Active particles in one dimension

Asymptotic behaviour and collective dynamics

M. Meskers

Technische Universiteit Delft

TU Delft

Active particles in one dimension

Asymptotic behaviour and collective dynamics

by

Marthe Meskers

to obtain the degree of Bachelor of Science

at the Delft University of Technology,

to be defended on Wednesday July 14, 2021 at 10:00 AM.

Supervisor applied mathematics: Prof. dr. ir. F. H. J. Redig

Supervisor applied physics: Dr. T. Idema

Student number: 4850521

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Abstract

In this thesis, we study the asymptotic behaviour and the dynamics of a one-dimensional active particle model with excluded volume interactions. The model is a version of run-and-tumble motion, where a particle performs both symmetric random walks and active transport. The direction and the speed of the transport are governed by an internal state process.

We show that this motion converges to Brownian motion upon diffusive scaling and determine the limiting diffusion coefficient. The internal state converges to a stationary distribution, by which it manifests itself in the diffusion coefficient. Furthermore, we prove that the active particle satisfies the large deviation principle. This allows us to derive an implicit expression for the rate by which the probability of rare events tends to zero.

Numerically, we investigate the influence of excluded volume interactions on the diffusion coefficient and the average velocity. We find that the velocity converges exponentially to its theoretical value as the number of particles allowed per position increases. In addition, this exclusion number strongly influences the manner in which the velocity decreases for high particle densities.

Predictions for the velocity as a function of particle density based on the model are compared to experimental data of the molecular motor kinesin-II. We find that model is not adequate for approximating the velocity of molecular motors in crowded environments and extensions in the form of Langmuir kinetics are suggested.

Contents

1

Introduction

Ensembles of many constituents that each dissipate energy from an internal or external source to perform directed motion are called active matter systems. As a result of the conversion of energy into motion, systems of active matter defy the laws of equilibrium statistical physics. Over the past decades, active systems have drawn increasing interest because of their rich emergent behaviour such as phase transitions and collective motion [22]. Many models have been developed in order to understand collective phenomena from the motile properties of the individual, interacting constituents of a system. Examples of active matter on the nanoscale include bacteria and molecular motors [17]. On a larger scale, emergent phenomena in active systems include the flocking of birds and the schooling of fish.

Considering active particles on a micro- and nanoscale, the motion of the particles has an active component due to dissipation and a stochastic component due to thermal fluctuations, also referred to as Brownian motion. In this thesis, we study a version of a one-dimensional active particle model known as the run-and-tumble model (see for example [6, 13, 20]). Run-and-tumble particles combine random and driven motion, where the direction of the driven motion shows sudden reversals. This type of motion has been observed for different types of bacteria [11, 17].

The questions which we ask ourselves for the active particle model are centered around four main limit theorems for stochastic processes. These form the foundation on which we build our analysis of the asymptotics of the run-and-tumble model. Let us give a brief introduction to the limit theorems for a random walk process X_t in continuous time. Throughout the report the following principles will be thoroughly explained.

1. Law of large numbers: $\lim_{t\to\infty} \frac{X_t}{t} = \lim_{t\to\infty} \frac{\mathbb{E}X_t}{t} = v$
   First of all, the law of large numbers states that X_t has a limiting asymptotic velocity v, which is almost surely constant.

2. Central limit theorem: $\frac{X_t - vt}{\sqrt{t}} \xrightarrow{d} \mathcal{N}(0, \sigma^2)$
   Secondly, the central limit theorem prescribes that after scaling by a factor $\frac{1}{\sqrt{t}}$, $X_t - vt$ converges in distribution to a Gaussian with mean zero and variance σ^2.

3. Invariance principle: $\varepsilon\left(X_{\varepsilon^{-2}t} - v\varepsilon^{-2}t\right) \xrightarrow{\varepsilon\to 0} W_{\sigma^2 t}$
   This principle is a strengthening of the central limit theorem. It states that under appropriate scaling of time and position, the entire process - in contrast to the distribution at $t = 1$ for the central limit theorem - converges to Brownian motion.

4. Large deviations: $\mathbb{P}\left(\frac{X_t}{t} \approx x\right) \approx e^{-tI(x)}$
   Finally, the large deviation principle is about the asymptotic behaviour of the moment generating function of a random variable. In short, it states that the probability that $\frac{X_t}{t}$ is approximately equal to x, decreases exponentially in t. The leading order in the exponent is $I(x)$, a function that depends on the value x and the moment generating function of X_t. Because by the law of large numbers, $\frac{X_t}{t}$ is approximately equal to v with probability 1, it holds that $I(v) = 0$.

Here we will adopt a version of the run-and-tumble model on a lattice. The aim of this study is to derive its asymptotic behaviour. In order to do so, we will derive the large deviation properties of the process and determine if the invariance principle applies. The setting which we consider is simple in comparison to actual one-dimensional active processes. However, in contrast to other models, this particular model has been shown to allow for exact computation of for instance the limiting diffusion coefficient and the asymptotic velocity [21].

Moreover, we aim to provide a better understanding of the dynamics of run-and-tumble particles in presence of excluded volume interactions and relate the dynamics to the Asymmetric Simple Exclusion Process (ASEP), for which analytical results are known [9]. Finally, we will compare predictions based on run-and-tumble motion with an exclusion principle to data for the biological process of intracellular transport by molecular motors.

This report is structured as follows. In chapter 2 we focus on the theory of Markov processes, a category under which the run-and-tumble process falls. Next, in chapter 3, we discuss Brownian motion and the invariance principle, followed by an explanation of large deviations in chapter 4. The next two chapters contain our analysis of the asymptotic behaviour of a one-dimensional active particle model, where we first derive the central limit theorem and subsequently determine the large deviation properties. In chapter 7 we study the influence of excluded volume interactions among active particles on the transport coefficients. An application of the model to the motion of the molecular motor kinesin-II is presented in chapter 8. Finally, our conclusions and directions for future research are given in chapter 9.

2

Markov processes

This chapter provides background information on Markov processes, a type of stochastic process to which the run-and-tumble process belongs. We will focus on theory of Markov processes in continuous time and include an explanation of the Poisson distribution and exponential distribution, which are closely related to each other and to continuous-time Markov processes. In addition, the concept of a Markov semigroup and generator will be introduced. Finally we discuss stationary distributions and the detailed balance conditions by which they can be found. This chapter is based on literature by G. Grimmett and D. Welsh [10] and F. Redig [18].

2.1. Markov property

A stochastic process $\{X_t, t \geq 0\}$ on a state space Ω is a Markov process if the distribution of future states of the process is only dependent on the current state and not on the further past. A process which satisfies this Markov property is also called memoryless. Mathematically, a Markov process is defined as follows [18]:

Definition 2.1 *If for all $t > 0$, $n \in \mathbb{N}$, $0 < t_1 < ... < t_n < t$ and for all $f : \Omega \to \mathbb{R}$ bounded and measurable*

$$\mathbb{E}(f(X_t)|X_{t_1}, X_{t_2}, ..., X_{t_n}) = \mathbb{E}(f(X_t)|X_{t_n}) \tag{2.1}$$

*then $\{X_t, t \leq 0\}$ is a **Markov process**.*

Different types of stochastic processes can have the Markov property. A well-known example is the random walk, where the next position that a particle takes does not depend on where it has been in the past but only on its current position and on random, independent increments. If the particle makes random jumps at fixed time intervals, this is a discrete-time Markov process. The events of a Markov process can also take place in continuous time. In that case the process is a family $\{X_t, t \geq 0\}$ of random variables indexed by continuous time.
Because this project is about continuous-time Markov processes, these will be covered in more detail in the rest of this chapter. We begin with the fundamentals of memoryless processes in continuous time: the exponential and Poisson distribution and their connection.

2.2. Poisson process

The Poisson process $(N_t : t \geq 0)$ models the counting of random events in continuous time. For example, the number of decays of a radioactive source or the number of incoming calls at a call center can be described by a Poisson process. $(N_t : t \geq 0)$ makes transitions of $+1$ and is constant and right-continuous in between two transitions, as represented by figure 2.1. The number of transitions in a time interval is assumed to depend only on the length of this interval and not on the moment in time. To summarise, a Poisson process satisfies the following conditions:

(a) $N_0 = 0$

(b) Stationarity of the increments: $N_{t+r} - N_{s+r}$ has the same distribution as $N_t - N_s$ for $r > 0$ and $t \geq s \geq 0$.

(c) Independence of the increments: for $0 \le t_1 < ... < t_n$, $n \ge 1$ the increments $N_{t_1} - N_0$, $N_{t_2} - N_{t_1}$, ..., $N_{t_n} - N_{t_{n-1}}$ are independently distributed.

(d) Arrival rate: the process has an arrival rate $\lambda > 0$, i.e., for small positive h:
$\mathbb{P}(N_{t+h} - N_t = 0) = 1 - \lambda h + o(h)$
$\mathbb{P}(N_{t+h} - N_t = 1) = \lambda h + o(h)$

Note that a function f satisfies $f(h) = \mathrm{o}(h)$, referred to as Landau's notation, if $f(h)/h \to 0$ in the limit $h \to 0$.

Using the probability generating function, it can be derived from conditions (a)-(d) that N_t has the Poisson distribution. For the derivation, we refer to chapter 9 of [16].

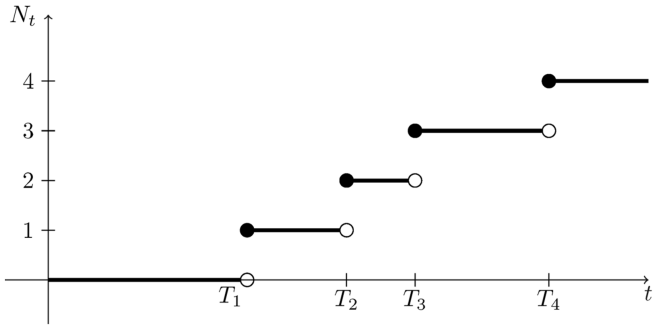


Figure 2.1: A representation of the Poisson process N_t with arrival times T_i of the events.

Poisson distribution A random variable N_t has a Poisson distribution with parameter λt if

$$P(N_t = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}. \tag{2.2}$$

The variance and mean of N_t are given by

$$\mathbb{E}(N_t) = \lambda t \qquad \mathrm{Var}(N_t) = \lambda t \qquad t > 0.$$

Another property of the Poisson process which we will use when deriving the generator of a continuous-time random variable is the following. For small h, the probability of two or more events happening within the time window $(t, t+h]$ is not significant, i.e. of order $o(h)$. This can be derived from property (d) of a Poisson process, namely:

$$\begin{aligned} \mathbb{P}(N_{t+h} - N_t \ge 2) &= 1 - \mathbb{P}(N_{t+h} - N_t = 0) - \mathbb{P}(N_{t+h} - N_t = 1) \\ &= 1 - (1 - \lambda h + o(h)) - (\lambda h + o(h)) \\ &= o(h). \end{aligned} \tag{2.3}$$

Let us now look at the exponential distribution, and establish its connection with a Poisson process.

Exponential distribution A random variable X has an exponential distribution with parameter $\lambda > 0$ if the probability density of X is

$$f(x) = \lambda e^{-\lambda x}, \qquad x > 0. \tag{2.4}$$

The arrival times T_i of Poisson events are defined by

$$T_0 = 0$$

$$T_i = \inf\{t : N_t = i\} \quad \text{for } i = 1, 2, ...,$$

such that the times between successive arrivals are

$$\tau_i = T_i - T_{i-1}.$$

Theorem 2.2 *For a Poisson process with parameter λt the inter-arrival times τ_i are independent random variables, having the exponential distribution with parameter λ.*

Proof. From (2.2) it immediately follows that τ_1 is exponentially distributed:

$$\mathbb{P}(\tau_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}.$$

For $n \geq 1$ we have:

$$\begin{aligned}
\mathbb{P}(\tau_n > t \,|\, \tau_{n-1} = s) &= \mathbb{P}(\text{no event in} (s, s+t] \,|\, \tau_{n-1} = s) \\
(\text{indep. of increments}) &= \mathbb{P}(\text{no event in } (s, s+t]) \\
(\text{stationarity of increments}) &= \mathbb{P}(\text{no event in } (0, t]) \\
&= e^{-\lambda t},
\end{aligned}$$

such that τ_n and τ_{n-1} are independent and $\tau_n \sim \exp(\lambda)$ for $n = 1, 2, \ldots$ □

2.3. Continuous-time Markov processes

In the beginning of this chapter we defined what memorylessness means for a stochastic process, a sequence of random variables (RV's). For just one RV, the following definition applies:

Definition 2.3 *A positive random variable X has the **lack-of-memory property** if:*

$$\mathbb{P}(X > a + b) | X > a) = \mathbb{P}(X > b) \quad \text{for } a, b \geq 0 \tag{2.5}$$

The following theorem gives a consequence for random variables that are continuous and lack memory:

Theorem 2.4 *A continuous random variable X satisfies the lack-of-memory property if and only if X is exponentially distributed.*

Proof. (based on the proof by Grimmet and Welsh in [10]) First, let X be exponentially distributed with parameter λ, such that its probability density function is given by (2.4). From this density function it follows that $\mathbb{P}(X > u) = e^{-\lambda u}$. For $v, w \geq 0$ we have:

$$\begin{aligned}
\mathbb{P}(X > v + w \mid X > v) &= \frac{\mathbb{P}(X > v + w \text{ and } X > v)}{\mathbb{P}(X > v)} \\
&= \frac{\mathbb{P}(X > v + w)}{\mathbb{P}(X > v)} \\
&= \frac{e^{-\lambda(v+w)}}{e^{-\lambda v}} \\
&= e^{-\lambda w} = \mathbb{P}(X > w),
\end{aligned} \tag{2.6}$$

such that X has the lack-of-memory property.
For the converse, suppose that X is continuous, positive and satisfies the lack-of-memory property. Let $H(v) := \mathbb{P}(X > v)$ for $v \geq 0$. We will show that $H(v) = e^{-\lambda v}$. As follows from (2.6), H satisfies $H(v + w) = H(v)H(w)$ for $v, w \geq 0$. By taking $v = w$ it follows that $H(2v) = H(v)^2$. By induction it can be shown that $H(nv) = H(v)^n$ for $n = 0, 1, 2, \ldots$. Letting $v = \frac{v}{n}$ gives $H(v) = H(\frac{v}{n})^n$ and thus $H(v)^{\frac{1}{n}} = H(\frac{v}{n})$. It now also follows that $H(v)^{\frac{m}{n}} = H(\frac{m}{n}v)$ for $m \in \mathbb{N}$. For $v = 1$ we have $H(1)^{\frac{m}{n}} = H(\frac{m}{n})$, which defines the function $e^{-\lambda v}$ for non-negative rationals v. We consider $\lambda > 0$ to obtain a well defined probability density. By taking limits from the right in the rational numbers, we obtain $H(v) = e^{-\lambda v}$ for for arbitrary $v \in \mathbb{R}$. □

Transition rates Now let X_t be a continuous-time Markov process on a finite state space Ω. For such a process we do not consider transition probabilities - as for a discrete process - but 'rates' $c(x, y) \geq 0$ for a transition from x to y. These can be interpreted as transition probabilities per unit of time. Because X_t is

Markov, the time between transitions is exponentially distributed. This can be seen as follows. Let T_x be the first time a jump away from position x takes place. We then have

$$\begin{aligned}\mathbb{P}(T_x > t+s | T_x > s) &= \mathbb{P}(X_v = x, \forall\, s \le v \le t+s | X_w = x, \forall\, 0 \le w \le s) \\ &= \mathbb{P}(X_v = x, \forall\, s \le v \le t+s | X_s = x) \\ &= \mathbb{P}(T_x > t).\end{aligned}$$

This implies T_x has the lack-of-memory property, according to definition 2.3. By theorem 2.4, the waiting time for a transition is exponentially distributed. The process with transition rates $c(x,y)$ is then described as follows. From position $X_t = x$, a jump takes place after an exponential time with rate $c_x = \sum_{y\in\Omega} c(x,y)$. With probability $p(x,y) = \frac{c(x,y)}{c_x}$ the position after this jump is y.

2.4. Markov semigroup

For a continuous-time Markov process $\{X_t, t \ge 0\}$ on a state space Ω, the semigroup S_t applied to a function $f : \Omega \to \mathbb{R}$ gives the expected value of $f(X_t)$.

$$S_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}_x(f(X_t)) = \sum_{y\in\Omega} \mathbb{P}(X_t = y|X_0 = x) f(y). \tag{2.7}$$

$\mathbb{P}(X_t = y|X_0 = x)$, the probability that the position at time t is y, given that the initial position is x, will be denoted by $p_t(x,y)$. For Ω finite, S_t can be viewed as a matrix with these probabilities as elements: $S_t(i,j) = p_t(i,j)$ for $i, j \in \Omega$. The function f can in this case be represented by the column vector $[f(y_1)f(y_2)...f(y_N)]^T$ for $\Omega = \{y_1, y_2, ..., y_N\}$. However, this representation only applies for processes on a finite state space, and in general S_t should be considered an operator instead of a matrix.
The transition probabilities of Markov process fulfill the Chapman-Kolmogorov equations, which read: $p_{t+s}(x,y) = \sum_{v\in\Omega} p_t(x,k) p_s(k,y)$. This ensures the semigroup property, which will be listed in the following proposition, along with other properties.

