On the general problem of the Pearson-Rayleigh Random Walk

Richard Oliveira

September 6, 2024

Contents

Chapter 1

Introduction

The Pearson-Rayleigh Random Walk, is an old problem that dates back to August 1905. Karl Pearson initially stated the problem with the following quote [**3**]: "A man starts from a point O and walks a yards in a straight line; he then turns through any angle whatever and walks another a yards in a second straight line. He repeats this process n times. I require the probability that after n of these stretches he is at distance between r and $r + \delta r$ from his starting point O." Karl Pearson, 1905.

Pearson wanted to study the movement of a mosquito in a dense forest, and in some way predict the endpoint of the mosquito after taking n uniformly oriented steps of varying length. Since, the insect, indeed does not have a well-specified goal in mind while taking these steps, the assumption of uniformly oriented directions is appropriate and almost logical. We will see that this assumption is the key feature in Random Walks of this kind, as it serves to greatly simplify the majority of calculations presented in this work. Another, and straightforward assumption in the modeling of this problem, is the fact that the walk takes place in continuous space. The coupling of uniformity and continuity gives rise to a "nice" geometrical structure in the walk. Conversely, when the continuity assumption is interchanged with discontinuity, we will see that computations quickly become intractable in higher dimensions. The choice of the walk taking place in discrete space, is not necessarily a natural one. Rather, it aims to study a different problem that is loosely connected to topics in Computer Science such as Random Walks on Graphs for example [**5**]. Furthermore, the choice of a discrete space reveals difficulties that the continuous counter part does not. In this work, we will explore different ways of interpreting the Discrete Pearson-

Rayleigh Random Walk. The exploration of this problem will reveal insights on how the model can be simplified and additionally serve as ”food for thought” for future direction in this field.

Chapter 2

Mathematical Preliminaries

In this section we briefly introduce a series of important definitions and properties associated to Probability Spaces and Random Vectors, such that all the Random Variables that will be mentioned later in this work are properly defined.

2.1 Probability Space

The Probability Space is completely defined by the triple $(\Omega, \mathcal{F}, \mu)$. Where Ω is the Sample Space, $\mathcal{F}$ is a σ-algebra of events ($\mathcal{F} \subset \Omega$) and μ is a probability measure on Ω.

The collection of events $\mathcal{F}$ is said to be a σ-algebra over Ω if it satisfies the following:

1. $\varnothing \in \mathcal{F}$
2. An event $E \in \mathcal{F}$ implies $E^C \in \mathcal{F}$
3. For countably many subsets $E_1,\ ..., E_k,\ ...\ \in \mathcal{F}$ we have that $\bigcup_{k=1}^{\infty} E_k \in \mathcal{F}$

A probability measure μ is a mapping of the form $\mu :\ \mathcal{F} \to [0, 1]$ that satisfies the following properties:

1. $\mu(\Omega) = 1$
2. For countably many subsets $\{E_k\}_{k=1}^{\infty} \in \mathcal{F}$ that are mutually disjoint (i.e. $E_i \cap E_j = \varnothing,\ \forall i \neq j$) we have:

$$\mu(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k) \tag{2.1}$$

Furthermore we list three important properties associated to σ-algebras:

1. μ and the trivial set $\{\varnothing, \Omega\}$ are always σ-algebras over Ω

2. For any collection of σ-algebras $\{\mathcal{F}_i\}_{i\in I}$ (I is the index set) over Ω we have that $\bigcap_{i\in I} \mathcal{F}_i$ is also a σ-algebra over Ω.

3. For any collection A of subsets of Ω we define the σ-algebra generated by A to be:
$$\sigma(A) = \bigcap_{A\subseteq \mathcal{F}_i} \mathcal{F}_i \tag{2.2}$$
In other words it is all the σ-algebras $\mathcal{F}_i$ that contain A.

The third property associated to σ-algebras allows us to define a so called Borel σ-algebra on $\mathbb{R}$, sometimes defined as $\Omega = \mathbb{R}$, $\mathcal{B}(\mathbb{R})$ or simply $\mathcal{B}$. By definition we have that:

$$\mathcal{B} = \sigma(\{(-\infty, a] \mid a \in \mathbb{R}\}) \tag{2.3}$$

Which is the smallest σ-algebra that contains all these open intervals. Next, we mention an important property associated to σ-algebra, which is the closeness under countable intersection. For any σ-algebra $\mathcal{F}$ on Ω, we have that for any collection of events $\{A_k\}_{k=1}^{\infty}$ in $\mathcal{F}$ the following holds:

$$\bigcap_{k=1}^{\infty} A_k \in \mathcal{F} \tag{2.4}$$

Finally, we mention some important properties related to the probability measure μ on $(\Omega, \mathcal{F})$:

1. **Monoticity**: if $A \subseteq B$ for A, $B \in \mathcal{F}$, then $\mu(A) \leq \mu(B)$.

2. **Sub-Additivity**: For events $A_1,\ A_2\ ... \in \mathcal{F}$ (not necessarily mutually exclusive) we have $\mu(\bigcup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \mu(A_k)$.

3. **Continuity from below** $\forall\ A_1 \subseteq A_2 \subseteq\ ...\ A_k \subseteq\ ...$ we have:
$$\lim_{k\to\infty} \mu(A_k) = \mu(\bigcup_{k=1}^{\infty} A_k) \tag{2.5}$$

4. **Continuity from above**: $\forall\ A_1 \supseteq A_2 \supseteq\ ...\ A_k \supseteq\ ...$ we have:

$$\lim_{k\to\infty} \mu(A_k) = \mu(\bigcap_{k=1}^{\infty} A_k) \tag{2.6}$$

2.2 Random Variables

We begin by defining a Random Variable in $\mathbb{R}$. A Random variable X on a Probability Space $(\Omega, \mathcal{F}, \mu)$ is a function with mapping $X : \Omega \to \mathbb{R}$ and with pre-image $X^{-1}((-\infty, a]) \in \mathcal{F}\ \forall\ a \in \mathbb{R}$.

In general, we say that X is a Random Variable on $(\Omega, \mathcal{F}, \mu)$ **if and only if**, for any Borel set B the following holds:

$$X^{-1}(B) \in \mathcal{F},\ \forall\ B \in \mathcal{B}(\mathbb{R}) \tag{2.7}$$

We say that $\mathbf{V} = (X_1,\ ...\ , X_n)$ is a Random Vector on $(\Omega, \mathcal{F}, \mu)$ if X_1, ..., X_n are random variables on that space.

