Universidad Autonoma de Madrid & ICMAT
Department of Mathematics

Reaction-Diffusion Processes and Their Interdisciplinary Applications

Alvaro Isidro Correales Fernandez

Advisor: Carlos Escudero Liebana

Tutor: Javier Cárcamo Urtiaga

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy, September 2019

Acknowledgements

- Thanks to my supervisor, Carlos Escudero Liébana, for the numerous hours of discussion, sharing of ideas, good references and for the filtering of attainable ideas and projects.
- Thanks to Javier Cárcamo Urtiaga and Jezabel Curbelo for accepting reading this document and help me through the final steps of my thesis.
- Thanks to Mariya Ptashnyk for her help and insights in our joint research paper that is an essential part of this work.
- Thanks to all researchers and colleagues from ICMAT and UAM for being an inspiration to me, for all the challenging discussions and seminars.
- Thanks to all researchers and colleagues from Stockholm University and NORDITA, for welcoming me so well in the land of winter. Special thanks to Joanna Tyrcha and John Hertz, for help me understand the hard concepts of physics regarding path integrals and their intimate relation with Probability. Thanks to Chun-Biu Li and Steffano Bo for their applied spirit that has inspired me to try to see the connection of theory and applications, especially in biology.

Dedication

As important, or even more, as the theoretical and research knowledge that I had to acquire for the completion of this thesis, has been the support, love, and understanding that so many people has shown to me during this long and, sometimes, difficult process.

Thanks to all my friends, both from ICMAT and UAM, that made me feel like home from the very beginning, from La Latina to Argüelles.

Thanks to Miguel, Pablo, Sonja and David, Jose Luis, Manu, Irene, Alvaro, Carlos, Adri, Nico, Maria Jose, and so many others. For all those beers and tapas that we have enjoyed together.

Thanks to all my friends from Stockholm, for all those trips around the nordic countries and for showing me the world was even bigger than Madrid. Thanks to Peter, Elin, Bara, Gino and Sanna, Babis and so many more. Without you, the cold would feel way worst.

Thanks to my family, for all their support, help, advice and unconditional love. Thanks to my parents Angel and Maria del Carmen, to my brothers Ernesto and Angel David, to Maria and Leo.

To Marta, my bigger inspiration and motivation for fighting against the theorems that do not want to be proved. For all the love. I do not have enough words for expressing all the hours, the doubts, the pain, that you shared with me along this journey. Thanks for everything.

‘We have not succeeded in answering all our problems.
The answers we have found only serve to raise a whole set
of new questions. In some ways we feel we are as confused
as ever, but we believe we are confused on a higher level
and about more important things.’
Tromso University

‘The only mathematics I have never used
are the ones I did not know.’
Anonymous

Contents

Chapter 1

Introduction

1.1 Motivation and Objectives

Reaction-diffusion models can be characterized as the study of the macroscopic dynamics of a physical variable, e.g., spatial concentration of chemical or biological species, in which the evolution of the system comes from the interplay of two sources: local interactions (reactions) and random movement (diffusion) in space.

This characterization leads, usually, to the study of a system of partial differential equations sketched as

$$\partial_t q(t,x) = D\Delta_x q(t,x) + R(q). \tag{1.1}$$

This equation has different names throughout the literature, in physics is a Schrödinger equation in imaginary time, in differential equations theory is a parabolic PDE and, in probability theory, is a Feynman-Kac type Kolmogorov equation.

The intention of the author of this thesis is not to be constraint by language, each point of view has its origins, motivations and, most importantly, consequences in the way a theory develops. To have a global point of view is a necessary condition when applying the machinery that has been developed during years, by different communities, to specific problems.

In reaction-diffusion models, one usually makes one significant assumption in writing (1.1): the local interactions are assumed to be deterministic. This assumption can be plausible in some specific cases, depending on the physics one wants to describe, by making a macroscopic (or thermodynamic)

approximation. Unfortunately, this assumption breaks down in cases like the ones we are interested in studying, e.g., the study of molecules in a biomolecular setting. The reduced number of molecular components involved in biomolecular networks makes noise an important actor in the study of local interactions. This feature makes the study of local interactions a challenge, but not just that, the noise appearing in this scale introduce new phenomena impossible to observe otherwise.

Actually, in this thesis, we focus on the study of local interactions. Still, equations of the type (1.1) can be found widespread in the text. The reason is that we still have noise, which arise from the micro or mesoscopic scales. If one wants to study both noises, that is, the one emerging from microscopic interactions and the spatial diffusion, one has to consider partial stochastic differential equations instead of deterministic ones, a subject that is still not clear how to tackle from a mathematical point of view.

The outline of the thesis is as follows; the main object of study is Doi-Peliti theory, which is a field-theoretic approach to Markov processes with discrete state space. This theory appeared in the theoretical physics literature at the end of the 70s [1, 2] and has been profoundly studied since then [3, 4]. Still today, certain aspects of this theory are criticized in the physics literature [5]. Our primary focus is to shed some light about this theory by using mathematical tools provided by probability theory and analysis. Along the way, we connect this theory with concepts familiar to probability theory: generating functions, duality in stochastic processes, and stochastic differential equations (SDEs).

Following this line of thought, in chapter (3), we focus on Doi-Peliti theory and establish the theoretical ground from which we look at Markov processes in discrete space. We mainly study abstract chemical reactions, an especial class of Markov processes, which have a unique combinatoric nature that makes their study more suitable from the perspective of Doi-Peliti theory. In particular, we write the evolution operator of the Markov process as combinations of creation and annihilation operators.

After our study of abstract chemical reactions, we derive three observations from it:

1. In chapter (4), we study the concept of duality in stochastic processes, which is related to coherent transformations, an object defined by Doi-Peliti theory. From there, we derive the connection between Doi-Peliti theory and the theory of sequences of polynomials, particularly, Sheffer sequences. As an interesting consequence, we can see the theory of generating functions of probability distributions is an instance of a Sheffer sequence representation, where the sequence of polynomials is $\{x^n\}_{n\in\mathbb{N}}$.

Applying the concept of duality to abstract chemical reactions, a particular type of second-order degenerated (in the boundaries) PDEs tend to appear. We characterize this duality as the extension of the classical Feynman-Kac theory to this context.

2. In chapter (5), and motivated by our previous study on degenerated PDEs and SDEs, we contribute to the lengthy discussion on the 'correct' stochastic interpretation. The so-called Itô vs. Stratonovich dilemma, which is usually addressed by saying that, after all, both formulations are equivalent to each other under certain conditions that are fulfilled in practice.

   Is in this last point, 'fulfilled in practice', where we see the weakness of such an explanation. As we present very famous examples of SDEs which do not fulfill that equivalence between the Itô and Stratonovich interpretation, in fact, we see that even new solutions can appear if one naively applies this 'equivalence'.

3. In chapter (6), and coming back to the coherent transformation from chapter (3), we will study a complex noise SDE. This SDE appears when applying the machinery of previous chapters to the coalescence reaction. Therefore, this motivates the study of the relation between negative diffusion PDEs and their formally equivalent complex SDEs. A key ingredient to uncover that relation is the Cauchy transformation, which is a way of representing certain spaces of distributions as holomorphic functions in the complex plane except for the real line. The distribution appears exactly because of the 'jump' of values of these functions over the real line.

Finally, we make use of field-theoretical methods to study a Markov process related to chemical biology. More concretely, we apply Martin-Siggia-Rose theory to a system of stochastic processes that model the interaction of a protein and mRNA molecules which behavior follows a positive feedback loop; leading to bi-stability. Through the application of WKB (Wentzel–Kramers–Brillouin) theory and the help of numerical methods, we end up with an expression that can be used to compute the average time of transition between the two locally stable solutions. The interest of this problem relies on the bi-stability of the system, which can only be explained by the stochastic nature of the interactions between the biological components inside the cell.

Some keywords are: **Markov chain**, **Generating function**, **Doi-Peliti field theory**, **Duality**, **Feynman-Kac**, **Itô vs. Stratonovich**, **Complex Itô equation**, **Cauchy transform**, **Distribution**, **Escape rate**, **Metastability**, **Martin-Siggia-Rose field theory**

1.2 Motivación y objetivos

Los modelos de reacción-difusión pueden ser definidos como el estudio del comportamiento macroscópico de una determinada variable física, e.g., la concentración espacial de agentes químicos o biológicos, en los que la evolución del sistema está dominada por dos fuentes principales: interacciones locales (reacciones) y movimiento aleatorio (difusión) por el espacio.

Esta caracterización deriva, normalmente, en el estudio de un sistema de ecuaciones differenciales parciales que suelen tener la forma

$$\partial_t q(t,x) = D\Delta_x q(t,x) + R(q). \tag{1.2}$$

Esta ecuación aparece con diversos nombres en la literatura, en física es una ecuación de Schrödinger en tiempo imaginario, en la teoría de ecuaciones diferenciales es una EDP parabólica y, en probabilidad, es una ecuación de Kolmogorov con término de Feynman-Kac.

La intención del autor de esta tesis es no estar constreñidos por el lenguaje, cada punto de vista tiene sus propios origenes, motivaciones y, más importante aún, sus consecuencias en la forma en la que una teoría se desarrolla. Tener una visión global del problema es una condición necesaria cuando queremos estudiar las distintas técnicas que se han ido desarrollando a lo largo de los años para resolverlos.

En modelos de reacción-difusión, usualmente se asume cuando se escribe (1.2), que las interacciones locales son deterministas. Esta hipótesis puede ser plausible en ciertos casos, dependiendo del problema que queramos describir, en límites macroscópicos (o termodinámicos) del sistema. Sin embargo, tratar de usar la ecuación (1.2) como aproximación para el caso en el que la hipótesis termodinámica no es razonable puede llevar a grandes errores de modelización y predicción. El reducido número de agentes químicos que tienen lugar en reacciones biomoleculares hace del ruido un elemento fundamental en las reacciones locales. Esto convierte la modelización de reacciones locales en un problema sumamente complejo.

En esta tesis, nuestro foco principal son precisamente las reacciones locales. Aun así, ecuaciones del tipo (1.2) aparecen a lo largo de este texto. La razón de ello deriva de que todavía tenemos ruido. El proveniente de la escala micro o mesoscópica. Si alguien quisiera estudiar el ruido proveniente de las reacciones locales junto al puramente difusivo, el elemento natural de estudio serían las ecuaciones estocásticas en derivadas parciales, que todavía son muy complejas incluso de formular desde un punto

de vista riguroso, matemáticamente hablando.

El pilar central de la tesis se base en el estudio de la teoría de Doi-Peliti, que es un ejemplo de teoría de campos aplicado al estudio de procesos de Markov con espacio discreto. Esta teoría apereció a finales de los años 70 [1, 2], y ha sido estudiada desde entonces generando un intenso campo de estudio [3, 4]. Incluso hoy en día, ciertos temas relacionados con esta teoría generan aún algunas discrepancias dentro de la propia comunidad [5]. El objetivo de esta tesis es proveer de herramientas matemáticas para el estudio de estos dilemas con técnicas que van desde la teoría de probabilidad al análisis, así como conectar esta teoría con temas clásicos de probabilidad como las funciones generatrices de probabilidad, la dualidad de processos estocásticos o las ecuaciones diferenciales estocásticas.

Siguiendo estas líneas, en el capítulo (3), nos centramos en el estudio propiamente dicho de la teoría de Doi-Peliti y establecemos las bases teoricas para el estudio de ciertos procesos de Markov. En particular, estudiamos reacciones químicas abstractas, que son especialmente subceptibles de ser formuladas bajo el prisma de la teoría de Doi-Peliti debido a sus propiedades combinatorias. En particular, vemos que podemos escribir el operador de evolución de estos procesos como combinación de los operadores de creación y destrucción.

A través de las ecuaciones que aparecen en nuestro estudio de las reacciones químicas abstractas, nos centramos en varias consecuencias de dicho estudio en el siguiente orden

1. En el capítulo (4), estudiamos el concepto de dualidad en el contexto de procesos estocásticos, que está relacionado con la teoría de Doi-Peliti a través de la transformada coherente. También estudiamos la relación de la teoría de Doi-Peliti con la teoría de secuencias de polinomios, en particular, con las secuencias de Sheffer. Como consecuencia de ello, la teoría de funciones generatrices de probabilidad puede verse como un caso particular de estas relaciones donde la secuencia de polinomios son los monomios $\{x^n\}_{n\in\mathbb{N}}$.

   También estudiamos en este capítulo ciertos tipos de EDPs de segundo orden degeneradas (en la frontera), que aparecen como consecuencia de la dualidad entre estos procesos de Markov discretos y ciertas EDPs dadas por la transformada coherente. Así mismo, vemos una extensión del clásico teorema de Feynman-Kac en este contexto de EDPs degeneradas en la frontera.

2. En el capítulo (5), y motivados por el estudio de las EDPs degeneradas, damos una nueva vuelta de tuerca a la vieja discusión de la correcta interpretación estocastica, la conocida dualidad Itô

vs Stratonovich. Este dilema suele cerrarse con la observación de que ambas formulaciones son equivalentes en ciertas condiciones que suelen cumplirse en la practica. Desde nuestro punto de vista, esto no es del todo cierto, y presentamos ecuaciones estocásticas clásicas para las que la transformación antes indicada entre las formulaciones de Itô y Stratonovich, llevan a la creacion artificial de soluciones, por lo que tal equivalencia no existe en el contexto de estas ecuaciones.

3. En el capítulo (6), y volviendo al estudio de la transformada coherente del capítulo (3), estudiamos ciertas ecuaciones estocásticas con ruido imaginario que aparecen de forma natural cuando aplicamos las herramientas antes estudiadas a la reacción de aniquilación. Motivados por este caso, estudiamos en general la relación entre ecuaciones estocásticas con ruido imaginario y sus correspondientes EDPs con difusión negativa. La clave para relacionar estos conceptos se encuentra en la transformada de Cauchy, que nos permite identificar un cierto espacio de distribuciones con el espacio de funciones holomorfas en el plano complejo salvo por la linea real. De hecho, es debido al salto de estas funciones en la linea real por la que distribuciones de todo tipo pueden ser representadas.

Finalmente, aplicamos métodos de teoría de campos al estudio de procesos de Markov que aparecen de forma natural en el campo de la química biológica. Aplicando la conocida teoría de Martin-Siggia-Rose sobre un sistema de ecuaciones estocásticas que representan la interacción de una proteína y ARNm, llegamos al comportamiento biestable de sus soluciones. Gracias a la aplicación de métodos como WKB (Wentzel–Kramers–Brillouin) y expansiones asintóticas, junto con la ayuda de métodos numéricos, encontramos una expresión para el tiempo medio de transición entre las dos soluciones localmente estables del sistema, que tiene un interés biológico claro. Es en el estudio de este tipo de reacciones donde el ruido que aparece de forma natural en el contexto biológico, hace emerger propiedades (biestabilidad) que se perderían en cualquier aproximación determinista de dicho sistema.

Algunas palabras clave de esta tesis son: **Cadenas de Markov**, **Función generatriz de probabilidad**, **Doi-Peliti**, **Duality**, **Feynman-Kac**, **Itô/Stratonovich**, **Ecuación de Itô compleja**, **Transformada de Cauchy**, **Distribuciones**,**Velocidad de escape**, **Metaestabilidad**, **Martin-Siggia-Rose**

1.3 Publications

1. Hertz J., Correales A. and Tyrcha J. **Stochastic activation in a genetic switch model** Phys. Rev. E 98, 052403 (2018)

2. Correales A., and Escudero Liebana C. **Absence of Itô/Stratonovich duality in physical theories with absorbing states** Journal of Mathematical Physics (Accepted Nov 2019)

3. Correales A., Escudero Liebana and Ptashnyk M. **Chemical Kinetics, Markov Chains, and the Imaginary Itô Interpretation** (Submitted Nov 2019 to Communications in Mathematical Physics)

Chapter 2

Background Theory

In this first chapter, we are going to introduce, in a succinct way, what we think of the essential background to the subsequent work. This background is needed for the sake of conceptualization and notation.

Therefore, we do not have the intention of developing a complete background theory. In particular, theorems are presented without proof. The only intention of this chapter is to be the seed for the later work.

2.1 Probability Essentials

We introduce the axiomatic view of Kolmogorov as the way of looking at probability. That is, we suppose, a priori, the existence of a set Ω which elements $\omega \in \Omega$ are, intuitively, the possible outcomes from a certain experiment.
However, as the "paradoxes" of the naive probability theory tells us, we need an additional structure. This structure is a σ-field [1] of events of $\mathcal{A}$, which is a set of subsets of Ω with the following properties:

1. $\Omega \in \mathcal{A}$.
2. If $A \in \mathcal{A}$ then also the complement $A^c \in \mathcal{A}$.
3. If $A_n \in \mathcal{A}\ \forall n \in \mathbb{N}$ then, $\cup_n A_n \in \mathcal{A}$.

[1] Sometimes called σ-algebra.

That is, the pair $(\Omega, \mathcal{A})$ is a measurable space. Now, we can properly define a probability measure P, i.e., a function from the σ-algebra $\mathcal{A}$ to the interval $[0, 1]$ such that:

1. $P(\Omega) = 1$.
2. $P(\cup_n A_n) = \sum_n P(A_n)$ whenever $A_i \cap A_j = \emptyset$ for $i \neq j$.

This structure can be summarize just saying the triplet $(\Omega, \mathcal{A}, P)$ is a probability space.

In almost all the cases we are interested in, the set Ω possesses a topological structure. In that case, the σ-algebra of interest should be the generated, somehow, by the topology. The intuition tells us that the smallest $\mathcal{A}$ that can be generated by finite iteration between complements and numerable unions from the topology is a σ-algebra. It is possible to prove that's the case. This situation covers, in particular, the typical case in which the set is $\Omega = \mathbb{R}^n$.

Once we have defined our structure of probability, we have all elements into play for the subsequent probabilistic concepts. In a significant part of the cases, an event ω contains more information we could ever need in practice. For example, if the event is "$1,000$ tosses of a dice", we could be interested in a particular outcome, or maybe we could measure the sum of the $1,000$ outcomes. That is, we want to define an observable from the space of events Ω to a particular space with the proper structure of "measurement", typically $\mathbb{R}^n$.
In this context, an observable just means a function between measurable spaces

$$X : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}'),$$

such that the probability of an observation is the probability of the possible events that would produce the same observation, that is, we define a probability measure in Ω' as the push-forward measure of P,

$$P'(A') := P(X^{-1}(A')).$$

To do that, we need the following compatibility condition

Definition 2.1 (Measurable function)**.** *X is called measurable whenever $X^{-1}(\mathcal{A}') \subset \mathcal{A}$*

In the special case $\Omega' = \mathbb{R}$ ($\mathbb{R}^n$) then X is known as a random variable (vector).

Definition 2.2. *The expectation of a random variable $X : \Omega \to \mathbb{R}$ is defined as*

$$E[X] := \int_{\Omega} X(\omega) dP(\omega) = \int_{\mathbb{R}} x dP'(x)$$

Similarly, the expectation of a random vector is defined component-wise.

In case the probability distribution P' could be written as the Radon-Nikodym derivative of the Lebesgue measure μ of $\mathbb{R}$, that is, $dP' = f d\mu$, the derivative f is called the density function of the random variable.

From the expectation, a collection of definitions appear, we are not going to define them all, but just a couple of interesting ones:

1. The n-th moment of X is defined as $E[X^n]$, the first moment is called mean.
2. The n-th centered moment is defined as $E[(X - E[X])^n)]$, the second one is called variance.
3. The characteristic function is defined as $E[e^{itX}]$, which reduce to the Fourier transform of the push-forward measure.
4. Given two random variables $X_{1,2}$ the covariance is defined as $E[(X_1 - E[X_1])(X_2 - E[X_2])]$.

