Diffusive behaviour and large deviations of one-dimensional active particles with drift

Bart van Gisbergen, 4604253

Bachelor Thesis,
Technical Mathematics and Physics,
TU Delft

Delft, June 28, 2019
Thesis Advisors: Prof. dr. F.H.J. Redig
and Dr. J.M. Thijssen

Contents

Introduction

Active matter is a subject of interest studied both experimentally as well as from the point of view of theoretical mathematics and physics. Active particles can model various phenomena such as: living particles (bacteria, unicellular), molecular motors, etc. [1][2][3]
We will look at a model for active particles for which each particle has an internal state which determines the drift of the particle. Next to that each particle can diffuse or they can switch their internal state [4]. First we will look at the discrete setting, which we will later extend to a continuous space and time. For the latter we will determine the diffusion coefficient of a single particle. After which we will look at the large deviations for such a particle. This model will then be applied to solid state physics by determining the diffusion constant in a semiconductor and by determining the ground state energy of the Bose-Hubbard model with an external magnetic field using diffusion Monte Carlo [5].
This paper is divided into three main sections, each spreading over three chapters. The first three chapters are mathematical background knowledge. In the second three chapters we analyse our model of active particles. In the final three chapters we take this model and apply it to physical phenomena.

1. Markov Processes

In this work we shall mainly be concerned with continuous time Markov Processes, which are defined as follows:

Definition 1.1. *A **Markov Process** on a measurable state space Ω, is a stochastic process $\{X_t, t \geq 0\}$ such that $\forall t > 0, n \in \mathbb{N}, 0 < t_1 < ... < t_n < t_{n+1}$ and $\forall f : \Omega \to \mathbb{R}$, with f bounded and measurable, it holds that $\mathbb{E}(f(X_{t_{n+1}})|X_{t_1}, ..., X_{t_n}) = \mathbb{E}(f(X_{t_{n+1}})|X_{t_n})$.*

In other words a Markov process is a stochastic process for which the distribution of a future state only depends on the current state and not on further states in the past. This is called the Markov property. It is often described by saying that the process is memoryless.

1.1 Poisson distribution

This section is largely based on chapter 12 in a book by Grimmett [6].
The Poisson distribution is inherently linked to continuous time Markov Processes. We will now see how and why this is the case. In order to do this we will first look at why the memoryless property is connected to the exponential distribution, then look at the Poisson distribution itself and finally look at how this Poisson distribution is linked to the exponential distribution.
So lets start of by looking at the exponential distribution. Its density function is defined as follows for parameter $\lambda > 0$:

$$f_\lambda(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{if } t > 0 \\ 0, & \text{if } t \leq 0 \end{cases}$$

From now on we will always assume that $t \geq 0$.
It is clear that $\mathbb{P}(T > t) = e^{-\lambda t}$. This brings us to the first theorem [6].

Theorem 1.1. *The continuous random variable X has the memoryless property if and only if X is exponentially distributed.*

Proof. First assume that X is exponentially distributed.

$$\begin{aligned} \mathbb{P}(X > s + t | X > s) &= \frac{\mathbb{P}(X > s + t \cap X > t)}{\mathbb{P}(X > s)} \\ &= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = \mathbb{P}(X > t) \end{aligned}$$

So X has the memoryless property, because $\mathbb{P}(X > s + t|X > s) = \mathbb{P}(X > t)$, thus the probability to jump is independent of earlier results.

Now assume that X has the memoryless property. Let $F(t) = \mathbb{P}(X > t)$ for $t \geq 0$. Then

$$F(t) = \mathbb{P}(X > s + t | X > s) = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} = \frac{F(s + t)}{F(s)} \tag{1.1}$$

Thus $F(s + t) = F(s)F(t)$. From here we prove that $F(t) = e^{-\lambda t}$.
Take $s = 0$, then $F(t) = F(0)F(t)$. Thus $F(t) = 0$ or $F(0) = 1$. The first doesn't lead to a probability distribution thus we conclude that $F(0) = 1$.
Now take $s = t$, then $F(2t) = F(t)^2$ and by induction also $F(nt) = F(t)^n$. Plugging in $t = t/n$ we find that $F(t) = F(\frac{t}{n})^n$, so $F(\frac{t}{n}) = F(t)^{\frac{1}{n}}$. Thus we can conclude that $F(\frac{m}{n}t) = F(t)^{\frac{m}{n}}$ and also $F(\frac{m}{n}) = F(1)^{\frac{m}{n}}$. This defines $F(t) = e^{-\lambda t}$ for $t \in \mathbb{Q}$. In order to get a well defined distribution function we must limit ourselves to $\lambda > 0$. Now also note that $F(t)$ must be right-continuous, which defines F for $t \in \mathbb{R}$. □

Now we want a distribution for the number of events which happens at a random point in time. An example of this is radioactive decay or a random walk. All we can say about these events is the likelihood of N events happening between time zero and $t > 0$. We call this number N_t. We want N_t to have the following properties:

1. N_t is a random variable taking values in $\mathbb{N} \cup \{0\}$
2. $N_0 = 0$
3. $N_s \leq N_t$ if $s \leq t$
4. if $0 \leq s < t$ then the number of events happening between time s and t is independent of the number of events which happened before time s
5. there exists a number $\lambda > 0$ such that for small positive h it holds that

$$\mathbb{P}(N_{t+h} = n + 1 | N_t = n) = \lambda h + \mathcal{O}(h)$$

$$\mathbb{P}(N_{t+h} = n | N_t = n) = 1 - \lambda h + \mathcal{O}(h)$$

This last property makes it so that 2 events happening in a time interval h is of order h ($\mathcal{O}(h)$).

Theorem 1.2. *Properties 1-5 lead to the Poisson distribution with parameter λt, that is for $t > 0$:*

$$\mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad \textit{for} \quad k = 0, 1, 2, ...$$

For visualization a Poisson process is demonstrated in the following figure.

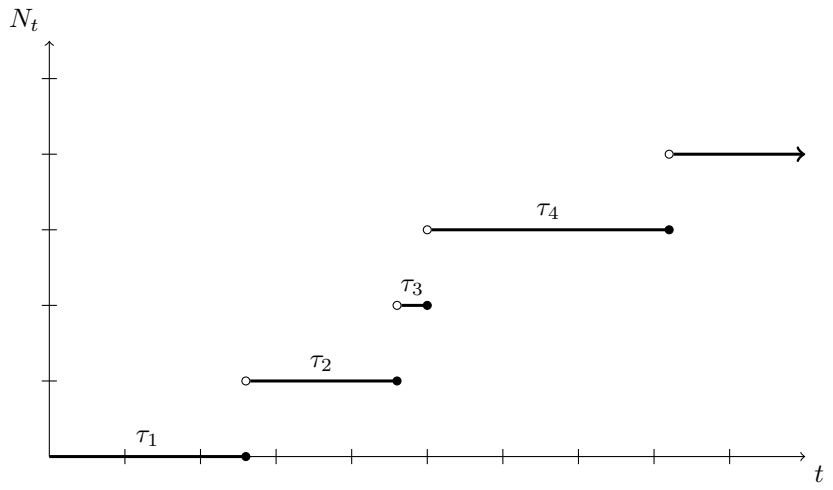

Figure 1: This is a visual representation of the Poisson process. The time increments τ_i are all exponentially distributed.

Proof. For the sake of brevity we will denote $p_k(t) = \mathbb{P}(N_t = k)$. We find, using property number 5 and the partition theorem, that:

$$\begin{aligned}\mathbb{P}(N_{t+h} = k) &= \sum_{i=0}^{k} \mathbb{P}(N_{t+h} = k | N_t = i)\mathbb{P}(N_t = i) \\ &= \mathbb{P}(N_{t+h}|N_t = k-1)\mathbb{P}(N_t = k-1) + \mathbb{P}(N_{t+h}|N_t = k)\mathbb{P}(N_t = k) + \mathcal{O}(h) \\ &= [\lambda h + \mathcal{O}(h)]\mathbb{P}(N_t = k-1) + [1 - \lambda h + \mathcal{O}(h)]\mathbb{P}(N_t = k) + \mathcal{O}(h) \\ &= \lambda h \mathbb{P}(N_t = k-1) + (1 - \lambda h)\mathbb{P}(N_t = k) + \mathcal{O}(h)\end{aligned}$$

Thus we find that for $k \in \mathbb{N}$:

$$p_k(t+h) - p_k(t) = \lambda h \left(p_{k-1}(t) - p_k(t)\right) + \mathcal{O}(h)$$

Now we can divide both sides by h and take the limit $h \downarrow 0$ giving a time derivative on the left hand side:

$$p_k'(t) = \lambda p_{k-1}(t) - \lambda p_k(t) \tag{1.2}$$

We can also do this for $k = 0$:

$$\begin{aligned}\mathbb{P}(N_{t+h} = 0) &= \mathbb{P}(N_{t+h} = 0 | N_t = 0)\mathbb{P}(N_t = 0) \\ &= [1 - \lambda h]\mathbb{P}(N_t = 0) + o(h)\end{aligned}$$

From which it follows as before that

$$p_0'(t) = -\lambda p_0(t)$$

From property 2 we derive the boundary condition for these two equations:

$$p_k(0) = \begin{cases} 1, & \text{if } k = 0 \\ 0, & \text{if } k \neq 0 \end{cases} \tag{1.3}$$

The way to solve these equations is using generating functions. Note that generating functions are unique and defined as follows:

$$G(s,t) = \mathbb{E}(s^{N_t}) = \sum_{k=0}^{\infty} p_k(t)s^k$$

From the definition we find the following properties:

$$\sum_{k=1}^{\infty} p_{k-1}s^k = sG(s,t) \quad \text{and} \quad \sum_{k=0}^{\infty} p_k' s^k = \frac{\partial G}{\partial t}$$

Now using equation (1.2) and the boundary conditions from equation (1.3) we obtain the following differential equation

$$\begin{cases} \frac{\partial G}{\partial t} = \lambda sG - \lambda G \\ G(s,0) = \sum_{k=0}^{\infty} p_k(0)s^k = 1 \end{cases}$$

Solving this we find that

$$G(s,t) = e^{\lambda t(s-1)} = \sum_{k=0}^{\infty} \left(\frac{(\lambda t)^k}{k!} e^{-\lambda t} \right) s^k$$

Thus

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \mathbb{P}(N_t = k)$$

We have now found that when we want properties 1-5 to hold we automatically end up with the Poisson distribution with parameter λt. □

Now let us link the Poisson distribution to the exponential distribution. For this we have the following theorem.

Theorem 1.3. *In a Poisson process with rate λ, the times at which an event occurs $X_1, X_2, ...$ are independent random variables, each having the exponential distribution with parameter λ.*

Proof. We will prove this by induction. First we will show that the probability for having not taken a step in the Poisson distribution is given by the exponential distribution. This follows directly from the definition of the Poisson distribution.

$$\mathbb{P}(N_t = 0) = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = p_{X_1}(t)$$

Where $p_{x_1}(t)$ denotes the probability density function of X_1.
Next we will show that when having k exponentially distributed independent random variables that together they form a Poisson process. Note for $n+1$ exponential distributions the probability density function is given by the convolution of all of them, which is

$$p_{X_1+X_2+...+X_n}(t) = \frac{t^n \lambda^{n+1}}{n!} e^{-\lambda t}$$

Now realize that finding the probability of having k events happen before time t is the same as having the $k+1$'st event happen after time t. Thus we find that

$$\begin{aligned}\mathbb{P}(N_t \leq k) &= \mathbb{P}(X_1 + ... + X_{k+1} > t) \\ &= \int_t^\infty \frac{x^k \lambda^{k+1}}{k!} e^{-\lambda x} dx \\ &= \sum_{i=1}^{k} \frac{(\lambda t)^i}{i!} e^{\lambda t}\end{aligned}$$

Where in the last step we use partial integration to find the sum. Note that this is the sum of Poisson distributions. Thus the exponential distribution is made up of independent exponentially distributed variables. □

We have now seen that in order to ensure the Markov property in a continuous time setting we use the exponential distribution in order to find when the next event occurs. We can then find the probability of k events occurring within a time window $[0, t)$ using the Poisson distribution.

1.2 Discrete state and continuous time Markov chains

We will now see how to create a Markov process [7]. We will again look at the discrete state and continuous time case. Consider a finite set $i, j \in \{1, \ldots, n\}$. After an exponential time t the particles jump to their next state. Then there exists a probability for particles beginning in state i to end up in a state j, this probability is defined as follows:

$$p_t(i, j) = \mathbb{P}(X(t) = j | X(0) = i) \tag{1.4}$$

Now assume we have n possible states, then these probabilities form a matrix with indices i and j:

$$\mathbf{S}_t = \begin{bmatrix} p_t(1,1) & p_t(1,2) & \ldots & p_t(1,n) \\ p_t(2,1) & p_t(2,2) & & \vdots \\ \vdots & & \ddots & \\ p_t(n,1) & \ldots & & p_t(n,n) \end{bmatrix}$$

We now define a function $f : \Omega \to \mathbb{R}$. Because Ω is finite (for now) we can see this as a column vector $\mathbf{f}$. If we multiply $\mathbf{S}_t$ by $\mathbf{f}$ we obtain:

$$\mathbf{S}_t\mathbf{f}(i) = \begin{bmatrix} p_t(1,1) & p_t(1,2) & \dots & p_t(1,n) \\ p_t(2,1) & p_t(2,2) & & \vdots \\ \vdots & & \ddots & \\ p_t(n,1) & \dots & & p_t(n,n) \end{bmatrix} \begin{bmatrix} f(1) \\ \vdots \\ \\ f(n) \end{bmatrix}$$
$$= \sum_j \mathbb{P}(X_t = j | X_0 = i) f(j) = \mathbb{E}(f(X_t)|X_0 = i)$$

Thus, multiplying $\mathbf{f}$ with our matrix $\mathbf{S}_t$ gives us the expected value of $\mathbf{f}$ at time t.
$\mathbf{S}_t$ is known as the **semigroup**, which is defined as the expected value of $f(X_t)$, thus it is defined as follows

$$\mathbf{S}_t f(x) = \mathbb{E}(f(X_t)|X_0 = x)$$

A semigroup has a couple of properties.

