

Mathematical Modeling of Subdiffusion on Cell Membranes

Tomoaki Fujii
Graduate School of Informatics, Kyoto University

April 27, 2018

Abstract

Various mathematical models of subdiffusion have already been established such as fractional calculus, continuous time random walk, and fractional Brownian motion. Many researchers tend to just choose one of the models to fit given data well. These models, however, assume that the variance of position follows the power-law: i.e., t^α for $0 < \alpha < 1$. This is not necessarily true of all of the subdiffusion cases. In this article, we work on modeling two-dimensional subdiffusion on the plasma cell membrane. We show that the asymptotic behavior of the process in the fence model is approximated by the compound Poisson process, which leads to transient subdiffusion rather than the power-law subdiffusion. We also show the numerical results that the hierarchical fence structure leads to the power-law subdiffusion as well as similar results of other hierarchical structures. These results imply that we require some hierarchical structures when using a model of the power-law subdiffusion.

Contents

1	Introduction	2
2	Derivation of the Langevin Equation	3
2.1	Model Description	4
2.2	Markov Approximation	7
2.3	The Closeness of Paths	12

3	Langevin Equation and Brownian Motion	18
3.1	The Strong Solution of the Langevin Equation	19
3.2	Approximation by Brownian Motion	20
4	Fence Model and Subdiffusion	22
4.1	Fence Model	22
4.2	Discretized Brownian Motion with Fences	23
4.3	Hopping Model	25
4.4	Transient Subdiffusion	25
4.5	Power-Law Subdiffusion	28
5	Discussion	30
6	Acknowledgements	33
	Appendices	33
A	Elastic Collision	33
B	Stochastic Fubini	36

1 Introduction

It is known that a pollen floating on pure water looks moving at random. This movement was named the Brownian motion after the botanist, Robert Brown, who discovered this phenomena in 1812. Almost one hundred years later, Einstein[8] established its mathematical explanation in 1905. That was the strong evidence to make sure that various medium consisted of many small particles. The mathematical theory for Brownian motion was formalized as differential equations by Langevin in 1908. He proposed the Langevin equation on the paper with the French title, "Sur la théorie du mouvement brownien", which meant "On the Theory of Brownian Motion" in English [15]. This model has widely been applied in various fields such as physics, biology, and finance [27] [20] [4] etc. In practical situation, a process \vec{X} governed by the Langevin equation is often identified by the variance of position proportional to t ; i.e., $E[\|\vec{X}_t - \vec{X}_0\|^2] \propto t$. We define a process with such variance of position as the normal diffusion.

This property, however, is not the case if there are inhomogeneous structures in the medium. For example, there are many reports that diffusion coefficient is not constant and is nonincreasing in the medium filled with obstacles [24] etc. To characterize these processes, we introduce the diffusion

coefficient:

$$D(t) = E[\|\vec{X}_t - \vec{X}_0\|^2]/t \text{ for } t > 0.$$

Note that $D(\cdot)$ is constant if \vec{X} is the normal diffusion. We define a process with $D(\cdot)$ that is not constant and is nonincreasing as subdiffusion.

In physics and mathematics literatures, subdiffusion is often defined as a process with the variance of position proportional to t^α for $0 < \alpha < 1$ [17]. For convenience, we call such process as the power-law subdiffusion in this paper. Various models consistent with properties of the power-law subdiffusion have been suggested such as continuous time random walk(CTRW) with infinite mean waiting time[16], fractional Brownian motion(FBM)[9], and fractional partial differential equation(FPDE)[17]. When establishing a mathematical model to explain experimental data, many researchers just choose one of these models that fits given data well. Not all of the subdiffusion, however, do not have the power-law diffusion coefficient[26]. Some may have a finite asymptotic diffusion coefficient as time goes to infinity. Thus, we have to look into precise mechanics for each phenomena before trying to use one of models for the power-law subdiffusion.

In this article, we will work on modeling subdiffusion on the plasma cell membrane. It is known that the plasma cell membrane works as a two-dimensional liquid solvent for diffusion of proteins [23]. This diffusion is known to be described as subdiffusion rather than the normal diffusion [7]. The "fence" structure is considered as one of the causes of subdiffusion [10]. There are filament substances called cytoskeletons inside a cell and these substances bind some proteins on a straight line. Proteins immobilized on the straight line works like a fence to hinder the diffusion. To specify how the fence structure has an effect on the diffusion, we first clarify underlying process hindered by fences. We derive that a molecule immersed in a liquid medium consisting of small particles is governed by the Langevin equation in the next section. As a matter of fact, this process can be approximated by a standard Brownian motion multiplied by a constant under certain condition. We show that the standard Brownian motion hindered by the fence has a finite asymptotic diffusion coefficient, which causes transient subdiffusion (defined in section 4) rather than the power-law subdiffusion. We also show numerical results that hierarchical fence structure leads to the power-law subdiffusion.

2 Derivation of the Langevin Equation

The mathematical model of nonequilibrium system with differential equations was suggested a century ago by Langevin and this model successfully

described various phenomena. Its mathematical derivation, however, had been left for years. In 1981, Dürr, Goldstein, and Lebowitz [6] led a mathematical derivation of the Langevin equation in the sense of a certain limit. In this section, we will show the precise derivation of [6] with some modifications and consider physical interpretations for assumptions of the model. Though we discuss three dimensional case here, we immediately have the same argument by fixing one coordinate to a constant.

2.1 Model Description

Consider a large particle(e.g., molecule) immersed in a fluid media consisting of small particles(e.g., atoms) with an equilibrium state; e.g., a pollen floating on pure water. It is known that collisions with atoms agitates the molecule and that the behavior of the molecule of mass M follows the Langevin equation:

$$Md\vec{V}_t = -\gamma\vec{V}_t dt + \sigma d\vec{W}_t, \quad (2.1)$$

where γ , σ are positive numbers; $W = \{W_t; 0 \leq t < \infty\}$ denotes a standard Brownian motion; \vec{V}_t denotes velocity of the molecule at t .

The right hand of the Langevin equation has two terms: a friction term and a noise one. Both of them originate from collisions with atoms. Thus, these two terms have to hold a certain relation. This relation is known as the fluctuation-dissipation theorem, which was explained at [12] up to extensions of generalized Langevin equations. For (2.1), this relation is described as

$$\sigma = \sqrt{2\gamma k_B T}, \quad (2.2)$$

where k_B is the Boltzmann constant and T is temperature of the heat bath. We will derive (2.1) satisfying (2.2) in the rest of part of this section.

Let us assume that the molecule and atoms interact only through elastic collisions. An elastic collision is an encounter between two bodies in which the total kinetic energies and momentum of two bodies are conserved. Suppose that all of the atoms have the same mass m and that the molecule is a sphere with radius r . We denote projections of vector \vec{X} to the normal direction and the tangent plane of the sphere of the molecule sphere by \vec{X}_n and \vec{X}_t respectively. We here take the direction from the sphere to the center as positive in the normal direction. As in appendix A, relations for velocities of pre- and post-collision are written as

$$\tilde{\vec{V}}_t^{(m)} = \vec{V}_t^{(m)}, \tilde{\vec{v}}_t = \vec{v}_t, \quad (2.3)$$

$$\tilde{\vec{V}}_n^{(m)} = \frac{M-m}{M+m} \vec{V}_n^{(m)} + \frac{2m}{M+m} \vec{v}_n, \quad (2.4)$$

$$\vec{v}_n = -\frac{M-m}{M+m}\vec{v}_n + \frac{2M}{M+m}\vec{V}^{(m)}, \quad (2.5)$$

where $(\tilde{\cdot})$ denotes a value of post-collision; $\vec{V}^{(m)}$ and \vec{v} are velocities of the molecule and an atom respectively. Each collision drives the following jumps:

$$\Delta\vec{V}_n^{(m)} = \frac{2m}{M+m}(\vec{v}_n - \vec{V}_n^{(m)}), \quad (2.6)$$

$$\Delta\vec{v}_n = -\frac{2M}{M+m}(\vec{v}_n - \vec{V}_n^{(m)}). \quad (2.7)$$

To specify the jump distribution, we assume that the heat bath always holds thermal equilibrium state; i.e., atoms are distributed as

$$p(d\vec{x}, d\vec{v}) \propto \exp\left[-\frac{E(\vec{x}, \vec{v})}{2k_B T}\right] d\vec{x}d\vec{v}, \quad (2.8)$$

where $E(\vec{x}, \vec{v})$ the energy of an atom with position \vec{x} and velocity \vec{v} . We have the precise derivation of (2.8) in [14]. Since we now consider any potential fields in the heat bath, all of atoms behave as free particles. Thus, we have $E(\vec{x}, \vec{v}) = \frac{1}{2}m\|\vec{v}\|^2$, which does not depend on \vec{x} . For every $\vec{x} \in \mathbb{R}^3$, the distribution of atoms follows

$$f_m(\vec{v}) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} \exp\left[-\frac{\|\vec{v}\|^2}{2k_B T}\right].$$

This is called the Maxwell-Boltzmann distribution.

Consider collisions that happen on the molecule's sphere in $[\theta, \theta + d\theta] \times [\phi, \phi + d\phi]$. We denote by $d\Omega = \sin\theta d\theta d\phi$ and let dS be sets of points that belong to the infinitesimal spherical area. When an atom collides with the molecule of velocity \vec{v} in $[t, t + dt]$, the atom has to be inside

$$A = \{\vec{x} + (\vec{v} - \vec{V}_\tau^{(m)})\tau | \vec{x} \in dS, \tau \in [t, t + dt]\}.$$

The volume of A is calculated as $(v_n - V_{t,n}^{(m)})_+ r^2 d\Omega dt$, where X_n is normal component of a vector \vec{X} . Since the distribution of free particles does not \vec{x} , the collision rate should be proportional to the volume of A . Note that we use $(v_n - V_{t,n}^{(m)})_+$ instead of $(v_n - V_{t,n}^{(m)})$ because a collision happens only when $v_n \geq V_{t,n}^{(m)}$. Let us introduce a constant λ_m , and we write down the collision distribution for $[t, t + dt] \times [\theta, \theta + d\theta] \times [\phi, \phi + d\phi] \times [\vec{v}, \vec{v} + d\vec{v}]$ as

$$p(dt, d\Omega, d\vec{v} | \vec{V}^{(m)}) = \lambda_m r^2 (v_n - V_{t,n}^{(m)})_+ f_m(\vec{v}) d\Omega dt d\vec{v}.$$

Marginalizing out the tangent plane component of \vec{v} leads us to

$$p(dt, d\Omega, v_n | \vec{V}^{(m)}) = \lambda_m r^2 (v_n - V_{t,n}^{(m)})_+ f_m^1(v_n) d\Omega dt dv_n, \quad (2.9)$$

$$\text{where } f_m^1(v_n) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left[-\frac{v_n^2}{2mk_B T}\right].$$

If collisions are independent of one another, the process becomes a Lèvy process. Recollisions, however, may keep the history and break the independence assumption. To avoid this, we assume that a relaxation time from nonequilibrium state to thermal equilibrium state is fast enough. Consider a heat bath consisting of N particles, they have $6N$ parameters: positions and momentums. Landau-Lifshitz[14] explained that the equilibrium state are attained after the state explore $6N$ dimensional for long enough. Interactions among atoms drive transition from states to states. Thus, short relaxation time is introduced as an assumption that an atom becomes a part of the equilibrium state immediately after interaction with other atoms. More precisely, we assume that an atom does not hold the history after the velocity changes through interactions with the heat bath.

Consider a transformation of variables, $\sqrt{m}v_n = y$ at (2.9), and we obtain

$$\tilde{p}(dt, d\Omega, y | \vec{V}^{(m)}) = \lambda_m r^2 \left(\frac{y}{\sqrt{m}} - V_{t,n}^{(m)}\right)_+ f^1(y) d\Omega dt dy,$$

where $f^1(y) = \frac{1}{\sqrt{2\pi k_B T}} \exp\left[-\frac{y^2}{2k_B T}\right]$. When taking the limit with m to 0, we should fix the energy scale flowing into the molecule. Otherwise, the probability measure of $\vec{V}^{(m)}$ does not converge. Then, we should take $\lambda_m = \lambda/\sqrt{m}$, where λ is a positive number independent of m .