Among the different probability distributions that appear in practice, two of them clearly predominate:

1. The delta distribution δ_a, which is equal one for all events A such that $a \in A$ and zero otherwise. This distribution allows an unified framework between continuous and discrete probability distributions.
2. The normal distribution $\mathcal{N}(\mu, \sigma^2)$, defined from the Gaussian density $f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, which emerge naturally as a "macroscopic" limit in cases where there are great quantity of random components, this will be made precise later when studying limiting theorems.

Intuitively, probability is a notion generated by two conditions:

- Symmetry a priori: Which makes a possible outcome quantitatively (a priori) more os less likely than others.

- Iteration: Which tell us that in a certain limit, the symmetry a priori emerge as a measurable quantity called probability.

We have learned how to measure probabilities[2], now we want what we measure matches with our notion of probability. Mathematically, two ingredients appear independence and convergence.

Definition 2.3 (Independent events)**.** *We say that an arbitrary set of events $A_\gamma \in \Omega$ with $\gamma \in \Gamma$ set of indexes, are independent when $P(A_{\gamma_1} \cap \cdots \cap A_{\gamma_n}) = P(A_{\gamma_1}) \cdots P(A_{\gamma_n})$ for all n with $\gamma_i \neq \gamma_j \in \Gamma$ whenever $i \neq j$.*

This notion can be expand to sub-sigma algebras

Definition 2.4 (Independent sub-sigma algebras)**.** *We say an arbitrary set of sub-sigma algebras $\mathcal{A}_\gamma \subset \mathcal{A}$ with $\gamma \in \Gamma$ are independent, when for all events $A_1 \in \mathcal{A}_{\gamma_1}, \ldots A_n \in \mathcal{A}_{\gamma_n}$ are independent.*

Every random variable (vector) generates a sub-sigma algebra in Ω; the pull-back of the natural sigma algebra in $\mathbb{R}$ ($\mathbb{R}^n$). We then define an arbitrary set of random variables (vectors) as independent whenever the sub-sigma algebras they generate in Ω are independent.

Another important probabilistic concept is the restriction of a random variable to a particular set $B \in \mathcal{A}$, which is a way of re-organizing our knowledge under privileged information

Definition 2.5 (Conditional probability)**.** *Let $B \in \mathcal{A}$ with $P(B) > 0$, we define the conditioned probability $P(A\,|B\,) := \dfrac{P(A \cap B)}{P(B)}$*

There is a probability distribution, that can be defined with the memoryless property, that is, a random variable $T : (\Omega, \mathcal{A}, P) \longrightarrow (0, \infty)$ is said to obey an exponential distribution if, and only if,

$$P(T > t + s\,|T > s\,) = P(T > t)$$

for all $0 < t, s$. This property defines uniquely the distribution if we set the value $E\,[T] = \lambda$, in that case its probability density is

$$f_\lambda(t) = \frac{1}{\lambda} e^{-\frac{t}{\lambda}}.$$

The exponential distribution is essential for defining Markov processes with discrete state space.

[2] That's why measure theory enters in play

For notions of convergence, one can define several concepts. Some of these notions are intrinsic to probability concepts already defined, while others, come from the general theory of measurable spaces.

Let X_n be a sequence of random variables (vectors) with $n \in \mathbb{N}$ and X be a random variable (vector):

Almost sure convergence $X_n \xrightarrow{\text{a.s.}} X$, if

$$\lim X_n(\omega) = X(\omega)$$

for all $\omega \notin N$ with $P(N) = 0$.

Convergence in probability $X_n \xrightarrow{\text{Pr}} X$, if

$$\lim P(|X_n - X| > \epsilon) = 0$$

for all $\epsilon > 0$.

Convergence in mean of order p(≥ 1) $X_n \xrightarrow{\text{p}} X$, if

$$\lim E(|X_n - X|^p) = 0$$

.

Convergence in law $X_n \xrightarrow{\text{L}} X$, if

$$\lim F_{X_n}(x) = F_X(x)$$

at any point x where the distribution function F_X is continuous.

Now we can understand the two critical theorems of probability theory properly

Theorem 2.1 (Strong law of large numbers)**.** *If $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of random variables*

$$X_n : (\Omega, \mathcal{A}, P) \longrightarrow \mathbb{R}$$

that are independent, identically distributed with finite mean $E[X_n] = \mu$ and finite variance $Var(X_n) = \sigma^2$, then the sample mean

$$S_n := \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{a.s.} \mu$$

Theorem 2.2 (Central limit theorem)**.** *If* $\{X_n\}_{n\in\mathbb{N}}$ *is a sequence of random variables*

$$X_n : (\Omega, \mathcal{A}, P) \longrightarrow \mathbb{R}$$

that are independent, identically distributed with finite mean $E[X_n] = \mu$ *and finite variance* $Var(X_n) = \sigma^2$*, then*

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{L} \mathcal{N}(0,1)$$

2.2 Stochastic Processes and Stochastic Differential Equations

In this section, the main ingredient is the dynamics of random variables, in which time will be a privilege variable.

A stochastic process can be defined as a parameterized family of random variables X_T, in which the parameter space has a structure of time, that is, the parametric family is an ordered set with a notion of past and future, with the condition that past affects future but not the other way around. The parameter space will be either discrete $T \in \mathbb{N}$ or continuous $\{T > 0\}$.

We are not going to define each object related to stochastic processes in a rigorous way but rather give some intuitive concepts jumping from one main definition to another in a rather chaotic way. For a more complete study of the topic, we can recommend, just to cite a few [6, 7, 8, 9, 10].

Unless otherwise stated, all the stochastic processes that will appear in this thesis are going to be Markov processes, which can be intuitively defined as processes in which the only important information is the most recent one (the rest in redundant) and can be schematically defined as

$$P(X_{t_{n+1}} | X_{t_n}, \ldots, X_{t_1}) = P(X_{t_{n+1}} | X_{t_n})$$

whenever $t_1 \leq \cdots \leq t_n \leq t_{n+1}$.

From this point of view, the best information about the future of a Markov process is its current state. The ultimate consequence of this assumption is that when studying these processes through their equivalent Forward Kolmogorov (or Fokker-Planck) equation, the differential equation will present just one-time derivative.

The state space of a Markov process can be either discrete or continuous. The methods and ideas from both cases are quite similar. They both possess Forward and Backward Kolmogorov equations, infinitesimal generators...

For the discrete case, let's take an intuitive example of a practical situation. Imagine we have a factory which produces some item which lifetime obeys an exponential distribution of parameter λ. The intuition behind it is that if the item is still working after, let's say, one year, the probability of the item to stop working in the future is the same as if the item were newly produced, this is the memoryless property which defines the exponential distribution.

Imagine a particular customer buys one item at one time and, whenever the item stops functioning it is immediately replaced by a new one (because of the memoryless property any item working is a "new" one). We describe the number of items that have been broken at time t by the random variable N_t. Then, it can be proved, that N_t is characterized by

1. $N_0 = 0$. That is, we begin our counting at time $t = 0$, when the first item is bought.
2. for $0 \leq t_1 \leq \cdots \leq t_m$ with $m \geq 2$, $N_{t_{i+1}} - N_{t_i}$ are independent random variables
3. for any $0 \leq s < t$ the random variable $N_t - N_s$ (which is discrete-valued) obeys a Poisson distribution of parameter $\frac{(t-s)}{\lambda}$ that is

$$P(N_t - N_s = k) = \frac{(t-s)^k}{\lambda^k \cdot k!} e^{-\frac{t-s}{\lambda}}$$

For the continuous state space case $\mathbb{R}^n$, the main object of study is the $n-$dimensional Brownian motion W_t. The definition is quite similar to the one we have just given for the discrete case, but with one important difference

Definition 2.6 (Brownian motion)**.** *1.* $W_0 = 0$

2. for $0 \leq t_1 \leq \cdots \leq t_n$ *with* $n \geq 2$, $W_{t_{i+1}} - W_{t_i}$ *are independent random variables*

3. for any $0 \leq s < t$ *the random variable* $W_t - W_s$ *obeys a normal distribution* $\mathcal{N}(0, \sigma^2(t-s))$ *that is*

$$P(W_t - W_s \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{\sigma^2(t-s)}} \exp\left[-\frac{y^2}{\sigma^2(t-s)}\right] dy$$

Brownian motion is certainly the most studied stochastic process in the literature. Its history begins in the first years of the XX century, when Albert Einstein defined it in the context of statistical mechanics and later on, Perrin used it to proof the existence of atoms. Actually the first time in which this process appeared was in the context of financial mathematics some years before Einstein discovery. For a complete mathematical guide of Brownian motion, we recommend [11, 12].

One important concept that appears in probability, physics, and finance are stochastic differential equations. Which are differential equations in which a special stochastic process, white noise, makes the evolution stochastic. white noise can be rigorously defined and, at least intuitively, it is the 'time derivative' of Brownian motion, which can not be literally interpreted. A particular realization of Brownian motion is almost everywhere, not differentiable.

A stochastic differential equation can be written then as

$$\dot{x} = a(x(t), t) + b(x(t), t)\xi(t), \tag{2.1}$$

where ξ is white noise. The usual approach is nevertheless to write (2.1) as

$$x(t) = x(t_0) + \int_{t_0}^{t} a(x(s), s)\, ds + \int_{t_0}^{t} b(x(s), s)\, dW_s, \tag{2.2}$$

where the object $\int_{t_0}^{t} b(x(s), s)\, dW_s$ is a stochastic integral.

Under some conditions, these stochastic integrals can be defined in an 'a la Rieman', but with one important inconvenience: one should take care of the way the limit is taken. Without some additional prescription, the stochastic integral is not uniquely defined. The two more important stochastic integrals are the Itô integral and the Stratonovich integral.

A stochastic differential equation (SDE) is usually written as

$$dx = a(x, t)dt + b(x, t)dW_t, \tag{2.3}$$

when using the Itô integral and

$$dx = a(x, t)dt + b(x, t) \circ dW_t, \tag{2.4}$$

when using the Stratonovich integral.

Luckily for us, there is a way of going (under certain conditions) from a Stratonovich SDE to a Itô SDE. In particular, whenever the conditions are satisfied, equation (2.4) is equivalent to the Itô equation

$$dx = \left\{ a(x,t) + \frac{1}{2} b(x,t) \partial_x b(x,t) \right\} dt + b(x,t) \circ dW_t. \tag{2.5}$$

The key condition that differentiates Itô and Stratonovich SDE's is that, for the Itô equation (2.3) the following identity always holds

$$\frac{d}{dt} E[x] = E[a(x,t)]. \tag{2.6}$$

Additionally, whenever doing a change of variable in an Itô equation, one can not just use naively the chain rule. Lets take a function sufficiently smooth f and define the variable $y_t = f(x_t)$ where x_t follows (2.3), then

$$dy_t = df(x_t) = f'(x_t) dx_t + \frac{1}{2} f''(x_t) (dx_t)^2 \tag{2.7}$$

where one should use, formally, the rules

$$dt^2 = 0, \qquad dt \cdot dW_t = 0, \qquad (dW_t)^2 = dt. \tag{2.8}$$

The ultimate consequence of this behavior has to do with the central limit theorem.

For the Stratonovich counterpart (2.4), one can just use the chain rule as in the classical case.

Stochastic differential equations are related to partial differential equations, in particular, for the Itô process (2.3) its probability density $p(x, t\,|x_0, t_0)$ [3] satisfies the so called forward Kolmogorov (or Fokker-Planck) equation

$$\frac{\partial}{\partial t} p(x, t\,|x_0, t_0) = -\frac{\partial}{\partial x} \left\{ a(x,t) p(x, t\,|x_0, t_0) \right\} + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left\{ b(x,t) p(x, t\,|x_0, t_0) \right\} \tag{2.9}$$

with the initial condition $p(x, t_0\,|x_0, t_0) = \delta(x - x_0)$.

[3] A stochastic process is known once having all the probability distributions $p(x_1, t_1; \ldots; x_n, t_n)$ where n is arbitrary. For Markov processes, this probabilities can be extrated from the function $p(x, t\,|y, s)$.

2.3 Sheffer Sequences

From probability, we now jump to sequences of polynomials. We connect these two concepts in our work.

A polynomial sequence is a set of polynomials [4] $\mathcal{B} = \{p_n(x) \,|n \in \mathbb{N}\}$ of increasing degree $deg(p_n(x)) = n$. It is easy to see that, for an arbitrary polynomial $q(x)$ there are always coefficients $\{c_n\}_{n\in\mathbb{N}}$ with all but a finite number of them been zero, such that $q(x) = \sum_{n=0}^{\infty} c_n p_n(x)$. Furthermore, these coefficients are unique for each such polynomial $q(x)$.

The set of polynomials $\mathcal{P}$ is, in fact, a vector space which also possesses a commutative product, which is bilinear, and there is unity for that product. But we are going to forget about the product for a moment and think of it as just an infinite-dimensional vector space.

We can define the algebraic dual $\mathcal{P}^*$, which is the set of linear applications from $\mathcal{P}$ to $\mathbb{R}$. From umbral calculus [13], $\mathcal{P}^*$ is the set of formal power series.

We are nevertheless interested in linear applications from $\mathcal{P}$ to itself. Once we have chosen a specific basis $\mathcal{B}$, we could define one of such linear applications from its behavior on $\mathcal{B}$. One of such applications is the annihilation operator

$$Ap_n(x) = np_{n-1}(x), \tag{2.10}$$

which definition reminds to the derivative operator (actually coincide when $\mathcal{B}$ is just the set of monomials) and which we call annihilation for its relation with our future work. Another such operator is

$$Cp_n(x) = p_{n+1}, \tag{2.11}$$

which we will call creation operator, and that is it the analogous of $x \cdot ()$ in the context of monomials. One property that these operators fulfill is

$$[A, C] = AC - CA = I, \tag{2.12}$$

where I represents the identity operator. It is easy to see that for each base $\mathcal{B}$ there would be such a pair of operators.

When the annihilation operator is such that commutes with all translation operators $T_a q(x) = q(x+a)$,

[4] We will just consider polynomials over $\mathbb{R}$

that is, $[T_a, A] = 0$ then A is called a delta operator and the sequence $\mathcal{B}$ is called a Sheffer sequence. The condition of commuting with all translation operators is slightly more restrictive than the condition that it commutes with the derivative operator, and its motivation comes from umbral calculus.

Sheffer sequences are characterized by its generating function

$$\sum_{n=0}^{\infty} \frac{p_n(x)}{n!} t^n = B(t) \exp(A(t)x). \tag{2.13}$$

Sheffer sequences form a group under the umbral composition, a subgroup of Sheffer sequences are Appell sequences, defined by the annihilation operator

$$A = \frac{d}{dx}$$

.

2.4 Large Deviations and Hamiltonian Dynamics

Hamiltonian dynamics appear in our thesis, when solving some optimization problems, as it is intimately related to variational calculus. The reason why this formulation is connected with stochastic dynamics has its origins in the problem of mean first passage times, WKB theory, and large deviations [14].

Lets study the equation

$$dX_t = f(X_t)dt + \sqrt{\epsilon g(X_t)} dW_t, \tag{2.14}$$

where we will let $\epsilon > 0$ go to zero. For simplicity, lets take the case $f(x) = 0$ and $g(x) = 1$. The probability distribution of X_t follows the forward Kolmogorov equation

$$\frac{\partial P(x, t\,|0, 0\,)}{\partial t} = \frac{\epsilon}{2} \frac{\partial^2 P(x, t\,|0, 0\,)}{\partial x^2} \tag{2.15}$$

with initial condition

$$P(x, 0\,|0, 0\,) = \delta_0. \tag{2.16}$$

We know the solution of this equation is

$$P(x,t\,|0,0\,) = \frac{1}{\sqrt{2\pi t\epsilon}} \exp\left[-\frac{x^2}{2t\epsilon}\right]. \tag{2.17}$$

In the limit of $\epsilon \to 0$, it is easy to see explicitly the analytical form of the large deviation functional [14]

$$I(x) := \lim_{\epsilon\to 0} \epsilon \log\left(P(X_t \geq x)\right) = -\frac{x^2}{2t} \tag{2.18}$$

where $x > 0$. [5]

In the sense of asymptotic expansions [15],

$$P(X_t \geq x) \sim e^{-\frac{x^2}{2\epsilon t}}. \tag{2.19}$$

It is important to note that functionals of the type $\exp\left(\frac{\cdot}{\epsilon}\right)$ are not Taylor expansible when $\epsilon \to 0$, that's why we need asymptotic expansions/ large deviation theory.

More importantly, this kind of functionals appears when dealing with semi-classical limits in quantum mechanics or low-temperature limits (or even large time limits) in statistical mechanics.

Equation (2.19) can be seen as the "infinite"-dimensional version of the Laplace's method

$$\lim_{\epsilon\to 0} \int e^{\frac{f(x)}{\epsilon}} dx \sim e^{\frac{max_{\mathbb{R}} f(x)}{\epsilon}} \tag{2.20}$$

if, of course, certain conditions about f are fulfilled [15].

When taking the limit $\epsilon \to 0$ usually one proceed taking an WKB ansatz of the form

$$P(x,t\,|0,0\,) \sim e^{-\frac{S(x,t)}{\epsilon}}. \tag{2.21}$$

From which, using equation (2.15) we see

$$-\frac{\partial S}{\partial t} = \frac{1}{2}\left(\frac{\partial S}{\partial x}\right)^2 + \frac{\epsilon}{2}\frac{\partial^2 S}{\partial x^2} \tag{2.22}$$

[5] For the case $x < 0$ one should use $P(X_t \leq x)$.

and so, at least formally, when $\epsilon \to 0$ we obtain

$$-\frac{\partial S}{\partial t} = \frac{1}{2}\left(\frac{\partial S}{\partial x}\right)^2 \tag{2.23}$$

which is the Hamilton-Jacobi equation for the free particle Hamiltonian

$$H(x,p) = \frac{1}{2}p^2 \tag{2.24}$$

where $\partial_x S = p$. As we know from Hamilton-Jacobi theory, one can solve the PDE (2.23) using the Hamiltonian flow

$$\dot{x} = \frac{\partial H}{\partial p} = p \tag{2.25}$$

$$\dot{p} = -\frac{\partial H}{\partial x} = 0, \tag{2.26}$$

with the initial conditions $x(0) = 0$ and $x(t) = x_f$, and then, using the identity

$$S(x_f, t) = \int_0^t \left(p\dot{x} - H(x(s), p(s))\right) ds \tag{2.27}$$

we obtain

$$S(x_f, t) = \frac{x_f^2}{2t} \tag{2.28}$$

as we already knew.

Chapter 3

Doi-Peliti theory

The studies made on this thesis can all be traced back, some way or another, to Doi-Peliti Field theory, which original ideas came from quantum optics and quantum field theory and that was developed during the 70's and 80's [1, 16, 2, 17, 18].

Applying ideas coming both from applied mathematics and physics, one can say that there are, at least, three equivalent approaches to study stochastic diffusion processes [7]. Two of them are well known to probability theory, namely, the equivalence in the descriptions between the forward Kolmogorov (or Fokker-Planck) equation, a second-order PDE, and the Itô (or Langevin) equation, a stochastic differential equation. This relationship is well established, even though there are some important subtleties to take into consideration, like the dependence (in the SDE approach) on the Itô integral versus some other equally valid stochastic integral (Stratonovich for example) which we will study separately in chapter 5.

Path integrals are the third element believed to be an equivalent description for stochastic diffusion processes. Path Integrals can be seen as something "in-between" a PDE, and an SDE as their formal properties resemble both approaches in some way. The idea was originally developed by Feynman in the context of quantum mechanics, which is intimately related with stochastic processes [1]. Nevertheless, this formalism is full of objections and subtleties that make it difficult to justify from a totally rigorous point of view, though some results are well known for the context of large deviations [14], which is intimately related to path integrals formalism.

[1] Actually, in stochastic processes, one can use the Wiener measure to construct a rigorous measure in a functional space while the situation for quantum mechanics is not so clear.

