Delft University of Technology

Particle systems with sources and sinks

Submitted in fulfillment of the requirements
for the degree of Bachelor of Science in
Applied Mathematics and Applied Physics

Jochem Hoogendijk

Delft, Nederland
July 2019

Copyright © 2019 by Jochem Hoogendijk. All rights reserved.

BSc report Applied Mathematics and Applied Physics

Exclusion as a mathematical model for prototype biological systems

Jochem Hoogendijk

Technische Universiteit Delft

Supervisors

Prof. Dr. F. H. J. Redig

Dr. S.M. Depken

Other committee members

Dr. A. Jakobi

Prof Dr. Y. M. Blanter

Prof. Dr. M.C. Veraar

July, 2019

Delft

Abstract

This project aims to describe crude models that work towards a mathematical model for transcription and translation. Transcription refers to the conversion of genetic code to a readable format for the cell, while translation refers to the conversion of this code into polypeptide chains using ribosomes. The main proposal for such a model is the totally asymmetric exclusion process which has already been applied to the case of translation [1]. In this report two simpler models and its mathematical abstractions are examined. These two are independent random walks on a grid $\mathbb{Z}^d$ and the totally symmetric exclusion process on $\mathbb{Z}^d$ where we allow a maximum of one particle per site. Since transcription and translation deal with particles being added and removed, we also consider the two models under which a source is placed in the origin for particles.

Starting with independent random walkers, we find a principle called duality. This says that we can describe the independent random walks of a whole configuration by letting a single particle execute a random walk starting from x and evaluate the random walk at a later time t. Since precise quantitative behaviour is rather difficult to extract without further assumptions, we instead turn towards invariant distributions. It is proven that the Poisson measure is invariant on $\mathbb{Z}^d$, with and without a source in the origin for independent random walkers.

The symmetric exclusion process is then constructed in which a source at the origin, sink at the origin and a general case with arbitrary sources and sinks on the grid $\mathbb{Z}^d$ are considered. We find that under a recurrent random walk, the grid fills up for the first case and empties for the second. In the third case, we examine the situation of having both a source and sink in the origin. In this situation, we find the rather surprising result that the grid still fills up in the recurrent case.

Further indication for the construction of the TASEP model for transcription and translation are roughly described and might be touched upon in a future project.

Contents

Chapter 1

Introduction

Approximately 50 years ago, the discovery of DNA was announced by James Watson and Francis Crick, laying the foundations for many great advances made in cellular biology in the past 50 years[2]. These advances include, but are not limited to, genetic engineering, protein synthesis and forensic analysis. Especially protein synthesis is a fascinating process in itself, which is still not completely well understood. Current ambitious efforts include the synthesis of an artificial cell. Applications of such a pursuit cannot be underestimated. If one were able to understand and alter arbitrary parameters in a cell, one could potentially turn cells into little factories creating whatever protein or more complex molecule we're interested in. The impact of such a development cannot be underestimated. For example, diabetes patients would not have to suffer from weekly visits to the hospital to get insulin injections. In this case, the newly made cells could be placed inside some device that will regulate production of these specific cells. This device can then placed in or outside the body to supply these artificially made cells. This will also lead to a decrease in production costs for medicines, which is crucial in a world where medicine prices are being raised every day. The benefits itself are worth a complete essay. The interested reader is referred to [3] for more information on possible applications.

Two of the most important processes in the eukaryotic cell are transcription and translation as part of the central dogma of the cell. These processes govern the synthesis of proteins based on the genetic code. Even though the qualitative steps of transcription and translation are well understood, the quantifiable effects are not. That is, there is no good model or mechanism to quantify the rate of protein production in the cell. The original main motivation for this project was to build a model that can properly quantify, predict and explain the rate of protein synthesis in a eukaryotic cell. Experimental observations show that after a short delay called is the initiation time, the production is roughly proportional to $\sim t^2$, reaching a steady state somewhat later [4]. A good model outputs an initiation time, a t^2 relation and prediction for when steady state is reached based on relevant parameters in the cell. Such parameters could be, but are not limited to, the length of the mRNA, density of ribosomes in the cell or amount of mRNA strands present or being synthesized. First attempts to reach such a model have already been made based on a mathematical technique called the totally asymmetric exclusion process (TASEP). Using TASEP, one can describe the rate of production related to translation [1]. However, since translation is intimately related to transcription, a complete model

that outputs our desired quantities cannot be given, yet.

The reason TASEP is a very good candidate model for transcription and translation is because of the following reasons. Both take fixed steps in only one direction, have defined start and stop sites, and since we work on the cellular level, we expect stochastic dynamics to occur due to thermal energies competing with determinstic ones. Now, since RNA polymerases move in one direction over the DNA and ribosomes move on the mRNA strings created during the process by the RNA polymerases, we find that the system might be described by two coupled TASEPs. Unfortunately, this effort turned out to be rather complicated on its way. The complications are described at the end of the results section (Chapter 4). Though, motivated by transcription and translation, different questions about the systems can be posed in a mathematical environment.

In the end this resulted in questions about different, but related models. The two models studied in this project are the independent random walker (IRW) model and symmetric exclusion. These models are then complicated by adding a source at the origin that emits particles at random, mimicking movement RNA polymerase or the arrival of ribosomes in translation. The main quantities of interest are the evolution of particles that all execute independent random walks, invariant distributions when starting with a certain distribution and whether grids overflow under the restrictions of symmetric exclusion.

We first give background and a mathematical introduction to the tools used in the next chapter. Chapter 3 will completely describe the IRW and exclusion models including a description for a source added at the origin. The last chapter will provide proofs for most of the claims made to outline the models given in the third. We assume that the reader has taken at least a basic course in probability, similar to the course TW1080 Introduction to probability taught at the Delft University of Technology. Knowledge about operators, real analysis and some stochastic processes is also helpful. While not a necessity, for a more formal background in probability we refer to the amazing book 'Probability with martingales' by David Williams. Most techniques used in this project are mostly based on two sets of lecture notes [5] [6].

Chapter 2

Background and preliminary theory

This chapter touches upon the main definitions and theorems that will be used in the models that will be build later on. Also, a description of the motivated biological system, transcription and translation, is given.

2.1 Markov processes

The corner stone property of many stochastic processes is the Markov property, which roughly says that the future state only depends on the current state, so not on its past. Since the definition of a Markov process can be incredibly hard depending on your probability space, we will first define the Markov process for countable state space and finite time.

Definition 2.1.1. (Markov Process) Let $\{X_n : n \in \mathbb{N}\}$ be a collection of random variables with countable state space E. We denote this collection by (X_n). We call (X_n) a Markov process if and only if

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, ..., X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n) \tag{2.1}$$

for all $n \in \mathbb{N}$ and all $i_0, i_1, ..., i_{n+1} \in E$. A Markov process is called homogeneous if for all $i, j \in E$, the conditional probability $\mathbb{P}(X_{n+1} = j | X_n = i)$ does not depend on n.

A Markov process in discrete time and on a countable state space is also often called a *Markov chain*. We see from the definition indeed that the future state only depends on the current state and not on its past for a Markov chain. The last probability in the homogeneous case is often also denoted as a transition probability $p(x, y)$, with $p(x, y) = \mathbb{P}(X_{n+1} = y | X_n = x)$. If we work on a finite state space, we can group these transition probabilities in a matrix P called a stochastic matrix. The rows of such a matrix need to add to 1. More generally, we find that for a countable state space E, $\sum_{y \in E} p(x, y) = 1$. Typical examples of Markov processes are random walks, branching processes and the Poisson process.

We now move towards the more complicated case of a continuous time (CT) Markov process, but still with countable state space. Intuitively, such a Markov process has to

move in jumps because the state space is countable. Also, the jump probabilities should only depend on the current location by the Markov property and not on the time already spent at x. So, we can regard a homogeneous continuous Markov process on a countable state space E to consist of two steps. We first spend a random amount of time in state x and then move to another state y dictated by the transition probability $p(x, y)$.

For the first step, one can derive that the time spent at x must exponentially distributed, since this is the only memoryless distribution. That is, for a random variable T, the exponential distribution is the only distribution that can satisfy $\mathbb{P}(T > s + t|T > s) = \mathbb{P}(T > t)$. We let $1/c(x)$ be the mean of the involved exponential distribution for the CT Markov process. Here, $c(x)$ is called the *jump rate* from position x. For the second step, the stochastic matrix P (possibly infinite) dictates the jump probability $p(x, y)$ depending on the current location x. We will not go further into a rigorous construction of a CT Markov process but hope to have shown the bare bones involved. For a formal construction, we refer to [6].

2.2 Semigroups and generators

A popular technique associated with Markov processes is semigroup analysis. While semigroups were originally used in the theory of ordinary differential equations (ODEs), they were found to be more useful for Markov processes. This section will give a heuristic approach to them. All properties used can be proven rigorously and we refer the interested reader to two sources [7] [8].

2.2.1 Motivation for semigroups

As a motivation, we first look at the concept of semigroups in ODes, from which it originated. Consider the first order ODE X_t^x on $\mathbb{R}^d$ given by:

$$\frac{dX_t}{dt} = g(X_t) \tag{2.2}$$

with g a continuous map $g : \mathbb{R}^d \to \mathbb{R}^d$ and the superscript x denoting the initial condition $X_0 = x$. To help solve this differential equation, we introduce an operator T_t that does the following on a map $f : \mathbb{R}^d \to \mathbb{R}$: $(T_t f(x)) = f(X_t^x)$. This operator essentially shifts time from 0 to t and then evaluates $f(X_t^x)$. This is called a *flow*. An evident property is $T_{t+s} = T_t T_s$, which is called the semigroup property.

Now, applying the time derivative to $T_t f(x)$ yields:

$$\begin{aligned}\frac{d}{dt}T_t f(x) &= \frac{d}{dt}f(X_t^x) \\ &= \nabla f \cdot \frac{dX_t^x}{dt} \\ &= \nabla f(X_t^x) \cdot g(X_t^x) \\ &= \mathcal{L}f(X_t^x) \\ &= \mathcal{L}T_t f(x)\end{aligned} \tag{2.3}$$

where $\mathcal{L}$ is defined to be $\mathcal{L} = g \cdot \nabla$. So, we are left with the following ODE:

$$\frac{d}{dt}T_t f(x) = \mathcal{L}T_t f(x) \tag{2.4}$$

It is immediately clear that the solution of this equation should be $T_t f = e^{t\mathcal{L}} f$. Here, $\mathcal{L}$ is called the generator of the ODE. Of course, it is not a priori obvious what it means to take the exponent of an operator. In the case of a bounded operator $\mathcal{L}$, we can take the infinite series representation of $e^{t\mathcal{L}}$. However, since differential operators are often unbounded, this does not have to converge necessarily and it should be defined differently.

