# *Asymptotic Behaviour of Moments in a Renewal Process Generated Random Walk*

TDK

*Author:*

*Ádám Bernát*

*Consultant:*

*Péter Bálint*

MŰEGYETEM 1782

Budapest University of Technology and Economics

Institute of Mathematics

**2022**

# Contents


**1 The Model** **2**

1.1 Motivation . . . . . . . . . . 2

1.2 Defining the Model . . . . . . . . . . 3

1.3 Why these distributions ? . . . . . . . . . . 4

1.4 Moments and Cumulants . . . . . . . . . . 5

**2 Renewal Theory** **6**

2.1 Overview . . . . . . . . . . 6

2.2 Laplace transform . . . . . . . . . . 8

**3 Renewal Process Calculations** **9**

3.1 Further Notations . . . . . . . . . . 9

**4 Fourth Cumulant** **14**

4.1 Second Moment . . . . . . . . . . 14

4.2 Fourth Moment . . . . . . . . . . 15

4.3 Fourth Moment $D$ event . . . . . . . . . . 16

4.4 Fourth Moment $B \cap C$ event - approximation . . . . . . . . . . 18

4.5 Fourth Moment $B \cap C$ event - dealing with the correctness of the approximation . . . . . . . . . . 25

**5 Summary and Outlook** **29**

# 1 The Model

## 1.1 Motivation

In this thesis we study a random walk generated by a renewal process that has polynomial tail distribution. The motivation comes from the study of chaotic billiards. This is an actively researched topic (see the monograph [6] for detail). A billiard ball is initialized on the boudaries of scatterers, typically placed in a periodic manner on $\mathbb{R}^2$. Some of the simpler cases are that they are non intersecting circles around the lattice points of $\mathbb{Z}^2$. The ball will travel at a constant speed and whenever it hits the scatterers it bounces off. The new direction can be obtained using the classical approach: angle of reflection = angle of incidence. Using convex scatterers after some time passed and many collisions later, one can consider the current position of the ball as random, independent of the starting point.

A system has the "infinite horizon" property if the possible travel lengths of a billiard ball, in between hitting scatterers is not bounded from above. A famous model, which might possess the infinite horizon property, is the Lorentz model where the billiard ball imitates a gas molecule bouncing around and hitting the scatterers. This model is discussed for example in [8] or [9].

A simple case, already mentioned above, is when the scatterers are circles around lattice points of $\mathbb{Z}^2$, which do not intersect and also possess a uniform radius (e.g $r = \frac{1}{4}$). The non-intersecting placement guarantees the infinite horizon property, as a molecule may bounce into a "corridor" where it could fly for an arbitrary long time, if the flight angle is almost aligned with the $x$ or $y$ axis .

We are going to study an even simpler version which corresponds to the analogy, that a particle is moving on $\mathbb{R}$ (on a straight line), changing directions randomly. Direction changes can be interpreted as collisions and the consecutive movement directions are either plus or minus one. That is it either continues onward or turns around (bounces back), with the two options being equally probable. Inter-collision times are i.i.d random variables so they (along with the collision times) give us a renewal process.We are going to consider such waiting times that have polynomial tail distributions. There are physical models where such distributions occur. In particular the infinite horizon case, where the tail distribution is $\mathbb{P}(L > x) \sim x^{-2}$, where L is the inter-collision time. Another example is a non-infinite horizon model (in [7] section 6.3), when instead of straight corridor motion long flights correspond to a more complicated repetitive bouncing motion, which causes different polynomial decay ($\mathbb{P}(L > x) \sim x^{-4}$).

Our model is simple, yet it captures the corridor movement, where exact vertical placement is negligible, all that is important are the hitting times and horizontal travel directions. The 2 dimensional billiard process is more complicated though, long flights in corridors, followed by a series of shorter flights. As a first approximation we could consider only the corridor motions and omit the short bouncing episodes.

Finally, let us note that distributions with polynomial decay occur in many areas of mathematics and applications, so studying them here might help us in other areas as well.

## 1.2 Defining the Model

Our model has the following setup. Let us consider a renewal process with arrival times $0 < T_0 < T_1 < T_2 < \dots$. Let $\xi(x)$ be the process, which has values 1 or -1 in between arrival times, with even probabilities (That is $+1/-1$ with probability $\frac{1}{2}$)

More precisely let $\xi_i$ $\forall i \in \mathbb{N}$ be i.i.d. random variables with probabilities: $\mathbb{P}(\xi_i = -1) = \mathbb{P}(\xi_i = 1) = \frac{1}{2}$. Now $\xi(x)$ is the process defined as $\xi(x) := \xi_i$ if $T_{i-1} < x < T_i$ $i \geq 1$ and $\xi(x) := \xi_0$ if $x < T_0$. Here $T_i$ are independent of $\xi_j$ for all $i, j \in \mathbb{N}$

Moreover, for some $a > 0$ and $\beta = 4$ fixed (could just as well be $\beta \geq 2$ but we are only focusing now on this special case), let $L_k = T_k - T_{k-1}$ be independent identically distributed random variables (the k'th inter-arrival time) with polynomial tail behavior $\mathbb{P}(L > x) \sim ax^{-\beta}$. For the desired results, in this thesis we assume a stronger condition, that is $\mathbb{P}(L > x) = ax^{-\beta} + \mathcal{O}(x^{-\beta-1})$ as $x \to \infty$ $(a > 0, \beta = 4)$ and let the $L_k$'s have distribution $L$. Also let $\mu = \mathbb{E}(L)$

Here, and throughout my thesis , $a(t) \sim b(t)$ means that :

$$\frac{a(t)}{b(t)} \to 1$$

as $t \to \infty$

Also $a(t) = \mathcal{O}(t^b)$ means that $a(t)$ is continuous and:

- $a(t)$ is bounded on $[0, 1]$

- $a(t) \leq C \cdot t^b$ for $t > 1$ and for some $C \in \mathbb{R}$

Now let us define the two quantities that we are interested in:

$$S_T := \int_0^T \xi(t)dt \tag{1}$$

and

$$X_k := \xi_k L_k, \quad S_n = \sum_{i=1}^{n} X_i \tag{2}$$

Where $k \geq 1$ and $n \geq 1$.
Also $X_k$ are all i.i.d. random variables, so let's call their common distribution $X$. Note that $|X|$ has the same distribution as $L$, yet the distribution of X is symmetric so $\mathbb{E}(X) = 0$ and $\mathbb{E}(X^3) = 0$
These two quantities are related. If $T = T_n$ is an arrival time for some $n \in \mathbb{N}$ then

$$(S_{T_n} =)S_T = S_n + \xi_0 T_0$$

### 1.3 Why these distributions ?

Now that the model has been defined, let us consider the reasons why we consider such distributions, with such a $\beta$ parameter.

**The original idea comes from [5] (Appendix A), where the case $\beta = 2$ is discussed**. In that case the Central Limit Theorem is not applicable, since the variance of X doesn't exist. The observation made there was, that using a modified Central limit theorem on the i.i.d. sum $S_n$ (2) it is true that (by [3] Section XVII.5, Theorem 2):

$$\frac{S_n}{\sqrt{n \cdot \ln(n)}} \to \mathcal{N}(0, a)$$

Here 'a' is the constant of the L distribution. From this Central Limit Theorem used for the i.i.d. sum $S_n$,we can conclude by the Law of Large Numbers that the renewal process also converges in distribution:

$$\frac{S_T}{\sqrt{T \ln(T)}} \to \mathcal{N}\left(0, \frac{a}{\mu}\right)$$

In contrast to the limit of the second moment:

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{S_T^2}{T\ln(T)}\right) = \frac{2a}{\mu}$$

Here this doubling effect is a direct analogue of what has been observed in infinite horizon Lorentz models ([7],[10]) (The $\mu$ is there because of the renewal process, but there we would expect to find $\frac{a}{\mu}$ ).

**We focus on the behavior of the 4th cumulant at** $\beta = 4$. Our motivation comes from [7] section 6.3. In this thesis we present a computation for this quantity, which is based on an approximation. The essence of the approximations is that certain quantities (see in Subsection 4.5) are replaced by their stationary values. Although we do not present a complete proof, we will also comment on the validity of this approximation. Also, in contrast to $\beta = 2$, here we are going to have the classical Central Limit Theorem , since the second moment of $X$ exists.

