force exerted by the tweezer hovers around a fixed strength $3.4\,k_{\mathrm{B}}T/\lambda$, except between $t = 1.4 \times 10^7$ to

3.4×10^7 resp. $t = 4.6 \times 10^7$ to 6.7×10^7, when the first and second (halfs of the) hairpins are pulled through the pore. Indeed, the top panel of Fig. 5.2 shows that at the onset of these intervals, $s(t)$ is almost constant, around 90 respectively 220, the starting locations of the hairpins.

5.2 Translocation without thermal fluctuations

First we discuss what sort of information on the secondary structure can *ideally* — i.e., in the *absence* of thermal fluctuations — be obtained from $F_{\rm tw}(t)$. We do this under the assumptions that the force extension curve of the RNA without any secondary structure is sequence-independent, and that the tweezer velocity is low enough for the force exerted by the tweezer to maintain a uniform chain tension ϕ all along the translocated part of the RNA. Then, ϕ is uniquely determined by the relative extension $x = z_{\rm b}/s$, and is balanced by $F_{\rm tw}(t)$, i.e.,

$$\phi = f(x) = F_{\rm tw} = k_{\rm tw}(z_{\rm tw} - z_{\rm b}). \tag{5.1}$$

Additionally, the equality of the rate of work done by the tweezer and the gain in free energy by the translocating nucleotides at the pore yields

$$F_{\rm tw} dz_{\rm b} = (\Delta U - T\Delta S)\, ds, \tag{5.2}$$

where ΔS is the entropic cost per nucleotide translocation due to the imbalance of the (entropic) chain tension across the pore, and ΔU is the energetic cost per nucleotide translocation. If translocation of the nucleotides does not involve breaking of CG-bonds at the pore, then $\Delta U = \Delta U_c \equiv 2qV$, otherwise $\Delta U = \Delta U_b \equiv 2qV + E_{CG}$. Thus, given $z_{\rm b}$ and $(\Delta U - T\Delta S)$, Eqs. (5.1) and (5.2) determine both the tweezer force and the relative extension.

During the translocation of the first 90 nucleotides of U_{30} $(U_{60}G_{32}U_6C_{32})_2\,U_{60}$, no secondary structure is broken at the pore — consequently, $\Delta U = \Delta U_c$ — and the tweezer force remains constant at $F_{\rm tw}(t) = 3.4\,k_{\rm B}T/\lambda$. The speed of translocation is then given solely by Eq. (5.1), with $\dot{z}_{\rm b} = v_{\rm tw}$, as $\dot{s} = v_{\rm tw}/f^{-1}[F_{\rm tw}(t)] = 1$ nucleotide per $252,000$ time steps, for which we have used the numerically obtained force extension curve (inset, Fig. 5.3); this speed $\dot{s}$ is shown in Fig. 5.2 by the upper blue line.

Following the arrival of the first hairpin at the pore, translocation requires breaking of the CG-bonds, and consequently, ΔU increases by $E_{CG} = 2.5\,k_{\rm B}T$. Both the tweezer force and the relative extension adjust to new values, determined by Eqs. (5.1) and (5.2) for the new value for ΔU. The resulting tweezer

force equals $F_{\rm tw}(t) = 7.5\, k_{\rm B}T/\lambda$; the correspondingly adjusted speed $\dot{s}$ is shown in Fig. 5.2 by the lower red line.

After the translocation of the first 32 G-nucleotides, ΔU returns to its base value ΔU_c. The tweezer force and the relative extension, too, fall back to their pre-hairpin values. This is seen in Fig. 5.2 by $s(t)$ leaving the red line *sharply* to re-coincide with the blue line; i.e., quite a few nucleotides at the end of the hairpin translocate nearly immediately. This chain of events repeats itself during translocation of the second hairpin: first the tweezer force increases gradually to its higher value and the translocated distance approaches the red line, then the tweezer force decreases steeply to its lower value and the translocated distance jumps to the blue line.

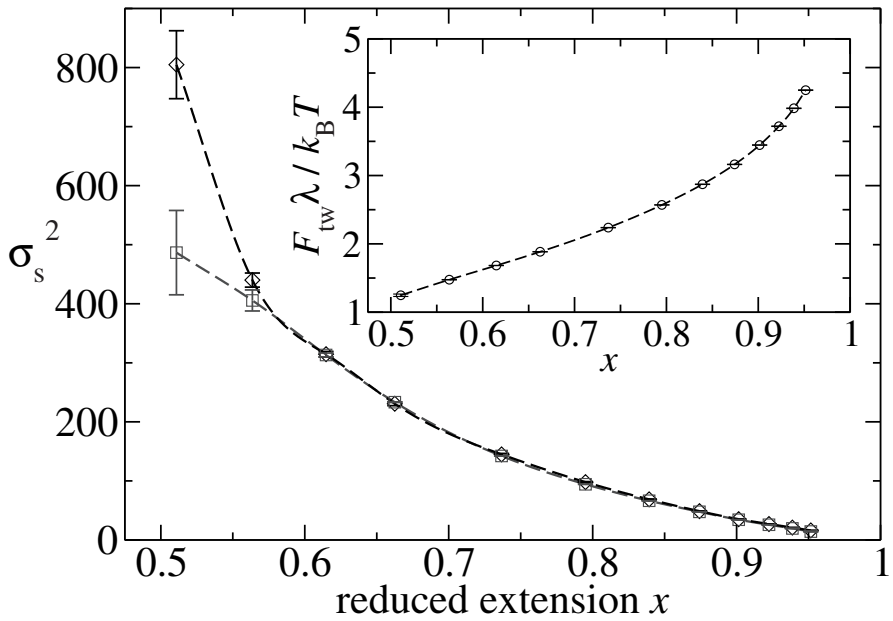

Figure 5.3: σ_s^2, mean squared displacement of s vs. reduced extension $x = z_b/s$, at constant $z_{\rm tw} = 300\,\lambda$. The chain tension ϕ is slowly increased by changing $\Delta U = 2qV$ from 0.40 to $2.75\, k_{\rm B}T$. Each data point required 80 independent polymers and simulation times of 20 million time steps, with a measurement every 100 time steps. Data points from direct simulations and Eq. (5.3) are represented as diamonds and squares. The dashed lines are cubic splines. The error bars represent statistical errors only. Inset: Rescaled force-extension curve for our model.

In conclusion, most features of Fig. 5.2 are qualitatively well understood. The above framework can be easily extended to a wider set of bond strengths and more elaborate secondary structures. Without thermal fluctuations, the setup is perfectly suited to determine the secondary structure up to the nucleotide resolution, under the restriction that the consecutive bonds along the backbone of the RNA are of increasing strength; if strong bonds are followed by weaker bonds that are not strong enough to halt the translocation process, breaking of the weaker bonds will not be accompanied by an increase of F_{tw}.