2.2.2 Semigroups in the Markov world

It turns out the technique of flows can also be exploited for Markov processes. We introduce the operator $S(t)$ together with a corresponding Markov process $(X_t)_{t\geq 0}$. The operator $S(t)$ is defined by $S(t)f(x) = \mathbb{E}_x[f(X_t)]$, where the subscript x denotes that the process starts from x. Using the Markov property, we can derive the following property:

$$\begin{aligned}S(t+s)f(x) &= \mathbb{E}_x[f(X_{t+s})] = \mathbb{E}[\mathbb{E}_x[f(X_{t+s})|\mathcal{F}_s] \\ &= \mathbb{E}_x[\mathbb{E}_{X_s}[f(X_t)] = \mathbb{E}_x[S(t)f(X_s)] = S(t)S(s)f(x)\end{aligned} \tag{2.5}$$

Here, we have used the definition of the operator $S(t)$ and tower property repeatedly. This motivates the following definition:

Definition 2.2.1. *Semigroup* Let $\{S(t) : t \geq 0\}$ be a collection of operators that work on the family of continuous function $f : \mathbb{R}^d \to \mathbb{R}$. If this collection satisfies the property $S(t+s) = S(t)S(s)$, we call $\{S(t) : t \geq 0\}$ a semigroup.

Specifically, every collection of operators associated with a Markov process, that is the random variable X_t associated with $S(t)f(x) = \mathbb{E}_x[f(X_t)]$ is part of Markov process, is also a semigroup as we have shown. In simpler Markov processes (finite state-space and discrete time), one can derive the Chapman-Kolmogorov equations, which is similar to the semigroup property. This says that for stochastic matrices P_s, P_t and time $s, t \in \mathbb{R}_{\geq 0}$, $P_{t+s} = P_t P_s$. These help to show the following relevant semigroup properties:

Proposition 1. Let $\{S(t) : t \geq 0\}$ be a semigroup. The following properties hold:

1. $S_0 f = f$ for all f
2. the map $t \to S_t f$ is right-continuous
3. (Semigroup property) $S_{t+s} f = S_t S_s f = S_s S_t f$ for all f
4. (Positivity) $f \geq 0$ implies $S_t f \geq 0$ for all t
5. (Normalization) $S_t 1 = 1$

To intuitively connect the generator and the semigroup, looking at the semigroup property, we see it is very similar to the functional equation $f(x+y) = f(x)f(y)$, which has solution $f(x) = e^{\mu x}$ for some $\mu \in \mathbb{R}$. Similarly, for the equation $S(t+s) = S(t)S(s)$, we know that there exists an operator $\mathcal{L}$ such that $S(t) = e^{t\mathcal{L}}$. We see that this operator satisfies the semigroup property. In the case that $S(t)$ is a bounded operator, recall that we were able to use the infinite series representation to define $S(t)$. That is, $S(t) = e^{t\mathcal{L}} = I + t\mathcal{L} + \mathcal{O}(t^2)$, where I is the identity operator. So, we can do the following to obtain an expression for $\mathcal{L}$.

$$\mathcal{L} f = \lim_{t \to 0} \frac{S_t f - f}{t} \tag{2.6}$$

In certain cases, it is indeed possible to explicitly calculate the generator $\mathcal{L}$ of the semigroup S. In a further section, this will be done for the random walk, which will be very important to our models.

2.3 The Feynman-Kac theorem

One of the major ways we will analyze and obtain results form our models is by using the Feynman-Kac theorem. This theorem allows us to analyze the effect of the semigroup when there is some perturbative, not necessarily Markovian, effect on the Markovian generator of a process $(X_t)_{t \geq 0}$. The reason this is useful is because we will examine situations where sources act as perturbations on the standard evolution given by the models we will define next chapter.

Theorem 1 (Feynman-Kac)**.** Let A_0 be a Markov generator of the process $\{X_t : t \geq 0\}$ on the state space Ω. Also, let $V : \Omega \to \mathbb{R}$ and define the semigroup $\tilde{S}(t) = e^{t(A_0+V)}$ which solves $\frac{d}{dt}\tilde{S}(t)f(x) = (A_0 + V)\tilde{S}(t)f(x) = \tilde{S}(t)(A_0 + V)f(x)$. Then,

$$\tilde{S}(t)f(x) = \mathbb{E}_x[e^{\int_0^t V(X_s)ds} f(X_t)] \tag{2.7}$$

The last formula is called the Feynman-Kac formula.[1] We will not prove this theorem, since it is outside the scope of this project [8]. However, it can be conveniently employed in our exclusion problem and we will use it repeatedly.

[1]The Feynman-Kac formula was a surprising collaboration. When Mark Kac attended a seminar by Richard Feynman, he noticed they were working on the same problem with very different approaches. Together they managed to prove this theorem which proves the real case of the famous Feynman path integrals in Quantum field theory (QFT). QFT still has a lot of questionable steps that need a rigorous framework.

2.4 Transcription and translation

We will provide a rough description of both transcription and translation as a recap based on [1] and [9]. Also, see figure 2.1. Consider a DNA strand in a eukaryotic cell. We will first consider transcription. Transcription refers to the process in which the genetic code of the DNA is transferred to another physical substrate. In this case, an mRNA string is made based on the DNA, coded by the 4 nucleic acids A, U, C and G. Biomakers bind to the DNA, allowing the enzyme RNA-polymerase to walk over the DNA and creating a complementary code based on the bases it reads out from the DNA. This results in an mRNA[2] string which is then transported to the outside of the nucleus and further processed.
After this, very complex machines called ribosomes can attach to the ends mRNA, walk over it, and search for the appropriate amino acids corresponding to coding triplets on the mRNA. By walking over the whole mRNA string, a chain of amino acids is created, forming a polypeptide chain that can be folded into a protein.

[2]The m stands for messenger. There are also other forms of RNA, such as tRNA which transports amino acids

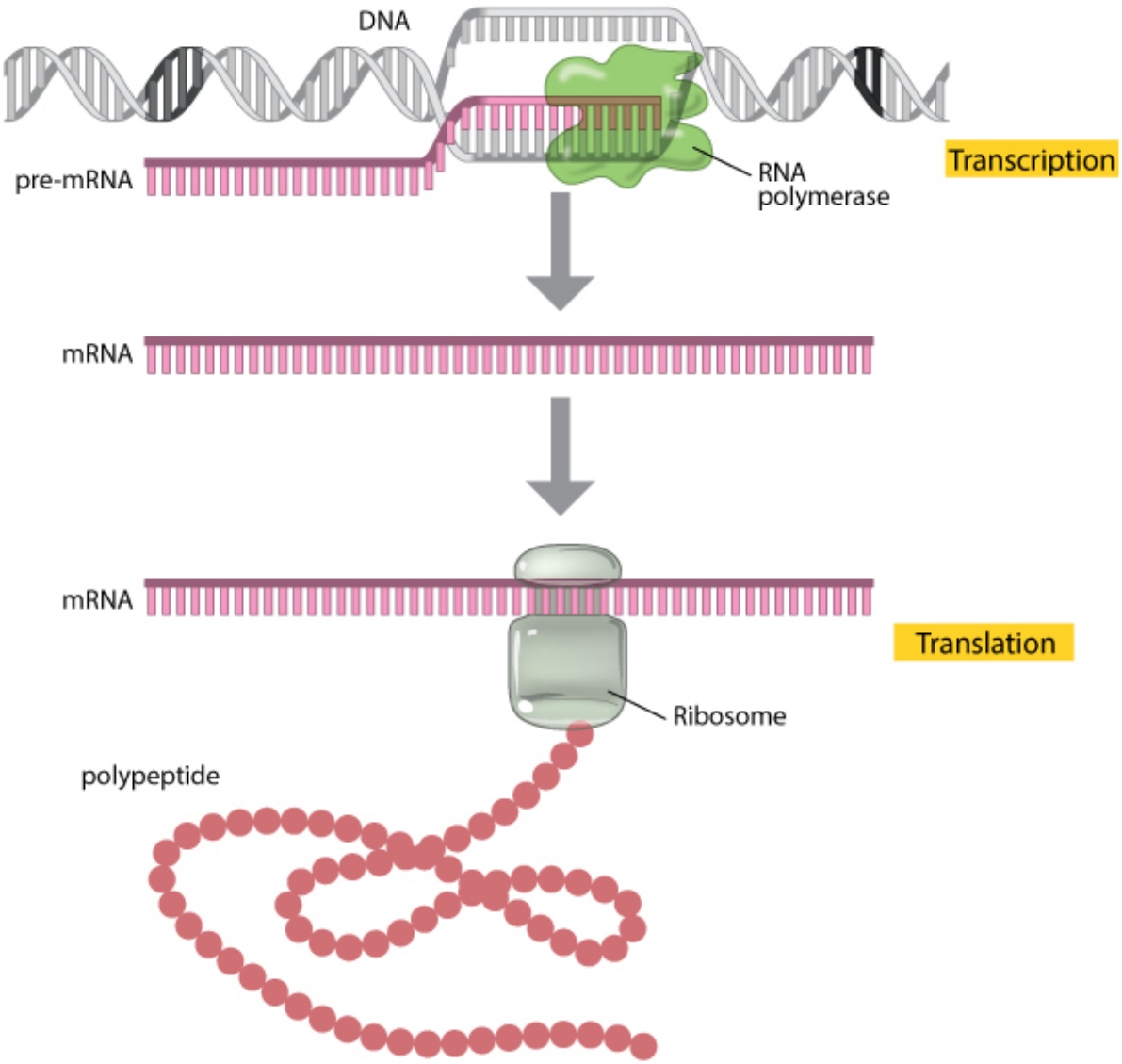

Figure 2.1: Rough sketch of transcription and translation. The steps show conversion from DNA to a polypeptide chain which forms a protein in the end. Figure taken from [10]

Chapter 3

Models

In this chapter, three models will be introduced, with a modification to all three afterwards. Specifically, the following models will be defined on a lattice $\mathbb{Z}^d$. We will first touch upon a standard continuous time random walk (CT RW). Expanding on this, we look at a second model incorporating a collection of independent CT random walkers on a lattice. The third model, moving towards a more sophisticated setting of the problem, is symmetric exclusion. All three models can be extended by incorporating sources and sinks that add or remove particles at a certain rate to the grid. The effect and results of adding sources to the models will be considered in the next chapter.

3.1 Random walks in continuous time

Suppose we have a grid $\mathbb{Z}^d$ with a particle starting in the origin. Our intuitive notion of a random walk tells us that the particle will spend some random time in $x = 0$ and attempt to jump some time later to one of its neighbours. The process then repeats itself at the new position ad infinitum. To make this notion precise, we introduce the following definition for a random walk on $\mathbb{Z}^d$.

Definition 3.1.1. Let $(X_t)_{t\geq 0}$ be a Markov process with state space $\mathbb{Z}^d$. We call $(X_t)_{t\geq 0}$ a random walk starting at x, if $X_0 = x$, the transition matrix corresponding to (X_t) is stochastic and has nonzero entries if y is a neighbour of a point x, and zero if y is not a neighbour of x. That is, $p(x, y) \geq 0$ if y is a neighbour of x and $p(x, y) = 0$ if y is not a neighbour of x. For clarity, we say that x is a neighbour of y if $|x - y| = 1$.