Proposition 1.4. *Let $\boldsymbol{S}_t$ be a semigroup and $t \geq 0$, then the following holds:*

(a) Identity at time zero: $\boldsymbol{S}_0 = I$

(b) Semigroup property: $\forall t, s > 0, f : \boldsymbol{S}_{t+s} f = \boldsymbol{S}_t(\boldsymbol{S}_s f) = \boldsymbol{S}_s(\boldsymbol{S}_t f)$

(c) Right continuity: the map $t \to \boldsymbol{S}_t f$ is right continuous

(d) Positivity: $f \geq 0 \implies \boldsymbol{S}_t f \geq 0$

(e) Normalization: $\boldsymbol{S}_t 1 = 1$

(f) Contraction: $\max_x |(\boldsymbol{S}_t f)(x)| \leq \max_x |f(x)|$

Proof. a) This follows from the fact that $\mathbf{S}_0 f(x) = \mathbb{E}(f(X_0)|X_0 = x) = f(x)$, from which follows that $\mathbf{S}_0 = I$.
b)

$$\begin{aligned} \mathbf{S}_{t+s} f(x) &= \mathbb{E}(f(X_{t+s}|X_0 = x)) \\ &= \sum_y f(y) p_{s+t}(x,y) = \sum_y f(y) \sum_z p_t(x,z) p_s(z,y) \\ &= \sum_z p_t(x,z) \sum_y f(y) p_s(z,y) = \sum_z p_t(x,z) \mathbb{E}(f(X_s|X_0 = z) \\ &= \mathbf{S}_t \mathbf{S}_s f(x) \end{aligned}$$

The second equality is proven by switching s and t.
c) A function is right continuous if $\lim_{x\downarrow x_0} f(x) - f(x_0) = 0$. Thus we find that

$$\begin{aligned}\lim_{t\downarrow t_0} \mathbf{S}_t f(x) - \mathbf{S}_{t_0} f(x) &= \lim_{h\downarrow 0} \mathbf{S}_{t+h} f(x) - \mathbf{S}_t f(x) \\ &= \mathbf{S}_t \lim_{h\downarrow 0} \mathbf{S}_h f(x) - f(x) \\ &= \mathbf{S}_t(\mathbf{S}_0 f(x) - f(x)) = \mathbf{S}_t(f(x) - f(x)) = 0\end{aligned}$$

d) If $f \geq 0$ then $\mathbf{S}_t f(x) = \sum_y p_t(x,y) f(y) \geq 0$ because $\forall i,j : p_t(i,j) \geq 0$ because it is a probability.
e) $\mathbf{S}_t 1 = \sum_y p_t(x,y) 1 = 1$, because it is the sum over all probabilities.
f)

$$\max_x |(\mathbf{S}_t f)(x)| = \max_x |\mathbb{E}(f(X_t)|X_0 = x)| \leq \max_x \mathbb{E}(|f(X_t)||X_0 = x) \leq \max_x |f(x)|$$

□

Because of property c), we can look at $\frac{d}{dt}\mathbf{S}_t|_{t=0}$. From the definition of the derivative we then define:

$$\mathbf{L} := \left.\frac{d}{dt}\mathbf{S}_t\right|_{t=0} = \lim_{t\to 0} \frac{\mathbf{S}_t - \mathbf{I}}{t} \tag{1.5}$$

$\mathbf{L}$ is known as the **generator matrix**. Combining $\mathbf{S}_{t+s} = \mathbf{S}_t \mathbf{S}_s$ with $\frac{d}{dt}\mathbf{S}_t|_{t=0} = \mathbf{L}$, we get $\mathbf{S}_t' = \frac{d}{dt}\mathbf{S}_{t+s}|_{s=0} = \mathbf{L}\mathbf{S}_t$, from which it follows that we can write

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

Note that this function is well defined if we take the following definition for the exponential matrix:

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

Let us take a closer look at the elements of our generator matrix, $\mathbf{L}$. We can just take the definition as given in equation (1.5) and see that we can take each individual element of our matrix and define it as:

$$\mathbf{L}_{i,j} = \lim_{t\to 0} \frac{\mathbb{P}(X_t = j|X_0 = i) - \delta_{i,j}}{t} = c(i,j)$$

In this equation $c(i,j)$ denotes the hopping rates from state i to state j. This leads to the following expression for the elements on the diagonal:

$$c(i,i) = -\sum_{k\neq i} c(i,k)$$

We can then write $\mathbf{L}$ as the following matrix:

$$\mathbf{L} = \begin{bmatrix} -\sum_{k\neq 1} c(1,k) & c(1,2) & \dots & c(1,n) \\ c(2,1) & -\sum_{k\neq 2} c(2,k) & & \vdots \\ \vdots & & \ddots & \\ c(n,1) & \dots & & -\sum_{k\neq n} c(n,k) \end{bmatrix} \tag{1.6}$$

From now on we want to think of $\mathbf{S}_t$ not as a matrix but as an operation on a function $\mathbf{f} : \Omega \to \mathbb{R}$, so from now on the notation will not be in bold. This also holds for the generator matrix $\mathbf{L}$. If Ω is locally compact, such as $\mathbb{R}^d$, then we typically choose $f \in \mathcal{C}_0(\Omega)$, which is the space of continuous functions which vanish at infinity.

1.3 Random Walk on $\mathbb{Z}$ in discrete time

Let us look at an example of a Markov Process. In this report we will be discussing a model which is an extension of a random walk, therefore it is a good idea to discuss the random walk first. When looking at a single particle, it can move from site i to $i+1$ with probability $\frac{1}{2}$ and to side $i-1$ with the same probability.
The particle can walk along an infinite axis, $\mathbb{Z}$. Say we have a single particle in the system and it starts at site i, we can then represent our initial state, x with a column vector of infinite size. This vector contains of only zeros except for a one in the i-th row, thus $f(j) = \delta_{i,j}$. We can also represent the probability matrix of moving from an initial state x to our next state y using the following matrix:

$$P = \begin{pmatrix} \ddots & \ddots & 0 & 0 & \cdots & 0 & \cdots \\ \ddots & 0 & \frac{1}{2} & 0 & \cdots & 0 & \cdots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \cdots & 0 & \cdots \\ & & \frac{1}{2} & & & & \\ & & & \ddots & \ddots & & \\ & & & & & \frac{1}{2} & \\ \cdots & 0 & \cdots & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \cdots & 0 & \cdots & 0 & 0 & \frac{1}{2} & \ddots \end{pmatrix}$$

We find that $y = P \cdot x$, y now represents the probabilities of finding a particle in a certain site after the particle has taken a single step. If we now want to find the probabilities of going to state z after state y we can then again multiply by this same matrix. We find that $z = P \cdot y = P^2 \cdot x$. Thus we see that $p_n(x,y) = P^n \cdot x$, where $p_n(x,y)$ is the probability of going from state x to state y after n steps.
In order to introduce the time we need to find how many jumps a particle makes between

time 0 and time $t > 0$. We call the number of jumps N_t. We find that for this process N_t is a rate one Poisson process. This means that:

$$\mathbb{P}(N_t = n) = \frac{t^n}{n!}e^{-t}$$

The fact that it is a rate one Poisson process means that, on average, the system takes one step every time step, thus $\mathbb{E}(N_t) = t$.
Now we can find S_t for the discrete space and continuous time random walk as follows

$$S_t f = \mathbb{E}(f(X_t)|X_0 = x) = \frac{1}{2}tf(x+1) + \frac{1}{2}tf(x-1) + (1-t)f(x)$$

Using equation (1.5) we now find that our generator L for this process becomes

$$Lf = \frac{1}{2}f(x+1) + \frac{1}{2}f(x-1) - f(x) = \frac{1}{2}\left(f(x+1) + f(x-1) - 2f(x)\right)$$

1.4 Brownian Motion

Brownian motion can be compared to a random walk in continuous space and time. Again the amount of steps taken between time 0 and time $t > 0$ is a rate one Poisson process.
Now we define Brownian motion

Definition 1.2. *Brownian motion is a process* $\{B_t, t \geq 0\}$ *with the following properties:*

1. $B_0 = 0$
2. $(B_{t_i} - B_{t_{i-1}}) \sim \mathcal{N}(0, t_i - t_{i-1})$ *for* $t_i \in \{0, t_1, ..., t_n\}$
3. $t \to B_t$ *is continuous*

with $\mathcal{N}(\mu, \sigma^2)$ *the normal distribution with expectation value* μ *and variance* σ^2

Thus we can define our semigroup for Brownian motion as such

$$S_t f(x) = \mathbb{E}_x[f(B_t)] = \mathbb{E}[f(x + \mathcal{N}(0,t))]$$

If we Taylor expand this we find that

$$\mathbb{E}[f(x+\mathcal{N}(0,t))] = f(x) + f'(x)\mathbb{E}[\mathcal{N}(0,t)] + \frac{1}{2}f''(x)\mathbb{E}[\mathcal{N}^2(0,t)] + \mathcal{O}(t^2)$$

Now we know that $\mathbb{E}[\mathcal{N}(0,t)] = 0$ and $\mathbb{E}[\mathcal{N}^2(0,t)] = t$. Using equation (1.5) we can determine our generator for Brownian motion, $\mathcal{L}_{BM}$

$$\mathcal{L}_{BM} = \lim_{t\to 0}\frac{S_t f(x) - f(x)}{t} = \lim_{t\to 0}\frac{\mathbb{E}_x[f(B_t)] - f(x)}{t} = \frac{1}{2}f''(x)$$

In the next chapter we will see that Brownian motion can be obtained as scaling limit of the continuous time random walk (X_t) through $\lim_{\varepsilon\to 0}\varepsilon X_{\varepsilon^{-2}t} = B_t$

2. Scaling

When looking at the convergence of a sequence of random variables, one often needs to scale the variables in a certain way in order to see anything non-trivial in the limit. A good example of this is the scaling in the central limit theorem [6].

Theorem 2.1 (Central Limit Theorem)**.** *Let $X_1, X_2, ...$ be independent and identically distributed random variables, each with mean μ and non-zero variance σ^2. The standardized version*

$$Z_N = \frac{S_N - N\mu}{\sigma\sqrt{N}} \quad \xrightarrow{N\to\infty} \quad \mathcal{N}(0,1)$$

with $S_N = X_1 + X_2 + ... + X_N$ and $\mathcal{N}(0,1)$ the normal distribution with mean 0 and variance 1.

In the central limit theorem we need to scale the mean with size N, while the variance is scaled with $\sqrt{N}$ in order for the distribution to converge to a standard Gaussian.

2.1 Random Walk

The random walk as discussed in section 1.3 can be expanded to a continuous space and time random walk. This is done by decreasing the space between the lattice points. However one can imagine that if we take a very small distance between the lattice points we will not see anything macroscopically for short times. In order to see something which resembles Brownian motion we must also let the process run for a longer time, thus scale the time.
First we introduce a scaling parameter $\varepsilon > 0$ and we scale the mesh in the lattice with this parameter. We want ε to be small.
We scale our lattice with this parameter, so $x \to \varepsilon x$. It will now be shown that the scaling for the time needs to be $t \to \frac{t}{\varepsilon^2}$.

Theorem 2.2. *If L is the generator of the discrete space random walk (as seen before) and $x \to \varepsilon x$. If time is scaled according to $t \to \frac{t}{\varepsilon^2}$ the generator will become $\frac{1}{2}\frac{\partial^2 f}{\partial x^2}$, which is the generator of Brownian motion.*

Proof. As seen before we know that the generator for the random walk where both particles move with probability $\frac{1}{2}$ is:

$$Lf = \frac{1}{2}\left(f(x+1) + f(x-1) - 2f(x)\right)$$

If we scale x with ε we obtain:

$$Lf = \frac{1}{2}\left(f(x+\varepsilon) + f(x-\varepsilon) - 2f(x)\right)$$

Now lets turn to the Taylor expansion:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

If we expand this to the second term for $x = x - \varepsilon$ and $a = x$ and combine this with the fact that $f'(x) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon)-f(x)}{\varepsilon}$ we obtain:

$$\frac{f(x+\varepsilon) + f(x-\varepsilon) - 2f(x)}{\varepsilon^2} = \frac{1}{2}f''(x) + \mathcal{O}(\varepsilon^3) \tag{2.1}$$

Thus we see that:

$$\frac{1}{\varepsilon^2}Lf = \frac{1}{2}f''(x) + \mathcal{O}(\varepsilon^3) \tag{2.2}$$

The last term can be neglected for $\varepsilon \to 0$, so we find that if we scale time by $1/\varepsilon^2$ then we obtain the generator of Brownian motion. This implies that also for the corresponding processes it holds that $X_t^\varepsilon \to B_t$ in the sense of weak convergence of finite dimensional distribution by the theorem of Trotter Kurtz. □

2.2 Fourier-Laplace scaling

Further on in the paper we will want to use the same scaling, however we will want to scale the Fourier-Laplace transform of this equation. We will see that the variables in the Fourier-Laplace domain need to be scaled differently in order for the convergence to Brownian motion to hold.
The Fourier-Laplace transform is defined as follows:

$$\hat{p}(k,z) = \int_0^\infty \sum_{x \in \mathbb{Z}} p(x,t)e^{-ikx}e^{-zt}dt \tag{2.3}$$

We will see that the diffusive scaling corresponds to the scaling $k \to \varepsilon k$ and $z \to \varepsilon^2 z$ in Fourier-Laplace variables. Note that $p(x,t)$ in this case means the probability that the walker is on site $x \in \mathbb{Z}$ at time t. This is calculated for a given initial state, x_0. This could be written as $p(x,t) = \mathbb{P}(X_t = x | X_0 = x_0)$. If we use this, we can say that:

$$\sum_{x \in \mathbb{Z}} p(x,t)e^{-ikx} = \mathbb{E}_{x_o}(e^{-ikX_t}) \tag{2.4}$$

For the sake of this argument we take $x_0 = 0$ and we write $\mathbb{E}_0(e^{-ikX_t})$ as $\mathbb{E}(e^{-ikX_t})$. We then get:

$$\hat{p}_{x,t}(k,z) = \int_0^\infty \mathbb{E}(e^{-ikX_t})e^{-zt}dt \tag{2.5}$$

Now we want to introduce our scaling of $\tau = \frac{t}{\varepsilon^2}$ and $Y_\tau = \frac{X_{\varepsilon^2 \tau}}{\varepsilon}$:

$$\hat{p}_{y,\tau}(k,z) = \varepsilon^2 \int_0^\infty \mathbb{E}(e^{-ik\varepsilon Y_\tau})e^{-z\varepsilon^2\tau}d\tau = \varepsilon^2 \hat{p}_{x,t}(\varepsilon k, \varepsilon^2 z) \tag{2.6}$$

As can be seen, if we introduce the scaling $z \to \varepsilon^2 z$ and $k \to \varepsilon k$ and multiply this all by ε^2, the same result is achieved as when scaling x and t.