5.3 Translocation with thermal fluctuations

In *reality*, thermal fluctuations are omnipresent in this nano-scale experiment. In fact, it is precisely the (thermal) fluctuations in $s(t)$ that serve to blur the coherence between $F_{\mathrm{tw}}(t)$ and $s(t)$, and thereby limit the resolution that can be achieved by this experiment. We will now study both the amplitude and the frequency spectrum of these thermal fluctuations in $s(t)$.

The amplitude $\sigma_s(t, \Delta U)$ of the fluctuations in $s(t)$ is that of an entropic spring of length z_{b} with one end tethered at the tweezer, while the number of nucleotides s in the spring are allowed to fluctuate through the pore. Now consider a *different* problem — an entropic spring of length z_{b}. From equipartition theorem, its length fluctuations are given by $\langle \delta z_{\mathrm{b}}^2 \rangle = k_{\mathrm{B}}T/c_{z_{\mathrm{b}}}$ [with spring constant $c_{z_{\mathrm{b}}} = (\partial f/\partial z_{\mathrm{b}})_s = f'(x)/s$]. For the present problem, such fluctuations in z_{b} can be thought of to be mediated by the fluctuations in s, yielding[1]

$$\sigma_s^2(t, \Delta U) = \langle \delta z_b^2 \rangle \left[\frac{\partial s}{\partial z_{\mathrm{b}}} \right]_x^2 = s(t) \big[x^2 f'(x) \big]^{-1} k_{\mathrm{B}}T. \tag{5.3}$$

In Fig. 5.3, Eq. (5.3) is compared to the simulation results for several values of ΔU, with constant $z_{\mathrm{tw}} = 300\,\lambda$. Note that ΔU only serves to set the value of ϕ.

For the frequency spectrum of $s(t)$, we build upon Ref. [104]. Therein it is shown that $\dot{s}(t)$ and the chain tension imbalance across the pore are related via a time-dependent memory kernel $a(t)$. This result, adapted to the notations

[1] The direct simulation results and Eq. (5.3) in Fig. 5.3 deviate from each other for $x \lesssim 0.6$, since in this case the assumption $\sigma_s \ll s$ does not hold any more. Also, tethering of nucleotide 1 requires $k_{\mathrm{tw}} \gg c_{z_b}$; this assumption is clearly valid in our simulations, but need not be made valid in real experiments.

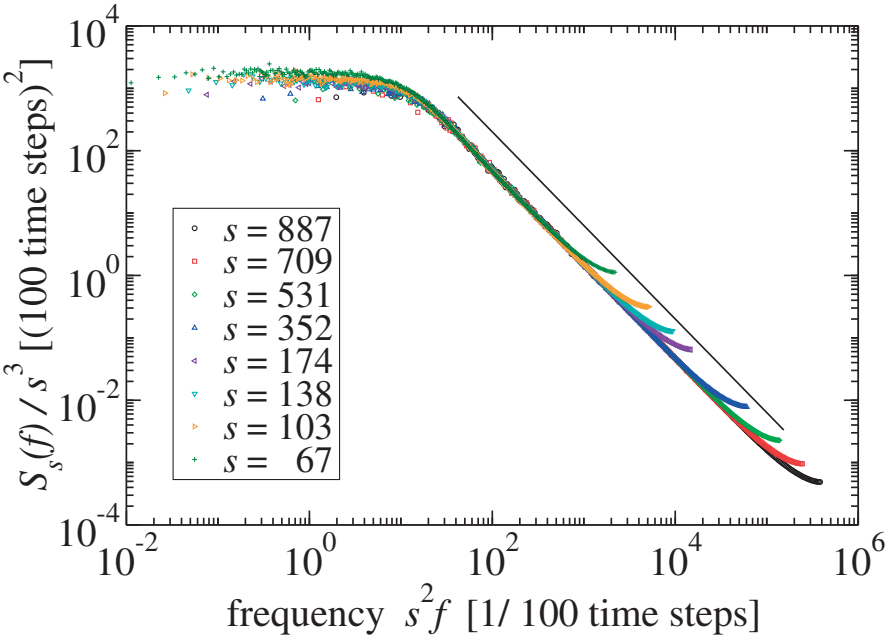

Figure 5.4: (Rescaled) power spectrum of $s(t)$, $S_s(f)$, versus (rescaled) frequency for $\Delta U = 1.5\ k_{\rm B}T$, and $z_{\rm tw}/\lambda$=60, 90, 120, 150, 300, 450, 600, 750. Each curve is composed of statistics from 80 polymers for 40 million time steps. The solid line $\sim f^{-3/2}$ is added as a guide to the eye.

in this chapter, is given by

$$\dot{s}(t) = \int_{-\infty}^{t} dt'\, a(t-t')\,[\phi(t') - \phi_{z<0}(t')], \tag{5.4}$$

where $\phi_{z<0}$ is the chain tension of the RNA at the $z < 0$ side of the pore, and $a(t) \sim t^{-3/2} \exp[-t/\tau_F]$ with $\tau_F \sim s^2$. The immediate consequence of Eq. (5.4) is that

$$\langle \dot{s}(t)\dot{s}(t')\rangle \sim |t-t'|^{-3/2} \exp\left[-|t-t'|/\tau_F\right], \tag{5.5}$$

following fluctuation-dissipation theorem[2], i.e.,

$$\begin{aligned}\langle s(t)s(t')\rangle &= \int_0^t dt_1 \int_0^{t'} dt_2 \langle \dot{s}(t_1)\dot{s}(t_2)\rangle \\ &= \begin{cases} 4\left[|t-t'|^{1/2} - |t|^{1/2} - |t'|^{1/2}\right] & |t-t'| \lesssim \tau_F \\ k & |t-t'| \gg \tau_F \end{cases}, \end{aligned} \tag{5.6}$$

with some constant k. This velocity autocorrelation function, for $|t-t'| \lesssim \tau_F$, has the same form as that of a fractional Brownian motion (FBM) [122, 123]

$$\langle s(t)s(t')\rangle = k'\left(|t|^{2H} + |t'|^{2H} - |t-t'|^{2H}\right), \tag{5.7}$$

with Hurst parameter $H = 1/4$. Since the r.h.s. of (5.7) is not a function of $(t-t')$ alone, the spectral density of FBM is not well-defined. However, by applying generalized concepts such as the Wigner-Ville spectrum [122], a limiting power spectrum can be obtained as follows: $S_s(f) = \frac{k''}{|f|^\alpha}$, with $\alpha = 2H+1 = 3/2$. We therefore expect a power spectrum, with $f_c \sim 1/\tau_F \sim 1/s^2$,

$$S_s(f) = \begin{cases} \frac{k''}{|f|^{3/2}} & (|f| > f_c) \\ \frac{k''}{f_c^{3/2}} & (|f| < f_c) \end{cases}. \tag{5.8}$$

As shown in Fig. 5.4, after having rescaled $S_s(f)$ and f by s^{-3} and s^2 respectively, the collapse of the power spectra confirms Eq. (5.8). The spectrum is dominated by the low-frequency fluctuations of $s(t)$, which are not easily suppressed by time averaging.