## 1.4 Moments and Cumulants

We are going to define the cumulants just like in [1] Chapter I §2

**Definition 1.1.** *Let $X$ be a random variable. It's logarithmic characteristic function is defined as $ln(\psi_X(t)) = \ln(\mathbb{E}(e^{itX}))$. If $ln(\psi_X(t))$ is $n$ times differentiable, then the $n$'th cumulant of $X$ is:*

$$\mathcal{C}_n(X) := \frac{1}{i^n}\left[\frac{d^n}{dt^n}\ln(\mathbb{E}(e^{itX}))\right]_{t=0}$$

It is easy to see that if the logarithmic characteristic function has a Taylor expansion around 0 then it is:

$$\ln(\mathbb{E}(e^{itX})) = \sum_{k=1}^{\infty}\frac{\mathcal{C}_k(X)}{k!}(it)^k \tag{3}$$

So the cumulants can be obtained from the Taylor expansion.
Some special cumulants are $\mathbb{E}(X) = \mathcal{C}_1(X)$ and $Var(X) = \mathcal{C}_2(X)$. Higher cumulants do not coincide with central moments but can be expressed in terms of them. In particular we have the following property:

**Lemma 1.2.** *Assume $X$ is a random variable, whose 4th moment exists and $\mathbb{E}(X) = 0$, then the 4th cummulant of $X$ can be expressed as:*

$$\mathcal{C}_4(X) = \mathbb{E}(X^4) - 3(\mathbb{E}(X^2))^2$$

*Proof.* See [1] Chapter I. □

Here it is worth noting that if X has standard normal distribution, then :

$$\ln(\mathbb{E}(e^{itX})) = -\frac{1}{2}t^2$$

So all cumulants (except the second , which is the variance) are 0. This is why cumulants, in some way, measure the "distance" from the standard normal distribution.

# 2 Renewal Theory

## 2.1 Overview

Let us recall the terminology and notation on renewal processes introduced in section 1.2.

Furthermore, for each $t > 0$ let's define: $Z_t = min\{m \in \mathbb{N} : t < T_m\}$ which is the number of arrivals before time $t$. Using this $Z_t$ random variable we can define some useful notations for ourselves. First one is called the **renewal function:**

$$M(t) = \mathbb{E}(Z_t) \tag{4}$$

**Further notation:** The derivative of M(t) will be denoted as $m(t)$ $(:= M'(t))$.
In addition, it is important to emphasize that these denote non-stationary quantities assuming there was an arrival at 0. Nevertheless as $t \to \infty$ $m(t) \to m^{stac}(t) = \frac{1}{\mu}$ that is $m(t)$ converges to the stationary case.

The second one is the **residual time:**

$$H^{stac} = H^{stac}(t) = T_{Z_t} - t \tag{5}$$

Where it is important to note that this is a distribution that assumes stationarity of the renewal process. So that is why actually, it is not a function of t. Whereas for non stationary cases we need a new notation (for $t \geq t_0$):

$$H(t - t_0) = T_{Z_{t-t_0}} - (t - t_0) \tag{6}$$

Where $t_0$ is an arrival time. After which we look at the process, as if it were a new starting point.

To ensure stationarity let $T_0$ have distribution:

$$\mathbb{P}(H^{stac}(t) > u) = \mathbb{P}(T_0 > u) \tag{7}$$

Also it follows from [2] Chapter 3 (Theorem 3.9) , (because of the sized biased distribution) , that:

$$\mathbb{P}(T_0 > t) = \frac{1}{\mu} \int\limits_t^\infty \mathbb{P}(L > x) dx$$

However we are going to need more specific functions, as stationarity does not always hold. Consider the function $\Phi(t, x)$ where t is going to be a given time after arrival happened at 0 and $x$ will be the required length of the residual. Then the **cumulative distribution functions of the residual** for a given $t$ is (see [3] Chapter XI):

$$1 - \Phi(t, x) = \mathbb{P}(H(t) > x) = \mathbb{P}(L > t + x) + \int\limits_{y=0}^{t} \left[\mathbb{P}(L > t + x - y)\right] dM(y)$$

As emphasized above, here M(y) is non-stationary as an arrival is assumed at time 0. From now on using the notation $F_L(x) = \mathbb{P}(L \leq x)$ for the cumulative distribution function of L:

$$\Phi(t, x) = \mathbb{P}(H(t) \leq x) = F_L(t + x) - \int\limits_{y=0}^{t} \left[1 - F_L(t + x - y)\right] dM(y) \tag{8}$$

For the sake of simplicity, for future references, we are going to use the formula derived for the density function. A simple consequence of the previous equation. Let:

$$\phi(t, x) = \frac{d}{dx} \Phi(t, x)$$

Then:

$$\phi(t,x) = f_L(t+x) + \int\limits_{y=0}^{t} [f_L(t+x-y)]\, dM(y) = f_L(t+x) + \int\limits_{y=0}^{t} [f_L(t+x-y)m(y)]\, dy \tag{9}$$

Where $f_L(s)$ is the probability density function of $L$ assuming it is absolutely continuous and $m(t)$ is the derivative of the renewal function.

Calculating $M(t)$ is not necessarily an easy task, however there is a very important "recursive" equation that $M(t)$ satisfies.(see [3] Chapter XI)

**The renewal equation:**

$$M(t) = F_L(t) + \int\limits_{s=0}^{t} M(t-s)dF_L(s) = F_L(t) + \int\limits_{s=0}^{t} M(t-s)f_L(s)ds \tag{10}$$

## 2.2 Laplace transform

These previous equations might seem nice as they have a relatively simple form. However solving them is not trivial. But using Laplace transform can help us derive an exact solution of the Laplace transform of the renewal function.

Recall that the Laplace transform of a function is defined the following way:

**Definition 2.1.** *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a real function then:*

$$\widehat{f}(s) = \mathcal{L}\{f(t)\}(s) = \int\limits_{t=0}^{\infty} f(t) \cdot e^{-st}dt$$

**Notation:** A function named after letter f will have Laplace transform whose name is the capped version of the same letter (e.g.: $f(t)$ has Laplace transform $\widehat{f}(s)$).

Let us mention the basic formulas that will be helpful to us.

Take the convolution of two functions $z(t) = x(t) * y(t)$ (supported on $t > 0$)

$$z(t) = \int\limits_{u=0}^{t} x(t-u) \cdot y(u)du$$

**The Laplace transform of the convolution function**:
(see [3] Chapter XIII Section 2)

$$\mathcal{L}\{z(t)\}(s) = \widehat{z}(s) = \widehat{x}(s) \cdot \widehat{y}(s) = \mathcal{L}\{x(t)\}(s) \cdot \mathcal{L}\{y(t)\}(s) \tag{11}$$

**Now let us apply this Laplace transform to our renewal equation (10):** Where first we differentiate both sides with respect to $t$

$$m(t) = f_L(t) + \int\limits_{s=0}^{t} m(t-s) f_L(s) ds$$

From this we get:

$$\mathcal{L}\{m(t)\}(r) = \widehat{m}(r) = \mathcal{L}\{f_L(t)\}(r) + \mathcal{L}\left\{\int\limits_{s=0}^{t} m(t-s) dF_L(s)\right\}(r) =$$

Using the convolution formula (11)

$$\widehat{m}(r) = \widehat{f_L}(r) + \widehat{m}(r) \cdot \widehat{f_L}(r)$$

From this:

$$\widehat{m}(r) = \frac{\widehat{f_L}(r)}{1 - \widehat{f_L}(r)} \tag{12}$$

# 3 Renewal Process Calculations

## 3.1 Further Notations

Let us introduce a notation $G_n(s)$ which will be useful for us. It is defined recursively in the following way:

**Definition 3.1.** *Let*

$$G_0(s) = \frac{1}{\mu}\mathbb{P}(L > s)$$

*The factor $\frac{1}{\mu}$ ensures that $G_0(s)$ is a density function supported on $\mathbb{R}^+$ (The density function of $T_0$). Now we can define:*

$$G_1(s) := \mathbb{P}(T_0 > s) = \frac{1}{\mu}\int_s^{\infty} \mathbb{P}(L > x) dx \qquad \textit{for} \quad s \geq 0$$

*If $G_n(s)$ is already defined then:*

$$G_{n+1}(s) := \int_s^\infty G_n(x)dx \qquad for \quad s \geq 0$$

Note that $G_n(s)$ for $n \geq 1$ is not a density function.