Proposition 2.5 *The Markov semigroup S_t satisfies properties:*

(a) Identity at time zero: $S_0 = I$, i.e. $S_0 f = f$ for all f

(b) Semigroup property: for all $t, s > 0, f : S_{t+s} f = S_t(S_s f) = S_s(S_t f))$

(c) Right continuity of the map $t \to S_t f$

(d) Positivity: $f \ge 0$ implies $S_t f \ge 0$

(e) Normalization: $S_t 1 = 1$

(f) Contraction: $\max_x |(S_t f)(x)| \le \max_x |f(x)|$

Proof. We prove the properties in the discrete case.

(a) This follows from the definition: $S_0 f(x) = \mathbb{E}(f(X_0)|X_0 = x) = f(x)$, which implies $S_0 = I$.

(b)

$$\begin{aligned}S_{s+t} f(x) &= \mathbb{E}(f(X_{s+t})|X_0 = x) = f(x) \\ &= \sum_{y\in\Omega} p_{t+s}(x,y) f(y) \\ &= \sum_{y\in\Omega} f(y) \sum_{v\in\Omega} p_t(x,k) p_s(k,y) \\ &= \sum_{v\in\Omega} p_t(x,k) \sum_{y\in\Omega} f(y) p_s(k,y) \\ &= \sum_{v\in\Omega} p_t(x,k) \mathbb{E}(f(X_s)|X_0 = k) \\ &= S_t(S_s f(x))\end{aligned}$$

By interchanging t and s it can be shown that the above expression also equals $S_s(S_t f(x))$.

(c) To show that the map $t \to S_t f(x)$ is right continuous, we need to show that $\lim_{t \downarrow c} S_t f(x) = S_c f(x)$.

We get $\lim_{t\downarrow c} S_t f = \lim_{\Delta t \downarrow 0} S_{c+\Delta t} f = S_c \left(\lim_{\Delta t \downarrow 0} S_{\Delta t} f(x) \right) = S_c(S_0 f(x)) = S_c f(x)$.

(d) For $f \geq 0$, it holds that $\mathbb{E}(f(x)) \geq 0$ for all x, so $S_t f(x) = \mathbb{E}(f(X_0)|X_0 = x) \geq 0$.

(e) $S_t 1 = \sum_{y\in\Omega} p_{t+s}(x,y)1 = 1$ because a sum of probabilities over the whole probability space is 1.

(f) $\max_x \left|(S_t f)(x)\right| = \max_x \left| \sum_{y\in\Omega} p_t(x,y) f(y) \right| \leq \max_x \sum_{y\in\Omega} p_t(x,y)|f(y)| \leq \max_x |f(x)| \sum_{y\in\Omega} p_t(x,y) = \max_x |f(x)|$

□

2.5. Generator

Let us consider $\frac{\mathrm{d}}{\mathrm{d}t} S_t|_{t=0}$. We define the domain $D(L)$ by

$$D(L) = \left\{ f : \lim_{t\to 0} \frac{S_t f - f}{t} \text{ exists} \right\}.$$

For functions $f \in D(L)$, the Markov generator corresponding to the Markov semigroup S_t, is given by [18]:

$$Lf := \frac{\mathrm{d}}{\mathrm{d}t} S_t f|_{t=0} = \lim_{t\to 0} \frac{S_t f - f}{t}. \tag{2.8}$$

From $S_{t+s} f = S_t S_s f = S_s S_t f$ (property (b) of the semigroup) and $\frac{\mathrm{d}}{\mathrm{d}t} S_t f|_{t=0} = Lf$, we can derive the following relation:

$$\frac{\mathrm{d}}{\mathrm{d}t} S_t = \frac{\mathrm{d}}{\mathrm{d}t} S_{t+s}|_{s=0} = LS_t = S_t L \tag{2.9}$$

In case X_t can take finitely many values, the semigroup and generator are matrices, related to each other by

$$\mathbf{S}_t = e^{t\mathbf{L}}, \tag{2.10}$$

where the exponent is defined in terms of its Taylor series:

$$e^{t\mathbf{L}} = \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n.$$

For L bounded, this exponent is always well-defined. In the next section we will see that for non-finite Ω, the set of functions for which (2.10) holds is restricted.

Let us return our attention to a general continuous-time Markov process where a transition from state x to state y takes place at a certain rate $c(x,y)$, as explained before. The total rate to exit state x is $c_x = \sum_{y\in\Omega} c(x,y)$. Since the number of transitions is Poisson-distributed, the probability that 2 or more transitions take place within the time interval $[0,t]$ is $o(t)$[1], as derived in (2.3).
This means the semigroup can be expressed for small t:

$$\begin{aligned} S_t f(x) &= \mathbb{E}_x(f(X_t)) = \mathbb{E}_x(f(X_t)\mathbb{1}_{\text{no jump}}) + \mathbb{E}_x(f(X_t)\mathbb{1}_{\text{one jump}}) + o(t) \\ &= f(x)\mathbb{P}(\text{no jump} \,|\, X_0 = x) + \sum_{y\in\Omega} f(y)\mathbb{P}(\text{one jump to } y \,|\, X_0 = x) + o(t) \\ &= f(x)e^{-c_x t} + \sum_{y\in\Omega} \frac{c(x,y)}{c_x} c_x t e^{-c_x t} f(y) + o(t) \\ &= f(x)(1 - c_x t) + \sum_{y\in\Omega} c(x,y) t f(y) + o(t) \\ &= f(x) + t \sum_{y\in\Omega} c(x,y)(f(y) - f(x)) + o(t). \end{aligned} \tag{2.11}$$

[1] This notation is called Landau's notation and represents a function that has a smaller order of magnitude than t as $t \to 0$. So $f(t) = o(t)$ when $f(t)/t \to 0$ as $t \to 0$.

From (2.8) and (2.11), it follows that the generator of a continuous-time random walk is then equal to

$$Lf(x) = \sum_{y\in\Omega} c(x,y)(f(y) - f(x)), \tag{2.12}$$

i.e., the sum of possible changes of the test function multiplied by the rate of the corresponding change.

Generators are particularly useful because convergence of the generator implies weak convergence of the corresponding process in topology on paths by the theorem of Trotter Kurtz. For more information about weak convergence in path space, which involves uniformity in time, see for example Billingsley [2].

2.6. Stationary distributions

Let us look at stationary distributions of continuous-time Markov processes. Consider the Markov process $\{X_t : t \geq 0\}$ with semigroup S_t, on the finite state space Ω. The probability distribution of X_t at time t is related to the initial distribution μ by $\mu_t = \mu S_t$ [18]. This can be written as

$$\mu_t(x) = \sum_{y\in\Omega} \mu(y) P_t(x,y), \tag{2.13}$$

where $P_t(x,y)$ is the matrix containing the transition probabilities $p_t(x,y)$, defined in section 2.4. Stated differently, μ_t is the unique probability distribution such that

$$\langle f, \mu_t\rangle = \langle S_t f, \mu\rangle \tag{2.14}$$

with $\langle f,\mu\rangle = \sum_{x\in\Omega} \mu(x) f(x)$.

Definition 2.6 *The probability distribution μ_t is stationary if $\mu_t = \mu$ for all t. This is the case if and only if*

$$\langle S_t f, \mu\rangle = \langle f, \mu\rangle \tag{2.15}$$

By dividing by t and taking the limit $t \to 0$, it follows from (2.15) that

$$\langle Lf, \mu\rangle = 0 \qquad \forall f \tag{2.16}$$

Conversely, (2.15) can be obtained from (2.16). Namely, if $\langle Lf,\mu\rangle = 0$ for all f, $\langle S_t f,\mu\rangle$ is constant by the relation $\frac{\mathrm{d}}{\mathrm{d}t} S_t f = L S_t f$.

From (2.12) we can derive that for a continuous-time Markov process, (2.16) implies

$$\sum_{y\in\Omega} [\mu(x)c(x,y) - \mu(y)c(y,x)] = 0. \tag{2.17}$$

A stronger condition for stationarity is to require that the terms in the sum of (2.17) are zero. This is called the detailed balance balance condition and is associated with equilibrium. For a process on a finite state space with transition rates $c(x,y)$, the stationary distribution μ can thus be found from the equation:

$$\mu(x)c(x,y) = \mu(y)c(y,x) \qquad \forall x,y \in \Omega \tag{2.18}$$

3

Brownian motion

This chapter is about Brownian motion, a process that is closely connected to continuous-time random walk. Robert Brown discovered the motion when he viewed pollen grains in water under a microscope. A mathematical description, idealizing the motion of the pollen, was later given by Wiener. The Wiener process, as it is also called, has a continuous state space and is defined as follows.

Definition 3.1 *Brownian motion* $\{W_t : t \geq 0\}$ *satisfies the following properties:*

(a) $W_0 = 0$

(b) The increments are independent: for $0 \leq t_1 \leq t_2 \leq ... \leq t_n : W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, ..., W_{t_n} - W_{t_{n-1}}$ *are independent*

(c) The increments are normally distributed: $W_{t_i} - W_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$, *with* $\mathcal{N}(\mu, \sigma)$ *the normal distribution with mean* μ *and variance* σ

(d) The paths are continuous: $t \to W_t$ *is continuous*

One can also consider Brownian motion starting from a position x, which is defined via $X_t = x + \mathcal{N}(0, t)$. From the independence of the increments it can be shown that $\{W_t, t \geq 0\}$ is a Markov process.

For Brownian motion on $\mathbb{R}$, the semigroup applied to a function $f \in C_0$, with $C_0 = \{f : \mathbb{R} \to \mathbb{R}$, continuous and $\lim_{x \to \pm\infty} f(x) = 0\}$, is given by

$$S_t f(x) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} e^{\frac{-(y-x)^2}{2t}} f(y)\, \mathrm{d}y. \tag{3.1}$$

As the integral of (3.1) consists of continuous functions, the semigroup is a mapping $S_t f : C_0 \to C_0$.

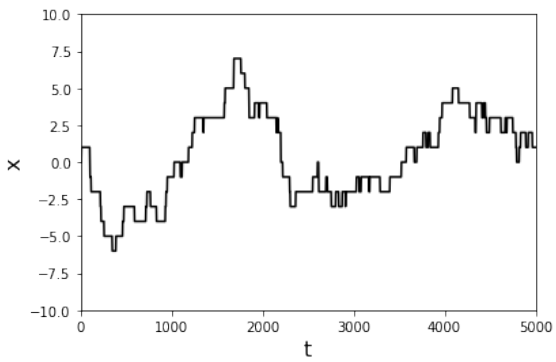
Let us determine the generator associated with this Markov process. To this end, the semigroup will be Taylor expanded. We make use of the known first and second moment of the normal distribution: $\mathbb{E}(\mathcal{N}(0, t)) = 0$ and $\mathbb{E}(\mathcal{N}^2(0, t)) = t$. Take $f : \mathbb{R} \to \mathbb{R}$, $f \in C_0^\infty$, i.e. f is in the class C_0 and is infinitely differentiable. Then,

$$\begin{aligned} S_t f(x) &= \mathbb{E}[f(x + \mathcal{N}(0, t)] \\ &= \mathbb{E}[f(x) + f'(x)\mathcal{N}(0, t) + \frac{1}{2} f''(x)\mathcal{N}(0, t)^2 + o(t)] \\ &= f(x) + \frac{1}{2} f''(x)\, t + o(t). \end{aligned}$$

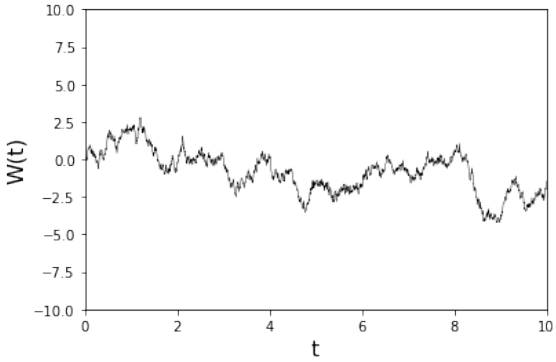
Using (2.8), it follows that the generator of Brownian motion is given by $L_B f(x) = \frac{1}{2} f''(x)$. We see that $L_B f(x)$ is not defined for all $f \in C_0$. It has a non-trivial dense domain $D(L_B) \subset C_0$. In this case $S_t f(x)$ is the unique solution to the partial differential equation

$$\frac{\partial}{\partial t} \Psi(t, x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \Psi(t, x),$$

with $\Psi(0, x) = f(x)$.



(a) Continuous-time random walk



(b) Wiener process

Figure 3.1: A continuous-time symmetric random walk with Poisson jump times and a realisation of Brownian motion.

3.1. Brownian motion as scaling limit

Suppose we have a symmetric random walk X_t, where a particle jumps to the right with rate $\frac{1}{2}$ and to the left with the same rate. According to (2.12), the generator of this continuous-time process is:

$$Lf(x) = \frac{1}{2}\big(f(x+1) + f(x-1) - 2f(x)\big). \tag{3.2}$$

Let us look at this walk on a continuous space instead of on a lattice. To do so, the spacing between the lattice sites should be decreased, i.e. the scaling $x \to \varepsilon x$ for $\varepsilon > 0$ will be applied and the limit $\varepsilon \to 0$ will be taken.
If not only space is scaled but time as well, the process converges weakly on path space, as stated in the following theorem:

Theorem 3.2 *The process $X_t^\varepsilon := \varepsilon X_{t/\varepsilon^2}$, obtained from X_t by scaling $x \to \varepsilon x$ and $t \to \frac{t}{\varepsilon^2}$, converges weakly to Brownian motion W_t in topology on paths.*

Proof. To show weak convergence of X_t^ε to W_t, we will show convergence of the generator of X_t^ε to the generator $L_B f(x)$ of Brownian motion, as described in section 2.5.
Since $Lf(x) = \lim_{t \downarrow 0} \frac{S_t f(x) - f(x)}{t}$, scaling time as $t \to \frac{t}{\varepsilon^2}$ is equivalent to considering $\frac{Lf(x)}{\varepsilon^2}$, the generator of the original process divided by ε^2. After additionally applying a scaling of space to this generator, we obtain the following generator for X_t^ε:

$$L_\varepsilon f(x) = \frac{\frac{1}{2}\big(f(x+\varepsilon) + f(x-\varepsilon) - 2f(x)\big)}{\varepsilon^2}. \tag{3.3}$$

Now let $f \in D(L_B)$ be arbitrary. We will prove that $\lim_{\varepsilon \to 0} L_\varepsilon f(x) = L_B f(x)$ by Taylor expansion of $f(x+\varepsilon)$ and $f(x-\varepsilon)$:

$$\begin{aligned} \lim_{\varepsilon \to 0} L_\varepsilon f(x) &= \lim_{\varepsilon \to 0} \frac{\frac{1}{2}\Big(f(x) + \varepsilon f'(x) + \frac{\varepsilon^2}{2} f''(x) + f(x) - \varepsilon f'(x) + \frac{\varepsilon^2}{2} f''(x) + o(\varepsilon^2)\Big)}{\varepsilon^2} \\ &= \frac{1}{2} f''(x) = L_B f(x). \end{aligned} \tag{3.4}$$

□

To form an intuition on Brownian motion as the scaling limit of random walk, see figure 3.1, which shows a realisation of a continuous-time random walk next to a realisation of the Wiener process.

3.2. Convergence of the characteristic function

It is not always possible to show convergence of a process by determining the limit of the generator. In section 5 we will for example encounter the following generator

$$L_\varepsilon f(x,\sigma) = \lambda f''(x,\sigma) + \frac{\eta_\sigma}{\varepsilon} f'(x,\sigma) + \frac{\gamma_\sigma}{\varepsilon^2}(f(x,-\sigma) - f(x,\sigma)).$$

When taking the limit $\varepsilon \to 0$, this generator will not converge, because of the ε^{-1} and ε^{-2} terms.