A random walk as a Markov process. This is intuitively clear, since a particle should not care about where it has been in the past for a random walk and should be able to jump at any time, independent of the time it has spent at the current position.

As noted in the background, the generator is a central tool to analyze and calculate the time evolution of random walks or other dynamical systems. The generator of a simple random walk is listed in the following theorem. It will be proven in chapter 5.

Lemma 2. Let $f : \mathbb{Z}^d \to \mathbb{Z}$ and $(X_t)_{t\geq 0}$ be a random walk of a single particle on $\mathbb{Z}^d$. The

generator $\mathcal{L}$ of a simple random walk is given by

$$\mathcal{L}f(x) = \sum_y p(x,y)[f(x) - f(y)] \tag{3.1}$$

where x and y denote sites, and $p(x,y)$ denotes the transition probability

3.2 Independent random walkers (IRW)

Instead of considering a single particle, we can also consider a whole collection of particles, placed as a configuration η over the grid $\mathbb{Z}^d$. That is, given a location $x \in \mathbb{Z}^d$, there are $\eta(x)$ particles at location x. We let all these particles execute individual random walks, independent of the other particles. This results in the following definition:

Definition 3.2.1. Let η be a configuration of particles on the grid $\mathbb{Z}^d$. We say that the particles in the configuration η are independent random walkers if for every particle $i \in \eta$, i executes a simple continuous time random walk. The resulting configuration at time t is denoted by η_t.

The generator of the configuration turns out to be very similar for a configuration of particles executing a simple random walk by a principle called duality where the behaviour of a configuration of many particles can be reduced to the behaviour of one single particle. The result follows immediately from independence.

Lemma 3. Let $f : \mathbb{Z}^d \to \mathbb{Z}$ and a configuration η on $\mathbb{Z}^d$ evolve as independent random walkers. The generator $\mathcal{L}$ of the IRW is given by

$$\mathcal{L}f(x) = \sum_y p(x,y)[f(x) - f(y)] \tag{3.2}$$

where x and y denote sites, and $p(x,y)$ denotes the transition probability

3.3 Symmetric exclusion

The main idea behind the exclusion principle is that one particle can occupy a site at maximum. This means that a configuration η on $\mathbb{Z}^d$ has the restriction that for every $x \in \mathbb{Z}^d$, $\eta(x) = 1$ or $\eta(x) = 0$. If we let this configuration evolve in time, we also set the restriction that a particle cannot jump to a new location x if the site is already taken by another particle. It turns out that, to consider the jump rates to be symmetric is much more easy than asymmetric jump rates, that is $p(x,y) = p(y,x)$ for all $x, y \in \mathbb{Z}^d$. This leads to the following definition for the symmetric exclusion process on $\mathbb{Z}^d$.

Definition 3.3.1. (Exclusion process) Let $\Omega = \{0,1\}^{\mathbb{Z}^d}$ be the configuration space. If x, $y \in \mathbb{Z}^d$ are neighbours $< xy >$ is defined to be the edge between x and y. Connect to each edge $< xy >$ a Poisson process $\{N_t^{<xy>} : t \geq 0\}$ with mean 1. Let $\mathcal{E}$ be the family of all Poisson clocks associated with all edges $< xy >$. Then, the symmetric exclusion process is defined by the following rule. If a Poisson clock $N_t^{<xy>} \in \mathcal{E}$ rings, i.e. an event

occurs, then if both x, y are empty or both occupied nothing will happen. However, if either x or y is occupied, the particles will swap. After this action, all clocks will reset and the process will repeat itself.

The generator of a symmetric exclusion process turns out to be the same as the one for random walks. This is caused by the symmetry requirements, as we will see in the proof in chapter 5.

Lemma 4. Let $f : \mathbb{Z}^d \to \mathbb{Z}$ and η a configuration $\{0,1\}^{\mathbb{Z}^d}$. Let this configuration evolve as dictated by the exclusion process. The generator $\mathcal{L}$ is given by

$$\mathcal{L}f(x) = \sum_y p(x,y)[f(x) - f(y)] \tag{3.3}$$

where x and y denote sites and $p(x,y)$ denotes the transition probability

The argument for this generator is that cross-symmetric terms exactly cancel each other. That is, $-p(x, x+1)\eta(x)(1-\eta(x+1))$ if a particle wants to move from x to $x+1$ and if a particle wants to move from $x+1$ to x, $p(x+1,x)\eta(x+1)(1-\eta(x))$. Adding these terms, we see a total contribution off $p(x,x+1))\eta(x+1) - \eta(x)$ by symmetry of $p(x,y)$.

3.4 Addition of sources and sinks

In two of the models defined, IRW and symmetric exclusion, we can add sources and sinks to the lattice. This means that particles get added at a certain location at some rate λ or removed at some rate μ. We consider the following effects this has on the generators in the case of the addition of a source. In case we add a sink, or multiple sources, a similar statement can be made.

Lemma 5. Let $f : \mathbb{Z}^d \to \mathbb{Z}$ and a configuration η on $\mathbb{Z}^d$ evolve as independent random walkers. The generator $\mathcal{L}$ of the IRW is given by

$$\mathcal{L}f(x) = \sum_y p(x,y)[f(x) - f(y)] + \lambda\delta_{x,0} \tag{3.4}$$

where x and y denote sites, and $p(x,y)$ denotes the transition probability

Lemma 6. Let $f : \mathbb{Z}^d \to \mathbb{Z}$, η a configuration $\{0,1\}^{\mathbb{Z}^d}$ and place a source in λ. Let this configuration evolve as dictated by the exclusion process. The generator $\mathcal{L}$ is given by

$$\mathcal{L}f(x) = \sum_y p(x,y)[f(x) - f(y)] + \lambda(1 - f(x))\delta_{x,0} \tag{3.5}$$

where x and y denote sites and $p(x,y)$ denotes the transition probability

Chapter 4

Results

This chapter aims to lay out the results of time evolution of the IRW model and symmetric exclusion. We will first touch upon a standard IRW on $\mathbb{Z}^d$ and after that add a source to the origin. Furthermore, we will touch upon invariant distributions for IRW over time. Specifically, it turns out that Poisson distributions stay invariant for IRWs. For exclusion, we will immediately place a source in the origin and will analyze the behaviour over time. We will then also consider a sink in the origin and lastly, arbitrary sources and sinks over the whole grid.

4.1 Evolution for IRW model

One of the simplest ways to approximate translation is to consider the mRNA to be an infinite string where particles are added and execute independent random walks. This is of course hardly accurate, but provides a very crude, simple model. Generalizing, we work on a grid $\mathbb{Z}^d$ with an initial configuration η where particles are added at rate λ to the origin. Since the particles execute random walks, we are interested in $\mathbb{E}_\eta[\eta_t(x)]$ which is the expected value of $\eta_t(x)$ if we start from a distribution η. The main problem of the IRW model is to actually determine the time evolution of this quantity and thus the time evolution of the whole IRW system. From this, we hope to determine if cluttering or different regimes occur depending on different initial conditions.

The standard technique to approach these problems is to look at the time derivative of this expectation. We will first consider one dimension, in which we obtain the following expression:

$$\frac{d}{dt}\mathbb{E}_\eta[\eta_t(x)] = \mathbb{E}_\eta[\eta_t(x+1) + \eta_t(x-1) - 2\eta_t(x)] + \lambda\delta_{x,0} \tag{4.1}$$

With $\mathbb{E}_\eta[\eta_t(x)]$, we mean the expected value of $\eta_t(x)$ as we start η. Note that the right term is exactly the generator. To intuitively justify this equation, consider figure 4.1. We see that for every $x \in \mathbb{Z}$, we get two inwards and two outwards contributions which together produce the required equation. Notice that there is also a source term which is only active at $x = 0$.

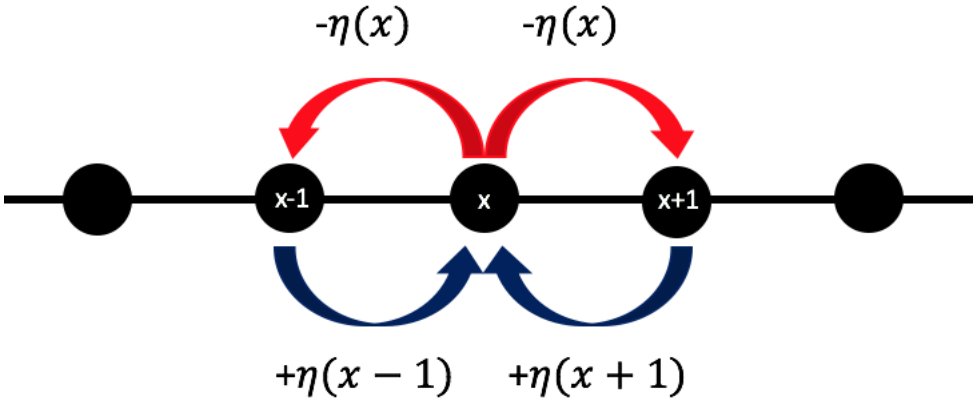

Figure 4.1: Figure showing the contributions per position x for the generator of a random walk. Notice that indeed per position x, 4 contributions are possible

Now, since we know standard random walks very well, we study the case $\lambda = 0$ first. This means that we start with a configuration η over the grid and look at the time evolution, where each particle does a random walk independent of the others.

4.1.1 IRWs where $\lambda = 0$: expected particle number

To simplify notation, we write $\mathbb{E}_\eta \eta_t(x) = \psi(t, x)$. Note that $\psi(0, x) = \eta(x)$. Then, rewriting our equation, we get:

$$\frac{d}{dt}\psi(t, x) = \psi(t, x + 1) + \psi(t, x - 1) - 2\psi(t, x) \tag{4.2}$$

One might recognize that this equation very much looks like a system $\frac{d}{dt}\psi = A\psi$, where A is some type of operator. Recalling the generator and semigroup tools that were developed in the previous chapter, if we were to find some generator $\mathcal{L}$ and corresponding semigroup S that were to reproduce this system, we can solve the system. Since all particles execute IRW, a possible candidate generator might be the generator of the random walk of a single particle. Using these tools, the following can be proven as a solution of the system. The complete proof itself will be given next chapter.

Theorem 7. Let the time evolution of the IRW model with initial configuration η on $\mathbb{Z}^d$ be dictated by 4.1. The solution of the system is given by:

$$\psi(t, x) \equiv \mathbb{E}_\eta[\eta_t(x)] = \mathbb{E}_x^{RW}[\eta(X_t)] \tag{4.3}$$

where $(X_t)_{t \geq 0}$ is the random walk of a single particle starting from x

This result is rather remarkable. It shows that the evolution of a whole collection of particles that execute independent random walks is given by the random walk of a single

particle during a time t and then evaluated at η. This reduces a many particles problem to a single particle problem. This principle/trick is also known as duality or self-duality in the literature [5].