We define the expected value of the Random vector $\mathbf{V}$ as:

$$\begin{aligned} E\left\{\mathbf{V}\right\} &= \left(E\left\{X_1\right\},\ ...\ , E\left\{X_n\right\}\right)^T \\ &= \left(\int x_1 p_{X_1}(x_1) dx_1,\ ...\ , \int x_n p_{X_n}(x_n) dx_n\right)^T \end{aligned} \tag{2.8}$$

The Covariance matrix of the Random Vector $\mathbf{V} = (X_1, X_2)$ (for $n = 2$) is defined as:

$$\begin{aligned} \Sigma_V &= E\left\{\left(\mathbf{V} - E\left\{\mathbf{V}\right\}\right)\left(\mathbf{V} - E\left\{\mathbf{V}\right\}\right)^T\right\} \\ &= \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Var(X_2) \end{pmatrix} \end{aligned} \tag{2.9}$$

Additionally, we note that Σ_V is symmetric, and positive semi-definite (PSD), meaning that:

$$\Sigma_V^T = \Sigma_V \ (\textbf{Symmetry})$$

$$\mathbf{x}^T \Sigma_V \mathbf{x} \geq 0 \ \forall \mathbf{x} \in \mathbb{R}^n \ (\textbf{PSD})$$

The Cumulative Distribution Function (cdf) of a random variable X is a function with mapping $F_X : \mathbb{R} \to [0, 1]$ defined as:

$$F_X(a) = \mu(X^{-1}((-\infty, a])) = \mu(\{\omega \in \Omega | X(\omega) \leq a\}) \tag{2.10}$$

Additionally, the cdf of any random variable X satisfies the following:

1. $\lim_{a\to-\infty} F_X(a) = \varnothing$
2. $\lim_{a\to\infty} F_X(a) = 1$
3. **Right-continuous**: $\lim_{x\to a^+} F_X(x) = F_X(a)$
4. **Non-decreasing**: $x \leq y, F_X(x) \leq F_X(y)$

2.2.1 Discrete Random Variables

We say that X is a Discrete Random variable if there exist a countable set $S \subseteq \mathbb{R}$ such that:

$$\mu(X \in S) = \mu(X^{-1}(S)) = 1$$

Furthermore, without loss of generality, we say that for any set $s \in S$ the associated measure $\mu(X = s) > 0$, leads us in defining the support of X, which is $S_X = \{s_1, s_2, \ ...\}$ such that:

$$\sum_{i=1}^{\infty} \mu(X = s_i) = 1$$

A Probability Mass Function(pmf), associated with X is given by:

$$\rho_X(x) = \mu(X = x)$$

2.2.2 Continuous Random Variables

We say that X is a continuous Random Variable if:

$$\mu(X \leq \alpha) = F_X(\alpha) = \int_{-\infty}^{\alpha} f_X(x)dx$$

for some non-negative function $f_X(.)$ which we define as a Probability Density Function(pdf).

The expected value of X is defined as:

$$E\{X\} = \int_{-\infty}^{\infty} x\ f_X(x)dx$$

Given a function $g : \mathbb{R} \to \mathbb{R}$, we have that the expected value of the Random Variable $Y = g(X)$ is given by the Law of the Unconscious Statistician (LOTUS):

$$E\{Y\} = E\{g(X)\} = \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

If a Random vector $\mathbf{X} = (X_1, ..., X_n)$ is distributed according to a multivariate Gaussian, with mean $\boldsymbol{\mu}$ and covariance matrix Σ then it has pdf:

$$p_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{2\pi|\Sigma|}} \exp\left((\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right) \tag{2.11}$$

Another interesting continuous random variable, which will be important in this context, is the Random Vector $\boldsymbol{Z} \in \mathbb{R}^d$ that is Uniformly distributed inside a Ball of radius r, more formally this is written as:

$$\boldsymbol{Z} \sim \frac{1}{\pi r^2},\ ||\boldsymbol{x}||_2^2 \leq r^2 \tag{2.12}$$

With expected value:

$$E\{\boldsymbol{Z}\} = \frac{d}{d+2} r^2$$

It is important to note, that random vectors of the form (2.12) have characteristic functions that are functions of only the modulus of the frequency variable $\boldsymbol{\omega}$. In other words:

$$E\left\{\exp\left(i\boldsymbol{w}^T\boldsymbol{z}\right)\right\} = \Phi_{\boldsymbol{Z}}(\boldsymbol{\omega}) = \Phi_{\boldsymbol{Z}}(||\boldsymbol{\omega}||_2) \tag{2.13}$$

Chapter 3

Random Walk in Continuous Space

The mathematical formulation of the Random Walk in the continuous domain has been introduced by Franceschetti in 2007 [**1**]. Before then, not much work has been done on this topic. We will see that after this period, a lot of work will be published. Authors such as Le Caër, investigated the Walk by means of a generalization [**2**]. Other authors such as Orsingher *et al* .[**4**], used Franceschetti's work as inspiration for topics such as Random Flights of varying length in higher dimensions. The work of Franceschetti serves as a stepping stone for Pearson-Rayleigh Random Walks in the continuous domain. A nice, but also unintuitive result arises from his work. The walker, after taking exactly 3 steps on the plane, can find itself uniformly anywhere inside a disc of radius equal to the total path length. This interesting result still remains not explained on why it happens, a geometric proof might serve as further insight on the behaviour of the Walk. Furthermore, the author suggests a *if-then* statement, at the end of section 3.1, which serves as a starting point for proving the non intuitive result just mentioned.

3.1 Exponentially Distributed Step Sizes

As mentioned in the introduction, Franceschetti introduces a model that describes the movement of a particle undergoing elastic collision. On the one dimensional real line $\mathbb{R}$ (and in higher dimensions), we place uniformly distributed obstacles. Once the particle hits one of these obstacles (in $\mathbb{R}$), it chooses to go either right or left with equal probability and then take a step of exponential size. Thus modeling the idea that by taking large enough step sizes, the probability of reaching a random coordinate on this line becomes less likely. The Walk becomes interesting when we study the probability of reaching a random coordinate after n steps while conditioned on the total path length [**1**].

On $\mathbb{R}$, the Walk is completely defined by the following sequence of n Random Vectors:

$$\mathcal{W}^n = \left\{ \begin{pmatrix} X_1 \\ L_1 \end{pmatrix}, \ \dots \ , \begin{pmatrix} X_n \\ L_n \end{pmatrix} \right\} \tag{3.1}$$

Where, each random vector is independent and identically distributed (i.i.d.). The probability of reaching a random coordinate $X \in \mathbb{R}$ with a stochastic step of Euclidean length $L \in \mathbb{R}^+$ is given by [**1**]:

$$f_{X,L}(x, l) = \frac{\beta}{2} e^{-\beta|x|} \delta(l - |x|) \tag{3.2}$$

The presence of the Delta function in (3.2) ensures that the random coordinate X is reached with non-zero probability. Now, by iterating the process of taking $m = n - 1$ steps, where each step has joint density with law (3.2), Bayes Theorem allows us to obtain the probability of reaching the endpoint of the Walk conditioned on the total path length l after m steps [**1**]:

$$f^m_{X|L}(x|l) = \frac{\pi}{(2\pi)^2 (2l)^m} e^{-\beta(x-l)} (g^{(1)}(x, l) + g^{(2)}(x, l)) \tag{3.3}$$

Where:

$$g^{(1)}(x, l) = 2\pi e^{-\beta(l-x)} \sum_{k=1}^{m} \frac{(m + k)! x^{m-k} (l - x)^{k-1}}{k!(m - k)!(k - 1)! 2^k}, \ l > x \tag{3.4}$$

$$g^{(2)}(x, l) = 2\pi x^m \delta(l - x) \tag{3.5}$$

After taking $m = 1$ steps, the conditional density is essentially a linear combination of two Delta's placed at the boundary $[-l, l]$ and a Uniform density. As Franceschetti points out, the one dimensional case does not give insights on what happens in higher dimensions but it serves as a glimpse on what could happen in dimensions ≥ 2.