However, one can also show weak convergence in terms of finite dimensional distributions of a process by showing convergence of its characteristic function [10]. The characteristic function of a random variable X is defined as

$$\phi_X(q) = \mathbb{E}\left(e^{iqX}\right)$$

and uniquely corresponds to the distribution of the random variable. By showing that the characteristic function of a process $X^\varepsilon(t)$ converges to the characteristic function of $X(t)$, we thus obtain

$$(X^\varepsilon(t_1), X^\varepsilon(t_2), ..., X^\varepsilon(t_k)) \xrightarrow{d} (X(t_1), X(t_2), ..., X(t_k)) \qquad \forall\, t_1, ..., t_k$$

Let us again consider scaled random walk and show that its finite dimensional distributions converge to those of Brownian motion. So far, we considered Brownian motion with diffusion coefficient 1, denoted as W_t. In general, Brownian motion W_{Dt} with diffusion coefficient D has a $\mathcal{N}(0, Dt)$ distribution. The characteristic function of W_{Dt} is given by

$$\mathbb{E}\left(e^{iq\mathcal{N}(0,Dt)}\right) = \int_{-\infty}^{\infty} e^{iqx} \frac{1}{\sqrt{2\pi Dt}} e^{-\frac{x^2}{2Dt}} \mathrm{d}x = e^{-\frac{1}{2}Dtq^2}. \tag{3.5}$$

If we view the symmetric random walk (3.2) as a difference between two independent Poisson processes with parameter $\frac{1}{2}t$, we can can compute the characteristic function directly. Let N_t^+ be the number of jumps to the right and N_t^- the number of jumps to the left in the time interval $[0,t]$. We then have $X_t = N_t^+ - N_t^-$, which gives

$$\begin{aligned}
\mathbb{E}\left(e^{iqX_t}\right) = \mathbb{E}\left(e^{iq(N_t^+ - N_t^-)}\right) &= \mathbb{E}\left(e^{iqN_t^+}\right)\mathbb{E}\left(e^{-iqN_t^+}\right) \qquad (N_t^+ \perp N_t^-) \\
&= e^{\frac{t}{2}(e^{iq}-1)} e^{\frac{t}{2}(e^{-iq}-1)} \\
&= e^{t(\cos(q)-1)}.
\end{aligned}$$

We used that the characteristic function of a Poisson random variable Y with parameter λ is given by $\mathbb{E}e^{iqY} = e^{\lambda(e^{iq}-1)}$ [10]. Applying the same scaling as in theorem 3.2, we obtain

$$\begin{aligned}
\mathbb{E}\left(e^{iq\varepsilon X_{t/\varepsilon^2}}\right) = e^{\varepsilon^{-2}t(\cos(q\varepsilon)-1)} &= e^{-2\varepsilon^{-2}t\sin^2\left(\frac{q\varepsilon}{2}\right)} \\
&\xrightarrow{\varepsilon\to 0} e^{-\frac{1}{2}tq^2},
\end{aligned}$$

the characteristic function of Brownian motion with $D = 1$.

In chapter 5 we will compute the Laplace transform of the characteristic function of a process X_t. To be able to compare this to Brownian motion and find the diffusion coefficient, we determine the Laplace transform of (3.5):

$$\int_0^\infty e^{-\frac{1}{2}Dtq^2} e^{-st} \mathrm{d}t = \frac{1}{s + \frac{q^2}{2}D} \tag{3.6}$$

4

Large deviations

The theory of large deviations deals with the probability of very rare events, where the outcome deviates strongly from the mean of the corresponding random variable. Think for example of flipping heads more than two hundred times after another. This is very unlikely, yet the probability is not zero. The probability that a sum of random variables S_n takes a value xn decays exponentially fast for large n:

$$\mathbb{P}(S_n \approx xn) \sim e^{-nI(x)+o(n)}.$$

We are interested in the leading order of the decay and its dependence on the value x. The function $I(x)$ in the leading order is also called the large deviation entropy and can be derived from the microscopic states of a system. It can then be used to describe the system's macroscopic laws.

To illustrate when deviations are said to be 'large', consider the sum S_n of n independent, identically distributed (i.i.d.) random variables with mean μ and variance $\sigma^2 < \infty$. The weak law of large numbers says that S_n converges to $n\mu$ in probability for large n. The deviation of S_n from its mean $n\mu$ is typically of the order $\sqrt{n}$, because the central limit theorem (CLT) states that $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ converges in distribution to $\mathcal{N}(0,1)$ as n tends to infinity. Now deviations are considered large when they are greater than the deviation which the CLT prescribes. For instance $\{S_n \leq n\mu - bn\}$ is an event with large deviation. The probability of such rare events tends to zero as n goes to infinity. The large deviation principle (LDP) is about the rate at which the probability decays to zero.

The theory presented in this chapter is based on literature by Dembo and Zeituni [7] and Den Hollander [8]. We will start by defining what a rate function is and when a probability measure satisfies the large deviation principle. Next we give a concrete example of the large deviation properties of a sequence of i.i.d. random variables. Finally the theorem by Gärtner and Ellis, which generalizes the LDP for non-i.i.d. random variables, will be stated.

Definition 4.1 *Let χ be a topological space, such that open and closed subsets are well-defined. A function $I : \chi \to [0,\infty]$ is called a good rate function if:*

(a) I is lower semi-continuous[1]

(b) the level sets $\Psi(\alpha) = \{x : I(x) \leq \alpha\}$ of I are compact

Note that the lower semi-continuity implies that I attains its infimum on a closed set.

[1] If χ is a metric space, a function $f : \chi \to [0,\infty]$ is lower semi-continuous if $\liminf_{n\to\infty} f(x_n) \geq f(x)$ for all $x_n, x \in \chi$ s.t. $x_n \to x$. Lower semi-continuity of a function is equivalent to closedness of the level sets.

Definition 4.2 *A family of probability measures $(P_n)_{n\geq 0}$ is said to satisfy the large deviation principle with rate n and rate function $I(x)$ if:*

(a) $I(x)$ *is a good rate function*

(b) $\limsup_{n\to\infty} \frac{1}{n}\log P_n(C) \leq -\inf_{x\in C} I(x) \quad \forall C \subset \chi$ *closed*

(c) $\liminf_{n\to\infty} \frac{1}{n}\log P_n(O) \geq -\inf_{x\in O} I(x) \quad \forall O \subset \chi$ *open*

Let us from now on consider a sequence Z_n of random variables on the probability space $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P})$, where $\mathscr{B}(\mathbb{R})$ denotes the Borel sigma algebra on $\mathbb{R}$. When the family (P_n) is defined as

$$P_n(\cdot) = \mathbb{P}(Z_n \in \cdot), \tag{4.1}$$

the large deviation principle can be summarized as

$$\mathbb{P}(Z_n \approx x) \approx e^{-nI(x)}.$$

Definition 4.2(b) and (c) give an upper and lower bound that make this approximate statement precise. They essentially state that the probability for Z_n to deviate from its mean decays exponentially with rate function $I(x)$. In order to explain what the rate function is, we first define the free energy function:

$$F(\alpha) := \lim_{n\to\infty} \frac{1}{n}\log\left(\mathbb{E}e^{\alpha Z_n}\right). \tag{4.2}$$

The rate function is related to $F(\alpha)$ by the Legendre transform:

$$I(x) = \sup_{\alpha\in\mathbb{R}}(x\alpha - F(\alpha)). \tag{4.3}$$

In situations where $F(\alpha)$ exists and is differentiable, the rate function can be found by determining the maximum of $x\alpha - F(\alpha)$. This is the case in the next example.

Example 4.3 *Suppose we have a sequence of independent, $\mathcal{N}(\mu,\sigma^2)$ distributed random variables X_i and are interested in the deviation of $\sum_{i=1}^n X_i$ from the expectation $n\mu$. Recall that the moment generating function of the X_i is given by $\mathbb{E}e^{\alpha X_i} = e^{\alpha\mu + \frac{\alpha^2\sigma^2}{2}}$. Because of the independence of the RV's, the free energy function can be computed directly:*

$$\begin{aligned} F(\alpha) &= \lim_{n\to\infty} \frac{1}{n}\log\left(\mathbb{E}e^{\alpha\sum_{i=1}^n X_i}\right) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \log\mathbb{E}e^{\alpha X_i} \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \log e^{\alpha\mu+\frac{\alpha^2\sigma^2}{2}} = \alpha\mu + \frac{1}{2}\alpha^2\sigma^2. \end{aligned}$$

Consequently, the rate function is:

$$I(x) = \sup_{\alpha}(x\alpha - \mu\alpha - \frac{1}{2}\alpha^2\sigma^2) = \frac{(x-\mu)^2}{2\sigma^2}.$$

Notice that $I(x)$ is symmetric around μ and that $I(\mu) = 0$.

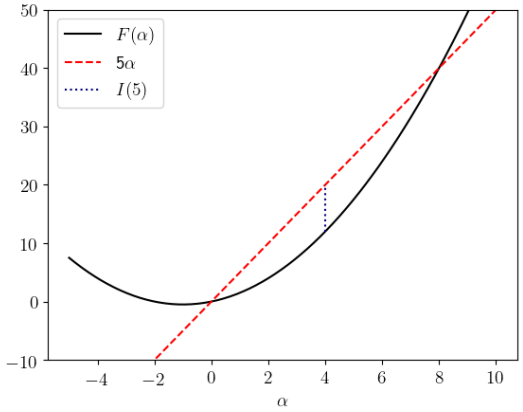


Figure 4.1: The function $F(\alpha)$ of example 4.3 with μ,σ equal to one. The value of the rate function $I(x)$ for $x = 5$ is shown as a dotted line.

Let us show that (4.3) indeed satisfies the properties of a good rate function. We will need that $F(\alpha)$ is convex, which we already obtained for example 4.3 but can show to be true in general.

Lemma 4.4 *The free energy function $F(\alpha)$ is convex.*

Proof. Let $\lambda \in [0,1]$. By Hölder's inequality with $p = \frac{1}{\lambda}$ and $q = \frac{1}{1-\lambda}$, we have

$$F(\lambda\alpha_1 + \alpha_2(1-\lambda)) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left(e^{\lambda\alpha_1 Z_n} e^{(1-\lambda)\alpha_2 Z_n}\right) \leq \lim_{n\to\infty} \frac{1}{n} \log\left[\left(\mathbb{E}e^{\alpha_1 Z_n}\right)^{\lambda} \left(\mathbb{E}e^{\alpha_2 Z_n}\right)^{1-\lambda}\right] = \lambda F(\alpha_1) + (1-\lambda)F(\alpha_2).$$

□

The convexity of the free energy ensures that its Legendre transform is always well-defined. In order to prove that the Legendre transform of $F(\alpha)$ is a good rate function we need to make two assumptions:

(1) $F(\alpha)$ exists

(2) $0 \in \text{int}(D_F)$ for the set $D_F = \{\alpha \in \mathbb{R} : F(\alpha) < \infty\}$

Lemma 4.5 $I(x) = \sup_{\alpha\in\mathbb{R}}(x\alpha - F(\alpha))$ *is a good rate function in the sense of definition 4.1.*

Proof. For part (a) of the definition note that αx as well as $F(\alpha)$ are lower semi-continuous functions, where the latter can be shown using Fatou's lemma. Since $I(x)$ is the supremum of $\alpha x - F(\alpha)$, it is also lower semi-continuous.
To prove part (b) of the definition it remains to show that the level sets are bounded, as the lower semi-continuity already implies they are closed. By assumption (2) it follows that there exists a $\delta > 0$ such that $B_{2\delta}(0) \subset \text{int}(D_F)$. As $F(\alpha)$ is convex, it is continuous on the open subset $\text{int}(D_F)$ of $\mathbb{R}$. Hence there is a $c < \infty$ such that $\sup_{\alpha\in B_\delta(0)} F(\alpha) = c$. This gives

$$I(x) \geq \sup_{\alpha\in B_\delta(0)} \{\alpha x - F(\alpha)\} \geq \delta|x| - c,$$

such that $|x| \leq \frac{I(x)+c}{\delta}$. Hence, the level sets of $I(x)$ are bounded and closed, and therefore compact. □

4.1. The Gärtner-Ellis theorem

We continue with the Gärtner-Ellis theorem that states the large deviation principle for dependent random sequences. In chapter 6 we are going to study the large deviations of the run-and-tumble model and will encounter a free energy function that is differentiable and strictly convex. This is why we present a slightly more simple version of the theorem, limited to strictly convex functions $F(\alpha)$ yet adequate for our purposes. Note that the Gärtner-Ellis theorem also applies to random sequences on a d-dimensional space, even though we work with α and x in $\mathbb{R}$.

Theorem 4.6 (Gärtner – Ellis theorem) *Let P_n, $F(\alpha)$ and $I(x)$ be defined as in (4.1), (4.2) and (4.3) respectively. If $F(\alpha)$ is continuously differentiable and strictly convex on $\mathbb{R}$, then:*

(a) $$\limsup_{n\to\infty} \frac{1}{n} \log P_n(C) \leq -\inf_{x\in C} I(x) \quad \forall C \subset \mathbb{R} \text{ closed}$$

(b) $$\liminf_{n\to\infty} \frac{1}{n} \log P_n(O) \geq -\inf_{x\in O} I(x) \quad \forall O \subset \mathbb{R} \text{ open}$$

For a proof of theorem 4.6, we recommend to read [8] or [7].

In chapter 6 we will apply the Gärtner-Ellis theorem. We then take the sets $C = x$ and $O = B_\varepsilon(x)$. For $\varepsilon \to 0$, the limit superior and limit inferior converge to one limit and the LDP gives:

$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_n \in B_\varepsilon) = -I(x).$$

Theorem 4.6 also applies to continuous-time random sequences. The discrete variable n can be replaced by t if a rate one Poisson process is used to determine when X_n takes place.

4.2. The free energy and moments of a random variable

Since the free energy function is the limit of the cumulant generating function of a random variable X_t, the moments of X_t can be obtained from $F(\alpha)$. By taking the derivative with respect to α and evaluating at $\alpha = 0$, the following scaled limits of the expectation and the variance are obtained:

$$\left.\frac{\mathrm{d}F}{\mathrm{d}\alpha}\right|_{\alpha=0} = \lim_{t\to\infty} \frac{1}{t} \left.\frac{\mathbb{E}\left(X_t e^{\alpha X_t}\right)}{\mathbb{E}\left(e^{\alpha X_t}\right)}\right|_{\alpha=0} = \lim_{t\to\infty} \frac{\mathbb{E}(X_t)}{t}$$

$$\left.\frac{\mathrm{d}^2F}{\mathrm{d}\alpha^2}\right|_{\alpha=0} = \lim_{t\to\infty} \frac{\mathbb{E}(X_t^2) - \mathbb{E}(X_t)^2}{t}$$

We used that $\lim_{\alpha\to 0}$ and $\lim_{t\to\infty}$ may be interchanged by the dominated convergence theorem.

If X_t is for example Brownian motion, then $F'(0)$ is equivalent to the velocity, which is zero, and $F''(0)$ to the diffusion constant of the motion.

5

One-dimensional active particle model

In this chapter we derive the asymptotic behaviour of a version of one-dimensional run-and-tumble motion. The dynamics of various active particles can be described by the run-and-tumble model, in which a particle performs symmetric random walk and additionally moves in a direction that is reoriented at random moments in time ('tumbles'). In different bacterial systems this motion has been observed, for example for myxobacteria [17]. Recent studies on run-and-tumble particles were performed by Demaerel and Maes [6], by Van Ginkel, Van Gisbergen and Redig [20] and by Malakar et al. [13]
The particular motion which we treat is described by the position on a discrete lattice and the internal state of a particle. This internal state may represent a chemomechanical state of a biological agent and is assumed to be independent of its position in this model. The state controls the preferred direction of motion, the rate at which the particle jumps to different positions and the rate at which the internal state changes. This is different from run-and-tumble models that have been studied before, where the internal state influences either the direction of motion, such as in [13], or the rate.

Here we will focus on two possible internal states: $\{\pm 1\}$, where $+1$, -1 correspond to a preference to move in the positive and negative direction respectively. To summarise the motion, a particle at position $x \in \mathbb{Z}$ and with internal state σ can make transitions to

- position $x+1$ or $x-1$, both with rate λ_σ (nearest neighbour random walk),
- position $x+\sigma$ with rate η_σ (transport),
- state $-\sigma$ with rate γ_σ (state change).

The process can be visualised as motion on two copies of $\mathbb{Z}$, each corresponding to an internal state, and rates γ_σ for switching between the lattices.

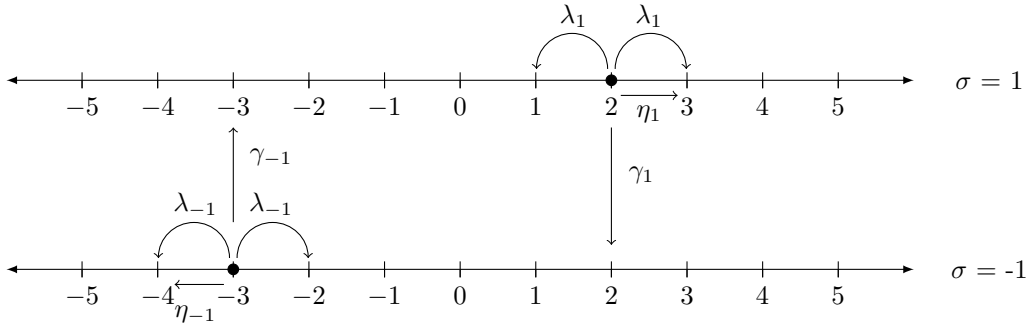


Figure 5.1: Schematic diagram of the possible transitions and corresponding rates of the run-and-tumble model. Particles at $(x,\sigma) = (2,1)$ and $(-3,-1)$ are depicted, but the rates apply to a particle at any position with the same internal state σ.