5.4 Discussion and conclusions

Having demonstrated the importance of thermal fluctuations in $s(t)$, let us now revisit the translocation of $U_{30}\left(U_{60}G_{32}U_6C_{32}\right)_2 U_{60}$, cf. Fig. 5.2. Except between times 1.4 and 3.4×10^7 (resp. 4.6 and 6.7×10^7), $F_{\rm tw}(t)$ is roughly $3.4\,k_{\rm B}T/\lambda$ with fluctuations $\langle \Delta F_{\rm tw}^2(t)\rangle \approx k_{\rm tw}k_{\rm B}T$, while Eq. (5.3) explains the growth of $\sigma_s^2(t)$ linearly with $s(t)$ (hence linearly with t). It is important to stress that the fluctuations in $F_{\rm tw}(t)$ may be removed to a reasonable degree

[2]Note here that Ref. [104] is about pulling the tweezer with a constant force, while in this Letter the tweezer is pulled with a constant velocity. This does not affect the memory kernel and the consequences of the RNA's memory.

by time averaging (e.g., by reducing the velocity of the tweezer), the fluctuations in $s(t)$ fundamentally reduce the accuracy with which RNA secondary structure can be observed. As soon as the hairpin is within a distance of order $\sigma_s(t, \Delta U)$ from the pore, i.e., $\langle s(t)\rangle + \sigma_s(t, \Delta U_c) = 90\,(\text{resp. } 220)$, the tweezer force already increases, to reach a plateau once $\langle s(t)\rangle - \sigma_s(t, \Delta U_b) = 90\,(\text{resp. } 220)$. Translocation at the pore then proceeds with a constant rate $v_{\text{tw}}/f^{-1}(F_{\text{tw}})$, until $\langle s(t)\rangle + \sigma_s(t, \Delta U_b) = 122\,(\text{resp. } 252)$ (in which we assume that the effective affinity E_{CG} is not altered much as the unzipping of the hairpin proceeds), at which time the chain tension is released by rapid translocation of the remaining nucleotides. These considerations clearly establish that only secondary structure elements with at least two times $\sigma_s(t, \Delta U_b)$ base pairs (≈ 8 nucleotides in Fig. 5.2) can be detected accurately, provided the affinities of consecutive base pairs are all alike. Moreover, the base pairs of the preceding stem of the secondary structure must be characterized either by weaker affinities, or by strong heterogeneity.

Although time averaging will not eliminate the fluctuations in $s(t)$, Eq. (5.3) does leave open avenues for higher accuracy, either by increasing x, or by increasing the stiffness $f'(x)$. The former can be achieved by applying a stronger potential difference $2V$ (the strength of which is of course limited by experimental considerations). On the other hand, the polymer's stiffness can be actively enhanced, e.g., by altering the salinity of the solution [64]; alternatively, if one is interested in doing single-molecule experiments with ssDNA, the addition of RecA proteins to the solution may be of help as well.

Bibliography

[1] M. Doi, *Introduction to Polymer Physics*, Clarendon Press, Oxford (1995).

[2] P.-G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, New York (1979).

[3] S. Caracciolo, M. S. Causo, and A. Pelissetto, *High-precision determination of the critical exponent for self-avoiding walks*, Phys. Rev. E **57**, R1215 (1998).

[4] P. J. Flory, *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, New York (1953).

[5] B. Li, N. Madras, and A. D. Sokal, *Critical exponents, hyperscaling, and universal amplitude ratios for two- and three-dimensional self-avoiding walks*, J. Stat. Phys. **80**, 661 (1995).

[6] B. Dünweg and K. Kremer, *Molecular dynamics simulation of a polymer chain in solution*, J. Chem. Phys. **99**, 6983 (1993).

[7] A. J. Banchio and J. F. Brady, *Accelerated Stokesian dynamics: Brownian motion*, J. Chem. Phys. **118**, 10323 (2003).

[8] P. Ahlrichs and B. Dünweg, *Simulation of a single polymer chain in solution by combining lattice Boltzmann and molecular dynamics*, J. Chem. Phys. **111**, 8225 (1999).

[9] O. Berk Usta, J. E. Butler, and A. J. C. Ladd, *Flow-induced migration of polymers in dilute solution*, Phys. Fluids, **18**, 031703 (2006).

[10] O. Berk Usta, J. E. Butler, and A. J. C. Ladd, *Transverse Migration of a Confined Polymer Driven by an External Force*, Phys. Rev. Lett. **98**, 098301 (2007).

[11] M. Doi and S. F. Edwards, *The Theory of Polymer Translocation*, Clarendon Press, Oxford (1986).

[12] P.-G. De Gennes, *Reptation of a Polymer Chain in the Presence of Fixed Obstacles*, J. Chem. Phys. **55**, 572 (1971).

[13] M. Rubinstein, *Discretized model of entangled-polymer dynamics*, Phys. Rev. Lett. **59**, 1946 (1987).

[14] T. P. Lodge, *Reconciliation of the Molecular Weight Dependence of Diffusion and Viscosity in Entangled Polymers*, Phys. Rev. Lett. **83**, 3218 (1999).

[15] M. Doi, *Explanation for the 3.4-power law for viscosity of polymeric liquids on the basis of the tube model*, J. Polym. Sci. Polym. Phys. Ed. **21**, 667 (1983).

[16] A. E. Likhtman and T. C. B. McLeish, *Quantitative Theory for Linear Dynamics of Linear Entangled Polymers*, Macromolecules **35**, 6332 (2002).

[17] A. Wischnewski, M. Monkenbusch, L. Willner, D. Richter, A. E. Likhtman, T. C. B. McLeish, and B. Farago, *Molecular Observation of Contour-Length Fluctuations Limiting Topological Confinement in Polymer Melts*, Phys. Rev. Lett. **88**, 058301 (2002).

[18] W. W. Graessley, *Entangled linear, branched and network polymer systems – Molecular theories*, Adv. Polym. Sci. **47**, 68 (1982).

[19] M. Rubinstein and R. H. Colby, *Self-consistent theory of polydisperse entangled polymers: Linear viscoelasticity of binary blends*, J. Chem. Phys. **89**, 5291 (1988).

[20] S. Q. Wang, *Chain dynamics in entangled polymers: Diffusion versus rheology and their comparison*, J. Polym. Sci. Part B: Polym. Phys. **41**, 1589 (2003).