On the plane, the Walk is defined by the following sequence:

$$\mathcal{V}^n = \left\{ \begin{pmatrix} X_1 \\ Y_1 \\ L_1 \end{pmatrix}, \ \dots \ , \begin{pmatrix} X_n \\ Y_n \\ L_n \end{pmatrix} \right\} \tag{3.6}$$

The setup is exactly the same as in the one dimensional case. The joint law between X, Y and L slightly differs. The probability of reaching a random coordinate pair $\boldsymbol{R} = (X, Y) \in \mathbb{R}^2$ with a stochastic step of Euclidean length $L \in \mathbb{R}^+$ is given by [**1**]:

$$f_{\boldsymbol{R},L}(\boldsymbol{r}, l) = \frac{\beta}{2\pi|\boldsymbol{r}|} e^{-\beta|\boldsymbol{r}|} \delta(l - |\boldsymbol{r}|) \tag{3.7}$$

Where $r = |\boldsymbol{r}| = \sqrt{x^2 + y^2}$ is the standard Euclidean length. By following the same exact analytical procedure as in the one dimensional case, we obtain the following conditional density after m steps [**1**]:

$$f^m_{\boldsymbol{R}|L}(\boldsymbol{r}|l) = \frac{m}{2\pi l^2}(1 - \frac{r^2}{l^2})^{\frac{m-2}{2}} \ , \ l > r, \ m = 1, 2, \ \dots \tag{3.8}$$

By taking exactly three steps on the plane ($m = 2$) the walker can find itself uniformly anywhere inside a disc of radius l. In other words the conditional vector $\boldsymbol{R}|L$ is uniformly distributed inside a circle of radius l. This interesting and rather surprising result is difficult to explain intuitively. By a process of reasoning in an "analytic" manner, (3.8) arises by surprise.

Franceschetti, then derives a necessary condition for the uniform density to arise in higher dimensions in n steps. The condition is given by [**1**]:

$$n = \frac{2(d + 2)}{d} - 1 \tag{3.9}$$

Where d corresponds to the number of dimensions. If the term $2(d + 2)/d$ is not an integer and not equal to $n + 1$, the uniform density does not arise in the Walk.

I presume that the generalization argument offered by Le Caër [**2**] was an attempt in explaining why the endpoint conditioned on the total path length is uniformly distributed after n steps.

The author also attempted a generalization argument. This will be the topic of the next subsection.

Decomposition of a Uniformly Distributed Random Vector

In the previous section, we have seen that in $d = 2$, when the Walker takes exactly 3 steps on the plane, it can find itself uniformly anywhere inside a disc of radius l.

As mentioned previously, a geometric proof can serve as way to explain why this happens when taking exactly 3 steps on the plane. When starting from the Random Vector **Z** uniformly distributed inside a circle of radius l, we wish to decompose **Z** into the sum of exactly 3 i.i.d. spherically symmetrical random vectors, whose distribution is exponential. More formally, we are looking to prove the following statement:

$$\begin{aligned} &\textbf{If} \;\; \boldsymbol{Z} = \boldsymbol{R}_1 + \boldsymbol{R}_2 + \boldsymbol{R}_3 | L \in \mathbb{R}^2 \sim \frac{1}{\pi l^2},\; z_1^2 + z_2^2 \leq l^2 \\ &\textbf{Then} \;\; \boldsymbol{R}_i \sim \frac{\beta}{2\pi ||\boldsymbol{r}_i||_2} \exp\left(-\beta ||\boldsymbol{r}_i||_2\right),\; \boldsymbol{r}_i \perp \boldsymbol{r}_j \;\forall\; i \neq j,\; i, j \in \mathbb{I} = \{1, 2, 3\} \end{aligned} \tag{3.10}$$

3.1.1 Laplacian Distributed Coordinates in $\mathbb{R}$

In this section the author offers a way to generalize the Walk in $\mathbb{R}$ by looking at the Walk in a slightly different way. We first notice the similarities between the Exponential density and the Laplacian density in $\mathbb{R}$. Now, instead of asking ourselves what is the probability of reaching a random coordinate $X \in \mathbb{R}$ with a step of Euclidean length $L \in \mathbb{R}^+$ with uniform orientation, we ask, what is the probability of reaching a Laplacian distributed coordinate $X \in \mathbb{R}$ with mean μ and variance $2b^2$ with a step of Euclidean length $L \in \mathbb{R}^+$ equal to deviation of the coordinate x from its mean. Essentially we are describing the Walk by omitting the fact that we are taking uniformly oriented directions, thus in some way eliminating the "nice" geometry in the Walk that the uniform direction offers. The uniformity in the direction of the walk, is already encoded in the Laplacian model, thus we are simply asking ourselves what is the probability of reaching a Laplacian distributed coordinate on the real line $\mathbb{R}$.

The model is given by:

$$f_{X,L}(x, l) = \frac{1}{2b} e^{\frac{|x-\mu|}{b}} \delta(l - |x - \mu|) \tag{3.11}$$

By setting the parameters $b = 1/\beta$ ($\beta > 0$), and $\mu = 0$, we have that (3.10) is exactly equal to (3.2). Furthermore, by having $\mu \neq 0$, the Walk doesn't actually start from the origin, instead the starting point is completely defined by μ. Now, by keeping $\mu \neq 0$ and $b > 0$, we perform the same computations as in section 3.1 to obtain the conditional density of the endpoint.

The Walk, in this case, is completely defined by the following sequence:

$$\mathcal{L}^n = \left\{ \begin{pmatrix} X_1 \\ L_1 \end{pmatrix}, \;\cdots\;, \begin{pmatrix} X_n \\ L_n \end{pmatrix} \right\} \tag{3.12}$$

Where $X_i \sim$ Laplace (μ, b) and $(X_i, L_i) \perp (X_j, L_j)\ \forall\ i \neq j$ (i.i.d.). Since we are interested in the density of the endpoint, we iterate the process by summing up all the random vectors in $\mathcal{L}^n$. Since, each random vector is i.i.d. the following relation is well known:

$$f^m_{X,L}(x, l) = f_{\sum_i (X_i, L_i)}(x, l) = f^{\star(m)}_{X,L}(x, l) \tag{3.13}$$

We let the Fourier Transform (Characteristic function) of $f_{X,L}(x, l)$ be $\Phi(\omega, \eta)$. By definition we have that:

$$\begin{aligned}
\Phi(\omega, \eta) &= \frac{1}{2b} \int_{-\infty}^{\infty} e^{\frac{-|x-\mu|}{b}}\ e^{i\omega x}\ e^{-i\eta|x-\mu|} dx \\
&= \frac{1}{2b} \left\{ \int_{-\infty}^{\mu} e^{\frac{x-\mu}{b}}\ e^{-i\omega x}\ e^{i\eta x}\ e^{-i\eta\mu} dx + \int_{\mu}^{\infty} e^{\frac{-(x-\mu)}{b}}\ e^{-i\omega x}\ e^{-i\eta x}\ e^{i\eta\mu} dx \right\} \\
&= \frac{1}{2b} \left\{ e^{-i\eta\mu} e^{-\frac{\mu}{b}} \int_{-\infty}^{\mu} e^{x(\frac{1}{b}+i\eta)}\ e^{-i\omega x}\ dx + e^{i\eta\mu} e^{\frac{\mu}{b}} \int_{\mu}^{\infty} e^{-x(\frac{1}{b}+i\eta)}\ e^{-i\omega x} dx \right\} \\
&= \frac{1}{2b} \left\{ \frac{e^{-\mu(\frac{1}{b}+i\eta)}}{(\frac{1}{b}+i\eta) - i\omega} + \frac{e^{\mu(\frac{1}{b}+i\eta)}}{(\frac{1}{b}+i\eta) + i\omega} \right\},\ \frac{1}{b} > 0
\end{aligned} \tag{3.14}$$

By the properties of the Fourier Transform, we know that:

$$\Phi^m(\omega, \eta) = \mathcal{F}\left\{ f^m_{X,L}(x, l) \right\} = (\Phi(\omega, \eta))^{m+1} \tag{3.15}$$