Doi-Peliti field theory can be seen as a way of deriving these equivalences in the context of discrete-state stochastic processes, in particular, it works for a special class of master equations, which are the forward Kolmogorov equation for a Markov process in continuous time and with discrete, but possibly infinite, state-space [6, 19, 20]. There are other related approaches in the context of discrete-state stochastic processes, as the Martin-Siggia-Rose field theory, which can be seen as equivalent to Doi-Peliti theory and which we will use in chapter 7.

The formalism of Doi-Peliti field theory allows to work with reaction-diffusion models, in particular, some abstract chemical reactions, in a rather simple way, as in its inner structure captures the combinatoric essence of this kind of processes. An abstract chemical reaction is not limited only to the context of chemistry; it is actually a rather abstract formulation that can describe processes from biology, ecology, sociology and has been used across lots of different fields for modeling a wide range of situations, from sand-pile models to predator-prey, or even for epidemic models. See, for example, [21, 5, 22, 23, 24, 25].

Even though the use of Doi-Peliti theory is widespread throughout the physical literature, there are still some mathematical subtleties to it. One of the aims of the present chapter is to contribute from a mathematically rigorous point of view to some of those questions. Some more references about Doi-Peliti Field theory and its applications are [26, 27, 28, 29, 30, 31, 32]

3.1 Doi-Peliti Field Theory

Let's suppose that we want to model a situation in which there are several molecules of different chemical species which number varies by some basic interactions and such that the molecules are placed in the nodes of a lattice (discrete space) but in which each point of the lattice can contain, in principle, an arbitrary number of molecules. We are tracing the exact number of molecules in each point of the state space.

We assume that the system evolves as a Markov process, that is, the only useful information to predict the future is the present. Our system is governed by rules that include local interactions and diffusion through space (the lattice). Inside each point of the state-space, the local interactions (reaction) can lead to creation, annihilation, reproduction, etc. of molecules. For the diffusion part, we assume that a molecule can hop from a particular point of the state space to its neighbors. That is, we consider a

reaction-diffusion process inside the lattice.

The state-space that we consider (our lattice) is of the form $\mathbb{Z}^d$ where $d \leq 3$. For describing all the information regarding the number of molecules in each point of the lattice, we need, at each time value t a probability of the states $\left\{(n_1, \ldots, n_m)_j\right\}_{j\in\mathbb{Z}^d}$ where $(n_1, \ldots, n_m)_j \in \mathbb{N}^{m2}$ and m stands for the number of chemical species inside the lattice point $j \in \mathbb{Z}^d$. That is, we need to know $P(t, \left\{(n_1, \ldots, n_m)_j\right\})$ with the condition

$$\sum_{\left\{(n_1,\ldots,n_m)_j\right\}} P(t, \left\{(n_1, \ldots, n_m)_j\right\}) = 1 \ , \ \forall t \geq 0. \tag{3.1}$$

All our future studies are related with the simplified version $d = 0, m = 1$, where $d = 0$ stands for taking into consideration only transition of number of molecules inside a point of space with no diffusion of molecules, still some terms similar to 'diffusion' will appear because of the noise involved in the interactions, i.e., elliptic operators. That said, the methods we are going to develop here can be generalized to the rest of the cases (d, m). We are going to define the main objects of Doi-Peliti theory, creation and annihilation operators, for general (d, m) and later come back to our case.

To summarize all the information in just one object, we are going to define a basis of an infinite dimensional vector space $\left|\left\{(n_1, \ldots, n_m)_j\right\}\right\rangle$ [3].

Once we have defined a vector for each possible state of our system, we are going to define operators in that vector space that carry the information of creating and eliminating a molecule of one species from a particular point of our state space. These operators are two per each node and species in the lattice. And so the notation for these operators, called creation and annihilation operators, is $a_k^\dagger(j)$ and $a_k(j)$, where $k \in \{1, \ldots, m\}$ and $j \in \mathbb{Z}^d$. The action of these operators is

- $a_k^\dagger(j) \left|\{(n_1, \ldots, n_m)_l\}\right\rangle = \left|\left\{(n'_1, \ldots, n'_m)_{l'}\right\}\right\rangle$ where $n'_s = n_s + 1$ for $l' = j, s = k$ and the rest of parameters does not change.
- $a_k(j) \left|\{(n_1, \ldots, n_m)_l\}\right\rangle = (n_k)_j \left|\left\{(n'_1, \ldots, n'_m)_{l'}\right\}\right\rangle$ where $n'_s = n_s - 1$ for $l' = j, s = k$ and the rest of parameters does not change. This also include the case in which $n_j = 0$ producing the null vector.

With this formalism we have now a vector in an abstract vector space for each possible state, the creation and annihilation operators allow us to describe the dynamics of the system through vector

[2] Throughout the text, we will follow the convention $0 \in \mathbb{N}$

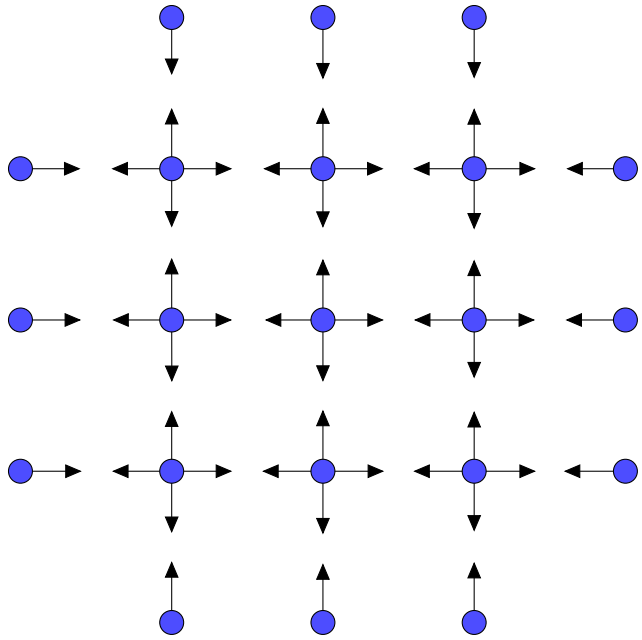
[3] The notation $|\cdot\rangle$ just stands for vector with index $\cdot$.

space notation and the coefficients of a vector of interest in this vector space on the basis of possible states is interpreted as the probability of that particular state.

Let's write the diffusion operator in this formalism. With that purpose in mind, let's first summarize in one object, the state vector, the whole information regarding our probabilities.

$$|\Psi(t)\rangle = \sum_{\left\{(n_1,\ldots,n_m)_j\right\}} P(t, \left\{(n_1,\ldots,n_m)_j\right\}) \left|\left\{(n_1,\ldots,n_m)_j\right\}\right\rangle . \tag{3.2}$$

Diffusion can be seen as a Poissonian waiting time for jumping onto one of the $2d$ neighbors that each point has. We take the rate at which these jumps happen as $\lambda \geq 0$, independent of the particular node, that is, we are modeling isotropic diffusion.



Following the pictured scheme, the probability variation along time can be written as

$$\frac{d}{dt}|\Psi(t)\rangle = -\frac{\lambda}{2d} \sum_{|j-j'|=1} \sum_{k=1}^{m} \left(a_k^\dagger(j') - a_k^\dagger(j)\right)\left(a_k(j') - a_k(j)\right)|\Psi(t)\rangle = H|\Psi(t)\rangle , \tag{3.3}$$

where $|j| = |(j_1,\ldots,j_d)| = \sum_{k=1}^{d}|j_k|$ represents the l^1 norm and H represents the linear evolution operator, also known as Hamiltonian in analogy with Quantum Field Theory. This formula is derived following the ideas from Master equations: for each possible transition there are two terms, one positive and another negative, that's why in our case, as between two neighbor nodes there are 2 possible

transitions we obtain in the sum four terms, two positives and two negatives. For example, in the case $d = m = 1$ we can write our states as $\left(\dots m_{j'}, m_j, \dots\right)$ with $j, j' \in \mathbb{Z}$, $j = j' + 1$ and $m_j, m_{j'} \in \mathbb{N}$, with the set of states represented as $\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}$. In this case the derivation goes as follows

$$
\begin{aligned}
\frac{d}{dt}\left|\Psi\left(t\right)\right\rangle &= \sum_{\{(\dots)\}} \frac{d}{dt} P(t, \left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}) \left|\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}\right\rangle = \\
&= \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} \left(m_j + 1\right) P(t, \dots m_{j'} - 1, m_j + 1, \dots) \left|\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}\right\rangle - \\
&\quad - \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} m_j P(t, \dots m_{j'}, m_j, \dots) \left|\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}\right\rangle + \\
&\quad + \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} \left(m_{j'} + 1\right) P(t, \dots m_{j'} + 1, m_j - 1, \dots) \left|\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}\right\rangle - \\
&\quad - \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} m_{j'} P(t, \dots m_{j'}, m_j, \dots) \left|\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}\right\rangle = \\
&= \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} m_j P(t, \dots m_{j'}, m_j, \dots) \left|\left\{\left(\dots m_{j'} + 1, m_j - 1, \dots\right)\right\}\right\rangle - \\
&\quad - \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} m_j P(t, \dots m_{j'}, m_j, \dots) \left|\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}\right\rangle + \\
&\quad + \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} m_{j'} P(t, \dots m_{j'}, m_j, \dots) \left|\left\{\left(\dots m_{j'} - 1, m_j + 1, \dots\right)\right\}\right\rangle - \\
&\quad - \frac{\lambda}{2} \sum_{|j-j'|=1} \sum_{m_j, m_{j'}} m_{j'} P(t, \dots m_{j'}, m_j, \dots) \left|\left\{\left(\dots m_{j'}, m_j, \dots\right)\right\}\right\rangle = \\
&= \frac{\lambda}{2} \sum_{|j-j'|=1} \left(a^\dagger(j')a(j) - a^\dagger(j)a(j) + a^\dagger(j)a(j') - a^\dagger(j')a(j')\right) \left|\Psi\left(t\right)\right\rangle = \\
&= -\frac{\lambda}{2} \sum_{|j-j'|=1} \left(a^\dagger(j') - a^\dagger(j)\right) \left(a(j') - a(j)\right) \left|\Psi\left(t\right)\right\rangle .
\end{aligned}
$$

Is in this type of reactions where the combinatoric nature of Doi-Peliti formalism realizes its full potential.

Remark. *One powerful aspect of Doi-Peliti formalism is that given a certain number of different reactions with diffusion, each reaction and the diffusion, can be treated separately. Meaning that the evolution Hamiltonian of the whole reaction-diffusion process can be calculated as the sum of each Hamiltonian.*

Up to this point, we have introduced calculations in a formal way. From now on, we are going to develop more the mathematics behind these concepts.

We will restrict ourselves to the case $d = 0$ and $m = 1$. In this type of processes, the only variable in the model is the number of particles $n \in \mathbb{N}$ of a particular species A.

Definition 3.1 (State Space)**.** *We define our state space as* $\mathbb{N}$*, representing the number of molecules of certain species in a 0-dimensional space. A basis of the vector space* $\mathcal{R}$ *generated by the state space is*

$$\mathcal{B} = \{|n\rangle : n \in \mathbb{N}\}.$$

And, finally, the abstract vector space $\mathcal{R}$ *is*

$$\mathcal{R} = \left\{ \sum_{n=0}^{\infty} a_n |n\rangle : n \in \mathbb{N}, \sum_{n=0}^{\infty} |a_n| < \infty \right\}.$$

That is, we define the abstract space as a formal series space isomorphic to the space of sequences l^1.

Sometimes, we will restrict the definition of our vector space even more depending on the particular case. Two important subspaces are:

Definition 3.2.

$$\mathcal{R}^{\infty} = \left\{ \sum_{n=0}^{\infty} a_n |n\rangle : n \in \mathbb{N}, \sum_{n=0}^{\infty} n^m |a_n| < \infty \forall m \in \mathbb{N} \right\}$$

$$\mathcal{R}^{0} = \left\{ \sum_{n=0}^{\infty} a_n |n\rangle : n \in \mathbb{N},\ s.t.\ \exists m \in \mathbb{N}\, a_j = 0\, \forall j \geq m \right\}$$

The next element to define is the state vector, which is the main object in Doi-Peliti field theory

Definition 3.3 (State vector)**.** *Let* $\{P_n\}_{n\in\mathbb{N}}$ *be a probability sequence, that is,* $P_n \geq 0\, \forall n$ *and* $\sum P_n = 1$*. We define the state vector* $|\Psi\rangle \in \mathcal{R}$ *as*

$$|\Psi\rangle := \sum_{n=0}^{\infty} P_n |n\rangle . \tag{3.4}$$

The creation and annihilation operators act in these vector spaces as

Definition 3.4 (Annihilation and creation operators)**.**

1. *The annihilation operator is defined to satisfy* $a|n\rangle := n|n-1\rangle$ *if* $n \geq 1$ *and* $a|0\rangle := 0$*.*

2. *The creation operator* [4] *is defined to satisfy* $a^{\dagger}|n\rangle := |n+1\rangle$*.*

[4] We call $a^{\dagger}$ to the creation operator just because of historical reasons, usually in the literature this means that there is a scalar product from which creation and annihilation operators are dual to each other, but we will not take this path here as we will see more carefully later on.

Theorem 3.1. *1. The annihilation operator is defined*

$$a : \mathcal{A} \subset \mathcal{R} \longrightarrow \mathcal{R}$$

from a dense subspace $\mathcal{A} \subset \mathcal{R}$ and it is a closed operator in $\mathcal{R}$

2. The creation operator is defined

$$a^\dagger : \mathcal{R} \longrightarrow \mathcal{R}$$

and it is a continuous operator.

Proof. 1. Defining $\mathcal{A} = \{\sum_{n=0}^{\infty} a_n |n\rangle : n \in \mathbb{N}, \sum_{n=0}^{\infty} n |a_n| < \infty\}$ it is clearly dense because the space $\mathcal{R}^0 \subset \mathcal{A} \subset \mathcal{R}$ and $\overline{\mathcal{R}^0} = \mathcal{R}$. Now, for proving that it is a closed operator, assume there is a sequence $|\psi_k\rangle \in \mathcal{R} \to |\psi\rangle$ with $|\psi_k\rangle \in \mathcal{A} \forall k \in \mathbb{N}$ such that $a|\psi_k\rangle \to |\psi_a\rangle$.

We need to prove that $|\psi\rangle \in \mathcal{A}$ and that $a|\psi_k\rangle = |\psi_a\rangle$. If $|\psi_k\rangle = \sum_n^{\infty} a_n^k |n\rangle$ and $|\psi\rangle = \sum_n^{\infty} a_n |n\rangle$ we can write

$$\sum_{n=0}^{M} n|a_n| = \sum_{n=0}^{M} n|a_n - a_n^k + a_n^k| \leq \sum_{n=0}^{M} n|a_n - a_n^k| + \sum_{n=0}^{M} n|a_n^k|$$

$\sum_{n=0}^{M} n|a_n^k| < \infty$ because $a|\psi_k\rangle \to |\psi_a\rangle$, now for $\sum_{n=0}^{M} n|a_n - a_n^k|$ if we fix M we can find a k big enough such that for all $0 \leq n \leq M$ we obtain $|a_n - a_n^k| < \frac{1}{M^3}$ having then, that for each M $\sum_{n=0}^{M} n|a_n - a_n^k| < \sum_{n=0}^{M} \frac{n}{M^3} = \frac{M(M+1)}{M^3} \to 0$.

This proofs that $|\psi\rangle \in \mathcal{A}$ and clearly follows that $a|\psi_k\rangle = |\psi_a\rangle$

2. $a^\dagger$ is clearly well defined in $\mathcal{R}$ and continuous as one can re-index

$$a^\dagger \sum_{n=0}^{\infty} a_n |n\rangle = \sum_{n=0}^{\infty} a_n |n+1\rangle = \sum_{n=1}^{\infty} a_{n-1} |n\rangle$$

□

Remark (Combinatorics and operators)**.** *Given n distinguisable particles, there are n possible ways of annihilate one of them, with the consequence of obtaining $n-1$ particles after that. One can clearly see then the combinatoric meaning of $a|n\rangle = n|n-1\rangle$.*

On the other hand, there is only one way to create a new particle $a^\dagger |n\rangle = |n\rangle$

We have been talking about the convenience of using Doi-Peliti field theory when studying abstract chemical reactions, but we haven't study yet any particular example.

As we have already said, we are going to study mainly systems of reactions of only one 'chemical' species in zero dimensions. Following the conventional notation, a general reaction of that type is

$$jA \xrightarrow{\lambda} lA. \tag{3.5}$$

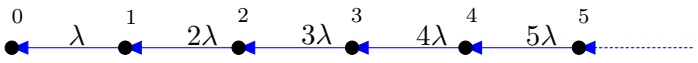
Where $j, l \in \mathbb{N}$.[5] This means that there is a local interaction in which j elements of a certain species, or particles as we are going to call from now on, can interact giving, as a result, l particles. The evolution of this system is Markovian, which means that the system evolves according to its current number of particles and not taking into consideration previous states.

The definition of λ is that it is the rate at which j particles combine in this system, that is, every $1/\lambda$ units of time, we expect to see at least one of those reactions, in average.

In particular, lets derive the equations that govern the following reaction:

$$A \xrightarrow{\lambda} \emptyset \tag{3.6}$$

This decaying process can be thought as a system of particles in which once reached a certain number n of particles, the following state of the system is $n-1$ particles with a rate of $n\lambda$ [6]



Calling $P(n,t)$ the probability of having n particles in the system at time t, and using the Markov property, we realize that the next state of the system follow the so-called Master equation

$$\frac{1}{\lambda}\frac{d}{dt}P(t,n) = (n+1)P(t,n+1) - nP(t,n). \tag{3.7}$$

Which meaning is that the state of n particles in an infinitesimal time from the present, is more likely if the present state has $n+1$ particles and less likely if the present state has n particles, and it is the only way the state of n particles in an infinitesimal time can be affected.

[5] Throughout this thesis this in particular implies the situation $j, l = 0$

[6] There are n possible particles that can decay, so it n times more 'faster' than when there is only one particle present.

As we have already seen when defining Doi-Peliti theory, we can actually define, through the state vector, a mathematical object containing the intuitive meaning of the state of the system at time t

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} P_n(t)\,|n\rangle\,, \tag{3.8}$$

where we use the notation $P_n(t) = P(t,n)$; the master equation translates into this formalism into

$$\frac{d}{dt}\,|\Psi(t)\rangle = \lambda(1-a^\dagger)a\,|\Psi(t)\rangle = H(a^\dagger,a)\,|\Psi(t)\rangle\,, \tag{3.9}$$

as one can easily see in this particular example just using the definitions we have made up to this point, and where H is the Hamiltonian or evolution operator. It should not be surprising by now the definition that we have made of creation and annihilation operators, they appear naturally in this context.

We can calculate the Hamiltonian of the reaction (3.5) in a general fashion.

Lemma 3.1. *The Doi-Peliti Master equation for a reaction of the type 3.5 is*

$$\frac{d}{dt}\,|\Psi(t)\rangle = H(a^\dagger,a)\,|\Psi(t)\rangle = \frac{\lambda}{j!}(a^{\dagger l}-a^{\dagger j})a^j\,|\Psi(t)\rangle\,, \tag{3.10}$$

where the previous differential equation should be defined in the dense subspace $\mathcal{R}^\infty \subset \mathcal{R}$.