We consider the convergence of the model in a right continuous left limits path space. Let $(\Omega, \mathcal{F}, \mathbb{P}_m)$ be a probability space. For any $I = [0, T], 0 < T < \infty$, for each $\omega \in \Omega$, the formulae (2.4) and (2.9) determine right continuous left limits process $\{\vec{V}_t^{(m)}(\omega); t \in I\}$. Let $D(I)$ denote the space of right continuous left limits defined on I and let $\mathcal{B}(D(I))$ be the Borel algebra of the Skorokhod-Topology. We have the precise definition and properties of the Skorokhod-Topology in [3]. The above description defines a stochastic process $\vec{V}^{(m)} = \{\vec{V}_t^{(m)}; t \in I\}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_m)$. We have an induced probability measure of $\vec{V}^{(m)}$, P_m on $D(I)$:

$$P_m(A) \triangleq \mathbb{P}_m(\{\omega \in \Omega | \vec{V}^{(m)}(\omega) \in A\}),$$

for all $A \in \mathcal{B}(D(I))$.

We set

$$\Phi_i = \int_0^\infty y^i f^1(y) dy.$$

In the above setting, we have the convergence theorem as follows:

Theorem 2.1. P_m weakly converges to P_0 on $D(I)$ as $m \rightarrow 0$, where P_0 corresponds to the path measure of (2.1) with parameters

$$\gamma = \frac{16\pi}{3} \lambda r^2 \Phi_1, \quad \sigma = \left(\frac{16\pi}{3} \lambda r^2 \Phi_3 \right)^2,$$

which satisfy (2.2).

Its proof has two step:

- (1) Markov approximation,
- (2) The closeness of Markov approximation toward $\vec{V}^{(m)}$.

In the rest of this section, we go into the precise discussion of (1) and (2).

2.2 Markov Approximation

We shall construct an appropriate Markov approximation that converges to $\vec{V}^{(0)}$ as m goes to 0. Consider the case where a slow atom collides with a molecule. The slow atom may wander around the molecule even after the first collision and it recollides with the molecule. Recollisions depends on past collisions; i.e., a process with recollisions is not a Markov process in general. Therefore, we consider a process that neglects collisions with slow atoms in order to satisfy the Markov condition.

Let $\vec{U} = \{\vec{U}_t; t \in I\}$ be a process that neglects collisions with atoms with the velocity $\|\vec{v}\| \leq c_m := m^{-1/5}$. This process has no recollisions until $\|\vec{U}_t\| > c_m$. We define

$$\tau_U^m := \{t \in I; \|\vec{U}_t\| > c_m\}.$$

Let $\vec{V}^{(m)}$ be a Markov process that coincides with \vec{U}_t for $t \leq \tau_U^m$. We set

$$N_m(\vec{V}) = \int d\Omega \int_{c_m}^{\infty} (v_n - V_n)_+ f_m^1(v_n) dv_n. \quad (2.10)$$

As a modified jump distribution for $\vec{V}^{(m)}$, we define

$$g_m(\vec{V}, \vec{v}_n) = \begin{cases} \frac{1}{N_m(\vec{V})} (v_n - V_n)_+ f_m^1(v_n) & v_n \geq c_m \\ 0, & v_n < c_m \end{cases} \quad (2.11)$$

Let $\bar{\lambda}_m = \lambda_m r^2 N_m(\vec{0})$ be a collision rate, and we obtain a new collision distribution for $\vec{V}^{(m)}$,

$$\bar{p}(dt, d\vec{v}, d\Omega | \vec{V}^{(m)}) = \bar{\lambda}_m g_m(\vec{V}^{(m)}, \vec{v}_n) d\Omega dt dv_n. \quad (2.12)$$

Suppose that no recollisions happen for the Markov process, which makes all of collisions independent of each other. The Markov process generated by (2.12) is understood as a compound process, where each jump follows (2.11) and a frequency of collisions is determined by a Poisson process with intensity $\bar{\lambda}_m$.

In the following discussion, we will show that

Lemma 2.2. $\vec{V}^{(m)}$ weakly converges to $\vec{V}^{(0)}$ as $m \rightarrow 0$.

A proof of Lemma 2.2 is based on a semigroup theory. Let (Ω, \mathcal{F}, P) be a probability space equipped with a filtration $\{\mathcal{F}_t; t \in I\}$. Let $\{X_t, t \in I\}$ be a Markov process, i.e., for all $f \in B_b(\mathbb{R}^d)$, we have

$$E(f(X_t)|\mathcal{F}_s) = E(f(X_t)|X_s) \text{ a.s. } P.$$

Recall that $B_b(\mathbb{R}^d)$ is a Banach space with the norm

$$\|f\| \stackrel{\Delta}{=} \sup_{x \in \mathbb{R}^d} \|f(x)\|.$$

With a Markov process X , we associate a family of operator $(T_{s,t}, s \leq t, s, t \in I)$ from $B_b(\mathbb{R}^d)$ to a Banach space by the prescription

$$(T_{s,t}f)(x) \stackrel{\Delta}{=} E[f(X_t)|X_s = x].$$

For a stationary Markov process $E[f(X_t)|X_s = x] = E[f(X_{t-s}|X_0 = x)]$ leads to the definition $T_{t-s} \stackrel{\Delta}{=} T_{0,t-s} = T_{s,t}$.

Let $C_0(\mathbb{R}^d)$ denote a subset of $B_b(\mathbb{R}^d)$ that consists of functions vanishing to 0 at $\|x\| \rightarrow \infty$. If X is a Lévy process, it has the following properties:

- (1) $T_t : C_0(\mathbb{R}^d) \subseteq C_0(\mathbb{R}^d)$ for all $t \in I$
 - (2) $\lim_{t \rightarrow 0} \|T_t f - f\| = 0$ for all $f \in C_0(\mathbb{R}^d)$
- (2.13)

A Markov process that satisfies (2.13) is called a Feller process. As a subset of range of

$$D_A = \{\psi; \exists \phi_\psi \in B \text{ such that } \lim_{t \downarrow 0} \left\| \frac{T_t \phi - \phi}{t} - \psi_\phi \right\| = 0\}.$$

Then, we can define

$$A\psi = \lim_{t \downarrow 0} \frac{T_t \phi - \phi}{t} \equiv \psi_\phi$$

Actually, we can take $C_0(\mathbb{R}^d)$ as D_A .

We use the following lemma,

Lemma 2.3. Let X_n be a Feller process with sample path in $D(I)$ and the generator A_n . Suppose that X is a Feller process with sample path in $D(I)$ and the generator A . Let K be a core for A and suppose that $f \in K$ implies that $f \in D_{A_n}$ for sufficiently large n . Suppose that the initial distribution of X_n converges weakly to the initial distribution of X and suppose that

$$\limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \|A_n f(x) - A f(x)\| = 0$$

for all $f \in K$. Then

$$X_n \xrightarrow{\text{weak}} X.$$

We see the proof of Lemma 2.3 in [25].

A core K of A is a linear subspace of the domain $\text{Dom}(A)$ such that A is the closure of the restriction of A to K . To see if a subset of the domain is a core, we have the following lemma from [25],

Lemma 2.4. Let A be a generator of a Markov process and let K be a subspace of the domain $\text{Dom}(A)$. Then, K is a core if K is dense and $T_t K \subseteq K$.

Let C_c^∞ be a continuous function space with a compact support. Since C_c^∞ is dense in C_0 , it follows that C_c^∞ is a core of A . Hence, it suffice to consider functions over C_c^∞ .

By Lemma 2.3, it suffices to show that for all $h \in C_c^\infty(\mathbb{R}^3)$,

$$\limsup_{m \downarrow 0} \sup_{x \in \mathbb{R}^3} \|A^{(m)} h(x) - A^{(0)} h(x)\| = 0,$$

where $A^{(m)}$ and $A^{(0)}$ are the generators of $\vec{V}^{(m)}$ and $\vec{V}^{(0)}$.

For diffusion process, it is known that we have

Lemma 2.5. Let X a process driven by the following diffusion process,

$$dX_t = \vec{b}(X_t)dt + \sigma \cdot dW_t,$$

where W is a standard Brownian motion, $\vec{b} \in \mathbb{R}^d$, and σ is a positive definite symmetric $d \times d$ matrix. Then, for $f \in C_0^2(\mathbb{R}^d)$,

$$Af(X(t)) = [b_i \partial^i + \frac{1}{2} \sigma_{ik} \sigma_j^k \partial^i \partial^j] f(X(t)).$$

The proof is given in [11].

By Lemma 2.5, we can calculate the generator $A^{(0)}$ of $\vec{V}^{(0)}$,

$$A^{(0)} h(\vec{V}_t^{(0)}) = [-\gamma \vec{V}_t^{(0)} \cdot \vec{\nabla} + \frac{1}{2} \sigma^2 \Delta] h(\vec{V}_t^{(0)}),$$

for all $h \in C_c^\infty(\mathbb{R}^3)$.

We denote by $G_m(\vec{V}^0, d\vec{V})$ transition probability from \vec{V}^0 to $[\vec{V}, \vec{V} + d\vec{V}]$ by one jump. By definition, we have

$$T_t^{(m)} h(\vec{V}^0) = e^{-\bar{\lambda}_m t} [h(\vec{V}^0) + \bar{\lambda}_m t \int G_m(\vec{V}^0, d\vec{V}) h(\vec{V}) + O(t^2)].$$

Then,

$$\begin{aligned} A^{(m)} h(\vec{V}^0) &= \lim_{t \downarrow 0} \frac{1}{t} [e^{-\bar{\lambda}_m t} [h(\vec{V}^0) + \bar{\lambda}_m t \int G_m(\vec{V}^0, d\vec{V}) h(\vec{V}) + O(t^2)] - h(\vec{V}^0)] \\ &= \bar{\lambda}_m h(\vec{V}^0) + \bar{\lambda}_m t \int G_m(\vec{V}^0, d\vec{V}) h(\vec{V}) \end{aligned} \quad (2.14)$$

We use $q_m \triangleq \frac{2}{M+m}$ and $\vec{u}_n \triangleq \vec{v}_n - \vec{V}_n^{(m)}$. A first collision drives velocity to $\vec{V}^0 + q_m \vec{u}_n$. By the Taylor expansion,

$$\begin{aligned} h(\vec{V}^0 + q_m \vec{u}_n) &= h(\vec{V}^0) + q_m (u_{n,i} \partial^i) h(\vec{V}^0) + \frac{1}{2} q_m^2 (u_{n,i} u_{n,j} \partial^i \partial^j) h(\vec{V}^0) \\ &\quad + \frac{1}{6} q_m^3 (u_{n,i} u_{n,j} u_{n,k} \partial^i \partial^j \partial^k) h(\vec{V}^0 + \rho q_m \vec{u}_n), \end{aligned} \quad (2.15)$$

where $\rho \in [0, 1]$. By using (2.15), (2.14) can be written by the following terms,

$$J_i \triangleq \frac{\bar{\lambda}_m}{N_m(\vec{V}^0)} \int d\Omega \int_{c_m}^\infty (v_n - V_n)_+^i \frac{q_m^{i-1}}{(i-1)!} e_{n,k_1} \dots e_{n,k_{i-1}} \partial^{k_1} \dots \partial^{k_{i-1}} h(\hat{V}_i) f_m^1(v_n) dv_n,$$

where $\hat{V}_i = \vec{V}^0$ for $i = 1, 2, 3$ and $\hat{V}_4 = \vec{V}^0 + q_m \vec{u}_n$.

We shall evaluate for $i = 4$ at first. Since $h \in C_c^\infty$, we have $\sup_{x \in \mathbb{R}^3} |(e_{n,k} \partial^k) h(x)| = M < \infty$. The symmetry of the distribution f_m^1 allows us to evaluate (2.10) as ,

$$\begin{aligned} N_m(\vec{V}) &= \int d\Omega \frac{1}{2} \left[\int_{c_m}^\infty (v_n + |V_n|) f_m^1(v_n) dv_n + \int_{\max(c_m, |V_n|)}^\infty (v_n - |V_n|) f_m^1(v_n) dv_n \right] \\ &= \int d\Omega \int_{c_m}^\infty v_n f_m^1(v_n) dv_n + \frac{1}{2} \int d\Omega \int_{c_m}^{\max(c_m, |V_n|)} (|V_n| - v_n) f_m^1(v_n) dv_n \end{aligned}$$

Thus, $N_m(\vec{V}) \leq N_m(0)$ and equality holds only when $|V_n| \leq c_m$.