The generator of this process on the state space $\Omega = \mathbb{Z} \times \{1,-1\}$, as follows from (2.12), is given by:

$$Lf(x,\sigma) = \lambda_\sigma(f(x+1,\sigma) + f(x-1,\sigma) - 2f(x,\sigma)) + \eta_\sigma(f(x+\sigma,\sigma) - f(x,\sigma)) + \gamma_\sigma(f(x,-\sigma) - f(x,\sigma)). \tag{5.1}$$

As defined before, the semigroup applied to f is $S_t f(x,\sigma) = \mathbb{E}_{x,\sigma} f(X_t,\sigma_t) = \mathbb{E}(f(X_t,\sigma_t) \mid x_0 = x, \sigma_0 = \sigma)$. Using the Kolmogorov backward equation (2.9), we can determine how the expectation of a test function $f(x,\sigma)$ evolves:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} S_t f(x,\sigma) = \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}_{x,\sigma} f(X_t,\sigma_t) = L\mathbb{E}_{x,\sigma} f(X_t,\sigma_t) \\
&= \lambda_\sigma \left(\mathbb{E}_{x,\sigma} f(X_t,\sigma_t) + \mathbb{E}_{x,\sigma} f(X_t,\sigma_t) - 2\mathbb{E}_{x,\sigma} f(X_t,\sigma_t)\right) + \eta_\sigma \left(\mathbb{E}_{x+\sigma,\sigma} f(X_t,\sigma_t) - \mathbb{E}_{x,\sigma} f(X_t,\sigma_t)\right) \\
&\quad + \gamma_\sigma \left(\mathbb{E}_{x,-\sigma} f(X_t,\sigma_t) - \mathbb{E}_{x,\sigma} f(X_t,\sigma_t)\right)
\end{aligned} \tag{5.2}$$

Because the generator, and therefore the semigroup, commutes with spatial translation (see appendix A), the following translation invariance property holds:

$$\mathbb{E}_{x+a,\sigma} f(X_t,\sigma_t) = \mathbb{E}_{x,\sigma} f(X_t + a, \sigma_t) \qquad \forall a \in \mathbb{R} \tag{5.3}$$

Since our goal is to derive the limit of the characteristic function, we take $f(x,\sigma) = e^{iqx}$. Applying (5.3) to (5.2) gives

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_{x,\sigma} e^{iqX_t} &= \lambda_\sigma \left(\mathbb{E}_{x,\sigma} e^{iq(X_t+1)} + \mathbb{E}_{x,\sigma} e^{iq(X_t-1)} - 2\mathbb{E}_{x,\sigma} e^{iqX_t}\right) + \eta_\sigma \left(\mathbb{E}_{x,\sigma} e^{iq(X_t+\sigma_t)} - \mathbb{E}_{x,\sigma} e^{iqX_t})\right) \\
&\quad + \gamma_\sigma \left(\mathbb{E}_{x,-\sigma} e^{iqX_t} - \mathbb{E}_{x,\sigma} e^{iqX_t}\right) \\
&= \lambda_\sigma (e^{iq} + e^{-iq} - 2)\mathbb{E}_{x,\sigma} e^{iqX_t} + \eta_\sigma (e^{iq\sigma} - 1)\mathbb{E}_{x,\sigma} e^{iqX_t} \\
&\quad + \gamma_\sigma \left(\mathbb{E}_{x,-\sigma} e^{iqX_t} - \mathbb{E}_{x,\sigma} e^{iqX_t}\right)
\end{aligned} \tag{5.4}$$

Let us first derive the limiting distribution of a simplification, where $\eta_1 = \eta_{-1} = 1$ and $\gamma_1 = \gamma_{-1} = 1$. We will prove the following theorem:

Theorem 5.1 *$\varepsilon X_{t\varepsilon^{-2}}$, the position and time scaled path of run-and-tumble motion given by (5.1) with $\eta_1 = \eta_{-1} = \gamma_1 = \gamma_{-1} = 1$, converges in finite dimensional distribution to $W_{t\mathbb{D}}$, Brownian motion with diffusion constant $\mathbb{D} = \lambda_1 + \lambda_{-1} + 2$, as $\varepsilon \to 0$.*

Proof. If we view $\mathbb{E}_{x,\sigma} e^{iqX_t}$ as a column vector indexed by the two possible internal states, (5.4) is a system of differential equations coupling $\mathbb{E}_{x,1} e^{iqX_t}$ and $\mathbb{E}_{x,-1} e^{iqX_t}$. We can write the system in matrix form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \mathbb{E}_{x,1} e^{iqX_t} \\ \mathbb{E}_{x,-1} e^{iqX_t} \end{pmatrix} = \begin{pmatrix} \lambda_1(2\cos q - 2) + (e^{iq} - 1) - 1 & 1 \\ 1 & \lambda_{-1}(2\cos q - 2) + (e^{-iq} - 1) - 1 \end{pmatrix} \begin{pmatrix} \mathbb{E}_{x,1} e^{iqX_t} \\ \mathbb{E}_{x,-1} e^{iqX_t} \end{pmatrix}, \tag{5.5}$$

where we shall denote the matrix by $M(q)$.

In order to transform this system of differential equations into a system of linear equations, we take the Laplace transform in t. Let $\mathscr{L}\{f(t)\} = F(s) = \int_0^\infty e^{-st} f(t)\mathrm{d}t$. The following property, which can be verified by partial integration, holds for the Laplace transform:

$$\mathscr{L}\left\{\frac{\mathrm{d}}{\mathrm{d}t} f(t)\right\} = sF(s) - f(0).$$

Returning to the evolution of the characteristic function, we define $S(q,s,\sigma) = \mathscr{L}\{\mathbb{E}_{x,\sigma} e^{iqX_t}\}$, such that

$$M(q)\begin{pmatrix} S(q,s,1) \\ S(q,s,-1) \end{pmatrix} = s\begin{pmatrix} S(q,s,1) \\ S(q,s,-1) \end{pmatrix} + \left.\begin{pmatrix} \mathbb{E}_{x,1} e^{iqX_t} \\ \mathbb{E}_{x,-1} e^{iqX_t} \end{pmatrix}\right|_{t=0}. \tag{5.6}$$

Upon assuming the initial position to be $X_0 = 0$ and the internal state to be $\sigma_0 = 1$, we find

$$S(q,s,1) = (1,0)\left(sI - M(q)\right)^{-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{5.7}$$

Note that the initial internal state does not influence the limiting distribution, so the limit can also be found by considering $S(q,s,-1)$.

Since $\big(sI - M(q)\big)$ is a two-by-two matrix, its inverse can be computed directly:

$$\big(sI - M(q)\big)^{-1} = \frac{1}{\det\big(sI - M(q)\big)} \begin{pmatrix} s - 2\lambda_1(\cos q - 1) - (e^{iq} - 1) + 1 & 1 \\ 1 & s - 2\lambda_{-1}(\cos q - 1) - (e^{-iq} - 1) + 1 \end{pmatrix}, \quad (5.8)$$

with

$$\begin{aligned} &\det\big(sI - M(q)\big) = s^2 + 2s + 4\lambda_1\lambda_{-1}(\cos q - 1)^2 + \\ &2(\cos q - 1)\Big[\lambda_1(e^{iq} - 1) + \lambda_{-1}(e^{-iq} - 1) - 1 - (s+1)[\lambda_1 + \lambda_{-1} + 1]\Big]. \end{aligned} \quad (5.9)$$

Next, we will apply the same scaling which we used to determine the limiting distribution of continuous-time random walk in theorem 3.2. The scaling of x by a factor ε and t by a factor ε^{-2} corresponds to the scaling of $S(q,s,\sigma)$ to $\varepsilon^2 S(\varepsilon q, \varepsilon^2 s, \sigma)$. Namely,

$$\int_0^\infty e^{-st}\mathbb{E}e^{iq\varepsilon X_{t\varepsilon^{-2}}}\,\mathrm{d}t = \varepsilon^2\int_0^\infty e^{-s\varepsilon^2\tau}\mathbb{E}e^{iq\varepsilon X_\tau}\,\mathrm{d}\tau = \varepsilon^2 S(\varepsilon q, \varepsilon^2 s).$$

From (5.7-5.9), it follows that

$$\begin{aligned} &\varepsilon^2 S(\varepsilon q, \varepsilon^2 s, 1) = \\ &\frac{\varepsilon^2\big[\varepsilon^2 s + 2 - 2\lambda_1(\cos\varepsilon q - 1) - (e^{i\varepsilon q} - 1)\big]}{\varepsilon^4 s^2 + 2\varepsilon^2 s + 4\lambda_1\lambda_{-1}(\cos\varepsilon q - 1)^2 + 2(\cos\varepsilon q - 1)\big[\lambda_1(e^{i\varepsilon q} - 1) + \lambda_{-1}(e^{-i\varepsilon q} - 1) - (\varepsilon^2 s + 1)[2\lambda_1 + \lambda_{-1} + 1]\big]} \\ &\xrightarrow{\varepsilon\to 0} \frac{1}{s + \frac{q^2}{2}(\lambda_1 + \lambda_{-1} + 2)}. \end{aligned} \quad (5.10)$$

As (3.6) indicates, this is the Laplace transform of $\mathbb{E}e^{iqW_{\mathbb{D}t}}$ with $\mathbb{D} = \lambda_1 + \lambda_{-1} + 2$. □

The first contribution, $(\lambda_1 + \lambda_{-1})$, is due to the random walk part of the motion, while the second contribution is due to the active transport at rate one in either direction and the switching between the states at rate one. With a different approach, Van Gisbergen derived that the limiting diffusion constant of a run-and-tumble process in which λ is independent of the internal state contains a term 2λ [21]. A similar term appears in (5.10), with an average of the random walk rates for the different internal states instead of λ.

5.1. State-dependent transport

The derivation can now be extended to run-and-tumble motion where not only the random walk rate λ_σ depends on the internal state, but the transport rate as η_σ well. Let us first determine what the mean displacement would be for this process. To accomplish that we consider a simplified process that only involves asymmetric transport. The jumps to the right and to the left are both Poisson distributed with parameters $t\eta^+$ and $t\eta^-$ respectively. The total position is

$$X_t = N_t^+ - N_t^-,$$

such that the mean equals $\mathbb{E}(X_t) = \mathbb{E}(N_t^+) - \mathbb{E}(N_t^-) = (\eta^+ - \eta^-)t$.

Applying this reasoning to the run-and-tumble model which involves transport to the right for $\sigma = 1$ and transport to the left for $\sigma = -1$, we obtain a mean displacement of $\mathbb{E}(X_t) = (\mu_1\eta_1 - \mu_{-1}\eta_{-1})t$. In this expression $[\mu_1, \mu_{-1}]$ is the stationary distribution of the internal state σ_t (a continuous-time Markov process with finite state space $\{1,-1\}$). This distribution is attained quickly relative to the limiting distribution of X_t. As described in 2.6, the stationary distribution can be found by solving (2.18). When the rates γ_1 and γ_{-1} to switch to the other internal state are equal, this simply results in $[\mu_1, \mu_{-1}] = \big[\frac{1}{2}, \frac{1}{2}\big]$.

We will now return to the original process, comprising symmetric random walk, active transport and switching of the internal state. Since the random walk motion has mean zero, the average displacement of this process with state-dependent rates is $\frac{1}{2}(\eta_1 - \eta_{-1})t$. As opposed to the derivation for the previous model, in which the transport was symmetric, we will now determine the limiting distribution of $X_t - vt$, where $v := \frac{1}{2}(\eta_1 - \eta_{-1})$.

In contrast to 5.5, the diagonal elements of $M(q)$ are now

$$2\lambda_\sigma(\cos(q) - 1) - 1 + \eta_\sigma\left(e^{\sigma i q} - 1\right).$$

The matrix $\left(sI - M(q)\right)^{-1}$ undergoes the same alteration, such that its determinant becomes:

$$\begin{aligned}&\det\left(sI - M(q)\right) = s^2 + 2s + 4\lambda_1\lambda_{-1}(\cos q - 1)^2 + \\ &2(\cos q - 1)\left[-\eta_1\eta_{-1} + \sum_{\sigma\in\{1,-1\}} \lambda_\sigma\eta_\sigma(e^{iq\sigma} - 1) - (s+1)[\lambda_\sigma + \eta_\sigma(e^{iq\sigma} - 1)]\right].\end{aligned} \tag{5.11}$$

The scaling which we apply to this process is $X_t - vt \to \varepsilon\left(X_{\varepsilon^{-2}t} - vt\varepsilon^{-2}\right)$ and the corresponding Laplace transform is:

$$\int_0^\infty e^{-st}\mathbb{E}e^{iq\varepsilon\left(X_{\varepsilon^{-2}t} - vt\varepsilon^{-2}\right)}dt = \varepsilon^2\int_0^\infty \mathbb{E}e^{iq\varepsilon X_\tau}e^{-iq\varepsilon\tau v}e^{-s\tau\varepsilon^2}d\tau = \varepsilon^2 S(\varepsilon q, \varepsilon^2 s + iq\varepsilon v).$$

Using (5.7) and (5.11), it can be derived that in this case:

$$\lim_{\varepsilon\to 0}\varepsilon^2 S(\varepsilon q, \varepsilon^2 s + iq\varepsilon v) = \frac{1}{s + \frac{q^2}{2}\left[\lambda_1 + \lambda_{-1} + \frac{\eta_1+\eta_{-1}}{2} + \left(\frac{\eta_1+\eta_{-1}}{2}\right)^2\right]} \tag{5.12}$$

The contribution to the diffusive coefficient due to transport contains the transport rates weighted by the stationary distribution of the internal state, similar to the random walk contribution.

If we do not require γ_1 and γ_{-1} to be equal, the averaging over the stationary distribution can be observed even more clearly. As follows from (2.18), the stationary distribution is given by $[\mu_1, \mu_{-1}] = \left[\frac{\gamma_{-1}}{\gamma_1+\gamma_{-1}}, \frac{\gamma_1}{\gamma_1+\gamma_{-1}}\right]$. As a sanity check we can take the limit of γ_1 to infinity and find a distribution of $[0, 1]$. This indeed corresponds to immediate switching from $\sigma = 1$ to -1.

Applying the same method as before, we find that the limiting diffusion coefficient of $X_t - (\mu_1\eta_1 - \mu_{-1}\eta_{-1})t$ is equal to

$$D = \frac{2\lambda_1\gamma_{-1} + 2\lambda_{-1}\gamma_1}{\gamma_1+\gamma_{-1}} + \frac{\eta_1\gamma_{-1} + \eta_{-1}\gamma_1}{\gamma_1+\gamma_{-1}} + \frac{2\gamma_1\gamma_{-1}(\eta_1+\eta_{-1})^2 + 2\eta_1\eta_{-1}(\gamma_1-\gamma_{-1})^2}{(\gamma_1+\gamma_{-1})^3}. \tag{5.13}$$

The first term of this expression is equivalent to the first term of (5.10), which can be rewritten as $2(\frac{1}{2}\lambda_1 + \frac{1}{2}\lambda_{-1})$. Instead of the factors $\frac{1}{2}$, the more general stationary distribution $[\mu_1, \mu_{-1}]$ in terms of the rates γ_1 and γ_{-1} appears. The second term of (5.13) contains the transport rates, also weighed by the stationary distribution of the internal state.

Upon setting $\gamma_1 = a\Gamma_1$, $\gamma_{-1} = a\Gamma_{-1}$ and taking the limit $a \to \infty$, we obtain

$$D = 2\left(\frac{\lambda_1\Gamma_{-1}}{\Gamma_1+\Gamma_{-1}} + \frac{\lambda_{-1}\Gamma_1}{\Gamma_1+\Gamma_{-1}}\right) + \frac{\eta_1\Gamma_{-1}}{\Gamma_1+\Gamma_{-1}} + \frac{\eta_{-1}\Gamma_1}{\Gamma_1+\Gamma_{-1}}.$$

The last term of (5.13) goes to zero. This limit is a version of the so called slow-fast limit. The name originates from the fact that the generator of a process with random walk and drift (without internal states), after scaling time and position, is given by

$$L_\varepsilon f(x) = \lambda f''(x) + \varepsilon^{-1}\eta f'(x)$$

Here we see that the random walk part converges slowly while the dissipative part converges quickly (note the high drift rate because of the factor ε^{-1}). The third term of (5.13) is an addition to this slow-fast limit due to the activity of the particle.

5.2. General random walk

In the previous models, the particle was assumed to make random walk jumps to nearest neighbour sites only. Let us generalize this to a process where a random walk jump to position $z \in \mathbb{Z}$ is made with rate $p(z,\sigma)$:

$$Lf(x,\sigma) = \sum_z p(z,\sigma)\big(f(x+z,\sigma) - f(x,\sigma)\big) + \eta_\sigma\big(f(x+\sigma,\sigma) - f(x,\sigma)\big) + \big(f(x,-\sigma) - f(x,\sigma)\big). \tag{5.14}$$

We will consider two different situations, where first $p(z,\sigma)$ is a product of $\sigma-$ and z-dependent rates and secondly $p(z,\sigma)$ can not be separated:

(a) $p(z,\sigma) = p(z)\lambda_\sigma$ with $p(z) = p(-z)$ and $\sum_z z^2 p(z) < \infty$.