Theorem 2.3. *For Brownian motion with diffusion constant $\mathbb{D}$ ($B_{t\mathbb{D}}$) it holds that*

$$\int_{t=0}^\infty \mathbb{E}(e^{ikB_{t\mathbb{D}}})e^{-zt}dt = \frac{1}{z + \frac{1}{2}k^2\mathbb{D}} \tag{2.7}$$

Proof. We know that $B_{\mathbb{D}t} \sim \mathcal{N}(0, \mathbb{D}t)$. From this it follows that the Fourier-transform of $B_{\mathbb{D}t}$ is given by

$$\begin{aligned}\mathbb{E}(e^{ikB_{\mathbb{D}t}}) &= \int_{-\infty}^{\infty} e^{ikx} \frac{1}{\sqrt{2\pi\mathbb{D}t}} e^{-\frac{x^2}{2\mathbb{D}t}} dx \\ &= e^{\frac{-k^2\mathbb{D}t}{2}}\end{aligned}$$

Thus taking the Fourier-Laplace of the distribution $B_{\mathbb{D}t}$ gives

$$\int_0^{\infty} e^{\frac{-k^2\mathbb{D}t}{2}} e^{-zt} dt = \frac{1}{z + \frac{1}{2}k^2\mathbb{D}}$$

□

So if we now scale the Fourier-Laplace transform as $\hat{p}(k, z) \to \varepsilon^2 \hat{p}(\varepsilon k, \varepsilon^2 z)$, we will get the same result as having scaled the initial function with $x \to \varepsilon x$ and $t \to \varepsilon^{-2} t$. Thus using this scaling in the Fourier-Laplace domain we can determine the diffusion coefficient (or variance) of the initial process.

3. Large Deviations

This chapter is based on the work of Den Hollander [8].
Large deviations is a part of probability which deals with the possibility where a sum of variables deviates from its expectation by more than what is described by the central limit theorem. This is crucial for the study of Markov processes in order to predict how often a system deviates from the norm. Large deviations particularly deals with events where the average over the processes $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ differs from the norm $\mu = \mathbb{E}(X_1)$ by an amount of order n.
The process is restricted through the rate function, which is defined as follows

Definition 3.1. *The function $I : \chi \to [0, \infty]$, with χ a Polish space, is called a rate function if*

1. *$I :\neq \infty$*
2. *I is lower semi-continuous, meaning it is continuous from below*
3. *I has compact level sets*

We can now also define the large deviation principle and its rate function

Definition 3.2. *A sequence of probability measures (P_n) on χ is said to satisfy the large deviation principle (LDP) with rate n and rate function I if*

1. *I is a rate function*
2. $\limsup_{n\to\infty} \frac{1}{n} \log P_n(C) \leq -I(C) \quad \forall C \subset \xi$ *closed*
3. $\liminf_{n\to\infty} \frac{1}{n} \log P_n(O) \leq -I(O) \quad \forall O \subset \xi$ *open*

where $I(S) = \inf_{x\in S} I(x)$

One could think of the rate function as a function which describes how fast P_n goes to 0 if it deviated from its expectation. If P_n satisfies the LDP then its deviation from the expectation goes as $\exp\{-nI(x)\}$.

3.1 Example: Coin tossing

Let us look at an example using coin tossing first. This will give a sketch of the type of things we are looking at [8].

Theorem 3.1. *Let (X_i) be independently identically distributed (i.i.d.) random variables with $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 1) = \frac{1}{2}$. Define $S_n = \sum_{i=1}^{n} X_i$. Then for all $a > \frac{1}{2}$*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -I(a)$$

with

$$I(a) = \begin{cases} \log 2 + a \log a + (1-a)\log(1-a), & \text{if } a \in [0,1] \\ \infty, & \text{else} \end{cases}$$

Proof. For $a > 1$ we obviously have the the probability of all of the coins combined to give something larger than n is impossible.
For $a \in (\frac{1}{2}, 1]$ we observe that

$$\mathbb{P}(S_n \geq an) = \sum_{k \geq an} \mathbb{P}(S_n = k) = \sum_{k \geq an} \binom{n}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{n-k} = 2^{-n} \sum_{k \geq an} \binom{n}{k}$$

Thus we can find an upper and lower bound of $\mathbb{P}(S_n \geq an)$:

$$2^{-n} Q_n(a) \leq \mathbb{P}(S_n \geq an) \leq (n+1) 2^{-n} Q_n(a) \tag{3.1}$$

with $Q_n(a) = \max_{k \geq an} \binom{n}{k}$. Note that this maximum lies at the smallest integer $\geq an$. We can use Stirling's formula ($n! \xrightarrow{n \to \infty} n^n e^{-n} \sqrt{2\pi n}(1 + \mathcal{O}(n^{-1}))$) to derive that:

$$\lim_{n \to \infty} \frac{1}{n} \log Q_n(a) = -a \log a - (1-a) \log(1-a)$$

Note that this derivation does require some work. If we take $n \to \infty$ we see that the upper and lower bounds of $\mathbb{P}(S_n \geq an)$ converge. We can then take the logarithm, divide by n and take the limit $n \to \infty$ for all the statements in equation (3.1) to reach the desired conclusion. Note that this also holds for $\alpha \in [0, \frac{1}{2})$. □

What we can take away from this example is that if we toss a large number of coins, we will still be left with some sort of probability that those coins will not average out to be heads and tails an equal amount of times. We see that this probability decays to 0 in the manner of $\exp\{-nI(a)\}$.

3.2 Gärtner-Ellis theorem

In this work it will restrict ourselves to a single theorem, the Gärtner-Ellis theorem. The theorem has also been slightly simplified, but this is sufficient for what we are trying to achieve.
Before we state the theorem we must first define a couple of things.
For a given sequence (X_n) of random variables on a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ where $\mathcal{B}(\mathbb{R})$ is the Borel sigma algebra on $\mathbb{R}$. We define

$$F(\alpha) = \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}(e^{\alpha n X_n})$$

We also define the Legendre transform:

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

Finally we define the Legendre transform on a set S as:

$$I(S) = \inf_{x \in S} I(x) \quad \textit{for} \quad S \subset \mathbb{R}$$

In order to prove the Gärtner-Ellis theorem we must first prove a couple of lemma's. Let us start by proving the exponential Chebyshev uncertainty.

Lemma 3.2. *Let X be any random variable and real numbers $c > 0$,$\varepsilon \in \mathbb{R}$. Then*

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}(e^{cX})}{e^{c\varepsilon}}$$

Proof. We will prove this for any f that is non-negative and increasing. This then obviously shows that it holds for $f(X) = e^{cX}$ with $c > 0$. Note that:

$$\mathbb{E}(f(X)) = \int_{-\infty}^{\infty} f(x)dP(x) \geq \int_{\varepsilon}^{\infty} f(x)dP(x) \geq \int_{\varepsilon}^{\infty} f(\varepsilon)dP(x) = f(\varepsilon)\mathbb{P}(f(X) \geq \varepsilon)$$

From this it follows that $\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}(e^{cX})}{e^{c\varepsilon}}$. □

Secondly we must prove that $I(x)$ is strictly convex under certain conditions:

Lemma 3.3. *$I(x)$ is strictly convex if $F(\alpha) \in C^1(\mathbb{R})$ and $F(\alpha)$ is strictly convex*

Proof. We must prove that $I(x)$ is strictly convex. This is done by realizing that we can find $\sup_{\alpha\in\mathbb{R}}(x\alpha - F(\alpha))$ by calculating $\frac{d}{d\alpha}(x\alpha - F(\alpha)) = 0$. We find that $x = \frac{d}{d\alpha}F(\alpha)$, thus $\alpha(x) = \left[\frac{d}{d\alpha}F(\alpha)\right]^{-1}(x)$. From now on we will refer to this function as $\alpha(x)$. We can then define the Legendre transform in a different way using this $\alpha(x)$:

$$I(x) = x\alpha(x) - F(\alpha(x))$$

Now we find that

$$\begin{aligned}\frac{d}{dx}I(x) &= \alpha(x) + x\frac{d\alpha(x)}{dx} - \frac{dF(\alpha(x))}{dx} \\ &= \alpha(x) + x\frac{d\alpha(x)}{dx} - \frac{\partial F(\alpha)}{\partial\alpha}\frac{d\alpha(x)}{dx} \\ &= \alpha(x) + x\frac{d\alpha(x)}{dx} - x\frac{d\alpha(x)}{dx} = \alpha(x)\end{aligned}$$

We can now see that:

$$\begin{aligned}\frac{d^2}{dx^2}I(x) &= \frac{d\alpha(x)}{dx} \\ &= \frac{1}{\frac{d}{dx}\alpha^{-1}(x)} \\ &= \frac{1}{\frac{\partial^2}{\partial x^2}F(x)} > 0\end{aligned}$$

For the final inequality we use that $F(\alpha)$ is strictly convex. Thus $I(x)$ is strictly convex. □

Now we can formulate the Gärtner-Ellis theorem [8].

Theorem 3.4 (Gärtner-Ellis theorem)**.** *Let $P_n(\cdot) = \mathbb{P}(Z_n \in \cdot)$ and $F(\alpha)$ and $I(x)$ be defined as above. If $F(\alpha)$ is strictly convex and $F(\alpha) \in C^1(\mathbb{R})$ with $C^1(\mathbb{R})$ being the continuously differentiable functions on $\mathbb{R}$. Then:*

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

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

This theorem states that if we find $F(\alpha)$ then we can determine the rate function $I(x)$. It turns out that often $F(\alpha)$ is solvable.

Proof. We will start by proving part (a) of the theorem.
For this we will first prove the theorem for compact sets. Let $\delta > 0$ and define for $x \in \mathbb{R}$:

$$I_\delta(x) = \min\{I(x) - \delta, \frac{1}{\delta}\}$$

For every $x \in \mathbb{R}$ there exists a α_x such that:

$$x\alpha_x - F(\alpha_x) \geq I_\delta(x)$$

Also for every $x \in \mathbb{R}$ there exists a neighborhood A_x for x such that

$$\inf_{y\in A_x} (\alpha_x(y - x)) \geq -\delta$$

We now use the exponential Chebyshev inequality (lemma 3.2) with $c = \delta > 0$ to get the following inequality:

$$\begin{aligned} P_n(A_x) = \mathbb{P}(Z_n \in A_x) &\leq \mathbb{P}((Z_n - x)\alpha_x \geq -\delta) \\ &\leq e^{\delta n}\mathbb{E}(e^{n\alpha_x(Z_n - x)}) = e^{\delta n}\mathbb{E}(e^{n\alpha_x Z_n})e^{-nx\alpha_x} \end{aligned}$$

Now let $K \subset \mathbb{R}$ be a compact set. Then the cover $\cup_{x\in K} A_x$ of K has a finite subcover $\cup_{i=1,\dots,N} A_{x_i}$. Thus we can deduce that

$$\begin{aligned} \frac{1}{n} \log P_n(K) &\leq \frac{1}{n} \log[N \max_{i=1,\dots,N} P_n(A_{x_i})] \\ &\leq \frac{1}{n} \log N + \delta - \min_{i=1,\dots,N}[x_i\alpha_{x_i} - \frac{1}{n} \log \mathbb{E}(e^{n\alpha_x Z_n})] \end{aligned}$$

Let us now take $n \to \infty$, then we obtain

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log P_n(K) &\leq \delta - \min_{i=1,\dots,N}[x_i\alpha_{x_i} - F(\alpha)] \\ &\leq \delta - \min_{i=1,\dots,N} I_\delta(x_i) \leq \delta - I_\delta(K) \end{aligned}$$

Finally let $\delta \downarrow 0$

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(K) \leq -I(K)$$

Now lets extend this to only closed sets. If C is closed, then $C \cap [-N, N]$ is compact for all $N > 0$. Thus if we use our last result we get

$$\limsup_{n\to\infty} \frac{1}{n} \log_n(C) \leq \max\{-I(C \cap [-N, N], -M_N\}$$

with

$$-M_N = \limsup_{n\to\infty} \frac{1}{n} \log P_n(\mathbb{R}\backslash[-N, N])$$

We will now prove that $\lim_{N\to\infty} M_N = \infty$. This will prove the claim because $\lim_{N\to\infty} I(C \cap [-N, N]) = I(C)$
Note that there exists a $\delta > 0$ such that $F(-\delta) < \infty$ and $F(\delta) < \infty$ because $F(\alpha) \in C^1(\mathbb{R})$. Thus we find by the exponential Chebyshev inequality with parameter $n\delta > 0$ that

$$\mathbb{P}(Z_n \leq -N) \leq e^{-n\delta N}\mathbb{E}(e^{-nZ_n\delta})$$
$$\mathbb{P}(Z_n \geq N) \leq e^{-n\delta N}\mathbb{E}(e^{nZ_n\delta})$$

We find that

$$\begin{aligned} -M_N &= \limsup_{n\to\infty} \frac{1}{n} \log P_n(Z_n \notin [-N, N]) \\ &\leq -\delta N + \max\{F(\delta), F(-\delta)\} \end{aligned} \tag{3.2}$$

Thus we find that $\lim_{N\to\infty} M_N = \infty$, this proves part (a) of the theorem.
Let us move on to part (b) of the theorem.
Let $B_\varepsilon(x)$ be an open ball around x with radius ε, thus $B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$. It is enough to prove that

$$\lim_{\varepsilon\downarrow 0} \liminf_{n\to\infty} \frac{1}{n} \log P_n(B_\varepsilon(x)) \geq -I(x) \quad \forall x \in O$$

Because for any open set O we have that

$$P_n(O) \geq P_n(B_\varepsilon(x)) \qquad \forall x \in O, \forall \varepsilon \leq \inf_{y\in\mathbb{R}\backslash O} |x - y|$$

Thus the claim will follow when we that $n \to \infty$ and $\varepsilon \downarrow 0$.
Because $I(x)$ is strictly convex (by lemma 3.3) we know that for all $y = x$ there is an α^* for which it holds that $I(y) - I(x) > \alpha^*(y - x)$.We also know that because $F(\alpha) \in C^1(\mathbb{R})$ that $\mathbb{E}(e^{n\alpha^* Z_n}) < \infty$. We now define a different probability measure $\hat{P}_n$ as

$$\frac{d\hat{P}_n}{dP_n}(y) = \frac{1}{\mathbb{E}(e^{n\alpha^* Z_n})} e^{ny\alpha^*}$$

Now we compute

$$\begin{aligned} \frac{1}{n} \log P_n(B_\varepsilon(x)) &= \frac{1}{n} \log \int_{B_\varepsilon(x)} P_n(dy) \\ &= \frac{1}{n} \log \mathbb{E}(e^{n\alpha^* Z_n}) + \frac{1}{n} \log \int_{B_\varepsilon(x)} e^{-ny\alpha^*} \hat{P}_n(dy) \\ &\geq \frac{1}{n} \log \mathbb{E}(e^{n\alpha^* Z_n}) - x\alpha^* - \varepsilon|\alpha^*| + \frac{1}{n} \log \hat{P}_n(B_\varepsilon(x)) \end{aligned}$$