For any $h \in C_c^\infty(\mathbb{R}^3)$, we can take sufficiently small m such that any vector \vec{X} that satisfies $\|X\| \geq c_m$ is valued as $h(\vec{X}) = 0$. Thus, it suffices to

consider the case $\|\vec{V}^0\| < c_m$. Under this condition, we have $(v_n - \vec{V}_n^0)_+ = v_n - \vec{V}_n^0$ on $\{v_n \geq c_m\}$. We always have $N_m(0) = N_m(\vec{V})$ and obtain $\frac{\bar{\lambda}_m}{N_m(\vec{V}^0)} = \lambda_m r^2$.

Lemma 2.6. *Let a, b a non-negative real number. For $p \geq 1$, we have $(a + b)^p \leq 2^{p-1}(a^p + b^p)$.*

Proof. We define $f(t) \triangleq \frac{(1+t)^p}{1+t^p}$ on $t \in [0, \infty)$. This function takes the maximum 2^{p-1} at $t = 1$. Taking $t = a/b$ leads to

$$2^{p-1} \geq \frac{(a+b)^p}{a^p + b^p}.$$

□

By Lemma 2.6, we have $(v_n - V_n)_+^4 \leq 8(v_n^4 + V_n^4) \leq 8(v_n^4 + c_m^4)$. Then,

$$\begin{aligned} J_4 &\leq \frac{\lambda r^2 M q_m^3}{6\sqrt{m}} \int d\Omega \int_{c_m}^\infty (v_n - V_n)_+^4 f_m^1(v_n) dv_n \\ &\leq \frac{8m^3 \lambda r^2 M}{6M^3 \sqrt{m}} \int d\Omega \int_{c_m}^\infty 8(v_n^4 + c_m^4) f_m^1(v_n) dv_n \\ &\leq \frac{32\lambda r/2}{3M^2} m^{5/2} \int d\Omega \int_{c_m m^{1/2}}^\infty (m^{-2} y^4 + V_n^4) f^1(y) dy \\ &= O(m^{1/2}) \end{aligned} \tag{2.16}$$

The symmetry of the distribution allows us to calculate as

$$\begin{aligned} J_3 &= \frac{\bar{\lambda}_m}{N_m(\vec{V}^0)} \frac{1}{2} \int d\Omega \int_{c_m}^\infty \frac{q_m^2}{2} [(v_n - \vec{V}^0 \cdot \vec{e}_n)^3 e_{n,k_1} e_{n,k_2} \partial^{k_1} \partial^{k_2} \\ &\quad + (v_n - \vec{V}^0 \cdot (-\vec{e}_n))^3 (-e_{n,k_1}) (-e_{n,k_2}) \partial^{k_1} \partial^{k_2}] h(\vec{V}^0) f_m^1(v_n) dv_n \\ &= \frac{\bar{\lambda}_m q_m^2}{2N_m(0)} \partial^{k_1} \partial^{k_2} h(\vec{V}^0) \int d\Omega \int_{c_m}^\infty (v_n^3 + 3v_n V_n^0) e_{n,k_1} e_{n,k_2} f_m^1(v_n) dv_n \\ &= \frac{2\lambda_m r^2 m^2}{M^2} \frac{4\pi}{3} \partial_k \partial^k h(\vec{V}^0) \int_{c_m}^\infty v_n^3 f_m^1(v_n) dv_n + O(m) \\ &= \frac{8\pi \lambda r^2}{3M^2} \Delta h(\vec{V}^0) \int_{c_m m^{1/2}}^\infty y^3 f^1(y) dy + O(m) \\ &\xrightarrow[m \rightarrow 0]{} \frac{8\pi \lambda r^2}{3M^2} \Phi_3 \Delta h(\vec{V}^0) \end{aligned} \tag{2.17}$$

where, we use $\int d\Omega e_{n,j} e_{n,k} = \frac{4\pi}{3} \delta_{j,k}$.

In the same way,

$$\begin{aligned}
J_2 &= \frac{\bar{\lambda}_m}{N_m(\vec{V}^0)} \frac{1}{2} \int d\Omega \int_{c_m}^{\infty} q_m [(v_n - \vec{V}^0 \cdot \vec{e}_n)^2 e_{n,k} \partial^k \\
&\quad + (v_n - \vec{V}^0 \cdot (-\vec{e}_n))^2 (-e_{n,k}) \partial^k] h(\vec{V}^0) f_m^1(v_n) dv_n \\
&= \frac{\bar{\lambda}_m q_m}{2N_m(0)} \partial^k h(\vec{V}^0) \int d\Omega \int_{c_m}^{\infty} (-2v_n V^{0,j} e_{n,j}) e_{n,k} f_m^1(v_n) dv_n \\
&= \frac{2\lambda_m r^2 m}{M} \frac{4\pi}{3} \partial_k h(\vec{V}^0) \int_{c_m}^{\infty} (-2v_n) V^{0,k} f_m^1(v_n) dv_n + O(m) \\
&= -\frac{16\pi\lambda r^2}{3M^2} \vec{V}^0 \cdot \vec{\nabla} h(\vec{V}^0) \int_{c_m m^{1/2}}^{\infty} y f^1(y) dy \\
&\xrightarrow[m \rightarrow 0]{} -\frac{16\pi\lambda r^2}{3M^2} \Phi_1 \vec{V}^0 \cdot \vec{\nabla} h(\vec{V}^0)
\end{aligned} \tag{2.18}$$

It is clear that

$$J_1 = \bar{\lambda}_m h(\vec{V}^0) \tag{2.19}$$

By using (2.16), (2.17), (2.18), (2.19) for (2.14), Lemma 2.3 completes the proof of Lemma 2.2.

2.3 The Closeness of Paths

In this section, we shall construct a Markov process $\vec{V}^{(m)}$ that has the same distribution as $\vec{V}^{(m)}$ and that is close to a mechanical process $\vec{V}^{(m)}$ in the following sense: for any $\epsilon > 0$,

$$\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(\{\bar{\omega} \in \bar{\Omega} \mid \sup_{t \in I} \|\vec{V}_t^{(m)}(\bar{\omega}) - \vec{V}_t^{(m)}(\bar{\omega})\| \geq \epsilon\}) = 0,$$

where $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}_m)$ is a probability space on which $\vec{V}^{(m)}$ and $\vec{V}^{(m)}$ are defined.

We denote by (Ma) the Markov process $\vec{V}^{(m)}$ and by (Me) the mechanical process $\vec{V}^{(m)}$. Note that (Ma) consists of a Poisson process with intensity $\bar{\lambda}_m$ and a jump distribution defined by $g_m(\vec{V}_t^{(m)}, \vec{v}_n) d\vec{v}_n d\Omega$. (Ma) is characterized by the collision rates

$$r(\vec{V}_t^{(m)}, \vec{v}_n) = \bar{\lambda}_m g_m(\vec{V}_t^{(m)}, \vec{v}_n). \tag{2.20}$$

In the same way, if $\sup_{0 \leq s \leq t} \|\vec{V}^{(m)}\| \leq c_m$, collisions with $\|\vec{v}_n\| > c_m$ is characterized by

$$r(\vec{V}_t^{(m)}, \vec{v}_n) = \bar{\lambda}_m g_m(\vec{V}_t^{(m)}, \vec{v}_n) \tag{2.21}$$

Slow atom collisions, however, make difference between (2.20) and (2.21). Then, they result to $r(\vec{V}_t^{(m)}, \vec{v}_n) \neq r(\vec{\tilde{V}}_t^{(m)}, \vec{v}_n)$. To offset this difference, we introduce the following modifications:

(a) The rate of extra collisions

$$R(\vec{V}_t^{(m)}, \vec{V}_t^{(m)}, \vec{v}_n) = \max(r(\vec{V}_t^{(m)}, \vec{v}_n) - r(\vec{\tilde{V}}_t^{(m)}, \vec{v}_n))_+, \quad (2.22)$$

(b) The acceptance rate

$$p_m(\vec{V}_t^{(m)}, \vec{\tilde{V}}_t^{(m)}, \vec{v}_n) = \min\left(\frac{r(\vec{V}_t^{(m)}, \vec{v}_n)}{r(\vec{\tilde{V}}_t^{(m)}, \vec{v}_n)}, 1\right). \quad (2.23)$$

Let $(\Omega, \mathcal{F}, \mathbb{P}_m)$ be a probability space, where the stochastic process $\vec{V}^{(m)}$ is defined. To introduce (2.22) and (2.23), we extend a probability space from $(\Omega, \mathcal{F}, \mathbb{P}_m)$ to $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}_m)$, where $\bar{\Omega} = \Omega \times H$, $\bar{\mathcal{F}} = \mathcal{F} \otimes \mathcal{B}(H)$, and $\bar{\mathbb{P}}_m(\cdot \times H) = \mathbb{P}_m(\cdot)$. We write as $\bar{\mathbb{P}}_m(\{\bar{\omega} \in \bar{\Omega} | \vec{V}^{(m)}(\bar{\omega}) \in F\}) = \bar{P}_m(F)$ for all $F \in \mathcal{B}(D(I))$, where \bar{P}_m is an induced probability measure.

Then, we shall show the following lemmas.

Lemma 2.7. *For any I and for all $\epsilon > 0$,*

$$\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(\{\bar{\omega} \in \bar{\Omega} | \sup_{t \in I} \|\vec{V}^{(m)}(\bar{\omega}) - \vec{\tilde{V}}^{(m)}(\bar{\omega})\| \geq \epsilon\}) = 0. \quad (2.24)$$

Lemma 2.8. *if for $t_0 \geq 0$, we have*

$$\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(\{\bar{\omega} \in \bar{\Omega} | \sup_{0 \leq t \leq t_0} \|\vec{V}^{(m)}(\bar{\omega}) - \vec{\tilde{V}}^{(m)}(\bar{\omega})\| \geq \epsilon\}) = 0, \quad (2.25)$$

for all $\epsilon > 0$, then,

$$\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(\{\bar{\omega} \in \bar{\Omega} | \sup_{0 \leq t \leq t_0+z} \|\vec{V}^{(m)}(\bar{\omega}) - \vec{\tilde{V}}^{(m)}(\bar{\omega})\| \geq \epsilon\}) = 0, \quad (2.26)$$

for all $\epsilon > 0$, where $z = \frac{M}{128\lambda r^2 \pi \Phi_1}$.

To prove Lemma 2.7, it suffices to prove Lemma 2.8.

Fix $\epsilon > 0$, and we introduce $t_0 \geq 0$ and stopping time:

$$t_m^* = \inf_{t \geq t_0} \{t; \|\vec{V}_t^{(m)} - \vec{\tilde{V}}_t^{(m)}\| > \epsilon\}.$$

Since $\vec{V}^{(m)}$ and $\vec{\tilde{V}}^{(m)}$ are right continuous, we have

$$\|\vec{V}_{t_m^*}^{(m)} - \vec{\tilde{V}}_{t_m^*}^{(m)}\| > \epsilon.$$

Observe that

$$\begin{aligned} \left\{ \sup_{0 \leq t \leq t_0+z} \|\vec{V}_t^{(m)} - \vec{\tilde{V}}_t^{(m)}\| > \epsilon \right\} &= \{t_m^* \leq t_0 + z\} \\ &= \{t_m^* \leq t_0 + z; \|\vec{V}_{t_m^*}^{(m)} - \vec{\tilde{V}}_{t_m^*}^{(m)}\| > \epsilon\} \\ \|\vec{V}_t^{(m)} - \vec{\tilde{V}}_t^{(m)}\| &\leq \epsilon \text{ for } t < t_m^*. \end{aligned} \quad (2.27)$$

By Lemma 2.2, we have

$$\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(G_p^m) = 1,$$

where

$$G_p^m = \{\bar{\omega} \in \bar{\Omega} \mid \sup_{0 \leq t \leq t_0+z} |\vec{V}_t^{(m)}(\bar{\omega})| < pc_m\}, p > 0.$$

Therefore, to establish (2.26), it suffices to show that

$$\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(\{t_m^* \leq \sup_{0 \leq t \leq t_0+z} \|\vec{V}_t^{(m)}(t_m^*) - \vec{\tilde{V}}_t^{(m)}(t_m^*)\| \geq \epsilon\} \cap G_p^m) = 0. \quad (2.28)$$

If we pick $p = 1/8$ and $\epsilon \leq c_m/8$, then on G_p^m

$$\sup_{0 \leq t \leq t_0+z} |\vec{V}_t^{(m)}| < c_m/4.$$

By (2.5) with sufficiently small m , $|\vec{v}_n| \geq c_m$ leads to

$$|\vec{v}_n| \geq \frac{M-m}{M+m} c_m - \frac{2M}{M+m} |V_n| > \frac{1}{4} c_m.$$

Therefore, atoms with $\|\vec{v}_n\| \geq c_m$ does not recollides with the molecule.