We can come up with some approximations for $G_n(s)$. However we need to be careful as these integrals are not finite for too large $n$ (depending on $\beta$). In fact the smaller $\beta$ is, the smaller $n$ needs to be in order for $G_n(s)$ to be finite and thus useful. Now we only consider $\beta = 4$ but these approximations also hold for every $\beta$.

**Lemma 3.2.** *if $n < \beta$ then $G_n(s)$ exists and:*

$$G_n(s) = \frac{a}{\mu \prod_{i=1}^{n}(\beta - i)} s^{-(\beta-n)} + \mathcal{O}(s^{-(\beta-n+1)})$$

*as $s \to \infty$*

*Proof.* Induction in n. The first step is for n = 1

$$G_1(s) = \mathbb{P}(T_0 > s) = \frac{1}{\mu}\int_s^\infty \mathbb{P}(L_k > x)dx = \frac{1}{\mu}\int_s^\infty ax^{-\beta} + \mathcal{O}(x^{-(\beta+1)})dx =$$
$$= \left[\frac{-a}{\mu(\beta-1)}x^{-(\beta-1)} + \mathcal{O}(x^{-(\beta)})\right]_{x=s}^\infty = \frac{a}{\mu(\beta-1)}s^{-(\beta-1)} + \mathcal{O}(s^{-(\beta)}).$$

Assume true until n-1, then:

$$G_n(s) = \int_s^\infty G_{n-1}(x)dx = \int_s^\infty \frac{a}{\mu \prod_{i=1}^{n-1}(\beta - i)}x^{-(\beta-n+1)} + \mathcal{O}(x^{-(\beta-n+2)})dx =$$
$$= \left[\frac{-a}{(\beta-n)\mu \prod_{i=1}^{n-1}(\beta - i)}x^{-(\beta-n)} + \mathcal{O}(x^{-(\beta-n+1)})\right]_{x=s}^\infty =$$
$$= \frac{a}{\mu \prod_{i=1}^{n}(\beta - i)}s^{-(\beta-n)} + \mathcal{O}(s^{-(\beta-n+1)})$$

□

**Meaning for $\beta = 4$:** only up until n=3 are these integrals finite. Meaning $G_4(s) =$ infinity

Another equality, which is not hard to see is the following:

**Lemma 3.3.** *if $n < \beta$ then*

$$G_n(0) = \frac{1}{(n-1)!}\mathbb{E}(T_0^{n-1}) = \frac{1}{n!\mu}\mathbb{E}(L^n)$$

*Where $L$ has the same distribution as the $L_k$'s , also $T_0$ has density function $G_0(t)$*

*Proof.*

$$\begin{aligned}
G_n(0) &= \int_{t_1=0}^{\infty} G_{n-1}(t_1)dt_1 = \int_{t_1=0}^{\infty}\int_{t_2=t_1}^{\infty} G_{n-2}(t_2)dt_2dt_1 = \\
&= \int_{t_2=0}^{\infty}\int_{t_1=0}^{t_2} G_{n-2}(t_2)dt_1dt_2 = \int_{t_2=0}^{\infty} t_2G_{n-2}(t_2)dt_2 = \cdots = \\
&= \int_{t_{n-1}=0}^{\infty} \frac{t_{n-1}^{n-2}}{(n-2)!}G_1(t_{n-1})dt_{n-1} = \int_{t_{n-1}=0}^{\infty} \frac{t_{n-1}^{n-2}}{(n-2)!}\mathbb{P}(T_0 > t_{n-1})dt_{n-1} = \\
&= \int_{t_{n-1}=0}^{\infty}\int_{t_n=t_{n-1}}^{\infty} \frac{t_{n-1}^{n-2}}{(n-2)!}G_0(t_n)dt_ndt_{n-1} = \int_{t_n=0}^{\infty}\int_{t_{n-1}=0}^{t_n} \frac{t_{n-1}^{n-2}}{(n-2)!}G_0(t_n)dt_{n-1}dt_n = \\
&= \int_{t_n=0}^{\infty} \frac{t_n^{n-1}}{(n-1)!}G_0(t_n)dt_n = \frac{1}{(n-1)!}\mathbb{E}(T_0^{n-1})
\end{aligned}$$

Since $G_0(t)$ is the density function of $T_0$

The second equation has the same proof save stopping at the last step. L has density function $f_L(t)$

$$\begin{aligned}
G_n(0) &= \int\limits_{t_n=0}^{\infty} \frac{t_n^{n-1}}{(n-1)!}G_0(t_n)dt_n = \int\limits_{t_n=0}^{\infty} \frac{t_n^{n-1}}{(n-1)!}\frac{1}{\mu}\mathbb{P}(L > t_n)dt_n = \\
&= \int\limits_{t_n=0}^{\infty}\int\limits_{t_{n+1}=t_n}^{\infty} \frac{t_n^{n-1}}{(n-1)!}\frac{f_L(t_{n+1})}{\mu}dt_{n+1}dt_n = \int\limits_{t_{n+1}=0}^{\infty}\int\limits_{t_n=0}^{t_{n+1}} \frac{t_n^{n-1}}{(n-1)!}\frac{f_L(t_{n+1})}{\mu}dt_ndt_{n+1} = \\
&= \int\limits_{t_{n+1}=0}^{\infty} \frac{t_n^{n}}{(n)!\mu}f_L(t_{n+1})dt_{n+1} = \frac{1}{n!\mu}\mathbb{E}(L^n)
\end{aligned}$$

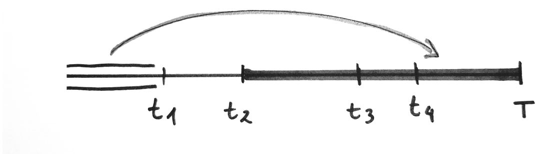
$\square$

Our goal is to calculate or approximate the 4th cumulant of the random variables $S_T$ (defined in section 1). This requires the second and the fourth moment (as stated in Lemma 1.2). Calculating the fourth moment of $S_T$ with a general $L$ tail distribution can be challenging. However some general observations can be made about the integral that will bring us closer to computing $\mathcal{C}_4(S_T)$. We are going to use 4 events to express the formulas.

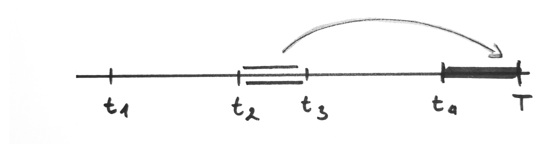
**Definition 3.4.** *Let* $0 < t_1 < t_2 < t_3 < t_4 < T$ *be four fixed numbers and let*

$$\begin{aligned}
A &= \{H^{stac}(t_1) > t_2 - t_1\} \\
B &= \{H(t_3) > t_4 - t_3\} \\
C &= \{t_3 - t_1 > H^{stac}(t_1) > t_2 - t_1\} \\
D &= \{H^{stac}(t_1) > t_4 - t_1\}
\end{aligned}$$

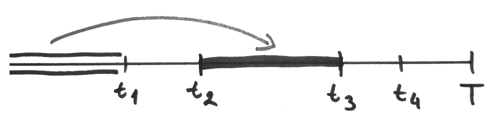
*be four events. See figure (1). From now on we may write* $H^{stac} = H^{stac}(t_1)$ *as it is independent of* $t_1$*. Note that in the definition of B we have* $H(t_3)$ *instead of* $H^{stac}$*. As we cannot always assume stationarity of* $H(t_3)$*. We will comment on this below.*



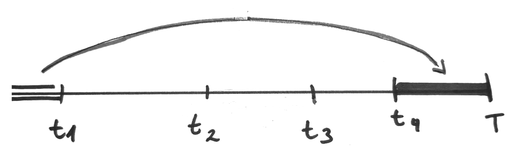

(a) A event




(b) B event




(c) C event




(d) D event

Figure 1: Events

It is important to note, that B and C are events concerning distinct renewal intervals.