In the last equation we use that $|y-x| \leq \varepsilon$ for $y \in B_\varepsilon(x)$. Thus we find that

$$\lim_{\varepsilon \downarrow 0} \liminf_{n\to\infty} \frac{1}{n} \log P_n(B_\varepsilon(x)) \geq [F(\alpha^*) - x\alpha^*] + \lim_{\varepsilon \downarrow 0} \liminf_{n\to\infty} \frac{1}{n} \log \hat{P}_n(B_\varepsilon(x))$$

Note that the first part is $\geq -I(x)$, it remains to show that

$$\lim_{\varepsilon \downarrow 0} \liminf_{n\to\infty} \frac{1}{n} \log \hat{P}_n(B_\varepsilon(x)) = 0$$

In order to do this we define the moment generating function associated with $\hat{P}_n$, we call this $\hat{\varphi}_n(n\alpha)$. Note that

$$\hat{\varphi}_n(n\alpha) = \int e^{n\alpha y} d\hat{P}_n(y) = \int e^{n\alpha y} \frac{e^{ny\alpha^*}}{\mathbb{E}(e^{n\alpha^* Z_n})} dP_n(y) = \frac{\mathbb{E}(e^{n(\alpha^*+\alpha)Z_n})}{\mathbb{E}(e^{n\alpha^* Z_n})}$$

Thus we find that

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{\varphi}_n(n\alpha) = \hat{F}(\alpha) = F(\alpha + \alpha^*) - F(\alpha^*)$$

Note that $\hat{F}(\alpha)$ is also continuously differentiable on $\mathbb{R}$, just as $F(\alpha)$. Finally we find that

$$\hat{I}(x) = \sup_{\alpha \in \mathbb{R}} [x\alpha - \hat{F}(\alpha)] = I(x) - x\alpha^* + F(\alpha^*)$$

We can now apply part (a) of the theorem to $\hat{P}_n$ to obtain

$$\limsup_{n\to\infty} \frac{1}{n} \log \hat{P}_n(\mathbb{R} \backslash B_\varepsilon(x)) \leq -\hat{I}(\mathbb{R} \backslash B_\varepsilon(x))$$

Also because $\hat{I}(x)$ is continuous is contains a minimum on a compact set, thus

$$\exists x_0 \neq x : \hat{I}(\mathbb{R} \backslash B_\varepsilon(x)) = \hat{I}(x_0)$$

Finally because $I(x)$ is strictly convex we find that

$$\begin{aligned} \hat{I}(x_0) &= I(x_0) - x_0\alpha^* + F(\alpha^*) \\ &\geq [I(x_0) - x_0\alpha^*] - [I(x) - x\alpha^*] > 0 \end{aligned}$$

Hence we find that

$$\limsup_{n\to\infty} \frac{1}{n} \log \hat{P}_n(\mathbb{R} \backslash B_\varepsilon(x)) < 0 \qquad \forall \varepsilon > 0$$

Finally since $\hat{P}(\mathbb{R}) = 1$ for all n we know that

$$\lim_{\varepsilon \downarrow 0} \liminf_{n\to\infty} \frac{1}{n} \log \hat{P}_n(B_\varepsilon(x)) = 0$$

This completes the proof $\square$

Note that this is a simplified version of the Gärtner-Ellis theorem which is suitable for what we are trying to achieve. In the actual theorem x and α can be multidimensional. Also, the assumption that $F(\alpha)$ has to be strictly convex is not needed there. As it turns out $F(\alpha)$ is always convex (though not always strictly convex), so the limitation that $F(\alpha) \in C^1(\mathbb{R})$ is enough.

In section 6 we will apply this theorem. We will do this by taking $C = x$ and $O = B_\varepsilon(x)$. Now take $\varepsilon \to 0$, in that case the lim sup and lim inf will converge to a single limit and we can say that

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

Now we can again replace our discrete values n by the continuous value t if we introduce a rate one Poisson process to determine when X_n occurs. Thus we find that

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

We will also from now on refer to the above limit using a shorthand notation, thus we define

$$P(X_t \approx x) \approx e^{-tI(x)} \iff \lim_{\varepsilon \to 0} \lim_{t \to \infty} \frac{1}{t} \log \mathbb{P}(X_t \in B_\varepsilon(x)) = -I(x)$$

4. Active particle on $\mathbb{Z}$

An active particle has a position $x \in \mathbb{Z}$ and an internal degree of freedom $\sigma \in \{1, -1\}$. The direction of motion is determined by σ together with noise. The way we will model active particles is by having 3 potential movements for a particle on the lattice. It can either make a random movement to the left or the right with rate κ. It can switch from one lattice to another with rate γ. Finally it can move towards the right with rate λ if the particle is located on lattice 1 and it can move towards the left with rate λ if the particle is located on lattice -1. A visual representation of this model is given in figure 2.

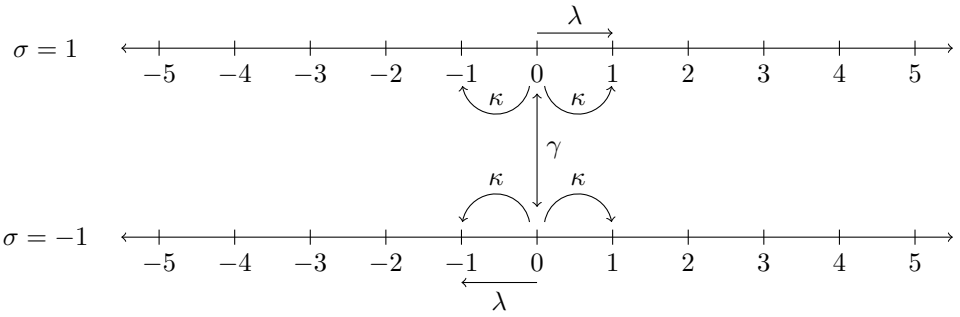

Figure 2: In this figure the possibilities the particle can move from either $(x, \sigma) = (0, 1)$ or $(x, \sigma) = (0, -1)$ are given with their rates. This extends to every site on the lattices.

Now fix a bounded $f : \mathbb{Z} \times \{-1, 1\} \to \mathbb{R}$, we can then formalize the above description via the generator L

$$Lf(x,\sigma) = \kappa[f(x+1,\sigma)+f(x-1,\sigma)-2f(x,\sigma)]+\lambda[f(x+\sigma,\sigma)-f(x,\sigma)]+\gamma[f(x,-\sigma)-f(x,\sigma)]$$

Now introduce a probability distribution $\mu = \binom{\mu_1}{\mu_{-1}}$ over our lattices. μ_1 and μ_{-1} are the distributions for the individual lattices $\sigma = 1$ and $\sigma = -1$ respectively. These are thus defined by $\mu_{\pm 1} = \mathbb{P}_\mu((X_t, \sigma_t = (x, \pm 1))$ with μ the initial distribution.
Using the definition of the generator we determine that:

$$\frac{d}{dt}\mu_1(x,t) = \kappa[\mu_1(x-1,t) + \mu_1(x+1,t)] + \lambda\mu_1(x-1,t) + \gamma\mu_{-1}(x,t) - 2\kappa\mu_1(x,t) - \lambda\mu_1(x,t) - \gamma\mu_1(x,t) \quad (4.1)$$

$$\frac{d}{dt}\mu_{-1}(x,t) = \kappa[\mu_{-1}(x-1,t) + \lambda\mu_{-1}(x+1,t) + \mu_{-1}(x+1,t)] + \gamma\mu_1(x,t) - 2\kappa\mu_{-1}(x,t) - \lambda\mu_{-1}(x,t) - \gamma\mu_{-1}(x,t) \quad (4.2)$$

Equations (4.1) and (4.2) combined is called the master equation or Kolmogorov forward equation of this system.
Ideally we would like to combine equations (4.1) and (4.2) to create a matrix such that

we can define the state of the process at time t given an initial state. However this is difficult for these two equations as the derivative with respect to t in the point (x,t) is dependent on the probability distribution in $x+1$ and $x-1$. This problem is however simplified when we take the Fourier-Laplace transform. This means that we define:

$$\hat{\mu}_{\pm 1}(k,z) := \int_0^\infty \sum_{x\in\mathbb{Z}} \mu_{\pm 1}(x,t)e^{-ikx}e^{-zt}dt \tag{4.3}$$

With $k \in \mathbb{R}$ and $z > 0$. This is a discrete Fourier transform in x followed by a continuous Laplace transform in t. We will also define only the discrete Fourier-transform:

$$\tilde{\mu}(k,t) := \sum_{x\in\mathbb{Z}} \mu_{\pm 1}(x,t)e^{-ikx}$$

This leads to the following theorem

Theorem 4.1. *For the process defined by equations* (4.1) *and* (4.2) *with initial distribution* $\mu = \alpha\delta_{x,0}\delta_{\sigma,1} + (1-\alpha)\delta_{x,0}\delta_{\sigma,-1}$ *for* $\alpha \in (0,1)$ *the following properties hold*

1. $\hat{\mu}(k,z) = \frac{i\lambda(2\alpha-1)\sin k - (2\kappa+\lambda)(\cos k-1)+2\gamma+z}{\lambda^2\sin^2 k + 2\gamma(z-(2\kappa+\lambda)(\cos k-1)) + (z-(2\kappa+\lambda)(\cos k-1))^2}$
2. $\hat{\mu}(\varepsilon k, \varepsilon^2 z) \xrightarrow{\varepsilon\to 0} \frac{1}{z+\frac{k^2}{2}(\frac{\lambda^2}{\gamma}+2\kappa+\lambda)}$

Where $\hat{\mu}$ *is the FL-transform as defined in* (4.3)*. As a consequence we find that for this process it holds that* $\varepsilon X_{\varepsilon^{-2}t} \to B_{\mathbb{D}t}$ *in the sense of weak convergence of finite dimensional distributions, with* $B_{\mathbb{D}t}$ *being Brownian motion with diffusion constant* $\mathbb{D} = \frac{\lambda^2}{\gamma} + 2\kappa + \lambda$

Proof. The FL-transformation allows us to get rid of the time derivative and space discrete derivatives in equations (4.1) and (4.2), thus turning a system of differential equations into a system of linear algebraic equations. Because of the properties of the Laplace transform we find that $z\hat{\mu}_{\pm1}(k,z) - \tilde{\mu}_{\pm1}(k,0) = \mathcal{L}\{\frac{d}{dt}\tilde{\mu}_{\pm1}\}(k,t)$. We can now rewrite equations (4.1) and (4.2) as follows

$$z\hat{\mu}_1 - \tilde{\mu}_1(k,0) = \lambda(e^{-ik}-1)\hat{\mu}_1 + 2\kappa(\cos k - 1)\hat{\mu}_1 + \gamma\hat{\mu}_{-1} - \gamma\hat{\mu}_1$$
$$z\hat{\mu}_{-1} - \tilde{\mu}_{-1}(k,0) = \lambda(e^{ik}-1)\hat{\mu}_{-1} + 2\kappa(\cos k - 1)\hat{\mu}_{-1} + \gamma\hat{\mu}_1 - \gamma\hat{\mu}_{-1}$$

Note that with $\mu_{\pm1}(k,t=0)$, the Fourier transform must first be executed followed by filling in $t=0$. The above equations can then be cast in matrix form, we call this matrix $\mathbf{M}(k)$

$$\begin{aligned}\mathbf{M}(k)\begin{bmatrix}\hat{\mu}_1\\ \hat{\mu}_{-1}\end{bmatrix} &= z\begin{bmatrix}\hat{\mu}_1\\ \hat{\mu}_{-1}\end{bmatrix} - \begin{bmatrix}\tilde{\mu}_1\\ \tilde{\mu}_{-1}\end{bmatrix}(k,0)\\ &= \begin{bmatrix}\lambda(e^{-ik}-1)+2\kappa(\cos k-1)-\gamma & \gamma\\ \gamma & \lambda(e^{ik}-1)+2\kappa(\cos k-1)-\gamma\end{bmatrix}\begin{bmatrix}\hat{\mu}_1\\ \hat{\mu}_{-1}\end{bmatrix}\end{aligned} \tag{4.4}$$

From (4.4) we solve

$$\begin{aligned}&\hat{\mu}(k,z) = (z\mathbf{I} - \mathbf{M}(k))^{-1}\tilde{\mu}(k,0)\\ &= \begin{bmatrix}z-\lambda(e^{-ik}-1)-2\kappa(\cos k-1)+\gamma & -\gamma\\ -\gamma & z-\lambda(e^{ik}-1)-2\kappa(\cos k-1)+\gamma\end{bmatrix}^{-1}\begin{bmatrix}\tilde{\mu_1}\\ \tilde{\mu_{-1}}\end{bmatrix}\end{aligned} \tag{4.5}$$

Now finding an inverse of a 2x2 matrix is easily done using Cramer's formula. It turns out that:

$$(z\mathbf{I}-\mathbf{M})^{-1} = \left(-\gamma^2+[\gamma+z-(2\kappa+\lambda)(\cos k-1)]^2+\lambda^2\sin^2 k\right)\cdot \begin{bmatrix} z-\lambda(e^{ik}-1)-2\kappa(\cos k-1)+\gamma & \gamma \\ \gamma & z-\lambda(e^{-ik}-1)-2\kappa(\cos k-1)+\gamma \end{bmatrix} \quad (4.6)$$

From here we can look at different initial distributions for μ and see what happens.
We first look at the simplest model, where the particles are evenly destributed among the lattices and start at $x=0$, thus $\mu(x,0)=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}\delta_{x,0}$. We are currently only interested in the phenomena of a single particle. The initial distribution in the Fourier domain is then $\tilde{\mu}(k,0)=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}\delta_{k,0}$. Now using equation (4.6) we find that

$$\hat{\mu}(k,z)=\frac{2\gamma+z-(\lambda+2\kappa)(\cos k-1)}{-\gamma^2+[\gamma+z-(2\kappa+\lambda)(\cos k-1)]^2+\lambda^2\sin^2 k}$$

In order to find the proper expression we need to scale these variables as in section 2.2. We scale z as $z\to\varepsilon^2 z$ and k as $k\to\varepsilon k$. In order to compensate for this we get an ε^2 in front of the whole equation, so it becomes:

$$\begin{aligned}\hat{\mu}(\varepsilon k,\varepsilon^2 z)=\varepsilon^2\frac{2\gamma+\varepsilon^2 z-(\lambda+2\kappa)(\cos(\varepsilon k)-1)}{-\gamma^2+[\gamma+\varepsilon^2 z-(2\kappa+\lambda)(\cos(\varepsilon k)-1)]^2+\lambda^2\sin^2(\varepsilon k)} \\ \xrightarrow{\varepsilon\to 0}\frac{1}{z+\frac{k^2}{2}(\frac{\lambda^2}{\gamma}+2\kappa+\lambda)} \quad (4.7)\end{aligned}$$

Recalling section 2.2, in particular equation (2.7) this proves the convergence to a Brownian Motion with diffusive coefficient $\mathbb{D}=\frac{\lambda^2}{\gamma}+2\kappa+\lambda$.
Now say we want an adjusted initial distribution, $\mu(x,0)=\begin{pmatrix}\alpha\\ 1-\alpha\end{pmatrix}\delta_{x,0}$, where $\alpha\in(0,1)$ and $\delta_{x,0}$ is the Dirac delta. We can then compute $\hat{\mu}(k,z)$ using $\tilde{\mu}(k,0)=\begin{pmatrix}\alpha\\ 1-\alpha\end{pmatrix}$. This then leads to the following result:

$$\hat{\mu}(k,z)=\frac{i\lambda(2\alpha-1)\sin k-(2\kappa+\lambda)(\cos k-1)+2\gamma+z}{\lambda^2\sin^2 k+2\gamma(z-(2\kappa+\lambda)(\cos k-1))+(z-(2\kappa+\lambda)(\cos k-1))^2}$$

When scaling this equation as done in equation (4.7) the same result is obtained for $\mathbb{D}$. □

We will end this chapter will a couple of remarks

Remark 4.2. Note that when taking $\gamma\to\infty$ the diffusion coefficient becomes $\mathbb{D}=2\kappa+\lambda$. This is to be expected because $\gamma\to\infty$ means that the particle switches lattices a lot, thus we can expect the particle to drift as much to the right as to the left on average. This is reflected with the extra λ term in the diffusion coefficient next to the already existing term for the diffusion.