Let $\mathcal{E}(t)$ denote the change of $\vec{V}^{(m)}$ during $[0, t]$ produced by collisions with slow atoms. We denote by $E(\dot{E})$ indices of extra collisions for (Me)((Ma)) and by B indices of common collisions. Let $\vec{W}_m^i(\vec{\tilde{W}}_m^i)$ denote the change in the velocity of (Me)((Ma)) due to index i collision with fast particles, and we write down as

$$\begin{aligned} \|\vec{V}_{t_m^*}^{(m)} - \vec{\tilde{V}}_{t_m^*}^{(m)}\| &= \left\| \sum_{i \in B} (\vec{W}_m^i - \vec{\tilde{W}}_m^i) + \sum_{i \in E} \vec{W}_m^i - \sum_{i \in \dot{E}} \vec{W}_m^i + \mathcal{E}(t_m^*) - \mathcal{E}(t_0) + \vec{V}_{t_0}^{(m)} - \vec{\tilde{V}}_{t_0}^{(m)} \right\| \\ &\leq W_m^{(1)}(t_m^*) + W_m^{(2)}(t_m^*) + W_m^{(3)}(t_m^*) + \|\vec{V}_{t_0}^{(m)} - \vec{\tilde{V}}_{t_0}^{(m)}\|, \end{aligned}$$

where

$$\begin{aligned} W_m^{(1)}(t_m^*) &= \|\mathcal{E}(t_m^*) - \mathcal{E}(t_0)\|, \\ W_m^{(2)}(t_m^*) &= \sum_{i \in E} \|\vec{W}_m^i\| + \sum_{i \in \tilde{E}} \|\vec{\bar{W}}_m^i\|, \\ W_m^{(3)}(t_m^*) &= \sum_{i \in B} \|\vec{W}_m^i - \vec{\bar{W}}_m^i\|. \end{aligned}$$

Then

$$\{\|\vec{V}_{t_m^*}^{(m)} - \vec{\bar{V}}_{t_m^*}^{(m)}\| > \epsilon\} \subset \bigcup_{i=1}^3 \{W_m^{(i)}(t_m^*) > \epsilon/4\} \cup \{\|\vec{V}_{t_0}^{(m)} - \vec{\bar{V}}_{t_0}^{(m)}\| > \epsilon/4\}$$

Note that $\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(\{\|\vec{V}_{t_0}^{(m)} - \vec{\bar{V}}_{t_0}^{(m)}\| > \epsilon/4\}) = 0$ from hypothesis. By (2.25), (2.28) follows if

$$\lim_{m \rightarrow 0} \bar{\mathbb{P}}_m(G_p^m \cap \{t_m^* \leq t_0 + z\} \cap \{W_m^{(i)}(t_m^*) > \epsilon/4\}) = 0 \text{ for } i = 1, 2, 3. \quad (2.29)$$

We now establish (2.29) for $i = 1, 2, 3$ with $p = 1/8$ and $\epsilon < c_m/8$. We assume that m is sufficiently small to satisfy inequalities mentioned above.

Consider the term of $i = 1$. Since the variation of the molecule's momentum caused by slow atoms should be equivalent to the variation of atoms' momentum. Thus, we are able to estimate $W_m^{(1)}(t_m^*)$ through an evaluation for the momentum of slow atoms.

Now, we will show that

$$\|\Delta p\| = m\|\vec{v}_n(t_m^*) - \vec{v}_n(t_0)\| \leq 7mc_m \quad (2.30)$$

Let $\vec{v}^{(i)}(\vec{v}^{(i)}(+))$ denote the slow atom's velocity of i th pre-(post-)collision . We divide the proof into two cases: $\|\vec{v}_t^{(1)}\| \geq c_m$ and $\|\vec{v}_t^{(1)}\| < c_m$.

(1) $\|\vec{v}_t^{(1)}\| \geq c_m$

We define Cartesian coordinates such that $v_z^{(1)} = -v_n^{(1)}$ and $v_x^{(1)} = v_t^{(1)}$. Let $(\theta^{(i)}, \phi^{(i)})$ denote spherical coordinates on the sphere of the molecule for i th collision. Suppose that we have k times collisions with an atom. Then

$$v_n^{(i+1)} = -v_t^{(i)}(+) \cos(\phi^{(i+1)} - \phi^{(i)}) \sin \theta^{(i+1)} - v_z^{(i+1)} \cos \theta^{(i+1)} \text{ for } 1 \leq i \leq N-1.$$

Since v_z is nondecreasing on $0 \leq \theta^{(i)} \leq \pi/2$, $v_z^{(i)} \geq v_z^{(1)}(+) = v_z^{(2)} \geq -c_m/4$, where we have used (2.5). Therefore, we have $v_z^{(i)} - V_{t,z}^{(m)} \geq -c_m/2$, which

concludes to $\frac{1}{2} \leq \cos\theta^{(i)}$. Then,

$$\begin{aligned}
\|\vec{v}(t_m^*) - \vec{v}(t_0)\| &\leq \sum_{i=1}^k |v_n^{(i)}(+)-v_n^{(i)}| \\
&\leq \sum_{i=1}^{k-1} \frac{v_z^{(i)}(+)-v_z^{(i)}}{\cos\theta^{(i)}} + v_n^{(k)}(+)-v_n^{(k)} \\
&\leq 2(v_z^{(k-1)}(+)-v_z^{(1)}) + c_m + c_m/4 \\
&\leq 2(c_m/4+c_m) + c_m + c_m/4 \\
&\leq 4c_m,
\end{aligned} \tag{2.31}$$

where we use the collision equation (2.7).

(2) $\|\vec{v}_t\| < c_m$
If $\|\vec{v}_t(t_1)\| \geq c_m$ for $t_1 \in [t_0, t_m^*]$, the above result 2.31 leads to

$$\|\vec{v}(t_m^*) - \vec{v}(t_0)\| \leq \|\vec{v}(t_m^*) - \vec{v}(t_1)\| + \|\vec{v}(t_1) - \vec{v}(t_0)\| \leq 4c_m + 2\sqrt{2}c_m. \tag{2.32}$$

If always $\|\vec{v}_t(t)\| \leq c_m$ on $[t_0, t_m^*]$, it is clear that we

$$\|\vec{v}(t_0) - \vec{v}(t_m^*)\| \leq 2\sqrt{2}c_m \tag{2.33}$$

(2.31), (2.32), and (2.33) concludes to (2.30).

Next, we consider the collision rate. Considering $(v_n - V_n)_+ \leq 2c_m$ and (2.9), we have

$$E(N_m) \leq 8\pi\lambda r^2 T c_m m^{-1/2}$$

By using (2.30), we evaluate as

$$\begin{aligned}
\bar{\mathbb{P}}_m(G_p^m \cap \{t_m^* \leq t_0 + z\} \cap \{W_m^{(i)}(t_m^*) > \epsilon/4\}) &\leq \bar{\mathbb{P}}_m(\{N_m |\Delta p| \geq \epsilon/4\}) \\
&\leq 28c_m E(N_m)/\epsilon \\
&\leq 224\pi T \lambda r^2 c_m^2 m^{-1/2} \\
&= O(m^{1/10})
\end{aligned}$$

We will establish (2.29) for $i = 3$. We describe an extra collision rate as

$$\bar{R}_m(\vec{V}_t^{(m)}. \vec{V}_t^{(m)}, \vec{v}_n) = |r(\vec{V}_t^{(m)}, \vec{v}_n) - r(\vec{V}_t^{(m)}, \vec{v}_n)|$$

(2.27) allows us to majorizes the rate $\bar{R}_m(\vec{V}_t^{(m)}. \vec{V}_t^{(m)}, \vec{v}_n)$ by

$$R_m^p(\epsilon, \vec{v}_n) = \lambda r^2 m^{-1/2} \epsilon f_m^1(v_n).$$

(2.4) and $\|\vec{V}^{(m)}\| \leq c_m$ give

$$\|\Delta \vec{V}^{(m)}\| \leq \frac{2m}{M+m}(|v_n| + c_m) \leq \frac{4m}{M+m}|v_n|$$

Then, we obtain

$$\bar{\mathbb{P}}_m(G_p^m \cap \{t_m^* \leq t_0 + z\} \cap \{W_m^{(i)}(t_m^*) > \epsilon/4\}) \leq \bar{\mathbb{P}}_m\left(\sum_{i=1}^N \frac{4m}{M+m} v_n^{(i)} \geq 4/\epsilon\right). \quad (2.34)$$

To evaluate an upper bound of the right hand side of (2.34), we calculate as

$$\begin{aligned} E_m\left(\sum_{i=1}^N \frac{4m}{M+m} v_n^{(i)}\right) &\leq \frac{4m}{M+m} 4\pi\lambda r^2 \epsilon m^{-1} z \int_{c_m m^{-1/2}}^{\infty} y f^1(y) dy \\ &\leq \frac{16\pi\lambda r^2 \epsilon}{M} \Phi_1 z \leq \epsilon/8. \end{aligned}$$

We denote by

$$J \triangleq \sum_{i=1}^N v_n^{(i)}.$$

Then right hand side of (2.34) can be estimated as follows: With

$$\begin{aligned} \bar{\mathbb{P}}_m\left(\frac{4m}{M+m} J \geq 4/\epsilon\right) &\leq \bar{\mathbb{P}}_m(J \geq 2E_m(J)) \\ &\leq \bar{\mathbb{P}}_m(J - E_m(J) \geq E_m(J)) \quad (2.35) \\ &\leq \frac{E_m[J - E_m(J)]^2}{E_m(J)^2} \end{aligned}$$

by Chebyshev's inequality.

We define

$$\frac{\int_{c_m}^{\infty} (v_n)^k f_m^1(v_n) dv_n}{\int_{c_m}^{\infty} f_m^1(v_n) dv_n} = \langle v_n^k \rangle$$

Based on the properties of the Poisson process, we calculate as

$$\begin{aligned} E_m(J) &= \langle v_n \rangle E_m(N) \\ E_m(J^2) &= E_m(N) \langle v_n^2 \rangle + E_m(N(N-1)) \langle v_n \rangle^2 = E_m(N) \langle v_n^2 \rangle + E_m(N)^2 \langle v_n \rangle^2 \end{aligned}$$

The right hand side of (2.35) is written as

$$E_m(N)^{-1} \langle v_n^2 \rangle / \langle v_n \rangle^2 = O(m^{1/2}),$$

where we use $E_m(N) = O(m^{-1/2})$.

We are left with (2.29) for $i = 3$. In the same way,

$$\begin{aligned}\|\vec{W}_i - \vec{W}_i\| &= \|(-\frac{2m}{M+m}\vec{V}_s^{(m)} + \frac{2m}{M+m}\vec{v}_n) - (-\frac{2m}{M+m}\vec{V}_s^{(m)} + \frac{2m}{M+m}\vec{v}_n)\| \\ &\leq \frac{2m}{M+m}\epsilon.\end{aligned}$$

Hence,

$$W_m^{(3)} \leq \frac{2m}{M+m}\epsilon N_m(z),$$

where $N_m(z)$ is the number of collisions within $[t_0, t_0 + z]$. Since collision occurrence is determined by a Poisson process with an intensity $\bar{\lambda}_m$, we obtain

$$\begin{aligned}E[N_m(z)] &= \bar{\lambda}_m z \leq 4\pi\lambda r^2 m^{-1} \Phi_1 z \\ &= \frac{\lambda r^2 M}{32m},\end{aligned}$$

Then, we obtain

$$\begin{aligned}&\bar{\mathbb{P}}_m(G_p^m \cap \{t_m^* \leq t_0 + z\} \cap \{W_m^{(3)}(t_m^*) > \epsilon/4\}) \\ &\leq \bar{\mathbb{P}}_m(\{\frac{2m}{M+m}\epsilon N_m(z) \geq \epsilon/4\}) \\ &\leq \bar{\mathbb{P}}_m(\{N_m(z) \geq \frac{M+m}{8m}\}) \\ &\leq \bar{\mathbb{P}}_m(\{N_m(z) \geq 2E(N_m(z))\}) \\ &\leq \frac{E[N_m(z) - E(N_m(z))]^2}{E(N_m(z))^2} \\ &\leq (EN_m(z))^{-1} \\ &\xrightarrow[m \rightarrow 0]{} 0\end{aligned}$$

We complete the proof of the Lemma 2.8. Then, the Theorem 2.1 follows from the Lemma 2.2 and the Lemma 2.7 [3].

