

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Corso di Laurea Triennale in Fisica

Tesi di Laurea

**Extreme Gaussian targeting as a possible explanation
for redundancy**

**Targeting gaussiano estremo come possibile spiegazione
della ridondanza**

Relatore

Prof./Dr. Fulvio Baldovin

Laureando

Jacopo Boccato

Anno Accademico 2022/2023

dedica

Abstract

A possible explanation of redundancy in biology is the need of optimizing extreme searching processes, where only one or few among many particles are requested to reach the target like in human fertilization. Inspired by the ligando-receptor binding mechanism on the surface of a cell, in this thesis it is addressed the targeting of a particle diffusing on a plane, with circular reflecting boundaries at the border and an absorbing small area at the center. After studying the solution for the ordinary mean first passage time, we address both analytically and numerically the fastest targeting among many searchers. We assess our results in the context of recent theorems for extreme first passage times.

Sommario

Una possibile spiegazione della ridondanza in biologia è la necessità di ottimizzare i processi di ricerca "estremi", nei quali è richiesto che solamente una o poche particelle raggiungano effettivamente l'obiettivo, come nel caso della fecondazione umana. Ispirandosi al legame ligando-recettore sulla superficie di una cellula, in questa tesi si affronta il targeting di una particella che diffonde su un piano, con condizioni al contorno riflettenti in un bordo circolare e una piccola area assorbente al centro. Dopo aver studiato la soluzione per l'ordinario tempo medio di primo passaggio, viene analizzato analiticamente e numericamente il targeting del più veloce tra molti ligandi. I risultati vengono valutati nel contesto di recenti teoremi per i tempi estremi di primo passaggio.

Contents

Introduction	1
1 Diffusive processes	2
1.1 Brownian motion	2
1.2 Langevin Equation	2
1.3 Fokker Planck equation	3
1.4 First Passage Time	5
1.5 Extreme first passage time	6
1.6 Lawley's theorem	6
2 2d diffusion on a surface	7
2.1 Ligando-receptor problem	7
2.2 Solution	7
2.3 Asymptotic behaviour of $J(x_0)$	10
2.4 Other dimensions	11
3 Simulations	12
3.1 1d Program	12
3.2 Simulation of free diffusion	12
3.3 Confirmation of asymptotic behaviour	14
3.4 Confirmation of Lawley's theorem	15
Conclusions	17
Appendix	18
3.4.1 Laplace transform of derivative	19
3.4.2 Bessel function	19
Bibliography	20

Introduction

Back in 1905 Einstein published paper explaining mathematically the origin of the seemingly random motion showed by small particle immersed in a liquid. This movement, that is caused by the water molecule hitting the particle, appears to be random, since at any instant the particle can get hit in any direction. The mathematical explanation of this trajectory, which became known as Brownian motion or diffusive motion has been used to model a wide range of phenomena, ranging from physics to economics. In particular Brownian motion has been used extensively in biology to model the movement of small cells in a variety of problem such as the ligando-receptor problem, of which we are going to present a simplified model in this thesis. The ligando-receptor problem is a problem that is concerned about describing the behaviour of a small diffusive particle that is trying to reach an activation site on a bigger biological structure. In this thesis we are going to focus more precisely on computing the so called first passage time of this process, which is the time needed for the particle to reach the absorption point for the first time. This quantity is crucial to determine the average time needed for the activation of a certain biological function [4]. After solving the probability density function of the first passage time problem we will compute the extreme first passage time of this process. This quantity, that is closely related to the first passage time, can be defined as the smallest time needed to reach a certain target in many realization of the same process. This quantity is interesting because in nature many situation are characterized by a large number of searcher trying to reach the target, this means that often, since for the activation of a biological function only one successful searcher is needed, the activation time is significantly smaller if there are a large number of searching particle, and it's dictated by the behaviour of the extreme first passage time. In the last years the study of extreme first passage time and its properties has gained a lot of interest, as reported for instance in Ref. [4]. In particular recent results published in literature has led scientist to formulate the so called redundancy principle which states that the seemingly redundant number of copies in a biological searching process serve the function of reducing the time needed to complete said process. In this thesis we are going to present these results and we are going to test their validity on our simple model of the ligando-receptor problem.

This thesis is articulated in three chapter. The first chapter aims at introducing the most relevant mathematical results necessary to understand the general first passage time problem. In this chapter we will introduce concept ranging from Brownian motion to first passage time probability. These concepts are needed in the thesis, later we are going to highlight results recently published in scientific literature, namely an universal formula for the asymptotic behaviour of extreme first passage time mentioned above, of which we are going to outline a simplified proof. In the last two chapters we are going to focus on a particular problem that requires the use of the techniques highlighted in the first chapter, the so called ligando-receptor problem. Firstly we will present the problem and offer a solution to the Fokker Planck equation on a simple circular 2d geometry. Secondly we will create a C++ program that allows us to simulate the behaviour of many diffusive particle and then we are going to run some computer simulation in order to compare them to the analytical results laid out in the second chapter of the thesis. Finally in the we are going to present some final comments and considerations.

Chapter 1

Diffusive processes

In this first chapter we are going to introduce the concept of diffusive search starting from the perspective of a single searcher and then we are going to progress towards the Fokker Planck equation. This equation will allow us to express the probability density function of any diffusive process with the appropriate boundary condition. Later we are going to introduce survival probability and mean extreme first passage time, that are two crucial concept to understand the mathematical explanation of the redundancy principle. After that we are going to define the so called extreme first passage time and lastly we will present a recently published theorem that proves the general asymptotic behaviour of this quantity. In order to do so we will firstly give a historical and practical overview of Brownian motion and the two approach used to study it, namely the Langevin and Fokker Planck equations.