[21] J. F. Vega, S. Rastogi, G. W. M. Peters, and H. E. H. Meijer, *Rheology and reptation of linear polymers. Ultrahigh molecular weight chain dynamics in the melt*, Journal of Rheology **48**, 663 (2004).

[22] C. Y. Liu, R. Keunings, and C. Bailly, *Do Deviations from Reptation Scaling of Entangled Polymer Melts Result from Single- or Many-Chain Effects?*, Phys. Rev. Lett. **97**, 246001 (2006).

[23] G. T. Barkema and D. Panja, to be published.

[24] M. Muthukumar, *Polymer translocation through a hole*, J. Chem. Phys. **111**, 10371 (1999).

[25] P. J. Park and W. Sung, *Polymer release out of a spherical vesicle through a pore*, Phys. Rev. E **57**, 730 (1998).

[26] M. Muthukumar, *Translocation of a Confined Polymer through a Hole*, Phys. Rev. Lett. **86**, 3188 (2001).

[27] D. K. Lubensky and D. R. Nelson, *Driven Polymer Translocation Through a Narrow Pore*, Biophys. J. **77**, 1824 (1999).

[28] P. J. Park and W. Sung, *Polymer translocation induced by adsorption*, J. Chem. Phys. **108**, 3013 (1998).

[29] E. Slonkina and A. B. Kolomeisky, *Polymer translocation through a long nanopore*, J. Chem. Phys. **118**, 7112 (2003).

[30] W. Sung and P. J. Park, *Polymer Translocation through a Pore in a Membrane*, Phys. Rev. Lett. **77**, 783 (1996).

[31] H. W. Diehl and M. Shpot, *Massive field-theory approach to surface critical behavior in three-dimensional systems*, Nucl. Phys. B **528**, 595 (1998).

[32] J. Chuang, Y. Kantor, and M. Kardar, *Anomalous dynamics of translocation*, Phys. Rev. E **65**, 011802 (2001).

[33] Y. Kantor and M. Kardar, *Anomalous dynamics of forced translocation*, Phys. Rev. E **69**, 021806 (2004).

[34] C. Chatalain, Y. Kantor, and M. Kardar, *Probability distributions for polymer translocation*, e-print arXiv: 0805.4168.

[35] I. Huopaniemi, K. Luo, T. Ala-Nissila, and S. C. Ying, *Langevin dynamics simulations of polymer translocation through nanopores*, J. Chem. Phys. **125**, 124901 (2006).

[36] K. Luo, S. T. T. Ollila, I. Huopaniemi, T. Ala-Nissila, P. Pomorski, M. Karttunen, S. C. Ying, and A. Bhattacharya, *Dynamical Scaling Exponents for Polymer Translocation through a Nanopore*, e-print arXiv: 0805.4312.

[37] J. Klein Wolterink, G. T. Barkema, and D. Panja, *Passage Times for Unbiased Polymer Translocation through a Narrow Pore*, Phys. Rev. Lett. **96**, 208301 (2006).

[38] J. L. A. Dubbeldam, A. Milchev, V. G. Rostiashvili, and T. A. Vilgis, *Polymer translocation through a nanopore: A showcase of anomalous diffusion*, Phys. Rev. E **76**, 010801(R) (2007).

[39] J. L. A. Dubbeldam, A. Milchev, V. G. Rostiashvili, and T. A. Vilgis, *Driven polymer translocation through a nanopore: A manifestation of anomalous diffusion*, Europhys. Lett. **79**, 18002 (2007).

[40] D. Panja, G. T. Barkema, and R. C. Ball, *Effect of Anomalous Dynamics on Unbiased Polymer Translocation*, e-print arXiv:cond-mat/0610671.

[41] D. Panja, G. T. Barkema, and R. C. Ball, *Anomalous dynamics of unbiased polymer translocation through a narrow pore*, J. Phys.: Condens. Matter **19**, 432202 (2007).

[42] A. van Heukelum and G. T. Barkema, *Reaching large lengths and long times in polymer dynamics simulations*, J. Chem. Phys. **119**, 8197 (2003).

[43] A. van Heukelum, G. T. Barkema, M. W. Edelman, E. van der Linden, E. H. A. de Hoog, and R. H. Tromp, *Fractionation in a Phase-Separated Polydisperse Polymer Mixture*, Macromolecules **36**, 6662 (2003).

[44] J. Klein Wolterink, G. T. Barkema, and M. A. Cohen Stuart, *Diffusion and Exchange of Adsorbed Polymers Studied by Monte Carlo Simulations*, Macromolecules **38**, 2009 (2004).

[45] J. Klein Wolterink and G. T. Barkema, *Polymer diffusion in a lattice polymer model with an intrinsic reptation mechanism*, Mol. Phys. **103**, 3083 (2005).

[46] P. H. Verdier and W. H. Stockmayer, *Monte Carlo Calculations on the Dynamics of Polymers in Dilute Solution*, J. Chem. Phys. **36**, 227 (1962).

[47] K. Kremer, A. Baumgartner, and K. Binder, *Collapse transition and crossover scaling for self-avoiding walks on the diamond lattice*, J. Phys. A: Math. Gen. **15**, 2879 (1982).

[48] K. Kremer, *Statics and dynamics of polymeric melts: a numerical analysis*, Macromolecules **16**, 1632 (1983).

[49] A. Kolinski, J. Skolnick, and R. Yaris, *Monte Carlo studies on the long time dynamic properties of dense cubic lattice multichain systems. I. The homopolymeric melt*, J. Chem. Phys. **86**, 7164 (1987).

[50] A. Kolinski, J. Skolnick, and R. Yaris, *Monte Carlo studies on the long time dynamic properties of dense cubic lattice multichain systems. II. Probe polymer in a matrix of different degrees of polymerization*, J. Chem. Phys. **86**, 7174 (1987).

[51] I. Carmesin and K. Kremer, *The bond fluctuation method: a new effective algorithm for the dynamics of polymers in all spatial dimensions*, Macromolecules **21**, 2819 (1988).

[52] I. Carmesin and K. Kremer, *Static and dynamic properties of two-dimensional polymer melts*, J. Phys. France **51**, 915 (1990).

[53] H. P. Deutsch and K. Binder, *Interdiffusion and self-diffusion in polymer mixtures: A Monte Carlo study*, J. Chem. Phys. **94**, 2294 (1991).

[54] T. Kreer, J. Baschnagel, M. Müller, and K. Binder, *Monte Carlo Simulation of Long Chain Polymer Melts: Crossover from Rouse to Reptation Dynamics*, Macromolecules **34**, 1105 (2001).

[55] K. E. Evans and S. F. Edwards, *Computer simulations of the dynamics of highly entangled polymers. I. Equilibrium dynamics*, J. Chem. Soc., Faraday Trans. 2 **77**, 1891 (1981).