Therefore, the inverse Transform is given by:

$$f^m_{X,L}(x, l) = \frac{1}{(2b)^{m+1}} \left(\sum_{k=0}^{m+1} \binom{m+1}{k} \mathcal{F}^{-1} \left\{ \frac{e^{-\mu(m+1-k)(\frac{1}{b}+i\eta)}}{[(\frac{1}{b}+i\eta) - i\omega]^{m+1-k}} \right\} \star \mathcal{F}^{-1} \left\{ \frac{e^{\mu k(\frac{1}{b}+i\eta)}}{[(\frac{1}{b}+i\eta) + i\omega]^{k}} \right\} \right)$$

The sum that runs from $k = 0$ to $m + 1$ arises due to the Binomial Theorem. To evaluate the inverse Transform we will need to compute the sum for when $k = 0$ and when $k = m + 1$. We look at the case when $k = 0$:

$$\begin{aligned}
&\mathcal{F}^{-1}\left\{\frac{e^{-\mu(m+1)(\frac{1}{b}+i\eta)}}{[(\frac{1}{b}+i\eta)-i\omega]^{m+1}}\right\} \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\eta\ e^{i\eta l} e^{\mu(m+1)(\frac{1}{b}+i\eta)}\ \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\ e^{i\omega x}\frac{1}{[(\frac{1}{b}+i\eta)-i\omega]^{m+1}} \qquad (3.16)\\
&= e^{\frac{1}{b}(x-\mu(m+1))}\frac{(-x)^m}{m!}\delta(l+[x-\mu(m+1)]) \\
&= \delta^{(1)}
\end{aligned}$$

For when $k = m + 1$, we get a similar expression:

$$\begin{aligned}
&\mathcal{F}^{-1}\left\{\frac{e^{-\mu(m+1)(\frac{1}{b}+i\eta)}}{[(\frac{1}{b}+i\eta)+i\omega]^{m+1}}\right\} \\
&= e^{-\frac{1}{b}(x+\mu(m+1))}\frac{(x)^m}{m!}\delta(l-[x+\mu(m+1)]) \qquad (3.17)\\
&= \delta^{(2)}
\end{aligned}$$

By combining the results from (3.15) and (3.16) we are left with:

$$\begin{aligned}
&f^m_{X,L}(x,l) = \frac{1}{(2b)^{m+1}}\left(\delta^{(1)}+\delta^{(2)}\right)\ + \\
&\frac{1}{(2b)^{m+1}}\ \left(\sum_{k=1}^{m}\binom{m}{k}\ \mathcal{F}^{-1}\left\{\frac{e^{-\mu(m+1-k)(\frac{1}{b}+i\eta)}}{[(\frac{1}{b}+i\eta)-i\omega]^{m+1-k}}\right\}\star\mathcal{F}^{-1}\left\{\frac{e^{\mu k(\frac{1}{b}+i\eta)}}{[(\frac{1}{b}+i\eta)+i\omega]^{k}}\right\}\right) \qquad (3.18)
\end{aligned}$$

Since the distribution is symmetric around the mean μ, we have that:

$$\delta^{(1)} = \delta^{(2)}$$

Now, for when $1 \leq k \leq m$ we need to compute the convolution in (3.17). Thus giving us the final result for the joint law of X and L:

$$
f^m_{X,L}(x,l) = \frac{1}{(2b)^{m+1}} \left(2\delta^{(2)}\right) +
$$

$$
\frac{1}{(2b)^{m+1}} \left(\frac{1}{2} e^{-\frac{1}{b}(2l-x)} \sum_{k=1}^{m} \binom{m}{k} \frac{1}{(m-k)!(k-1)!} \left(e^{\frac{\mu}{b}}(l-x-\mu(m+1))\right)^{m-k} \left(e^{\frac{\mu}{b}}(l-\mu(m+1))\right)^{k-1} \right) \tag{3.19}
$$

Now it is easy to check that for $m = 0$, $b = 1/\beta$ and $\mu = 0$ we arrive at the joint law of (3.2) for $x > 0$. We are interested in computing the conditional density. By Bayes Theorem we have that:

$$
f^m_{X|L}(x|l) = \frac{f^m_{X,L}(x,l)}{f^m_L(l)} \tag{3.20}
$$

By marginalization we have that:

$$
\begin{aligned}
f_L(l) &= \int_0^\infty f_{X,L}(x,l)\ dl \\
&= \frac{1}{2b} e^{-\frac{l}{b}}
\end{aligned}
$$

With Fourier Transform equal to:

$$
\Phi(\eta) = \frac{1}{2b}\ \frac{1}{(\frac{1}{b} + i\eta)},\ \ \frac{1}{b} > 0
$$

Finally we arrive at:

$$
f^m_L(l) = \frac{1}{(2b)^{m+1}} \frac{l^m}{m!} e^{-\frac{l}{b}} \tag{3.21}
$$

By substituting (3.18) and (3.20) into (3.19) we obtain the conditional density of the endpoint, conditioned on the total path length:

$$
\begin{aligned}
f^m_{X|L}(x|l) = \frac{m!}{(2l)^m} &\left\{ 2\frac{(2x)^m}{m!} \delta(l - [x + \mu(m+1)]) e^{-\frac{1}{b}(x - l + \mu(m+1))} \right\} + \\
&\frac{m!}{(2l)^m} 2 \sum_{k=1}^{m} \binom{m}{k} \frac{1}{(m-k)!(k-1)!} \frac{[(l-x) - \mu(m-1)]^{m-k}}{2^{-k}} \\
&\times \frac{[(l+x) - \mu(m-1)]^{k-1}}{4^k}
\end{aligned} \tag{3.22}
$$

Again, it is easy to check that (3.21) becomes (3.3) by setting $\mu = 0$ and $b = 1/\beta$. We notice that this expression does indeed share some similarities with (3.3). It is a linear combination of Delta functions placed at the boundary $[-(l - \mu(m+1)), (l - \mu(m+1))]$ and the Uniform density. We notice in fact that the position of the Delta functions, for $\mu \neq 0$ depends on m. In the case of 3.3 ($\mu = 0$), we have that the Delta functions will always be placed at $[-l, l]$. Again, this model does not give much insights on what happens in dimensions ≥ 2.

Chapter 4

Random Walk in Discrete Space

In this section, we will discuss the Discrete analog of the continuous Pearson-Rayleigh Random Walk. The author has modelled two different versions of the Discrete Pearson-Rayleigh Random Walk. One that is, in some way, the Discrete analog for exponentially distributed step sizes and one that models the Walk with Poisson distributed step sizes. Since the Walk takes place in the Discrete domain, we say that the Walk takes place in the following space:

$$\Lambda_d = \{\boldsymbol{x} = (x_1, \ \dots \ , x_d) \mid x_i \in \mathbb{Z}, \ \forall \ i = 1, \ \dots \ , d \ \} \tag{4.1}$$

We will look at the cases for $d = 1$ and for $d = 2$. Even though, majority of the computations for the one-dimensional case are non-trivial, we still obtain valid conditional densities for any number of steps. For the two-dimensional case however, obtaining conditional densities still remains unsolved due to the intractability of the computations. This type of Walk is completely new in the literature, thus offering a new direction of research in this field.

4.1 Geometrically Distributed Step Sizes

We begin our study of the Discrete Pearson-Rayleigh Random Walk in $d = 1$ (i.e. in $\mathbb{Z}$) with Geometrically distributed step sizes. This model, is a straightforward discretization of the model that Franceschetti introduces [**1**]. A more sophisticated version of this Walk will be studied in section 4.2.