Remark 4.3. Finally we note that the diffusion coefficient is independent of the initial distribution of σ. This means that if all the particles start on one lattice this will not have an effect on the diffusion coefficient.

5. Active particle on $\mathbb{R}$

We now take the same model into consideration, but make the time continuous. In this model we also scale our variables so that they become noticeable in the continuous time. For this we need to use the scaling $\lambda \to \varepsilon\lambda$ and $\gamma \to \varepsilon^2\gamma$. The generator then becomes:

$$\begin{aligned} L^{(\varepsilon)}f(x,\sigma) =& \lambda\varepsilon[f(x+\varepsilon\sigma,\sigma) - f(x,\sigma)] + \kappa[f(x+\varepsilon,\sigma) + f(x-\varepsilon,\sigma) - 2f(x,\sigma)] \\ &+ \varepsilon^2\gamma[f(x,-\sigma) - f(x,\sigma)] \end{aligned} \tag{5.1}$$

Using the Taylor expansion in order to reduce the equations to derivatives, equation (5.1) will reduce to the following:

$$\begin{aligned} &\varepsilon^{-2}L^{(\varepsilon)}f(x,\sigma) = \\ &\varepsilon^{-1}\lambda[f(x+\varepsilon\sigma,\sigma)-f(x,\sigma)]+\varepsilon^{-2}\kappa[f(x+\varepsilon,\sigma)+f(x-\varepsilon,\sigma)-2f(x,\sigma)]+\gamma[f(x,-\sigma)-f(x,\sigma)] \\ &\xrightarrow{\varepsilon\to 0} \sigma\lambda\frac{\partial f}{\partial x}(x,\sigma) + \frac{1}{2}\kappa\frac{\partial^2 f}{\partial x}(x,\sigma) + \gamma(f(x,-\sigma) - f(x,\sigma)) \end{aligned} \tag{5.2}$$

Now scale time as before, so $t \to \varepsilon^{-2}t$, this turns the $\varepsilon^{-2}L^{(\varepsilon)}$ into a derivative in time ($\frac{\partial}{\partial t}$). We see that we had to scale the parameters λ and γ in order for them to appear in the differential equation.

Now in chapter 1.2 we have found that $\frac{\partial}{\partial t}S_t f = LS_t f = L\mathbb{E}(f(X_t)|X_0 = x)$ and we want to translate this to an equation where we say something about a probability distribution which evolves in time. This is done using an initial probability measure $\mu(dx) = h(x)dx$. In this case $h(x)$ would represent our initial probability distribution. Now we want to look at our process, specifically what the probability measure of our process looks like after a time t. We define this as μ_t and define it as such:

$$\int S_t f d\mu = \int f d\mu_t$$

We now write $\mu_t = h(t,x)dx$, which is our new probability distribution at time t. We now find that

$$\int f d\mu_t = \int S_t f d\mu = \int S_t f(x)h(x)dx = \int f(x)S_t^* h(x)dx$$

In this equation S_t^* represents the Hermitian conjugate of the operator S_t. Thus we find that in order to develop a process from its initial distribution we need to use the Hermitian conjugate of S_t. This way we obtain the equation for the probability distribution in time. Thus we can determine the master equation for the probability distribution using the generator L through the following formula

$$\frac{\partial \rho_t(x,\sigma)}{\partial t} = L^*\rho_t(x,\sigma)$$

Where $\rho_t(x,\sigma)$ is the probability density a time t.

We will now show that for our operator it only changes the sign for our first derivative term.
Let us take a look at the operator $\frac{\partial}{\partial x}$. Define two functions $f(x), g(x) \in \mathcal{C}_0(\mathbb{R})$

$$\int_{\mathbb{R}} \left[\frac{\partial}{\partial x} f(x)\right] g(x)dx = f(x)g(x)\Big|_{x=-\infty}^{\infty} - \int_{\mathbb{R}} f(x)\frac{\partial}{\partial x} g(x)dx = \int_{\mathbb{R}} f(x)\cdot -\frac{\partial}{\partial x} g(x)dx$$

This last step is valid because both f and g vanish at infinity. Thus we see that $\left(\frac{\partial}{\partial x}\right)^* = -\frac{\partial}{\partial x}$. I will state without proof that both the second derivative and the constant in our S_t are Hermitian (thus their Hermitian conjugate and the operator itself are the same). This can easily be checked using the above method.

Let us return to the model at hand. We now define a function ρ, which represents the density probability of the particle. Define two probabilities, one for the $\sigma = 1$ particles, ρ^+ and another for the $\sigma = -1$ particles, ρ^-. The Kolmogorov forward or master equation corresponding to the generator (5.2) then reads

$$\begin{cases} \frac{\partial \rho^+}{\partial t} = -\lambda \frac{\partial \rho^+}{\partial x} + \frac{1}{2}\kappa \frac{\partial^2 \rho^+}{\partial x^2} + \gamma(\rho^- - \rho^+) \\ \frac{\partial \rho^-}{\partial t} = \lambda \frac{\partial \rho^-}{\partial x} + \frac{1}{2}\kappa \frac{\partial^2 \rho^-}{\partial x^2} + \gamma(\rho^+ - \rho^-) \end{cases} \tag{5.3}$$

Where we changed the sign for the first derivatives because we are now looking at probability densities.

6. Large Deviations

It turns out that the model described in chapter 4 satisfied the LDP and has a rate function. This is stated in the following theorem

Theorem 6.1. *For the process defined by equations* (4.1) *and* (4.2) *we find that*

$$F(\alpha) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}(e^{\alpha X_t}) = \left(\frac{1}{2}(e^{\alpha} + e^{-\alpha}) - 1\right)(2\kappa+\lambda) - \gamma + \sqrt{\frac{\lambda^2}{4}(e^{2\alpha} + e^{-2\alpha} - 2) + \gamma^2}$$

As a consequence $\frac{X_t}{t}$ *satisfies the LDP with rate function*

$$I(x) = \sup_{\alpha}(\alpha x - F(\alpha))$$

Proof. We now take an approach to determine the large deviations of the model. We do this using the Gärtner-Ellis theorem, from which we concluded that $\mathbb{P}(X_t \approx x) \approx e^{-tI(x)}$ (for large t) with $I(x) = \sup_\alpha(\alpha x - F(\alpha))$ and $F(\alpha) = \lim_{t\to\infty} \frac{1}{t}\log(\mathbb{E}(e^{\alpha X_t}))$. $\mathbb{E}(e^{\alpha X_t})$ can be determined by using the computation of $\mathbb{E}(e^{ikX_t})$ and substituting $\alpha = ik$. Notice that $\mathbb{E}(e^{ikX_t}) = \frac{1}{2}(\sum_x \mu_1(x,t)e^{ikx} + \sum_x \mu_{-1}(x,t)e^{ikx})$ is the Fourier transform of $\mu(x,t)$ with initial distribution $\mu_{\pm1}(x,0) = \frac{1}{2}\delta_{x,0}$. In chapter 4 we took the Fourier-Laplace transform, leading to expression (4.5). Now we only want to take the Fourier transform and keep time $t > 0$. We can however reuse our matrix $\mathbf{M}(k)$ as stated in equation (4.4). This would lead to the following expression in the Fourier domain:

$$\frac{d}{dt}\begin{bmatrix}\tilde{\mu}_1\\ \tilde{\mu}_{-1}\end{bmatrix}(k,t) = \mathbf{M}(k)\begin{bmatrix}\tilde{\mu}_1\\ \tilde{\mu}_{-1}\end{bmatrix}(k,t) \tag{6.1}$$

This leads to the solution:

$$\begin{bmatrix}\tilde{\mu}_1\\ \tilde{\mu}_{-1}\end{bmatrix}(k,t) = e^{t\mathbf{M}(k)}\begin{bmatrix}\tilde{\mu}_1\\ \tilde{\mu}_{-1}\end{bmatrix}(k,0) \tag{6.2}$$

Now we would like to determine $e^{t\mathbf{M}(k)}$. To do this we can diagonalize the matrix $\mathbf{M}(k) = PDP^{-1}$ with D a diagonal matrix. It can then be said that $e^{t\mathbf{M}(k)} = Pe^{tD}P^{-1}$ where D is a diagonal matrix so we can take the exponent of every argument on the diagonal to compute e^{tD}. A proof of this statement is included in Appendix A. Diagonalizing $\mathbf{M}(k)$ gives the following matrices:

$$P = \begin{bmatrix} -\lambda i\sin(k) + \sqrt{-\lambda^2\sin^2(k) + \gamma^2} & -\lambda i\sin(k) - \sqrt{-\lambda^2\sin^2(k) + \gamma^2} \\ \gamma & \gamma \end{bmatrix}$$

$$D = \begin{bmatrix} (\cos(k)-1)(2\kappa+\lambda) - \gamma + \sqrt{-\lambda^2\sin^2(k) + \gamma^2} & 0 \\ 0 & (\cos(k)-1)(2\kappa+\lambda) - \gamma - \sqrt{-\lambda^2\sin^2(k) + \gamma^2} \end{bmatrix}$$

$$P^{-1} = \frac{1}{2\gamma\sqrt{-\lambda^2\sin^2(k) + \gamma^2}} \begin{bmatrix} \gamma & \lambda i\sin(k) + \sqrt{-\lambda^2\sin^2(k) + \gamma^2} \\ -\gamma & -\lambda i\sin(k) + \sqrt{-\lambda^2\sin^2(k) + \gamma^2} \end{bmatrix} \tag{6.3}$$

In order to simplify the following equations we now define $A = (\cos(k) - 1)(2\kappa + \lambda) - \gamma$ and $B = \sqrt{-\lambda^2 \sin^2(k) + \gamma^2}$. This results in the following expression for $e^{t\mathbf{M}(k)}$:

$$e^{t\mathbf{M}(k)} = Pe^{tD}P^{-1} = \\ \frac{e^{tA}}{2\gamma B}\begin{bmatrix} -\gamma\lambda i\sin(k)(e^{tB} - e^{-tB}) + \gamma B(e^{tB} + e^{-tB}) & \gamma^2(e^{tB} - e^{-tB}) \\ \gamma^2(e^{tB} - e^{-tB}) & \gamma\lambda i\sin(k)(e^{tB} - e^{-tB}) + \gamma B(e^{tB} + e^{-tB}) \end{bmatrix} \tag{6.4}$$

Now to compute $\mathbb{E}(e^{ikX_t})$ we multiply $e^{t\mathbf{M}(k)}$ by $\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ and add the components together.

$$\mathbb{E}(e^{ikX_t}) = \frac{1}{2}e^{tA}(e^{tB} + e^{-tB}) + \frac{1}{2}\frac{\gamma e^{tA}}{B}(e^{tB} - e^{-tB}) \tag{6.5}$$

Now substituting in $\alpha = ik$ leaves us with:

$$\begin{aligned} \mathbb{E}(e^{\alpha X_t}) &= \frac{e^{tA}}{2}(e^{tB} + e^{-tB}) + \frac{\gamma e^{tA}}{2B}(e^{tB} - e^{-tB}) \\ A &= \left(\frac{1}{2}(e^{\alpha} + e^{-\alpha}) - 1\right)(2\kappa + \lambda) - \gamma \\ B &= \sqrt{\frac{\lambda^2}{4}(e^{2\alpha} + e^{-2\alpha} - 2) + \gamma^2} \end{aligned} \tag{6.6}$$

We can now finally compute $F(\alpha)$:

$$F(\alpha) = \lim_{t\to\infty}\frac{1}{t}\log(\mathbb{E}(e^{\alpha X_t})) = A + B = \\ \left(\frac{1}{2}(e^{\alpha} + e^{-\alpha}) - 1\right)(2\kappa + \lambda) - \gamma + \sqrt{\frac{\lambda^2}{4}(e^{2\alpha} + e^{-2\alpha} - 2) + \gamma^2} \tag{6.7}$$

In this derivation we used that $\lim_{t\to\infty}\log(A_t + B_t) = \log(\max_{t\to\infty}\{A_t, B_t\})$, the proof of this is also included in Appendix A. This concludes the proof.

□

Here are some remarks regarding the result:

Remark 6.2. We find that $F(\alpha) = F(-\alpha)$. This indicates that $I(x)$ is symmetric around 0.