[56] S. F. Edwards and K. E. Evans, *Computer simulations of the dynamics of highly entangled polymers. II. Static properties of the primitive chain*, J. Chem. Soc., Faraday Trans. 2 **77**, 1913 (1981).

[57] K. E. Evans and S. F. Edwards, *Computer simulations of the dynamics of highly entangled polymers. III. Dynamics of the primitive chain*, J. Chem. Soc., Faraday Trans. 2 **77**, 1929 (1981).

[58] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics*, Clarendon Press, Oxford (1999).

[59] G. T. Barkema, J. F. Marko, and B. Widom, *Electrophoresis of charged polymers: Simulation and scaling in a lattice model of reptation*, Phys. Rev. E **49**, 5303 (1994).

[60] J. M. J. van Leeuwen and A. Kooiman, *The drift velocity in the Rubinstein-Duke model for electrophoresis*, Physica A **184**, 79 (1992).

[61] A. Kooiman and J. M. J. van Leeuwen, *Reptation models for electrophoresis*, Physica A **194**, 163 (1993).

[62] A. van Heukelum, *Simulation of Polymer Dynamics in Gels and Melts*, Ph.D. thesis, Utrecht University (2003).

[63] A. Drzewiński and J. M. J. van Leeuwen, *Crossover from reptation to Rouse dynamics in the cage model*, Phys. Rev. E **74**, 061801 (2006).

[64] Y. Seol, G. M. Skinner, and K. Visscher, *Elastic Properties of a Single-Stranded Charged Homopolymeric Ribonucleotide*, Phys. Rev. Lett. **93**, 118102 (2004).

[65] Y. Seol, G. M. Skinner, and K. Visscher, *Stretching of Homopolymeric RNA Reveals Single-Stranded Helices and Base-Stacking*, Phys. Rev. Lett. **98**, 158103 (2007).

[66] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, *Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM*, Science **276**, 1109 (1997).

[67] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, *Reversible Unfolding of Single RNA Molecules by Mechanical Force*, Science **292**, 733 (2001).

[68] B. Onoa, S. Dumont, J. Liphardt, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, *Identifying Kinetic Barriers to Mechanical Unfolding of the T. thermophila Ribozyme*, Science **299**, 1892 (2003).

[69] C. Bustamante, Z. Bryant, and S. B. Smith, *Ten years of tension: single-molecule DNA mechanics*, Nature **421**, 423 (2003).

[70] U. F. Keyser, B. N. Koeleman, S. van Dorp, D. Krapf, R. M. M. Smeets, S. G. Lemay, N. H. Dekker, and C. Dekker, *Direct force measurements on DNA in a solid-state nanopore*, Nature Physics **2**, 473 (2006).

[71] R. S. Hoy and M. O. Robbins, *Effect of equilibration on primitive path analyses of entangled polymers*, Phys. Rev. E **72**, 061802 (2005).

[72] R. Auhl and R. Everaers, *Equilibration of long chain polymer melts in computer simulations*, J. Chem. Phys. **119**, 12718 (2003).

[73] A. K. Kron, Vysokomol. Soyed. **7**, 1228 (1965) [Polymer Science USSR **7**, 1361 (1965)].

[74] A. K. Kron, O. B. Ptitsyn, A. M. Skvortsov, and A. K. Fedorov, Molek. Biol. **1**, 576 (1967) [Molec. Biol. **1**, 487 (1967)].

[75] F. T. Wall and F. Mandel, *Macromolecular dimensions obtained by an efficient Monte Carlo method without sample attrition*, J. Chem. Phys. **63**, 4592 (1975).

[76] F. Mandel, *Macromolecular dimensions obtained by an efficient Monte Carlo method: The mean square end-to-end separation*, J. Chem. Phys. **70**, 3984 (1979).

[77] A. D. Sokal, *Monte Carlo Methods for the Self-Avoiding Walk*, e-print arXiv:hep-lat/9405016.

[78] J. M. Deutsch, *Towards an Explanation of the 3.4-Power Dependence of the Viscosity on Molecular Weight*, Phys. Rev. Lett. **54**, 56 (1985).

[79] L. Mattioni, J. P. Wittmer, J. Baschnagel, J.-L. Barrat, and E. Luijten, *Dynamical properties of the slithering-snake algorithm: A numerical test of the activated-reptation hypothesis*, Eur. Phys. J. E **10**, 369 (2003).

[80] J. M. Deutsch, *The dynamics of entangled polymers*, J. Phys. France **48**, 141 (1987).

[81] A. N. Semenov, *Concentration fluctuations and reptation dynamics in polymer solutions and melts*, Physica A **171**, 517 (1991).

[82] A. N. Semenov, in *Theoretical Challenges in the Dynamics of Complex Fluids* edited by T. McLeish, pp. 6386, Kluwer, Dordrecht (1997).

[83] A. N. Semenov and M. Rubinstein, *Dynamics of strongly entangled polymer systems: activated reptation*, Eur. Phys. J. B **1**, 87 (1998).

[84] B. Dreiseikelmann, *Translocation of DNA across bacterial membranes*, Microbiol. Rev. **58**, 293 (1994).

[85] J.-P. Henry, J.-F. Chich, D. Goldschmidt, and M. Thieffry, *Blockade of a mitochondrial cationic channel by an addressing peptide: An electrophysiological study*, J. Membr. Biol. **112**, 139 (1989).

[86] J. Akimaru, S. Matsuyama, H. Tokuda, and S. Mizushima, *Reconstitution of a Protein Translocation System Containing Purified SecY, SecE, and SecA from Escherichia coli*, Proc. Natl. Acad. Sci. USA **88**, 6545 (1991).

[87] D. Görlich and T. A. Rapoport, *Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane*, Cell **75**, 615 (1993).

[88] G. Schatz and B. Dobberstein, *Common Principles of Protein Translocation Across Membranes*, Science **271**, 1519 (1996).

[89] I. Szabò, G. Bàthori, F. Tombola, M. Brini, A. Coppola, and M. Zoratti, *DNA Translocation Across Planar Bilayers Containing Bacillus subtilis Ion Channels*, J. Biol. Chem. **272**, 25275 (1997).

[90] B. Hanss, E. Leal-Pinto, L. A. Bruggeman, T. D. Copeland, and P. E. Klotman, *Identification and characterization of a cell membrane nucleic acid channel*, Proc. Natl. Acad. Sci. USA **95**, 1921 (1998).

[91] Y.-L. Tseng, J.-J. Liu, and R.-L. Hong, *Translocation of Liposomes into Cancer Cells by Cell-Penetrating Peptides Penetratin and Tat: A Kinetic and Efficacy Study*, Mol. Pharmacol. **62**, 864 (2002).