4.1.1 Geometric in One dimension

Since the Geometric distribution is a probability mass function (pmf) that models the number of failures before the first success occurs with some probability p. We look at the Manhattan distance of a random coordinate $x \in \mathbb{Z}$ as $|x|$ being the "number of hops" the Walker takes on $\mathbb{Z}$ before the coordinate x is reached (which is seen as a first success). In other words the successful event is denoted as $E = \{$ "$x \in \mathbb{Z}$ is reached" $\}$. It follows, that the probability, of reaching a random coordinate $x \in \mathbb{Z}$ with a stochastic step of Integer length $L \in \mathbb{Z}^+$ and with uniform orientation, is given by:

$$g_{X,L}(x,l) = \frac{1}{2}(1-p)^{|x|}\ p\ \delta(l-|x|),\ p \in [0,1] \tag{4.2}$$

As before, we are interested in the joint pmf of X and L after m steps, therefore the *modus operandi* is exactly the same as it was in the previous sections. We start with the Discrete-Time Fourier Transform of (4.2). By definition, it is given by:

$$\begin{aligned}
\phi(\omega,\eta) &= \sum_{x=-\infty}^{\infty}\sum_{l=0}^{\infty}\frac{p}{2}(1-p)^{|x|}\ \delta(l-|x|)\ e^{-i\omega x}\ e^{-i\eta l} \\
&= \frac{p}{2}\sum_{x=-\infty}^{\infty}(1-p)^{|x|}\ e^{-i\omega x}\ e^{-i\eta|x|} \\
&= \frac{p}{2}\left(\sum_{x=0}^{\infty}(1-p)^{x}e^{-i\omega x}e^{-i\eta x} + \sum_{x=-\infty}^{0}(1-p)^{-x}e^{-i\omega x}e^{i\eta x}\right) \\
&= \frac{p}{2}\left(\sum_{x=0}^{\infty}[(1-p)\ e^{-i(\eta+\omega)}]^{x} + \sum_{x=0}^{\infty}[(1-p)e^{-i(\omega-\eta)}]^{x}\right) \\
&= \frac{p}{2}\left(\frac{1}{1-(1-p)e^{-i(\omega+\eta)}} + \frac{1}{1-(1-p)e^{i(\omega-\eta)}}\right),\ p \neq 0
\end{aligned} \tag{4.3}$$

The last equality in (4.3) was obtained by using the Geometric Series. The sum that runs from $-\infty$ to ∞ can be decomposed into two sums that run from 0 to ∞ due to the symmetry of the Walk. By the property of Fourier Transforms we have:

$$(\phi(\omega,\eta))^{m+1} = \frac{p^{m+1}}{2^{m+1}}\left(\frac{1}{1-(1-p)e^{-i(\omega+\eta)}} + \frac{1}{1-(1-p)e^{i(\omega-\eta)}}\right)^{m+1} \tag{4.4}$$

By the Binomial Theorem (4.4) can be written as:

$$(\phi(\omega,\eta))^{m+1} = \frac{p^{m+1}}{2^{m+1}}\left(\sum_{k=0}^{m+1}\binom{m+1}{k}\frac{1}{[1-(1-p)e^{-i(\omega+\eta)}]^{m+1-k}}\frac{1}{[1-(1-p)e^{i(\omega-\eta)}]^{k}}\right) \tag{4.5}$$

The inverse Transform can be easily applied due to its Linearity:

$$g^{m}_{X,L}(x,l) = \frac{p^{m+1}}{2^{m+1}}\left(\sum_{k=0}^{m+1}\binom{m+1}{k}\mathcal{F}^{-1}\left\{\frac{1}{[1-(1-p)e^{-i(\omega+\eta)}]^{m+1-k}}\right\}\star\mathcal{F}^{-1}\left\{\frac{1}{[1-(1-p)e^{i(\omega-\eta)}]^{k}}\right\}\right) \tag{4.6}$$

Similarly with what we did in the previous sections, we start by evaluating the sum in (4.6) by looking at the cases for when $k=0$, $k=m+1$ and for when $1 \le k \le m$.

For when $k=0$ we have that:

$$\begin{aligned}&\mathcal{F}^{-1}\left\{\frac{1}{[1-(1-p)e^{-i(\omega+\eta)}]^{m+1}}\right\}\\ &= \frac{(l+m)!}{l!\;m!}(1-p)^{l}\delta(x-l)\end{aligned} \tag{4.7}$$

For when $k=m+1$ we have that:

$$\begin{aligned}&\mathcal{F}^{-1}\left\{\frac{1}{[1-(1-p)e^{-i(\eta-\omega)}]^{m+1}}\right\}\\ &= \frac{(l+m)!}{l!\;m!}(1-p)^{l}\delta(x+l)\end{aligned} \tag{4.8}$$

Next, we perform the Discrete-Time convolution for when $1 \le k \le m$:

$$\mathcal{F}^{-1}\left\{\frac{1}{[1-(1-p)e^{-i(\omega+\eta)}]^{m+1-k}}\right\}\star\mathcal{F}^{-1}\left\{\frac{1}{[1-(1-p)e^{i(\omega-\eta)}]^{k}}\right\}$$

$$=\left(\frac{l+(m-k)!}{l!\ (m-k)!}(1-p)^l\delta(x-l)\right)\star\left(\frac{l+k-1!}{l!\ (k-1)!}(1-p)^l\delta(x+l)\right)$$

$$=\sum_{\theta}\sum_{\gamma}\frac{\gamma+m-k!}{\gamma!\ m-k!}(1-p)^{\gamma}\delta(\gamma-\theta)(1-p)^{l-\gamma}\frac{l-\gamma+k-1!}{l-\gamma!\ k-1!}\delta((x-\theta)+(l-\gamma))$$

$$=\frac{1}{m-k!\ k-1!}\sum_{\theta}\frac{\theta+m-k!}{\theta!}(1-p)^{\theta}\frac{l-(x-\theta+l)+k-1!}{(l-(x-\theta+l))!}(1-p)^{l-(x-\theta+l)}\delta(2\theta-(x+l))$$

$$=\begin{cases}0 & \forall\ \theta\in\mathbb{Z}\\ \neq 0 & \text{if } (x+l)\%2=0\end{cases}$$

$$=\frac{1}{m-k!\ k-1!}\frac{x+l+m-k!}{x+l!}(1-p)^{x+l}\frac{l+k-1!}{l!}(1-p)^{l}\sum_{\theta}\delta(\theta-(x+l)),\ (x+l)\%2=0 \tag{4.9}$$

The convolution yields a result that inherits the nature of the discrete behaviour of the Walk on the discrete line. The condition that $x+l$ must be divisible by 2 such that we have non-zero probability of reaching certain sites on $\mathbb{Z}$ makes intuitive sense. Meaning that if we are in fact taking "jumps" of integer length on $\mathbb{Z}$ there will be some locations that will never be reachable with probability 1, since we are walking on Λ_1. This behaviour can be verified by computer simulation in MatLab in section 4.2.1-Transition of the Poisson Mass Function. Independent of the model used, the unreachability of certain locations on Λ_1 is a key feature of Walks of this kind.