1.1 Brownian motion

Brownian motion was firstly used to describe the seemingly random movement of mesoscopic particle suspended in a fluid, this phenomena was observed for the first time by the botanist Robert Brown in 1827. Brownian motion has since become a fundamental tool in Physics for the description of random motion, i.e. random walk. Applications of this concept can be found in a wide range of fields, including Chemistry, Biology, Finance, and Engineering. In modern physics Einstein was the first to describe mathematically the concept of Brownian motion in his famous work "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In this paper (1905) Einstein lays out for the first time a conclusive proof of the existence of atoms and molecule. The first one who developed a rigorous mathematical theory of Brownian motion was Norbert Wiener in 1923.

Eulerian and Lagrangian coordinates

There are two different way to mathematically model a diffusive process. One could either look at the trajectory of a particle in each and every individual realization of a random walk or describe the probability density function of finding the diffusive particle in a region of space. In the first part of this chapter we are going to present the first view, that yields the Langevin equation, in this section the spatial coordinates are Lagrangian coordinates $x(t)$, this means that the position of each realization of the random walk is a function of time. On the other hand, in the second part of the chapter, the one describing the Fokker Planck equation the spatial coordinates are fixed in the frame of reference and they point to an exact place in space that is not moving, in this case the coordinates used are the Eulerian coordinates.

1.2 Langevin Equation

Following Einstein work on Brownian motion P. Langevin in 1908 wrote the law of motion of a mesoscopic particle in a fluid of microscopic particles without external forces. This particle is performing

a Brownian motion. Using Newton's second law we get:

$$m \frac{d\vec{v}}{dt} = -\gamma\vec{v} + F_r\vec{\tau}(t) \quad (1.1)$$

Where $-\gamma\vec{v}$ is the deterministic damping Stokes term and $F_r\vec{\tau}$ is a normal random variable with the following characteristics:

$$\langle F_r\vec{\tau}(t') \cdot F_r\vec{\tau}(t'') \rangle = C_0\delta(t' - t'') \quad \langle F_r\vec{\tau}(t) \rangle = 0$$

Starting from the asymptotic behaviour of $\vec{v}(t)$ and applying the equipartition theorem one gets the so called fluctuation dissipation theorem (Note that the number 6 is sensible to the dimensionality of the problem $6 = 2 \cdot d$):

$$\langle F_r\vec{\tau}(t') \cdot F_r\vec{\tau}(t'') \rangle = 6k_B T \gamma \delta(t' - t'') \quad (1.2)$$

It's important to remark that equation (1.1) is not mathematically rigorous since the velocity we have defined is continuous everywhere but not differentiable (This can be proven using the definition of Brownian motion, for reference see [5]). In order to be mathematically rigorous we are going to define the 3d Langevin equation as an update condition. In this thesis the two formulations are used interchangeably:

$$\vec{v}(t + dt) = \vec{v}(t) - \gamma\vec{v}(t)dt + \sqrt{6m\gamma k_B T dt} \vec{N}(0, 1) \quad (1.3)$$

Where $\vec{N}(0, 1)$ is a normally distributed 3d random variable with standard deviation $\sigma = 1$.

Over-damped Langevin Equation

We observe that in equation (1.1) the information about initial velocity is lost on timescale bigger than $\tau \sim m/\gamma$ which is known as Smoluchowski timescale. We are interested about the law of motion of our particle in said timescale. One easy way to obtain such information is to simply put $\frac{dv}{dt} = 0$. Now the equation, assuming spatially uniform diffusion coefficient, becomes:

$$\gamma\vec{v} = F_r\vec{\tau}(t) \quad (1.4)$$

Or, rearranged in the more mathematically precise update formulation:

$$\vec{x}(t + dt) = \vec{x}(t) + \sqrt{6Ddt} \vec{N}(0, 1) \quad (1.5)$$

Where D is the diffusion coefficient. Note that the diffusion coefficient has the unit of $[\mathbb{T}]^2[\mathbb{L}]^{-1}$. This equation clearly defines a continuous memory-less stochastic process. Which is a process that has the following property: its future evolution is independent of its history, to get a more mathematically rigorous definition see [5]. Furthermore the equation above is a stochastic ordinary differential equation. (for a mathematical definition see Ref. [3])

The other approach

After presenting the Lagrangian way of studying Brownian motion we are now going to present the Fokker Planck equation which is the results of the Eulerian perspective. This equation is a regular partial differential equation and can be solved with no knowledge of stochastic differential equation even though its describing an inherently probabilistic process.

1.3 Fokker Planck equation

In this section we are going to present the Fokker Planck equation and some of its properties. Since we are interested in diffusive process with no drift the Fokker Planck equation in 1d reads as follows [3]:

$$\frac{dp(x, t)}{dt} = D \frac{d^2}{dx^2} [p(x, t)] \quad (1.6)$$

Where D is the diffusion factor that in this thesis we have assumed to be spatially independent. The 3d equation is a natural extension that is obtained simply swapping $\frac{d^2}{dx^2}$ with ∇^2 . Note also that in 3d D becomes a tensor.

Probability flux

We want now to define the probability flux or probability current in 1d as:

$$j(x, t) = \frac{dp(x, t)}{dx}$$

Obviously one can generalize this in any dimension by transforming $\frac{dp(x, t)}{dx}$ in $\nabla p(x, t)$

Boundary condition

The Fokker Planck is a partial differential equation so in order to be a well posed problem it must have a set of boundary condition as well as some initial condition. In this section we are going to present the most common type of 1d boundary condition.

Natural boundary

In this case the process is defined on \mathbb{R} and the condition is the one in which the probability current vanishes at the boundaries $x_{min} = -\infty$ and $x_{max} = \infty$. This would imply the conservation of the normalization for $p(x, t)$ since:

$$\int_{-\infty}^{\infty} p(x, t) dx = 1$$

Clearly the decay must be sufficiently rapid to ensure the normalization of the integral above.