**Lemma 3.5.** *([5], Appendix A) Without loss of generality let* $0 < t_1 < t_2 < T$*. Then*

$$\mathbb{E}(S_T^2) = 2! \int_{t_2=0}^{T} \int_{t_1=0}^{t_2} \mathbb{P}(A) dt_1 dt_2$$

*Proof.* We start with the expansion.

$$\mathbb{E}(S_T^2) = 2! \int_{t_2=0}^{T} \int_{t_1=0}^{t_2} \mathbb{E}[\xi(t_1)\xi(t_2)] dt_1 dt_2$$

Now $\mathbb{E}[\xi(t_1)\xi(t_2)]$ can be written in terms of conditional expectation, where we condition on the $t_i$'s placement in relation with the arrival times. Because if $t_1$ and $t_2$ are in the same renewal period then the product of $\xi(t_1)$ and $\xi(t_2)$ will certainly be 1. (Indeed if $t_1, t_2 \in [T_{k-1}, T_k]$ then the respective $\xi$ values are the same). However, if $t_1$ and $t_2$ are in different renewal blocks, then because of the independence of the $\xi_i$'s, the product's expected value will become the expected values' product. Also from the distribution of $\xi_i$'s the expected values are 0.

$$\begin{aligned}\mathbb{E}[\xi(t_1)\xi(t_2)] &= \mathbb{E}[\xi(t_1)\xi(t_2) \mid t_2 - t_1 > H^{stac}(t_1)] \cdot \mathbb{P}(t_2 - t_1 > H^{stac}(t_1)) \\ &+ \mathbb{E}[\xi(t_1)\xi(t_2) \mid H^{stac}(t_1) > t_2 - t_1] \cdot \mathbb{P}(H^{stac}(t_1) > t_2 - t_1) = \\ &= \mathbb{P}(H^{stac}(t_1) > t_2 - t_1) = \mathbb{P}(A)\end{aligned}$$

Exactly because $t_2 - t_1 > H^{stac}(t_1)$ means that $\xi_{t_1}$ and $\xi_{t_2}$ have values $\xi_i$ with different $i$'s. Meaning their product's expected value is 0. Also $\mathbb{E}[\xi(t_1)\xi(t_2) \mid H^{stac}(t_1) > t_2 - t_1]$ is exactly 1 because from the condition follows: $\xi_{t_1} = \xi_{t_2} = \xi_{m(t_1)}$

□

**Lemma 3.6.** *Without loss of generality let $0 < t_1 < t_2 < t_3 < t_4 < T$. Then*

$$\mathbb{E}(S_T^4) = 4! \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \Big(\mathbb{P}(D) + \mathbb{P}(C \cap B)\Big) dt_1 dt_2 dt_3 dt_4$$

*Proof.*

$$\mathbb{E}(S_T^4) = 4! \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \mathbb{E}[\xi(t_1)\xi(t_2)\xi(t_3)\xi(t_4)] dt_1 dt_2 dt_3 dt_4$$

This is very similar to the proof of lemma 3.5. With extra possible divisions of the $t_i$'s. But the gist of it is the same. $\mathbb{E}[\xi(t_1)\xi(t_2)\xi(t_3)\xi(t_4)]$ is only non-zero in those cases when there is an even number of $t$'s in each renewal period, then their product will be positive. Because of independence if an odd number of $t_i$'s are in one renewal period then the product's expected value will be zero (here then it is certainly true that there exists a $t_i$ which sits alone in a renewal block, whose expected value is 0 ). It is easy to see that $t_i$'s can only be placed in 2 ways for this product to have non-zero expected value. Either all four of them are in 1 big renewal period ($D$) or 2-2 of them are in two distinct renewal periods ($C \cap B$). From this follows:

$$\mathbb{E}[\xi(t_1)\xi(t_2)\xi(t_3)\xi(t_4)] =$$
$$= \mathbb{E}[\xi(t_1)\xi(t_2)\xi(t_3)\xi(t_4) \mid D\ ] \cdot \mathbb{P}(D) + \mathbb{E}[\xi(t_1)\xi(t_2)\xi(t_3)\xi(t_4) \mid C \cap B\ ] \cdot \mathbb{P}(C \cap B)$$

Because of the conditions these expected values will be 1 , that leaves us with the sum of their probabilities, which is what we were looking for.

□

Let us comment finally on having $H(t_3)$ instead of $H^{stac}$ in event B. Note that B occurs in the form of $\mathbb{P}(B \cap C)$ thus we have to consider the conditional probabilities $\mathbb{P}(B|C)$. From event C follows that there has been an arrival between $t_2$ and $t_3$, meaning we no longer can assume that the process is in stationary state at time $t_3$.

# 4 Fourth Cumulant

## 4.1 Second Moment

**Let's start by calculating** $\mathbb{E}(S_T^2)$ , where according to lemma 3.5:

$$\frac{1}{2!}\mathbb{E}(S_T^2) = \int\limits_{t_2=0}^{T}\int\limits_{t_1=0}^{t_2} \mathbb{E}[\xi(t_1)\xi(t_2)]dt_1dt_2 = \int\limits_{t_2=0}^{T}\int\limits_{t_1=0}^{t_2} \mathbb{P}(A)dt_1dt_2 =$$

Because of stationarity (7) and the definition of $G_n(x)$ (def 3.1)

$$= \int\limits_{t_2=0}^{T}\int\limits_{t_1=0}^{t_2} \mathbb{P}(T_0 > t_2 - t_1)dt_1dt_2 = \int\limits_{t_2=0}^{T}\int\limits_{t_1=0}^{t_2} G_1(t_2 - t_1)dt_1dt_2 =$$

With the substitution $s = t_2 - t_1$ and by the definition of $G_2(x)$ (def 3.1)

$$= \int\limits_{t_2=0}^{T}\int\limits_{s=0}^{t_2} G_1(s)dsdt_2 = \int\limits_{t_2=0}^{T} \Big(G_2(0) - G_2(t_2)\Big)dt_2 = TG_2(0) - (G_3(0) - G_3(T))$$

So we can conclude that :

$$\mathbb{E}(S_T^2) = 2\,(TG_2(0) - G_3(0) + G_3(T)) \tag{13}$$

Here it is worth noting that during the argument for (13), we did not use that $\beta > 4$. Also I would like to mention that in [5] similar calculations are done however, without the $G_n(x)$ notation. So our argument for $\mathbb{E}(S_T^2)$ is depending on the existence of $G_3(x)$ and $G_2(x)$, to be more precise:

- for $\beta > 3$ (13) is true
- for $3 \geq \beta > 2$ this argument stands, save $G_3(x)$'s existence. There we should use:

$$\int\limits_{t_2=0}^{T} \Big(G_2(0) - G_2(t_2)\Big)dt_2 = TG_2(0) + o(T^{3-\beta}) = TG_2(0) + o(T)$$

- for $\beta = 2$ the explicit argument in [5] gives us:

$$\mathbb{E}(S_T^2) = \frac{2a}{\mu} T \ln(T) + \mathcal{O}(T)$$

Which is exactly the doubling effect mentioned in section 1.3.

This concludes our discussion of $\mathbb{E}(S_T^2)$.

## 4.2 Fourth Moment

**Now lets focus on the more complicated** $\mathbb{E}(S_T^4)$. The calculations are elementary however there is a lot of them and as we approach the end result, we will be dividing it into smaller integrals, solving it one bit at a time. But first let's draw the big picture using lemma 3.6:

$$\frac{1}{4!}\mathbb{E}(S_T^4) = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} \mathbb{E}[\xi(t_1)\xi(t_2)\xi(t_3)\xi(t_4)]dt_1dt_2dt_3dt_4 =$$

$$= \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} \Big(\mathbb{P}(D) + \mathbb{P}(C \cap B)\Big) dt_1dt_2dt_3dt_4 =$$

$$= \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} \mathbb{P}(D)dt_1dt_2dt_3dt_4 + \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} \mathbb{P}(C \cap B)dt_1dt_2dt_3dt_4 \qquad (14)$$

So here we have to determine these 2 integrals.