Remark 6.3. In order to compute the supremum and find $I(x)$ we can take the derivative of $F(\alpha)$ and set this equal to x. If we can then calculate the inverse $F'^{-1}(x) = \alpha$ we have found the α for the supremum and thus we can calculate $I(x)$.

$$F'(\alpha) = \frac{1}{2}(2\kappa + \lambda)(e^{\alpha} - e^{-\alpha}) + \frac{\frac{\lambda^2}{4}(e^{2\alpha} - e^{-2\alpha})}{\sqrt{\frac{\lambda^2}{4}(e^{2\alpha} - e^{-2\alpha} - 2) + \gamma^2}} \tag{6.8}$$

From this $F'(\alpha)$ however it is very hard to find the inverse. We can plot $I(x)$

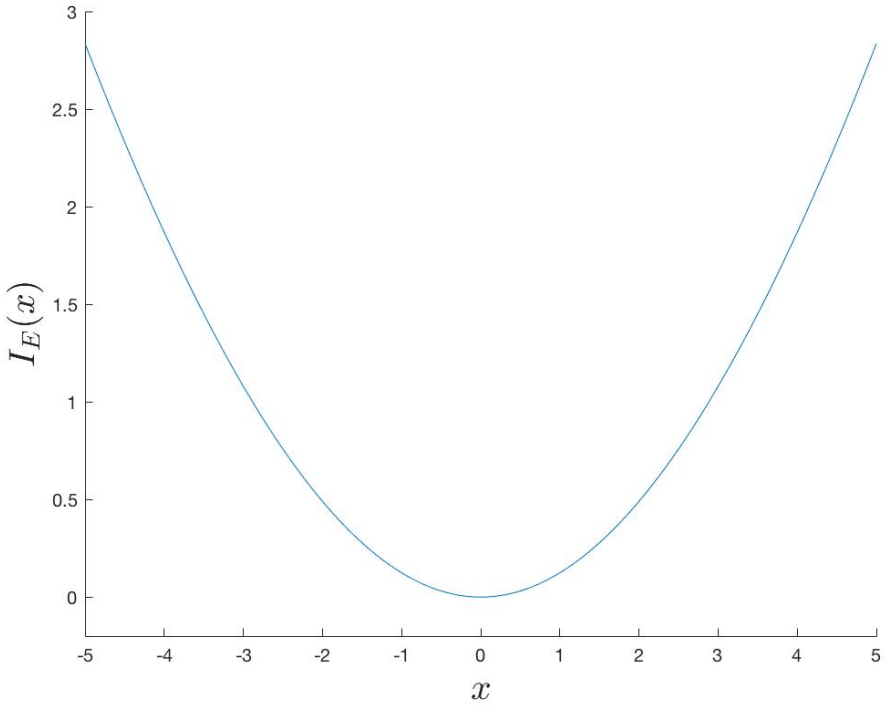

Figure 3: The Legendre transform of this model with $\gamma = 1$, $\kappa = 1$ and $\lambda = 1$

Remark 6.4. Now we check the limit $\gamma = 0$. This is where the particles stay on the same lattice. $F(\alpha)$ then becomes

$$F(\alpha) = \lambda(e^{\alpha} - 1) + \kappa(e^{\alpha} + e^{-\alpha} - 2)$$

If the particle stays on the same lattice it can be modeled as three independent Poisson processes. For these processes it holds that

$$\mathbb{E}(e^{\alpha X_t}) = \frac{1}{2}\left(e^{\kappa(e^{\alpha}-1)}e^{\kappa(e^{-\alpha}-1)}e^{\lambda(e^{\alpha}-1)}\right) + \frac{1}{2}\left(e^{\kappa(e^{\alpha}-1)}e^{\kappa(e^{-\alpha}-1)}e^{\lambda(e^{-\alpha}-1)}\right)$$

Thus the $\gamma = 0$ limit seems to coincide with what we expect

Remark 6.5. Finally we check the limit $\gamma \to \infty$. This is the case where particles flip lattices very often and the spin of the particle does not play a role. In this case

$$F(\alpha) = \left(\frac{1}{2}(e^{\alpha} + e^{-\alpha}) - 1\right)(2\kappa + \lambda)$$

This is equivalent to diffusion with diffusion constant $\kappa + \frac{\lambda}{2}$. This is also what we expect.

6.1 Asymmetric diffusion

We might also split the diffusion term up into two parts which are not equal to each other. This might correspond to an electric field E which causes the diffusion to be asymmetric.

$$\kappa_+ = \frac{2\kappa e^E}{e^E + e^{-E}}$$
$$\kappa_- = \frac{2\kappa e^{-E}}{e^E + e^{-E}}$$

This corresponds with non-symmetric diffusion when $E \neq 0$. For these parameters we can also define a generator L_E.

$$\begin{aligned} L_E f(x,\sigma) =& \kappa_+[f(x+1,\sigma) - f(x,\sigma)] + \kappa_-[f(x-1,\sigma) - f(x,\sigma)] \\ &+ \gamma[f(x,-\sigma) - f(x,\sigma)] + \lambda[f(x+\sigma,\sigma) - f(x,\sigma)] \end{aligned} \tag{6.9}$$

We will only be looking at small E, so we can approximate $e^E \approx 1 + E$. Now we can rewrite the following

$$\begin{aligned} &\kappa_+[f(x+1,\sigma) - f(x,\sigma)] + \kappa_-[f(x-1,\sigma) - f(x,\sigma)] \\ &= \frac{2\kappa}{e^E + e^{-E}}[e^E f(x+1,\sigma) - e^E f(x,\sigma) + e^{-E} f(x-1,\sigma) - e^{-E} f(x,\sigma)] \\ &\approx \frac{2\kappa}{1+E+1-E}[\{f(x+1,\sigma) + f(x-1,\sigma) - 2f(x,\sigma)\} + \\ &\qquad \{f(x+1,\sigma) - f(x-1,\sigma)\}] \end{aligned} \tag{6.10}$$

Now if we want to take the continuous space and time limit just as for the model with symmetric diffusion we find that:

$$\begin{cases} \frac{\partial \rho^+}{\partial t} = \kappa \frac{\partial^2 \rho^+}{\partial x^2} - 2E\kappa \frac{\partial \rho^+}{\partial x} - \lambda \frac{\partial \rho^+}{\partial x} + \gamma(\rho^- - \rho^+) \\ \frac{\partial \rho^-}{\partial t} = \kappa \frac{\partial^2 \rho^-}{\partial x^2} - 2E\kappa \frac{\partial \rho^-}{\partial x} + \lambda \frac{\partial \rho^-}{\partial x} + \gamma(\rho^+ - \rho^-) \end{cases} \tag{6.11}$$

Here we again use the scaling $\lambda \to \varepsilon\lambda$, $\gamma \to \varepsilon^2\gamma$ and $E \to \varepsilon E$. This last one is needed for the first order derivative dependant on E (similar to the λ scaling). Also note that we again use L^{E*}, identical to how we used L^* in equation (5.3). We can again pose the same theorem as at the start of this chapter.

Theorem 6.6. *For the process defined by equation* (6.11) *we find that*

$$F(\alpha) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}(e^{\alpha X_t}) = \alpha^2\kappa + 2\alpha\kappa E - \gamma + \sqrt{\lambda^2\alpha^2 + \gamma^2}$$

As a consequence $\frac{X_t}{t}$ *satisfies the LDP with rate function*

$$I(x) = \sup_\alpha(\alpha x - F(\alpha))$$

Proof. This proof is analogous to the proof of theorem 6.1. We can again take the Fourier-transform in x and write this as a matrix representation. We will denote this as $\mathcal{F}_x\{\rho(x,t)\} = \tilde{\rho}(q,t)$. We use q instead of k to indicate that this is a continuous Fourier transform in x.

$$\begin{aligned}\frac{\partial}{\partial t}\begin{bmatrix}\tilde{\rho}^+\\ \tilde{\rho}^-\end{bmatrix}(q,t) &= \mathbf{M}(q)\begin{bmatrix}\tilde{\rho}^+\\ \tilde{\rho}^-\end{bmatrix}\\ &= \begin{bmatrix}-q^2\kappa + 2iqE\kappa + iq\lambda - \gamma & \gamma\\ \gamma & -q^2\kappa + 2iqE\kappa - iq\lambda - \gamma\end{bmatrix}\begin{bmatrix}\tilde{\rho}^+\\ \tilde{\rho}^-\end{bmatrix} \quad (6.12)\end{aligned}$$

We can now take the same steps to find $e^{t\mathbf{M}(q)}$ by diagonalizing $\mathbf{M}(q)$ and then use $iq \to \alpha$ to find $F_E(\alpha)$ for this new situation. Doing this leads to the following results:

$$\begin{aligned}\mathbb{E}(e^{\alpha X_t}) = \frac{e^{t(\alpha^2\kappa + 2\alpha\kappa E - \gamma)}}{\sqrt{\lambda^2\alpha^2 + \gamma^2}}\Big[&\gamma\sinh(t\sqrt{\lambda^2\alpha^2+\gamma^2})\\ &+\sqrt{\lambda^2\alpha^2+\gamma^2}e^{-t\sqrt{\lambda^2\alpha^2+\gamma^2}} + \lambda\alpha e^{t\sqrt{\lambda^2\alpha^2+\gamma^2}}\Big] \quad (6.13)\end{aligned}$$

$$F_E(\alpha) = \lim_{t\to\infty}\frac{1}{t}\log(\mathbb{E}(e^{\alpha X_t})) = \alpha^2\kappa + 2\alpha\kappa E - \gamma + \sqrt{\lambda^2\alpha^2+\gamma^2} \quad (6.14)$$

□

We again have some remarks

Remark 6.7. Again finding the Legendre transform of $F_E(\alpha)$ is difficult to find, though we can plot $I_E(x)$.

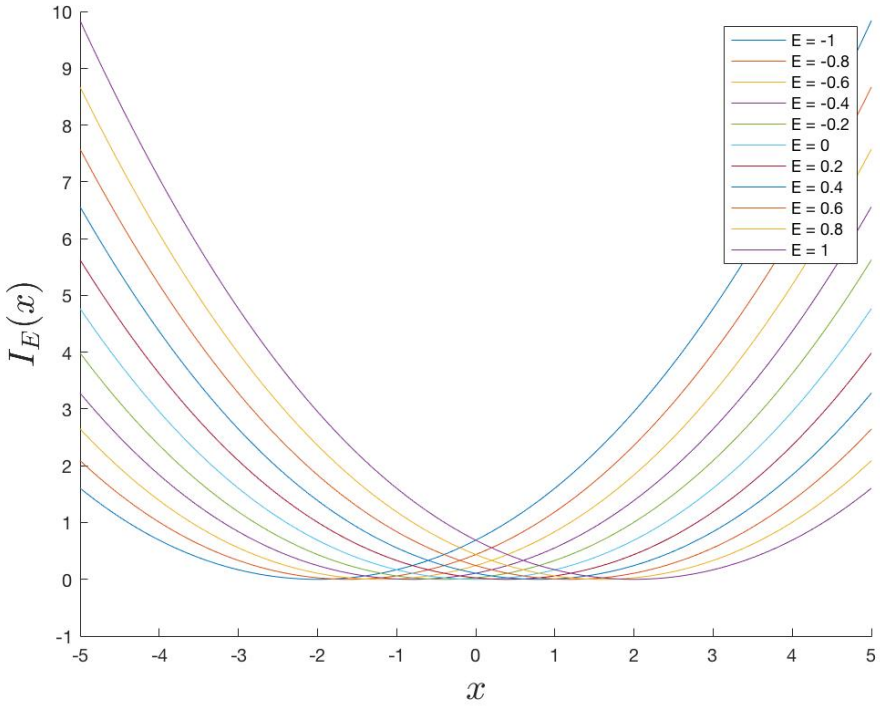

Figure 4: The Legendre transform of the non-symmetric model for different values of E with $\gamma = 1$, $\kappa = 1$ and $\lambda = 1$

What we can observe from this figure is that the minimum of $I_E(x)$ shifts as E does. This means that on average the particle will shift an amount of $\min_x I(x)$ per time step.

7. Electrons and Holes on a Ring

Now it would be nice if the model from chapter 4 describes a physical system. Since we have the internal characteristic related to a drift it seems natural to consider electrons, or electrons and holes. The fact that these particles have opposite charge suggests that they can easily be given a drift force which acts in opposite directions for the two species, using an electric field. There is however an extra condition which we need to fulfill in order for this to represent an electron-hole model, which is that the charge must be conserved. This means that if we replace a hole for an electron, then we must also replace an electron for a hole.

7.1 Electron-Hole recap

In order to understand this model we must first understand what electrons and holes are within a semiconductor. A quick recap will be given in this section. Let us start of by reviewing what a semiconductor is. A semiconductor is a solid for which the electrical conductivity lies between that of a conductor and a resistor. Whether a semiconductor conducts electricity is usually determined by its temperature, if the temperature of a semiconductor increases so will its conductivity.
A metal conducts electricity when it has electrons which can move freely. A semiconductor only has a few of these electrons, those which are excited to a state where they can move freely. This is only possible if this excitation requires only a small amount of energy, which is supplied through thermal excitation. The band where electrons can move freely is called the conduction band, while the band where electrons cannot move freely is called the valence band.
When an electron is excited from the valence band to the conduction band it leaves an empty state behind in the valence band. We call this empty state a hole. Holes have a positive charge. So we see that when an electron is excited it also creates a hole [9].
This is very similar to the model discussed previously because we now have two different particles with opposite charges. So if we introduce a magnetic field this means these particles will start to drift.

7.2 Physical model

In the model we have chosen the electrons can move around a ring, thus have periodic boundary conditions. This can be realised using an increasing magnetic flux such that there is a constant drift of electrons and holes in opposite directions. This is all done in a semiconductor such that there are only few excited electrons, in which case we can neglect interactions. Do note that because the charge needs to be conserved, the electron and hole can only annihilate if they both decide to jump to the other lattice. In that case a particle disappears from both populations. Likewise, the generation of an electron-hole pair implies the birth of a particle in both. To avoid an empty lattice a new electron-hole pair will be created randomly on the lattice when the electron-hole pair decide to

annihilate.
We will now see how an increasing magnetic flux will lead to an electric field. Let the particles move along a ring with radius R. These particles move in discrete space over a lattice of size $L \in \mathbb{N}$, thus the space between lattice sites is $\frac{2\pi R}{L}$. Letting a magnetic field pass through this ring imposes an electromotive force $\mathcal{E}$, given by

$$\mathcal{E} = -\frac{d\Phi_B}{dt}$$

with Φ_B being the magnetic flux. The magnetic flux through our ring is given by $\alpha R^2 \pi t$, in which α is determined by the strength of the magnetic field. From this we can find the electric field over the ring through $\vec{E} = \mathcal{E} 2\pi R$, which in turn is related to the drift of charged particles.

7.3 Computation Method

We will take into consideration a slightly different model than the one described in chapter 4. We will still have the parameters κ for the diffusion and λ for the drift. We introduce a new parameter $\hat{\gamma}$. $\hat{\gamma}$ is the rate at which both of the particles switch lattices. Thus in the electron-hole case, the rate at which the particles annihilate. We also assume that $\lambda + 2\kappa + \hat{\gamma} = 1$.
Now we want to find the diffusion constant for a process numerically. The 1D diffusion constant is determined through the following formula

$$\mathbb{D} = \frac{\langle r^2 \rangle}{2t}$$

With $\langle r^2 \rangle$ being the mean squared displacement and t the time. The mean squared displacement is the squared value of the displacement of a walker from its mean. On average a walker will be at position λt at time t because of the drift, this is thus the mean displacement of a walker.
The computation for determining the diffusion coefficient is done by placing two walkers on different lattices and letting them walk. When they annihilate a new set of particles is created. In the end the total displacement of all particles is added up and we find the diffusion coefficient using the following formula

$$\mathbb{D} = \frac{1}{2t} \frac{1}{N} \sum_{i=1}^{N} \left(x_i(t) \pm t\lambda \right)^2 \tag{7.1}$$

Where N is the total amount of experiments run and t will be the run time of each experiment. The $+$ is for electrons and the $-$ for holes. $x_i(t)$ is the sum of the displacements of all particles during experiment i. Each experiment runs for a time t.