Proof. The Master equation for the reaction (3.5) is well-defined for an interval $0 \le t < T$ due to Theorem 3.4 [33], and can be written as

$$\frac{1}{\lambda}\frac{dP_n(t)}{dt} = \binom{n-l+j}{j}P_{n-l+j}(t) - \binom{n}{j}P_n(t),$$

where $P_n(t)$ is the probability of the state $n \in \mathbb{N}$ at time $t > 0$. The state vector $|\Psi(t)\rangle$ is defined as

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} P_n(t)\,|n\rangle\,.$$

It is clear that

$$\sum_{n=0}^{\infty}\left\{\binom{n-l+j}{j}P_{n-l+j}(t) + \binom{n}{j}P_n(t)\right\}|n\rangle \in \mathcal{R}^\infty$$

in the subspace $|\Psi(t)\rangle \in \mathcal{R}^\infty$ which implies

$$\sum_{n=0}^{\infty} \left| \frac{d}{dt} P_n(t) \right| |n\rangle \in \mathcal{R}^\infty$$

which, by dominated convergence theorem, implies

$$\frac{d}{dt} |\Psi(t)\rangle \in \mathcal{R}^\infty$$

so, in the space $\mathcal{R}^\infty$ we can calculate

$$\frac{1}{\lambda}\frac{d}{dt}|\Psi(t)\rangle = \frac{1}{\lambda}\frac{d}{dt}\sum_{n=0}^{\infty} P_n |n\rangle = \sum_{n=0}^{\infty} \binom{n-l+j}{j} P_{n-l+j} |n\rangle - \sum_{n=0}^{\infty} \binom{n}{j} P_n |n\rangle =$$
$$= \sum_{n=0}^{\infty} \binom{n}{j} P_n |n+l-j\rangle - \sum_{n=0}^{\infty} \binom{n}{j} P_n |n\rangle = \frac{1}{j!}(a^{\dagger l} a^j - a^{\dagger j} a^j) |\Psi(t)\rangle = \frac{1}{j!}(a^{\dagger l} - a^{\dagger j}) a^j |\Psi(t)\rangle$$

where $a, a^\dagger$ can be interchanged with the sum in the subspace $\mathcal{R}^\infty$. □

Annihilation and creation operators do not commute, actually, their commutator is the identity operator

Lemma 3.2. *The commutator of the annihilation and creation operators is*

$$\left[a, a^\dagger\right] := aa^\dagger - a^\dagger a = 1. \tag{3.11}$$

in the set $\mathcal{A} \subset \mathcal{R}$.

Proof. For the base $\mathcal{B}$ it is clear just following the definitions

$$\left[a, a^\dagger\right] |n\rangle = aa^\dagger |n\rangle - a^\dagger a |n\rangle = (n+1) |n\rangle - n |n\rangle = |n\rangle .$$

For an element $|\Psi\rangle \in \mathcal{A}$, as we know both a and $a^\dagger$ are well defined in that subset and also that we can always write

$$|\Psi(t)\rangle = \lim_{M\to\infty} \sum_{n=0}^{M} \Psi_n |n\rangle$$

the result follows just taking into account that the operator $\left[a, a^\dagger\right]$ is closed and that for finite sums it acts as the identity operator. □

Remark (Normal order)**.** *We use the normal order convention, that is, all annihilation operators are to the right of all creation operators. This is not a problem as we know by* (3.11) *how to change the order and so, how to bring all annihilation operators to the right.*

From now on, when writing $H(a^\dagger, a)$ we are thinking in the normal representation of the Hamiltonian.

We now proceed to introduce a fundamental ingredient of probability theory, the expected value. Doi-Peliti formalism is often introduced in a very similar fashion as quantum mechanics, in particular, a scalar product is normally defined between state vectors[7]. We are not going to follow this path here, mainly because of two reasons:

1. Scalar products are very useful when working with an L^2 structure, that is, when working with amplitudes. However, in our case, as we are trying to make a theory of probability the natural structure is L^1. This problem can be "solved" defining a vector which scalar product with an arbitrary vector generates de expected value. So the following question arises: why to define a scalar product just for using it against a privileged vector?

2. In chapter 4, we introduce the connection between Doi-Peliti field theory and the theory of Sheffer sequences. In that connection, one can see that introducing a scalar product implies restrictions on the Sheffer basis.

Definition 3.5 (Expected value)**.** *We define the expected value operator w.r.t the basis $\mathcal{B}$ as the linear extension of*

$$\begin{aligned} E: \quad \mathcal{R} &\longrightarrow \mathbb{R} \\ |n\rangle &\longrightarrow 1 \end{aligned}$$

$\forall n \in \mathbb{N}$.

Lemma 3.3. *The expected value E is a continuous linear operator in $\mathcal{R}$ with $\|E\| = 1$.*

Proof. E is well defined for each

$$|\Psi\rangle = \sum_{n=0}^{\infty} \phi_n |n\rangle \in \mathcal{R}$$

defining

$$E\left[|\Psi\rangle\right] = \sum_{n=0}^{\infty} \phi_n,$$

[7]That is the reason for the notation $|n\rangle$, as the scalar product would be written as $\langle m|n\rangle$.

in particular $\sum_{n=0}^{\infty}\phi_n < \infty$ because

$$\| \, |\Psi\rangle \, \| = \sum_{n=0}^{\infty} |\phi_n| < \infty$$

Now $|E\left[|\Psi\rangle\right]| = |\sum_{n=0}^{\infty}\phi_n| \leq \sum_{n=0}^{\infty}|\phi_n| < \infty$ which implies $\|E\| \leq 1$, for proving the equality we have just to pick any element $|n\rangle \in \mathcal{B}$, as $\| \, |n\rangle \, \| = 1$ and $[E] \, |n\rangle = 1$ □

This operator can be of great help when dealing with creation and annihilation operators as the following Lemma proofs

Lemma 3.4. *Let*

$$|\Psi\rangle = \sum_{n=0}^{\infty} P_n \, |n\rangle$$

be a state vector in $\mathcal{R}$. The following identities are fulfilled

1. $E\left[a^{\dagger k} \, |\Psi\rangle\right] = E\left[|\Psi\rangle\right] = 1 \ \forall k \in \mathbb{N}$
2. *If $a^k \, |\Psi\rangle \in \mathcal{R}$ then $E\left[a^k \, |\Psi\rangle\right]$ = kth factorial moment of the probability distribution.*

The proof of this lemma is trivial, and consist in writing down the definitions in each case.

A more interesting property is the property of conservation of probability

Lemma 3.5. *Probability is conserved under the flux*

$$\frac{d}{dt} \, |\Psi(t)\rangle = H(a^{\dagger}, a) \, |\Psi(t)\rangle \tag{3.12}$$

if and only if $H(1, a) = 0$, where H is normal ordered and whenever (3.12) is such that $|\Psi(t)\rangle \in \mathcal{R}^{\infty}$.

Proof. As H is normal ordered, it can be written in a general fashion as $H(a^{\dagger}, a) = \sum_{m,n} c_{mn} a^{\dagger m} a^n$ where $c_{mn} = 0$ for sufficiently large n, m. Taking expected values in the flux equation

$$0 = \frac{d}{dt} E\left[|\Psi(t)\rangle\right] = E\left[H(a^{\dagger}, a) \, |\Psi(t)\rangle\right] = \sum_{m,n} c_{mn} E\left[a^{\dagger m} a^n \, |\Psi(t)\rangle\right]$$

where the first equality is just conservation of probability. We can apply lemma 3.4 and eliminate all the creation operators

$$\sum_n \sum_m c_{mn} E\left[a^n \, |\Psi\rangle\right] = 0$$

for all $|\Psi\rangle$. In particular, taking vectors of the form $|j\rangle$, from $j = 0$ and increasing the order we start to obtain this chain of equalities:

- For $j = 0$ we obtain $\sum_m c_{m0} |0\rangle = 0$ which implies $\sum_m c_{m0} = 0$
- For $j = 1$ we obtain $\sum_m c_{m0} |1\rangle + \sum_m c_{m1} |0\rangle = 0$ which implies $\sum_m c_{m0} = \sum_m c_{m1} = 0$

Following this argument, and because there is a finite number of terms, we clearly obtain $\sum_m c_{mn} = 0$ $\forall n \in \mathbb{N}$ which is the same as saying

$$H(1, a) = 0.$$

Note that this proves both directions of our Lemma. □

We have defined the annihilation and creation operators from relations involving our basis $\mathcal{B}$. One could ask if, once defined, the operators themselves also are uniquely associated with the basis.

Lemma 3.6. *Lets take another basis for $\mathcal{R}$ different from $\mathcal{B}$ and call it $\mathcal{B}_0 = \{|n\rangle_0\}$. Then the relations*

$$a |n\rangle_0 = n |n-1\rangle_0$$
$$a^\dagger |n\rangle_0 = |n+1\rangle_0$$

are fulfilled if and only if there is a constant $K \neq 0$ with $\mathcal{B}_0 = K\mathcal{B}$.

Proof. The two basis are related through some structure constants

$$|n\rangle_0 = \sum_m a_{nm} |m\rangle$$

Imposing the condition about creation and annihilation operators we obtain the following identities between coefficients

$$a_{n+1,m} = a_{n,m-1}$$
$$(m+1)a_{n,m+1} = n a_{n-1,m}$$

which implies $(m+1)a_{n-1,m} = n a_{n-1,m}$ $\forall n, m$. From there we deduce $a_{n,m} = K_{n,m}\delta_{n,m}$ where $\delta_{m,n}$ represents the Kronecker delta. Using the first relation, we can stablish $K_{n,n} \equiv K$ so $a_{n,m} = K\delta_{n,m}$ where K is an arbitrary constant. Of course, this is the same as to say $|n\rangle_0 = K |n\rangle$ $\forall n \in \mathbb{N}$. □

3.2 Coherent State Representation

Following the analogy of Doi-Peliti theory with quantum field theory, we can study a special family of vectors in $\mathcal{R}$ called coherent vectors. The motivation behind coherent vectors in the context of quantum mechanics, is that the equations describing their evolution remind to classical mechanics. In principle, we have not a 'classical' analogy of our Doi-Peliti theory, but it is a good guess to follow the analogy and say that classical behavior in a probabilistic setting would have to do with the central limit theorem and thermodynamic limits.

Another useful property of coherent vectors is one can define a representation of every other vector, the coherent representation that could potentially give useful insights.

Definition 3.6 (Coherent vectors)**.** *In the context of Doi-Peliti field theory a coherent vector is a vector $|\phi\rangle \in \mathcal{R}$ that is an eigenvector of the annihilation operator a, that is,*

$$a\,|\phi\rangle = \phi\,|\phi\rangle\,, \tag{3.13}$$

[8] *where $\phi \in \mathbb{R}$.*

Lemma 3.7. *Let $|\phi\rangle \in \mathcal{R}$ be a coherent vector and $\phi \in \mathbb{R}$, then $|\phi\rangle$ can be written in our basis as*

$$|\phi\rangle = \alpha \sum_{n=0}^{\infty} \frac{\phi^n}{n!}\,|n\rangle \tag{3.14}$$

where $\alpha \in \mathbb{R}$ is an arbitrary constant.

Proof. We know the coherent state is a vector of our vector space $\mathcal{R}$, so it can be expressed as

$$|\phi\rangle = \sum_{n=0}^{\infty} a_n\,|n\rangle$$

as this vector has to fulfill (3.13); this implies the following relations

$$\phi a_n = (n+1)a_{n+1}$$

and if we call $a_0 = \alpha$ then $a_n = \alpha\dfrac{\phi^n}{n!}$. □

[8] This notation could induce to confusion as the same symbol can represent either a element of the basis $\mathcal{B}$ or a coherent state, but it is not difficult to infer from the context which object we referring to.

We are interested in a special subset of coherent vectors, the coherent states that are a representation of a probability vector. This determine the otherwise arbitrary coefficient α.

Definition 3.7 (Coherent states)**.** *For any $\phi \in [0, \infty)$, we define the coherent state*

$$|\phi\rangle := e^{-\phi} \sum_{n=0}^{\infty} \frac{\phi^n}{n!} |n\rangle .$$

That represents a Poisson distribution with parameter ϕ.

Remark. *When writing $|\phi\rangle$ we refer to the coherent state, which is uniquely defined.*

Lemma 3.8. *The coherent states fulfill the properties*

$$\begin{aligned} a |\phi\rangle &= \phi |\phi\rangle , \\ a^\dagger |\phi\rangle &= \left(1 + \frac{\partial}{\partial \phi}\right) |\phi\rangle . \end{aligned}$$

Proof. The first property is just the definition of a coherent state.

For the second compute

$$\begin{aligned} a^\dagger |\phi\rangle &= e^{-\phi} \sum_{n=0}^{\infty} \frac{\phi^n}{n!} a^\dagger |n\rangle = e^{-\phi} \sum_{n=0}^{\infty} \frac{\phi^n}{n!} |n+1\rangle = e^{-\phi} \frac{\partial}{\partial \phi} \sum_{n=0}^{\infty} \frac{\phi^{n+1}}{(n+1)!} |n+1\rangle = \\ &= e^{-\phi} \frac{\partial}{\partial \phi} \sum_{n=1}^{\infty} \frac{\phi^n}{n!} |n\rangle = e^{-\phi} \frac{\partial}{\partial \phi} \sum_{n=0}^{\infty} \frac{\phi^n}{n!} |n\rangle = \\ &= \left(1 + \frac{\partial}{\partial \phi}\right) \left(e^{-\phi} \sum_{n=0}^{\infty} \frac{\phi^n}{n!} |n\rangle \right) = \left(1 + \frac{\partial}{\partial \phi}\right) |\phi\rangle . \end{aligned}$$

□

There is a relation between coherent states and "macroscopic states", that is, when the thermodynamic limit follow

$$E\left[|\Psi\rangle\right] >> 1$$

where we expect, by the central limit theorem, that the probabilistic quantities converge in the macroscopic scale to a deterministic (Dirac delta) distribution.

When the expected number of particles is big enough, we expect, in this regime, that moments behave according to the following relations

$$E\left[N^k\right] \approx E\left[N\right]^k , \tag{3.15}$$

where N represents a random variable in our discrete state space. Formally this relation implies[9]

$$E\left[N(N-1)\cdots(N-k+1)\right] = E\left[\binom{N}{N-k}\right] \approx E\left[N\right]^k. \tag{3.16}$$

Because of the properties of the annihilation operator Lemma (3.4)

$$E\left[a^k\left|\phi\right\rangle\right] = \phi^k, \tag{3.17}$$

which is just the kth factorial moment of the Poisson distribution.

Actually, this property of coherent states characterize them:

Theorem 3.2. *If there is a constant $\phi \geq 0$ and a state vector $|\Psi\rangle \in \mathcal{R}$ such that $|\Psi\rangle = \sum_{n=0}^{\infty} \Psi_n |n\rangle$ with $\Psi_n \geq 0$, and fulfilling the identities*

$$E\left[a^k\left|\Psi\right\rangle\right] = \phi^k, \tag{3.18}$$

for all $k \in \mathbb{N}$. Then $|\Psi\rangle = |\phi\rangle$.

Proof. As $|\Psi\rangle \in \mathcal{R}$ we can write $|\Psi\rangle = \sum_{n=0}^{\infty} \Psi_n |n\rangle$.

From

$$E\left[a^k\left|\Psi\right\rangle\right] = \phi^k \ \forall k \geq 0$$

it follows

$$E\left[a^k\left(\frac{a}{\phi}\left|\Psi\right\rangle\right)\right] = \phi^k, \ \forall k \geq 0 \tag{3.19}$$

or

$$E\left[a^k\left(\frac{a}{\phi} - 1\right)\left|\Psi\right\rangle\right] = 0, \ \forall k \geq 0 \tag{3.20}$$

It is easy to see that $\left(\frac{a}{\phi} - 1\right)|\Psi\rangle \in \mathcal{R}$ due to the identity

$$\left(\frac{a}{\phi} - 1\right)|\Psi\rangle = \sum_{n=0}^{\infty} \left(\Psi_{n+1}(n+1)\phi^{-1} - \Psi_n\right)|n\rangle$$

[9] Of course, this is just formal and we only expect to obtain this result in the cases in which the behavior for large expected value behaves as in the thermodynamic regime. We nevertheless expect that the chemical reactions we will study with this formalism will have a similar behavior.

If we call $c_n = \Psi_{n+1}(n+1)\phi^{-1} - \Psi_n$ these coefficients satisfy

$$\sum_{n=0}^{\infty} n(n-1)\dots(n-k+1)c_n = 0,\ \forall k \tag{3.21}$$

as

$$E\left[a^k\left(\sum_{n=0}^{\infty} c_n |n\rangle\right)\right] = 0,\ \forall k \geq 0.$$

From where we can extract

$$\sum_{n=0}^{\infty} n^k c_n = 0,\ \forall k \tag{3.22}$$

and also

$$\sum_{n=0}^{\infty} f(n)c_n = 0 \tag{3.23}$$

for any analytical function in $\mathbb{R}$.

In particular, we can choose $f(n)$ as a Gaussian function centered in any natural number and with variance as small as wanted, from which we derive $c_n = 0\ \forall n$, implying

$$a|\Psi\rangle = \phi|\Psi\rangle \tag{3.24}$$

which is the definition of the coherent state $|\phi\rangle$. □

Definition 3.8 (Coherent representation)**.** *Given a function $\Psi(\phi) \in L^1(0,\infty)$ and a state vector $|\Psi\rangle \in \mathcal{R}$, with $|\Psi\rangle = \sum_{n=0}^{\infty} \Psi_n |n\rangle$, such that*

$$\Psi_n = \frac{1}{n!}\int_0^{\infty} \phi^n e^{-\phi}\,\Psi(\phi)\,d\phi, \tag{3.25}$$

for all $k \in \mathbb{N}$ we will say that the function $\Psi(\phi)$ is the coherent representation of the state vector $|\Psi\rangle \in \mathcal{R}$.

This definition may seem a bit out of place, but just thinking formally for a minute, we could interpret equation (3.25) as deriving from

$$|\Psi\rangle = \int_0^{\infty} \Psi(\phi)\,|\phi\rangle\,d\phi. \tag{3.26}$$

Theorem 3.3. *If $\Psi(\phi) \in L^1(0,\infty)$ is the coherent representation of $|\Psi\rangle \in \mathcal{R}$ then*

$$E\left[|\Psi\rangle\right] = \int_0^{\infty} \Psi(\phi)\,d\phi. \tag{3.27}$$

Proof. First of all

$$E\left[|\Psi\rangle\right] = \sum_{n=0}^{\infty} \Psi_n = \lim_{N\to\infty} S_N$$

with $S_N = \sum_{n=0}^{N} \Psi_n$, also $S_N = \int_0^\infty I_N(\phi)\, d\phi$ with

$$I_N(\phi) = \Psi(\phi) e^{-\phi} \sum_{n=0}^{N} \frac{\phi^n}{n!}$$

Now

$$|I_N(\phi)| = |\Psi(\phi)| e^{-\phi} \sum_{n=0}^{N} \frac{\phi^n}{n!} \leq |\Psi(\phi)| e^{-\phi} \sum_{n=0}^{\infty} \frac{\phi^n}{n!} = |\Psi(\phi)| \in L^1(0,\infty)$$

As $\lim_N I_N(\phi) = \Psi(\phi)$ point-wise, the dominated convergence theorem implies the result. □

$$E\left[a^k |\Psi\rangle\right] = \int_0^\infty \Psi(\phi, t)\phi^k d\phi \tag{3.28}$$

and taking annihilation operators in both sides of the equation. Also, by linearity, we should obtain

$$P_n = \frac{1}{n!} \int_0^\infty \phi^n\, e^{-\phi}\, \Psi(\phi)\, d\phi. \tag{3.29}$$

And we would like that the mapping from $L^1(0,\infty)$ to l^1 the space of summable sequences behave in a controllable manner. It turns out that things will not be so simple

Theorem 3.4. *The operator T defined as*

$$\begin{aligned} T : L^1_+(0,\infty) &\longrightarrow \ell^1 \\ \Psi(\phi) &\longmapsto T[\Psi(\phi)] := \left\{ \frac{1}{n!} \int_0^\infty \phi^n\, e^{-\phi}\, \Psi(\phi)\, d\phi \right\}_{n=0}^{\infty}, \end{aligned}$$

where $L^1_+(0,\infty) = \{\Theta \in L^1(0,\infty) : \Theta(\cdot) \geq 0\}$. is well-defined, linear, continuous, isometric and thus injective, but not surjective.