4.1.2 Expected particle number with a source ($\lambda \neq 0$)

By adding $\lambda \neq 0$ to the previous equation in the origin, we have basically added a source, emitting particles at rate λ. This results in the equation:

$$\frac{d}{dt}\phi(t,x) = \phi(t,x+1) + \phi(t,x-1) - 2\phi(t,x) + \lambda\delta_{x,0} \tag{4.4}$$

As seen, the system becomes non-linear and our previous technique fails and we are therefore unable to deploy duality immediately. However, we do have a homogeneous solution to the system, so we might be able to obtain a solution by using variation of constants. In higher dimensions, this leads to the following solution of the system with a full proof in the next chapter:

Theorem 8. Let η be a configuration on $\mathbb{Z}^d$, λ a source in the origin and let this system evolve by the IRW model dictated by a higher dimensional version of 4.4. The expected particle number at time t is given by:

$$\phi(t,x) = \mathbb{E}_\eta[\eta_t(x)] = \int_0^t e^{sA}\lambda\delta_{x,0}ds + e^{tA}\psi(0,x) = \int_0^t e^{sA}\lambda\delta_{x,0}ds + \mathbb{E}_x^{RW}[\eta(X_t)] \tag{4.5}$$

where A is the generator of the IRW model.

This solution shows that the independent random walkers evolve as usual, indicated by the presence of the term $e^{tA}\psi(0,x)$. However, there is also an addition of a term accounting for the particles that are added by the source and might move to the location of interest x in time t.

We will now analyze the solution qualitatively. It roughly says that source attenuates the distribution over time. As λ grows, this attenuation also grows larger. Our main problem is whether the solution explodes as t tends to infinity. Now, the right most term will not diverge a.s. by recurrence of random walks. So, we are interested whether the term $\int_0^\infty e^{sA}\lambda\delta_0 ds$ converges. This is entirely dependent on the mean time spent at 0 by the random walk of a single particle. Since we have no information about this distribution, we can't say more about this term.

4.1.3 Invariance of the Poisson distribution

Since we are not able to determine quantitative results on the evolution unless the underlying distribution of the configuration is specified, we have to take a different approach. Suppose we start with a known distribution e.g. a product of Bernoulli RVs on the grid $\mathbb{Z}^d$. Will this still be a product of Bernoulli distributions at time t? The intuitive answer

is no. When we let the system evolve, it seems very likely that at least at some location, two particles will occupy the same site. However, one of the distributions that turns out stay invariant over time is the Poisson distribution with independent values over the grid. Even more interesting, it turns out that even if we add a source at the origin the distribution will stay Poisson over time. This is captured in the following two theorems.

Theorem 9. ($\lambda = 0$) Let $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i}$ be a product of Poisson measures. Then, $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i,t}$ is still a Poisson measure for all $t \geq 0$.

Theorem 10. ($\lambda \neq 0$) Given an initial distribution $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i}$ and a source term at $z = 0$ which adds particles at rate λ, $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i,t}$ is still a Poisson measure for all $t \geq 0$.

The proofs of both theorems are quite intricate and even quite different from each other, showing that the second theorem is a nontrivial fact. Both will be proven in next chapter.

4.2 Symmetric exclusion

We now set the restriction that one particle can occupy a site at maximum. This is similar to two ribosomes not being able to occupy the same RNA acid triplet in translation or two cars not being able to occupy the same spot in a traffic jam. Now, given a product of Bernoulli RVs over the grid $\mathbb{Z}^d$ with evolution dictated by the symmetric exclusion rule, we are interested in the behaviour over time when we add a source to the origin. Since only one particle is allowed per site, will the grid completely fill up, or will there always be empty spots near the origin? We will also consider the behaviour of a sink in the origin. In this case we are interested whether all particles will disappear from the grid over time. Finally, we will combine both cases and place arbitrarily many sources and sinks at different positions in the grid. In this case we also want to know what happens over time. Will there be some form of cluttering? Will there be regimes with high density and low densities?

In order to deal with exclusion, we have to utilize different tools than before. Previously, we were able to exploit duality and independence to solve our problems. This is not possible anymore, since we have lost independence through the exclusion rule. This is exactly what makes exclusion much harder to deal with.

4.2.1 Symmetric exclusion with a source

For simplicity, we will again first consider the 1D case and easily generalize to higher dimensions later. Using the generator for symmetric exclusion in 1D, we observe the time derivative of our expected distribution again.

$$\frac{d}{dt}\mathbb{E}_\eta[\eta_t(x)] = \begin{cases} \mathbb{E}_\eta\left[\eta_t(x+1) + \eta_t(x-1) - 2\eta_t(x)\right] & \text{if } x \neq 0 \\ \mathbb{E}_\eta\left[\eta_t(1) + \eta_t(-1) - 2\eta_t(0)\right] + \lambda(1-\eta_t(0)) & \text{if } x = 0 \end{cases} \tag{4.6}$$

Theorem 11. (Exclusion with a source) Suppose a symmetric exclusion process is given on $\mathbb{Z}^d$ with initial configuration $\eta = \{0,1\}^{\mathbb{Z}^d}$ and a source term emitting particles at rate λ. That is, if the time evolution satisfies a higher dimensional version of ODE (4.6), then the time evolution of η is given by

$$\psi(t,x) = \mathbb{E}_\eta[\eta_t(x)] = 1 - \mathbb{E}_x[e^{-\lambda l_t(0)}(1-\eta_0(X_t)] \tag{4.7}$$

where $l_t(0) = \int_0^t I(X_s = 0)ds$ is defined to be the local time spent at 0 of a particle executing a random walk $(X_t)_{t\geq 0}$ starting from x.

The theorem is consistent with the exclusion rule. For every $x \in \mathbb{Z}^d$, there can be a maximum of 1 particle, since $\mathbb{E}_x[e^{-\lambda l_t(0)}(1-\eta_0(X_t)]$ is a positive quantity between 0 and 1. Also, comparing with our analysis for IRWs, we see that duality indeed doesn't hold. Previously, we saw that $\mathbb{E}_\eta[\eta_t(x)] = \mathbb{E}_x[\eta(X_t)]$. But now, the term $\mathbb{E}_\eta[\eta_t(x)]$ is entirely different. This is because the system needs to incorporate a source term and the exclusion rule.

Now, to properly analyze the behaviour of the solution, we set a Bernoulli distribution over the grid, $\nu_\rho = \otimes_{x\in\mathbb{Z}^d} Ber(\rho)$. This means that every gridpoint has an equal probability to either contain, or not contain a particle. Using this distribution, we can find the distribution over time of our grid:

$$\psi(t,x) = 1 - (1-\rho)\mathbb{E}_x[e^{-\lambda l_t(0)}] \tag{4.8}$$

Therefore, the behaviour of the distribution entirely depends on the time spent at 0 of particles. This is in line with our intuition. By the law of large numbers, a fraction ρ of the grid is exactly filled with particles. For the grid to overflow with particles eventually, there must be enough particles added at 0. This means that the 0 site must be occupied for a certain total time to obtain overflow behaviour. We will see which cases for $l_t(0)$ lead to overflow behaviour.

Limit distributions for $l_t(0)$

Theorem 12. Suppose we are given a Bernoulli distribution $\nu_\rho = \otimes_{x\in\mathbb{Z}^d} Ber(\rho)$ that evolves by the exclusion rule with a source in the origin as specified in theorem 11. Under recurrence in the origin, that is, $l_t(0) \to \infty$ as $t \to \infty$, then $\lim t \to \infty \psi(t,x) = 1$. For transience in the origin, convergence of $\psi(t,x)$ as $\lim t \to \infty$, depends on the form $l_\infty(0)$. Then,

$$\lim_{t\to\infty} \psi(t,x) = 1 - (1-\rho)\mathbb{E}_x[e^{-\lambda l_\infty(0)}] \tag{4.9}$$

If $l_t(0) \to \infty$ a.s. (recurrence), we observe that $\psi(t,x) \to 1$ as $t \to \infty$. At first glance, this might seem for some people as counter intuitive, while for others it seems trivial. For both camps something can be said. One might expect that as $l_t(0) \to \infty$ as $t \to \infty$, only a few particles that reside near the origin, occupy the origin for long enough so that the grid never fill up completely. However, the analysis that has been done shows this is not

the case. It turns out, as particles are generated at the origin at every chance possible, the grid will eventually fill up completely. So, as $l_t(0) \to \infty$ as $t \to \infty$, new particles will appear at the origin when possible to make sure the grid completely fills up.

Now, another very interesting case is when $l_t(0)$ converges a.s. to some finite random variable as $t \to \infty$ (transience).While we are not able to find an explicit expression, we can do some estimations for the limiting distribution. Observing that e^{-x} is a convex function, we apply Jensen's inequality to find

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}_{\nu_\rho}[\eta_t(x)] &= 1 - (1-\rho)\mathbb{E}_x[e^{-\lambda l_\infty(0)}] \\ &\leq 1 - (1-\rho)e^{-\lambda G(x,0)} \end{aligned} \tag{4.10}$$

Here, $G(x,0) = \int_0^\infty \rho_t(x,0)$. Now, since we have no specific knowledge of the distribution $\rho_t(x,0)$, we cannot make further estimations, though it is likely that this term will depend on the dimension we work in.

4.2.2 Symmetric exclusion with a sink

In a similar way as the previous section, we can consider symmetric exclusion on a grid $\mathbb{Z}^d$ with a sink placed in the origin. This means that a particle is removed at rate λ if a particle is present at the origin and nothing is done if there is no particle. This results in a similar system to (4.6):

$$\frac{d}{dt}\mathbb{E}_\eta[\eta_t(x)] = \begin{cases} \mathbb{E}_\eta\left[\eta_t(x+1) + \eta_t(x-1) - 2\eta_t(x)\right] & \text{if } x \neq 0 \\ \mathbb{E}_\eta\left[\eta_t(1) + \eta_t(-1) - 2\eta(0)\right] - \lambda\eta_t(0) & \text{if } x = 0 \end{cases} \tag{4.11}$$

The solution of the time evolution is captured by a similar theorem:

Theorem 13. Suppose a totally symmetric exclusion process is given on $\mathbb{Z}^d$ with initial configuration $\eta = \{0,1\}^{\mathbb{Z}^d}$ and a sink absorbing particles at rate λ if they are present at the origin. That is, if the time evolution satisfies a higher dimensional form of the ODE (4.11), then the time evolution of η is given by

$$\psi(t,x) = \mathbb{E}_\eta[\eta_t(x)] = \mathbb{E}_x[e^{-\lambda l_t(0)}\eta_0(X_t)] \tag{4.12}$$

where $l_t(0) = \int_0^t I(X_s = 0)ds$ is defined to be the local time spent at 0 of a particle executing a random walk (X_s) starting from x.

We therefore see that a similar equation holds for when a sink is added to the origin. Again, we will analyze the limiting behaviour as $t \to \infty$.

Limit distributions for $l_t(0)$

We first set another product of Bernoulli variables $\nu_\rho = \otimes_{x\in\mathbb{Z}^d} Ber(\rho)$ over the grid all with equal probability ρ. This leads to the form

$$\psi(t,x) = \rho\mathbb{E}_x[e^{-\lambda l_t(0)}] \tag{4.13}$$

Behaviour as $t \to \infty$ is captured in the following theorem.