Reflecting boundary

An absorbing condition at $x = b$ on a 1d process means that a particle cannot go past $x = b$. With this definition the mathematical expression of said condition is:

$$j(a, t) = 0 \quad \forall t$$

Absorbing boundary

An absorbing condition at $x = a$ on a 1d process means that when particles hit $x = a$ for the first time they are removed from the interval. The appropriate BC for this type of scenario is :

$$p(a, t) = 0 \quad \forall t$$

It's crucial to note that with this BC $\int_{-\infty}^{\infty} p(x, t) dx < 1$ for $t > 0$.

Generalization

All the boundary condition expressed above can be easily generalized. It's interesting to highlight how the reflecting boundary becomes in higher dimensions. Suppose a diffusive process defined on a domain Ω with an impenetrable wall on the edge of the set, the hypersurface $\partial\Omega$. Writing \hat{n} as the (outward) unit normal vector at the wall, the no flux condition becomes:

$$\hat{n} \cdot \vec{j}|_{\partial\Omega} = 0$$

We have now presented all the necessary mathematical tools to understand the Fokker Planck equation.

1d free diffusion

Since we will need this theoretical solution in the following chapter we now want to show that a 1d free diffusion with the following initial and boundary condition:

Initial condition

$$p(x, 0) = \delta(x)$$

Boundary condition

$$\int_{-\infty}^{\infty} p(x, t) dx = 1$$

Has the following probability density function:

$$p(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}$$

In order to verify the equation we will compute the derivative

$$\begin{aligned}\frac{\partial p(x, t)}{\partial t} &= \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \left(\frac{x^2}{4Dt^2} - \frac{1}{2t} \right) \\ \frac{\partial p(x, t)}{\partial x} &= \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \left(-\frac{2x}{4Dt} \right) \\ \frac{\partial^2 p(x, t)}{\partial x^2} &= \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \left(\frac{x^2}{4D^2t^2} - \frac{1}{2Dt} \right)\end{aligned}$$

By equating the two we get:

$$\frac{p(x, t)}{2t} \left(\frac{x^2}{2Dt} - 1 \right) = D \frac{p(x, t)}{2tD} \left(\frac{x^2}{2Dt} - 1 \right)$$

We can easily observe that the integral $\int_{-\infty}^{\infty} p(x, t) dx = 1$ for all t .

1.4 First Passage Time

In this section we are going to define the concept of first passage time and then highlight some way to compute its probability density function.

Definition

Suppose a 1d stochastic process $x(t)_{t \geq 0}$ defined on a space Ω with boundaries $\partial\Omega$, and initial condition x_0 . Then the first passage time is defined as:

$$T = \inf\{t \geq 0 | X(t) \in \partial\Omega\} \quad (1.7)$$

$T(\partial\Omega|x_0)$ is clearly a random variable so it's natural to look at its probability distribution $\mathcal{T}(\partial\Omega; t|x_0)$:

$$\mathcal{T}(\partial\Omega; t|x_0) = \mathbb{P}\{T(\partial\Omega|x_0) < t | X(0) = x_0\}$$

Where \mathbb{P} is the probability that the first passage time is smaller than a certain t . Roughly speaking $\mathcal{T}(\partial\Omega; t|x_0)$ represents the fraction of realizations "escaped" from the region Ω during the time interval $[0, t]$. From this we can easily define the probability density function:

$$p_T(\partial\Omega; t|x_0) dt = \{t \geq T(\partial\Omega|x_0) \geq t + dt | X(0) = x_0\} = d\mathcal{T} \quad (1.8)$$

Survival probability

One can then define the survival probability as:

$$\mathfrak{S}(\partial\Omega; t|x_0) = 1 - \mathcal{T}(\partial\Omega; t|x_0)$$

We can now compute the mean first passage time or MFPT (to ease notion we will drop the dependency on $\partial\Omega$):

$$\mathbb{E}\{T(x_0)\} = \int_0^\infty t p_T(t) dt = - \int_0^\infty t \frac{dS}{dt}(t) dt = \int_0^\infty dt S(t) \quad (1.9)$$

We are now going to present some recently published results that set the framework for the explanation of the redundancy principle in a mathematically rigorous way. In order to do so we will look at extreme first passage time rather than the normal first passage time. Firstly we are going to define what an extreme first passage time is.

1.5 Extreme first passage time

Consider $N \gg 1$ independent and identically distributed realization of a Brownian motion. Let $\{\tau_1, \tau_2, \dots, \tau_N\}$ be their N identically distributed first passage time to find some target, we define the extreme first passage time as:

$$T_{1,N} := \min\{\tau_1, \tau_2, \dots, \tau_N\} \quad (1.10)$$

The quantity defined above is extremely interesting in fields like biology because searching processes are often characterized by a huge number of particle trying to reach the target. Furthermore since these processes are considered successful when only a bunch or just one of these searchers reach the target, it's more interesting to look at the behaviour of the extreme first passage time among many particle rather than looking at the mean first passage time.

We are now going to state the following theorem, that very strongly dictate the asymptotic behaviour of the extreme first passage time.

1.6 Lawley's theorem

For any fixed $m \geq 1$ and $k \geq 1$ the m th moment of the k th fastest first passage time satisfies:

$$\mathbb{E}\{(T_{k,N})^m\} \sim \left(\frac{L^2}{4D \ln N}\right)^m \quad (1.11)$$

Where L is the geodesic distance between the initial position of the particle and the closest point of the absorbing boundary. This powerful theorem holds for very generic circumstances. For instance it's true in any dimension d and for any space dependent diffusion coefficient. In appendix we provide a proof, for a more in depth explanation see Ref. [2].

Chapter 2

2d diffusion on a surface

2.1 Ligando-receptor problem

In biology many functions or processes are activated when a particle called ligando reaches an activation site, which is often a protein called receptor. This general interaction can vary widely depending on the context, and obtaining a complete understanding of it will have important effects on biophysics, chemistry and pharmacy. In this chapter we will try to solve a particular ligando-receptor problem. In our model the small cell is a dimensionless point, moving randomly on the surface of a bigger cell that we have modeled as a circle. This geometry is almost equivalent to a periodic boundary condition imposed on a infinite plane.