Combining (4.9) with (4.6), we obtain:

$$\begin{aligned}g^{m}_{X,L}(x,l)=&\frac{p^{m+1}}{2^{m+1}}\left(\frac{(l+m)!}{l!\ m!}(1-p)^l[\delta(x-l)+\delta(x+l)]\right)\\&+\frac{p^{m+1}}{2^{m+1}}\left(\frac{(1-p)^{2l+x}}{l!\ x+l!}\sum_{k=1}^{m}\binom{m}{k}\frac{x+l+(m-k)!}{m-k!}\frac{l+(k-1)!}{k-1!}\right),(x+l)\%2=0\end{aligned} \tag{4.10}$$

Next, we marginalize the joint pmf, in order to obtain the following:

$$g_L(l) = p(1-p)^l \tag{4.11}$$

After m steps, (4.11) becomes:

$$g_L^m(l) = p^{m+1}\frac{l+m!}{l!\ m!}(1-p)^l \tag{4.12}$$

By using Bayes Theorem, we obtain the density of the endpoint conditioned on the total path length:

$$g_{X|L}^m(x|l) = \frac{1}{2^{m+1}}\left([\delta(x-l)+\delta(x+l)] + \frac{(1-p)^{x+l}\ m!}{x+l!\ l+m!}\sum_{k=1}^{m}\binom{m}{k}\frac{x+l+(m-k)!\ l+(k-1)!}{m-k!\ (k-1)!}\right)$$
$$(x+l)\%2 = 0, p \neq 0 \tag{4.13}$$

For $m = 1$ (two steps), we obtain a linear combination of Delta functions placed at the boundaries of the interval $[-l, l]$ weighted by $1/4$ and a term that almost resembles a Geometric pmf:

$$g_{X|L}^1(x|l) = \frac{1}{4}[\delta(x+l)+\delta(x-l)] + \frac{1}{4(l+1)}(1-p)^{x+l},\ (x+l)\%2 = 0, p \neq 0$$

In general the one-dimensional case does not give insights on what happens in higher dimensions. The author has created a model in dimension equal to 2 and it will be discussed in the following section.

4.1.2 Geometric in Two Dimensions

In the two dimensional case, the model will vary slightly. When starting from the origin, the Walker chooses a direction uniformly at random and takes a step of Geometrically distributed length. Since we are in Λ_2, we have precisely $4|\boldsymbol{x}|_1$ locations that can be reached. Accordingly, we let the probability of reaching a random coordinate $\boldsymbol{X} = (X, Y) \in \Lambda_2$ in a stochastic step of integer length $L \in \mathbb{Z}^+$ with uniform orientation as:

$$g_{\boldsymbol{X},L}(\boldsymbol{x}, l) = \frac{1}{4(|\boldsymbol{x}|)}(1-p)^{|\boldsymbol{x}|}p\ \delta(l - |\boldsymbol{x}|),\ p \in [0,1] \tag{4.14}$$

Where $|\boldsymbol{x}| = |\boldsymbol{x}|_1$ denotes the Manhattan distance. The main difficulty in obtaining a conditional density for the two-dimensional case, is that the Discrete-Time Fourier Transform of (4.14) is intractable computationally. This intractability, in the continuous case, is avoided since the l_2 distance is easily parameterized by using polar coordinates. In fact, this change of coordinates highlights the "nice" geometry that the Walk exhibits in the continuous case. Conversely, the discrete case is not easily parameterizable since the Manhattan distance has a graph of a diamond. Therefore, due to this lack of "nice" geometry in the discrete case we are left with computing the following sums, as a consequence of the Discrete-Time Fourier Transform definition:

$$
\begin{aligned}
\phi(\gamma,\omega,\eta) &= \sum_{x=-\infty}^{\infty}\sum_{y=-\infty}^{\infty}\sum_{l=0}^{\infty}\frac{1}{4(|\boldsymbol{x}|)}(1-p)^{|\boldsymbol{x}|}p\ \delta(l-|\boldsymbol{x}|)e^{-i\gamma x}\ e^{-i\omega y}\ e^{-i\eta l} \\
&= \sum_{x=-\infty}^{\infty}\sum_{y=-\infty}^{\infty}\frac{1}{4(|x|+|y|)}(1-p)^{|x|+|y|}p\ e^{-i\gamma x}\ e^{-i\omega y}\ e^{-i\eta|x|}\ e^{-i\eta|y|} \\
&= \sum_{x,y=0}^{\infty}\frac{1}{4(x+y)}(1-p)^{x+y}p\ e^{-i\gamma x}\ e^{-i\omega y}\ e^{-i\eta x}\ e^{-i\eta y} \\
&+ \sum_{x,y=0}^{\infty}\frac{1}{4(x+y)}(1-p)^{x+y}p\ e^{i\gamma x}\ e^{-i\omega y}\ e^{-i\eta x}\ e^{-i\eta y} \\
&+ \sum_{x,y=0}^{\infty}\frac{1}{4(x+y)}(1-p)^{x+y}p\ e^{i\gamma x}\ e^{i\omega y}\ e^{-i\eta x}\ e^{-i\eta y} \\
&+ \sum_{x,y=0}^{\infty}\frac{1}{4(x+y)}(1-p)^{x+y}p\ e^{-i\gamma x}\ e^{i\omega y}\ e^{-i\eta x}\ e^{-i\eta y}
\end{aligned} \tag{4.15}
$$

The above sums lead to intractable computations, therefore leaving the problem unsolved.

4.2 Poisson Distributed Step Sizes

As mentioned in the introduction, when the steps sizes are Poisson distributed we are essentially counting the number of times the following event occurs: $W = \{\text{"One hop occurs"}\}$. A "hop" occurs when the Walker makes a "jump" to adjacent discrete sites once in either direction in order to reach a discrete site on Λ_d at a rate that is described by the number of hops over a single step $\lambda/(m+1)$. A step, on the other hand, is the number of hops that have occurred when reaching the discrete site x. In other words, the number of hops are counted by computing the Manhattan distance of the reachable discrete site x.

4.2.1 Poisson in One Dimension

In the one-dimensional case, the setting is exactly the same as in the previous section. The pmf has rate λ since $m = 0$ (one step). Therefore we have:

$$h_{X,L}(x,l) = \frac{1}{2}\lambda^{|x|}\frac{e^{-\lambda}}{|x|!}\ \delta(l - |x|),\ m = 0 \tag{4.16}$$

Again, we follow the same analytical procedure that is outlined in the previous sections. The Discrete-Time Fourier Transform of (4.16) is:

$$\Omega(\omega,\eta) = \frac{1}{2}e^{-\lambda}\left(\exp\left(\lambda e^{i(\omega-\eta)}\right) + \exp\left(\lambda e^{-i(\omega+\eta)}\right)\right) \tag{4.17}$$

Over m steps, the joint law between X and L is given by:

$$h^m_{X,L}(x,l) = \frac{e^{-\lambda(m+1)}}{2^{m+1}}\sum_{k=0}^{m+1}\binom{m+1}{k}\mathcal{F}^{-1}\left\{\exp\left(\lambda(m+1-k)e^{i(\omega-\eta)}\right)\right\}\star\mathcal{F}^{-1}\left\{\exp\left(\lambda k e^{-i(\omega+\eta)}\right)\right\} \tag{4.18}$$

When evaluating the sum at both $k = 0$ and $k = m+1$ we have the following inverse transforms for both $k = 0$ and $k = m+1$ respectively:

$$\mathcal{F}^{-1}\left\{\exp\left(\lambda(m+1-k)e^{i(\omega-\eta)}\right)\right\} = \frac{(\lambda(m+1))^l}{l!}\delta(x+l) \tag{4.19}$$

$$\mathcal{F}^{-1}\left\{\exp\left(\lambda k e^{-i(\omega+\eta)}\right)\right\} = \frac{(\lambda(m+1))^l}{l!}\delta(x-l) \tag{4.20}$$