Proof. It is clear that this operator is linear if well-defined. To see it is well-defined and continuous

compute

$$\begin{aligned}
\|T(\Psi)\|_{\ell^1} &= \sum_{n=0}^{\infty} \frac{1}{n!} \left| \int_0^{\infty} \phi^n \, e^{-\phi} \, \Psi(\phi) \, d\phi \right| \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^{\infty} \phi^n \, e^{-\phi} \, |\Psi(\phi)| \, d\phi \\
&= \int_0^{\infty} \sum_{n=0}^{\infty} \frac{1}{n!} \, \phi^n \, e^{-\phi} \, |\Psi(\phi)| \, d\phi \\
&= \int_0^{\infty} |\Psi(\phi)| \, d\phi \\
&= \|\Psi\|_{L^1} < \infty,
\end{aligned}$$

where we have used the monotone convergence theorem in order to commute the integral and the sum. Note this string of inequalities also implies T is isometric, and thus injective.

It reminds to show that T is not surjective; of course if we are able to show that there is no L^1_+ function which image is $e_0 = (1, 0, 0, 0....)$ $(\equiv |0\rangle)$ then we are done. Lets proceed by contradiction: suppose that there exists such a function Ψ_0 with $T(\Psi_0) = e_0$, then

$$\frac{1}{n!} \int_0^{\infty} \phi^n \, e^{-\phi} \, \Psi_0(\phi) \, d\phi = \delta_{n0},$$

or equivalently

$$\int_0^{\infty} \phi^n \, e^{-\phi} \, \Psi_0(\phi) \, d\phi = \delta_{n0}.$$

Therefore for any polynomial $P(\phi)$

$$\int_0^{\infty} P(\phi) \, e^{-\phi} \, \Psi_0(\phi) \, d\phi = P(0).$$

Now assume $f(\phi) \in C_c(\mathbb{R}_+)$ and compute

$$\begin{aligned}
\int_0^\infty f(\phi)\, e^{-\phi}\, \Psi_0(\phi)\, d\phi &= \int_0^L f(\phi)\, e^{-\phi}\, \Psi_0(\phi)\, d\phi \\
&= \int_0^L P(\phi)\, e^{-\phi}\, \Psi_0(\phi)\, d\phi + \int_0^L [f(\phi) - P(\phi)]\, e^{-\phi}\, \Psi_0(\phi)\, d\phi \\
&\leq \int_0^L |P(\phi)|\, e^{-\phi}\, |\Psi_0(\phi)|\, d\phi + \int_0^L |f(\phi) - P(\phi)|\, e^{-\phi}\, |\Psi_0(\phi)|\, d\phi \\
&\leq \int_0^\infty |P(\phi)|\, e^{-\phi}\, |\Psi_0(\phi)|\, d\phi + \int_0^L |f(\phi) - P(\phi)|\, e^{-\phi}\, |\Psi_0(\phi)|\, d\phi \\
&\leq |P(0)| + \|f(\phi) - P(\phi)\|_{L^\infty(0,L)} \|\Psi_0(\phi)\|_{L^1(0,\infty)} \\
&= |P(0) - f(0)| + \|f(\phi) - P(\phi)\|_{L^\infty(0,L)} \|\Psi_0(\phi)\|_{L^1(0,\infty)},
\end{aligned}$$

where $L > 0$ is large enough so $[0, L]$ contains the support of $f(\phi)$ and we have used that $f(0) = 0$. Now the Weierstrass approximation theorem assures us that we can choose a polynomial $P(\phi)$ such that

$$\int_0^\infty f(\phi)\, e^{-\phi}\, \Psi_0(\phi)\, d\phi \leq \epsilon \quad \forall\, \epsilon > 0.$$

An analogous argument yields the reversed inequality, so we conclude

$$\int_0^\infty f(\phi)\, e^{-\phi}\, \Psi_0(\phi)\, d\phi = 0,$$

and this equality holds for any $f(\phi) \in C_c(\mathbb{R}_+)$. This clearly implies that $\Psi_0(\phi) = 0$ a.e., a contradiction.

□

Remark. *$L^1_+(0,\infty)$ (neither $L^1(0,\infty)$) is not sufficient to describe the space ℓ^1 via representation* (3.29)*, for that, we should consider $\Psi(\phi)$ to be an ultra distribution, that is, we would need an arbitrary number of derivatives of Dirac's delta functions to construct the full representation. From this observation, we derive that the natural space of definition of the PDEs is an ultra distribution space.*

3.3 Analytical Representation and Examples

One common approach in Doi-Peliti field theory is the use of the analytical representation

$$\mathcal{R} \equiv C^\omega\,(-1, 1)\,, \tag{3.30}$$

which is also known as the generating function approach. In this approach, we make use of the following representation

$$\begin{aligned} |n\rangle &\longleftrightarrow x^n, \\ a &\longleftrightarrow \frac{d}{dx}(\cdot), \\ a^\dagger &\longleftrightarrow x \cdot (\cdot). \end{aligned} \tag{3.31}$$

that is,

$$|\Psi\rangle = \sum_{n=0}^{\infty} \Psi_n |n\rangle \equiv \sum_{n=0}^{\infty} \Psi_n x^n := G(x) \quad \text{where } x \in (-1,1) \tag{3.32}$$

In particular, if $\Psi_n \geq 0$ as in the case of probabilities states that we are most interested in, then, $G(x) \in C^\omega(-1,1) \cap C[-1,1]$. In other words, the analytic representation of probabilistic states will be an analytic function in the interval $(-1,1)$ and continuous in the closed interval $[-1,1]$.

Lemma 3.9. *Let $\{P_n\}_n$ be a probability sequence with N moments finite, then the generating function*

$$G(x) = \sum_{n=0}^{\infty} P_n x^n \tag{3.33}$$

is derivable N times in $x = 1$ with

$$\frac{\partial^N}{\partial x^N} G(1) = \sum_{n=0}^{\infty} n \cdot (n-1) \cdots (n-N+1) P_n < \infty, \tag{3.34}$$

which is the N-factorial moment of $\{P_n\}_n$

Proof. The N-factorial moment is finite if and only if the N-moment is finite (which implies that also smaller order moments are finite).

The rest of the proof can be found in classical books like Chapter XI Theorem 2 [34]. □

That is, for distributions with all moments finite the generating function

$$G(x) \in C^\omega(-1,1) \cap C^\infty[-1,1].$$

Lemma 3.10 (Generating function PDE)**.** *Giving the general reaction*

$$jA \xrightarrow{\lambda} lA. \tag{3.35}$$

The generating function $G(x,t)$ that represents the probability distribution solving the Markov process follows the PDE

$$\frac{\partial G}{\partial t} = \frac{\lambda}{j!}(x^l - x^j)\frac{\partial^j G}{\partial x^j}. \tag{3.36}$$

The next object to study is the generating function representation of coherent states

Lemma 3.11. *The coherent state $|\phi\rangle$ has the generating function representation*

$$G_\phi(x) = e^{\phi(x-1)}. \tag{3.37}$$

Proof.

$$G_\phi(x) = e^{-\phi}\sum_{n=0}^{\infty}\frac{\phi^n}{n!}x^n = e^{\phi(x-1)} \tag{3.38}$$

□

The coherent state representation of a coherent vector can be written in the generating function approach as

Theorem 3.5. *Given a state vector $|\Psi\rangle$ that has a coherent representation $\Psi(\phi) \in L^1(0,\infty)$ the generating function*

$$G_\Psi(x) := \sum_{n=0}^{\infty}\Psi_n x^n \tag{3.39}$$

can be written as

$$G_\Psi(x) = \int_0^\infty e^{\phi(x-1)}\Psi(\phi)\,d\phi. \tag{3.40}$$

Proof. We know that

$$\Psi_n = \frac{1}{n!}\int_0^\infty \phi^n e^{-\phi}\Psi(\phi)\,d\phi \tag{3.41}$$

Following a similar argument to the proof of Theorem (3.3), we can prove (3.40). □

Remark. *We use relation* (3.40) *as the definition of Ψ being the coherent transformation of the function $G_\Psi(x)$.*

Lemma 3.12. *Let* $\Psi(\phi) \in L^1(0,\infty) \cap C^\infty[0,\infty)$ *be the coherent transform of* $G_\Psi(x) \in C^\omega(-1,1) \cap C^\infty[-1,1]$ *and such that* $\phi\Psi(\phi), \partial_\phi\Psi(\phi) \in L^1(0,\infty)$ *then*

1. $\phi\Psi(\phi)$ *is the coherent transform of* $\dfrac{\partial}{\partial x}G_\Psi(x)$
2. $\left(1-\dfrac{\partial}{\partial\phi}\right)\Psi(\phi)+\delta_0\Psi(0)$ *is the coherent transform of* $xG_\Psi(x)$ [10]

Proof. We know that G_Ψ and Ψ follow the relation (3.40)

1.
$$\frac{\partial}{\partial x}G_\Psi = \int_0^\infty \frac{\partial}{\partial x}e^{\phi(x-1)}\Psi(\phi)\,d\phi = \int_0^\infty e^{\phi(x-1)}\left(\phi\Psi(\phi)\right)\,d\phi \tag{3.42}$$

2.
$$xG_\Psi = \int_0^\infty xe^{\phi(x-1)}\Psi(\phi)\,d\phi = \int_0^\infty (x-1)e^{\phi(x-1)}\Psi(\phi)+e^{\phi(x-1)}\Psi(\phi)\,d\phi =$$
$$= \int_0^\infty \frac{\partial}{\partial\phi}e^{\phi(x-1)}\Psi(\phi)+e^{\phi(x-1)}\Psi(\phi)\,d\phi = -\Psi(0)+\int_0^\infty e^{\phi(x-1)}\left(1-\frac{\partial}{\partial\phi}\right)\Psi(\phi)\,d\phi$$

□

Corollary 3.1 (Coherent PDE)**.** *Given the process*

$$jA \xrightarrow{\lambda} lA, \tag{3.43}$$

with $l<j$*. The coherent representation* $\Psi(\phi,t)$ *of the Doi-Peliti solution* $|\Psi(t)\rangle$ *of* (3.43)*, if it exists and have a coherent representation at all times with* $\Psi(\phi,t) \in L^1(0,\infty)\cap C^\infty[0,\infty)$*, fulfill the PDE*

$$\frac{\partial\Psi}{\partial t} = \frac{\lambda}{j!}\left\{\left(1-\frac{\partial}{\partial\phi}\right)^l-\left(1-\frac{\partial}{\partial\phi}\right)^j\right\}\left(\phi^j\Psi\right). \tag{3.44}$$

Proof. Using Lemma (3.12) join with the form of the Generating function PDE (3.36) and the condition $l<j$ makes that all Dirac's delta evaluate to zero. □

Remark (PDE order)**.** *Given the reaction* (3.35) *with* $0 \le l < j$*, both the generating function and the coherent transform follows a linear PDE of order* j *.*

[10] δ_0 is the Dirac's delta in $\phi = 0$

In particular, it is clear from the previous remark that for obtaining a second order PDE, for a reaction of the type (3.35), there are just two possibilities

$$\begin{aligned} A + A \xrightarrow{\lambda} \varnothing \quad &\rightarrow \quad \frac{1}{\lambda}\frac{\partial \Psi}{\partial t} = \frac{\partial}{\partial \phi}(\phi^2 \Psi) - \frac{1}{2}\frac{\partial^2}{\partial \phi^2}(\phi^2 \Psi) \\ A + A \xrightarrow{\lambda} A \quad &\rightarrow \quad \frac{1}{\lambda}\frac{\partial \Psi}{\partial t} = \frac{1}{2}\frac{\partial}{\partial \phi}(\phi^2 \Psi) - \frac{1}{2}\frac{\partial^2}{\partial \phi^2}(\phi^2 \Psi) \end{aligned}$$

Where they are actually related by the change of variables $\phi \rightarrow \frac{1}{2}\phi$, so we will study only the first one from now on.

The formal Itô equation that this equation follow is

$$A + A \xrightarrow{\lambda} \varnothing \quad \Longrightarrow \quad d\phi = -\phi^2 dt + i\phi dW, \tag{3.45}$$

which has imaginary noise, and so it in particular makes that the PDE has to be posed in some non-classical space, and makes the relation between the PDE and the SDE only formal for the moment. We will study more about this representation of a complex SDE in chapter 6.

3.3.1 Back to the Decaying Process

We come back once more to our first example of the decaying reaction (3.6), also known as pure death reaction

$$A \xrightarrow{\lambda} \varnothing.$$

Which generating function follows the PDE

$$\frac{\partial G}{\partial t} = \lambda(1 - x)\frac{\partial G}{\partial x}. \tag{3.46}$$

This PDE can be solved, given the initial condition $G_0(x)$, as

$$G(x, t) = G_0\left(1 + (x - 1)e^{-\lambda t}\right).$$

The coherent representation Ψ solve

$$\frac{\partial \Psi}{\partial t} = \lambda \frac{\partial}{\partial \phi}(\phi \Psi), \tag{3.47}$$

which yields the deterministic SDE $d\phi = -\lambda\phi dt$. Its solution is $\Psi(\phi, t) = e^{\lambda t}\Psi_0(\phi e^{\lambda t})$, and thus

$$G(x,t) = \int_0^\infty e^{\lambda t}\Psi_0(\phi e^{\lambda t})e^{\phi(x-1)}d\phi. \tag{3.48}$$

3.3.2 Creation and Destruction: The Squared Bessel Process

Consider now the set of reactions

$$\begin{aligned} A &\xrightarrow{\alpha} \emptyset, \\ \emptyset &\xrightarrow{\beta} A, \\ A &\xrightarrow{\gamma} A + A, \end{aligned}$$

which can be described via the forward Kolmogorov equation

$$\frac{dP_n}{dt} = \gamma[(n-1)P_{n-1} - nP_n] + \beta(P_{n-1} - P_n) + \alpha[(n+1)P_{n+1} - nP_n]. \tag{3.49}$$

For the sake of analytical tractability we restrict ourselves to the case $\alpha = \gamma = \beta$; then we find the equation

$$\alpha^{-1}\frac{\partial G}{\partial t} = (x-1)G + (x-1)^2\frac{\partial G}{\partial x} = (x-1)\frac{\partial}{\partial x}\left((x-1)G\right), \tag{3.50}$$

to be solved for the generating function G. Its solution reads

$$G(x,t) = \frac{1}{1-\alpha t(x-1)}\, G_0\left(\frac{x-\alpha t(x-1)}{1-\alpha t(x-1)}\right). \tag{3.51}$$

The coherent transform Ψ then follows the PDE

$$\alpha^{-1}\partial_t\Psi = \partial_\phi\phi\partial_\phi\Psi, \tag{3.52}$$

which is a dilatation of the solution to the squared bessel process of dimension $\delta = 2$ [35]. The SDE associated to the process (3.52) is

$$d\phi = \alpha dt + \sqrt{2\alpha\phi}dW_t \tag{3.53}$$

It is important to note that the differential operator $A = \partial_\phi \phi \partial_\phi$ is symmetric, so formally $A = A^T$. Which in particular means that if we define

$$\begin{aligned} \alpha^{-1}\partial_t \xi(t,\phi) &= A\xi = \partial_\phi \phi \partial_\phi \xi \\ \xi(0,\phi) &= e^{\phi(x-1)} \end{aligned}$$

then the integral

$$I(t,s) = \int_0^\infty \xi(t-s,\phi)\Psi(s,\phi)d\phi \tag{3.54}$$

is independent of s, to see this we take derivative of I w.r.t s to obtain

$$\begin{aligned} \alpha^{-1}\frac{d}{ds}I(t,s) &= -\int_0^\infty \alpha^{-1}\partial_t \xi(t-s,\phi)\Psi(s,\phi)d\phi + \int_0^\infty \alpha^{-1}\xi(t-s,\phi)\partial_t\Psi(s,\phi)d\phi = & (3.55)\\ &= -\int_0^\infty (A\xi)(t-s,\phi)\Psi(s,\phi)d\phi + \int_0^\infty \xi(t-s,\phi)(A\Psi)(s,\phi)d\phi = & (3.56)\\ &= \int_0^\infty \xi(t-s,\phi)\left(A - A^T\right)\Psi(s,\phi)d\phi = 0 & (3.57) \end{aligned}$$

Note that as $\xi(\phi,t) \in L^1(0,\infty) \cap L^\infty(0,\infty)$ the integral I is well-defined once $\Psi(\phi,0) \in L^1(0,\infty)$.

In particular,

$$\int_0^\infty e^{\phi(x-1)}\Psi(t,\phi)d\phi = \int_0^\infty \xi(\phi,t)\Psi_0(\phi)d\phi. \tag{3.58}$$

And, as can be checked by direct computation the solution is

$$\xi(t,\phi) = \frac{1}{1-\alpha t(x-1)} e^{\frac{\phi(x-1)}{1-\alpha t(x-1)}} \tag{3.59}$$

which implies, substituting in the definition of coherent transform, that

$$\begin{aligned} G(x,t) &= \int_0^\infty e^{\phi(x-1)}\Psi(t,\phi)d\phi = \\ &= \frac{1}{1-\alpha t(x-1)}\int_0^\infty e^{\frac{\phi(x-1)}{1-\alpha t(x-1)}}\Psi_0(\phi)d\phi = \frac{1}{1-\alpha t(x-1)}\int_0^\infty e^{\phi\left(\frac{x-\alpha t(x-1)}{1-\alpha t(x-1)}-1\right)}\Psi_0(\phi)d\phi \\ &= \frac{1}{1-\alpha t(x-1)} G_0\left(\frac{x-\alpha t(x-1)}{1-\alpha t(x-1)}\right) \end{aligned}$$

in perfect agreement with (3.51). Note that this last result makes sense even if Ψ_0 is not a probability measure.

3.3.3 Coalescence reaction

In this section we are going to apply the Doi-Peliti field theory to study the coalescence reaction in zero space dimensions

$$A + A \xrightarrow{\lambda} \emptyset, \tag{3.60}$$

As we have seen in (3.45), its coherent transform follows, formally, a complex SDE.

This observation has been studied profoundly in [36, 37, 4, 38, 39, 40, 41, 42, 43, 44] and we will study that particular relation in chapter 6.

The system (3.60) is described by the differential equation

$$\frac{d}{dt}P_n(t) = \frac{\lambda}{2}[(n+2)(n+1)P_{n+2}(t) - n(n-1)P_n(t)]. \tag{3.61}$$

The probability vector state

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} P_n |n\rangle$$

follows the evolution equation

$$\frac{d}{dt}|\Psi(t)\rangle = H(a^\dagger, a)|\Psi(t)\rangle = \frac{\lambda}{2}(1 - a^{\dagger 2})a^2 |\Psi(t)\rangle \tag{3.62}$$

with Hamiltonian

$$H(a^\dagger, a) = \frac{\lambda}{2}(1 - a^{\dagger 2})a^2. \tag{3.63}$$

And the generating function PDE is

$$\frac{\partial G}{\partial t} = \frac{\lambda}{2}\left(1 - x^2\right)\frac{\partial^2 G}{\partial x^2}. \tag{3.64}$$

It is interesting to note that the elliptic operator vanishes in the boundary, so it is a degenerated case of the classical theory of second order differential PDE's. The theory regarding the existence and unicity of this kind of equations can be seen in [45, 46, 47], in fact, we will study this kind of processes in more detail in the next chapter when talking about duality as this kind of degenerated elliptic operator at the boundary is a general phenomenon related with those types of processes.

The degeneracy of the elliptic operator is related to the following conservation laws

1. *Conservation of probability* $G(1,t) = 1$ for all $t \geq 0$.