2.2 Solution

We want to solve the Fokker Planck equation on a 2d geometry with circular symmetry, we define two radii, R_b and r_a . The first is the radius where we impose the reflective boundary, this circular boundary $S+$ will work almost as a periodic boundary condition, simulating the presence of many absorption point. The second radius is the one where we will enforce the absorbing boundary condition, the small circumference $S-$ that simulates the dimension of the absorption point. We translate the boundary condition on the two radii as follows:

$$\frac{d}{dt}p(\vec{r}, t) = D\Delta p(\vec{r}, t) \quad (2.1)$$

Boundary condition

$$p(r_a, t) = 0 \quad \hat{n} \cdot j_r(R_b, t) = 0$$

Initial condition:
spherical shell

$$p(\vec{r}, t = 0) = \frac{\delta(r - r_0)}{2\pi r_0}$$

Firstly we recall the 2d Laplacian operator

$$\Delta f = \frac{df}{dr^2} + \frac{1}{r} \frac{df}{dr} + \frac{1}{r^2} \frac{df}{d\theta}$$

Firstly we note the symmetry of the problem that enables us to neglect the angular variable. This symmetry allows us to transform the variable \vec{r} into a scalar variable r . We want to solve this partial differential equation written below using the green's function method for solving differential equation. For reference see [1].

$$D\left[\frac{dp}{dr^2} + \frac{1}{r} \frac{dp}{dr}\right] = \frac{d}{dt}p(r, t)$$

To solve the problem firstly we will Laplace transform the equation. We can use the well-known relation (proof in Appendix):

$$\mathcal{L}\left[\frac{d}{dt}p(r, t)\right] = s\tilde{p}(r, s) - p(r, 0)$$

$$D[\tilde{p}''(r, s) + \frac{1}{r}\tilde{p}'(r, s)] - s\tilde{p}(x, r) = -\frac{\delta(r - r_0)}{2\pi r_0}$$

Where the prime denotes differentiation with respect to r . To simplify the notation we introduce the dimensionless parameter $x = r\sqrt{s/D}$. The equation now reads:

$$\tilde{p}''(x, s) + \frac{1}{x}\tilde{p}'(x, s) - \tilde{p}(x, s) = \frac{\delta(x - x_0)}{2\pi x_0 D} \quad (2.2)$$

Where now the prime denotes differentiation with respect to x . We rewrite the two radii where we have imposed the boundary condition to match the new dimensionless parameter: $x_+ = R_b\sqrt{s/D}$ e $x_- = r_a\sqrt{s/D}$. We can now solve the equation using the Green's method. Firstly we solve the equation on each of the two subdomain $x > x_0$ e $x < x_0$, where x_0 is the initial condition. In each of the two the region the solution is given by the modified Bessel function of the first and second type.(In Appendix)

The equation is second order differential equation, so its general solution is a linear combination of two linearly independent solutions. In order to be clear we will denote with $\tilde{p}_<$, $\tilde{p}_>$ the probability density function in the two subdomains, $x < x_0$ and $x > x_0$ respectively.

$$\tilde{p}_< = AI_0(x) + BK_0(x)$$

$$\tilde{p}_> = CI_0(x) + DK_0(x)$$

We now want to obtain the impose the boundary condition on the two radii. Firstly we will impose it in x_- . The equation we need to solve is:

$$\tilde{p}_<(x_-) = A[I_0(x_-) + \frac{B}{A}K_0(x_-)] = 0$$

Now we can redefine the constant A e and by solving the equation we find that $\tilde{p}_<$ has the following form:

$$\tilde{p}_<(x) = A[I_0(x)K_0(x_-) - K_0(x)I_0(x_-)]$$

We now repeat the same process for x_+ . Since the Bessel functions depends solely on the scalar variable x , the boundary condition becomes $\frac{d}{dx}\tilde{p}_>(x_+) = 0$.

We now recall the following property of the Bessel functions: $I_0'(x) = I_{-1}(x)$ e $K_0'(x) = -K_{-1}(x)$

We want to solve

$$\frac{d\tilde{p}_>}{dx}(x_+) = C[I_0(x_+) + \frac{D}{C}K_0(x_+)] = 0$$

Using the property written above we obtain the following solution:

$$\tilde{p}_>(x) = C[I_0(x)K_{-1}(x_+) + K_0(x)I_{-1}(x_+)]$$

We are going to introduce the following quantities to ease notation: $f_{0,0}(a, b) = I_0(a)K_0(b) - K_0(a)I_0(b)$
 $f_{0,-1}(a, b) = I_0(a)K_{-1}(b) + K_0(a)I_{-1}(b)$.

After this definition we can rewrite the probability density function we found above as:

$$\tilde{p}_>(x) = Af_{0,-1}(x, x_+)$$

$$\tilde{p}_<(x) = Bf_{0,0}(x, x_-)$$

We now define the following quantities: $x_> = \max(x, x_0)$ and $x_< = \min(x, x_0)$. We have two solutions, one for each subdomain. We are interested in solution which are continue in all the domain, so we need to adjust the two constants to get a continue solution in x_0 . The results we get is the following:

$$\tilde{p}(x, s) = A' f_{0,0}(x_<, x_-) f_{0,-1}(x_>, x_+) \quad (2.3)$$

We now show that Eq (2.3) is continuous:

$$\begin{aligned} \lim_{x \rightarrow x_0^+} A' f_{0,0}(x_0, x_-) f_{0,-1}(x, x_+) &= A' f_{0,0}(x_0, x_-) f_{0,-1}(x_0, x_+) \\ \lim_{x \rightarrow x_0^-} A' f_{0,0}(x, x_-) f_{0,-1}(x_0, x_+) &= A' f_{0,0}(x_0, x_-) f_{0,-1}(x_0, x_+) \end{aligned}$$

The equation we have found is continuous in x_0 and has a peculiar behaviour. The two terms behave as a constant when the particle is in the other subdomain and as the probability we have found above in the other case.