7.4 Results

We calculate the diffusion coefficient for different values of λ and $\hat{\gamma}$. Note that κ is fixed because of the condition $\lambda + 2\kappa + \hat{\gamma} = 1$. The results for $5 \cdot 10^4$ walkers on a lattice of

size $L = 100$ for a time of $t = 10^4$ is given with their line of best fit. The line of best fit was determined using the curve fitting tool from Matlab.

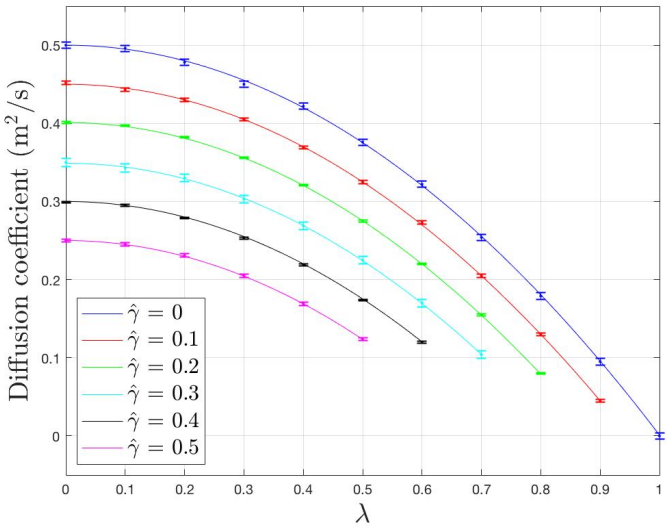

Figure 5: The diffusion coefficient and its error determined for different values of λ and γ. The data is presented in table 1 in appendix B.

The lines of best fit all have the form $\mathbb{D} = a\lambda^2 + b$, the values of a and b are presented in table 2 in appendix B. From this table conclude that for all $\hat{\gamma}$, a is a constant with value $a = -0.501 \pm 0.004$. If we can find how b is related to $\hat{\gamma}$ we can then derive an equation for $\mathbb{D}$.

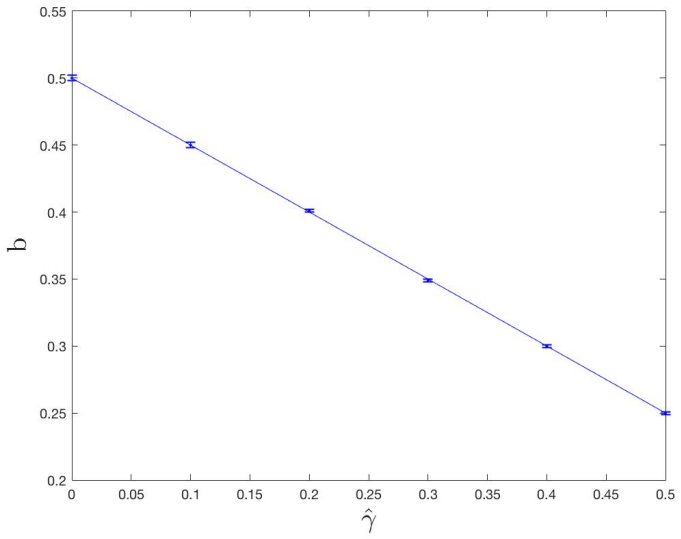

Figure 6: b for different values of $\hat{\gamma}$ and its curve of best fit. The data is presented in table 2 in appendix B

We find that $b = -0.50 \pm 0.005\hat{\gamma} + 0.50 \pm 0.001$ from the line of best fit shown in figure 6. Thus giving us the following equation for the diffusion coefficient

$$\mathbb{D} = -0.50\lambda^2 - 0.50\hat{\gamma} + 0.50 = \frac{1}{2}(1 - \lambda^2 - \hat{\gamma})$$

7.5 Discussion

When looking at the case of $\lambda = 0$, where there is only diffusion we find that $\mathbb{D} = \frac{1}{2}(1 - \hat{\gamma}) = \frac{1}{2}(2\kappa) = \kappa$. This is what we expect from a diffusion equation.
We can also check the case for $\hat{\gamma} = 0$, which results in $\mathbb{D} = \frac{1}{2}(1 - \lambda^2)$. This means that as λ increases the diffusion coefficient decreases. An easy way to picture this is by realizing that if the particle is moving due to the drift it is not moving due to diffusion. Thus the particle will diffuse less, resulting in a smaller diffusion coefficient. Also note that λ is squared, which ensures that the diffusion coefficient is independent of the direction of the drift.
Thus we conclude that the obtained result is in line with what is expected. We also find that all results have a decently small computation error, which again indicates that formula for $\mathbb{D}$ is correct.
In the future it might be interesting to research a model where the electron and hole can only annihilate if they are in close proximity. One could even increase the chance of annihilation according to how close the particles are to each other.

8. Bose-Hubbard Model

The Bose-Hubbard model was introduced to take the Coulomb interaction between bosons into account. It does this by adding a term to the Hamiltonian which increases the energy if more than two or more bosons are on the same site. The model can be used to describe the workings of a Mott insulator. The reason why we only take bosons into consideration is described at the end of section 9.2.

8.1 Bose-Hubbard model

The Hamiltonian for the Bose-Hubbard model is given by

$$H_{BH} = -t \sum_{<i,j>} b_j^\dagger b_i + \mu \sum_i n_i + U \sum_i n_i(n_i - 1) \tag{8.1}$$

Here b_i and $b_i^\dagger$ are the creation and annihilation operators and n_i is the number of particles at site i. The first term in this Hamiltonian corresponds to the hopping term, the second to the chemical potential and the third to the Hubbard term. The Hubbard term can be thought of as a penalty which particles have to pay in order for two or more of them to be on the same lattice site, thus the particles will want to avoid this.
The chemical potential term in this Hamiltonian is to control the amount of particles when considering a grand canonical ensemble. In our model we will fix the amount of particles, so the chemical potential will have no effect [10]. Thus we neglect this term in our final Hamiltonian, which becomes

$$H_{BH} = -t \sum_{<i,j>} b_j^\dagger b_i + U \sum_i n_i(n_i - 1) \tag{8.2}$$

The Bose-Hubbard model is a variation of the Hubbard model, in the Hubbard model fermions are considered instead of bosons. For more information on the Hubbard model I recommend reading chapter 23 of [9].

8.2 Mott Insulator

The Bose-Hubbard model works well for simulating a Mott insulator or a superfluid depending on the ratios t/U and μ/U as can be seen in the following phase diagram

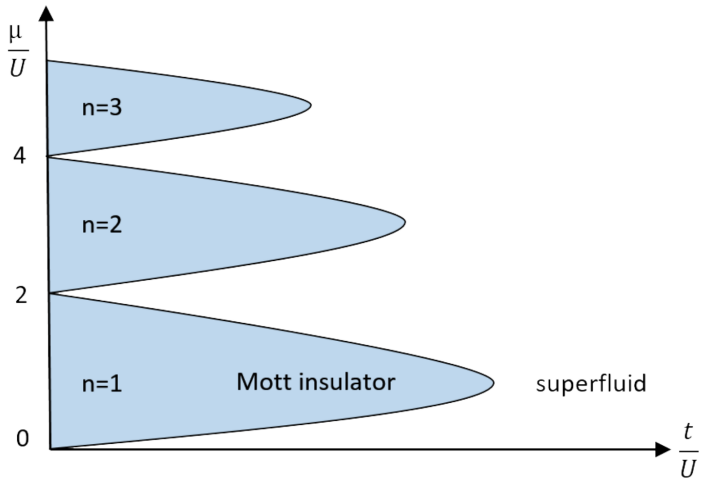

Figure 7: The phase diagram for the Bose-Hubbard model. n refers to the amount of bosons per site [11]

Since we neglect μ and consider a large U we will be looking at a Mott insulator. A Mott insulator is best imagined as a metal with a half-filled valence band. This means that on average every site contains a single electron. For a high enough Hubbard-U this will indeed be the case because there is a high penalty for two electrons to be on the same site. Therefore the electrons cannot move to neighbouring sites, thus making it an insulator. This situation is represented in the following figure.[9]

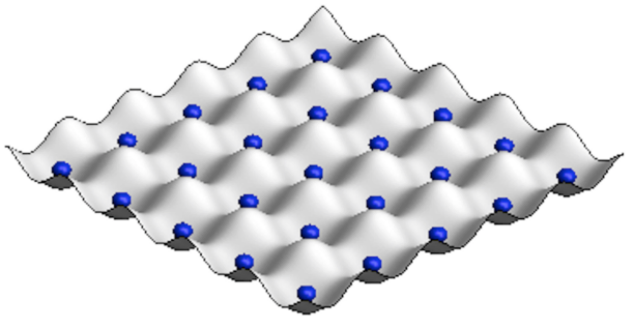

Figure 8: A diagram of what the final state looks like in a Mott insulator [12]

9. Bose-Hubbard Model Computations

Our goal now is to perform computations for the Bose-Hubbard model. We want to find the ground state energy for three particles which can diffuse with rate κ, drift with rate λ and can interact through the Bose-Hubbard interaction. Note that the direction of the drift is determined by the charge of the particle.

9.1 Hamiltonian

We first need to find the Hamiltonian for this model. We can find the Hamiltonian through the time dependent Schrödinger equation (in 1D)

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = \frac{-\hbar}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x,t)\psi(x,t) = H\psi(x,t)$$

We can now incorporate our drift term and the Bose-Hubbard Hamiltonian in continuous space. This way our Hamiltonian becomes

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) - \lambda\frac{\partial}{\partial x}\psi(x,t) + V_C\psi(x,t)$$

In this equation V_C is the potential from the Coulomb interaction.
Now the trick is to look at this system in imaginary time, which means that we take an imaginary variable $\tau = it$. Our system then becomes a PDE, for which we will solve lowest energy such that $H\psi = E\psi$. Simulating this PDE using diffusion Monte Carlo will give us its ground state energy.
In our final equation we will neglect all prefactors and we will scale accordingly later in order to simplify the explanation. Our equation then becomes

$$\frac{\partial}{\partial \tau}\psi(x,\tau) = \kappa\frac{\partial^2}{\partial x^2}\psi(x,\tau) + \lambda\frac{\partial}{\partial x}\psi(x,\tau) + V_C\psi(x,\tau) \tag{9.1}$$

Note that if we take $\lambda = 0$ then this equation is the continuous space equivalent equation of the Bose-Hubbard Hamiltonian in equation (8.2). This is because the hopping term in continuous space is equivalent to the diffusion term and the Hubbard term is equivalent to the potential due to Coulomb interaction between particles.
We can now solve for the ground state energy using the diffusion Monte Carlo method.

9.2 Diffusion Monte Carlo

This section is based on sections 12.3 and 12.4 of a book by J. M. Thijssen [13].
Diffusion Monte Carlo is a method where we take an initial density distribution and let this distribution develop using a PDE. This is done by placing walkers on a grid and letting these walkers diffuse. Eventually all the walkers will have diffused and we will be left with the ground state of the PDE.
Diffusion Monte Carlo uses Green's functions combined with a Monte Carlo simulation

to move the walkers. A Monte Carlo simulation in this case just means a simulation using random numbers. We want to use Green's functions because they are the result of a delta-peak forcing. This means that a Green's function describes what happens at site y when a delta-peak is introduced on a site x after a certain time $\Delta\tau$, the Green's function is denoted as $G(x, y; \Delta\tau)$.
We will use the Green's function as a probability distribution of where the particle can be after a time $\Delta\tau$. We will then use a random number to decide where this particle lands.
In our case we are dealing with a Fokker-Planck equation (to which we will add the Hubbard term in a moment):

$$\frac{\partial\rho(x,\tau)}{\partial\tau} = \kappa\frac{\partial}{\partial x}\left[\frac{\partial}{\partial x} - F(x)\right]\rho(x,\tau)$$

In the case where we only assume drift and diffusion we get $F(x) = -\lambda/\kappa$. The Green's function for such an equation is given by

$$G_{FP}(x, y; \Delta\tau) = \frac{1}{\sqrt{4\pi\kappa\Delta\tau}} e^{-[y-x+\lambda\Delta\tau/\kappa]^2/(4\kappa\Delta\tau)} \tag{9.2}$$

Thus we will use this equation together with a random number to determine where a particle starting at location x will end up after a short time $\Delta\tau$.

Now there is another aspect to diffusion Monte Carlo. This is the aspect of finding the ground state energy. The method for finding this ground state energy is based on the fact that when adding a potential to the Fokker-Planck equation the resulting Green's function will not be normalized. This is because (as we will see next) the walkers in this system follow a birth death process and thus the number of walkers is not preserved. Our new Green's function will look something like this:

$$G(x, y; \Delta\tau) = G_{FP}(x, y; \Delta\tau)e^{-\Delta\tau V(y)}$$

In this equation $V(y)$ is the potential of the new location of the walker. It is clear that this Green's function is not normalized if $V(y) \neq 0$.
We can normalize this function by introducing a new term E_T (which will be our ground state energy) and adding this to the potential of our original equation. This means that we multiply the Greens function with $e^{E_T\Delta\tau}$, where E_T is tuned to give the proper normalization. We then obtain a normalized Greens function. One could think of E_T as the constant which averages the potential around 0.

There is one final aspect to diffusion Monte Carlo, that is a birth and death process (also referred to as a branching process). This is a process where walkers with a high potential have the possibility of dying and walkers with a low potential have the possibility of giving birth to another walker. This way we get rid of all the walkers we do not want (because they have a high potential) and keep the walkers we do want because they are in a low energy state.
The way this is implemented is by evaluating $q = \exp[-\Delta\tau(V(y) - E_T)]$.
If $q < 1$ (equivalent to $V(y) > E_T$) then the walker survives with probability q and dies

with probability $1 - q$. This way if the potential of this walker is a lot higher than the current estimated ground state energy then the walker has a high probability of dying. Now if $q > 1$ then a walker gives birth to $[q]$ new walkers with probability $q - [q]$, where $[q]$ is the integer part of q. And with probability $1 - q + [q]$ the walker gives birth to $[q - 1]$ walkers. This ensures that as $V(y)$ is lower than E_T that there will be walkers spawned accordingly. This means that there will be many walkers in parts with a lower potential and less particles in parts with higher potentials.
Finally determining a new E_T after each process is then done using

$$E_T = E_0 + \alpha \ln\left(\frac{\tilde{M}}{M}\right)$$

In this equation E_0 is an estimate of the ground state energy, α is a small parameter, $\tilde{M}$ is a target number of walkers and M is the actual number of walkers in the simulation. These three concepts make up the simulation. For sake of clarity a pseudo-code is given [13].

```
REPEAT
    FOR all walkers DO
        Move walker to a new location y using equation (9.2)
        Evaluate q = exp{−Δτ(V(y) − E_T)}
        Kill or give birth to walkers as explained
    END FOR
    Update E_T
UNTIL finished
```

We will be considering a one dimensional system with three particles. In order to model this using diffusion Monte Carlo we take a walker walking on a three dimensional grid. This walker takes a step according to equation (9.2) in each direction, which is equivalent to all three walkers taking a step on a one dimensional lattice. If the walker lands on a site (i, j, k) where $i = j$, $i = k$ or $j = k$ it is equivalent to two particles being on the same site, thus a Hubbard-U term will be added to the potential.