[92] J. J. Nakane, M. Akeson, and A. Marziali, *Nanopore sensors for nucleic acid analysis*, J. Phys.: Condens. Matter. **15**, R1365 (2003).

[93] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, *Characterization of individual polynucleotide molecules using a membrane channel*, Proc. Natl. Acad. Sci. USA **93**, 13770 (1996).

[94] S. E. Henrickson, M. Misakian, B. Robertson, and J. J. Kasianowicz, *Driven DNA Transport into an Asymmetric Nanometer-Scale Pore*, Phys. Rev. Lett. **85**, 3057 (2000).

[95] A. Meller, L. Nivon, and D. Branton, *Voltage-Driven DNA Translocations through a Nanopore*, Phys. Rev. Lett. **86**, 3435 (2001).

[96] M. Akeson, D. Branton, J. J. Kasianowicz, E. Brandin, and D. W. Deamer, *Microsecond Time-Scale Discrimination Among Polycytidylic Acid, Polyadenylic Acid, and Polyuridylic Acid as Homopolymers or as Segments Within Single RNA Molecules*, Biophys. J. **77**, 3227 (1999).

[97] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, and D. Branton, *Rapid nanopore discrimination between single polynucleotide molecules*, Proc. Natl. Acad. Sci. USA **97**, 1079 (2000).

[98] A. Meller and D. Branton, *Single molecule measurements of DNA transport through a nanopore*, Electrophoresis **23**, 2583 (2002).

[99] A. J. Storm, C. Storm, J. Chen, H. Zandbergen, J.-F. Joanny, and C. Dekker, *Fast DNA Translocation through a Solid-State Nanopore*, Nano Lett. **5**, 1193 (2005).

[100] I. Szabò, G. Bàthori, F. Tombola, A. Coppola, I. Schmehl, M. Brini, A. Ghazi, V. De Pinto, and M. Zoratti, *Double-stranded DNA can be translocated across a planar membrane containing purified mitochondrial porin*, FASEB J. **12**, 495 (1998).

[101] S. Howorka, L. Movileanu, O. Braha, and H. Bayley, *Kinetics of duplex formation for individual DNA strands within a single protein nanopore*, Proc. Natl. Acad. Sci. USA **98**, 12996 (2001).

[102] S. Howorka, S. Cheley, and H. Bayley, *Sequence-specific detection of individual DNA strands using engineered nanopores*, Nature Biotechnol. **19**, 636 (2001).

[103] D. Panja, G. T. Barkema, and R. C. Ball, *Polymer translocation out of planar confinements*, J. Phys.: Condens. Matter **20**, 075101 (2008).

[104] D. Panja and G. T. Barkema, *Passage Times for Polymer Translocation Pulled through a Narrow Pore*, Biophys. J. **94**, 1630 (2008).

[105] K. Luo, I. Huopaniemi, T. Ala-Nissila, and S. C. Ying, *Polymer translocation through a nanopore under an applied external field*, J. Chem. Phys. **124**, 114704 (2006).

[106] A. Cacciuto and E. Luijten, *Confinement-Driven Translocation of a Flexible Polymer*, Phys. Rev. Lett. **96**, 238104 (2006).

[107] D. Wei, W. Yang, J. Xigao, and Q. Liao *Unforced translocation of a polymer chain through a nanopore: The solvent effect*, J. Chem. Phys. **126**, 204901 (2007).

[108] A. Milchev, K. Binder, and A. Bhattacharya, *Polymer translocation through a nanopore induced by adsorption: Monte Carlo simulation of a coarse-grained model*, J. Chem. Phys. **121**, 6042 (2004).

[109] R. Metzler and J. Klafter, *When Translocation Dynamics Becomes Anomalous*, Biophys. J. **85**, 2776 (2003).

[110] A. Gopinathan and Y. W. Kim, *Polymer Translocation in Crowded Environments*, Phys. Rev. Lett. **99**, 228106 (2007).

[111] A. B. Fulton, *How crowded is the cytoplasm?*, Cell **30**, 345 (1982).

[112] A. P. Minton, *Effect of a Concentrated "Inert" Macromolecular Cosolute on the Stability of a Globular Protein with Respect to Denaturation by Heat and by Chaotropes: A Statistical-Thermodynamic Model*, Biophys. J. **78**, 101 (2000).

[113] H. X. Zhou, *Loops, Linkages, Rings, Catenanes, Cages, and Crowders: Entropy-Based Strategies for Stabilizing Proteins*, Acc. Chem. Res. **37**, 123 (2004).

[114] W. H. Coulter, *Means for Counting Particles Suspended in a Fluid*, U.S. Patent 2,656,508 (1953).

[115] J. W. F. Robertson, C. G. Rodrigues, V. M. Stanford, K. A. Rubinson, O. V. Krasilnikov, and J. J. Kasianowicz, *Single-molecule mass spectrometry in solution using a solitary nanopore*, Proc. Natl. Acad. Sci. USA **104**, 8207 (2007), and the references therein.

[116] R. E. Thompson and E. D. Siggia, *Physical Limits on the Mechanical Measurement of the Secondary Structure of Bio-molecules*, Europhys. Lett. **31**, 335 (1995).

[117] U. Gerland, R. Bundschuh, and T. Hwa, *Translocation of structured polynucleotides through nanopores*, Phys. Biol. **1**, 19 (2004).

[118] R. Bundschuh and U. Gerland, *Coupled Dynamics of RNA Folding and Nanopore Translocation*, Phys. Rev. Lett. **95**, 208104 (2005).

[119] S. Kotsev and A. B. Kolomeisky, *Effect of orientation in translocation of polymers through nanopores*, J. Chem. Phys. **125**, 084906 (2006).

[120] H. Vocks, D. Panja, G. T. Barkema, and R. C. Ball, *Pore-blockade Times for Field-Driven Polymer Translocation*, J. Phys.: Condens. Matter **20**, 095224 (2008).

[121] I.-C. Yeh and G. Hummer, *Diffusion and Electrophoretic Mobility of Single-Stranded RNA from Molecular Dynamics Simulations*, Biophys. J. **86**, 681 (2004).

[122] C. Heneghan and G. McDarby, *Establishing the relation between detrended fluctuation analysis and power spectral density analysis for stochastic processes*, Phys. Rev. E **62**, 6103 (2000).

[123] H. E. Hurst, *Long-term storage capacity of reservoirs*, Trans. Am. Soc. Civ. Eng. **116**, 770 (1951).

Samenvatting

(summary in Dutch)

In dit proefschrift bestuderen we de translocatie van polymeren door een klein gaatje in een membraan met behulp van Monte Carlo simulaties.