We are now interested in computing the value of the constant A' . In order to do so we have to integrate equation (2.2) over the discontinuity.

$$\tilde{p}' > |_{x_0^+} - \tilde{p}' < |_{x_0^-} = \frac{1}{2\pi x_0 D}$$

Using the already cited property of the Bessel function $I_0'(x) = I_{-1}(x)$ $K_0'(x) = -K_{-1}(x)$ and the so called Wronskian relation, which we present below:

$$I_{-1}(x_0)K_0(x_0) + K_{-1}(x_0)I_0(x_0) = \frac{1}{x_0}$$

we get:

$$\tilde{p}(x, s) = \frac{f_{0,0}(x_<, x_-) f_{0,-1}(x_>, x_+)}{D 2\pi f_{0,-1}(x_-, x_+)} \quad (2.4)$$

Equation (2.4) is the Laplace transform of the probability density function of our diffusive process. The results obtained cannot be anti-transform easily. Nonetheless with this result we can obtain interesting quantity. For instance we can obtain the Laplace transform of the total current by integrating over the absorbing boundary the derivative of the probability density function:

$$J(x_0) = \int_{S^-} D \frac{dp}{dr} |_{r=R} dr \quad (2.5)$$

The result we get is: (for Reference see [1])

$$J(x_0) = \frac{f_{0,-1}(x_0, x_+)}{f_{0,-1}(x_-, x_+)} \quad (2.6)$$

. The expression we obtained is difficult to anti-transform, but can be used to predict the asymptotic behaviour of particular case. For instance we are going to focus on the asymptotic behaviour of the first passage time probability as the reflecting boundary $R_b \rightarrow \infty$.

In this limit the current simplifies to: (for Reference see [1])

$$J_-(x_0) \rightarrow \frac{K_0(x_0)}{K_0(x_-)}$$

In this case expanding $J(x_0)$ for small s gives us:

$$J(x_0) \simeq 1 - 2 \frac{\ln(x_0/r_a)}{\ln(s)}$$

We observe now that since the coefficient of zeroth power is equal to 1 the probability of reaching the absorption point is 1, since all the other terms of the expansion vanish as $s \rightarrow 0$ this means that every realization of the Brownian motion will reach the absorption point.

2.3 Asymptotic behaviour of $J(x_0)$

We are now going to present a mathematical method to obtain information about the asymptotic behaviour of the function when only the Laplace transform is known. In particular we will use this method to obtain an asymptotic behaviour as $t \rightarrow \infty$ of $p_T(t, x_0)$. For a complete mathematical explanation with full justification of approximation see [1].

Firstly we define the generating function as follows:

$$F(z) = \sum_{t=0}^{\infty} F(t)z^t$$

Where $F(t)$ can be any function but in our particular example is the first passage time probability. We now transform the sum in integral (we assume that $F(t)z^t$ does not creates problem as $t \rightarrow 0$). We can then disregard the lower limit of integration since we are interested in the asymptotic behaviour:

$$\begin{aligned} &\simeq \int_0^{\infty} F(t)e^{-t \ln(1/z)} dt \\ &\sim \int_0^{t^*} F(t) dt \end{aligned}$$

with $t^* = [\ln(1/z)]^{-1}$. These simple algebraic steps involve several important simplifications and approximations that we can justify as follows:

- From the second line of the equation above, we see that $F(z)$ is essentially the Laplace transform of $F(t)$ with $s = \ln(1/z)$. We recall the definition of the Laplace transform:

$$\mathcal{L}[f(x, t)](s) = \int_0^{\infty} e^{-st} f(x, t) dt =$$

Because the $s \rightarrow 0$ limit of the Laplace transform corresponds to long-time limit of $F(t)$, we immediately infer that long-time behavior may also be obtained from the limit $z \rightarrow 1$ from below in the generating function.

- Sharp cutoff. In the last step, we replace the exponential cutoff in the integral, with characteristic lifetime $t^* = [\ln(1/z)]^{-1}$, with a step function at t^* . Although this crude approximation introduces numerical errors of the order of 1, we can prove (for complete proof see Ref. [1]) that the essential asymptotic behavior is preserved.

This allows us with a simple change of variable $s = (1 - z)$ to write the limit as $z \rightarrow 1$ in the following way:

Remembering that $t^* = [\ln(1/z)]^{-1} \sim 1/s$ for $s \rightarrow 0$ We can write:

$$F(s) \sim \int_0^{1/s} F(t) dt \tag{2.7}$$

Now we want to obtain information about the relationship between $\mathcal{F}(t) = \int_0^t F(t') dt'$ and the generating function $F(s)$. For $s = 1/t^*$ with $t^* \rightarrow \infty$ equation (2.7) becomes

$$F(s = 1/t^*) \sim \int_0^{t^*} F(t) dt = \mathcal{F}(t^*)$$

Now we rewrite the expansion for $s \rightarrow 0$ of the current $J(x_0)$:

$$J(x_0, s) \simeq 1 - 2 \frac{\ln(x_0/r_a)}{\ln(s)}$$

Using the equation written above we get:

$$J(s = 1/t^*) \sim \int_0^{t^*} J(t) dt = 1 - S(t^*)$$

Substituting $t^* = 1/s$ (this substitution yields a minus sign which we are going to ignore) we get: (we are going to call $t^* \rightarrow t$)

$$S(t) \sim \frac{A_2}{\ln(t)}$$

Where we are not interested in the value of A_2 , we now use equation (1.9) to get:

$$p_T(t, x_0) = -\frac{dS(t)}{dt} \propto \frac{1}{t \ln^2 t}$$

(for an in depth mathematical explanation see Section 1.4 of [2]). This results is important because it allows us to have a prediction that is easy to test and can be used to verify the results we obtain from the simulations of the following chapter.

2.4 Other dimensions

It's now interesting to look at the problem from another perspective, we want to highlight some observation that make the 2d case peculiar and obtain general information about the n-dimensional problem.