3 Langevin Equation and Brownian Motion

The section 2 showed that nonequilibrium system was described by the stochastic differential equation (2.1) under certain assumptions. Since we can consider each component of the derived equation independently, we have the same equation for two dimension. We now consider properties of the strong solution of two-dimensional version of the Langevin equation and its relation to a standard Brownian motion.

3.1 The Strong Solution of the Langevin Equation

Let us start with Borel-measurable functions $b_i(t, x), \sigma_{ij}(t, x); 1 \leq i \leq d, 1 \leq j \leq r$, from $[0, \infty) \times \mathbb{R}^d$ into \mathbb{R} , and we define a $(d \times 1)$ drift vector $b(t, x) = \{b_i(t, x)\}_{i \leq i \leq d}$ and a $(d \times r)$ dispersion matrix $\sigma(t, x) = \{\sigma_{ij}\}_{i \leq i \leq d, 1 \leq j \leq r}$. Then, we consider the following stochastic differential equation:

$$d\vec{X}_t = b(t, \vec{X}_t)dt + \sigma(t, \vec{X}_t) \cdot d\vec{W}_t, \quad (3.1)$$

where $\vec{W} = \{\vec{W}_t; 0 \leq t < \infty\}$ is an r -dimensional Brownian motion.

For existence and uniqueness of strong solution for (3.1), the following lemma from [11] is known.

Lemma 3.1. *Suppose that coefficients $b(t, \vec{x})$ and $\sigma(t, \vec{x})$ satisfy the global Lipschitz and linear growth conditions:*

$$\|b(t, \vec{x}) - b(t, \vec{y})\| + \|\sigma(t, \vec{x}) - \sigma(t, \vec{y})\| \leq K\|\vec{x} - \vec{y}\|,$$

$$\|b(t, \vec{x})\|^2 + \|\sigma(t, \vec{x})\|^2 \leq K^2(1 + \|\vec{x}\|^2),$$

for every $0 \leq t < \infty, \vec{x} \in \mathbb{R}^d, \vec{y} \in \mathbb{R}^d$, where K is a positive constant. On some probability space $(\Omega, \mathcal{F}, \mathbb{P}_m)$, let $\vec{\xi}$ be an \mathbb{R}^d -valued random vector, independent of the r -dimensional Brownian motion $\vec{W} = \{\vec{W}_t, \mathcal{F}^W; 0 \leq t < \infty\}$, and with finite second moment:

$$E\|\vec{\xi}\|^2 < \infty.$$

Let $\{\mathcal{F}_t; 0 \leq t < \infty\}$ be a filtration augmented by $\{\mathcal{F}_t^W; 0 \leq t < \infty\}$ and $\sigma(\vec{\xi})$. Then there exists a continuous, adapted process $\vec{X} = \{\vec{X}_t, \mathcal{F}_t; 0 \leq t < \infty\}$, which is the unique strong solution of (3.1) relative to \vec{W} , with initial condition $\vec{\xi}$.

Now, we are back to consider the Langevin Equation,

$$Md\vec{V}_t = -\gamma dt + \sqrt{2\gamma k_B T}d\vec{W}_t, \quad (3.2)$$

$$d\vec{X}_t = \vec{V}_t dt. \quad (3.3)$$

We combine (3.2) and (3.3) into the matrix formulation as

$$d \begin{bmatrix} X_t^1 \\ X_t^2 \\ V_t^1 \\ V_t^2 \end{bmatrix} = \begin{bmatrix} V_t^1 \\ V_t^2 \\ -\frac{\gamma}{M} \\ -\frac{\gamma}{M} \end{bmatrix} dt + \sqrt{2\gamma k_B T} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} d\vec{W}_t \quad (3.4)$$

Since we assume that M and γ are constants, the Lemma 3.1 makes sure the existence and uniqueness of strong solution for (3.4).

Let X^i be i-th component of \vec{X} . By Itô formula, we derive

$$V_t^i = V_0^i e^{-\frac{\gamma}{M}t} + \frac{\sqrt{2\gamma k_B T}}{M} \int_0^t e^{-\frac{\gamma}{M}(t-s)} dW_s^i.$$

Then, we obtain

$$\begin{aligned} E[V_t^i] &= V_0^i e^{-\frac{\gamma}{M}t} \\ E[(V_t^i - E[V_t^i])^2] &= \frac{2k_B T}{M} [1 - e^{-\frac{2\gamma}{M}t}] \\ &\rightarrow \frac{2k_B T}{M} \text{ as } t \rightarrow \infty \end{aligned}$$

When taking $V_0^i = 0$, we have $\frac{1}{2}ME(V_0^i)^2 = k_B T$, which coincides with the statistics when the molecule stays in the equilibrium state.

We also have

$$X_t^i = X_0^i + \frac{M}{\gamma} V_0 [1 - e^{-\frac{\gamma}{M}t}] + \sqrt{\frac{2k_B T}{\gamma}} \int_0^t dW_s^i [1 - e^{-\frac{\gamma}{M}(t-s)}] \quad (3.5)$$

, where we use stochastic Fubini theorem (appendix B). Then, we obtain

$$\begin{aligned} E[X_t] &= X_0 + \frac{M}{\gamma} V_0 [1 - e^{-\frac{\gamma}{M}t}] \\ E[(X_t^i - E[X_t^i])^2] &= \frac{2k_B T}{\gamma} [t - \frac{2M}{\gamma} (1 - e^{-\frac{\gamma}{M}t}) + \frac{M}{2\gamma} (1 - e^{-\frac{2\gamma}{M}t})] \end{aligned}$$

If we have $\vec{V}_0 = \vec{X}_0 = 0$, then

$$E[\|\vec{X}_t\|^2] \sim \begin{cases} \frac{\gamma k_B T}{M^2} t^3, & t \ll 1, \\ \frac{4k_B T}{\gamma} t, & t \gg 1. \end{cases}$$

3.2 Approximation by Brownian Motion

Some may prefer treating nonequilibrium system with a standard Brownian motion due to its simplicity. Intuitively, we can derive this approximation with the assumption that the small mass of the molecule allows us to neglect the inertial term. That is, (3.2) changes into

$$0 = -\gamma \vec{V}_t dt + \sqrt{2\gamma k_B T} d\vec{W}_t.$$

Then, we obtain

$$\vec{X}_t = \sqrt{\frac{2k_B T}{\gamma}} \vec{W}_t, \quad (3.6)$$

where we assume $\vec{V}_0 = \vec{X}_0 = 0$ for the sake of simplicity.

Now, our question is in what situation the mass of the molecule can be regarded as small enough to approximate (3.5) with (3.6). Since both (3.5) and (3.6) belong to continuous square integrable martingale, Doob's maximal inequality [5] gives

$$\begin{aligned} E[\sup_{0 < s < t} \|\vec{X}_s - \vec{\tilde{X}}_s\|^2] &\leq 4 \cdot 3 \cdot \frac{2k_B T}{\gamma} \int_0^t e^{-\frac{2\gamma}{M}(t-u)} du \\ &= \frac{12k_B T}{\gamma} \frac{M}{\gamma} [1 - e^{-\frac{2\gamma}{M}t}] \\ &\leq \frac{12k_B T}{\gamma} \frac{M}{\gamma} \end{aligned}$$

By using Fatou's lemma, we obtain

$$E[\sup_{0 < s < \infty} \|\vec{X}_s - \vec{\tilde{X}}_s\|^2] \leq \frac{12k_B T}{\gamma} \frac{M}{\gamma} \quad (3.7)$$

Hence, under the condition, $M \ll \gamma$, the stochastic process governed by the Langevin equation is approximated by a Brownian Motion.

Consider the following transformation:

$$\vec{X}_{c,t} = \frac{1}{\sqrt{c}} \vec{X}_{ct},$$

where $c > 0$. (3.5) gives

$$\begin{aligned} \hat{X}_{c,t}^i &= \frac{1}{\sqrt{c}} \sqrt{\frac{2k_B T}{\gamma}} \int_0^{ct} dW_s^i [1 - e^{-\frac{\gamma}{M}(ct-s)}] \\ &= \sqrt{\frac{2k_B T}{\gamma}} \int_0^t dW_s^i [1 - e^{-\frac{\gamma}{M}c(t-s)}] \\ &= \sqrt{\frac{2k_B T}{\gamma}} \int_0^t dW_s^i [1 - e^{-\frac{\gamma}{M}(t-s)}], \end{aligned} \quad (3.8)$$

where we denote $\tilde{M} = M/c$. (3.8) implies that \hat{X}_c corresponds to the process driven by the Langevin equation with the molecule of mass M/c . By taking

the limit $c \rightarrow \infty$, (3.7) makes the process to the standard Brownian motion multiplied by the constant. Multiplication of $\frac{1}{\sqrt{c}}$ and transform from t to ct correspond to the change of typical length scale and time interval of observations respectively. Let's take Δt as time interval and $l\sqrt{\Delta t}$ as the typical length scale. The observed process is defined as $\frac{1}{l\sqrt{\Delta t}}\{\vec{X}_{j\Delta t}\}_{j=1}^N$. Then, We obtain

$$\begin{aligned} \frac{1}{l^2\Delta t}E[\max_{1 \leq j \leq N}\|\vec{X}_{j\Delta t} - \tilde{\vec{X}}_{j\Delta t}\|^2] &\leq E[\sup_{0 < s < t}\|\vec{X}_s - \tilde{\vec{X}}_s\|^2] \\ &\leq \frac{12k_B T}{\gamma l^2} \frac{M}{\gamma \Delta t} \end{aligned} \quad (3.9)$$

Therefore, large typical length scale and long time interval observation environment makes the process look the standarnd Brownian motion multiplied by the constant.

4 Fence Model and Subdiffusion

As we discussed in the section 3, the variance of free particles with $\vec{V}_0 = 0$ is proportional to t when $t \gg 1$. This is, however, not the case when there are obstacles and inhomogeneous structure in medium. On the plasma cell membranes, immobilized membrane proteins constitute the fence structure, which slow diffusion. This slowed diffusion does not have a constant diffusion coefficient and is classified as subdiffusion. In this section, we introduce the fence model to treat such diffusion mathematically and clarity how the fence structure causes subdiffusion.

4.1 Fence Model

Each cell membrane is mainly composed of bilayer of phospholipids, membrane proteins and cholesterol. The classic membrane model introduced by Singer and Nicholson[23] suggested that the bilayer is in a fluid state in almost all membranes and that the bilayer acts as a two-dimensional solvent for membrane proteins. Though this model was succeeded to describe a short time scale dynamics, it failed in predicting the behavior in long time scale, which was slower than predicted. This comes from the fact that diffusion is hindered by proteins immobilized by cytoskeletons, which are filament substances helping cells maintain their shape and internal organization. [10] explains that this structure arranges immobilized proteins on straight lines and produces compartments.

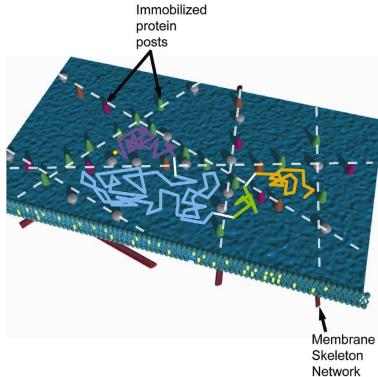


Figure 1: This is an example figure of fence model from [10]. The color of the trajectory of a molecule changes depending on which compartments the molecule belong to

4.2 Discretized Brownian Motion with Fences

As we discussed in the section ??, (3.9) implies that coarser and larger time interval observations lead to the Brownian motion approximation. Since we now have interest in an asymptotic behavior, we use the standard Brownian motion that has the same variance as free particles for $t \gg 1$; i.e., $\frac{4k_B T}{\gamma} t$. For the sake of the simplicity, we assume $\frac{2k_B T}{\gamma} = 1$.