Here we need to mention that we focus on the case of $\beta = 4$, meaning $G_n(t)$ will only be finite for $n \leq 3$. Thus we need an approximation formula for the integral, as we are not allowed to use the Lemmas 3.3 and 3.2 involving $G_4(t)$. Here the following modification of lemma 3.2 ($n = 3, \beta = 4$) will be most helpful to us:

$$\int_{t=0}^{T} G_3(t)dt = \int_{t=1}^{T} \left[ \frac{a}{\mu \prod_{i=1}^{n}(\beta - i)} t^{-(\beta-n)} + \mathcal{O}(t^{-(\beta-n+1)}) \right] dt + \mathcal{O}(1) =$$

$$= \int_{t=1}^{T} \left[ \frac{a}{6\mu} t^{-1} + \mathcal{O}(t^{-2}) \right] dt + \mathcal{O}(1) =$$

$$= \frac{a}{6\mu} \ln(T) + \mathcal{O}(1). \tag{15}$$

This is a valid approximation, because we only look at the results when T is big. In fact our approximation formula in lemma 3.2 only works as $T \to \infty$ Also there is some rule for the early behavior (e.g $[0, 1]$ interval), where this integral $G_3(t)$ of L has to be finite. We can assume that we integrate from 1 to $T$ since only the tail distribution is known to us and the remaining term is $\mathcal{O}(1)$.

## 4.3 Fourth Moment $D$ event

Let's start by solving the term in (14) involving $\mathbb{P}(D)$:

$$\int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \mathbb{P}(D) dt_1 dt_2 dt_3 dt_4 = \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \mathbb{P}(H^{stac}(t_1) > t_4 - t_1) dt_1 dt_2 dt_3 dt_4 =$$

Because of the stationarity of $T_0$ (7):

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \mathbb{P}(T_0 > t_4 - t_1) dt_1 dt_2 dt_3 dt_4 =$$

By the definition of $G_n(k)$'s (def 3.1):

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} G_1(t_4 - t_1) dt_1 dt_2 dt_3 dt_4 =$$

By $s = t_4 - t_1$ substitution:

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{s=t_4-t_2}^{t_4} G_1(s) ds dt_2 dt_3 dt_4 = \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \Big( G_2(t_4 - t_2) - G_2(t_4) \Big) dt_2 dt_3 dt_4 =$$

By $u = t_4 - t_2$ substitution:

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \left( -G_2(t_4) t_3 + \int_{u=t_4-t_3}^{t_4} G_2(u) du \right) dt_3 dt_4 =$$

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} [-G_2(t_4) t_3 + G_3(t_4 - t_3) - G_3(t_4)] dt_3 dt_4 =$$

By $v = t_4 - t_3$ substitution:

$$= \int_{t_4=0}^{T} \left( -\frac{1}{2} G_2(t_4) t_4^2 - G_3(t_4) t_4 + \int_{v=0}^{t_4} G_3(v) dv \right) dt_4 =$$

With the help of (15):

$$= \int_{t_4=0}^{T} \Big( -\frac{1}{2} G_2(t_4) t_4^2 - G_3(t_4) t_4 \Big) dt_4 + \int_{t_4=1}^{T} \left( \frac{a}{6\mu} ln(t_4) + \mathcal{O}(1) \right) dt_4 + \mathcal{O}(1) =: U$$

Where we have used that:

$$\int_{t_4=0}^{1} \left( \frac{a}{6\mu} \ln(t_4) + \mathcal{O}(1) \right) = \mathcal{O}(1).$$

Lets denote this integral by U and do some smaller calculations. Now it helps to know that:

$$\int_{t_4=0}^{T} -\frac{1}{2} G_2(t_4) t_4^2 dt_4 = \left[ \frac{t_4^2}{2} G_3(t_4) \right]_{t_4=0}^{T} + \int_{t_4=0}^{T} -t_4 G_3(t_4) dt_4 = \frac{T^2}{2} G_3(T) + \int_{t_4=0}^{T} -t_4 G_3(t_4) dt_4$$

And also we can calculate the following with the help of lemma 3.2:

$$\int_{t_4=0}^{T} -t_4 G_3(t_4) dt_4 = \int_{t_4=1}^{T} -t_4 \Big[ \frac{a}{6\mu} t_4^{-1} + \mathcal{O}(t_4^{-2}) \Big] dt_4 + \mathcal{O}(1) = -\frac{a}{6\mu} T + \mathcal{O}(\ln(T))$$

So now we can calculate the integrals in U:

$$U = \frac{T^2}{2}G_3(T) + 2\int_{t_4=0}^{T} -t_4 G_3(t_4)dt_4 + \int_{t_4=0}^{T}\left(\frac{a}{6\mu}\ln(T) + \mathcal{O}(1)\right)dt_4 =$$

$$= \frac{T^2}{2}G_3(T) - 2\frac{a}{6\mu}T + \mathcal{O}(1) + \frac{a}{6\mu}T\ln(T) - \frac{a}{6\mu}T + \mathcal{O}(T)$$

From the fact $G_3(T) = \mathcal{O}(T^{-1})$ follows that $T^2G_3(T) = \mathcal{O}(T)$. Thus:

$$U = \frac{a}{6\mu}T\ln(T) + \mathcal{O}(T) \tag{16}$$

## 4.4 Fourth Moment $B \cap C$ event - approximation

Now we are going to consider the integral containing $\mathbb{P}(C \cap B)$. In this subsection we will use the following approximation: When considering $\mathbb{P}(B|C)$ we will pretend as if the process was in stationary state at time $t_3$. In other words, we pretend that C and B are independent. This is a strong approximation, but it makes the calculations easier. We intend to justify the results in the next subsection.

But in fact if we know that an arrival happened just now or long ago, then prior knowledge will, in most cases, change the next arrival's distribution. This is easier to see for distributions whose domain is bounded. (e.g: Uniform on [0,1] if I know that the last arrival happened at 0.9 ago, then it is certain that in the next 0.1 it will happen again. On the other hand if it happened 0.5 ago, then on the next 0.1 it will only happen with 20% probability )

Nevertheless, we first approximate these events, albeit not giving us exact values for the moments and cumulants of $S_T$'s in the $\beta = 4$ polynomial tail distribution case.

The case of $\mathbb{P}(C \cap B)$ pretending the independence of B and C:

$$\int_{t_4=0}^{T}\int_{t_3=0}^{t_4}\int_{t_2=0}^{t_3}\int_{t_1=0}^{t_2} \mathbb{P}(C \cap B)dt_1dt_2dt_3dt_4 = \int_{t_4=0}^{T}\int_{t_3=0}^{t_4}\int_{t_2=0}^{t_3}\int_{t_1=0}^{t_2} \mathbb{P}(C)\mathbb{P}(B)dt_1dt_2dt_3dt_4 =$$

$$= \int_{t_4=0}^{T}\int_{t_3=0}^{t_4}\int_{t_2=0}^{t_3}\int_{t_1=0}^{t_2} \mathbb{P}(H^{stac}(t_3) > t_4 - t_3)\mathbb{P}(t_3 - t_1 > H^{stac}(t_1) > t_2 - t_1)dt_1dt_2dt_3dt_4 =$$

Because of stationarity of $T_0$ (7):

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \mathbb{P}(T_0 > t_4 - t_3)\mathbb{P}(t_3 - t_1 > T_0 > t_2 - t_1) dt_1 dt_2 dt_3 dt_4 =$$

Also using the definition of $G_n(k)$'s (def 3.1):

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} G_1(t_4 - t_3) \left[G_1(t_2 - t_1) - G_1(t_3 - t_1)\right] dt_1 dt_2 dt_3 dt_4 =$$

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \Big[G_1(t_4 - t_3)G_1(t_2 - t_1) - G_1(t_4 - t_3)G_1(t_3 - t_1)\Big] dt_1 dt_2 dt_3 dt_4$$

Here let's start introducing notations for the subsequent terms:

$$I = \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} G_1(t_4 - t_3)G_1(t_2 - t_1) dt_1 dt_2 dt_3 dt_4$$

$$II = \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} G_1(t_4 - t_3)G_1(t_3 - t_1) dt_1 dt_2 dt_3 dt_4$$

Rewriting our result we get that :

$$\int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \mathbb{P}(C \cap B) dt_1 dt_2 dt_3 dt_4 = I - II \tag{17}$$