Theorem 14. Set $\nu_\rho = \otimes_{x\in\mathbb{Z}^d} Ber(\rho)$ as an initial configuration for symmetric exclusion with a sink. Time evolution is dictated by theorem 13. Then, as $\lim t \to \infty$, in the case that $l_t(0) \to \infty$, $\psi(t,x) \to 0$. In the case that $l_t(0)$ is transient, it converges a.s. to some random variable $l_\infty(0)$. Then, $\lim t \to \infty \psi(t,x) = \rho\mathbb{E}_x[e^{-\lambda l_\infty(0)}]$

The recurrent case seems to make sense. If there are always particles that occupy the sink site, but they are only able to disappear, they should be removed eventually. Taking long enough, this results in a 0 distribution over the whole grid.

In the case that $l_t(0) \to \infty$ converges a.s. to some random variable $l_\infty(0)$, we are only sure[1] that $\lim_{t\to\infty} \psi(t,x) = \rho\mathbb{E}_x[e^{-\lambda l_\infty(0)}]$, unfortunately. Since we are not working with a minus sign, Jensen's inequality will only give underestimations in this case. Taylor's theorem will not work either unfortunately. So unless we have more information about $l_\infty(0)$, we have little knowledge about the distribution on the grid.

4.2.3 Symmetric exclusion with multiple sources and sinks

Now, suppose we are again on $\mathbb{Z}^d$ and decide to place sources and sinks at arbitrary positions (sources and sinks can even be situated at the same place!). We hope to say something about the long term behaviour as the system evolves. Let $\mathcal{B}$ be the collection of source positions, while $\mathcal{S}$ is the collection of a sink positions. Then, we can write the following system that determines the symmetric exclusion evolution:

$$\frac{d}{dt}\psi(t,x) = A_{RW}\psi(t,x) - \sum_{y\in\mathcal{S}} \lambda\delta_{x,y}\psi(t,x) + \sum_{y\in\mathcal{B}} \lambda\delta_{x,y}(1-\psi(t,x)) \tag{4.14}$$

Here, A_{RW} is the standard random walk generator which also is valid for symmetric exclusion. The other contributions dictate whether a particle can be added or removed at a source or sink site. As usual, $\psi(t,x) = \mathbb{E}_\eta[\eta_t(x)]$. We capture the solution to this system in our last theorem:

Theorem 15. Let $\mathbb{Z}^d$ be a grid with an exclusion process on it. Now, let $\mathcal{S}$ and $\mathcal{B}$ be the sink and source collections respectively that also take part in the exclusion process through the system (4.14). Then, the solution to this system is given by:

$$\psi(t,x) = 1 + \mathbb{E}_x\left[e^{-\lambda\left(\sum_{y\in\mathcal{S}\cup\mathcal{B}} l_t(y)\right)}\eta_0(X_t)\right] - \mathbb{E}_x\left[e^{-\lambda\sum_{y\in\mathcal{B}} l_t(y)}\right] \tag{4.15}$$

where $l_t(y) = \int_0^t I(X_s = y)ds$ is defined as the local time the random walk (X_t) spends at position y during $[0,t]$.

We will first do some sanity checks for the validity of this equation. In case we set the sink collection to be empty and take one source in the origin, we observe that we resolve the original equation to the source in origin case considered before. Similarly, if we set the source set to be empty, the right term becomes 1 and cancels out the first term. Placing only a sink in the origin, we are left with the original equation of the previous section.

[1] no pun intended

This points towards a justification for this equation. A formal proof of the correctness of this equation is given in the next section.

There are many different combinations of sources and sinks that can be observed. Specifically, if we place a source and sink in the origin, the time evolution is given by:

$$\psi(t,x) = 1 + \mathbb{E}_x[e^{-2\lambda l_t(0)}\eta_0(X_t)] - \mathbb{E}_x[e^{-\lambda l_t(0)}] \tag{4.16}$$

We capture the resulting effects in the transient in a theorem, because it is so remarkable and unexpected.

Theorem 16. Suppose we take the same setting as in theorem 15, put a Bernoulli distribution on the grid, but put only a source and sink in the origin. Under a recurrent occupation of the origin, that is, $l_t(0) \to \infty$ as $t \to \infty$:

$$\lim_{t\to\infty} \psi(t,x) = 1 \tag{4.17}$$

This is quite surprising and a non trivial fact. It seems that when one places a source and a sink at the origin that remove or add a particle when possible, the grid will still overflow in the recurrent case. This is not what one expects a priori, because particles have the same probability of being removed as being added.

In the transient case, not much interesting can be said unless one knows the distribution. If we assume that $l_t(0) = 0$, which means that when particles arrive at 0, they must immediately move away, the distribution takes form $\psi(t,x) = \mathbb{E}_x[\eta_0(X_t)]$. This shows that the distribution mimics the random walk of a single particle and is then evaluated at $\eta_0(X_t)$.

4.3 The leap towards a TASEP model for transcription and translation

In this project, very simple prototype models have been created for modelling biological systems. The main motivation for this project was to mathematically explain the mechanism behind the speed at which proteins are produced as a cell has been created after cell division. As explained before, producing proteins requires two processes: transcription and translation. Both processes have a lot of similarities with totally asymmetric exclusion (TASEP). Fixed start-stop sites, one directional behaviour and stochastic dynamics cause TASEP to be a suitable model. For E. Coli[**?**], translation occurs while transcription is not yet finished. This means that the two processes are coupled in some way, meaning that the TASEP needs to be coupled.
Contrary to symmetric exclusion, transcription is obviously not symmetric at all, so requires TASEP. TASEP has already been applied to model the random walk of RNA-polymerase [[1]]. More interestingly, a possibility to explain the rate of protein synthesis might be to not only apply TASEP to the walk of RNA polyemerases, but also to the ribosomes that travel on mRNA sticking out from the RNA polymerases.

The reason that TASEP is so much more difficult in a mathematical sense than symmetric exclusion is the following. At the start of the symmetric exclusion section, we stated that particles move in all directions with same probability, if a site is vacant. This was described by contributions $-\eta(x)(1-\eta(x+1)$ if a particle wants to move away from location x to $x+1$. Similarly, if a particle wanted to move from $x+1$ to x, this was described by $\eta(x+1)(1-\eta(x))$. Adding these terms results in a total contribution $\eta(x+1)-\eta(x)$, which shows that the cross terms disappear. This is because we are working with symmetric exclusion. In the case of asymmetric exclusion, these terms don't disappear anymore. As used before, notice that these terms exactly look like the generator of a random walk. Losing this generator form, we also use all the tools we have used previously. Therefore, TASEP is much more difficult than normal exclusion and utilizes different methods.

Chapter 5

Proofs and detailed calculations

We will show the proofs of various claims made in the previous chapter. Amongst others, we will calculate the generators, show invariance for Poisson measures and show proofs for the solution operators for symmetric exclusion.

5.1 Generator

Lemma 17. Let $f : \mathbb{Z}^d \to \mathbb{Z}$ and $(X_t)_{t\geq 0}$ be a random walk of a single particle on $\mathbb{Z}^d$. The generator $\mathcal{L}$ of a simple random walk is given by

$$\mathcal{L}f(x) = \sum_y p(x,y)[f(x) - f(y)] \tag{5.1}$$

where x and y denote sites, and $p(x,y)$ denotes the transition probability

Proof. Suppose the process X_t starts at $x \in \mathbb{Z}^d$. The probability that the particle makes more than one jump in the time window $[0,t]$ is bounded by Ct^2 where C is some constant in $\mathbb{R}$. Now, let K_t denote the number of jumps in $[0,t]$. We attempt to invoke the expression for the generator $\mathcal{L}$ given in (8). So, consider $S_t f - f$. Then,

$$\begin{aligned} S_t f(x) - f(x) &= \mathbb{E}_x(f(X_t)) - f(x) \\ &= \mathbb{E}_x(f(X_t) \cap K_t = 0) + \mathbb{E}_x(f(X_t) \cap K_t = 1) + R_t - f(x) \end{aligned} \tag{5.2}$$

Here, R_t is some process bounded by Ct^2. Now, using the transition rates of a random walk, we get:

$$\begin{aligned} S_t f(x) - f(x) &= (1 - e^{-c_x t}) \sum_y \frac{c(x,y)}{c_x} f(y) + e^{-c_x t} f(x) - f(x) + R_t \\ &= (1 - e^{-c_x t}) \sum_y \frac{c(x,y)}{c_x} (f(y) - f(x)) + R_t \end{aligned} \tag{5.3}$$

Dividing the left side by t and taking the limit to 0, we obtain the desired result. □

5.2 IRWs without a source

We will first prove theorem 7. For completeness, we restate it here again:

Theorem 18. Let the time evolution of the IRW model with initial configuration η on $\mathbb{Z}^d$ be dictated by 4.1. The solution of the system is given by:

$$\psi(t,x) \equiv \mathbb{E}_\eta[\eta_t(x)] = \mathbb{E}_x^{RW}[\eta(X_t)] \tag{5.4}$$

where $(X_t)_{t\geq 0}$ is the random walk of a single particle starting from x

Proof. The underlying idea is to look at the equalities of the generators for a random walk of a single particle and multiple independent random walkers. So, we turn to the evolution of a single particle starting at x. Now, let $S(t)$ be the semigroup operator corresponding to $S(t)\phi(x) = \mathbb{E}_x^{RW}[\phi(X_t)]$ which is associated with the random walk $(X_t)_{t\geq 0}$. Using the property $\frac{d}{dt}S(t)f(x) = S(t)\mathcal{L}f(x)$ and the generator for the random walk of a single particle, we find the system:

$$\frac{d}{dt}\mathbb{E}_x^{RW}[\phi(X_t)] = \mathbb{E}_{x+1}^{RW}[\phi(X_t)] + \mathbb{E}_{x-1}^{RW}[\phi(X_t)] - 2\mathbb{E}_x^{RW}[\phi(X_t)] \tag{5.5}$$

Because the generators for IRWs and single particles are equal, the corresponding ODEs and therefore solutions are also equal. This allows us to exploit the generator tools to find the time evolution of the configuration η. So, we find the following solution of the system in (4.2):

$$\psi(t,x) = e^{t\mathcal{L}}\psi(0,x) = \mathbb{E}_x^{RW}[\eta(X_t)] \tag{5.6}$$

where the last equality follows from $\psi(0,x) = \eta(x)$ and $e^{t\mathcal{L}} = S(t)$.
A simplified solution is hard to obtain, since that would mean that the solution operator $e^{t\mathcal{L}}$ must be computed. Even for matrices X, it is often hard to calculate e^X unless X is a diagonalizable matrix. Now, we know that $S(t) = e^{t\mathcal{L}}$, so this tells us that:

$$\psi(t,x) = \mathbb{E}_x^{RW}[\eta(X_t)] \tag{5.7}$$

□

5.3 IRWs with a source

The appropriate equation of interest was given by:

$$\frac{d}{dt}\phi(t,x) = \phi(t,x+1) + \phi(t,x-1) - 2\phi(t,x) + \lambda\delta_{x,0} \tag{5.8}$$

The solution was captured in the following theorem of which we give the proof now:

Theorem 19. Let η be a configuration on $\mathbb{Z}^d$, λ a source in the origin and let this system evolve by the IRW model dictated by a higher dimensional version of 4.4. The expected particle number at time t is given by:

$$\phi(t,x) = \mathbb{E}_\eta[\eta_t(x)] = \int_0^t e^{sA}\lambda\delta_{x,0}ds + e^{tA}\psi(0,x) = \int_0^t e^{sA}\lambda\delta_{x,0}ds + \mathbb{E}_x^{RW}[\eta(X_t)] \quad (5.9)$$

where A is the generator of the IRW model.