The Brownian motion are known to fluctuate locally infinitely many times, which makes it difficult to treat with the fence structure. Instead, we use a discrete time model. Let $\tau > 0$ and let $\{\vec{\xi}_i\}_{i \in \mathbb{N}}$ be i.i.d. two dimensional normal distributed random variables with mean 0 and two dimensional identity matrix covariance. Let us denote summations of steps by $\vec{S}_0 = 0$ and $\vec{S}_k = \sum_{i=1}^k \vec{\xi}_i$, $k \geq 1$. We define a discrete time version of the Brownian motion as

$$\vec{X}_t = \sqrt{\tau} \vec{S}_{[t/\tau]} + (t - \tau[t/\tau]) \vec{\xi}_{\lceil t/\tau \rceil},$$

where $[\cdot]$ and $\lceil \cdot \rceil$ are the Gauss symbol and the ceiling function respectively. We assume that describe the behavior of a molecule by this process.

To simplify discussion, we here set all of the compartments squares with length 1 edges. There are ten equivalent size walls aligned on each edge with equivalent intervals between walls. The size and interval of walls are determined by a parameter $0 < p < 1$ – the occupation of walls on each edge – and are set as $\frac{1-p}{10}$ and $\frac{p}{10}$ respectively.

These distributed walls make an effect especially on an asymptotic behavior of the molecule. When $t \ll 1$, the molecule hardly reach any walls, and the diffusion is close to the normal diffusion. On the other hand, when $t \gg 1$, walls confine the molecule in each compartment for long time and slow diffusion.

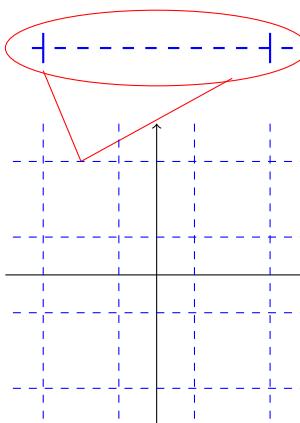


Figure 2: This figure is an example when $p = 0.5$. Walls are distributed uniformly.

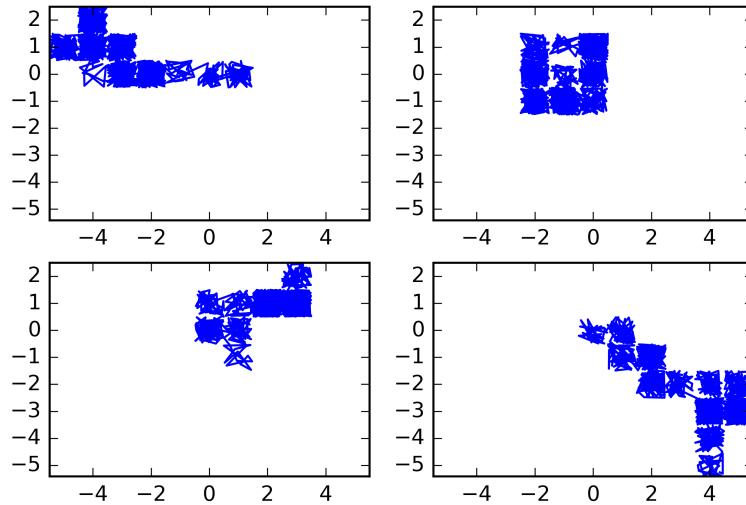


Figure 3: This is the figure of four sample paths of $p = 10^{-2}$. To make it easy to see, we plot every 100τ . Trajectories show the staying in a compartment with infrequent transition.

4.3 Hopping Model

After long enough time has passed, which compartment the molecule belongs to is more dominant factor to determine the variance than the precise position inside a compartment. Hence, the process hopping among the centers of compartments is able to describe the asymptotic behavior. Besides that, if p is small, the molecule stays in each compartment for long time with infrequent transition to neighbor compartments, which reduce the dependency of an entrance point of the compartment. Thus, each hopping can be regarded as independent jumps. We define this process as the hopping model.

Mathematically, this model is one of the Continuous Time Random Walk (CTRW). CTRW consists of two sets of random variables: sojourn times and jumps. The sojourn times between jumps are determined by nonnegative i.i.d.(independent and identically distributed) random variables J_1, J_2, \dots . We set $T(0) = 0$ and $T(n) = \sum_{j=1}^n J_j$, the time of the n -th jump. The jumps are given by i.i.d. random vectors $\vec{Y}_1, \vec{Y}_2, \dots$ with $E\vec{Y}_i = 0$, which are assumed to be independent of $\{J_j\}_{j \in \mathbb{N}}$. Let $\vec{S}_0 = 0$ and $\vec{S}_n = \sum_{i=1}^n \vec{Y}_i$, the position of the particle after the n -th jump. For $t \geq 0$, let us denote the number of jumps up to time t by

$$N_t = \max\{n \leq 0 : T(n) \leq t\}.$$

We define the stochastic process $\{\vec{X}_t; t \geq 0\}$ by

$$\vec{X}_t = \vec{S}_{N_t} = \sum_{i=1}^{N_t} \vec{Y}_i.$$

According to [16], if sojourn time distribution belongs to the certain class with infinite mean, the scaling limit of CTRW converges to process with the variance proportional to t^α in the Skorokhod topology. If the sojourn time distribution has a finite mean, we have

$$E\|\vec{X}_t\|^2 = \frac{E\|\vec{Y}_1\|^2}{EJ_1}t.$$

Hence, if the sorjourn time has a finite mean, the hopping model is seen as the normal diffusion rather than subdiffusion.

4.4 Transient Subdiffusion

In some physical phenomena, subdiffusion is observed in only certain time range and the normal diffusion is observed for $t \gg 1$ and $t \ll 1$. This

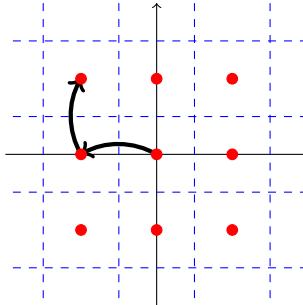


Figure 4: This is the figure of hopping model. The molecule jumps among the center of compartments – red points.

phenomena is called transient subdiffusion. More formally, we define as the subdiffusion that satisfies the following: $\exists D_1, \exists D_2 \in \mathbb{R}$ such that $\lim_{t \rightarrow +0} D(t) = D_1$, $\lim_{t \rightarrow \infty} D(t) = D_2$, and $D_1 > D_2$. Transition between two different diffusion coefficients causes transient anomalous behavior.

Since we assume that each jump is independent of each other if p is small, it is natural to consider the hopping model with a symmetry jump distribution; i.e., $P(\vec{Y}_1 = (1, 0)) = P(\vec{Y}_1 = (0, 1)) = P(\vec{Y}_1 = (-1, 0)) = P(\vec{Y}_1 = (0, -1)) = 1/4$.

For each p , a sojourn time is characterized by

$$\begin{aligned}\lambda_p(t) &\stackrel{\Delta}{=} \lim_{\Delta t \rightarrow 0} \frac{P(J_i \leq t + \Delta t | J_i > t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(J_i \leq t + \Delta t)}{\Delta t P(J_i > t)} \\ &= -\frac{d}{dt} P(J_i \leq t).\end{aligned}$$

$\lambda_p(t)$ is a quantity to indicate the intensity of jump occurrence at t . If p is so small that the position distribution spreads inside a compartment, the rate of jump does not depend on time; i.e. $\lambda_p(t) = \text{constant}$. Besides that, this constant intuitively has to be proportional to p . Thus, we have $\lambda_p(t) = p\lambda$, where λ is a parameter determined by observations, and the sojourn time has an exponential distribution with intensity $p\lambda$. Hence, the hopping process corresponds to a compound Poisson process with the intensity of jump distribution $p\lambda$, which has $D(t) = p\lambda$. In the diffusion of transient subdiffusion, $D_1 = 2$ and $D_2 = p\lambda$.

The Figure 5 is a numerical result of diffusion coefficients at $t = 10^4$ for different parameters p . Points from $p = 10^{-5}$ to 10^{-2} are fit on a straight line $p\lambda$, where λ is determined by $p = 10^{-3}$. On the other hand, points of large and small p are away from the line. When specifying the distribution of jumps and sojourn times, we assume that the molecule stays at one compartment for long time. That is, for large p , this assumption is not satisfied. For small

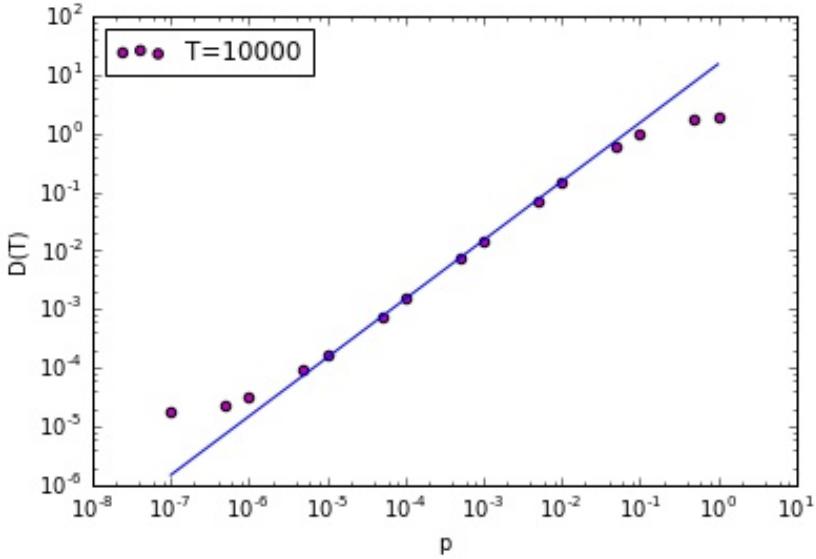


Figure 5: We estimate λ with data of $p = 10^{-3}, t = 10^4$. Blue line is $p\lambda$

p , since transition is too infrequent, and the process does not still converge to slower normal diffusion.

Considering up to how large the approximation works is difficult because we have to consider a process depending on long range history. We just focus on estimating a valid range of the approximation for small p : the left range in the Figure 5. Let \vec{X} and $\tilde{\vec{X}}$ be the original process and the hopping model. We assume that \vec{X} and $\tilde{\vec{X}}$ stay at the same compartment; i.e., both processes move to neighbor compartments at the same time.

$$\|\tilde{\vec{X}}_t\| - \frac{1}{\sqrt{2}} \leq \|\vec{X}_t\| \leq \|\tilde{\vec{X}}_t\| + \frac{1}{\sqrt{2}}$$

Then,

$$\begin{aligned} E\|\tilde{\vec{X}}_t\|^2 - \frac{1}{2} &\leq E\|\vec{X}_t\|^2 \leq E\|\tilde{\vec{X}}_t\|^2 + \frac{1}{2} \\ \tilde{D}(t) - \frac{1}{2t} &\leq D(t) \leq \tilde{D}(t) + \frac{1}{2t} \end{aligned}$$

We obtain

$$\begin{aligned}
\frac{|D(t) - \tilde{D}(t)|}{D(t)} &\leq \frac{|D(t) - \tilde{D}(t)|}{\tilde{D}(t)} \left(\frac{D(t)}{\tilde{D}(t)} \right)^{-1} \\
&\leq \frac{1}{2\tilde{D}(t)t} \left(\frac{D(t)}{\tilde{D}(t)} \right)^{-1} \\
&\leq \frac{1}{2p\lambda t} \left(1 - \frac{1}{2p\lambda t} \right)^{-1} \\
&\leq \frac{1}{2p\lambda t} + \frac{\theta}{4p^2\lambda^2 t^2},
\end{aligned}$$

where $0 < \theta < 1$.

Consider a discretized Brownian motion that moves with only probability p for every step. In this case, we calculate as,

$$D(t) = 2p.$$

While the discreteized Brownian motion with fences hits a wall and bounces back with ratio p only near edges, this process accept a step with the ratio p for the entire time. We deduce that this process have slower diffusion. Thus, we have

$$\lambda \geq 2.$$

Then, we obtain

$$\frac{|D(t) - \tilde{D}(t)|}{D(t)} \leq \frac{1}{4pt} + \frac{\theta}{16p^2t^2},$$

where $0 < \theta < 1$. Hence, the error rate is mainly determined by pt .

Look at the Figure 6. This figure is a numerical result of diffusion coefficients for different t and p . As you see in the figure, for ten times larger t , the points of ten times smaller p are fit on the straight line; they converge to the converges to normal diffusion. This result justified that pt is a dominant factor for the error of the approximation.

4.5 Power-Law Subdiffusion

We shall show some results of numerical experiments to imply the connection between the transient subdiffusion and the power-law subdiffusion.