(b) $p(z,\sigma)$ with $\sum_z z p(z,\sigma) = 0$, such that the mean of the random walk part is still zero.

In both cases we will again determine the asymptotic behaviour of $X_t - \frac{1}{2}(\eta_1 - \eta_{-1})t$, meaning that the analysis of the previous section still largely applies. The off-diagonal elements of the matrix $M(q)$ remain the same, but the diagonal elements are now

(a) $\lambda_\sigma \sum_z p(z)(\cos(qz) - 1) - 1 + \eta_\sigma\left(e^{\sigma i q} - 1\right)$

(b) $\sum_z p(z,\sigma)(e^{iqz} - 1) - 1 + \eta_\sigma\left(e^{\sigma i q} - 1\right)$,

where in case (a) the symmetry of the transition rate $p(z)$ was used.

Similarly to how (5.12) was derived, we obtain the following limiting scalar behavior for case (a):

$$\lim_{\varepsilon\to 0} \varepsilon^2 S(\varepsilon q, \varepsilon^2 s + i q \varepsilon v) = \frac{1}{s + \frac{q^2}{2}\left[\frac{1}{2}(\lambda_1 + \lambda_{-1})\sum_z z^2 p(z) + \frac{\eta_1+\eta_{-1}}{2} + \left(\frac{\eta_1+\eta_{-1}}{2}\right)^2\right]}. \tag{5.15}$$

The factor $\frac{1}{2}$ in front of the sum of λ's comes from the fact that one jump to position $X_t + z$ is considered, compared to a jump to $X_t + 1$ and one to $X_t - 1$ in (5.1).

For case (b) we obtain:

$$\lim_{\varepsilon\to 0} \varepsilon^2 S(\varepsilon q, \varepsilon^2 s + i q \varepsilon v) = \frac{1}{s + \frac{q^2}{2}\left[\sum_z z^2 (p(z,1) + p(z,-1)) + \frac{\eta_1+\eta_{-1}}{2} + \left(\frac{\eta_1+\eta_{-1}}{2}\right)^2\right]}. \tag{5.16}$$

Alike (5.10) and (5.12), the contribution to the diffusion constant due to the random walk part contains an average of the rates over the internal states: $\frac{1}{2}(\lambda_1 + \lambda_{-1})\sum_z z^2 p(z)$ in case (a) and $\sum_z z^2 (p(z,1) + p(z,-1))$ in case (b). Moreover, we observe that $\left(\frac{\eta_1+\eta_{-1}}{2}\right)^2 + \frac{\eta_1+\eta_{-1}}{2}$ is an addition solely due to the activity of the particles. It is unaffected by the form of the symmetric random motion.

6

Large deviations of the run-and-tumble model

From the characteristic function of run-and-tumble motion which we derived in chapter 5 it is a small step towards studying the large deviation properties of the position X_t. We will derive the large deviation free energy function $F(\alpha)$. As described in chapter 4, this function allows to determine the rate function $I(x)$ and offers a different method to find the limiting diffusion coefficient and asymptotic velocity of the motion. The next theorem states the large deviation results of the run-and-tumble model considered in chapter 5.

Theorem 6.1 *The position corresponding to run-and-tumble motion with generator (5.1) and $\gamma_1 = \gamma_{-1} = \gamma$ satisfies the large deviation principle with free energy function:*

$$F(\alpha) = (\cosh(\alpha) - 1)(\lambda_1 + \lambda_{-1}) + \frac{\eta_1}{2}(e^{\alpha} - 1) + \frac{\eta_{-1}}{2}(e^{-\alpha} - 1) - \gamma + A, \tag{6.1}$$

where

$$A = \sqrt{\left[(\cosh(\alpha) - 1)(\lambda_1 - \lambda_{-1}) + \frac{\eta_1}{2}(e^{\alpha} - 1) - \frac{\eta_{-1}}{2}(e^{-\alpha} - 1)\right]^2 + \gamma^2}. \tag{6.2}$$

The rate function $I(x)$ is equal to the Legendre transform of the free energy:

$$I(x) = \sup_{\alpha \in \mathbb{R}} (\alpha x - F(\alpha)).$$

Proof. In order to compute the large deviation free energy

$$F(\alpha) = \lim_{t \to \infty} \frac{1}{t} \log\left(\mathbb{E}_{x,\sigma} e^{\alpha X_t}\right),$$

we first derive $\mathbb{E}_{x,\sigma} e^{\alpha X_t}$ from equation (5.5). This can be achieved by exponentiating the matrix $M(-i\alpha)$, where the argument q has been replaced by $-i\alpha$.

$$M(-i\alpha) = \begin{pmatrix} \lambda_1(2\cosh\alpha - 2) + \eta_1(e^{\alpha} - 1) - \gamma & \gamma \\ \gamma & \lambda_{-1}(2\cosh\alpha - 2) + \eta_{-1}(e^{-\alpha} - 1) - \gamma \end{pmatrix} = \begin{pmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{pmatrix}$$

To determine $e^{tM(-i\alpha)}$, we diagonalise the matrix $M(-i\alpha) = PDP^{-1}$, with D a diagonal matrix, such that

$$e^{tM(-i\alpha)} = \sum_{n=0}^{\infty} \frac{t^n}{n!} (PDP^{-1})^n = Pe^{tD}P^{-1}.$$

The matrices P and D are given by:

$$P = \begin{pmatrix} \frac{M_{1,1} - M_{2,2} - 2A}{2\gamma} & \frac{M_{1,1} - M_{2,2} + 2A}{2\gamma} \\ 1 & 1 \end{pmatrix}$$

$$D = \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(M_{1,1} + M_{2,2}) - A & 0 \\ 0 & \frac{1}{2}(M_{1,1} + M_{2,2}) + A \end{pmatrix}$$

We again assume $X_0 = 0$ and find the following expressions for $\mathbb{E}_{0,1}e^{\alpha X_t}$ and $\mathbb{E}_{0,-1}e^{\alpha X_t}$:

$$\begin{aligned}\mathbb{E}_{0,1}e^{\alpha X_t} &= (1,0)e^{tM(-i\alpha)}\begin{pmatrix}1\\1\end{pmatrix}\\ &= e^{\phi_1 t}\left(\frac{2A^2-\gamma-A(M_{1,1}-M_{2,2})}{2A}\right)+e^{\phi_2 t}\left(\frac{2A^2+\gamma+A(M_{1,1}-M_{2,2})}{2A}\right)\end{aligned} \tag{6.3}$$

$$\begin{aligned}\mathbb{E}_{0,-1}e^{\alpha X_t} &= (0,1)e^{tM(-i\alpha)}\begin{pmatrix}1\\1\end{pmatrix}\\ &= \frac{1}{4A}\left(e^{\phi_1 t}(2A+M_{1,1}-M_{2,2}-2\gamma)+e^{\phi_2 t}(2A-(M_{1,1}-M_{2,2})+2\gamma)\right)\end{aligned} \tag{6.4}$$

Finally, we need to take the logarithm, divide by t and apply the limit $t\to\infty$.
Note that the following inequalities hold:

$$2A = 2\sqrt{\frac{1}{4}(M_{1,1}-M_{2,2})^2+\gamma^2} > (M_{1,1}-M_{2,2}) \qquad \text{and} \qquad 2A > 2\gamma.$$

As a result, the terms in front of $e^{\phi_1 t}$ and $e^{\phi_2 t}$ in (6.4) are strictly positive. Hence, we can use that for $a,b,c,d\in\mathbb{R}$, $a,c>0$

$$\lim_{t\to\infty}\frac{1}{t}\log(ae^{bt}+ce^{dt}) = \max\{b,d\},$$

a lemma which we prove in appendix A. We then obtain $\lim_{t\to\infty}\frac{1}{t}\log\left(\mathbb{E}_{0,-1}e^{\alpha X_t}\right)=\phi_2$, the largest eigenvalue of $M(-i\alpha)$.

To determine the same limit for (6.3) we cannot use lemma A.1 directly, because the term before $e^{\phi_1 t}$ is only positive for small γ, $\gamma \le \frac{M_{1,1}-M_{2,2}}{2\sqrt{\frac{1}{4}(M_{1,1}-M_{2,2})^2+\gamma^2}}$ to be precise. However, $\phi_1<0$, $\phi_2>0$ and the term in front of $e^{\phi_2 t}$ is strictly positive, such that the first term of (6.3) goes to zero and the second term goes to infinity as t goes to infinity. In this way, the limit of $\frac{1}{t}$ multiplied by the logarithm of (6.3) also returns the largest of the two eigenvalues.

For both initial internal states we therefore find

$$F(\alpha)=\phi_2=(\cosh(\alpha)-1)(\lambda_1+\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)+\frac{\eta_{-1}}{2}(e^{-\alpha}-1)-\gamma+A,$$

with A defined by (6.2). □

To gain more insight in the found free energy function, let us look at the limit $\gamma\to\infty$. This corresponds to the particles switching their internal state at a very high rate, such that the stationary distribution will quickly be attained. We find

$$\begin{aligned}\lim_{\gamma\to\infty}F(\alpha) &= (\cosh(\alpha)-1)(\lambda_1+\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)+\frac{\eta_{-1}}{2}(e^{-\alpha}-1)-\gamma\\ &\quad+\gamma\sqrt{1+\frac{(\cosh(\alpha)-1)(\lambda_1-\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)-\frac{\eta_{-1}}{2}(e^{-\alpha}-1)}{\gamma^2}}\\ &\approx (\cosh(\alpha)-1)(\lambda_1+\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)+\frac{\eta_{-1}}{2}(e^{-\alpha}-1)-\gamma+\gamma\\ &\quad+\frac{(\cosh(\alpha)-1)(\lambda_1-\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)-\frac{\eta_{-1}}{2}(e^{-\alpha}-1)}{2\gamma}+o\left(\frac{1}{\gamma}\right)\\ &= (\cosh(\alpha)-1)(\lambda_1+\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)+\frac{\eta_{-1}}{2}(e^{-\alpha}-1).\end{aligned}$$

This is indeed the free energy function of diffusion with rate $\lambda_1 + \lambda_{-1}$ and transport in the positive direction with rate $\frac{\eta_1}{2}$ and in the negative direction with rate $\frac{\eta_{-1}}{2}$.

If we consider $\gamma = 0$, on the other hand, the free energy becomes

$$F(\alpha) = (\cosh(\alpha)-1)(\lambda_1+\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)+\frac{\eta_{-1}}{2}(e^{-\alpha}-1)-\left|(\cosh(\alpha)-1)(\lambda_1-\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)-\frac{\eta_{-1}}{2}(e^{-\alpha}-1)\right|.$$

We see that in this case the free energy and therefore the rate function only depends on one of the two internal states at the same time. The free energy is either $(\cosh(\alpha)-1)\lambda_1+\eta_1(e^{\alpha}-1)$ or $(\cosh(\alpha)-1)\lambda_{-1}+\eta_{-1}(e^{-\alpha}-1)$, depending on the sign of the term in absolute value. Since setting $\gamma = 0$ results in a constant internal state, it is in line with our expectations that only the state with the highest rate influences the free energy.

Furthermore, under the condition that the transport rates η_σ are zero and the random walk rates λ_σ are equal, the free energy function reduces to $2\lambda(\cosh(\alpha)-1)$, the free energy of symmetric random walk with rate λ.

Finally, remark that $F(\alpha) \neq F(-\alpha)$ when the transport rates η_1 and η_{-1} are unequal. This means that the rate function $I(x)$ is asymmetric, as one would expect when there is net transport.

Rate function The rate function $I(x)$ can be found by maximizing αx–$F(\alpha)$. Since F is continuously differentiable, we can set $\hat{\alpha} = F'^{-1}(x)$ and find the rate function as: $I(x) = \hat{\alpha}x - F(\hat{\alpha})$.

The derivative of $F(\alpha)$ is equal to

$$\frac{\mathrm{d}F}{\mathrm{d}\alpha} = \sinh(\alpha)(\lambda_1+\lambda_{-1})+\frac{1}{2}\left(\eta_1 e^{\alpha}+\eta_1 e^{-\alpha}\right)+ \frac{\left[(\cosh(\alpha)-1)(\lambda_1-\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)-\frac{\eta_{-1}}{2}(e^{-\alpha}-1)\right]\left[\sinh(\alpha)(\lambda_1+\lambda_{-1})+\frac{1}{2}\left(\eta_1 e^{\alpha}+\eta_1 e^{-\alpha}\right)\right]}{\sqrt{\left[(\cosh(\alpha)-1)(\lambda_1-\lambda_{-1})+\frac{\eta_1}{2}(e^{\alpha}-1)-\frac{\eta_{-1}}{2}(e^{-\alpha}-1)\right]^2+\gamma^2}}.$$

An analytic expression of the inverse of this function can not be obtained, yet we can determine its values numerically. The following figure shows the rate function $I(x)$ for different parameters η_1 and γ.

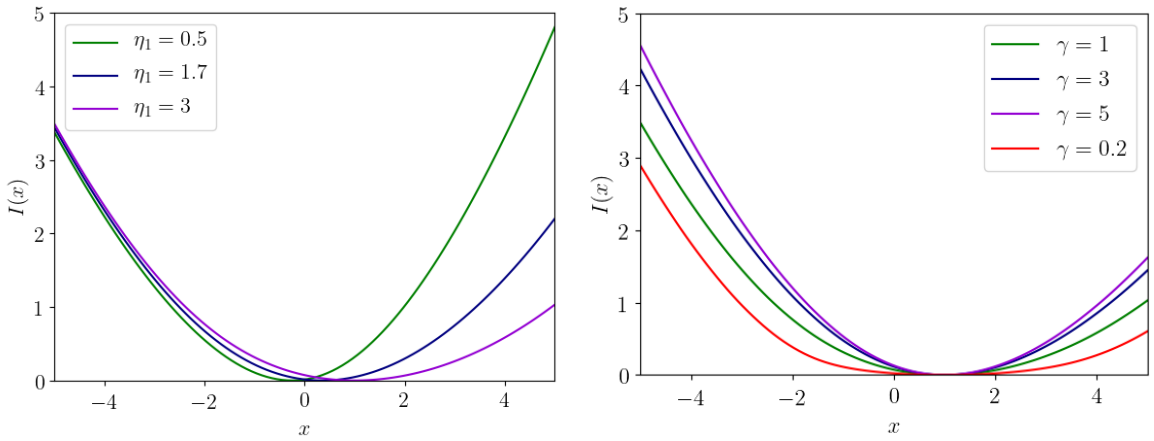


Figure 6.1: The rate function $I(x)$ from theorem 6.1. For both subfigures, the rates $\lambda_1 = \lambda_{-1} = 0.6$ and $\eta_{-1} = 1$ apply. We took $\gamma = 3$ and varied η_1 in the left plot, while we took $\eta_1 = 3$ and varied γ in the right plot.

Asymptotic velocity and diffusion constant Let us determine $F'(0)$ and $F''(0)$ and compare the results to the asymptotic velocity and diffusion constant found previously by scaling of the characteristic function. Like in section 5.1, we take $\gamma = 1$ and obtain

$$\left.\frac{\mathrm{d}F}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}(\eta_1 - \eta_{-1}) \tag{6.5}$$

and

$$\left.\frac{\mathrm{d}^2F}{\mathrm{d}\alpha^2}\right|_{\alpha=0} = \lambda_1 + \lambda_{-1} + \frac{\eta_1 + \eta_{-1}}{2} + \left(\frac{\eta_1 + \eta_{-1}}{2}\right)^2. \tag{6.6}$$

The first derivative of F at zero is indeed in accordance with the asymptotic velocity that we derived by means of the stationary distribution of the internal state. Moreover, (6.6) is equal to the limiting diffusion constant of the run-and-tumble motion (see (5.12)).
Because $F(\alpha)$ can be analytically extended in neighbourhood of the origin in the complex plane, it follows from [4] that the central limit theorem applies to the process, with mean and variance given by (6.6) and (6.5).

7

Collective dynamics under exclusion interactions

So far, we investigated the properties of an individual particle on a lattice exhibiting active motion. We derived the limiting diffusion coefficient and the asymptotic velocity of a run-and-tumble particle that moves without spatial restrictions. In reality, active particles on a linear filament, such as motor proteins, can often not move freely. At high motor densities they are for example hindered by fellow motors which they cannot overtake. When multiple particles are not able to occupy the same position, the process is subject to an exclusion principle. This exclusion affects the motility of the particles and can give rise to jams. In this chapter we study the effect of exclusion on the motility of run-and-tumble particles by means of numerical simulations. We will determine how the average diffusion coefficient and average velocity depend on the particle density ρ and the maximum allowed number of particles per site k. Let us start by introducing an existing and well-studied exclusion process.