2. *Conservation of parity* $G(-1,t) = \wp$ for all $t \geq 0$, where $\wp = \sum_{n=0}^{\infty} P_{2n}(0) - \sum_{n=0}^{\infty} P_{2n+1}(0)$.

While the existence of the first conserved quantity is ensured by the nature of the Markov process, the existence of the second one is special of the binary annihilation (3.60). As particles are annihilated in pairs, the total probability of having an odd number of particles and that of having an even number, are separately conserved.

Theorem 3.6. *The factorial moments $\mathcal{M}_m$ fulfill the system of coupled differential equations*

$$\frac{1}{\lambda}\frac{d}{dt}\mathcal{M}_m = -\frac{m(m-1)}{2}\mathcal{M}_m - m\mathcal{M}_{m+1} \tag{3.65}$$

Proof. Taking m derivatives with respect to x in equation (3.64)

$$\frac{\partial}{\partial t}\frac{\partial^m G}{\partial x^m} = \frac{\lambda}{2}\sum_{j=0}^{m}\binom{m}{j}\frac{\partial^{m-j}}{\partial x^{m-j}}(1-x^2)\frac{\partial^{j+2}G}{\partial x^{j+2}}$$

Evaluating in $x = 1$ we obtain our result. □

We can solve the problem explicitly using separation of variables:

$$G(x,t) = T(t)X(x) \quad \text{so} \quad T'(t)X(x) = \frac{\lambda}{2}(1-x^2)T(t)X''(x). \tag{3.66}$$

Then

$$\frac{2T'}{\lambda T} = \frac{(1-x^2)X''}{X} = \mu, \quad \text{for some constant } \mu. \tag{3.67}$$

For T we obtain $T(t) = \exp\left(\frac{\mu\lambda t}{2}\right)$ for $t > 0$. In the equation for X

$$(1-x^2)X'' = \mu X \tag{3.68}$$

as we already know, we can consider the derivatives as annihilation of an index and multiplication by x as a creation of an index. Following that logic it is easy to see that polynomials of any degree are solutions of equation (3.68). If we assume that a polynomial g_n of degree n is a solution of (3.68) for some $\mu = C_n$, then $C_n = -n(n-1)$. Hence we obtain a sequence $\{g_n(x)\}$ of polynomials of degree n that satisfies equation (3.68). Additionally from the structure of (3.68) we obtain that polynomials of degree 0 and 1 solve this equation only if $X(1) = 1$ and $X(-1) = \wp$, or correspondently if $G(1,t) = 1$

and $G(-1,t) = \wp$ for some constant $\wp$ and $t \geq 0$. Furthermore, if g_n has a degree greater or equal to 2 then $g_n(1) = g_n(-1) = 0$ and using the boundary conditions we can calculate the constants A_0 and A_1 in the expansion

$$G(x,t) = \sum_{n=0}^{\infty} A_n g_n(x) e^{-n(n-1)\lambda t/2}.$$

The polynomials $g_n(x)$ for all $n \geq 2$ are orthogonal. To show it define

$$\langle g_n, g_m \rangle = \int_{-1}^{1} \frac{g_n g_m}{1-x^2} dx$$

and using (3.68) obtain that

$$\langle g_n, g_m \rangle = \frac{-1}{n(n-1)} \int_{-1}^{1} g_n'' g_m dx = \frac{-1}{n(n-1)} \int_{-1}^{1} g_n g_m'' dx = \frac{m(m-1)}{n(n-1)} \int_{-1}^{1} \frac{g_n g_m}{1-x^2} dx.$$

Then

$$(n(n-1) - m(m-1))\langle g_n, g_m \rangle = 0$$

and, hence, for $n, m \geq 2$ and $n \neq m$ we have $\langle g_n, g_m \rangle = 0$. Each monomial x^n can be obtained as a linear combination of polynomials g_n, with $n \in \mathbb{N}$.

Using the orthogonality of g_n we can determine the coefficients A_n, for $n \geq 2$, as

$$A_n = \frac{1}{\langle g_n, g_n \rangle} \int_{-1}^{1} \frac{G(x,0) g_n}{1-x^2} dx = -\frac{1}{n(n-1)} \frac{1}{\langle g_n, g_n \rangle} \int_{-1}^{1} G_0(x) g_n'' dx < \infty. \tag{3.69}$$

Even though the solution is fully known, there are some problems with it. The most important one being that we have solved this equation in an expansion that does not have "probabilistic" meaning, or at least not in a direct way. The reason for that is that the coefficients of the Gegenbauer polynomials can be both negative and positive.

3.4 Relation with Martin-Siggia-Rose Field Theory

As we stated previously, Doi-Peliti field theory appeared as one way to tackle, with field-theoretic methods (e.g., path integrals), stochastic relations were the state space is discrete, in contrast to the previous field-theoretic methods that assumed a topologically continuous state space.

One possible alternative to achieve this goal is to use Martin-Siggia-Rose Field theory, a method that dates back to the 70's [48]. We apply this technique in chapter 7 for solving a problem regarding a system of discrete stochastic processes appearing in the context of biochemistry.

Before that, we present in this section the relation of these two field-theoretic representations, using for that a particular use case, the familiar decaying reaction

$$A \xrightarrow{\lambda} \emptyset. \tag{3.70}$$

Using the analytical representation of Doi-Peliti field theory, the equation followed by this system is

$$\frac{\partial}{\partial t} G(x,t) = \lambda\,(1-x)\,\partial_x G(x,t), \tag{3.71}$$

i.e., the reaction (3.70) can be described through the Doi-Peliti Hamiltonian

$$H(x,p) = \lambda(1-x)p, \tag{3.72}$$

where $p = \partial_x$ and $[p,x] = 1$.

In order to see the relation between this approach and the one of Martin-Siggia-Rose we will use the following definition of the exponential of the momentum

$$e^p f(x) := f(x+1),$$

which can be related to the following chain of identities for the case $f(x) \in \mathcal{C}^\omega(\mathbb{R})$

$$e^p f(x) = \sum_{n=0}^{\infty} \frac{p^n}{n!} f(x) = \sum_{n=0}^{\infty} \frac{\partial_x^n}{n!} f(x) = f(x+1).$$

Coming back to the process (3.70), its Markov chain, as defined by the probabilities $P(n,t)$, follows the system of ODE's

$$\frac{\partial}{\partial t} P(n,t) = \lambda(n+1)P(n+1,t) - \lambda n P(n,t) \tag{3.73}$$

which can be seen now as an 'exponential' PDE

$$\frac{\partial}{\partial t}f(x,t) = \lambda(x+1)f(x+1,t) - \lambda x f(x,t) = \lambda e^{\partial_x}\{xf(x,t)\} - \lambda x f(x,t) = \tag{3.74}$$

$$= \lambda(1-e^{-\partial_x})e^{\partial_x}xf(x,t) \tag{3.75}$$

in which the continuous function $f(x,t)$ fulfills the property $f(n,t) = P(n,t)$ for each $n \in \mathbb{N}$.

The Martin-Siggia-Rose Hamiltonian for the reaction (3.70) is defined to be exactly

$$H(x,p) = (1-e^{-\partial_x})e^{\partial_x}x.$$

The relation between Doi-Peliti and Martin-Siggia-Rose Hamiltonians rely upon a canonical change of variables, that is, both make use of the 'same' Hamiltonian, but in a different representation. The change of variables that lead to the aforementioned relation is

$$X = e^{-p} \tag{3.76}$$

$$P = e^{p}x, \tag{3.77}$$

in which $p = \partial_x$. One realizes the Hamiltonian for (3.74) is

$$H(X,P) = \lambda(1-X)P,$$

additionally

$$[P,X] = PX - XP = e^{p}xe^{-p} - e^{-p}e^{p}x = 1.$$

The relation (3.76) is a canonical change of variables in a classical way as well, i.e., if we think of $\{x,p\}$ as commutative variables.

As we have seen, Doi-Peliti and Martin-Siggia-Rose formalism are utterly analogous in the sense that they relate to each other through a canonical change of variables; nevertheless, the convenience of the use of each one depends on a case by case basis. As a rule of thumb, Doi-Peliti formalism is more convenient when the coefficients of the Markov process have a combinatoric meaning while Martin-Siggia-Rose could be more useful for highly non-linear coefficients as we do in chapter 7.

Chapter 4

Duality

The concept of duality in stochastic processes appeared for the first time in the late 40s, and early 50s [49, 50, 51], and was developed in the following decades, nowadays is a well-establish and mature field of research.

Duality in stochastic processes can be defined in layman terms as: two stochastic processes are said to be dual if it is possible to extract information of one of the processes knowing some information about the other one.

Duality has, at least, two important consequences. Firstly, it can be used for solving one problem knowing the solution to a different one. Secondly, it tells us something about the structure we are studying that at first sight was not evident.

The fact that Doi-Peliti formalism is related to some types of duality allows us to find new dualities that were not obvious at first sight.

The first concept of this connection between Doi-Peliti theory and duality has been already studied in the previous chapter, as the generating function approach is related to moment duality. The theory developed in the previous chapter can be seen as a way of writing duality abstractly. The introduction of Sheffer sequences allow us to find all possible representations in a specific class of functions, the polynomials, that generalize the concept of generating function as a particular example in which the polynomial basis is just the set of monomials $\{x^n\}$.

Often, the duality of Markov processes is view as a way of representing a "backward-time behavior" that can be formally stated in a path integral formalism [5]. However, we are not going to use this

point of view here, as path integrals are useful for mental representation, but are difficult to justify from a mathematically rigorous point of view.

4.1 Duality in Stochastic Processes

Given two Markov processes in continuous time[1] $X_1(t)$ and $X_2(t)$ in probability spaces $(\Omega_1, \mathcal{A}_1, \mathcal{P}_1)$ and $(\Omega_2, \mathcal{A}_2, \mathcal{P}_2)$, and a function

$$F : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R},$$

we are going to say that $X_1(t)$ is dual to $X_2(t)$ with duality function F, if the following identities holds

Definition 4.1 (Duality)**.**

$$E_{1,2}\left[F(X_1(t), X_2(s))\right] = E_{1,2}\left[F(X_1(t'), X_2(s'))\right] \tag{4.1}$$

for all $0 \leq t' \leq t, 0 \leq s \leq s'$ *with* $t + s = t' + s' = T$ *where* $T \geq 0$ *is an arbitrary constant. And where* $E_{1,2}[\cdot] = E_1[E_2[\cdot]]$

In particular, the identity above states

$$E_{1,2}\left[F(X_1(T), X_2(0))\right] = E_{1,2}\left[F(X_1(0), X_2(T))\right] \tag{4.2}$$

for all $T \geq 0$.

Joining equation (4.2) with the right initial conditions [2] we can obtain information of $X_1(t)$ if we are able to solve $X_2(t)$ as well as the other way around.

Note that the state spaces Ω_1 and Ω_2 need not be similar, topologically speaking. For instance, one could be discrete while the other could be continuous.

If we restrict duality to the class of finite state Markov processes, their probability distributions

$$P = (p_1, \ldots, p_n)^T$$

[1] We are going to define duality just for continuous time, though it is possible to do it in the discrete time case as well.
[2] Dirac's delta or Kronecker's delta

and

$$Q = (q_1, \dots, q_m)^T$$

fulfill their respective Master equations

$$\begin{cases} \frac{d}{dt}P = L_1 P \\ \frac{d}{dt}Q = L_2 Q \end{cases}$$

where $L_1 \in \mathcal{M}_{n\times n}$ and $L_2 \in \mathcal{M}_{m\times m}$. In this context the function F is just a matrix $M \in \mathcal{M}_{n\times m}$

Lemma 4.1. *In the case of a finite state space, any function $F : \{1, \dots, n\} \times \{1, \dots, m\} \longrightarrow \mathbb{R}$ can be represented as a matrix $M \in \mathcal{M}_{n\times m}$ with $M_{i,j} = F(\{i\}, \{j\})$ such that the duality relation*

$$E_{1,2}\left[F(X_1(t), X_2(s))\right] = E_{1,2}\left[F(X_1(t'), X_2(s'))\right] \tag{4.3}$$

can be written as

$$P^T(t)MQ(s) = P^T(t')MQ(s')$$

Proof. It is enough to proof the connection between the left-hand side of both equations, as the right-hand side is analogous

$$E_{1,2}\left[F(X_1(t), X_2(s))\right] = \sum_{i,j} F(\{i\}, \{j\})P_i(t)Q_j(s) = \sum_{i,j} M_{i,j}P_i(t)Q_j(s) = P^T(t)MQ(s) \tag{4.4}$$

□

The following Lemma characterizes duality in this context

Lemma 4.2. *The processes $P(t) \in \mathbb{R}^n$ and $Q(t) \in \mathbb{R}^m$ are dual with respect to the matrix $M \in \mathcal{M}_{n\times m}$ if and only if*

$$L_1^T M = M L_2 \tag{4.5}$$

Proof. As we have seen, duality in the context of finite Markov processes states that

$$P^T(t)MQ(s) = P^T(t')MQ(s')$$

whenever $t + s = t' + s'$.

Defining

$$\gamma(t,s) = P^T(t)MQ(s)$$

the duality relation can be written as the transport PDE

$$\frac{d}{dt}\gamma(t,s) = \frac{d}{ds}\gamma(t,s)$$

which translates into

$$P^T(t)L_1^T MQ(s) = P^T(t)ML_2Q(s)$$

for all $P(t), Q(s)$ from which the identity follows. □

Of course, because of the properties of expectation, a constant function $F \equiv C$ will make all stochastic processes dual to each other. But if we restrict ourselves to say that two Markov processes are dual if there exist some non-constant function that makes them dual then the picture change quite a lot.

Just for the sake of completeness lets see an example of a very simple finite state Markov process duality, we will see more complex situations when dealing with non-finite state spaces in the next section.

We define the Markov process $P(t)$ with two states $S = \{1, 2\}$ and solving the Master equation

$$\begin{aligned}\frac{d}{dt}P_1(t) &= 2\alpha P_2(t)\\ \frac{d}{dt}P_2(t) &= -2\alpha P_2(t)\end{aligned}$$

which makes

$$L_1 = \begin{pmatrix} 0 & 2\alpha \\ 0 & -2\alpha \end{pmatrix}$$

While the Markov process $Q(t)$ has three states $S = \{1, 2, 3\}$ and follows the Master equation

$$\begin{aligned}\frac{d}{dt}Q_1(t) &= \alpha Q_2(t)\\ \frac{d}{dt}Q_2(t) &= -2\alpha Q_2(t)\\ \frac{d}{dt}Q_3(t) &= \alpha Q_2(t)\end{aligned}$$

with matrix

$$L_2 = \begin{pmatrix} 0 & \alpha & 0 \\ 0 & -2\alpha & 0 \\ 0 & \alpha & 0 \end{pmatrix}$$

We claim these two processes are dual with respect to the matrices

$$M_\Gamma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \Gamma & 1 \end{pmatrix} \tag{4.6}$$

where Γ is an arbitrary constant. The duality relation (4.2) translates into

$$Q_1(0) + Q_2(0)\left[P_1(t) + \Gamma P_2(t)\right] + Q_3(0) = Q_1(t) + Q_2(t)\left[P_1(0) + \Gamma P_2(0)\right] + Q_3(t) \tag{4.7}$$

4.2 Generating Function Representation and Duality

As we stated in the previous chapter, we are interested in the study of reactions of the type

$$jA \xrightarrow{\lambda_{lj}} lA, \tag{4.8}$$

that is, local reactions with one chemical species involved. We are mainly interested in studying pairwise interactions, i.e., $j \leq 2$. The reason for studying only pairwise interactions is that it is the common framework in physics where one supposes that a n-wise interaction could be broken down into a sequence of pairwise interactions.

In particular, we restrict ourselves to the set of reactions

$$\begin{aligned} A + A &\xrightarrow{\lambda} A, \\ A &\xrightarrow{\alpha} lA, \\ \emptyset &\xrightarrow{\beta} pA. \end{aligned} \tag{4.9}$$

Thanks to Theorem 3.4 in [33], we know that the forward and backward Chapman-Kolmogorov equations are well defined in this case, except for an eventual explosion at a finite time. We are going to prove that such an explosion is not possible in our setting so that the equations can be defined for all

times.

Theorem 4.1. *With probability one, there is no explosion in finite time for the Markov process that follows the system* (4.9).

Proof. Following Proposition 2.3 in [33], if we define the pathwise solution Y_n of the system of reactions (4.9), where Y_n stands for the state of the path after n jumps, and we define

$$R = \sum_n \Lambda(Y_n)^{-1}$$

with $\Lambda(Y_n)$ the absolute value of the diagonal element in the position Y_n of the intensity matrix. Then the set of explosive paths

$$\{Y_n : Y_N = \infty \text{ for a finite } N \in \mathbb{N}\}$$

has non-zero measure if and only if $R = \infty$

In our case the state space is $S = \mathbb{N}$, and the set of $\Lambda_n(\equiv \Lambda(n))$ is an increasing function of n. It is easy to see then that if system (4.9) is explosive then the system defined only by the birth rates will be explosive as well. In particular, if we proof that the system with birth rates

$$A \xrightarrow{\alpha} lA \tag{4.10}$$

$$\emptyset \xrightarrow{\beta} pA \tag{4.11}$$

where $l \geq 2, p \geq 1$ is non explosive then the theorem follows. Such system has

$$\Lambda_n = \alpha n + \beta$$

so we need to proof

$$R = \sum_n \frac{1}{\alpha Y_n + \beta} = \infty.$$

In our case, $Y_n \leq \gamma n + Y_0$, where $\gamma = \max\{l - 1, p\}$ which implies $\alpha Y_n + \beta \leq \gamma\alpha n + \beta + \alpha Y_0$ and so $(\alpha Y_n + \beta)^{-1} \geq (\gamma\alpha n + \beta + \alpha Y_0)^{-1}$

As we know from basic theory of series

$$\sum_n \frac{1}{\gamma\alpha n + \beta + \alpha Y_0} = \infty$$

which proofs our result. □

Remark. *The previous result implies the existence of a unique solution* $P_n(t)\ \forall t > 0$ *to the system of Chapman-Kolmogorov equations defined by the system of reactions* (4.9).

From Lemma (3.10), the generating function of the reaction (4.8) is

$$\frac{\partial G}{\partial t} = \frac{\lambda_{lj}}{j!}\left(x^l - x^j\right)\frac{\partial^j G}{\partial x^j}. \tag{4.12}$$

One of the advantages of this formalism is that, calling

$$D_{lj} := \frac{\lambda_{lj}}{j!}\left(x^l - x^j\right)\frac{\partial^j}{\partial x^j}, \tag{4.13}$$

we just need to add these operators in the generating function PDE for studying a system of reactions of the form (4.8) from which we obtain that the form of the generating function PDE in our case is

$$\frac{\partial G}{\partial t} = b(x)\frac{\partial^2 G}{\partial x^2} + a(x)\frac{\partial G}{\partial x} + c(x)G, \tag{4.14}$$

with a, b, c polynomials with degree(b)=2 if $b \neq 0$.

Of course, these polynomials cannot be arbitrary, in particular $b(x) = \frac{\lambda}{2}x(1-x)$. Also, they fulfill some properties that simplify our task.

Lemma 4.3. *1. Probability conservation*

$$a(1) = b(1) = c(1) = 0 \tag{4.15}$$

2. Compatibility boundary condition

$$b(0) = 0 \tag{4.16}$$

$$a(0) = \delta_{p,0} \geq 0 \tag{4.17}$$

3. The coefficient $c(x) \leq 0$ *in its natural interval of definition.*

Proof. 1. The coefficients of the PDE are linear combinations of monomials of the form $\frac{\lambda}{j!}\left(x^l - x^j\right)$, so $x = 1$ is always a root.

2. $b(0) = 0$ because $b(x) = \dfrac{\lambda}{2}x(1-x)$.

   $a(x)$ is a linear combination (with positive coefficients) of monomials of the form $x^n - x$ with $n \geq 0$ which is null for $n > 0$ and positive for $n = 0$.