As in the previous disucussion and in the Figure 7, the error gets larger and $\lim_{t \rightarrow \infty} D(\infty)$ gets smaller as p becomes smaller. Intuitively, a certain distribution of different p fences may lead to power law subdiffusion.

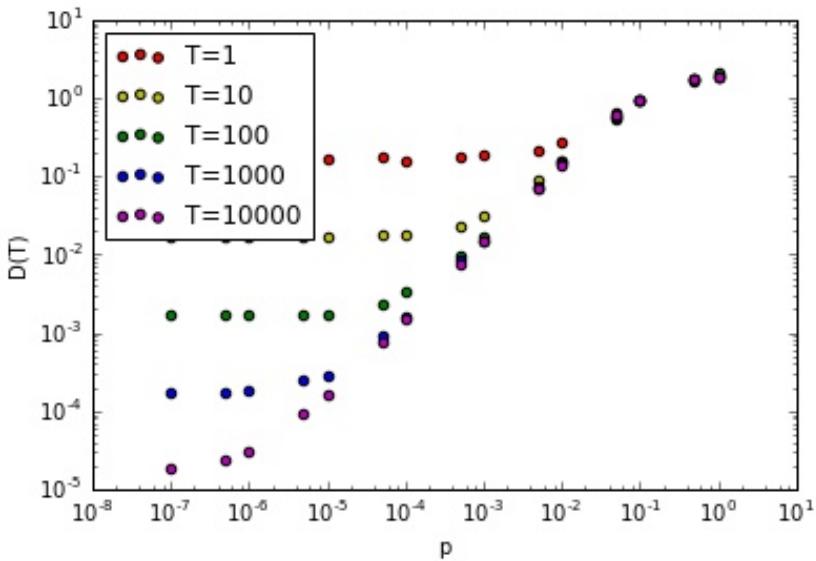


Figure 6: This is the plot of diffusion coefficient for various time. We take points, $p = 0.5, 0.1, 0.05, 0.01, \dots, 5.0 * 10^{-7}, 1.0 * 10^{-7}$. Every multiplication of 10 makes another two points converge.

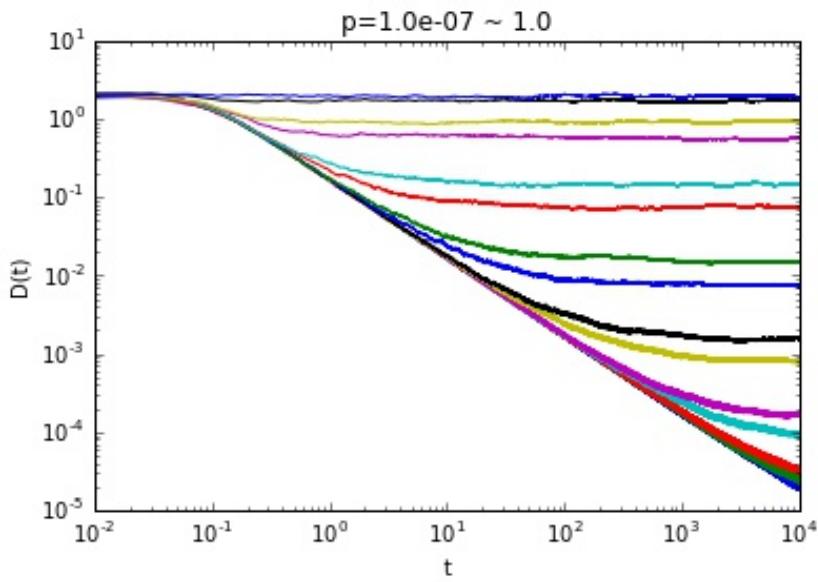


Figure 7: This is the diffusion coefficient plot with different p . We take $p = 1.0, 0.5, 0.1, 0.05, 0.01, \dots, 5.0 * 10^{-7}, 1.0 * 10^{-7}$

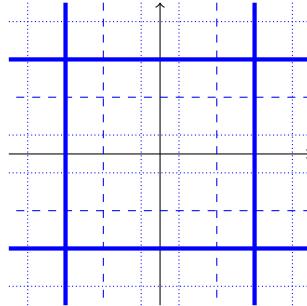


Figure 8: This is an figure of the hierarchical fence model. Different type line indicate different ratio distribution of walls.

Then, we introduce the following model.

Hierarchical Fence Model:

Fix $p_0 > 0$ and $s > 0$. Let i th wall means a wall that has $i - 1$ other walls between itself and the origin. Assume that $i = 2^n \times (\text{oddnumber})$. For i th fence, we define

$$p_i = p_0 \times s^n.$$

In the Figure 9, black lines $\propto t^{0.51}$ fit data with $s = 0.1$ and different three p_0 for $t > 10$. Thus, parameter α for power law subdiffusion is determined by s . In the Figure 10, different s show the different parameter α . For $0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$, we obtain $0.49, 0.61, 0.71, 0.86, 0.94, 0.96, 0.98, 0.98$.

These implies that the power of diffusion coefficient is determined by s and not depends on p_0 . We considered the hierarchical structure; i.e. the same structure in different scales. The property of hierarhical structure have to be determined by the speed for fence parameters to get small the value – s – rather than the first wall paramter p_0 . For example, Consider structures 1 and 2, and let p_0^1 is twice as large as p_0^2 for each structures. We set $s = 0.5$. In this setting, by taking twice larger space scale for 1, we obtain the same structure as 2. Hence, it is natural that only parameter s determines the power of diffusion coefficients.

5 Discussion

We worked on modeling subdiffusion on the plasma cell membrane. We derived that the molecule floating on a liquid state medium consisting of atoms was governed by the Langevin equation in the limit with the mass of an atom to zero. Though this implies that the Langevin equation is a proper model to describe nonequilibrium phenomena, many researchers tend to assume that a molecule rather behaves as a standard Brownian motion. In section 3, we

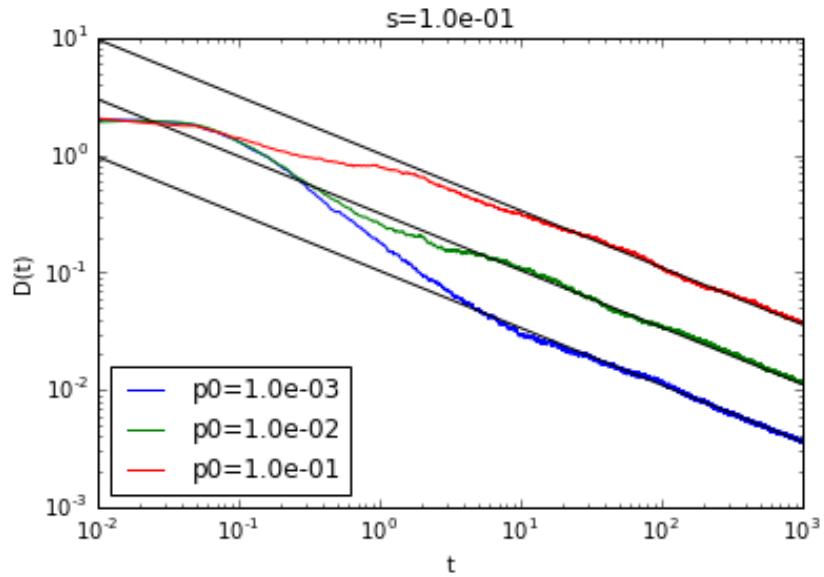


Figure 9: This is the figure of diffusion coefficient for $s = 0.1$ with different p_0 . Black lines are plots of $\beta t^{0.51}$, where β are some constants determined by each p_0 .

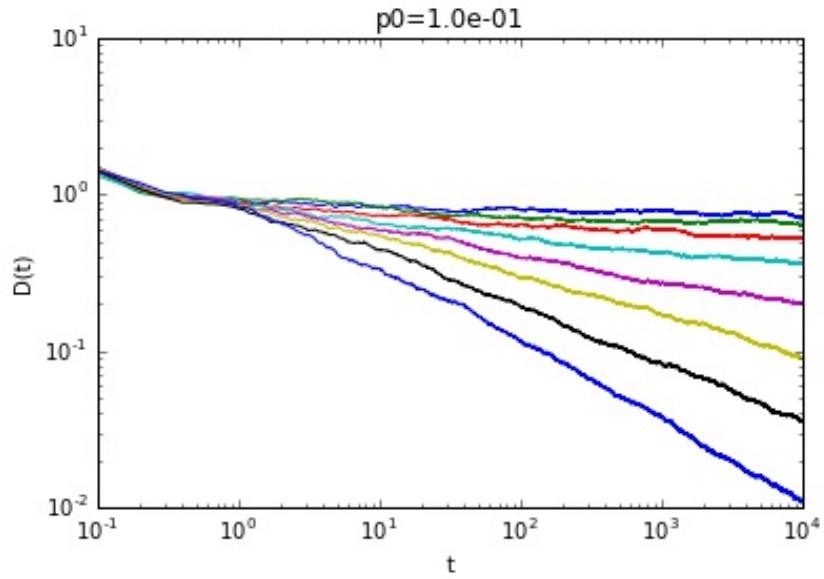


Figure 10: This is the figure of different s with fixed $p_0 = 0.1$. We take $s = 0.1, 0.2, 0.3, \dots, 0.8$.

verified this assumption by showing that the process could be approximated by a standard Brownian motion multiplied by the constant if we have large enough time and length scale observation.

For large enough time t , variances of both the approximation of Brownian motion and the process governed by the Langevin equation are proportional to t . This property is often used to characterize the normal diffusion. Diffusion on the plasma cell membrane, however, does not satisfy this property. The variance grows slower than the normal diffusion. $E(X_t)^2/t$ is not constant and is nonincreasing. To establish a mathematical model to explain this phenomena, most of research papers chose one of the models of the power-law subdiffusion and optimized parameters in order to fit given data [2] [22]. Furthermore, some introduce quantitative criteria to determine which model is proper[1]. Though these approaches may lead good results when the process has a variance decaying with the power-law, a variance of subdiffusion does not necessary follow the power-law. As we showed in the section 4, the Brownian motion hindered by the fence structure leads to transient subdiffusion. This comes from the fact that the process can be approximated by compound Poisson process with jump distribution intensity $p\lambda$ when $pt \gg 1$, where λ is a constant determined by observed data. The transition of diffusion coefficient along with time from 2 – standard Brownian motion – to $2p\lambda$ – compound Poisson process – leads to transient subdiffusion. This decaying diffusion coefficient is not explained by the power-law. Thus, there is no reason to fit all of the subddfisuation data with models of the power-law subdiffusion.

We also showed the numerical results that hierarchical distributed fence parameter p realized the power-law subdiffusion. There are already similar results that hierarchical structure leads to the power-law subdiffusion. Kutasov, Teitel, and Domany [19] showed that the process with self-similar transition probability led to the power-law subdiffusion. It is, however, impossible to introduce infinite hierarchical structures in real physical systems due to the uncertainty principal of quantum mechanics. We cannot define such structures in infinitesimal scale. Saxton [21] showed that even finite hierarchical trapping distribution led to the power-law subdiffusion in the finite time range. These results imply the possibility that finite hierarchical structure in physical systems results in the power-law subdiffusion in a finite observation time range.

There is another possible factor to cause the power-law subdiffusion: long range history. Roughly speaking, the Markov property is the property that the process depends only on the latest information; i.e., the distribution of X_n depends only on the information of X_{n-1} for observation sequence $\{X_n\}_{n \in \mathbb{N}}$. On the other hand, models of the power-law subdiffusion do not

have the Markov property. If we consider the process that has collision with negligible finite mass atoms, we have to consider their long range history. Note that the key of the derivation of the Langevin equation in section 3 is the Markov property. The Markov property allows us to bring discussion to a semigroup theory, which simplifies the proof of convergence. Thus, we are unable to extend our proof to non-Markov processes. We need other approaches from different perspectives.

6 Acknowledgements

I would first like to thank my supervisor Professor Iso of the Graduate School of Informatics at Kyoto University. He not only taught me approaches of mathematical modeling through rigid mathematics and numerical analysis but also helped me make the most of my comprehension in physics for mathematical analysis. I also appreciate useful comments and advices by the other faculty members in our laboratory. They gave me new perspectives on my research.

I have to acknowledge Professor Shigekawa of the Department of Mathematics at Kyoto University, too. He kindly allowed me to join his seminar on stochastic analysis. He helped me enhance my understanding in probability theory.