Asymptotic behaviour

Generalizing this problem with spherical symmetry in any dimension the 2d system is interesting because in the limit $R_b \rightarrow \infty$ 2d is the highest dimension for which the probability of the particle reaching the absorption site is equal to 1. This means that in higher dimension $d \geq 3$ some realization of the Brownian motion wander off to infinity without ever reaching the target. It's also worth noting that for $d > 4$ the hitting time for the returnees is finite. This means that a particle must reach the absorber in a finite time for large spatial dimension if it is to be absorbed at all. For a more complete discussion see [1].

Chapter 3

Simulations

3.1 1d Program

The program we created simulates the trajectory of a particle moving according to the 1d Langevin equation. In order to do so we use discrete time steps. The time evolution will be defined in a recursive manner, using the following equation:

$$x_{i+1} = x_i + 0.1 \cdot N(0, 1) \quad (3.1)$$

Where $N(0, 1)$ is a gaussian variable with mean 0 and variance $\sigma=1$. We now recall the over-damped 1D Langevin equation presented in the first chapter:

$$\frac{dx}{dt} = v_d(x) + \eta(t) \quad (3.2)$$

Where $v_d(x)$ is the drift velocity and $\eta(t)$ is a gaussian stochastic variable with the following property:

$$\langle \eta(t')\eta(t'') \rangle = 2D\delta(t' - t'') \quad (3.3)$$

Rearranging the equation in terms of an update condition and assuming $v_d(x) = 0$ we get:

$$x_{t+dt} = x_t + \eta(t)\sqrt{dt} \quad (3.4)$$

In conclusion, using the equations above, we get that the time evolution of our program can be written as follows:

$$x_{t+dt} = x_t + \sqrt{2Ddt}N(0, 1) \quad (3.5)$$

This allows us to fix the value of D by changing the multiplicative constant.

3.2 Simulation of free diffusion

In order to verify that the program was able to simulate the evolution of a particle moving according to a Brownian motion we developed a simulation of a 1d free diffusion. The analytical result expected is reported in the first chapter. We will test the results we get by fitting the simulation data with the equation obtained above:

$$p(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \quad (3.6)$$

In conclusion we expect the particle to be distributed in a gaussian way centered in 0 with standard deviation $\sigma = \sqrt{2Dt}$. The simulations have $N = 10000$ particles and are stopped at different time instant t , the time steps are $\delta t = 0.1$

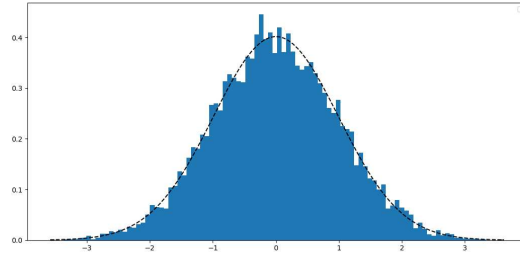


Figure 3.1: Particle distribution at time $t = 10$

After plotting the results we fit them using a gaussian curve. The parameters of the curve are:

	Mean	σ
Simulation	0.0008	0.985
Theoretical expectation	0	1

Table 3.1: Fit

We now repeat the simulation, looking at the particle distribution at a different time ($t = 5$). The results are the following:

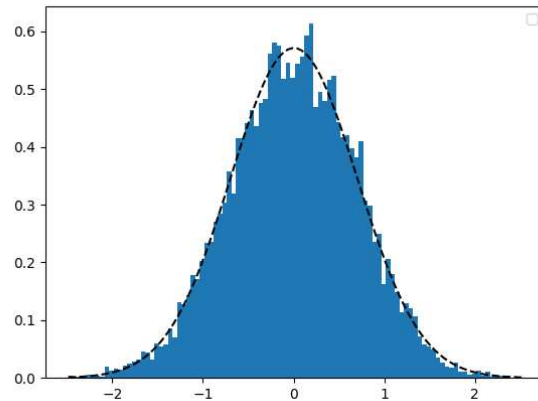


Figure 3.2: Particle distribution at $t = 5$

	Mean	σ
Simulation	0.003	0.487
Theoretical expectations	0	0.5

Table 3.2: Fit

Conclusion

These results allows us to say with great certainty that the theoretical expectations are matched, leading us to conclude that the program we wrote is indeed able to simulate the motion of a particle following the 1d Langevin equation.

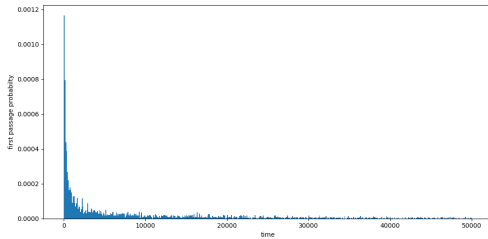
3.3 Confirmation of asymptotic behaviour

Now that we can confidently say that we can simulate the trajectory of a diffusive particle we are now interested in verifying the results we obtained on the asymptotic behaviour of the first time passage probability of the 2d problem on circular geometry that we solved analytically in the previous chapter. In order to do so we generalized the time evolution of the particle as follows, going from 1d to 2d:

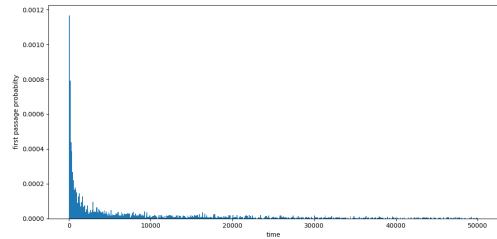
$$x_{i+1} = x_i + N(0, 1) \cdot 0.1$$

$$y_{i+1} = y_i + N(0, 1) \cdot 0.1$$

We then implemented a new part of the program that controls if the particle is inside the circle after every iteration and stops the simulation when the absorption point is reached. Since the asymptotic behaviour we have obtained is a limit as the reflecting boundary goes to infinity we compared two simulation: one where we enforced the boundary at a distance $R = 50$ and one where we let the particle move freely. Our intention with this simulation is to verify if the distance chosen is enough to justify the assumption of $R_b \rightarrow \infty$. In both cases the initial condition where $x_0 = 1, y_0 = 1$. We now report the results:



(a) No boundary



(b) Boundary

As one can see the results are really similar. We are now interested in observing the expected $\propto \frac{1}{t \ln^2 t}$ asymptotic behaviour. In order to confront data with expectations we perform a least square fitting on big time, excluding from the fit the small time regime where we have no information. The results we get are the following:

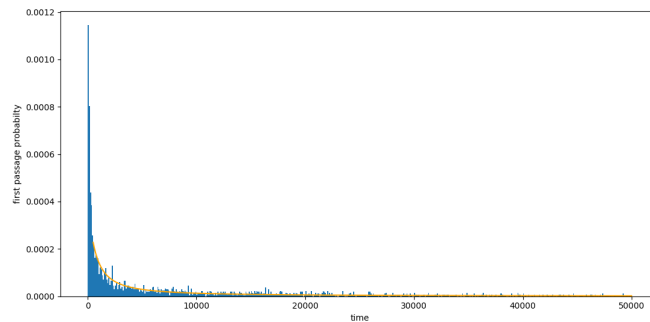


Figure 3.4: Least square fitting

Conclusion

The simulation we have run shows the asymptotic behaviour expected. This accordance with the theoretical expectation allows us to say that the program we wrote behave according to expectation.

3.4 Confirmation of Lawley's theorem

After verifying all the preparatory test we now are interested in providing a simple empirical evidence for the mathematical explanation of the redundancy principle. More specifically we want to show a confirmation of the Lawley's theorem on the simple 2d model of the ligando-receptor problem we have created. We expect to see a particular asymptotic behaviour as the number of ligandos increase. In order to do so we have created a large number of simulations with particle number ranging from $N = 100$ to $N = 10000$. We created a population of simulation that we used to extract a mean extreme first passage time for each particle number. The parameter that we chose for the simulations are the following: reflecting radius $R = 5$ absorbing radius $r = 0.1$ initial position $x_0 = 2.5$, $y_0 = 2.5$ and the time steps where of size $\Delta t = 0.1$.

Particle number	Mean extreme first passage time	σ
100	19.86	1.05
200	16.33	0.99
300	15.11	0.69
500	13.7	0.58
1000	12.39	0.55
5000	9.18	0.49
7000	8.67	0.35
10000	8.38	0.25

Table 3.3: Simulation data

These are the results of the simulations. We are now interested in studying the relationship between the mean extreme first passage time and the number of particle. In order to verify the theorem we are going to do a linear regression between the mean extreme first passage time and the inverse of the logarithm of the number of particle. The results are the following:

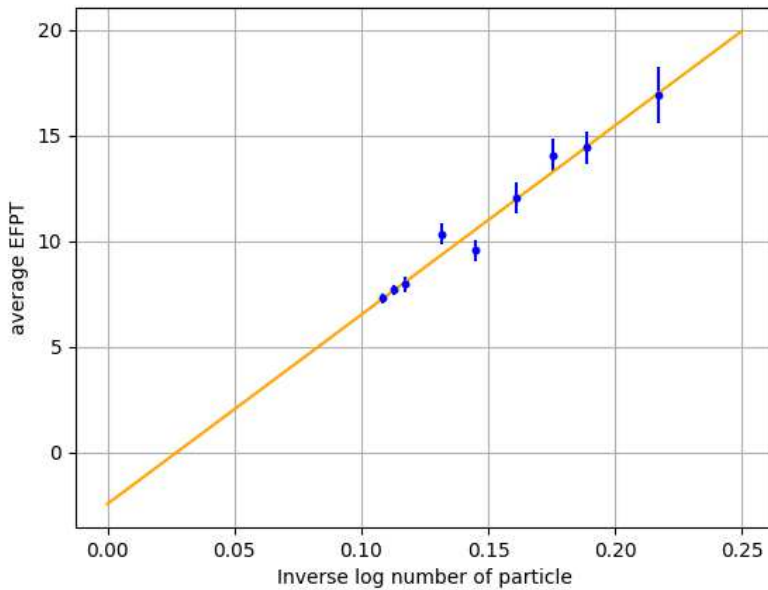


Figure 3.5: Linear regression

	m	q
Value	103.76	-2.93
σ	5.94	0.77

Table 3.4: Fit parameters

The data clearly follows a linear trend. In order for the theorem to be confirmed we now need to verify that the linear coefficient we obtained empirically is compatible with the theoretical expectation. According to the Lawley's theorem we can calculate the linear coefficient as $\frac{L^2}{4D}$. L is the geodesic distance between the starting point and the closest absorption point, that we can compute as follows:

$$L^2 = (x_0 - x_{cerchio})^2 + (y_0 - y_{cerchio})^2 = 2 * (2.5 - 0.224)^2 \simeq 10.36$$

Looking at the time evolution equation that we have written above we can compute the value of the diffusion coefficient remembering that the equation (3.5) can be written as (where 4 is sensible to dimensionality: $2 \cdot d$):

$$x_{t+dt} = x_t + \sqrt{4Ddt}N(0, 1) \quad (3.7)$$

We can now calculate the value of 4D as:

$$4D = \frac{0.01}{0.1} = 0.1$$

This yield us:

$$m_{exp} = \frac{L^2}{4D} = 10.36/0.1 = 103.6$$

As we can see the results are incredibly compatible between one another.