7.1. Exclusion processes

Equipped with an exclusion principle, the run-and-tumble model becomes closely related to a paradigmatic model in non-equilibrium statistical physics: the asymmetric simple exclusion process (ASEP). In this one-dimensional interacting particle model, particles enter the first site of a finite lattice at rate k_{en}. From there, they can hop forward or backward with rates k_f and $k_b < k_f$ respectively. Upon arriving at the final site of the lattice, a particle exits at rate k_{ex}. Throughout the process, only one particle can occupy a site at a given time. This model is studied extensively, since it can be used to describe traffic flow as well as biological processes such as the translation of mRNA by ribosomes. Moreover, this model is exactly solvable: Derrida constructed a formalism that allows the current and density profiles to be determined exactly [9]. The density profiles depend on the boundary conditions and can be categorized into different phases, where the flux is either maximal, limited by the entrance rate or limited by the exit rate of particles.

When only forward motion is considered ($k_b = 0$), one obtains a Totally Asymmetric Simple Exclusion Process (TASEP). For the TASEP model with particle density ρ, the probability for a site to be empty is $1-\rho$. The velocity of the particles is then given by the forward hopping rate multiplied by the probability that the next site is empty: $v = k_f(1-\rho)$.

A main difference between the ASEP model and run-and-tumble motion under influence of exclusion is that run-and-tumble motion is governed by an independent state process which determines if a particle is in a forward or backward mode of motion. Also, run-and-tumble motion incorporates symmetric random walk (diffusion) whereas the ASEP and TASEP models describe asymmetric motion. Finally, we will focus on motion in a closed system instead of a system with open boundaries in this chapter.

7.2. Method

With a rejection-based Monte Carlo algorithm, the paths of N particles that move according to equation (5.1), figure 5.1 and an exclusion limit of k particles per position, will be numerically computed. The paths are assumed to obey periodic boundary conditions on a linear segment of size L, such that the process actually takes place on a ring. The process will be sampled by the highest total rate at which transitions take place: $R = N \cdot \max_\sigma\{2\lambda(\sigma) + \eta(\sigma) + \gamma(\sigma)\}$. Recall that $\lambda(\sigma)$ is the random walk rate, $\eta(\sigma)$ the rate of active transport and $\gamma(\sigma)$ the rate at which the internal state σ changes. The Monte-Carlo algorithm for simulating the paths of N particles, made with help of Dr. S. Grosskinsky, is described in pseudocode below:

Algorithm 1 Calculate $\mathbf{x}(t) = \{x_i(t); i \in 1, ..., N\}$

set $t = 0$
define k, the maximum number of particles allowed per site
let $R = N \cdot \max_\sigma\{2\lambda(\sigma) + \mu(\sigma) + \gamma(\sigma)\}$
generate an equidistant distribution for $\mathbf{x}(0)$ over the lattice of length L
draw $\sigma(0) \in \{1, -1\}^N$ from a uniform distribution.
while $t < t_{max}$ **do**
  $t \leftarrow t + \text{Exp}(R)/N$
  pick a particle index i from $\{1, ..., N\}$
  compute the total transition rate $r(\sigma_i)$ of particle i in state σ_i: $r(\sigma_i) = 2\lambda(\sigma_i) + \mu(\sigma_i) + \gamma(\sigma_i)$,
  if Unif(0,1)$\leq \frac{r(\sigma_i)}{R}$ **then**
    determine the move of particle i by a draw from a multinomial distribution with probabilities $\left\{\frac{\lambda(\sigma_i)}{r(\sigma_i)}, \frac{\lambda(\sigma_i)}{r(\sigma_i)}, \frac{\eta(\sigma_i)}{r(\sigma_i)}, \frac{\gamma(\sigma_i)}{r(\sigma_i)}\right\}$ for respectively: one step to the left, to the right, to the direction indicated by $\sigma_i(t)$ or for a flip of the internal state
    if the move is a state change **then**
      $\sigma_i(t) = -\sigma_i(t-1)$
    end if
    if the move is a step of length 1 to site l **then**
      $x_i(t) = l$ if the occupation of site l is less than k, otherwise no move is made
    end if
  end if
end while

From the simulated paths $\mathbf{x}(t)$ we will determine the average diffusion coefficient and the velocity. Since one-dimensional Brownian motion with diffusion coefficient D has a $\mathcal{N}(0, Dt)$ distribution, D will be determined by dividing the sample variance of the position by the elapsed time. Let $r_i = x_i(t) - x_i(0)$ denote the displacement of particle i after time t and $\overline{r}$ denote the sample mean of the r_i. We determine the average diffusion coefficient and average velocity of a set of N particles according to

$$D_{av} = \frac{S}{t} = \frac{1}{t}\frac{\sum_{i=1}^N (r_i - \overline{r})^2}{N-1} \qquad \text{and} \qquad v_{av} = \frac{1}{Nt}\sum_{i=1}^N r_i, \tag{7.1}$$

where S denotes the sample variance of the displacements.

7.3. Results

Unless stated otherwise, the following rates for the run-and-tumble motion (in s^{-1}) were used to obtain the results:

$$\lambda_1 = 0.6 \quad \eta_1 = 1.7 \quad \gamma_1 = 0.7$$
$$\lambda_{-1} = 0.6 \quad \eta_{-1} = 1 \quad \gamma_{-1} = 1.2$$

such that there is net transport in the positive x-direction. The velocity of the transport and the diffusion constant can be calculated from (2.18) and (5.13). Rounded to two decimal places, they are given by $v = 4.55$ sites/s and $D = 0.71$ sites^2/s. Next, let us look at the influence of an exclusion constraint on these quantities.

7.3.1. Diffusion coefficient

The diffusion coefficient is only well-defined if the variance of the displacement scales linearly with time. If this relation is nonlinear, the motion is called anomalous diffusion, for which we do not speak of a diffusion coefficient. We thus begin by examining the variance as a function of time.

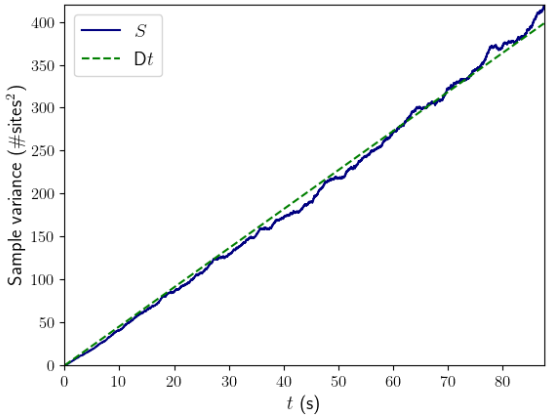
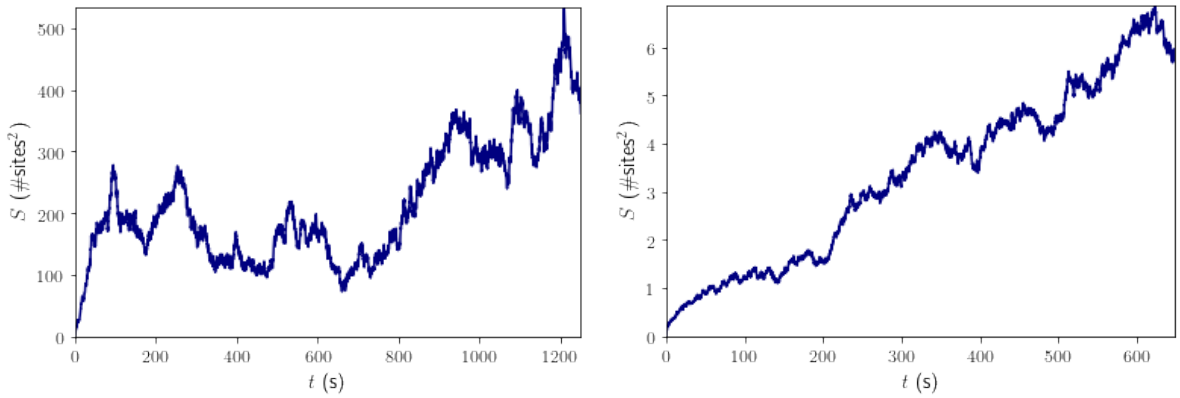


Figure 7.1: The sample variance of the displacement against time. For these data 400 particles were placed on a segment of 1000 sites with exclusion constraint $k = 4$. The dashed line shows the theoretical variance: Dt, where $D = 0.71$ sites^2/s is the analytically derived diffusion constant.

Figure 7.1 shows the sample variance in a situation where the exclusion constraint has a small influence on the motion: 350 particles were simulated on a lattice of 1000 sites under the condition that no more than 4 particles could occupy a site. Next to the sample variance, the theoretical variance of the motion is plotted.



(a) $N = 40$ (b) $N = 540$

Figure 7.2: The sample variance S of the displacement of 40 (left) and 540 particles (right) against time. A lattice of 600 sites and exclusion constraint of $k = 1$ were used.

On the other hand, figure 7.2 shows the sample variance of motion that is highly influenced by the exclusion constraint. The plots were obtained using the parameters $L = 600$ and $k = 1$ and correspond to a particle number of (a): $N = 40$ and (b): $N = 540$.

Presumably the fluctuations and non-linearity of the sample variance are caused by the fact that particles cannot overtake each other. Let us verify this assumption by changing the nearest neighbour random walk to next-nearest neighbour random walk and leaving all other parameters and characteristics of the model unaltered. In other words, the steps that are taken at rate $\lambda(\sigma)$ have length 2 instead of length 1. Figure (7.3) shows the sample variance for this motion:

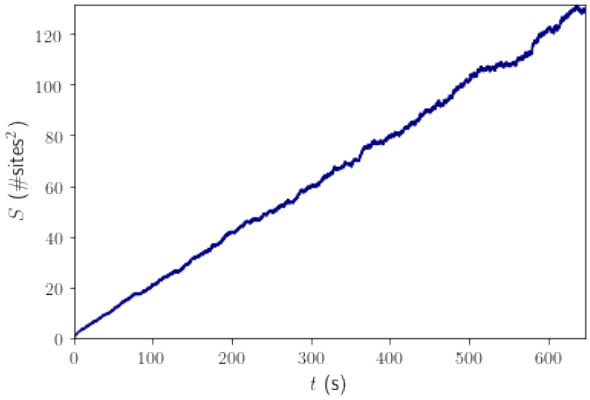


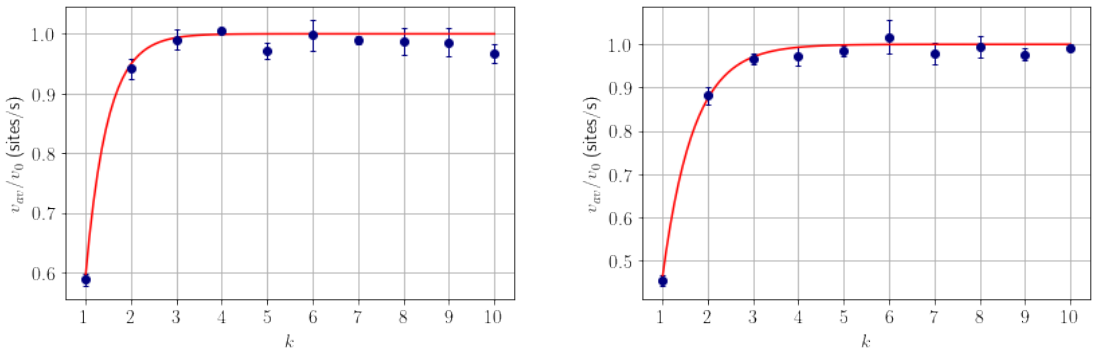
Figure 7.3: The sample variance S of the displacement of 540 particles against time. A lattice of 600 sites and exclusion constraint of $k = 1$ were used. For this simulation the particles did not perform nearest neighbour random walk but next-nearest neighbour random walk.

As figure 7.3 shows, the sample variance behaves linearly as a function of time, in contrast to the sample variances shown in figure 7.2. The gradient is approximately 0.2, lower than the theoretical gradient of 0.71 sites^2/s.

We conclude that run-and-tumble particles with an exclusion limit of $k = 1$ do not exhibit normal diffusion, when simulated on a ring. In the same situation but with next-nearest neighbour random walk, it is possible to determine the average diffusion coefficient, described by (7.1). This diffusion coefficient is found to be lower than the theoretical value.

7.3.2. Velocity

Unlike the diffusion coefficient, the average velocity of the particles is always well-defined. We can therefore determine the behaviour of the velocity as a function of the allowed number of particles per site k, ranging from $k = 1$ (hard-core exclusion) to $k = 10$. Below, the average velocity scaled by the expected velocity is shown for 350 and 500 particles.



(a) $N = 350$ (b) $N = 500$

Figure 7.4: The scaled average velocity $\frac{v_{av}}{v_0}$ against the maximum number of particles per site k for particle numbers $N = 350$ and $N = 500$ and lattice size $L = 1000$. Each data point is an average of three simulations, of which the standard deviation is shown as an error bar. The continuous lines show a least square error fit of the form $\frac{v_{av}}{v_0}(k) = 1 - \frac{b}{v_0} \cdot e^{-k/a}$ to the data. The obtained fit parameters are $a = 0.48 \pm 0.07$, $b = 2.3 \pm 0.7$ and $a = 0.69 \pm 0.06$, $b = 1.4 \pm 0.2$ for 350 and 500 particles respectively.

Based on the data in figure 7.4, we presume that the velocity depends on the exclusion limit k in an exponential manner. Indeed, we can fit the data with a functional form $\frac{v_{av}}{v_0}(k) = 1 - \frac{b}{v_0} \cdot e^{-k/a}$. In order to verify that the velocity v_{av} is well described by this proposed function, we consider a log-linear plot of $\frac{v_{av}}{v_0}$ against

k at a high particle density (figure 7.5). The exponential nature of the increase can best be observed for high densities, where the expected velocity v_0 is attained at larger values of k compared to low particles densities. At particle densities below 0.25, the maximum velocity is already attained at $k = 2$, such that no information can be extracted on how the velocity increases to its maximum.

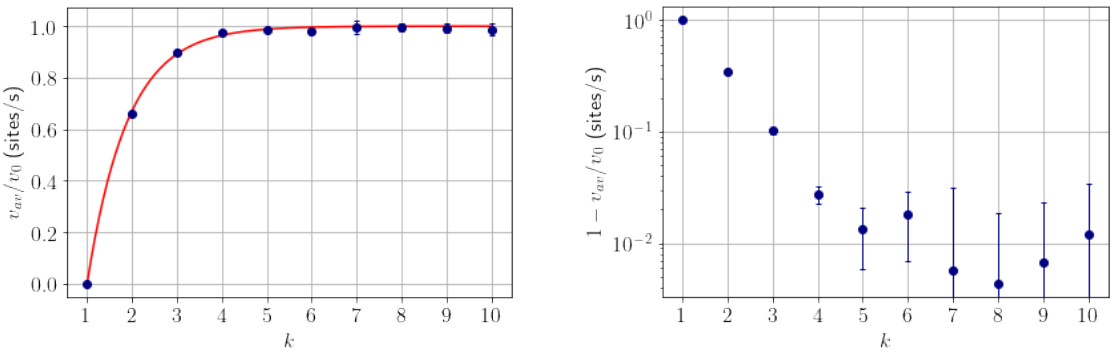


Figure 7.5: The scaled average velocity $\frac{v_{av}}{v_0}$ against the maximum number of particles per site k on a linear y-scale (left) and $1 - \frac{v_{av}}{v_0}$ on a logarithmic y-scale (right). Parameters $N = 1000$ and $L = 1000$ were used to obtain these results. Each data point is an average of three simulations, of which the standard deviation is shown as an error bar. The left plot includes a least square error fit of the form $v(k) = 1 - \frac{b}{v_0} \cdot e^{-k/a}$ to the data. The obtained fit parameters are $a = 0.90 \pm 0.01$ and $b = 2.14 \pm 0.04$.

From the fitted exponential curves we can determine the characteristic number of particles per site, a, for different particle densities. The a-values and the corresponding uncertainties obtained from the least squares approximations are shown in figure 7.6. The values for b, ranging between 1.4 and 3.4, can be found in appendix B.

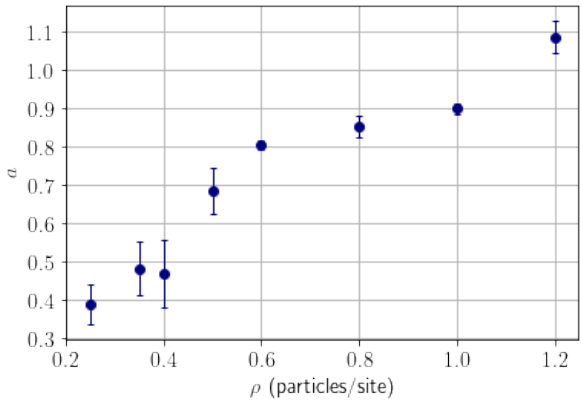


Figure 7.6: The values of a corresponding to the curve of best fit of the form $\frac{v_{av}}{v_0} = 1 - \frac{b}{v_0} e^{-k/a}$ to the obtained average velocities. The obtained uncertainty in a is shown as an error bar

Note that for particles densities greater than one, the fit function can be fully described by one parameter. In that case, the x-intercept of the function is equal to the particle density.