We will start by calculating $I$:

$$I = \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} G_1(t_4 - t_3)G_1(t_2 - t_1) dt_1 dt_2 dt_3 dt_4 =$$

$$= \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} G_1(t_4 - t_3) \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} G_1(t_2 - t_1) dt_1 dt_2 dt_3 dt_4 =$$

observe that $\int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} G_1(t_2 - t_1) dt_1 dt_2$ is nothing else, than $\frac{1}{2!}\mathbb{E}(S_{t_3}^2)$. So by (13)

$$I = \int_{t_4=0}^{T} \int_{t_3=0}^{t_4} G_1(t_4 - t_3) \left(t_3 G_2(0) - G_3(0) + G_3(t_3)\right) dt_3 dt_4 = I_1 - I_2 + I_3$$

Now here we have 3 integrals, where $I_1$, $I_2$ and $I_3$ stands for :

$$I_1 = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) t_3 G_2(0) dt_3 dt_4$$

$$I_2 = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) G_3(0) dt_3 dt_4$$

$$I_3 = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) G_3(t_3) dt_3 dt_4$$

Then also decompose II:

$$II = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} G_1(t_4 - t_3) G_1(t_3 - t_1) dt_1 dt_2 dt_3 dt_4 =$$

$$= \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) \int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} G_1(t_3 - t_1) dt_1 dt_2 dt_3 dt_4$$

We need to calculate the inner integral first (substitute $s = t_3 - t_1$):

$$\int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} G_1(t_3 - t_1) dt_1 dt_2 = \int\limits_{t_2=0}^{t_3} \int\limits_{s=t_3-t_2}^{t_3} G_1(s) ds dt_2 = \int\limits_{t_2=0}^{t_3} \Big[ G_2(t_3 - t_2) - G_2(t_3) \Big] dt_2 =$$

Substituting $u = t_3 - t_2$ :

$$= \int\limits_{u=0}^{t_3} G_2(u) du - G_2(t_3) t_3 = G_3(0) - G_3(t_3) - t_3 G_2(t_3)$$

So now we know that:

$$II = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) \left[ G_3(0) - G_3(t_3) - t_3 G_2(t_3) \right] dt_3 dt_4 =$$

$$\int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) G_3(0) dt_3 dt_4 - \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) G_3(t_3) dt_3 dt_4$$

$$-\int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_1(t_4-t_3)t_3G_2(t_3)dt_3dt_4 = II_1 - II_2 - II_3$$

Where the names are:

$$II_1 = \int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_1(t_4-t_3)G_3(0)dt_3dt_4$$

$$II_2 = \int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_1(t_4-t_3)G_3(t_3)dt_3dt_4$$

$$II_3 = \int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_1(t_4-t_3)t_3G_2(t_3)dt_3dt_4$$

Now first let's make the observations that:

$$I_2 = II_1 \ and \ I_3 = II_2$$

From this and (17) follows, that :

$$\int_{t_4=0}^{T}\int_{t_3=0}^{t_4}\int_{t_2=0}^{t_3}\int_{t_1=0}^{t_2} \mathbb{P}(C\cap B)dt_1dt_2dt_3dt_4 = (I_1 - I_2 + I_3) - (II_1 - II_2 - II_3) =$$

$$= I_1 - 2I_2 + 2I_3 + II_3 \tag{18}$$

We need to compute these four integrals, **starting with $I_1$:**

$$I_1 = \int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_1(t_4-t_3)t_3G_2(0)dt_3dt_4 = G_2(0)\int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_1(t_4-t_3)t_3dt_3dt_4 =$$

Substitute $s = t_4 - t_3$:

$$G_2(0)\int_{t_4=0}^{T}\int_{s=0}^{t_4} G_1(s)(t_4-s)dsdt_4 =$$

$$= G_2(0)\left(\int_{t_4=0}^{T} t_4\int_{s=0}^{t_4} G_1(s)dsdt_4 - \int_{t_4=0}^{T}\int_{s=0}^{t_4} sG_1(s)dsdt_4\right) =$$

$$= G_2(0) \int\limits_{t_4=0}^{T} \Big[t_4 G_2(0) - t_4 G_2(t_4)\Big] dt_4 - G_2(0) \int\limits_{t_4=0}^{T} \int\limits_{s=0}^{t_4} s G_1(s) ds dt_4 =$$

Integration by parts:

$$= G_2^2(0) \int\limits_{t_4=0}^{T} t_4 dt_4 - G_2(0) \int\limits_{t_4=0}^{T} t_4 G_2(t_4) dt_4$$

$$-G_2(0) \int\limits_{t_4=0}^{T} \left( \Big[ - s G_2(s) \Big]_0^{t_4} + \int\limits_{s=0}^{t_4} G_2(s) ds \right) dt_4 =$$

$$= G_2^2(0) \frac{1}{2} T^2 - G_2(0) \int\limits_{t_4=0}^{T} t_4 G_2(t_4) dt_4$$

$$+G_2(0) \int\limits_{t_4=0}^{T} t_4 G_2(t_4) dt_4 - G_2(0) \int\limits_{t_4=0}^{T} [G_3(0) - G_3(t_4)] dt_4 =$$

$$= G_2^2(0) \frac{1}{2} T^2 - G_2(0) \int\limits_{t_4=0}^{T} [G_3(0) - G_3(t_4)] dt_4$$

$$= G_2^2(0) \frac{1}{2} T^2 - T G_2(0) G_3(0) + G_2(0) \int\limits_{t_4=0}^{T} G_3(t_4) dt_4 =$$

Now with the help of (15)

$$I_1 = G_2^2(0) \frac{1}{2} T^2 - T G_2(0) G_3(0) + G_2(0) [\frac{a}{6\mu} \ln(T) + \mathcal{O}(1)] = G_2^2(0) \frac{1}{2} T^2 + \mathcal{O}(T)$$

.

With that finished, **let's solve** $I_2$ (which we already have). By noticing that:

$$I_2 = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) G_3(0) dt_3 dt_4 = G_3(0) \frac{1}{2!} \mathbb{E}(S_T^2)$$

$$= G_3(0)[T G_2(0) - G_3(0) + G_3(T)] = \mathcal{O}(T) \tag{19}$$

Follows from (13).

**Now only** $I_3$ **and** $II_3$ **are left to be calculated**. However we are going to approximate them since as it turns out their leading terms are negligible to us. For this we need to notice 2 facts. First, the 2 integrals are very similar in fact $G_3(t)$ and $tG_2(t)$ are both of the same order. So there exist a constant $C$ for which $tG_2(t) < CG_3(t)$ $(\forall t > 0)$ So by this we can also conclude that:

$$II_2 = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) t_3 G_2(t_3) dt_3 dt_4 < \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) C G_3(t_3) dt_3 dt_4 =$$

$$= C \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) G_3(t_3) dt_3 dt_4 = C \cdot I_3$$

Now we have that $2I_3 + II_3 < (2 + C) \cdot I_3$.

By rewriting the original $I_3$ integral with integration by parts we see that:

$$I_3 = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_1(t_4 - t_3) G_3(t_3) dt_3 dt_4 =$$

$$= \int\limits_{t_4=0}^{T} \left( \Big[ G_2(t_4 - t_3) G_3(t_3) \Big]_{t_3=0}^{t_4} + \int\limits_{t_3=0}^{t_4} G_2(t_4 - t_3) G_2(t_3) dt_3 \right) dt_4 =$$

Now here we can't write $G_4(0) - G_4(T)$ as these don't exist but rather use (15)

$$I_3 = \int\limits_{t_4=0}^{T} \Big[ G_2(0) G_3(t_4) - G_2(t_4) G_3(0) \Big] dt_4 + \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_2(t_4 - t_3) G_2(t_3) dt_3 dt_4 =$$

$$= G_2(0) \left( \frac{a}{6\mu} \ln(T) + \mathcal{O}(1) \right) - G_3(0) \Big( G_3(0) - G_3(T) \Big) + \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_2(t_4 - t_3) G_2(t_3) dt_3 dt_4 =$$

$$\mathcal{O}(T) + \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} G_2(t_4 - t_3) G_2(t_3) dt_3 dt_4$$

Where the last step is valid as $G_3(T) = \mathcal{O}(T)$ and also $ln(T) = \mathcal{O}(T)$.