For when $0 < k < m+1$ we obtain the following result (here we omit the computations of the Discrete-Time convolution since it is the exact same as in section 4.9):

$$\begin{aligned}&\left(\frac{(\lambda(m+1-k))^l}{l!}\delta(x+l)\right)\star\left(\frac{(\lambda k)^l}{l!}\delta(x-l)\right)\\&=\frac{(\lambda(m+1-k))^{l-x}}{l-x!}\frac{(\lambda k)^l}{l!},\ (x-l)\%2=0\end{aligned}\tag{4.21}$$

Next, by substituting (4.19), (4.20) and (4.21) into (4.18) we obtain the joint law between X and L after taking m steps on Λ_1:

$$h^m_{x,L}(x,l)=\frac{e^{-\lambda(m+1)}}{2^{m+1}}\left\{\frac{(\lambda(m+1))^l}{l!}[\delta(x-l)+\delta(x+l)]+\frac{\lambda^{2l-x}}{l-x!\ l!}\sum_{k=1}^{m}\binom{m}{k}(m+1-k)^{l-x}k^l\right\}\tag{4.22}$$

Now, since $L_i\perp L_j\ \forall i\neq j$ we have that:

$$h^m_L(l)=h_{\sum_i^n L_i}(l)=\frac{e^{-\lambda(m+1)}(\lambda(m+1))^l}{l!}\tag{4.23}$$

By Bayes rule we obtain the conditional pmf of the endpoint conditioned on the total path length:

$$\begin{aligned}h^m_{X|L}(x|l)&=\frac{1}{2^{m+1}}\left\{[\delta(x-l)+\delta(x+l)]+\frac{\lambda^{l-x}}{l-x!\ (m+1)^l}\sum_{k=1}^{m}\binom{m}{k}(m+1-k)^{l-x}k^l\right\}\\&(x-l)\%2=0,\ l>x\end{aligned}\tag{4.24}$$

We look at some special cases:

1. $m=0$ (**one step**): $\frac{1}{2}[\delta(x-l)+\delta(x+l)]$. When conditioned on the total length path and the Walker can only take one step, the Walker is limited in only reaching the boundaries of the interval $[-l,l]$ as expected.
2. $m=1$ (**two steps**): $\frac{1}{4}\left\{\delta(x-l)+\delta(x+l)+\frac{\lambda^{l-x}}{l-x!\ 2^l}\right\}$. Here, not only we can reach the boundaries of the interval $[-l,l]$ in two steps with 1/4 weight, but we have the contribution of the Poisson term $\lambda^{l-x}e^{-\lambda}/(l-x)!$ with weight $1/(2^l e^{-\lambda})$.

3. $m = 2$ (**three steps**):

$$\frac{1}{8}\left\{\delta(x-l)+\delta(x+l)+\frac{\lambda^{l-x}}{l-x!\;3^l}[2\times 2^{l-x}+2^l]\right\}$$

$$=\frac{1}{8}\left\{\delta(x-l)+\delta(x+l)\right\}+\frac{1}{8}\left\{2\times\frac{(2\lambda)^{l-x}}{l-x!\;3^l}+\left(\frac{2}{3}\right)^l\frac{\lambda^{l-x}}{l-x!}\right\}$$

In this case, we start seeing how the sum in (4.24) contributes to the overall conditional pmf. It serves as a linear combination of terms that add up to the usual probability of reaching the interval $[-l, l]$, for which, as m increases it becomes less likely for the Walker to reach the boundaries.

As $m \to \infty$ the Walker gets entangled around the origin and stays there almost surely, as expected. This entanglement around the origin will be confirmed with simulation in the next section.

Transition of the Poisson Mass Function

In this section, we analyze the conditional pmf obtained in section 4.2.1 and observe it's behaviour as we vary the number of steps m and the rate parameter λ. When conditioned on the total path length l, the probability of reaching the endpoint $x \in \Lambda_1$ after $m = n-1$ steps is given by:

$$h^m_{X|L}(x|l)=\frac{1}{2^{m+1}}\left\{[\delta(x-l)+\delta(x+l)]+\frac{\lambda^{l-x}}{l-x!\;(m+1)^l}\sum_{k=1}^{m}\binom{m}{k}(m+1-k)^{l-x}k^l\right\}$$

$$(x-l)\%2=0,\ \ l>x$$

As mentioned at the beginning of this section the term $\lambda/(m+1)$ describes the number of hops the Walker takes on Λ_1 over the number of steps $m+1$. We can rewrite (4.24) as:

$$h^m_{X|L}(x|l) = \frac{1}{2^{m+1}}\left\{\delta(x-l)+\delta(x+l)\right\}$$

$$+\frac{1}{2^{m+1}}\left\{\frac{(m+1)^{-x}}{e^{-\lambda/(m+1)}}\frac{e^{-\lambda/(m+1)}\left(\frac{\lambda}{m+1}\right)^{l-x}}{l-x!}\sum_{k=1}^{m}\binom{m}{k}(m+1-k)^{l-x}k^l\right\}$$

$$(x-l)\%2=0,\ l>x \tag{4.25}$$

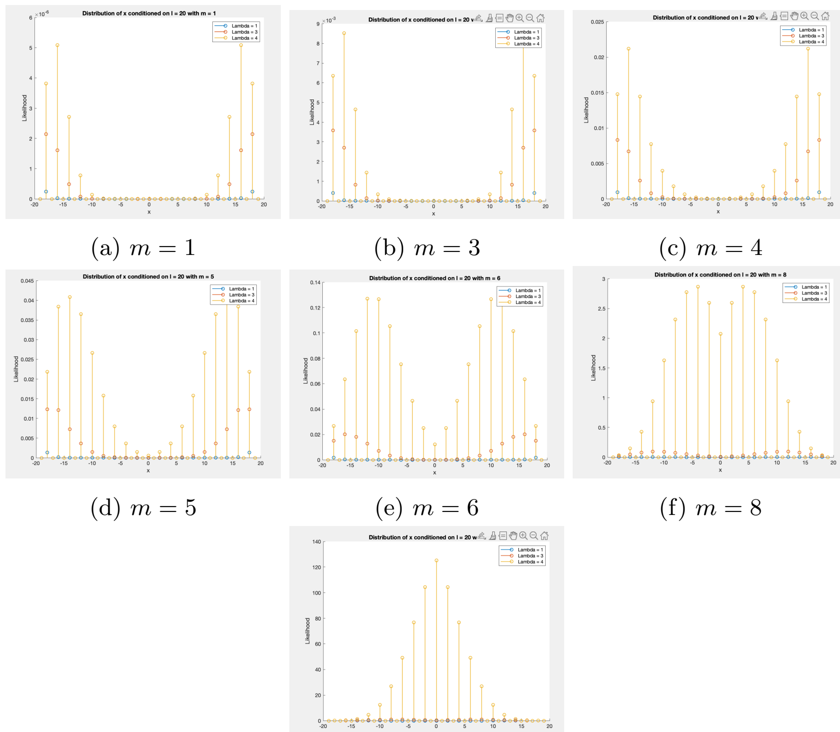
By rewriting (4.24) in the above way, we can clearly see the influence of $\lambda/(m+1)$ on the entire conditional mass function.