Next, we will study the effect of particle density on the average velocity more directly. Below, v_{av} is shown as a function of the normalised density $\frac{\rho}{k}$ for exclusion limits of one to four particles per site.

For all considered values of k, the velocity v_{av} approaches the expected velocity v_0 in the limit $\frac{\rho}{k} \to 0$ and decreases to 0 as $\frac{\rho}{k} \to 1$. The way in which these limits are attained varies strongly per value of k. The relation between the velocity and normalised density is almost linear for $k = 1$ while it could obey a power law with power greater than one for 2, 3 and 4 particles allowed per site. Let us investigate if this is the case by applying a least squares approximation of the form $\frac{v_{av}}{v_0} = 1 - \left(\frac{\rho}{k}\right)^q$ to the data. The results are summarized in the following figure and table:

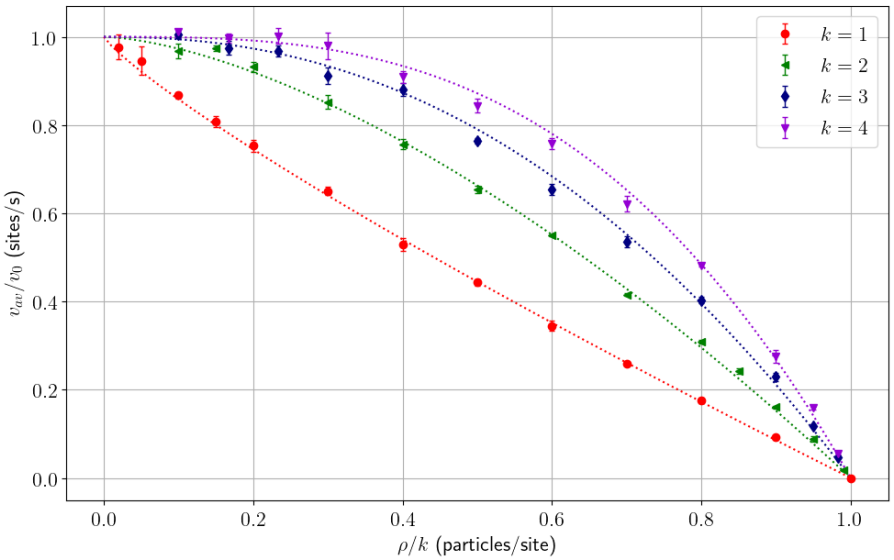


Figure 7.7: The scaled average velocity $\frac{v_{av}}{v_0}$ against the normalised particle density $\frac{\rho}{k}$ for $k = 1$ to 4. The data for $k = 1,2$ were obtained using $L = 1000$ while for $k = 3,4$ a smaller lattice of length 800 was used. Each data point is an average over three simulations, of which the standard deviation is shown as an error bar. Least squares approximations of the form $\frac{v_{av}}{v_0} = 1 - \left(\frac{\rho}{k}\right)^q$ are included as dotted lines.

Table 7.1: The values of q that minimize the square error between the function $\frac{v_{av}}{v_0} = 1 - \left(\frac{\rho}{k}\right)^q$ and the data from figure 7.7. The obtained uncertainty in q is given by $u(q)$.

k	q	$u(q)$
1	0.849	0.005
2	1.569	0.005
3	2.25	0.02
4	2.97	0.01

Note that the differences in q between subsequent k-values lie very close to each other: 0.72, 0.68 and 0.72. Furthermore remark that in the limit $k \to \infty$, v_{av} will approach the constant function v_0.

7.4. Discussion

By simulating next-nearest neighbour random walk, we tested the hypothesis that the observed non-linearity and fluctuations of the sample variance are mainly a result of particles not being able to overtake each other. When a site can still only be occupied by one particle at a time and the particles diffuse to next-nearest neighbour sites, the sample variance is indeed found to increase linearly with time. The fact that the gradient is much lower than the expected diffusion constant, indicates that the exclusion constraint still strongly influences the motion.
When further investigating the diffusive properties of this model, the accuracy of D_{av} can still be improved by determining the displacement in multiple time intervals of the same trajectory. In this way, the sample variance for a certain value of t is an average of multiple displacements instead of just one.

Next to the diffusive behaviour, we analysed the influence of the number of particles allowed per site k on the average velocity. The decrease in v_{av} for decreasing k appears to be well described by an exponential function, as the linearity of the data points in figure 7.5(b) for low k confirms. The characteristic number of particles per site a which we derived from the fitted curves increases as a function of density, as expected. For low particle densities, the uncertainties in a and b are relatively large, since the velocity increases very rapidly as a function of k. The values of the fit parameters are more accurate for higher particle densities.

The manner in which the velocity decreases for higher particle densities differs strongly as a function of k. Power laws of the normalised density were fitted to the data to quantify and better understand the difference in velocity behaviour. For $k = 1$, all data points lie on the fitted curve within their uncertainty. Moreover, the relative uncertainty in the obtained power is 0.6% and indicates that the curve fit is of good quality. For $k = 3$ and $k = 4$, the curves correspond slightly less well to the data, especially for normalised densities around 0.5 particles/site. The differences in obtained powers for subsequent k-values lie close to each other. Simulations for higher k could be performed to investigate if in this case a power law remains a good approximation and if the increase in power for increasing k remains approximately constant.
As described in section 7.1, the velocity depends linearly on the density in the TASEP model. Remarkably, our model gives a power of 0.849 ± 0.005 for $k = 1$. A possible explanation for the faster velocity decrease as a function of density is the fact that our model includes bidirectional motion, whereas motion in the TASEP model is strictly forward. By determining the velocity for very low rates of diffusion and backward transport compared to the forward transport rate, it can be verified if the bidirectional motion accounts for the deviation from the TASEP model.

For further research on the run-and-tumble model, it is interesting to simulate motion where the rates have a spatial dependence. In that situation, the analysis of section 5 does not apply and it is more difficult to analytically derive the limiting distribution. However, via simulations it may be possible to for example determine the steady state density and velocity profile.

8

Modelling kinesin motion

In this chapter we will turn to an application of the one-dimensional transport model for which we have already derived the limiting distribution, determined the large deviation properties and studied the dynamics under influence of excluded volume interactions. We will apply the run-and-tumble model in combination with an exclusion constraint to the motion of a molecular motor, responsible for transport of large components along linear filaments within cells. The particle velocity as a function of density which we obtain from numerical simulations will be compared to experimental data of kinesin-II motors to investigate how well the model can describe kinesin motility in a crowded environment.

8.1. Molecular motors

Within the dynamical environment of a biological cell, there is a constant need of transportation of cell components. Small components can diffuse within the cytoplasm but larger components like organelles or vesicles are taken to their destination by molecular motors. Molecular motors are enzymes that convert chemical energy, obtained from hydrolysing ATP, into mechanical energy. Most molecular motors perform linear motion along filaments that are part of the cytoskeleton of the cell. For example the motor proteins dynein and kinesin move along microtubule filaments, while myosin uses actin filaments as its track. The motor proteins have a defined direction in which they move, due to the polarity of the filament they 'step' on. Kinesins move towards the plus end of microtubules (away from the nucleus) while dyneins move towards the minus end. Myosins, that are among other things responsible for muscle contraction, move in the plus direction on actin filaments (with an exception for myosin-IV).

As already mentioned, the molecular motor which we will apply our model to is kinesin-II. This motor drives intraflagellar transport in cilia [3]. In order to model its motion well, we need to be familiar with the structural characteristics of microtubules as well as the stepping motion of kinesin molecules. Therefore, we will provide an introduction to microtubules and to the motility of kinesin.

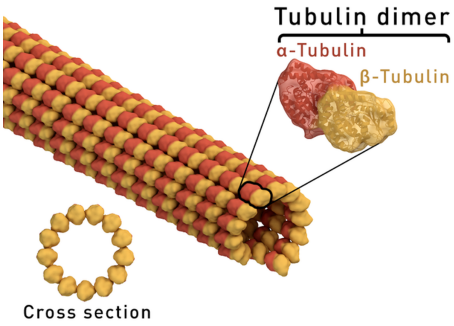


Figure 8.1: A representation of a microtubule and the $\alpha\beta$-tubulin dimers from which it is constructed.
Published by Thomas Splettstoesser under the Creative Commons Attribution-Share Alike 4.0 International license, on `https://commons.wikimedia.org/wiki/File:Microtubule_structure.png`.

8.1.1. Microtubules

Microtubules are hollow cylinders with 13 protofilaments [1]. Each of the protofilaments is built up of $\alpha\beta$-tubulin dimers with a length of 8 nm, that are positioned head to tail. Figure 8.1 shows the alignment of the tubulin dimers in a microtubule.

8.1.2. Mechanochemical cycle of kinesin

Kinesin molecules have two motor domains (heads) that are connected to a tail domain by a coiled-coil stalk [1]. To the tail domain, cellular cargo can bind. The motor domains step forward on a microtubule in a head-over-head manner. During a step, the lagging motor head unbinds from its tubulin dimer and passes the leading head, before attaching to the next available tubulin binding site. At first, the lagging head is tightly bound to the microtubule and to ATP, which it hydrolyses in order to detach from the tubulin. Meanwhile, dissociation of ADP which was bound to the leading head, followed by binding of ATP, drives the motion of the rear head. This process repeats itself, where with each step the molecule takes itself and its cargo 8 nm closer to the plus end of the protofilament.

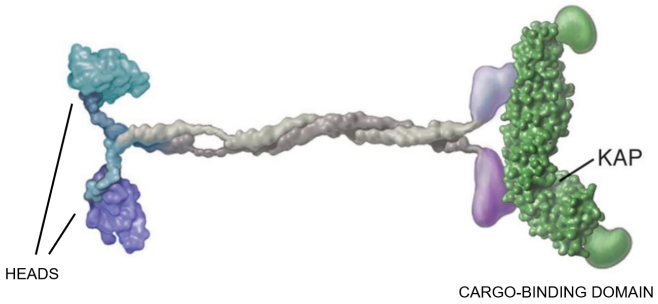


Figure 8.2: Illustration of heterotrimeric kinesin-II with two distinct motor subunits on the left, the coiled-coil stalk in the middle and a subunit called KAP bound to the tail domain.
From the Vale lab, `https://valelab.ucsf.edu/motors/motortoolboxkinesin/`, that permits the use of its illustrations for educational purposes.

Even though kinesin mostly moves away from the cell nucleus, steps in the opposite direction also take place. The frequency of such backsteps is higher at larger loads. Currently, backsteps are thought to be a forward step with reversed direction. However, recent research suggests that movement in the opposite direction should be considered slips rather than steps [19]. Multiple different methods to incorporate backstepping in the mechanochemical cycle of kinesin have been proposed, but it is still a topic of debate what the right explanation is and whether or not ATP is required for a backstep [5].

When applying the run-and-tumble model to kinesin motors, we can account for backstepping by the $\sigma = -1$ state and let $\sigma = 1$ correspond to the normal forward motion. In the next section, we will further elaborate on how the model should be adjusted to the motion of kinesin on microtubules.

8.2. Method

The simulation algorithm explained in section 7.2 needs to be slightly adapted in order to apply it to the motion of kinesin. Since kinesin is a dimeric protein, a motor is assumed to occupy two subsequent lattice sites in the model. This corresponds to the assumption that both motor heads are bound to the microtubule filament. Furthermore, the stepping rates need to be chosen such that the net velocity corresponds to the velocity of kinesin at low densities. As a reference, we will use the results from recent research on the crowding dynamics of heterotrimeric kinesin-II by Kushwaha, Acar, Miedema, Denisov, Schall and Peterman [12]. In this research they performed kinesin-II motility assays using Total Internal Reflection Fluorescence (TIRF) microscopy. Microtubules were attached to a glass slide, to which labeled and unlabeled kinesin-II molecules (in a PEM12 buffer) could bind. For further details of the experimental method, we refer to [12].
Applying particle tracking to a single molecule assay, Kushwaha et al. obtained an average velocity of $0.33 \pm 0.01\,\mu$m/s. Correlation analysis for multiple motors yielded a velocity of $0.30 \pm 0.02\,\mu$m/s for densities below

10 motors/μm.
Setting the length of a lattice site in our model to 8 nm, the length of a tubulin dimer, this latter velocity is equal to 38 ± 3 sites/s.

To model kinesin-II , we use the following rates (in s^{-1}):

$$\begin{aligned} \lambda_1 &= 10 \quad \eta_1 = 100 \quad \gamma_1 = 37 \\ \lambda_{-1} &= 10 \quad \eta_{-1} = 48 \quad \gamma_{-1} = 50 \end{aligned} \tag{8.1}$$

The rates η_σ and γ_σ were chosen such that the asymptotic velocity is in accordance with the velocity of kinesin-II at low densities. The asymptotic velocity, obtained from the stationary distribution (2.18), is 37.1 sites/s. The non-zero random walk rates λ_σ account for displacements due to thermal noise. At the nanoscale these fluctuations are not negligible. Since thermal noise is a result of collisions with small molecules and is independent of the mechanochemical state of a particle, λ_1 and λ_{-1} are taken equal.

The way to model a microtubule should be carefully considered. In section 7.3.2, we found that modelling a single lane with a certain k-value is not equivalent to using the same amount of independent lanes. In order to model the motility of kinesin-II , we decide to model the microtubule protofilaments as independent lanes. The reason behind this choice is that at high densities, motors can hardly make sidesteps, since the neighbouring tubulins will already be occupied by other motors.

With the rates of (8.1) and an exclusion parameter of $k = 1$, we will run simulations for different particle numbers. In this way we can study the motility of kinesin-II as predicted by our model and compare this to experimentally observed data.

8.3. Results

The following figure shows the average velocity obtained from the simulations for different number of particles on a lattice of length 1000. Using that one site represents a tubulin dimer of 8 nm long, the velocity has been given in μm/s and the particle density in particles/μm. Additionally, the kinesin-II velocity determined by Kushwaha et al. [12] from an *in vitro* motility assay is shown in figure 8.3. From time series of TIRF images they extracted average velocities and corresponding densities of kinesin-II motors per μm of microtubule. With our Monte Carlo algorithm we simulated one linear filament. Best agreement between the predicted velocities and the experimental data was obtained when dividing the densities from the experimental data by 13. This is the amount of protofilaments of a microtubule.

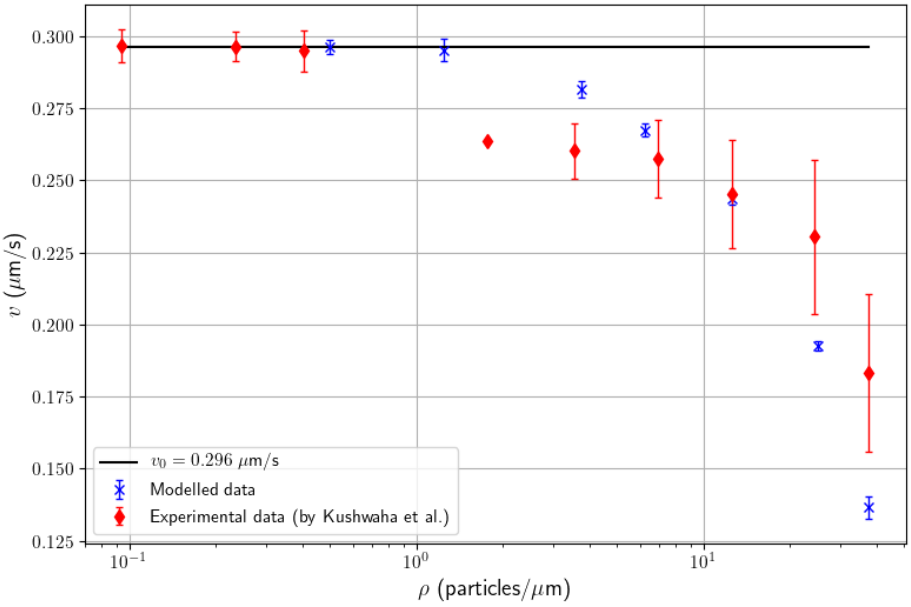


Figure 8.3: The velocity v against the particle density ρ. The data marked with a cross are the average velocities of three simulations of the Monte Carlo algorithm, with standard deviations shown as error bars. The data marked with a diamond show the velocity and its uncertainty of kinesin-II motors, as experimentally determined by Kushwaha et al. [12].