The second observation is that the inner part of this last double integral is a convolution of $G_2(x)$ with itself. Convolution has the property that it is a closed operation in $L^1$ space. That is if $f, g \in L^1$ are convolved, then the resulting function will also be in $L^1$. And since $G_3(0)$ exists (for $\beta > 3$) it is true that $G_2(x)$ is integrable , thus the convolution is also in $L^1$, let's call it $W(t_4)$. And thus:

$$\int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_2(t_4-t_3)G_2(t_3)dt_3dt_4 = \int_{t_4=0}^{T} W(t_4)dt_4 = \mathcal{O}(1)$$

With this we can see that:

$$I_3 = \int_{t_4=0}^{T}\int_{t_3=0}^{t_4} G_2(t_4-t_3)G_2(t_3)dt_3dt_4 + \mathcal{O}(T) = \mathcal{O}(1) + \mathcal{O}(T) = \mathcal{O}(T)$$

From this it is clear that:

$$2I_3 + II_3 < (2+C)\cdot I_3 = (2+C)\cdot\mathcal{O}(T) = \mathcal{O}(T) \tag{20}$$

Now using (18) we know that

$$\int_{t_4=0}^{T}\int_{t_3=0}^{t_4}\int_{t_2=0}^{t_3}\int_{t_1=0}^{t_2} \mathbb{P}(C\cap B)dt_1dt_2dt_3dt_4 = I_1 - 2I_2 + 2I_3 + II_3$$

$$= G_2^2(0)\frac{1}{2}T^2 + \mathcal{O}(T) \tag{21}$$

So now we know with the help of (14), (16) and (21):

$$\mathbb{E}(S_T^4) = 24\Big(G_2^2(0)\frac{1}{2}T^2 + \frac{a}{6\mu}T\ln(T) + \mathcal{O}(T)\Big) = 12G_2^2(0)T^2 + \frac{4a}{\mu}T\ln(T) + \mathcal{O}(T)$$

Now as for the role of the second moment (13)

$$3(\mathbb{E}(S_T^2))^2 = 3\Big[2\left(TG_2(0) - G_3(0) + G_3(T)\right)\Big]^2 =$$

$$= 12\Big(T^2G_2^2(0) + G_3^2(0) + G_3^2(T) - 2TG_2(0)G_3(0) + 2TG_2(0)G_3(T) - 2G_3(0)G_3(T)\Big) =$$

$$= 12T^2G_2^2(0) + \mathcal{O}(T)$$

So using lemma 1.2 and the two equations above, we get that:

$$\mathcal{C}_4(S_T) = \mathbb{E}(S_T^4) - 3(\mathbb{E}(S_T^2))^2 = 12G_2^2(0)T^2 + \frac{4a}{\mu}T\ln(T) + \mathcal{O}(T) - 12T^2G_2^2(0) + \mathcal{O}(T)$$

$$\mathcal{C}_4(S_T) = \frac{4a}{\mu}T\ln(T) + \mathcal{O}(T) \tag{22}$$

## 4.5 Fourth Moment $B \cap C$ event - dealing with the correctness of the approximation

Expanding the event $\mathbb{P}(B \cap C)$:

$$\mathbb{P}(B \cap C) = \int_{t_2-t_1}^{t_3-t_1} \mathbb{P}(B|H^{stac}(t_1) = z) f_{H^{stac}(t_1)}(z)dz =$$

Using the definition of B and "setting $t_1 + z$ to be our new start", that is we know it is an arrival time and using the definition of $\Phi(x, t)$ (from equation (8) ):

$$= \int_{t_2-t_1}^{t_3-t_1} [1 - \mathbb{P}(H[t_3 - (t_1 + z)] \leq t_4 - t_3)] \, f_{H^{stac}(t_1)}(z)dz =$$

$$= \int_{t_2-t_1}^{t_3-t_1} [1 - \Phi(t_4 - t_3, t_3 - z - t_1)] \, f_{H^{stac}(t_1)}(z)dz$$

Now let us consider this last form of the event probability. If $t_3 - z \to \infty$ then stationarity will appear and also using (8) we get:

$$\mathbb{P}(B \cap C) = \int_{t_2-t_1}^{t_3-t_1} [1 - \Phi(t_4 - t_3, t_3 - z - t_1)] \, f_{H^{stac}(t_1)}(z)dz =$$

$$= \int_{t_2-t_1}^{t_3-t_1} [1 - \mathbb{P}(H[t_3 - (t_1 + z)] \leq t_4 - t_3)] \, f_{H^{stac}(t_1)}(z)dz \approx$$

$$\approx \int_{t_2-t_1}^{t_3-t_1} \Big[1 - \mathbb{P}(H^{stac}(t_3) \leq t_4 - t_3)\Big] \, f_{H^{stac}(t_1)}(z)dz =$$

Where $H(t_3)$ is no longer a function of $z$ nor of $t_3 - (t_1 + z)$, because of stationarity. So from this follows:

$$= \Big[1 - \mathbb{P}(H^{stac}(t_3) \leq t_4 - t_3)\Big] \int_{t_2-t_1}^{t_3-t_1} f_{H^{stac}(t_1)}(z)dz = \mathbb{P}(B) \int_{t_2-t_1}^{t_3-t_1} 1dF_{H^{stac}(t_1)}(z) =$$

$$= \mathbb{P}(B) \cdot \mathbb{P}(t_3 - t_1 > H^{stac}(t_1) > t_2 - t_1) = \mathbb{P}(B) \cdot \mathbb{P}(C)$$

So from stationarity the independence of the events B and C follows. However in reality stationarity is a strong statement (one that is not true in general). In fact what we say is that $\mathbb{P}(H(t) > x) = \mathbb{P}(T_0 > x)$ which is a strong approximation. Recall that $\mathbb{P}(T_0 > x) \sim x^{-3}$. We suspect that the following bound could be enough and correct:

$$\mathbb{P}(H(t) > x) = \mathbb{P}(T_0 > x) + \mathcal{O}\left(\frac{1}{(t+x)^3}\right) \tag{23}$$

**Is it enough ?** Now writing this instead of the independence approximation we get that

$$\mathbb{P}(B \cap C) = \int_{t_2-t_1}^{t_3-t_1} \left[1 - \Phi(t_4 - t_3, t_3 - z - t_1)\right] f_{H^{stac}(t_1)}(z)dz =$$

$$= \int_{t_2-t_1}^{t_3-t_1} \left[1 - \mathbb{P}(H[t_3 - (t_1 + z)] \leq t_4 - t_3)\right] f_{H^{stac}(t_1)}(z)dz =$$

$$= \int_{t_2-t_1}^{t_3-t_1} \mathbb{P}(H[t_3 - t_1 - z] > t_4 - t_3) f_{H^{stac}(t_1)}(z)dz =$$

Using the approximation we get that:

$$= \int_{t_2-t_1}^{t_3-t_1} \Big[\mathbb{P}(H^{stac} > t_4 - t_3) + \mathcal{O}\left((t_3 - t_1 - z + t_4 - t_3)^{-3}\right)\Big] f_{H^{stac}(t_1)}(z)dz =$$

Which is a sum of two integrals, where the first term is exactly like the approximate case. Thus:

$$= \mathbb{P}(B) \cdot \mathbb{P}(C) + \int_{t_2-t_1}^{t_3-t_1} \mathcal{O}\left((t_4 - t_1 - z)^{-3}\right) f_{H^{stac}(t_1)}(z)dz$$

In order for equation (22) to hold, even after we dispose of the approximation, we need that:

$$\int_{t_4=0}^{T} \int_{t_3=0}^{t_4} \int_{t_2=0}^{t_3} \int_{t_1=0}^{t_2} \int_{t_2-t_1}^{t_3-t_1} \mathcal{O}\left((t_4 - t_1 - z)^{-3}\right) f_{H^{stac}(t_1)}(z)dzdt_1dt_2dt_3dt_4 = \mathcal{O}(T) \tag{24}$$

Where $f_{H^{stac}(t_1)}(z) = \mathcal{O}(z^{-4})$. Using this and a $v = z + t_1$ substitution we get:

$$\int\limits_{t_2-t_1}^{t_3-t_1} \mathcal{O}\left((t_4 - t_1 - z)^{-3}\right) \mathcal{O}\left((z)^{-4}\right) dz = \int\limits_{t_2}^{t_3} \mathcal{O}\left((t_4 - v)^{-3}\right) \mathcal{O}\left((v - t_1)^{-4}\right) dv =$$

$$= \int\limits_{t_2}^{t_3} \mathcal{O}\left((t_4 - v)^{-3}(v - t_1)^{-4}\right) dv$$