The author shows a series of different plots, where we can observe how the mass function changes when varying the number of steps for different values of λ. We notice, that by varying m and λ for some fixed total path length l, it is not clear when the mass functions starts to change drastically for some m.

First, we keep the rate low. Meaning that we keep the parameter λ small, and we increase the number of steps. We look at $\lambda = 1, 3, 4$ and we fix the total path length to be $l = 20$ for $m = 1, 3, 4, 5, 6, 8, 10$ (Figure 4.1).

We notice that as soon as the rate decreases, the mass function starts to get concentrated around the origin. Thus suggesting that as soon as the number of steps is greater than the set parameter λ, a transition starts to occur.

On the other hand, if we keep the rate high, we have the following plots for $\lambda = 5, 15, 20$ with a fixed length path of $l = 50$ for $m = 1, 2, 3, 4, 5$ (Figure 4.2).



(a) $m = 1$ (b) $m = 3$ (c) $m = 4$

(d) $m = 5$ (e) $m = 6$ (f) $m = 8$

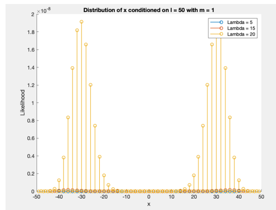
(g) $m = 10$

Figure 4.1: Plots low rate for $m = 1, 3, 4, 5, 6, 8, 10$

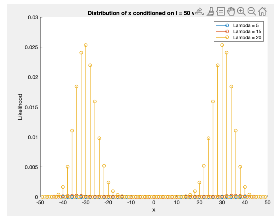
For when the rate is high, we notice something different. We have a sudden transition in a drastically different plot at an early step, specifically for $m = 3$ (Figure 4.2). Thus suggesting that even if the rate is already high, by only making an extra additional step, we have a very different mass function. This unexpected result leads to searching for the number of steps required such that a drastic transition occurs. This problem can be formulated by searching for the optimal m that maximizes the dissimilarity between two conditional mass functions at m and $m + 1$ respectively. The following maximization problem captures this idea:

$$m^{\star} = \arg\max_{m} \; D_{KL}\left(h^{m}_{X|L}(x|l) \; || \; h^{m+1}_{X|L}(x|l)\right) \tag{4.26}$$

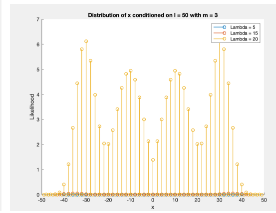
Although (4.26) seems like a natural question to pose, an analytical solution for $m^{\star}$ is not possible to compute. On the other hand, a numerical result may be feasible, but not too useful for our purposes.



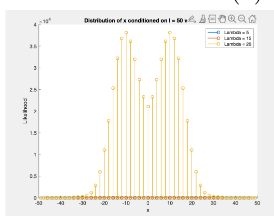
(a) $m = 1$



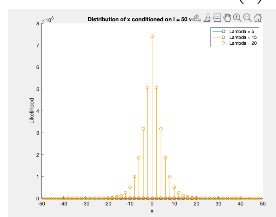
(b) $m = 2$



(c) $m = 3$



(d) $m = 4$



(e) $m = 5$

Figure 4.2: Plots high rate for $m = 1, 2, 3, 4, 5$

4.2.2 Poisson in Two Dimensions

In the two-dimensional case, we find ourselves in the same situation as in section 4.1.2, where the computation of the Discrete-Time Fourier Transform of the joint law between $\boldsymbol{X} = (x, y) \in \Lambda_2$ and $L \in \mathbb{Z}^+$ is intractable. Furthermore, the model that describes this joint law for $m = 0$ is given by:

$$h_{\boldsymbol{X},L}(\boldsymbol{x}, l) = \frac{1}{4|\boldsymbol{x}|!} \frac{e^{-\lambda}\lambda^{|\boldsymbol{x}|}}{|\boldsymbol{x}|!} \delta(l - |\boldsymbol{x}|) \tag{4.27}$$

We *conjecture*, that on Λ_2, we can only reach $4|\boldsymbol{x}|$ discrete sites on the boundaries of a rhomboid when conditioned on a integer length l in one step. When the number of steps increases, the behaviour of the Walk remains unknown. A possible approach in solving this kind problem is using second moments. More specifically, studying the second moment of the endpoint conditioned on L could be useful:

$$
\begin{aligned}
&E\left\{\begin{pmatrix} R_1 \\ R_2 \end{pmatrix}^T \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \mid L\right\} \\
&= E\left\{\left(\sum_{i=1}^{n}\begin{pmatrix} X_i \\ Y_i \end{pmatrix}\right)^T \left(\sum_{i=1}^{n}\begin{pmatrix} X_i \\ Y_i \end{pmatrix}\right) \mid L\right\} \\
&= E\left\{(X_1 + \ \ldots \ + X_n)^2|L\right\} + E\left\{(Y_1 + \ \ldots \ + Y_n)^2|L\right\} \\
&= \sum_{i=1}^{n} E\left\{X_i^2|\ L\right\} + \sum_{i=1}^{n} E\left\{Y_i^2|\ L\right\} + \sum_{i\neq j} E\left\{X_i X_j|\ L\right\} + \sum_{i\neq j} E\left\{Y_i Y_j|L\right\}
\end{aligned}
\tag{4.28}
$$

Although, this alternative approach may seem valid, an intractable computation hides in this simple formulation.

Chapter 5

Conclusion

In this work we have introduced and discussed the general problem of the Pearson-Rayleigh Random Walk in both the continuous and discrete domain. Starting from its inception in 2007, we have seen that a lot of work has been published in continuous space, but there is little to almost no work in the discrete space for walks of this kind. Inspired by the works of Le Caër, the author suggests a reformulation of the walk with Exponentially distributed step sizes in one dimension as a way to further generalize the problem. Additionally, the author computed conditional mass functions of the endpoint of the walk in one dimension for both Geometric and Poisson distributed step sizes. The one dimensional discrete walks reveal how certain locations on Λ_1 are unreachable with probability 1 and how the conditional mass functions are equal to linear combinations of Delta's placed at the boundary of the walk and polynomials that are functions of the number of steps m and total path length l. Although the computations in one dimension are non trivial, they do not reveal any insights on what happens in dimensions ≥ 2. On Λ_1 it is not clear when a transition occurs when the number of steps increases. The author suggests a way of formulating the problem of finding the optimal m such that the mass functions are maximally dissimilar between consecutive steps. On Λ_2, the computations for finding conditional mass functions are intractable, therefore playing around with 2^{nd} moments can help study the behaviour of the walk without explicitly computing conditional mass functions in dimensions ≥ 2. Unfortunately, most of the work on Discrete Pearson-Rayleigh Random Walks still remain unknown due to the intractability of certain computations. Nevertheless, this paper can serve as a new research direction for walks of this kind and can be further studied for applications in Applied Probability Theory, Computer Science, Robotics and Communication over Networks.

Bibliography

[1] Massimo Franceschetti. *When a Random Walk of Fixed Length can Lead Uniformly Anywhere Inside a Hypersphere.* Journal of Statistical Physics, 2007.

[2] Gérard Le Caër. *A New Family of Solvable Pearson-Dirichlet Random Walks.* Journal of Statistical Physics, 2011.

[3] Karl Pearson. *The Problem of The Random Walk.* Nature, 1905

[4] Enzo Orsingher, Alessandro De Gregorio. *Random Flights in Higher Spaces.* Journal of Theoretical Probability, 2007

[5] L. Lovász. *Random Walks on Graphs.* Combinatorics, Paul Erdös is Eighty Keszthely, 1993