Proof. Set $\phi_h(t,x) = e^{t\mathcal{L}}\psi(x,0)$ and rewrite the system more clearly as

$$\frac{d}{dt}\phi(t,x) = A\phi(t,x) + \lambda\delta_{x,0} \quad (5.10)$$

Note that this equation is inhomogeneous in the t variable. Our proposed $\phi_h(t,x)$ is the homogeneous solution of this system, since it is the solution of the situation discussed in theorem 15. We attempt variation of constants. Therefore, suppose that the solution of this system is given by $\phi_t = e^{t\mathcal{L}}\chi_t$. We take χ_t to be some unspecified function.

We temporarily suppress position dependence to reduce notation. So, $\phi(t,x) = \phi_t$. Differentiating and applying the chain rule results in a new equation:

$$\frac{d}{dt}\phi_t = Ae^{tA}\chi_t + e^{tA}\frac{d}{dt}\chi_t = A\phi_t + \lambda\delta_0 \quad (5.11)$$

Here, the right side is the original system and the middle is the result of applying the chain rule to our ansatz. Combining both, we get the following ODE, which can be integrated then be integrated.

$$e^{tA}\frac{d\chi_t}{dt} = \lambda\delta_0 \quad (5.12)$$

$$\chi_t = \int_0^t e^{-sA}\lambda\delta_0 ds + \psi_0 \quad (5.13)$$

We can insert this expression back into our ansatz to obtain the following solution:

$$\phi_t = \int_0^t e^{(t-s)A}\lambda\delta_0 ds + e^{tA}\psi_0 \quad (5.14)$$

A quick substitution $u = t - s$ grants us the final solution:

$$\phi_t = \int_0^t e^{sA}\lambda\delta_0 ds + e^{tA}\psi_0 \quad (5.15)$$

Now, the methods by which this equation was obtained might seem questionable, since we have nothing guarantees the existence of the inverse of e^{tA}. However, to get around this, we pose the solution we obtain as *the* solution to the system and check that this solution satisfies the ODE. This turns out to be indeed the case and invite the reader to check this by similar steps as stated above.

□

5.4 Invariance for Poisson measures under IRWs

In this section, we show three proofs about the invariance of Poisson measures on a lattice $\mathbb{Z}^d$. Two proofs are when no source is present in the origin and the last proof shows that Poisson measures are even invariant when a source is present emitting particles at some rate λ. The reason we show two proofs is because the proof of invariance for sources is fundamentally different from the first proof. We therefore repeat the proof of the first theorem in the same spirit as the second.

5.4.1 $\lambda = 0$

To prove the case $\lambda = 0$, we first need some small tools. Let ξ be a finite configuration on $\mathbb{Z}$ and define the following polynomial: $D(\xi, \eta) = \prod_x \frac{\eta_x!}{(\eta_x - \xi_x)!}$. Here, η_x and ξ_x are both the number of particles of the distribution at location x where x ranges over the grid $\mathbb{Z}^d$. Since ξ is a finite configuration, we can write it as a sum of of Kronecker deltas. Therefore, we examine the action of this polynomial for some specific choices related to Kronecker deltas. For example, let $\xi = \delta_x$. Then, all terms in the numerator and denominator are equal for all x, except at the location specified by the Kronecker delta. So, $D(\delta_x, \eta) = \eta_x$. If $\xi = 2\delta_x$, we find in a similar way that $D(2\delta_x, \eta) = \eta_x(\eta_x - 1)$. As a last example, let $\xi = \delta_x + \delta_y$. Then, $D(\delta_x + \delta_y, \eta) = \eta_x \eta_y$. We return now to create the setting for our Poisson problem. So, let ν_ρ be a Poisson distribution with density ρ. What we are interested in, is whether the measure $\otimes_{i\in\mathbb{Z}}\nu_{\rho_i}$ is invariant over time.

Theorem 20. Let $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i}$ be a product of Poisson measures. Then, $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i,t}$ is still a Poisson measure for all $t \geq 0$.

In order to prove this, we make use of the following property of the Poisson measure. The defining property for the Poisson measure is $\int D(\xi, \eta)(\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i})(d\eta) = \prod_{i\in\mathbb{Z}^d} \rho_i^{\xi_i}$ [5]. Therefore, if we show that at time t, the distribution still has this form, we know it's a Poisson measure.

Proof. At time t, we have a distribution $\otimes_{i\in\mathbb{Z}}\nu_{\rho_i}S(t)$, where S is the IRW semigroup. So,

$$\int D(\xi, \eta)\left[(\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i})S(t)\right](d\eta) = \int \mathbb{E}_\eta D(\xi, \eta_t)(\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i}))(d\eta) \tag{5.16}$$

Here, we have applied the semigroup operator to obtain the conditional expectation. Now comes the crucial step, where we invoke duality and Fubini's theorem to obtain:

$$\begin{aligned}\int \mathbb{E}_\eta D(\xi, \eta_t)(\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i}))(d\eta) &= \int \mathbb{E}_\xi D(\xi_t, \eta)(\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i})(d\eta) \\ &= \mathbb{E}_\xi \int D(\xi_t, \eta)(\otimes_{i\in\mathbb{Z}^d})(d\eta) \\ &= \mathbb{E}_\xi \left(\prod_i \rho_i^{\xi_t(i)}\right)\end{aligned} \tag{5.17}$$

Now, we simplify notation a bit in order to continue manipulation. We know that ξ is a finite configuration, so we can rewrite the product in a different way: $\prod_{i\in\mathbb{Z}} \rho_i^{\xi_i} = \prod_{j=1}^n \rho_{x_j}$. Here, the locations x_j are the occupied sites and we know the particle number at x_j, so we simply write ρ_{x_j}. Then, we can rewrite the conditional expectation to yield:

$$\begin{aligned}\mathbb{E}_\xi\left(\prod_i \rho_i^{\xi_t(i)}\right) &= \mathbb{E}^{IRW}_{x_1,\dots,x_n}\left(\prod_{j=1}^n \rho_{X_j(t)}\right) \\ &= \prod_{j=1}^n \mathbb{E}^{IRW}_{x_j}\rho_{X_j(t)} \\ &= \prod_{j=1}^n (\rho_t)_{x_j} = \prod_i (\rho_t)_i^{\xi_i}\end{aligned} \tag{5.18}$$

We have used independence and the expectation for Poisson distributions successively. Indeed, we see that we have obtained the desired form. So the distribution is still Poisson at time t. □

5.4.2 $\lambda \neq 0$

Theorem 21. Given an initial distribution $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i}$ and a source term at the origin which adds particles at rate λ, $\otimes_{i\in\mathbb{Z}^d}\nu_{\rho_i,t}$ is still a Poisson measure for all $t \geq 0$.

This theorem requires a nontraditional proof. In order to illustrate this type of proof, we will also illustrate it for the case $\lambda = 0$. However, we first need to introduce a technique used in this proof. This technique is called intertwining where we do the following. Suppose we have two generators $\mathcal{L}$ and $\hat{\mathcal{L}}$ working on functions $f \in C(\Omega), C(\Omega')$ respectively. We say these operators are intertwined with the intertwining operator $G : C(\Omega) \to C(\Omega')$ such that $G\mathcal{L} = \hat{\mathcal{L}}G$. Without proof, we also state the result $GS(t) = \hat{S}(t)G$ for semigroups. In the case that $\mathcal{L}$ is bounded, we can show this using $G\mathcal{L}^n = \hat{\mathcal{L}}^n G$

Proof. (*Theorem 17 - 2nd proof*) Let $f : \mathbb{N}_0 \to \mathbb{R}$. We now introduce an object similar to the generating function which we call a generalized generating function: $Gf(z) = \sum_{n=0}^\infty f(n)\frac{z^n}{n!}$. Observe that this generalized generating function is very similar to a function integrated against the Poisson measure. That is, $\nu f = e^{-z}Gf(z) = \sum_{n=0}^\infty f(n)\frac{z^n}{n!}e^{-z}$.

We also define 3 linear operators: $Af(n) = nf(n-1)$, $Bf(n) = f(n+1)$ and $Cf(n) = nf(n)$. We look for the intertwining operators with G, the effect of the operators on Gf is as follows:

$$G(Af) = zGf(z) \quad , \quad G(Bf) = \frac{\partial}{\partial z}Gf(z) \quad \text{and} \quad G(Cf) = z\frac{\partial}{\partial z}Gf(z) \tag{5.19}$$

Having defined these operators, we will immediately generalize to define the multivariate cases for these operators. Let $\mathcal{G}f(z) = \sum f(\bar{n})\frac{z^{\bar{n}}}{\bar{n}!}$. Here, $z^{\bar{n}} = \prod z_i^{n_i}$ and $\bar{n}! = \prod n_i!$. It is important to also realize that z_i is the i-th component of z. We also need to reconsider the operators A, B and C. We'll let them work on a single component i. That is,

$A_i f(n) = n_i f(n - \delta_i)$. Similar things can be said for B_i and C_i.

Now, recall the generator of our random walk. It can immediately be rewritten as:

$$\begin{aligned}\mathcal{L}_{RW} f(n) = \frac{1}{2} \sum_{i,j} p(i,j) \left[n_i [f(n - \delta_i + \delta_j) - f(n)] + n_j [f(n - \delta_j + \delta_i) - f(n)]\right] \\ + \lambda (f(n + \delta_0) - f(n))\end{aligned} \tag{5.20}$$

We can use the previous defined operators to rewrite this expression to:

$$\begin{aligned}\mathcal{L}_{RW} f(n) &= \frac{1}{2} \sum_{i,j} p(i,j) [(B_j A_i - B_i A_i) + (B_i A_j - B_j A_j)] f(n) \\ &= \frac{1}{2} \sum_{i,j} p(i,j) [(A_i - A_j)(B_i - B_j)] f(n)\end{aligned} \tag{5.21}$$

Now, let $\mathcal{G}$ work on $f(n)$ as an operator. This same operator can also be applied to $\mathcal{L}_{RW} f(\eta)$, we get:

$$\mathcal{G}\mathcal{L}_{RW} f(n) = \frac{1}{2} \sum_{i,j} -p(i,j) \left[(z_i - z_j) \left(\frac{\partial}{\partial z_i} - \frac{\partial}{\partial z_j} \right) \right] \mathcal{G} f(n) \tag{5.22}$$

We can interpret the above as a new operator, namely $\hat{\mathcal{L}}$ for which we have the identity $\mathcal{G}(\mathcal{L}_{RW} f) = \hat{\mathcal{L}}(\mathcal{G} f)$. So we have shown that the generator of a random walk is intertwined with the generator of a deterministic ODE through the operator G. This also implies that $\mathcal{G}(S_{RW}(t)) f) = \hat{S}(t)(\mathcal{G} f)$. Using the background material on flows, we find that the corresponding ODE is:

$$\dot{z}_i = \sum_j p(i,j)(z_j - z_i) \tag{5.23}$$

where $i \in \mathbb{Z}^d$. This observation will be used later.