As stated at the start of chapter 8, we only take bosons into consideration. This is because diffusion Monte Carlo cannot be performed when considering fermions. To see why one has to realize that fermions have anti symmetric wave functions. Meaning that for a wave function dependent on two fermions it holds that

$$\psi(x_1, x_2) = -\psi(x_2, x_1)$$

This in turn means that the wave function can be negative. In diffusion Monte Carlo we estimate the ground state wave function using probability densities of the particles. These can obviously not be negative, thus it cannot represent a fermion wave function. This is often referred to as the fermion problem.

9.3 Results

We program the procedure outline in the previous section for three particles with the same charge. The initial distribution for the simulation is constant. The parameters are the hopping parameter (κ), the drift parameter (λ) and the Hubbard-U (U). All models are run for a time of $\tau_{max} = 10^5$ with time step $\Delta\tau = 0.1$. The target number of walkers is $\tilde{M} = 10^4$ with $\alpha = 5 \cdot 10^{-5}$. The walkers walk along a lattice of size $L = 10$.
It turns out that if all three particles have the same charge then the ground state energy is constant for any value of κ, λ and U. This constant is $E_T = 5.94 \cdot 10^{-7} \pm 0.05 \cdot 10^{-7}$. All results are presented in the tables in appendix C.
The following plot shows the density of two dimensional walkers in the ground state.

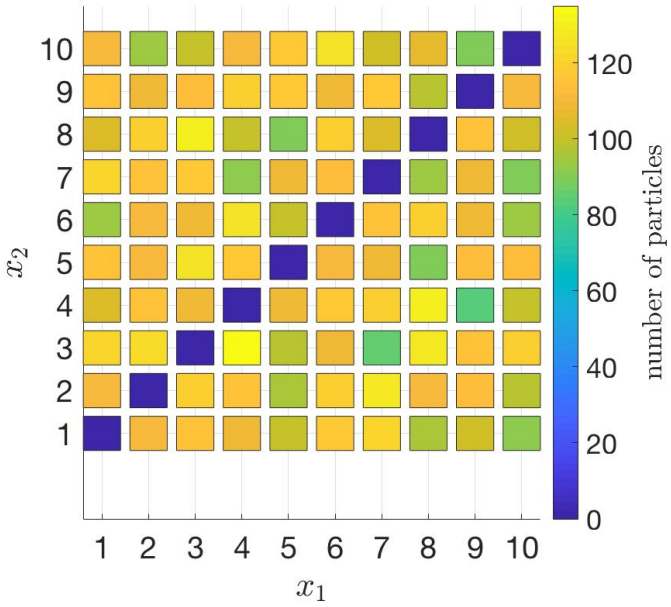

Figure 9: The particle density for walkers in the ground state on a 2D grid is shown. The parameters used are $\kappa = 0.5$, $\lambda = 1$ and $U = 10$. All walkers have the same charge.

As it turns out that if we change the charge for one of the three particles, the ground state energy still stays the same for all values of κ, λ and U.

9.4 Discussion

In figure 9 we see that none of the walkers are on the sites (i, i) with $i \in \{1, ..., 10\}$, these are the sites that represent the situation where two particles are on the same site. This is what we expect in the ground state for any value of U. We also see that the rest of

the sites have roughly the same amount of particles, this is to be expected because the particles are evenly distributed at the start and a particle can move around freely when it is not in range of another particle.
It is likely that this ground state energy is close to 0 because this is the potential of a single particle in the ground state. That the ground state energy is larger than 0 is also to be expected. In order for the ground state energy to be 0, the ground state wave function has to be constant. This is not the case because of the restriction that the particles cannot be in the same position. This restriction causes the ground state wave function to be zero at certain points, thus it is not constant. This means that the first and second derivative in space in equation (9.1) give a contribution. This contribution then leads to the increase of the ground state energy. This is in line with the ground state energy found in the computation.
However ground state energy is also independent of κ and λ. This is interesting, one might expect these parameters to influence the ground state energy because we expect this energy to come from the first and second derivative terms. This however is not the case according to our results. In further research it would be interesting to solve the ground state wave function for this model. This way we might discover why κ and λ do not have an influence on the ground state energy.

10. Conclusion

In conclusion we have used the Fourier-Laplace transform together with scaling of the parameters in order to determine the diffusion coefficient of the model described in chapter 4. In this model a particle can drift, with rate λ, the direction of this drift determined by an internal state. This particle can also diffuse with rate κ and change its internal state with rate γ. It turns out that the diffusion coefficient for such a model is $\mathbb{D} = \frac{\lambda^2}{\gamma} + 2\kappa + \lambda$.
Next we have used the Gärtner Ellis theorem to prove that this model satisfies the LDP and we have determined a function $F(\alpha)$ for which it holds that the Legendre transform of $F(\alpha)$ is the rate function of such a particle.
Then we have looked at this particle as a particle on a ring in a semiconductor moved by an electric field. A few of the properties changed in order for two of these particles to represent an electron-hole pair in a semiconductor. Using a simulation we have determined that the diffusion coefficient for an electron-hole pair is $\mathbb{D} = \frac{1}{2}(1 - \lambda^2 - \hat{\gamma})$. This is under the condition that $2\kappa + \lambda + \hat{\gamma} = 1$, where $\hat{\gamma}$ denotes the rate at which the electron-hole pair annihilates.
Finally we have looked at the Bose-Hubbard model and added a drift factor. We then determined a ground state energy for three particles on a ring using diffusion Monte Carlo. From this we found that the ground state energy is $E_T = 5.94 \cdot 10^{-7} \pm 0.05 \cdot 10^{-7}$ independent of the drift coefficient or the diffusion coefficient. Further research might clarify why these two parameters do not influence the ground state energy.

Acknowledgements

I would like to thank my advisors Frank Redig and Jos Thijssen for helping me come up with ideas and guiding through this project. They were both very open to answering all my questions and helped me understand difficult concepts.

The plotting in this report was done using Matlab [14], other computations were done using Fortran 95 [15]. For any questions regarding the code or other aspects regarding this thesis please email me at *b.l.vangisbergen@student.tudelft.nl*

Appendix A

Here is are the proofs of a couple of lemma's used in the text.

Lemma A.1. *Let A be a diagonizable matrix with diagonalization $A = PDP^{-1}$, then $e^{tA} = Pe^{tD}P^{-1}$.*

Proof. By induction we find that $(PDP^{-1})^n = PD^nP^{-1}$, now we find that

$$\begin{aligned} e^{tA} &= \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} (PDP^{-1})^n \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} PD^nP^{-1} \\ &= Pe^{tD}P^{-1} \end{aligned} \tag{A.1}$$

In the third line we use that $PP^{-1} = I$. □

Lemma A.2. *Let $A_t, B_t : \mathbb{R}^+ \to \mathbb{R}^+ \setminus \{0\}$ then*

$$\lim_{t\to\infty} \frac{1}{t} \log(A_t + B_t) = \lim_{t\to\infty} \frac{1}{t} \log(A_t \vee B_t)$$

Where the assume that the right hand side exists

Proof. Note that the following two properties hold

$$A_t + B_t = A_t \vee B_t + A_t \wedge B_t = (A_t \vee B_t)\left(1 + \frac{A_t \wedge B_t}{A_t \vee B_t}\right) \quad \leq \quad 2(A_t \vee B_t) \tag{A.2}$$

$$(A_t + B_t) \geq (A_t \vee B_t) \tag{A.3}$$

Now we consider $\frac{1}{t}\log(A_t + B_t)$. From (A.2) we get

$$\limsup_{t\to\infty} \frac{1}{t} \log(A_t + B_t) \leq \limsup_{t\to\infty} \frac{1}{t} \log(A_t \vee B_t)$$

From (A.3) we get

$$\liminf_{t\to\infty} \frac{1}{t} \log(A_t + B_t) \geq \limsup_{t\to\infty} \frac{1}{t} \log(A_t \vee B_t)$$

So if $\lim_{t\to\infty} \frac{1}{t}\log(A_t \vee B_t)$ exists is follows that

$$\limsup_{t\to\infty} \frac{1}{t} \log(A_t + B_t) = \liminf_{t\to\infty} \frac{1}{t} \log(A_t + B_t) = \lim_{t\to\infty} \frac{1}{t} \log(A_t \vee B_t)$$

□

Appendix B

$\gamma = 0$

λ	$\mathbb{D}$
0	0.500
0.1	0.496
0.2	0.478
0.3	0.450
0.4	0.422
0.5	0.376
0.6	0.322
0.7	0.254
0.8	0.179
0.9	0.095
1.0	0.000

$\gamma = 0.1$

λ	$\mathbb{D}$
0	0.452
0.1	0.443
0.2	0.430
0.3	0.405
0.4	0.369
0.5	0.325
0.6	0.273
0.7	0.205
0.8	0.130
0.9	0.045

$\gamma = 0.2$

λ	$\mathbb{D}$
0	0.401
0.1	0.397
0.2	0.382
0.3	0.356
0.4	0.321
0.5	0.275
0.6	0.220
0.7	0.155
0.8	0.080

$\gamma = 0.3$

λ	$\mathbb{D}$
0	0.350
0.1	0.343
0.2	0.330
0.3	0.303
0.4	0.269
0.5	0.225
0.6	0.170
0.7	0.104

$\gamma = 0.4$

λ	$\mathbb{D}$
0	0.299
0.1	0.295
0.2	0.279
0.3	0.253
0.4	0.219
0.5	0.174
0.6	0.124

$\gamma = 0.5$

λ	$\mathbb{D}$
0	0.250
0.1	0.245
0.2	0.231
0.3	0.205
0.4	0.169
0.5	0.124

Table 1: Tables showing the results of the diffusion constant ($\mathbb{D}$) for different values of γ and λ. All results have an uncertainty in $\mathbb{D}$ of ± 0.004 and are the results for a simulation of $5 \cdot 10^4$ particles walking for a time of $t = 10^4$. The uncertainties were calculated using 6 simulations

λ	a	b
0	-0.499 ± 0.001	0.500 ± 0.002
0.1	-0.500 ± 0.004	0.450 ± 0.002
0.2	-0.503 ± 0.002	0.401 ± 0.001
0.3	-0.499 ± 0.005	0.349 ± 0.001
0.4	-0.499 ± 0.005	0.300 ± 0.001
0.5	-0.506 ± 0.007	0.350 ± 0.001

Table 2: The results for drawing a line of best fit through the results from Table 1 for different values of γ. The line of best fit was drawn using the formula $\mathbb{D} = a \cdot \lambda^2 + b$. All lines of best fit have an R^2 value of 0.9999 or higher.

Appendix C

For all tables, the middle column represents the ground state energy for all three particles with the same charge. The right column represents the ground state energy for one of the particles with opposite charge.

κ	$E_{T,\uparrow\uparrow\uparrow}$	$E_{T,\uparrow\downarrow\uparrow}$
0.5	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
1	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
2	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
3	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
10	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$

Table 3: Results for the diffusion Monte Carlo simulation. All results end with $M = 9882$. The parameters used were $\lambda = 2$, $U = 10^4$.

L	$E_{T,\uparrow\uparrow\uparrow}$	$E_{T,\uparrow\downarrow\uparrow}$
10	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
20	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
40	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
100	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$

Table 4: Results for the diffusion Monte Carlo simulation. All results end with $M = 9882$. The parameters used were $\lambda = 2$, $\kappa = 1$.

λ	$E_{T,\uparrow\uparrow\uparrow}$	$E_{T,\uparrow\downarrow\uparrow}$
0	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
0.1	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
0.2	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
0.3	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
1	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
10	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$
50	$5.935 \cdot 10^{-7}$	$5.935 \cdot 10^{-7}$

Table 5: Results for the diffusion Monte Carlo simulation. All results end with $M = 9882$. The parameters used were $\kappa = 0.5$, $U = 10^4$.

Bibliography

[1] Étienne Fodor and M Cristina Marchetti. The statistical physics of active matter: From self-catalytic colloids to living cells. *Physica A: Statistical Mechanics and its Applications*, 504:106–120, 2018.

[2] Thibaut Demaerel and Christian Maes. Active processes in one dimension. *Physical Review E*, 97(3):032604, 2018.

[3] Giacomo Gradenigo and Satya N Majumdar. A first-order dynamical transition in the displacement distribution of a driven run-and-tumble particle. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(5):053206, 2019.

[4] Kanaya Malakar, V Jemseena, Anupam Kundu, K Vijay Kumar, Sanjib Sabhapandit, Satya N Majumdar, S Redner, and Abhishek Dhar. Steady state, relaxation and first-passage properties of a run-and-tumble particle in one-dimension. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(4):043215, 2018.

[5] Ioan Kosztin, Byron Faber, and Klaus Schulten. Introduction to the diffusion monte carlo method. *American Journal of Physics*, 64(5):633–644, 1996.

[6] Geoffrey Grimmett and Dominic Welsh. *Probability: an introduction*. Oxford University Press, 2014.

[7] Basic techniques in interacting particle systems.

[8] Frank Den Hollander. *Large deviations*, volume 14. American Mathematical Soc., 2008.

[9] Steven H Simon. *The Oxford solid state basics*. OUP Oxford, 2013.

[10] Jos Thijssen. Lecture notes statistical physics, Spring 2018.

[11] Luka Pavešic. Simulating the bose-hubbard model in optical lattices.

[12] Eric Duchon, Yen Lee Loh, and Nandini Trivedi. Optical lattice emulators: Bose and fermi hubbard models. *arXiv preprint arXiv:1311.0543*, 2013.

[13] Jos Thijssen. *Computational physics*. Cambridge university press, 2007.

[14] MATLAB. *version 9.5.0 (R2018b)*. The MathWorks Inc., Natick, Massachusetts, 2018.

[15] GNU Fortran. *version 8.1.0*. Free Software Foundation, Inc., 2018.