Finally, I must express my very profound gratitude to my parents and my friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Appendices

A Elastic Collision

In the framework of classical physics, an elastic collision is an encounter between two bodies where the total kinetic energy and the momentum of the two bodies are equivalent between before and after the encounter. This physical phenomena can be understood as an interaction through a potential wall with infinite height.

Consider an inertial system where a molecule stays at the origin; i.e., its velocity \vec{V} is equal to 0 before a collision. Let a denote the radius of molecule.

Suppose that we regard the molecule has the following potential wall,

$$U_n(r) = n(r - a)_+ \quad (\text{A.1})$$

In this fields, an atom has the following energy,

$$E = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r\sin^2\theta\dot{\phi}^2) + U_n(r) \quad (\text{A.2})$$

In [13], Landau explains that when the potential in (A.2) does not depend on ϕ, θ , the Euler-Lagrange equation results to

$$\begin{aligned} mr^2\dot{\theta} &= M \\ mr^2\sin^2\theta\dot{\phi} &= L, \end{aligned} \quad (\text{A.3})$$

where M and L are constant.

Before encountering the molecule, an atom behaves as a free particle and travel on a straight line. Thus, we can choose a plane that contains the atom's orbit and the center of the molecule. Then, we take a coordinate such that $L = 0$. We can rewrite (A.2) by (A.3) as

$$E = \frac{m}{2}\dot{r}^2 + \frac{M^2}{2mr^2} + U_n(r)$$

Then, we obtain

$$\dot{r} = \pm\left(\frac{2}{m}(E - U_n(r)) - \frac{M^2}{2m^2r^2}\right)^{1/2},$$

where we take minus sign before collision and plus sign after collision.

Let r^* denote a positive real number such tat $(\frac{2}{m}(E - U_n(r*)) - \frac{M^2}{2m^2r_*^2})^{-1} = 0$. Before the collision, the radius component r decreases until r^* . When $r < a$, the molecule velocity has interaction with the atom and change molecule's velocity. Therefore, the molecule does not stay at the origin of the system anymore and the description (A.1) is not anymore correct. However, by taking limit n to ∞ , r^* goes to a and $\dot{r} = \pm\left(\frac{2}{m}(E - \frac{M^2}{2m^2r^2})\right)^{1/2}$ only except for infinitely short time. Then the potential description is still available.

Considering $\theta(t) \equiv \theta(r(t))$, we have $\frac{d\theta}{dt} = \frac{d\theta}{dr}\frac{dr}{dt}$, which gives

$$\begin{aligned} \frac{d\theta}{dr} &= \frac{d\theta}{dt}\left(\frac{dr}{dt}\right)^{-1} \\ &= \frac{M}{mr^2} \pm \left(\frac{2}{m}E - \frac{M^2}{2m^2r^2}\right)^{-1/2} \end{aligned}$$

Let θ_1 denote the angular coordinate when the collision happens and let θ_2 denote the angular coordinate after infinite time has passed. We can calculate as

$$\begin{aligned}\theta_1 &= \int_{\infty}^a -\frac{M}{mr^2} \pm \left(\frac{2}{m}E - \frac{M^2}{2m^2r^2}\right)^{-1/2} dr \\ \theta_2 &= \theta_1 + \int_a^{\infty} \frac{M}{mr^2} \pm \left(\frac{2}{m}E - \frac{M^2}{2m^2r^2}\right)^{-1/2} dr.\end{aligned}$$

We have

$$\theta_2 = 2\theta_1. \quad (\text{A.4})$$

Clearly, θ_1 and θ_2 are equivalent to the angular with respect to normal direction of the collision place. Thus, the angle of the incidence and the reflection are equivalent.

Let \vec{V} and \vec{v} denote velocities of the molecule and the atom. We denote by $(\cdot)^L$ and $(\cdot)^0$ values of at the laboratory system and the center-of-mass system, i.e.; we have

$$\begin{aligned}\vec{V}^L &= \vec{V}^0 + \frac{M\vec{V}^L + m\vec{v}^L}{M+m}, \\ \vec{v}^L &= \vec{v}^0 + \frac{M\vec{V}^L + m\vec{v}^L}{M+m}.\end{aligned} \quad (\text{A.5})$$

Since the angle of the incidence and the reflection are determined by the relative velocity between the molecule and the atom, Galilean transformations does not ruin the relation (A.4). Hence, the kinetic energy conservation leads to

$$\begin{aligned}\tilde{\vec{v}}_t^0 &= \vec{v}_t^0 \\ \tilde{\vec{v}}_n^0 &= -\vec{v}_n^0,\end{aligned} \quad (\text{A.6})$$

where we denote by $(\tilde{\cdot})$ the value of after the collision. We define

$$\vec{v}_{diff} = \vec{v}^L - \vec{V}^L = \vec{v}^0 - \vec{V}^0. \quad (\text{A.7})$$

$$M\vec{V}^0 + m\vec{v}^0 = M\vec{V}^0 + m\vec{v}^0 = 0 \quad (\text{A.8})$$

By (A.7) and (A.8), we obtain

$$\begin{aligned}\tilde{\vec{V}}^L - \vec{V}^L &= \tilde{\vec{V}}^0 - \vec{V}^0 = -\frac{m}{M+m}(\tilde{\vec{v}}_{diff} - \vec{v}_{diff}) \\ \tilde{\vec{v}}^L - \vec{v}^L &= \tilde{\vec{v}}^0 - \vec{v}^0 = \frac{M}{M+m}(\tilde{\vec{v}}_{diff} - \vec{v}_{diff})\end{aligned} \quad (\text{A.9})$$

(A.6) reduces (A.9) to

$$\begin{aligned}\vec{\tilde{v}}_t^L &= \vec{v}_t^L, \quad \vec{\tilde{V}}_t^L = \vec{V}_t^L, \\ \vec{\tilde{V}}_n^L - \vec{V}_n^L &= \frac{2m}{M+m}(\vec{v}_n^L - \vec{V}_n^L)\end{aligned}\tag{A.10}$$

$$\vec{\tilde{v}}_n^L - \vec{v}_n^L = -\frac{2M}{M+m}(\vec{v}_n^L - \vec{V}_n^L)\tag{A.11}$$

(A.10) and (A.11) can be converted to

$$\vec{V}_n^L = \frac{M-m}{M+m}\vec{V}_n^L - \frac{2m}{M+m}\vec{v}_n^L\tag{A.12}$$

$$\vec{v}_n^L = -\frac{M-m}{M+m}\vec{v}_n^L + \frac{2M}{M+m}\vec{V}_n^L\tag{A.13}$$

B Stochastic Fubini

It is known that certain conditions allow you to change the order between Lebesgue integral and stochastic integral as follows,

Theorem B.1. *Let (S, Σ, μ) be a σ -finite measure space and (Ω, \mathcal{F}, P) be a probability space. We have $\phi : S \times [0, T] \times \Omega \rightarrow \mathbb{R}$ which is $\Sigma \otimes \mathcal{B} \otimes \mathcal{F}$ -measurable. If $(t, \omega) \rightarrow \phi_s(t, \omega) := \phi(s, t, \omega)$ is integrable with respect to semi-martingale X for all $s \in S$, if the process $(t, \omega) \rightarrow \int_S \phi_s d\mu(s)$ is well defined integrable with respect to X , and if*

$$\int_S \left| \int_0^T \phi_s(t) dX(t) \right| d\mu(s) < \infty \text{ almost surely,}$$

then, we have

$$\int_S \int_0^T \phi_s(t) dX(t) d\mu(s) = \int_0^T \int_S \phi_s(t) d\mu(s) dX(t) \text{ almost surely.}$$

We can see the proof at [18].

References

- [1] M. M. Tamkunc A. V. Weigela B. Simonb and D. Krapfa. “Ergodic and Nonergodic Processes Coexist in the Plasma Membrane as Observed by Single-Molecule Tracking”. In: *PNAS* 108 (2011), pp. 6438–6443.

- [2] D. S. Banks and C. Fradin. “Anomalous Diffusion of Proteins Due to Molecular Crowding”. In: *Biophys. J.* 89 (2005), 2960–2971.
- [3] P. Billingsley. *Convergence of Probability Measures*. Wiley-Interscience publication, 1999. ISBN: 0471197459.
- [4] J. P. Bouchaud and R. Cont. “A Langevin Approach to Stock Market Fluctuations and Crashes”. In: *R. Eur. Phys. J. B* 6 (1998), 543–550.
- [5] K. L. Chung. *Course in Probability Theory*. Academic Press, 2001. ISBN: 0121741516.
- [6] J. L. Lebowitz D. Durr S. Goldstein. “Mechanical Model of Brownian Motion”. In: *Commun. Math. Phys.* 78 (1981), pp. 507–530.
- [7] JA. Dix and AS. Verkman. “Crowding Effects on Diffusion in Solutions and Cells”. In: *Annu. Rev. Biophys.* 37 (2008), pp. 247–263.
- [8] A. Einstein. “The Theory of the Brownian Movement”. In: *Ann. der Physik* 17 (1905), p. 549.
- [9] JH. Jeon and R. Metzler. “Fractional Brownian Motion and Motion Governed by the Fractional Langevin Equation in Confined Geometries”. In: *Phys. Rev. E* 81 (2010), p. 11.
- [10] T. Fujiwara K. Murase A. Kusumi K. Ritchie R. Lino. “The Fence and Picket Structure of the Plasma Membrane of Live Cells as Revealed by Single Molecule Techniques (Review)”. In: *Molecular Membrane Biology* (2003).
- [11] I. Karazas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate texts in mathematics. Springer-Verlag, 1991. ISBN: 0387976558.
- [12] R. Kubo. “The Fluctuation-Dissipation Theorem”. In: *Rep. Prog. Phys.* 29 (1966).
- [13] L. D. Landau and E. M. Lifshitz. *Mechanics*. Pergamon Press, 1976. ISBN: 0750628960.
- [14] L. D. Landau and E. M. Lifshitz. *Statistical Physics Part 1*. Pergamon Press, 1980. ISBN: 0080230393.
- [15] S. Lemonsa and A. Gythiel. “Paul Langevin’s 1908 paper “On the Theory of Brownian Motion” [“Sur la theorie du mouvement brownien,” C. R. Acad. Sci. (Paris) 146, 530–533 (1908)]”. In: *Am. J. Phys.* 65 (1997), p. 1079.
- [16] M. M. Meerschaert and H. P. Scheffler. “Limit Theorem for Continuous-Time Random Walks with Infinite Mean Waiting Times”. In: *J. Appl. Prob.* 42 (2004), pp. 623–638.

- [17] R. Metzler and J. Klafter. “The Random Walk’s Guide to Anomalous Diffusion: a Fractional Dynamics Approach”. In: *Phys. Rep.* 399 (2000), pp. 1–77.
- [18] J. Van Neerven and M. Veraar. “On the Stochastic Fubini Theorem in Infinite Dimensions”. In: *Lect. Notes Pure Appl. Math.* 245 (2006), pp. 323–336.
- [19] D. Kutasov S. Teitel and E. Domany. “Diffusion and Dynamical Transition in Hierarchical Systems”. In: *Phys. Rev. B* 36 (1987), pp. 684–698.
- [20] P. G. Saffman and M. Delbrück. “Brownian Motion in Biological Membranes”. In: *PNAS* 72 (1975), pp. 3111–3113.
- [21] M. J. Saxton. “A Biological Interpretation of Transient Anomalous Subdiffusion. I. Qualitative Model”. In: *J. Biophys.* 92 (2007), 1178–1191.
- [22] M. J. Saxton. “Anomalous Diffusion Due to Obstacles: a Monte Carlo study”. In: *Biophys. J.* 66 (1994), pp. 394–401.
- [23] S. J. Singer and G. L. Nicholson. “Fluid Mosaic Model of Structure of Cell-Membranes”. In: *Science* (1972).
- [24] I. M. Sokolov. “Models of Anomalous Diffusion in Crowded environments”. In: *Soft Matter* 8 (2012), pp. 9043–9052.
- [25] T.K.Kurtz. “Semigroups of Conditioned Shifts and Approximation of Markov Process”. In: *Ann Prob* 3 (1975), pp. 618–642.
- [26] A. Wedemeier, H. Merlitz, and J. Langowski. “Anomalous Diffusion in the Presence of Mobile Obstacles”. In: *EPL* 88 (2009), p. 38004.
- [27] R. Zwanzig. “Nonlinear Generalized Langevin Equations”. In: *J. Stat. Phys.* 9 (1973), 215–220.