To continue the proof, we make another observation. The generalized generating function $\mathcal{G} f$ can be rewritten as

$$\begin{aligned}\mathcal{G} f(z) &= \sum_{\bar{n}} f(\bar{n}) \frac{z^{\bar{n}}}{\bar{n}!} \\ &= e^{\sum z_i} \left(\sum_{\bar{n}} f(\bar{n}) \frac{z^{\bar{n}}}{\bar{n}!} e^{-\sum z_i} \right) \\ &= e^{\sum z_i} \nu_z(f)\end{aligned} \tag{5.24}$$

with ν_z denoting the Poisson measure $\nu_z = \otimes_{i \in \mathbb{Z}^d} Poisson(z_i)$. The expression $\nu_z(f)$ is the standard Poisson integral. So, we have managed to relate the generalized generating

function to the Poisson measure.

Combining these observations with the intertwining relation $\mathcal{G}(S_{RW}(t))f) = \hat{S}(t)(\mathcal{G}f)$, we find that

$$\begin{aligned} e^{\sum_i z_i}\nu_z(S(t)_{RW}f) &= \hat{S}(t)\mathcal{G}f = \mathcal{G}f(z^z(t)) \\ &= e^{\sum z_i^z(t)}\nu_{z^z(t)}(f) \\ &= e^{\sum z_i(t)}\nu_{z^z(t)}(f) \end{aligned} \tag{5.25}$$

The notation $z_i^z(t)$ might be confusing. We mean the position of the coordinate z_i when started from position z and evaluated at time t. Using that the total sum over z_i is invariant over time, we can cross out the term $e^{\sum_i z_i}$ on both sides and are therefore left with

$$\nu_z(S_{RW}f) = \nu_{z^z(t)}(f) \tag{5.26}$$

This shows that the Poisson measure is invariant if we propagate using an IRW over time t.

□

Proof. (*Theorem 18*) We use the same notation and operators from the previous proof with some slight modifications. The main idea is to utilize intertwining of the random walk generator including a source with some deterministic ODE generator. It turns out that addition of a source at the origin still results in an intertwining with a deterministic ODE.

Recall the generator of the IRW model with a source in the origin, as in last proof, written in a slightly different way:

$$\begin{aligned} \mathcal{H}_{RW}f(n) = \frac{1}{2}\sum_{i,j} p(i,j)\left[\eta_i[f(n-\delta_i+\delta_j)-f(n)] + \eta_j[f(n-\delta_j+\delta_i)-f(n)]\right] \\ + \lambda(f(n+\delta_0)-f(n)) \end{aligned} \tag{5.27}$$

Using the operators that we defined, we can rewrite this as:

$$\mathcal{H}_{RW}f(n) = \frac{1}{2}\sum_{i,j} p(i,j)[(B_i-B_j)(A_i-A_j)]f(n)+\lambda(f(\eta+\delta_0)-f(n)) = \mathcal{L}_{RW}f(n)+\lambda(f(n+\delta_0)-f(n)) \tag{5.28}$$

where $\mathcal{L}_{RW}$ is the generator of a random walk without a source in the origin. Applying the operator $\mathcal{G}$ to $\mathcal{H}f(n)$ results in

$$\begin{aligned} \mathcal{G}(\mathcal{H}_{RW}f(n)) &= \mathcal{G}(\mathcal{L}_{RW}f(n)) + \mathcal{G}(\lambda f(n+\delta_0)-f(n)) \\ &= \hat{\mathcal{L}}\mathcal{G}f + \lambda\frac{\partial}{\partial z_0}\mathcal{G}f \end{aligned} \tag{5.29}$$

So, we see that the generator of the IRW with a source is intertwined with $\hat{\mathcal{H}} \equiv \hat{\mathcal{L}} + \lambda \frac{\partial}{\partial z_0}$ through $\mathcal{G}$. It turns out that this new generator $\mathcal{H}$ is also the generator of an ODE.

Now, letting the system evolve and integrating against the Poisson measure, we see that:

$$\begin{aligned} \nu_z[S(t)_{RW+source} f] &= e^{-\sum z_i}(\mathcal{G} S(t)_{RW+source} f)(\xi) \\ &= e^{-\sum z_i}(\hat{S}_{RW+source}(t)[\mathcal{G} f])(\xi) \\ &= e^{-\sum z_i}(\mathcal{G} f)(z^{\xi}(t)) \\ &= \nu_{z^{z}(t)}(f) \end{aligned} \tag{5.30}$$

These steps are not obvious at all, so we go through all steps. The first equality is given by the definition of integration against the Poisson measure starting from z and the relation of the Poisson measure with $\mathcal{G}$. The second equality is given by the intertwining relation we proved previously. Since the generator of the ODE works on the argument of the function $(\mathcal{G} f)(\xi)$, we get the third equality. This term $z^{\xi}(t)$ is the solution of the ODE starting from ξ and moving up to time t. Then, the last equality is again by definition of integration against the Poisson measure. □

5.5 Symmetric exclusion

5.5.1 Exclusion with a source

The time evolution for exclusion with a source was captured in theorem 11. We restate it here, with a complete corresponding proof.

Theorem 22. (Exclusion with a source) Suppose a totally symmetric exclusion process is given on $\mathbb{Z}^d$ with initial configuration $\eta = \{0,1\}^{\mathbb{Z}^d}$ and a source term emitting particles at rate λ. That is, if the time evolution satisfies an arbitrary dimensional form of the ODE (4.6), then the time evolution of η is given by

$$\psi(t,x) = \mathbb{E}_{\eta}[\eta_t(x)] = 1 - \mathbb{E}_x[e^{-\lambda l_t(0)}(1 - \eta_0(X_t)] \tag{5.31}$$

where $l_t(0) = \int_0^t I(X_s = 0) ds$ is defined to be the local time spent at 0 of a particle executing a random walk (X_s) starting from x.

Proof. In order to solve the system, we must first solve the homogeneous equation. Now, writing $\mathbb{E}_{\eta}[\eta_t(x)] = \psi(t,x)$, we rewrite the homogeneous form of the system as follows:

$$\frac{d}{dt}\psi_h(t,x) = A_{RW}\psi_h(t,x) - \lambda \delta_{x,0}\psi_h(t,x) = (A_{RW} - \lambda \delta_{x,0})\psi_h(t,x) \tag{5.32}$$

Here, $\psi_h(t,x)$ is the homogeneous solution. Both A_{RW} and $\lambda\delta_{x,0}$ are linear operators, where A_{RW} is the standard generator of the exclusion process. Since it has the same form as the generator for the random walk process, we can apply the same technique for the solution operator. So, the associated semigroup is of the form $e^{t(A_{RW} - F(x))}$ with $F(x) = \lambda \delta_{x,0}$. To interpret this solution operator, we can use the very advanced Feynman-Kac theorem on our ODE to get the following solution:

$$\psi_h(t,x) = e^{t(A_{RW}-F(x))} = \mathbb{E}_x^{RW}\left[\eta_0(X_t)e^{-\lambda\int_0^t dsI(X_s=0)}\right] \tag{5.33}$$

As noted before, we use a shorthand notation $l_0(t) = \int_0^t I(X_s = 0)ds$ for the expression in the exponent, which is called the local time spent at 0 during $[0,t]$. The notation $I(X_s = 0)$ is just another way to write the Kronecker delta $\delta_{X_s,0}$. We proceed to solve the inhomogeneous equation

$$\frac{d}{dt}\psi(t,x) = A\psi(t,x) + \lambda\delta_{x,0} \tag{5.34}$$

where $A = A_{RW} - \lambda\delta_{x,0}$, using variation of constants, which is the same method as we used before. So, suppose that the solution of (5.34) is given by $\psi(t,x) = e^{tA}\chi_t$ with χ_t again being an unspecified function. One can check using the same procedure that we need to integrate the equation $e^{tA}\frac{d\chi_t}{dt} = \lambda\delta_{x,0}$. This leads to

$$\chi_t = \int_0^t e^{-sA}\lambda\delta_{x,0}ds + \psi_h(0,x) \tag{5.35}$$

and solution of the total equation

$$\psi(t,x) = \int_0^t e^{(t-s)A}\lambda\delta_{x,0}ds + e^{tA}\psi_h(0,x) \tag{5.36}$$

Using a quick substitution $u = t - s$ and applying the previous Feynman-Kac form, we can simplify this solution to:

$$\begin{aligned}\psi(t,x) &= \int_0^t e^{sA}\lambda\delta_{x,0}ds + e^{tA}\psi_h(0,x)\\ &= \mathbb{E}_x[e^{-\lambda l_t(0)}\eta_0(X_t)] + \lambda\int_0^t \mathbb{E}_x[e^{-\lambda l_s(0)}\delta_{X_s,0}]ds\end{aligned} \tag{5.37}$$

To make sense out of the last term, we observe the following. If we set $f(x) = 1$ in the ODE $\frac{d}{dt}(\tilde{S}(t)f(x)) = \tilde{S}(t)(A_{RW} - \lambda\delta_{x,0})f(x)$, we can set $A_{RW}f(x) = 0$ by using the expression previously derived. So, we are left with $\frac{d}{dt}\tilde{S}(t) = \tilde{S}(t)(-\lambda\delta_{x,0}1)$, where 1 is the 1 function. Applying Feynman-Kac, we see:

$$\frac{d}{dt}\tilde{S}(t)1 = -\lambda\tilde{S}(t)\delta_{x,0}1 = -\lambda e^{tA}\delta_{x,0}1 = -\lambda\mathbb{E}_x[e^{-\lambda l_t(0)}\delta_{X_t,0}] \tag{5.38}$$

We can conveniently substitute this back into our previous solution. So,

$$\begin{aligned}\psi(t,x) &= \mathbb{E}_x[e^{-\lambda l_t(0)}\eta_0(X_t)] - \int_0^t \frac{d}{dt}\tilde{S}(s)1ds\\ &= \mathbb{E}_x[e^{-\lambda l_t(0)}\eta_0(X_t)] - (\tilde{S}(t)1 - \tilde{S}(0)1)\\ &= \mathbb{E}_x[e^{-\lambda l_t(0)}\eta_0(X_t)] + 1 - \mathbb{E}_x[e^{-\lambda l_t(0)}]\\ &= 1 + \mathbb{E}_x[e^{-\lambda l_t(0)}(\eta_0(X_t) - 1)]\end{aligned} \tag{5.39}$$

Since we can have at most one particle per site, it is actually more enlightening to pose the final solution as:

$$\psi(t,x) = 1 - \mathbb{E}_x[e^{-\lambda l_t(0)}(1-\eta_0(X_t)] \tag{5.40}$$

□

5.5.2 Symmetric exclusion with a sink

Let the following system be the system associated to symmetric exclusion with a sink added to the origin.

$$\frac{d}{dt}\mathbb{E}_\eta[\eta_t(x)] = \begin{cases} \mathbb{E}_\eta\left[\eta_t(x+1) + \eta_t(x-1) - 2\eta_t(x)\right] & \text{if } x \neq 0 \\ \mathbb{E}_\eta\left[\eta_t(1) - \eta(0)\right] - \lambda\eta_t(0) & \text{if } x = 0 \end{cases} \tag{5.41}$$

Theorem 23. Suppose a totally symmetric exclusion process is given on $\mathbb{Z}^d$ with initial configuration $\eta = \{0,1\}^{\mathbb{Z}^d}$ and a sink absorbing particles at rate λ if they are present at the origin. That is, if the time evolution satisfies ODE (5.41), then the time evolution of η is given by

$$\psi(t,x) = \mathbb{E}_\eta[\eta_t(x)] = \mathbb{E}_x[e^{-\lambda l_t(0)}\eta_0(X_t)] \tag{5.42}$$

where $l_t(0) = \int_0^t I(X_s = 0)ds$ is defined to be the local time spent at 0 of a particle executing a random walk (X_s) starting from x.