Polymeren vormen een belangrijke klasse van materialen. Er zijn natuurlijke polymeren, zoals eiwitten, RNA, DNA, natuurlijk rubber, cellulose, zetmeel. Maar er zijn ook synthetische polymeren. Overal om ons heen worden deze synthetische polymeren toegepast. Dit is gemakkelijk te verklaren: ze zijn immers eenvoudig te fabriceren, ze hebben goede mechanische eigenschappen en ze zijn bovenal goedkoop. Het bekendste voorbeeld wordt gevormd door polyetheen: één procent (> 60 miljoen ton per jaar) van de wereldwijd gewonnen aardolie gaat op aan de productie ervan.

Een polymeer is een molecuul dat bestaat uit een groot aantal identieke of soortgelijke delen die met elkaar verbonden zijn middels covalente bindingen. De herhalende eenheid wordt "monomeer" genoemd. In dit proefschrift kijken we uitsluitend naar flexibele, lineaire polymeren. Lineaire polymeren hebben slechts twee uiteinden. Polymeerfysici drukken flexibiliteit uit in termen van de "persistentie lengte", $l_{\rm p}$. Een polymeer van lengte L bestaat dan effectief uit een keten van $N \equiv L/l_{\rm p}$ staafjes elk met lengte $l_{\rm p}$. Het grote aantal vrijheidsgraden ($N \gg 1$) zorgt ervoor dat een enkele keten zeer veel verschillende conformaties kan aannemen, wat uitnodigt tot een statistische beschrijving van het systeem. Fysici zijn dan ook voornamelijk geïnteresseerd in gemiddelde grootheden, zoals de gemiddelde afmeting van een enkele keten ("gyratiestraal").

De gyratiestraal van flexibele, lineaire polymeren voldoet aan de schalingswet $R_{\rm g} \sim N^{\nu}$, met de "Flory-exponent" $\nu = 0.588$: knip je een lange keten in

twee stukken, dan zal de gemiddelde afmeting reduceren met een factor (van ongeveer) 1.5. Dit geldt zowel voor DNA, RNA, ketens van polyetheen, maar ook schoenveters, gekookte spaghetti, tuinslangen (allemaal in oplossing) en zelfs "virtuele polymeren" in een computergeheugen. De Flory-exponent is voor alle ketens hetzelfde, universeel. In hoofdstuk 1 zullen we zien dat vele eigenschappen van polymeren universeel zijn.

Theoretische fysici gebruiken pen en papier (en het softwarepakket mathematica) om fundamentele (fysische) problemen op te lossen. Fysici zijn net gewone mensen: goede ideeën komen nooit helemaal vanzelf. Een manier om nieuwe wegen in te slaan is door gebruik te maken van computersimulaties. Computers zijn heel erg goed in het oneindig vaak herhalen van eenvoudige berekeningen. Dat maakt ze heel geschikt om virtuele polymeren te simuleren: in elke tijdstap worden alle virtuele monomeren bewogen volgens eenvoudige regels, die we tezamen dynamica noemen. Deze modelpolymeren zullen dezelfde universele eigenschappen vertonen als echte polymeren: ze hebben bijvoorbeeld een gyratiestraal die ook schaalt met N als N^{ν}.

Computersimulaties zijn een fantastisch hulpmiddel: de programmeur kiest namelijk zelf de regels volgens welke de monomeren zich voortbewegen. Als we de dynamica aanpassen, zullen zekere dynamische eigenschappen van de polymeren veranderen. Op die manier is het mogelijk om beter inzicht te verkrijgen in de fysische oorsprong van polymeereigenschappen die in het laboratorium gemeten worden. Ook kunnen "experimenten" worden uitgevoerd, die nog nooit in een laboratorium gedaan zijn. Deze vorm van fysica is experimenteel van aard, maar zonder de noodzaak om de handen vuil te maken. In hoofdstuk 2 wordt uitgebreid aandacht besteed aan het extended repton model. Alle simulaties van de translocatie van polymeren zijn verricht met dit Monte Carlo algoritme.

Voor veel biologische processen speelt het transport van moleculen door een membraan een grote rol. Messenger RNA moleculen bijvoorbeeld worden uit de celkern getransporteerd door het zogenaamde "nuclear pore complex". Denk ook aan de injectie van viraal DNA in gastheercellen. Vaak zijn de moleculen erg lang, en is het gaatje in het membraan juist erg klein. Zo'n molecuul kan er alleen maar doorheen gaan, door zich te vervormen. In dit proefschrift bestuderen we de translocatie van polymeren door een gaatje dat zo klein is dat er maar één monomeer tegelijk in past.

Gezien de grote rol die polymeertranslocatie speelt in vele levensprocessen, is het niet verwonderlijk dat er al heel veel praktisch onderzoek naar de

translocatie van polymeren is gedaan. Maar dit empirisch onderzoek is voornamelijk verricht vanuit het gezichtspunt van de biologie. De laatste jaren is de studie van translocatie als fysisch proces ook tot wasdom gekomen. Met behulp van computersimulaties is onderzocht [37, 41] hoe lang een polymeer van lengte L er over doet om zich door een gaatje heen te wringen. Zodra een monomeer van links naar rechts door het gaatje gaat, wordt de spanning links in de keten iets groter, en rechts juist kleiner: de monomeer wordt dus teruggetrokken. Tengevolge van dit spanningsverschil zal de translocatie "anomale diffusie" vertonen.

Sinds kort is het mogelijk om (gecontroleerd) gaatjes in membranen te maken met een diameter van enkele nanometers. Tengevolge hiervan is er veel interesse voor translocatie vanuit het oogpunt van de technologie. Een mogelijke toepassing wordt gevonden in de analyse en detectie van afzonderlijke macromoleculen. In deze experimenten worden geladen polymeren in een elektrolytische oplossing aan één zijde van een membraan geplaatst. De moleculen kunnen het membraan alleen passeren door door het kleine gaatje heen te gaan. Door nu een elektrisch spanningsverschil aan te brengen over de twee zijden van het membraan, zullen de geladen moleculen de opening willen passeren. Wanneer een molecuul het gat binnengaat, dan heeft dit invloed op de elektrische weerstand van het systeem. De verandering van deze elektrische weerstand leidt tot een dip in de elektrische stroom. De diepte en duur van de dip blijken een goede maat te zijn voor de grootte en lengte van het molecuul dat door het gat heengaat. In hoofdstuk 3 bestuderen we hoe translocatie van polymeren verloopt als deze gedreven wordt door een elektrisch spanningsverschil.

Biologische cellen zijn geen zakjes gevuld met slechts water: het cytoplasma is gevuld met allerlei organellen en ook een groot aantal biologische moleculen. In hoofdstuk 4 bestuderen we hoe de translocatie van polymeren verloopt in een omgeving gevormd door andere polymeren. Het zal blijken dat het van belang is of deze omgevingspolymeren lang zijn of juist kort. Een kluwen van lange polymeren lijkt een beetje op een bord spaghetti: trek je een enkele sliert hieruit, dan zal deze zich door het netwerk van spaghetti heen moeten bewegen. Tengevolge hiervan zal de tijd die het polymeer nodig heeft om het membraan te passeren sterk toenemen.