8.4. Discussion

From figure 8.3, it follows that our model predicts a steeper velocity decrease when the kinesin density increases than experimentally observed. In addition, the experimentally determined velocity starts to decrease at lower densities: at a density of 1.8 particles/μm, the measured velocity has decreased by 11%, while the predicted velocity has only decreased a similar amount at a density of 6.3 particles/μm.
A model in which particles move under hard core exclusion with expected velocity equal to the maximum motor velocity does not seem sufficient to capture the dynamics of kinesin-II under crowded circumstances. Note that this can still be better established by comparing the predictions from the model to results of multiple experimental studies.

A main characteristic of kinesin motility that is not incorporated in the model is the binding and unbinding of motors to the filament. In reality, molecular motors attach to a location on the track and detach again after a period of time. Since this results in a more dynamic distribution of the particles over the lattice, it will likely influence the average particle velocity as well. In addition, the waiting time of kinesin molecules before a detachment or a backstep is longer than before a forward step [19], which of course affects the average velocity. It is of interest to determine if including attachment and detachment in the model (referred to as Langmuir kinetics, see for example [14, 15]) would increase the resemblance to the experimental data. The detachment rates could in this case be determined from a measured average run-length of the molecular motor. Moreover, the average run length would provide an extra quantity, next to the velocity, to calibrate our model with. For example, a better estimate of the diffusion rates λ_1 and λ_{-1} could then be made. Namely, these rates do not influence the average velocity but will influence the average run length of the particles.

The rates η_σ and γ_σ for making forward or backward steps and for changing the direction of motion respectively, were taken such that the asymptotic velocity corresponds to the maximum motor velocity. The ratio between forward and backward motion for the rates given by (8.1) is 2.8. As of yet, research data on the conditions and frequency of molecular motor backstepping is rather limited and a reference for the ratio between forward and backward steps of kinesin-II could not be obtained. If such data would be available, the predictions of our model for kinesin-II could be improved by taking more suitable η_σ and γ_σ, while keeping the expected velocity constant. However, an adaptation of this kind is not expected to eliminate the deviation

from the experimental data on its own.

Another point of discussion when comparing our simulated results to experimental data is sidestepping of kinesin-II motors. In our model we neglected sidestepping of motors to a neighbouring protofilament. The reason behind this is that we ran simulations with high motor densities relative to the maximum motor density of ~ 39 motors per μm per lane observed by Kushwaha et al.[12] At densities close to this maximum, the occupation of sites on a protofilament is greater than 50% and it will be difficult for motors to step to a parallel protofilament. It therefore does not seem likely that the assumption of no sidestepping causes the deviation between the simulated and experimental data.
This deviation was smallest when setting the number of accessible lanes to 13. An optimal number of lanes less than 13 would be more accurate, however. Namely, not all protofilaments could be occupied by motors in the motility assay, because of hindrance by the glass slide and neighbouring motors. The actual number of accessible lanes depends on the experimental setup as well as the size of the molecular motors and their cargo.

A final remark is that the extent to which a molecular motor is affected by crowding varies largely among motors, even within the kinesin family. Kushwaha et al. also performed motility assays for kinesin-I and OSM-3[1], whose velocities both decrease at significantly lower densities than for kinesin-II . This difference can not be obtained only by varying the parameters of the currently used run-and-tumble model under exclusion interactions. Further extensions to the proposed model are required in order to give a good description of molecular motor motion in a crowded environment, which does not appear to be governed merely by simple exclusion.

[1] OSM-3 is homodimeric kinesin-II from *C. elegans* [12]

9

Conclusion

Run-and-tumble motion as studied in this thesis, converges to Brownian motion under diffusive scaling of the position and time. We showed this by convergence of the characteristic function of the scaled process. We computed the limiting diffusion coefficient, which exhibits homogenisation over the stationary distribution of the internal state. In addition, this run-and-tumble model is shown to satisfy the large deviation principle, from which it follows that the central limit theorem applies to the process. By involving the free energy function, we derived an implicit expression for the rate by which the probability of rare events tends to zero.

By means of simulations, we studied the influence of excluded volume interactions between the active particles on a lattice and discovered that no normal diffusion takes place when particles cannot overtake each other. The average velocity of the particles is found to converge exponentially to its expected value as the number of particles allowed per position increases. Furthermore, as the lattice becomes crowded with particles, the velocity decreases, where the manner of this decrease depends strongly on the number of particles allowed per site.

The active particle model with excluded volume constraints was applied to the motion of the molecular motor kinesin-II. We predicted the average velocity for different motor densities and compared this to experimental data from Kushwaha et al. [12]. The velocity decrease deviates from the measured velocity of kinesin-II. From this we conclude that the motion of molecular motors in crowded circumstances is an intricate process that cannot merely be described by a model with active transport, thermal noise and simple exclusion. By also including binding and unbinding of motors to the lattice in the model, a better approximation of the dynamical behaviour of molecular motor ensembles may be obtained. Finally, as a direction of further research we suggest to (numerically) analyse the case of spatial inhomogeneity in the rates of motion.

Acknowledgements

I would like to thank my supervisors Prof. Frank Redig and Dr. Timon Idema for their guidance during this project. They explained new concepts, gave valuable feedback on my findings and helped me determine good follow-up questions. Furthermore, they were very open to answering questions and took time to think along when I encountered a problem.
I am also grateful for the help of dr. Stefan Grosskinsky, who shared his expertise on stochastic Monte Carlo methods with me.
Finally, I would like to acknowledge the helpful suggestions and critical view of the members of the Idema group during the weekly meetings.

A

A.0.1. Commutation of the semigroup and translation

The generator of the run-and-tumble process with internal state dependence (5.1) commutes with translation in space. Define the translation operator $\tau_a f(x,\sigma) = f(x+a,\sigma)$. The commutativity can be derived as follows

$$\begin{aligned}(L\tau_a f)(x,\sigma) =& \lambda(\sigma)(\tau_a f(x+1,\sigma) + \tau_a f(x-1,\sigma) - 2\tau_a f(x,\sigma)) + \eta(\sigma)(\tau_a f(x+\sigma,\sigma) - \tau_a f(x,\sigma)) \\ &+ \gamma(\sigma)(\tau_a f(x,-\sigma) - \tau_a f(x,\sigma)) \\ =& \lambda(\sigma)(f(x+a+1,\sigma) + f(x+a-1,\sigma) - 2f(x+a,\sigma)) + \eta(\sigma)(f(x+a+\sigma,\sigma) - f(x+a,\sigma)) \\ &+ \gamma(\sigma)(f(x+a,-\sigma) - f(x+a,\sigma)) \\ =& (\tau_a L f)(x,\sigma).\end{aligned} \tag{A.1}$$

From the relation $S_t = e^{tL}$, which can be defined in terms of its Taylor series, it follows that the semigroup commutes with spacial translation as well. Namely, when L commutes with τ_a, so does L^n.

A.0.2. Limit and logarithm

In the the derivation of $F(\alpha)$ a limit of the form $\lim_{t\to\infty} \frac{1}{t}\log(f(t)+g(t))$ appears, where f and g are positive, real-valued functions. With the following lemma we can determine such limits:

Lemma A.1 *Let $f(t), g(t) : [0,\infty] \to (0,\infty]$ and assume that $\lim_{t\to\infty} \frac{1}{t}\log\Big(\max_{t\to\infty}\{f(t),g(t)\}\Big)$ exists, then*

$$\lim_{t\to\infty} \frac{1}{t}\log(f(t)+g(t)) = \lim_{t\to\infty} \frac{1}{t}\log\Big(\max_{t\to\infty}\{f(t),g(t)\}\Big).$$

Proof. Suppose, without loss of generality, that $\max_{t\to\infty}\{f(t),g(t)\} = f(t)$. To prove the lemma, it suffices to show that

$$\limsup_{t\to\infty} \frac{1}{t}\log(f(t)+g(t)) \le \lim_{t\to\infty} \frac{1}{t}\log f(t) \le \liminf_{t\to\infty} \frac{1}{t}\log(f(t)+g(t)).$$

The first inequality follows from the fact that

$$f(t)+g(t) \le 2\max\{f(t),g(t)\} \quad \forall\, t \ge 0,$$

which implies

$$\limsup_{t\to\infty} \frac{1}{t}\log(f(t)+g(t)) \le \lim_{t\to\infty} \frac{1}{t}\big(\log f(t) + \log 2\big) = \lim_{t\to\infty} \frac{1}{t}\log f(t).$$

For the second inequality, note that $f(t) + g(t) > f(t)$. As a consequence,

$$\liminf_{t\to\infty} \frac{1}{t} \log(f(t) + g(t)) \geq \lim_{t\to\infty} \frac{1}{t} \log f(t).$$

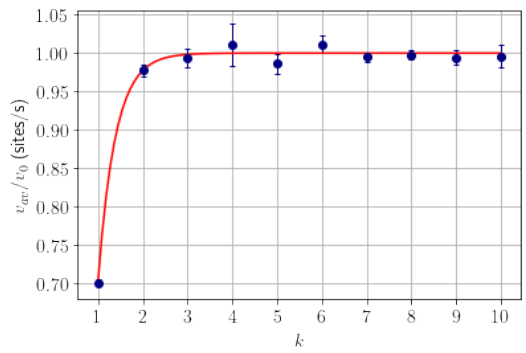
□

For the derivation of the free energy function $F(\alpha)$, lemma (A.1) can be used to determine that

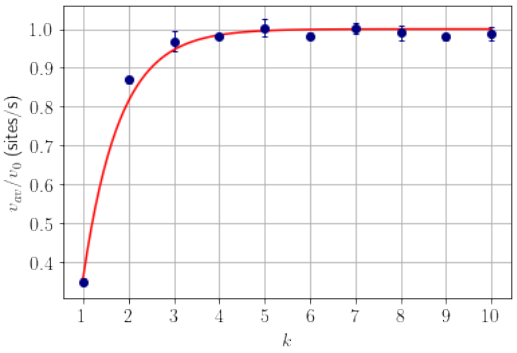
$$\lim_{t\to\infty} \frac{1}{t} \log(ae^{bt} + ce^{dt}) = \max\{b, d\},$$

with $a, b, c, d \in \mathbb{R}$ and $a, c > 0$.

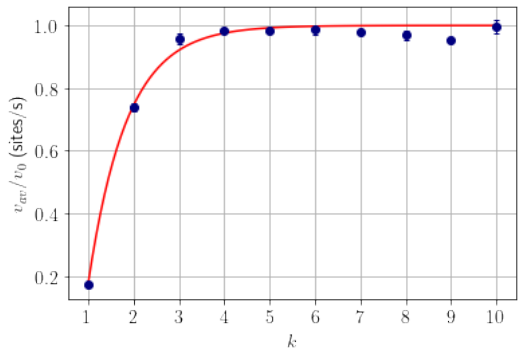
B



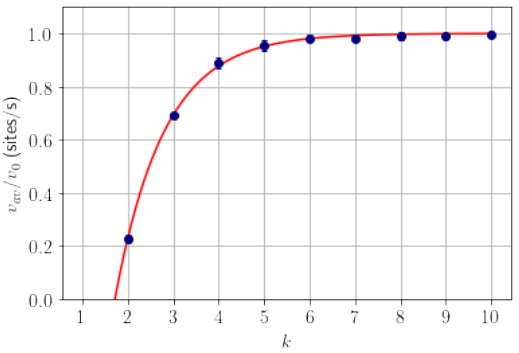
(a) $N = 250$



(b) $N = 600$



(c) $N = 800$



(d) $N = 1200$

Figure B.1: The average velocity scaled by the expected velocity $\frac{v_{av}}{v_0}$ against the maximum number of particles per site k for particle numbers $N = 250, 600, 800, 1200$. A lattice size of $L = 1000$ was used for the data of subfigures (a)-(c); the data in subfigure (d) were obtained using $L = 700$. Each data point is an average of three simulations, of which the standard deviation is shown as an error bar. Least squares approximations of the form $\frac{v_{av}}{v_0} = 1 - \frac{b}{v_0} \cdot e^{-k/a}$ were applied to the data. The values of a for the best fit are plotted in figure 7.6 and the values for b can be found in table B.1 below.

Table B.1: The values of b that minimize the square error between the function $\frac{v_{av}}{v_0} = 1 - \frac{b}{v_0} e^{-k/a}$ and the average velocities for different k-values, obtained from simulations. The obtained uncertainty in b is given by $u(b)$.

N	b	$u(b)$
250	3	1
350	2.3	0.7
400	3	1
500	1.4	0.2
600	1.57	0.04
800	1.88	0.08
1000	2.14	0.04
1200	3.4	0.3

Bibliography

[1] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter. *Molecular biology of the cell*. Garland Science, New York, 6th edition, 2015.

[2] P. Billingsley. *Convergence of probability measures*. Wiley, New York, 2nd edition, 1999.

[3] J. R. Broekhuis, W. Y. Leong, and G. Jansen. Regulation of Cilium Length and Intraflagellar Transport. In *International Review of Cell and Molecular Biology*, volume 303, pages 101–138. Elsevier Inc., 1 2013. doi: 10.1016/B978-0-12-407697-6.00003-9.

[4] W. Bryc. A remark on the connection between the large deviation principle and the central limit theorem. *Statistics & Probability Letters*, 18(4):253–256, 1993.

[5] B. E. Clancy and S. M. Block. Kinesin single-molecule mechanics. In *Comprehensive Biophysics*, volume 4. Academic Press, Amsterdam, 2012.

[6] T. Demaerel and C. Maes. Active processes in one dimension. *Physical Review E*, 97(3), 3 2018. ISSN 24700053. doi: 10.1103/PhysRevE.97.032604.

[7] A. Dembo and O. Zeituni. *Large Deviations Techniques and Applications*. Springer, New York, 2nd edition, 1998.

[8] F. den Hollander. *Large Deviations*. American Mathematical Society, Providence, Rhode Island, 2000.

[9] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier. Exact solution of a 1d asymmetric exclusion model using a matrix formulation. *Journal of Physics A: Mathematical and General*, 26:1493, 1993.

[10] G. Grimmet and D. Welsh. *Probability: an introduction*. Oxford University Press, Oxford, 2nd edition, 2014.

[11] R. Großmann, F. Peruani, and M. Bär. Diffusion properties of active particles with directional reversal. *New Journal of Physics*, 18(4), 4 2016. ISSN 13672630. doi: 10.1088/1367-2630/18/4/043009.

[12] V. S. Kushwaha, S. Acar, D. M. Miedema, D. V. Denisov, P Schall, and E. J. G. Peterman. The crowding dynamics of the motor protein kinesin-ii. *PLoS ONE*, 15(2), 2 2020. ISSN 19326203. doi: 10.1371/journal.pone.0228930.

[13] K. Malakar, V. Jemseena, A. Kundu, K. V. Kumar, S. Sabhapandit, S. N. Majumdar, S. Redner, and A. Dhar. Steady state, relaxation and first-passage properties of a run-and-tumble particle in one-dimension. 11 2017. doi: 10.1088/1742-5468/aab84f. URL `http://arxiv.org/abs/1711.08474`. arXivId: 1711.08474.

[14] D. M. Miedema, V. S. Kushwaha, D. V. Denisov, S. Acar, B. Nienhuis, E. J. G. Peterman, and P. Schall. Correlation imaging reveals specific crowding dynamics of kinesin motor proteins. *Physical Review X*, 7 (4), 11 2017. ISSN 21603308. doi: 10.1103/PhysRevX.7.041037.

[15] A. Parmeggiani, T. Franosch, and E. Frey. Phase coexistence in driven one-dimensional transport. *Physical Review Letters*, 90(8):4, 2003. ISSN 10797114. doi: 10.1103/PhysRevLett.90.086601.

[16] N. Privault. *Understanding Markov Chains*. Springer, Singapore, 2nd edition, 2018.

[17] A. Rabani, G. Ariel, and A. Be'er. Collective motion of spherical bacteria. *PLoS ONE*, 8(12), 12 2013. ISSN 19326203. doi: 10.1371/journal.pone.0083760.

[18] F. Redig. *Basic techniques in interacting particle systems*. September 2014. Lecture notes, TU Delft.

[19] A. Toleikis, N. J. Carter, and R. A. Cross. Backstepping mechanism of kinesin-1. *Biophysical journal*, 119: 1984–1994, 2020. doi: 10.1016/j.bpj.2020.09.034. URL https://doi.org/10.1016/j.bpj.2020.09.034.

[20] B. van Ginkel, B. van Gisbergen, and F. Redig. Run-and-Tumble Motion: The Role of Reversibility. *Journal of Statistical Physics*, 183(3):44, 6 2021. ISSN 0022-4715. doi: 10.1007/s10955-021-02787-1. URL https://link.springer.com/10.1007/s10955-021-02787-1.

[21] B. van Gisbergen and F. Redig. Central limit theorem and large deviations for run and tumble particles on a lattice. 10 2019. URL http://arxiv.org/abs/1910.03350. arXivId: 1910.03350.

[22] T. Vicsek and A. Zafeiris. Collective motion. 10 2010. doi: 10.1016/j.physrep.2012.03.004. URL http://arxiv.org/abs/1010.5017. arXivId: 1010.5017.