Proof. This fortunately turns out to be a homogeneous equation. Setting $\mathbb{E}_\eta[\eta_t(x)] = \psi(t,x)$, we can rewrite the systems as:

$$\frac{d}{dt}\psi(t,x) = (A_{RW} - \lambda\delta_{x,0})\psi(t,x) \tag{5.43}$$

This ODE has solution $\psi(t,x) = e^{tA}\psi(0,x)$ with $A = A_{RW} - \lambda\delta_{x,0}$. We can invoke the Feynman-Kac theorem to see that

$$\psi(t,x) = \mathbb{E}_x[e^{-\lambda\int_0^t I(X_s=0)ds}\eta_0(X_t)] = \mathbb{E}_x[e^{-\lambda l_t(0)}\eta_0(X_t)] \tag{5.44}$$

□

5.5.3 Exclusion with multiple sources and sinks

We will now take the previous two exclusion cases to the extreme, where we place arbitrarily many sources and sinks throughout the grid $\mathbb{Z}^d$. Recall the ODE we posed for this system with $\mathcal{B}$ the source set and $\mathcal{S}$ the sink set:

$$\frac{d}{dt}\psi(t,x) = A_{RW}\psi(t,x) - \sum_{y\in\mathcal{S}}\lambda\delta_{x,y}\psi(t,x) + \sum_{y\in\mathcal{B}}\lambda\delta_{x,y}(1-\psi(t,x)) \tag{5.45}$$

The corresponding theorem states the solution of this system:

Theorem 24. Let $\mathbb{Z}^d$ be a grid with an exclusion process on it. Now, let $\mathcal{S}$ and $\mathcal{B}$ be the sink and source collections respectively that also take part in the exclusion process through the system (4.14). Then, the solution to this system is given by:

$$\psi(t,x) = 1 + \mathbb{E}_x\left[e^{-\lambda\left(\sum_{y\in\mathcal{S}\cup\mathcal{B}} l_t(y)\right)}\eta_0(X_t)\right] - \mathbb{E}_x\left[e^{-\lambda\sum_{y\in\mathcal{B}} l_t(y)}\right] \tag{5.46}$$

where $l_t(y) = \int_0^t I(X_s = y)ds$ is defined as the local time the random walk (X_t) spends at position y during $[0,t]$.

Proof. We can again apply the same techniques and split the equation in a homogeneous and inhomogeneous part. So, for the homogeneous equation, we have:

$$\frac{d}{dt}\psi_h(t,x) = A_{RW}\psi_h(t,x) - \sum_{y\in\mathcal{S}} \lambda\delta_{x,y}\psi_h(t,x) - \sum_{y\in\mathcal{B}} \lambda\delta_{x,y}\psi_h(t,x) = A\psi_h(t,x) \tag{5.47}$$

The solution to this equation is $\psi_h(t,x) = e^{tA}\psi_h(0,x)$. Using the Feynman-Kac theorem, we can express the solution operator as:

$$\psi_h(t,x) = \mathbb{E}_x\left[e^{\int_0^t V(X_s)ds}\eta_0(X_t)\right] \tag{5.48}$$

Here, $V(x) = -\left(\sum_{y\in\mathcal{S}} \lambda\delta_{x,y} + \sum_{y\in\mathcal{B}} \lambda\delta_{x,y}\right)$. Using linearity of the integral, we can also express the integral in the exponent as a sum of local times. So, let $l_t(y) = \int_0^t I(X_s = y)$ be the local time spent during $[0,t]$ at location y. Therefore, we can reform the solution as

$$\psi_h(t,x) = \mathbb{E}_x\left[e^{-\lambda\left(\sum_{y\in\mathcal{S}} l_t(y) + \sum_{y\in\mathcal{B}} l_t(y)\right)}\eta_0(X_t)\right] \tag{5.49}$$

To solve the inhomogeneous equation, we again use variation of constants which results in the following solution:

$$\psi(t,x) = e^{tA}\psi_h(0,x) + \sum_{y\in\mathcal{B}} \lambda\int_0^t e^{sA}\delta_{x,y}ds \tag{5.50}$$

Since we are only left with inhomogeneous source contributions, consider the ODE

$$\frac{d}{dt}\tilde{S}(t)f(x) = \tilde{S}(t)\left(A_{RW} - \lambda\sum_{y\in\mathcal{B}}\delta_{x,y}\right)f(x) \tag{5.51}$$

If we choose $f(x) = 1$, then we are left with:

$$\frac{d}{dt}\tilde{S}(t)1 = -\lambda\sum_{y\in\mathcal{B}}\tilde{S}(t)\delta_{x,y}1 = -\lambda\sum_{y\in\mathcal{B}} e^{tA'}\delta_{x,y}1 \tag{5.52}$$

Here, $A' = A_{RW} - \lambda\sum_{y\in\mathcal{B}}\delta_{x,y}$. Using this result, we can simplify our solution as:

$$\begin{aligned}\psi(t,x) &= e^{tA}\psi_h(0,x) - \int_0^t \frac{d}{ds}\tilde{S}(s)1ds = e^{tA}\psi_h(0,x) - (\tilde{S}(t) - \tilde{S}(0))1 \\ &= 1 + e^{tA}\psi_h(0,x) - e^{tA'}1 \\ &= 1 + \mathbb{E}_x\left[e^{-\lambda\left(\sum_{y\in\mathcal{S}\cup\mathcal{B}} l_t(y)\right)}\eta_0(X_t)\right] - \mathbb{E}_x\left[e^{-\lambda\sum_{y\in\mathcal{B}} l_t(y)}\right]\end{aligned} \tag{5.53}$$

□

So, this is our final solution for the cases where we add arbitrarily many sources and sinks to the grid. Of course, we made some hidden assumptions on the way, since we have changed sums and integrals on the way. Specifically, the place where it could lead to conflict is the function $V(x) = -\left(\sum_{y\in\mathcal{S}} \lambda\delta_{x,y} + \sum_{y\in\mathcal{B}} \lambda\delta_{x,y}\right)$. Since Feynman-Kac shows that we need to take the integral $\int_0^t V(X_s)ds$, we can only express this integral as a sum of local times if we allow the integrals and sums to change under certain conditions. This is however not a problem, since any grid $\mathbb{Z}^d$ is countable, so the sums are countable, too.

Concluding remarks

As we have set the stage for movement towards a coupled TASEP, we look back on our analysis. The simplest model one can come up with to model the random walks particles execute in biological systems, is by considering the random walks to be independent of each other. By adding a source point, the process becomes somewhat more complicated and the resulting distribution is attenuated by the source term.

We have also found that Poisson measures stay invariant over time, whether a source emits particles at the origin at rate λ or not. This also points towards possible configurations that must be satisfied if we start a process from the empty configuration. In this case, it turns out that we always end up with a Poisson distribution of particles if we start from the zero configuration on a grid.

With this information, we moved towards the more complicated case of exclusion, where only one particle is allowed per site at maximum. Consider a completely symmetric situation, we were able to reuse the tools developed before in combination with the Feynman-Kac theorem. Looking at a grid $\mathbb{Z}^d$ and placing arbitrarily many sources and sinks we were able to determine whether the grid overflowed with particles after some time or not. Specifically, we found the mathematically surprising result that when we add a source and sink to the origin, the grid will overflow with particles under the assumption that the local time tends to infinity as $t \to \infty$.

Unfortunately, this was not enough to build a TASEP model for transcription. Since asymmetric exclusion causes us to lose the previous tools, that have been developed. To properly describe transcription using TASEP, one needs to thoroughly study TASEP first and might therefore be explored in a further project. Nevertheless, the mathematical models that have been explored are very fascinating themselves from a mathematical point of view.

5.6 Acknowledgements

I would first of all like to thank my supervisors Frank Redig and Martin Depken for their advice and kindness. They provided a very interesting problem to work on and even though I felt intimidated at first, they helped with the struggles and gave direction to the project. I would also like to thank my parents for their kindness, love and support for this project. Even though you don't understand what I'm doing, it makes me incredibly happy that you stand behind me and support me every single day. I would also like to

thank my friends for their patience and the silly jokes about this project. Lastly, I am also thankful to God for letting me take on this project. It has taught me a lot about research in mathematics and physics, but also a lot about myself. I feel incredibly blessed and thankful to have had the opportunity to work on this problem.

Bibliography

[1] R. K. Zia, J. J. Dong, and B. Schmittmann, "Modeling Translation in Protein Synthesis with TASEP: A Tutorial and Recent Developments," *Journal of Statistical Physics*, vol. 144, no. 2, pp. 405–428, 2011.

[2] F. Crick, *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. Atheneum press, 1968.

[3] C. Xu, S. Hu, and X. Chen, "Artificial cells: from basic science to applications," *HHS Public Access*, 2016.

[4] B. Amin Espah, A. Channarasappa, and H. M. Salis, "Translation rate is controlled by coupled trade-offs between site accessibility, selective rna unfolding and sliding at upstream standby sites," vol. 42, p. 2646–2659, 2014.

[5] F. Redig, "Basic techniques in interacting particle systems," 2014.

[6] T. Seppalainen, "Translation Invariant Exclusion Processes," 2003.

[7] J. K. Hunter, "Notes on Partial Differential Equations," pp. 1–24, 2009.

[8] R. Schilling and L. Partzsch, *Brownian Motion - An introduction to Stochastic Processes*. De Gruyter, 2001.

[9] R. Philips, J. Kondev, J. Theriot, and H. G. Garcia, *Physical biology of the cell*.

[10] S. Clancy and W. Brown, "Translation: DNA to mRNA to Protein," *Nature Education*, 2008.