Ribonucleïnezuur (RNA) is samengesteld uit vier verschillende nucleotiden. Elke nucleotide bevat een fosfaatgroep, een suiker groep (ribose) en een base. De polymeer wordt gevormd door de binding van twee fosfaatgroepen,

en de ribose ring verbindt een base aan de fosfaatgroep. De nucleotiden verschillen alleen van elkaar door hun basen. De vier varianten zijn adenine (A), cytosine (C), guanine (G) en uracil (U). In RNA kunnen twee soorten basenparen gevormd worden: guanine kan zich binden aan cytosine, en adenine kan zich binden aan uracil, met drie respectievelijk twee waterstofbruggen. Deze worden de Watson-Crick basenparen genoemd. De drie-dimensionale structuur van het RNA-molecuul die ontstaat door de vorming van basenparen, noemt men de secundaire structuur. Voor het correct functioneren van veel RNA moleculen is de secundaire structuur van zeer groot belang, vaak nog belangrijker dan de exacte volgorde van de nucleotiden.

Sinds kort is het mogelijk om afzonderlijke RNA moleculen te manipuleren met bijvoorbeeld een optisch pincet. Daarom bestuderen we in hoofdstuk 5 het volgende translocatie experiment: we trekken, met behulp van een optisch pincet, een RNA molecuul door een gaatje dat zo klein is dat er maar één nucleotide tegelijk doorheen kan. Eventuele aanwezige basenparen zullen verbroken moeten worden, en dat vereist een zekere kracht. Door te meten hoe de benodigde kracht gedurende het experiment varieert, kunnen we afleiden hoeveel basenparen er verbroken zijn, en daarmee is het mogelijk om een idee te krijgen van de secundaire structuur van het molecuul. Althans, dat is de theorie. In hoofdstuk 5 laten we zien dat de nauwkeurigheid van dit experiment fundamenteel beperkt wordt door thermische fluctuaties. Het zal blijken dat de gewenste ideale resolutie van een enkele nucleotide, onder normale experimentele omstandigheden, niet gehaald kan worden.

List of Publications

This thesis is based on the following articles:

- H. Vocks, D. Panja, G. T. Barkema, and R. C. Ball, *Pore-blockade Times for Field-Driven Polymer Translocation*, J. Phys.: Condens. Matter **20**, 095224 (2008).
- H. Vocks, D. Panja, and G. T. Barkema, *Can Translocation Be Used to Determine RNA Secondary Structures?*, e-print arXiv: 0801.3947; submitted.
- H. Vocks, D. Panja, and G. T. Barkema, *Polymer Translocation in a Crowded Environment*. To be submitted.

Other articles to which the author has contributed:

- H. Vocks, M. V. Chubynsky, G. T. Barkema, and N. Mousseau, *Activated sampling in complex materials at finite temperature: The properly obeying probability activation-relaxation technique*, J. Chem. Phys. **123**, 244707 (2005).
- N. Mousseau, M. Chubynsky, H. Vocks, and G. T. Barkema, *Exploiting memory in event-based simulations*, J. Non-Crys. Sol. **352**, 4424 (2006).
- K. D. Meisel, H. Vocks, and P. A. Bobbert, *Polarons in semiconducting polymers: Study within an extended Holstein model*, Phys. Rev. B **71**, 205206 (2005).

Dankwoord

(acknowledgements)

Promoveren doe je niet alleen en ik wil iedereen die er op de een of andere manier aan heeft bijgedragen dan ook hartelijk bedanken. Een aantal mensen wil ik nu speciaal vermelden.

First I would like to thank Normand Mousseau. Thanks to your unwavering support and enthousiasm we were very close to making POP-ART a success. Your hospitaly during my visit to Montreal was more than I could have asked for.

I would also like to thank Mykyta Chubynsky. Your contributions to POP-ART were of great importance.

Natuurlijk dank ik ook mijn promoter Gerard Barkema. Dank voor je steun gedurende de afgelopen vier jaar, maar in het bijzonder voor de levendige discussies die wij heden ten dage voeren over allerlei aspecten van de polymeerfysica.

Many thanks to Debabrata Panja. Without your support I would never have been able to make these rapid strides in the world of polymer physics. We often disagree on just about everything, but I think your critical attitude has truly contributed to our rapid advances in the understanding of several phenomena.

Ook bedank ik Henk van Beijeren voor zijn vele opbouwende op- en aanmerkingen over de tekst van dit proefschrift.

Natuurlijk dank ik de promovendi en post-docs die deel uitmaakten van het ITF toen ik daar voor het eerste voet binnen zette. Men name Bas, Michiel, Mathijs, Joanne, Masud en Astrid, die allen altijd wel een interessante mening hadden over zo'n beetje alles. Uiteraard dank ik Dennis en mijn eerste

kamergenoot Vladimir. Wij drieën hebben samen menig biertje gedronken. Van de “tweede” generatie dank ik Jos, mijn kamergenoot Markus, en bovenal Eelco voor die dagelijkse kop koffie.

Geertje, Biene en Wilma, jullie vormen de sociale lijm die het ITF bijeenhoudt.

Bovenal dank ik mijn Lieve Pief Suzanne. Dank je voor je steun, en de blijvende lol die we altijd samen beleven. Hier heb je ’m dan eindelijk!

Curriculum Vitae

April 1, 1979	Born in Venlo, The Netherlands
1991 - 1997	Grammar school, Collegium Marianum, Venlo
1997 - 2004	M.Sc. (cum laude) in applied physics, Eindhoven University of Technology, Eindhoven
	Traineeship at Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology, USA. Subject: *Junction magnetoresistance in free-electron-like models and artificially doped magnetic tunnel junctions*, supervised by dr. J. S. Moodera and prof. dr. B. Koopmans.
	Master thesis project at the Theoretical and Polymer Physics group (PFY) of the departement of applied physics, Eindhoven University of Technology, Eindhoven. Subject: *Semi-classical and Quantum polarons in π-conjugated organic materials*, supervised by dr. P. Bobbert.
2004 - 2008	Ph.D. project at the Institute for Theoretical Physics, Utrecht University, Utrecht
2004 - 2006	Subject: *properly-obeyed probability activation-relaxation technique (POP-ART)*, supervised by prof. dr. N. Mousseau (University of Montreal, Montreal, Canada) and prof. dr. G. T. Barkema.
2007 - 2008	Subject: *Simulations of Polymer Translocation*, supervised by dr. D. Panja (Universiteit van Amsterdam, Amsterdam) and prof. dr. G. T. Barkema.