Now using elementary integral calculations from the indefinite integral $\int (t_4 - v)^{-3}(v - t_1)^{-4} dv$, we obtain that the leading term of this integral $\int\limits_{t_2}^{t_3} \mathcal{O}\left((t_4 - v)^{-3}(v - t_1)^{-4}\right) dv$ is of $\mathcal{O}\left((t_4 - t_1)^{-3}\right) \mathcal{O}\left((t_2 - t_1)^{-3}\right)$. So substituting back into equation (24) we get:

$$\int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \int\limits_{t_1=0}^{t_2} \mathcal{O}\left((t_4 - t_1)^{-3}\right) \mathcal{O}\left((t_2 - t_1)^{-3}\right) dt_1 dt_2 dt_3 dt_4) \leq$$

By applying that $\mathcal{O}\left((t_4 - t_1)^{-3}\right)$ is upper bounded by $\mathcal{O}\left((t_4 - t_2)^{-3}\right)$:

$$\leq \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \mathcal{O}\left((t_4 - t_2)^{-3}\right) \int\limits_{t_1=0}^{t_2} \mathcal{O}\left((t_2 - t_1)^{-3}\right) dt_1 dt_2 dt_3 dt_4) =$$

where the inner integral is finite so:

$$= \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \mathcal{O}\left((t_4 - t_2)^{-3}\right) \mathcal{O}(1) dt_2 dt_3 dt_4) = \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \int\limits_{t_2=0}^{t_3} \mathcal{O}\left((t_4 - t_2)^{-3}\right) dt_2 dt_3 dt_4) =$$

$$= \int\limits_{t_4=0}^{T} \int\limits_{t_3=0}^{t_4} \mathcal{O}\left((t_4 - t_3)^{-2}\right) dt_3 dt_4)$$

Again the inner part is a finite integral, leaving us with:

$$= \int\limits_{t_4=0}^{T} \mathcal{O}(1) dt_4 = \mathcal{O}(T)$$

Exactly what is required by equation (24).

**Is it true ?** Lets discuss why would the approximation (23) be correct. Some complicated calculations (and not yet rigorous ones) lead us to believe that they are true. Recall for the density $\phi(t, x)$ equation (23) reduces to:

$$\phi(t,x) = \frac{1}{\mu}\mathbb{P}(L > x) + \mathcal{O}\left((t+x)^{-4}\right)$$

To see our idea we need to inspect the formula in (9).

$$\phi(t,x) = f_L(t+x) + \int\limits_{y=0}^{t} \left[f_L(t+x-y)m(y)\right] dy$$

here $f_L(t+x) = \mathcal{O}(x+t)^{-5}$, which trivially fits into $\mathcal{O}(x+t)^{-4}$ . Then secondly, by using the formula (12) in $\widehat{m}(r)$ we expect $\widehat{m}(r) = \frac{1}{\mu r} + \widehat{m_0}(r)$ where $m_0(t)$ is a probability density function with $m_0(t) \sim t^{-4}$. This requires further justification.

Now expanding the expression in (9) linearly:

$$\int\limits_{y=0}^{t} \left[f_L(t+x-y)m(y)\right] dy = \int\limits_{y=0}^{t} \left[f_L(t+x-y)\frac{1}{\mu}\right] dy + \int\limits_{y=0}^{t} \left[f_L(t+x-y)m_0(y)\right] dy =$$

Using the same trick as with the $G_n(t)$'s with a $s = t - y$ integral substitution:

$$= \frac{1}{\mu}\left[\int\limits_{y=0}^{\infty} f_L(x+s)ds - \int\limits_{y=t}^{\infty} f_L(x+s)ds\right] + \int\limits_{y=0}^{t} \left[f_L(t+x-y)m_0(y)\right] dy =$$

$$\frac{1}{\mu}\left[\mathbb{P}(L > x) - \mathbb{P}(L > x+t)\right] + \int\limits_{y=0}^{t} \left[f_L(t+x-y)m_0(y)\right] dy$$

So back to $\phi(t,x)$ and putting this together:

$$\phi(t,x) = f_L(t+x) + \frac{1}{\mu}\left[\mathbb{P}(L > x) - \mathbb{P}(L > x+t)\right] + \int\limits_{y=0}^{t} \left[f_L(t+x-y)m_0(y)\right] dy$$

- $f_L(t+x) = \mathcal{O}\left((x+t)^{-5}\right)$
- $\mathbb{P}(L > x+t) = \mathcal{O}((x+t)^{-4})$
- $\int\limits_{y=0}^{t} \left[f_L(t+x-y)m_0(y)\right] dy$ If it was an integral until t+x, then it would be convolution of the form (through Laplace transforming reasoning) $f_L * m_0(x+t) \sim (t+x)^{-4}$. (That is $= \mathcal{O}((t+x)^{-4})$)

These 3 approximations would be taken care of, which would mean that $\mathcal{O}((t+x)^{-3})$ in (23) would hold.

# 5 Summary and Outlook

In this thesis we studied a random walk $S_T$ generated by a renewal process with its inter-arrival times having a tail distribution $\mathbb{P}(L > x) \sim ax^{-\beta}$ (with $\beta = 4$). The goal was to determine the 4th cumulant of this process, which we did under strong approximations. In particular, we introduced four parameters $0 \leq t_1 < t_2 < t_3 < t_4 \leq T$, along with events B and C (see figure 1). When constructing $\mathbb{P}(B|C)$ we pretended that the process is in stationary state at time $t_3$, although an arrival occurred in the interval $[t_2, t_3]$.

The heuristics behind this approximation is that as $T$ grows, $t_3 - t_2$ also grows in average. We have also included some discussion based on the exact formulas (9) and (12), to argue that the error terms in this approximation indeed contribute a correction of lower order. Further analysis of the error term is the subject of future research.

**Different goals** $\beta = 6$**:** Now, using a more complicated event system than A,B,C and D, we could also determine the 6th cumulant in the case of $\beta = 6$. However not only would this make the events more complicated, but also handling the error term would be a significantly more complex problem than that of the $\beta = 4$ case. Though harder, we do not expect the requirement of further theoretical insight than for cases $\beta = 2$ and $\beta = 4$


# Acknowledgements

We are thankful for Imre Péter Tóth for his useful comments on the earlier version of this work.


# References


[1] V.V.Petrov. *Sums of Independent Random Variables.* Ergebnisse der Mathematik und ihren Grenzgebiete, Band 82. *Springer - Verlag*, 1975

[2] Rick Durrett. *Essentials of Stochastic Processes.* 2nd Edition, ISBN-10: 1461436141, Springer, 2012

[3] William Feller. *An Introduction To Probability Theory And Its Applications. Volume II* . 2nd edition. John Wiley & Sons, Inc. New York-London-Sydney, 1971

[4] Sébastien Gouëzel. *Stable Laws for the Doubling Map.*, https://perso.univ-rennes1.fr/sebastien.gouezel/articles/DoublingStable.pdf, 2008

[5] Bálint, P.; Chernov, N.; Dolgopyat, D. *Convergence of moments for dispersing billiards with cusps.* Contemporary Mathematics 698: pp. 35-67, 2017

[6] Chernov, Nikolai, and Roberto Markarian. *Chaotic billiards.* Mathematical Surveys and Monograph. No. 127. American Mathematical Soc, 2006

[7] Dettmann, Carl P. *Diffusion in the Lorentz gas.* Communications in Theoretical Physics 62.4: 521, 2014

[8] Bálint Péter, Henk Bruin, and Dalia Terhesiu. *Lorentz gas with small scatterers.* arXiv:2107.10529, 2021

[9] Szász Domokos, and Tamás Varjú. *Limit laws and recurrence for the planar Lorentz process with infinite horizon.* Journal of Statistical Physics 129.1: 59-80, 2007

[10] G. Cristadoro, T. Gilbert, M. Lenci, & D. P. Sanders. *Measuring logarithmic corrections to normal diffusion in infinite-horizon billiards.* Physical Review E 90(2): 022106, 2014