Conclusion

The results shown above allows to say with great confidence that the results we have obtained confirm the Lawley's theorem. This simple simulation gives an idea of how having a large number of searcher can impact the average time needed for the fastest particle to reach the activation site. Looking at this example makes it easy to understand why many biological searching process seem to be characterized by a redundant number of particle. This excess number has the clear function of reducing the time needed, allowing for a faster activation or suppression of a particular biological function. For a more complete discussion that encompasses the role of redundancy also in other field like chemistry see Ref. [4]

Conclusions

In this thesis we have presented the concept of Brownian motion, with a particular interest in the first passage time and its application to the ligando-receptor problem, a specific problem in biology concerned with computing the dynamical properties of exchanges between small cells and bigger biological structure. After highlighting the necessary mathematical tool we solved analytically a 2d version of the Fokker Planck equation, inspired by the ligando-receptor problem. This analytical solution of the simple model we developed gave us some theoretical prediction that we later verified using a C++ simulation of the same problem, the results we got where encouraging since the analytical expectation strongly agreed with the computational results. At the same time we presented the concept of extreme first passage time, which is an advanced concept closely linked with the first passage time, and some recently published results that try to use the asymptotic behaviour of this quantity to explain mathematically the so called redundancy principle. The redundancy principle is a concept in biology that expresses the need of many copies of the same entity to fulfill a biological function. Finally in the last section using the simulation developed for the ligando-receptor problem we show mathematically the effect that many ligandos have on the time scale of the simple problem we solved.

The results presented in this thesis offer some interesting insights. To obtain more general and physically more relevant results one could relax some of the assumption we made, for instance one could work with a spatially dependent diffusion coefficient, this would allow to better model the ligando-receptor problem, since often the ligando travels trough different substances to reach the absorption site. Furthermore it could test the Lawley's theorem in a more complex setting.

Appendix

We are now going to outline a simplified proof for the Lawley's theorem Eq (1.11):

Proof

Let $S(t) := P(\tau_1 > t)$ denote the survival probability of a single first passage time. The survival probability of the fastest first passage time for τ_1, \dots, τ_N independent and identically distributed realization is $P(T_1, N > t) = P(\min\{\tau_1, \dots, \tau_N\} > t) = [S(t)]^N$ since the probability for each one to survive is $[S(t)]$, the probability for all N to survive is the product. The mean of any non-negative random variable $Z > 0$ is $\mathcal{E}[Z] = \int_0^\infty P(Z > z) dz$. Therefore, the mean fastest first passage time:

$$\mathcal{E}[T_1, N] = \int_0^\infty [S(t)]^N dt$$

Since $S(t)$ is a decreasing function of time, it is clear from that the large N asymptotic of $\mathcal{E}[T_1, N]$ are determined by the short time behavior of $S(t)$. The following theorem determines these asymptotic in terms of the short time behavior of $S(t)$ on a logarithmic scale. Let $\tau_{n=1}^\infty$ be a sequence of independent and identically distributed non-negative random variables with survival probability $S(t) := P(\tau_1 > t)$. Assume that

$$\int_0^\infty [S(t)]^N dt < \infty \quad (1)$$

for some $N \gg 1$, and assume that there exists a constant $C > 0$ so that:

$$\lim_{t \rightarrow 0^+} t \ln[1 - S(t)] = -C < 0 \quad (2)$$

We now sketch the proof of Theorem. The assumption in (2) means roughly that:

$$S(t) \simeq 1 - e^{-C/t} \quad \text{for } t \ll 1$$

Now, for a one-dimensional, pure diffusion process with unit diffusivity starting at the origin, let $\tau(l)$ denote the first time the process escapes the interval $(-2l, 2l)$. The survival probability $S_l(t) := P(\tau(l) > t)$ is shown to satisfies (for Reference see [2]):

$$S_l(t) \simeq 1 - e^{-l^2/t} \quad \text{for } t \ll 1$$

Therefore, taking $l_\pm = \sqrt{C} \pm \epsilon$ for small $\epsilon > 0$ yields

$$S_{l_+}(t) \geq S(t) \geq S_{l_-}(t) \quad \text{for } t \ll 1$$

Hence, for sufficiently large N we have the bounds:

$$\int_0^\infty [S(t)_{l_+}]^N dt \geq \int_0^\infty [S(t)]^N dt \geq \int_0^\infty [S(t)_{l_-}]^N dt$$

Since the large N and the fact that $S(t)$ is a monotonically decreasing function the behaviour of these integrals is determined by the short time behavior of their integrands. Furthermore, it's known (we are not going to prove it, for Reference see [2]) that:

$$\int_0^\infty [S_{l_\pm}(t)]^N dt \simeq \frac{C \pm \epsilon}{\ln N} \quad \text{as } N \rightarrow \infty$$

Noting that ϵ is arbitrary completes the argument. [2]

3.4.1 Laplace transform of derivative

$$\begin{aligned}\mathcal{L}\left[\frac{d}{dt}f(r,t)\right](s) &= \int_0^\infty e^{-st} \frac{df}{dt} dt = \\ &= [e^{-st}f(r,t)]_0^\infty - \int_0^\infty (-s)e^{-st}f(r,t)dt = s\tilde{f}(r,s) - f(r,0)\end{aligned}$$

3.4.2 Bessel function

Bessel functions are defined as the function that solve this second order differential equation:

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - \alpha^2)y = 0$$

Being a second order differential equation this one has two linearly independent solutions. These two solutions can be written in different ways depending on the problem. In this case the two functions we are using are the modified Bessel functions, defined from regular Bessel functions in the following way:

$$\begin{aligned}I_\alpha &= i^{-\alpha} J_\alpha(ix) \\ K_\alpha &= \frac{\pi}{2} \frac{I_{-\alpha} - I_\alpha}{\sin(\alpha\pi)}\end{aligned}$$

Bibliography

- [1] Redner S. A guide to first-passage processes. Cambridge: Cambridge University press; 2001.

- [2] Lawley SD. Universal formula for extreme first passage statistics of diffusion. *Physical Review E* 101 (1), 012413

- [3] Enzo Orlandini and Fulvio Baldovin. Lecture note on stochastic mechanics

- [4] *Physics of Life Reviews*, Volume 28, 2019

- [5] Daniel T. Gillespie. The mathematics of brownian motion and johnson noise. *American Journal of Physics*, 64(3):225–240, 1996.