3. $c(x)$ is of the form $\lambda\left(x^n - 1\right)$ with $n \geq 1$ and $\lambda \geq 0$. So it is negative in $[0, 1]$.

□

Additionally, our set of reactions have an important property

Theorem 4.2. *If all the moments of the initial condition are finite, the moments of the stochastic process defined through the evolution of the Markov process are finite.*

Proof. We are going to prove that all factorial moments $\mathcal{M}_n := E\left[n \cdot (n-1) \cdots (n-m+1)\right]$ are finite, which imply that all moments are finite and vice-versa as they relate through an algebraic relation. Additionally, as our state space is $\mathbb{N}$ all moments and all factorial moments are positive (≥ 0).

The factorial moments, whenever finite, can be extracted from the generating function as

$$\mathcal{M}_n(t) = \frac{\partial^m G}{\partial x^m}(1, t) \tag{4.18}$$

As initially all moments are finite, and we know the explicit form of the evolution of G we could derive a system of ODE's with the following form

$$\partial_t \mathcal{M}_n = m b_x(1) \mathcal{M}_{n+1} + F_n\left(\mathcal{M}_n, \mathcal{M}_{n-1}, \cdots, \mathcal{M}_1\right) \tag{4.19}$$

where F_n is linear and we have used Lemma 4.15. Additionally, $\partial_x b(1) < 0$ which implies that

$$\partial_t \mathcal{M}_n \leq F_n\left(\mathcal{M}_n, \mathcal{M}_{n-1}, \cdots, \mathcal{M}_1\right) \tag{4.20}$$

which, joint with an iterative application of Gronwall's Lemma give us that all factorial moments are finite for all times. □

Remark. *As the (factorial) moments of the Markov chain can be recovered as derivatives of the generating function*

$$E\left[n \cdot (n-1) \cdots (n-m+1)\right] = \frac{\partial^m G}{\partial x^m}(1, t) \tag{4.21}$$

and all moments are finite, following Chapter XI Theorem 2 [34] we obtain that the generating function, and its derivatives, are continuous in the closed interval $[0,1]$ *which makes that the generating function PDE is well defined in strong sense on that interval.*

Using the classical version of Feynman-Kac theorem, one is tempted to say that, for the PDE (4.14) the following identity will be true

$$G(x,t) = E^x\left[G_0(X_t)e^{\int_0^t ds\ c(X_s)}\right] \tag{4.22}$$

where X_t is the stochastic process followed by the Itô SDE

$$dX_t = a(X_t)dt + \sqrt{2b(X_t)}dW_t \tag{4.23}$$

on $[0,1]$.

This identity is valid, but the classical theory does not fully cover it on Feynman-Kac. From Lemma 4.15 this Itô SDE is degenerate on the boundary and does not fulfill the Lipschitz condition. Luckily for us, recent developments [46, 45, 47] on the theory of degenerate elliptic operators, improve the classical results for some instances with degenerate boundaries.

In particular, it can be stated that the Itô stochastic process (4.23) possesses a unique strong solution with $\{1\}$ as absorbing state. Additionally, $\{0\}$ will be an absorbing state if, and only if, $a(0) = 0$. So, in particular, Dirac's deltas appear at a finite time at the boundaries of the Itô SDE that are absorbing and, eventually, all probability gets stick at those absorbing boundaries.

Theorem 4.3 (Feynman-Kac)**.** *Given polynomials* $a(x), b(x), c(x)$ *defined as in the equation* (4.14)*. The solution* $G(x,t)$ *to the initial value problem*

$$\partial_t G = a(x)\frac{\partial G}{\partial x} + b(x)\frac{\partial^2 G}{\partial x^2}G + c(x)G \tag{4.24}$$

with an analytic initial condition $G_0(x) \in C^\omega(\mathbb{R})$*, can be written in terms of the process followed by the Itô equation*

$$dX_t = a(X_t)dt + \sqrt{2b(X_t)}dW_t \tag{4.25}$$

as

$$G(x,t) = E^x\left[G_0(X_t)e^{\int_0^t ds\ c(X_s)}\right] \tag{4.26}$$

Proof. The Itô equation (4.25) is well defined and have a unique, strong solution even though the coefficients are not Lipschitz, due to a theorem of Watanabe and Yamada [45, 46, 52, 53, 35, 47]. Additionally, as a fulfill 4.15 and 4.16, the boundary is absorbing in this case if and only if a is null on it. In particular, $\{1\}$ is always an absorbing boundary. That is, there is no need for additional boundary conditions. Furthermore, its probability distribution behave as an absolutely continuous measure (with smooth density) with respect to Lebesgue measure in the interior of $(0,1)$; plus a Dirac delta with the corresponding weight (probability is conserved) in each absorbing boundary [46]. As has been seen before, the initial value problem (4.24) is well defined, that is, there exist an unique solution that is $C^{\omega}(0,1) \cap C\,[0,1]$ and which fulfill $\partial_x^n G(0,t) \geq 0 \ \forall t$. The reason for that been the Markov process associated, from which G is its generating function, is well defined and have a unique solution as has been seen before.

Additionally, in the case of absorbing boundary, as a and b are null, we can easily calculate the value of G on that boundary

$$G(\cdot,t) = G_0(\cdot)e^{c(\cdot)t} \leq G_0(\cdot) \tag{4.27}$$

Following [45, 46] ,and the usual Feynman-Kac theorem, we define the stopping times

$$\zeta_0 = inf\,\{t > 0 \,:\, X_t = 0\}$$

and

$$\zeta_1 = inf\,\{t > 0 \,:\, X_t = 1\}\,.$$

We also define the stopping time $\zeta := \zeta_0 \wedge \zeta_1$.

We then apply Itô's lemma to the process

$$Y_s = e^{-R(s)}G(t-s, X_s) \tag{4.28}$$

with

$$R(s) := \int_0^s c(X_\gamma)d\gamma. \tag{4.29}$$

It is easy to see that Y_s is a martingale; and evaluating its expected value at $s = 0$ and $s = t$ we obtain

$$G(x,t) = E^x\left[G_0(X_t)e^{\int_0^t ds\ c(X_s)}\right], \tag{4.30}$$

where x in $E^x[\cdot]$ represents that the Itô equation is possed with initial condition $X_0 = x$.

Relation (4.30) can be written, in case $\{0\}$ were an absorbing boundary, as

$$G(x,t) = E^x\left[e^{\int_0^t c(X_s)ds}1_{\zeta_0<t}\right]G_0(0,t) + E^x\left[e^{\int_0^t c(X_s)ds}1_{\zeta_1<t}\right]G_0(1) + E^x\left[G_0(X_t)e^{\int_0^t ds\ c(X_s)}1_{\zeta>t}\right] =$$
$$= E^x\left[e^{\int_0^t c(X_s)ds}1_{\zeta_0<t}\right]G_0(0)e^{c(0)t} + E^x\left[e^{\int_0^t c(X_s)ds}1_{\zeta_1<t}\right] + E^x\left[G_0(X_t)e^{\int_0^t ds\ c(X_s)}1_{\zeta>t}\right]$$

It is important to note that in case $\{0\}$ were not an absorbing boundary the identity will reduce to

$$G(x,t) = E^x\left[e^{\int_0^t c(X_s)ds}1_{\zeta_1<t}\right]G_0(1) + E^x\left[G_0(X_t)e^{\int_0^t ds\ c(X_s)}1_{\zeta>t}\right] =$$
$$= E^x\left[e^{\int_0^t c(X_s)ds}1_{\zeta_1<t}\right] + E^x\left[G_0(X_t)e^{\int_0^t ds\ c(X_s)}1_{\zeta>t}\right]$$

□

There is a clear relation between equation (4.24) and duality between stochastic processes.

Firstly we explicitly write the definition of the generating function in both sides of the equality (4.26)

$$\sum_n P_n(t)x^n = \sum_n P_n(0)E^x\left[X_t^n e^{\int_0^t ds\ c(X_s)}\right], \tag{4.31}$$

where X_t is the Itô process following (4.25) and $P_n(t)$ solves the Chapman-Kolmogorov equations for the system (4.9).

For the case $c(x) \equiv 0$, we can rewrite (4.31) as

$$\sum_n \int P_n(t)x^n p(x,0)dx = \sum_n \int P_n(0)x^n p(x,t)dx, \tag{4.32}$$

where we include the possibility of Dirac's delta in the definition of the probability distribution $p(x,t)$.

Theorem 4.4. *The Markov process defined by the system* (4.9) *is dual to the Itô stochastic process* (4.25) *with*

$$F(x,n) = x^n$$

that is, these two processes are moment dual to each other. [3]

[3] See [54] for a complete definition and examples for moment duality

For the case $c(x) \neq 0$ we can rewrite (4.31)

$$\sum_n P_n(t) E^x \left[X_0^n e^{\int_0^0 ds\ c(X_s)} \right] = \sum_n P_n(0) E^x \left[X_t^n e^{\int_0^t ds\ c(X_s)} \right] \tag{4.33}$$

which says that if we make a stochastic representation of the Doi-Peliti usual basis $|n\rangle$ [2] as a time dependent stochastic process, then we could fix the coefficients and just evolve the basis.

Let's finally this Section with some examples of this type of duality.

4.2.1 Coalescence vs. Wright-Fisher Diffusion

The coalescence reaction

$$A + A \xrightarrow{\lambda} A \tag{4.34}$$

satisfy the generating function PDE

$$\frac{\partial G}{\partial t} = \frac{\lambda}{2} \left(x - x^2 \right) \frac{\partial^2 G}{\partial x^2} \tag{4.35}$$

in the interval $[0, 1]$. The dual Itô process is the Wright-Fisher diffusion

$$dX_t = \sqrt{\lambda \left(x - x^2 \right)} dW_t \tag{4.36}$$

And we can write the duality relation in the limit $t \to \infty$ as

$$G_\infty(x) = p_0^x(\infty) P_0 + p_1^x(\infty) \tag{4.37}$$

where $p_{0,1}^x(t)$ is the probability of the absorbing states in the Itô process at time t, and P_0 is the probability of being in the state zero for the Markov process, which is conserved in the evolution. As (4.36) is a martingale, we obtain $p_0^x(\infty) = 1 - x$ and $p_1^x(\infty) = x$, so

$$G_\infty(x) = P_0 + x(1 - P_0) \tag{4.38}$$

that is, the large time limit of the reaction is to maintain the probability of having 0 particles and the rest of the probability evolves into the probability of obtaining 1 particle.

4.2.2 Coalescence plus Division

$$A + A \xrightarrow{\alpha} A \tag{4.39}$$

$$A \xrightarrow{\beta} A + A \tag{4.40}$$

with generating function satisfying

$$\frac{\partial G}{\partial t} = \frac{\alpha}{2}\left(x - x^2\right)\frac{\partial^2 G}{\partial x^2} + \beta\left(x^2 - x\right)\frac{\partial G}{\partial x} \tag{4.41}$$

which has as dual

$$dX_t = \beta\left(x^2 - x\right) + \sqrt{\alpha\left(x - x^2\right)}dW_t \tag{4.42}$$

We first realize, using Itô's lemma, that $e^{\frac{2\beta}{\alpha}X_t}$ is a martingale. So following similar lines as before, the limit at large time will be

$$G_\infty(x) = P_0\frac{e^{\frac{2\beta}{\alpha}} - e^{\frac{2\beta}{\alpha}x}}{e^{\frac{2\beta}{\alpha}} - 1} + \frac{e^{\frac{2\beta}{\alpha}x} - 1}{e^{\frac{2\beta}{\alpha}} - 1} \tag{4.43}$$

4.2.3 Coalescence and Creation

$$A + A \xrightarrow{\alpha} A \tag{4.44}$$

$$\emptyset \xrightarrow{\beta} A \tag{4.45}$$

with generating function PDE

$$\frac{\partial G}{\partial t} = \frac{\alpha}{2}\left(x - x^2\right)\frac{\partial^2 G}{\partial x^2} + \beta\left(x - 1\right)G \tag{4.46}$$

and dual Itô equation associated

$$dX_t = \sqrt{\alpha\left(x - x^2\right)}dW_t \tag{4.47}$$

which is a martingale. In this case, we are going to use the information about the stationary G to extract information about the Itô process, it is easy to solve the PDE

$$\frac{\alpha}{2}\left(x - x^2\right)\frac{\partial^2 G_\infty}{\partial x^2} + \beta\left(x - 1\right)G_\infty = 0 \tag{4.48}$$

by using this relation iteratively for obtaining the values $\frac{\partial^n G_\infty}{\partial x^n}(0)$ and then Taylor expand around $x = 0$

$$G_\infty(x) = \left(\sum_{n=1}^{\infty} \frac{1}{n!(n-1)!} \left(\frac{2\beta}{\alpha}\right)^{n-1}\right)^{-1} \sum_{n=1}^{\infty} \frac{1}{n!(n-1)!} \left(\frac{2\beta}{\alpha}\right)^{n-1} x^n \tag{4.49}$$

and so,

$$E^x\left[e^{\int_0^\infty (X_t-1)dt}\right] = \left(\sum_{n=1}^{\infty} \frac{1}{n!(n-1)!} \left(\frac{2\beta}{\alpha}\right)^{n-1}\right)^{-1} \sum_{n=1}^{\infty} \frac{1}{n!(n-1)!} \left(\frac{2\beta}{\alpha}\right)^{n-1} x^n \tag{4.50}$$

4.3 A Generalize Duality: Sheffer Sequences and Doi-Peliti

As we have seen in previous sections, the Doi-Peliti formalism is particularly useful when a continuous time Markov process in discrete state space possesses a "combinatorial structure". In our context, this means that the evolution operator of the Markov process can be written in terms of the annihilation-creation operators a and $a^\dagger$,

$$\partial_t |\Psi(t)\rangle = H\left(a^\dagger, a\right) |\Psi(t)\rangle . \tag{4.51}$$

In particular we know that all reactions of the type (4.9) can be written in this fashion, see Lemma (3.1).

By Lemma (3.11) we know that these annihilation-creation operators satisfy the commutation relation

$$\left[a, a^\dagger\right] = aa^\dagger - a^\dagger a = 1 \tag{4.52}$$

where 1 represents here the identity operator.

It is often useful to think of annihilation/creation operators in Doi-Peliti theory as a Weyl algebra, which is the ring of differential operators with polynomial coefficients in one variable. In particular one can clearly make the identifications

$$a^\dagger \equiv x \tag{4.53}$$

$$a \equiv \partial_x \tag{4.54}$$

which leads to the generating function representation as we already have studied, in particular, this

identification allows us to unveil some duality relations between abstract chemical Markov chains and Itô SDE's.

The purpose of this section is to explore other ways of identifying these operators as differential polynomial operators. In particular, if we find two polynomials $p(x,y), q(x,y)$ such that

$$[p(x,\partial_x), q(x,\partial_x)] = 1, \tag{4.55}$$

then, we have another representation of annihilation and creation operators

$$a^\dagger \equiv q(x,\partial_x) \tag{4.56}$$

$$a \equiv p(x,\partial_x). \tag{4.57}$$

Theorem 4.5. *Let $\{p(x,y), q(x,y)\} \in \mathbb{R}[x,y]$ be two polynomials satisfying*

$$[p(x,\partial_x), q(x,\partial_x)] = 1 \tag{4.58}$$

$$p(x,\partial_x)\cdot 1 = 0 \tag{4.59}$$

where $\cdot$ represents the product of the operator with the constant polynomial .

The following sequence of polynomials

$$s_n(x) := q(x,\partial_x)^n \cdot 1, \tag{4.60}$$

satisfy the property

$$p(x,\partial_x)s_n(x) = n s_{n-1}(x) \tag{4.61}$$

Proof. Relation (4.62) is just

$$p(x,\partial_x)q(x,\partial_x) - q(x,\partial_x)p(x,\partial_x) = 1 \tag{4.62}$$

$$p(x,\partial_x)\cdot 1 = 0 \tag{4.63}$$

One can proof easily that

$$p(x,\partial_x)q(x,\partial_x)^n \cdot 1 = q(x,\partial_x)^{n-1}\cdot 1 + q(x,\partial_x)\cdot p(x,\partial_x)\cdot q(x,\partial_x)^{n-1}\cdot 1 \tag{4.64}$$

which by induction implies

$$p(x,\partial_x)q(x,\partial_x)^n\cdot 1 = nq(x,\partial_x)^{n-1}\cdot 1 + q(x,\partial_x)^n\cdot p(x,\partial_x)\cdot 1 = nq(x,\partial_x)^{n-1} \tag{4.65}$$

□

Definition 4.2 (Sheffer Sequence)**.** *A Sheffer sequence is a sequence of polynomials* $\{s_n(x) : n\in\mathbb{N}\}$*, each one of degree* n*, that fulfill the relation*

$$Qs_n(x) = ns_{n-1}(x),$$

where Q *is a linear operator in the space of polynomials. Without loss of generality, we can take* $s_0(x) = 1$ *and* $Qs_0(x) = 0$*.*

Lemma 4.4 (Sheffer representation)**.** *We can represent the states* $|n\rangle$ *of the Doi-Peliti Field Theory as elements of a Sheffer sequence* $s_n(x)$*, where the operator* Q *is the annihilation operator.*

We will call this representation a Sheffer representation.

Proof. The previous Theorem with the notion of Sheffer sequence clearly implies the result. □

Definition 4.3 (Appell sequences)**.** *Appell sequences are Sheffer sequences such that* $a = \partial_x$*.*

Remark. *Examples of Appell sequences are monomials, Hermite sequence, Bernoulli sequence and Euler sequence, among others.*

Lemma 4.5. *All Appell sequences can be defined, at least formally, from the operators* $a = \partial_x$ *and*

$$a^\dagger = x + \sum_{n=0}^{\infty} c_n \frac{d^n}{dx^n}$$

where c_n *are arbitrary real numbers as*

$$s_n(x) = \left(x + \sum_{m=0}^{\infty} c_m \frac{d^m}{dx^m}\right)^n \cdot 1 \tag{4.66}$$

Proof. The creation operator can be written generally as

$$a^\dagger = \sum_{m,n} c_{m,n} x^n \partial_x^m, \tag{4.67}$$

with $c_{n,m}$ real numbers. Imposing the commutation relation (4.52) we obtain

$$\sum_{m,n} c_{m,n}\partial_x \cdot x^n \cdot \partial_x^m - \sum_{m,n} c_{m,n}x^n \cdot \partial_x^{m+1} = 1. \tag{4.68}$$

Also using the formula

$$\partial_x x^n = nx^{n-1} + x^n\partial_x \tag{4.69}$$

we obtain

$$\sum_{m,n} c_{m,n}n\partial_x \cdot x^{n-1} \cdot \partial_x^m + \sum_{m,n} c_{m,n}x^n \cdot \partial_x^{m+1} - \sum_{m,n} c_{m,n}x^n \cdot \partial_x^{m+1} = \tag{4.70}$$

$$= \sum_{m,n} c_{m,n}n\partial_x \cdot x^{n-1} \cdot \partial_x^m = 1. \tag{4.71}$$

In particular, the only possibilities are $n = 0$ for any m or $n = 1, m = 0$. This can be written as

$$a^\dagger = x + \sum_{n=0}^{\infty} c_n \frac{d^n}{dx^n} \tag{4.72}$$

Using Theorem (4.5) we can write

$$s_n(x) = \left(x + \sum_{m=0}^{\infty} c_m \frac{d^m}{dx^m} \right)^n \cdot 1 \tag{4.73}$$

Also, finding the coefficients c_n for each Sheffer sequence is a solvable problem using the previous equality as increasing n will give one coefficient at a time. □

Using Lemma (4.4) the Sheffer representation of the set of probabilities $\{P_n(t)\}_n$ is the function

$$S(x,t) = \sum_{n=0}^{\infty} P_n(t)s_n(x) \tag{4.74}$$

whenever $S(x,t)$ is well-defined for all times $t > 0$.

Now, for every Itô process X_t satisfying the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{4.75}$$

$$X_0 = z \tag{4.76}$$

we define the Sheffer moments as

Definition 4.4 (Sheffer moments)**.** *We define the n-Sheffer moment of the Itô process X_t defined by* (4.75) *as*

$$E^z\left[s_n(X_t)\right] = \int_\alpha^\beta s_n(x)p(x,t)dx \tag{4.77}$$

where the integration limits will depend on the interval of definition of (4.75)*. And where $p(x,t)$ solves the forward Komogorov (or Fokker-Planck) equation*

$$\frac{\partial p}{\partial t} = -\partial_x\left\{b(x)p(x,t)\right\} + \frac{1}{2}\partial_x^2\left\{\sigma^2(x)p(x,t)\right\} \tag{4.78}$$

with initial condition a Dirac's delta at z.

Definition 4.5 (Sheffer duality)**.** *We define the discrete state Markov process $\{P_n(t)\}$ with $n \in \mathbb{N}$ to be Sheffer dual to the Itô process X_t whenever*

$$\sum_{n=0}^{\infty}\int_\alpha^\beta P_n(t)s_n(x)p(0,x)dx = \sum_{n=0}^{\infty}\int_\alpha^\beta P_n(0)s_n(x)p(t,x)dx \tag{4.79}$$

and where the integration limits $\{\alpha,\beta\}$ depends on the interval of definition of (4.75)

Remark. *This definition is just a restricted case of the usual definition of duality* (4.1) *with $F(x,n) = s_n(x)$*

The definition of Sheffer Duality, has a clear connection with the Feynman-Kac theorem from (4.3). Solving the Itô process with a Dirac delta initial distribution one obtains

$$S(x,t) = E^x\left[S(0,X_t)\right], \tag{4.80}$$

where S is given by (4.74) and $E^x[\cdot]$ means expected value with $X_0 = x$.

4.3.1 Hermite Duality

In this section, we are going to study a particular example of Sheffer duality related to Hermite polynomials, a type of Appell sequences.

If one chooses as annihilation operator $a = \partial_x$ and as creation operator $a^\dagger = x - \partial_x$, one obtains the

Hermite sequence $\{h_n(x)\}$ representation as $|n\rangle = (a^\dagger)^n |0\rangle$ thanks to Theorem (4.5)

$$|n\rangle \equiv h_n(x) = (x - \partial_x)^n \cdot 1 = (-1)^n e^{x^2/2} \cdot \partial_x^n \cdot e^{-x^2/2} \tag{4.81}$$

Lemma 4.6. *$h_n(x) \in L^2\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx\right)$ and*

$$\int_{-\infty}^{\infty} h_n(x) h_m(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = n! \delta_{n,m} \tag{4.82}$$

Proof. $h_n(x)$ is a polynomials of degree n so clearly $h_n(x) \in L^2\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx\right)$.

Let's assume now that $n > m$ then

$$\int_{-\infty}^{\infty} h_n(x) h_m(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = \int_{-\infty}^{\infty} (-1)^n e^{\frac{x^2}{2}} \left(\partial_x^n e^{-\frac{x^2}{2}}\right) h_m(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = \tag{4.83}$$

$$= \int_{-\infty}^{\infty} \partial_x^n h_m(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = 0 \tag{4.84}$$

where in the last equation we are using the fact that $h_m(x)$ is a polynomial of degree m.

We only need to proof now that

$$\int_{-\infty}^{\infty} \partial_x^n h_n(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = n! \tag{4.85}$$

which is clear when one realizes that, because $h_n(x) = (x - \partial_x)^n \cdot 1$ the maximum degree monomial of $h_n(x)$ is x^n from which the result follows. □

Let $P_n(t)$ be a Markov process following the set of reactions (4.8). Making use of the Sheffer representation for the present case of Hermite polynomials, the Hermite representation of $P_n(t)$ is

$$H(x,t) = \sum_{n=0}^{\infty} P_n(t) h_n(x). \tag{4.86}$$

Following the previous result (4.6), the natural space of definition for the Markov process $\{P_n(t)\}_{n\in\mathbb{N}}$ is

$$\mathcal{H} = \left\{ |\Psi\rangle = \sum_{n=0}^{\infty} a_n |n\rangle : \|\Psi\| := \left(\sum_{n=0}^{\infty} n! |a_n|^2\right)^{1/2} < \infty \right\},$$

which can be related with the space $\mathcal{R}^\infty$ given in Definition (3.2)

Theorem 4.6. $\mathcal{H} \subset \mathcal{R}^\infty$

Proof. Let $\{a_n\}$ be a sequence with $a_n \in \mathcal{H}$ then, we can clearly divide the sequence $\{a_n\}$ between two non-intersecting subsets $\{n^m < n!a_n\}$ and $\{n!a_n \leq n^m\}$.

1. In the first case, multiplying by a_n this subset has finite sum.
2. In the second, we have $a_n n^m \leq n^{2m}/n!$ which sum is bound by $\dfrac{\partial}{\partial x^{2m}} \exp(\exp(x)) \mid_{x=0}$

Which clearly implies that the sequence $a_n \in \mathcal{R}^\infty$ □

Lemma 4.7. *Hermite polynomials* $\{h_n(x)\}$ *are a complete orthogonal basis of* $L^2\left(\dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx\right)$

Proof. We already know that Hermite polynomials are orthogonal to each other by lemma (4.6), so we just need to proof the completeness.

Lets suppose there is a function $f \in L^2\left(\dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx\right)$ such that

$$\int_{-\infty}^{\infty} f(x) h_n(x) e^{-\frac{x^2}{2}} = 0 \ \forall n \in \mathbb{N} \tag{4.87}$$

In particular, this implies the following relation by finite linear combinations

$$\int_{-\infty}^{\infty} f(x) x^n e^{-\frac{x^2}{2}} = 0 \ \forall n \in \mathbb{N} \tag{4.88}$$

Which implies that the entire function

$$F(z) = \int_{-\infty}^{\infty} f(x) e^{zx - \frac{x^2}{2}} dx = \sum_{n=0}^{\infty} \frac{z^n}{n!} \int_{-\infty}^{\infty} f(x) x^n e^{-\frac{x^2}{2}} dx = 0 \tag{4.89}$$

is null. In particular, the Fourier transform of the function $f(x)e^{-\frac{x^2}{2}} \in L^2(\mathbb{R})$ is null, implying that the function itself is null. □

Corollary 4.1. *The Hermite representation* $H(x,t)$ *of the Markov process* $\{P_n(t)\}_{n\in\mathbb{N}}$ *follows the property* $H(x,t) \in L^2\left(\dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx\right)$ *if, and only if,* $P_n(t) \in \mathcal{H}$.

Proof. Because we know that Hermite polynomials are a complete orthogonal basis of $L^2\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx\right)$ then we it is trivial to obtain the result given the definition of the Hermite representation

$$H(x,t)=\sum_{n=0}^{\infty}P_n(t)h_n(x) \tag{4.90}$$

which implies

$$\parallel H(x,t)\parallel=\parallel P_n(t)\parallel_{\mathcal{H}} \tag{4.91}$$

□

Decaying Process

From the reaction

$$A\xrightarrow{\lambda}\emptyset \tag{4.92}$$

the Hermite representation $H(x,t)$ will follow the PDE

$$\frac{1}{\lambda}\partial_t H=(1-x)\partial_x H+\partial_x^2 H. \tag{4.93}$$

Where we have used Lemma (3.1) plus the Hermite representation

$$\begin{aligned} a &= \partial_x \\ a^\dagger &= x-\partial_x \end{aligned}$$

This is the Backwards-Kolmogorov equation of the Itô SDE

$$dX_t=\lambda(1-X_t)dt+\sqrt{2\lambda}dW_t \tag{4.94}$$

which have as Forward-Kolmogorov equation

$$\frac{1}{\lambda}\partial_t p=\partial_x\left\{(x-1)p\right\}+\partial_x^2 p. \tag{4.95}$$

Following a parallel argument to the one given in the monomial duality case, we obtain that (4.92) and (4.94) are Hermite duals of each other. Our aim is to extract valuable information of (4.92) from

information in (4.94) through the duality relation.

Choosing $P_n(0) = \delta_{n,m}$ as the initial condition for (4.92), where $m \in \mathbb{N}$, and Theorem (4.3) the following identity holds

$$H(x,t) = E^x \left[H(0, X_t) \right] = E^x \left[h_m(X_t) \right]. \tag{4.96}$$

In particular, as for $t \to \infty$ the Itô process converges in law to the distribution $p(x,\infty) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}$ and, as from the identity (4.82),

$$\int_{-\infty}^{\infty} h_m(x) \frac{h_n(x)}{n!} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \delta_{m,n} \tag{4.97}$$

one can sum in n having into account that $\sum_{n=0}^{\infty} (n!)^{-1} h_n(x) = e^{x-\frac{1}{2}}$ and obtain

$$E^x \left[h_m(X_\infty) \right] = \int_{-\infty}^{\infty} h_m(y) \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}} dy = 1 \ \ \forall \ m \in \mathbb{N} \tag{4.98}$$

and so, $H(x,\infty) = 1$ which just means that $P_n(\infty) = \delta_{0,n}$, that is, there are no particles left.

Decaying Process plus Creation

A more interesting set of reactions is

$$A \xrightarrow{\alpha} \emptyset \tag{4.99}$$

$$\emptyset \xrightarrow{\beta} A \tag{4.100}$$

which in Hermite representation fulfills the PDE

$$\frac{\partial H}{\partial t} = \{\alpha(1-x) - \beta\} \partial_x H + \alpha \partial_x^2 H + \beta(x-1) H. \tag{4.101}$$

Feynman-Kac theorem (4.3) allows us again to study this kind of equations with 'potential'-like terms, in particular,

$$H(x,t) = E^x \left[H(0, X_t) e^{\beta \int_0^t (X_s - 1) ds} \right] \tag{4.102}$$

where X_t is the Itô process

$$dX_t = (\alpha(1 - X_t) - \beta)\, dt + \sqrt{2\alpha} dW_t \tag{4.103}$$

with $X_0 = x$. Now we choose, as before, an initial condition with exactly 'm' particles $H(0,x) = h_m(x)$.

For studying the large time limit of both reactions, one can easily seen that the Itô process (4.103) converges in law to a Gaussian distribution with mean $1 - \frac{\beta}{\alpha}$ and variance 1. Additionally, the Markov process (4.99) will converge to a unique stationary distribution as can be seen from classical theory of Markov processes. Our aim is to calculate this large time distribution of the Markov process using the stochastic duality (4.102). We obtain

$$H(x,\infty) = \lim_{t\to\infty} E^x \left[h_m(X_t) e^{\beta \int_0^t (X_s - 1) ds} \right]. \tag{4.104}$$

Using (4.103) it is easy to see

$$\beta \int_0^t (X_s - 1) ds = \frac{\beta}{\alpha} x - \frac{\beta}{\alpha} X_t - \frac{\beta^2 t}{\alpha} + \sqrt{\frac{2\beta^2}{\alpha}} W_t \tag{4.105}$$

Substituting in (4.104) we get

$$H(x,\infty) = e^{\frac{\beta}{\alpha} x} \lim_{t\to\infty} E^x \left[h_m(X_t) e^{-\frac{\beta}{\alpha} X_t} e^{-\frac{\beta^2 t}{\alpha} + \sqrt{\frac{2\beta^2}{\alpha}} W_t} \right]. \tag{4.106}$$

As we know $H(x,\infty)$ should not depend on m (the stationary distribution of (4.99) is unique), we could multiply both sides of (4.106) by $\frac{e^{-\frac{\beta}{\alpha}}}{m!} \left(\frac{\beta}{\alpha}\right)^m$ and sum over all $m \in \mathbb{N}$ for obtaining

$$\begin{aligned} H(x,\infty) &= e^{\frac{\beta}{\alpha}(x-1)} \lim_{t\to\infty} E^x \left[e^{\frac{\beta}{\alpha} X_t - \frac{\beta^2}{2\alpha^2}} e^{-\frac{\beta}{\alpha} X_t} e^{-\frac{\beta^2 t}{\alpha} + \sqrt{\frac{2\beta^2}{\alpha}} W_t} \right] = & (4.107) \\ &= e^{\frac{\beta}{\alpha}(x-1) - \frac{\beta^2}{2\alpha^2}} \lim_{t\to\infty} E^x \left[e^{-\frac{\beta^2 t}{\alpha} + \sqrt{\frac{2\beta^2}{\alpha}} W_t} \right]. & (4.108) \end{aligned}$$

From classical stochastic calculus it is easy to see that

$$E^x \left[e^{-\frac{\beta^2 t}{\alpha} + \sqrt{\frac{2\beta^2}{\alpha}} W_t} \right] = 1 \quad \forall\, t \geq 0. \tag{4.109}$$

And so we finally obtain

$$H(x,\infty) = e^{\frac{\beta}{\alpha}(x-1) - \frac{\beta^2}{2\alpha^2}} = e^{-\frac{\beta}{\alpha}} \sum_{n=0}^{\infty} \frac{h_n(x)}{n!} \left(\frac{\beta}{\alpha}\right)^n \tag{4.110}$$

from which the stationary distribution for the Markov process is

$$P_n(\infty) = \frac{e^{-\frac{\beta}{\alpha}}}{n!}\left(\frac{\beta}{\alpha}\right)^n$$

as can be also state from direct computations with the Markov process.

Chapter 5

Itô vs. Stratonovich Revisited

From our previous study on Doi-Peliti Field theory and duality we have found several examples of chemical reactions of the form (4.8) that are dual to an Itô equation

$$dX_t = f(X_t)dt + \sqrt{2g(X_t)}dW_t \tag{5.1}$$

where $f(x)$ and $g(x)$ are polynomials. In particular, for the coalescence reaction, that we have studied in section (4.2.1), the coefficients are

$$\begin{aligned} f(x) &= 0 \\ g(x) &= \frac{\lambda}{2}x(1-x), \end{aligned}$$

i.e.,

$$dX_t = \sqrt{\lambda\left(X_t - X_t^2\right)}dW_t. \tag{5.2}$$

In our previous study we were mainly interested in the long term dynamics rather than in the solution at all times. Thinking in the methods one ought to try for solving these kind of reactions one easily first think in transforming the Itô equation into a Stratonovich one where the usual rules of calculus apply.

The formal Stratonovich analogous of equation (5.2) is

$$dX_t = -\frac{\lambda}{4}\left(1 - 2X_t\right)dt + \sqrt{\lambda\left(X_t - X_t^2\right)} \circ dW_t. \tag{5.3}$$

But here we can observe what seems to be a contradiction:

1. Equation (5.2) does have $X_t = 0$ and $X_t = 1$ as constant solutions, as $\{0, 1\}$ are absorbing states.
2. Equation (5.3) does not have $X_t = 0$ or $X_t = 1$ as constant solutions, neither they are absorbing states.

In this chapter, we study the theory of Itô and Stratonovich interpretations applied to equations of the form (5.1), which opens again the seemingly endless discussion about Itô and Stratonovich interpretations in the context of modeling stochastic phenomena [55, 56].

For equations with absorbing states of a certain type; this transformation is not well defined even though the formal stochastic equation that we end up with is well defined in itself.

For a deeper look into the topics of this chapter see [57].

5.1 About Itô and Stratonovich Interchangeability

Stochastic differential equations can be defined as an extension of an ODE in which a term of 'Gaussian' noise (white noise) is added as an external influence to the deterministic ODE.

This intuition about an SDE was transformed into a rigorous theory by Itô and others from the 1950s. Still, there are some misunderstandings about the limitations of this rigorous theory and, particularly, in the limitations of the relation between Itô SDEs and Stratonovich SDEs.

In the intuitive picture about SDEs mentioned above, one could come up with the following formal definition of an SDE

$$dX_t = \mu(X_t, t)\, dt + \sigma(X_t, t)\, \xi_t, \tag{5.4}$$

where ξ_t is a random function satisfying $E\left[\xi_t \cdot \xi_s\right] = \delta(|t - s|)$.

The influence of the function $\sigma(x, t)$ is the main culprit of the complication around stochastic integration and, in particular, the culprit that the expression (5.4) is not uniquely defined unless additional conditions apply.

The reason for $\sigma(x, t)$ to have such a big influence on this picture is that it is the way X_t influences the stochastic behavior of its future self. If time were discrete then there would be no problem at all, but

the fact that time is continuous makes that depending on the grid definition of the SDE the Central Limit Theorem can bring unexpected extra terms (or not) due to change of variables, making that Itô SDEs do not follow the formal change of variables rules of usual ODEs. That is not the case for its Stratonovich counterpart [58], making it the favorite definition in certain domains as Physics.

The striking fact is that there is (when conditions apply!) a well-known relation between Itô SDE's and Stratonovich SDE's so one could use one of the two formulations that are more convenient for the particular problem that one is working on.

The Itô SDE

$$dX_t = \mu(X_t, t)\, dt + \sigma(X_t, t)\, dW_t, \tag{5.5}$$

is equivalent to the Stratonovich SDE

$$dX_t = \left[\mu(X_t) - \frac{1}{2}\sigma'(X_t)\sigma(X_t)\right] dt + \sigma(X_t) \circ dW_t, \tag{5.6}$$

with an analogous formula for the inverse transformation, under very general circumstances [55]. Indeed, it is the case for regular enough drift and diffusion terms. Nevertheless, to highlight the importance of the counterexamples, let us state this fact in a precise manner.

Theorem 5.1. *Let $\mu(\cdot,\cdot) : \mathbb{R}\times\mathbb{R}_+ \longrightarrow \mathbb{R}$ and $\sigma(\cdot,\cdot) : \mathbb{R}\times\mathbb{R}_+ \longrightarrow \mathbb{R}$ be two globally Lipschitz continuous functions such that $\sigma(\cdot,\cdot)$ is continuously differentiable with respect to the first variable and $\sigma'(\cdot,\cdot)\sigma(\cdot,\cdot)$ is also globally Lipschitz continuous, where the prime denotes differentiation with respect to the first variable. Then the Itô SDE*

$$dX_t = \mu(X_t, t)\, dt + \sigma(X_t, t)\, dW_t, \qquad X_0 = x_0 \in \mathbb{R},$$

and the Stratonovich SDE

$$dY_t = \left[\mu(Y_t, t) - \frac{1}{2}\sigma'(Y_t, t)\sigma(Y_t, t)\right] dt + \sigma(Y_t, t) \circ dW_t, \qquad Y_0 = x_0 \in \mathbb{R},$$

both possess a unique solution such that $X_t = Y_t$ for all $t \in [0, T]$ almost surely, for any $T > 0$. Correspondingly the Stratonovich SDE

$$dX_t = \mu(X_t, t)\, dt + \sigma(X_t, t) \circ dW_t, \qquad X_0 = x_0 \in \mathbb{R},$$

and the Itô SDE

$$dY_t = \left[\mu(Y_t,t) + \frac{1}{2}\sigma'(Y_t,t)\sigma(Y_t,t)\right]dt + \sigma(Y_t,t)\,dW_t, \qquad Y_0 = x_0 \in \mathbb{R},$$

both possess a unique solution such that $X_t = Y_t$ for all $t \in [0,T]$ almost surely, for any $T > 0$.

This statement and its corresponding proof can be essentially found in [59].

Things become interesting at this point if we realized that equations that appeared naturally in our study of stochastic duality like (5.2) do not follow the previous theorem, as we discussed in the introduction of the chapter making that the condition '$\sigma(x,t)$ globally Lipschitz' cannot be relaxed from the previous theorem!

Theorem 5.2. *Let $f(\cdot,\cdot) : \mathbb{R}\times\mathbb{R}_+ \longrightarrow \mathbb{R}$ and $g(\cdot,\cdot) : \mathbb{R}\times\mathbb{R}_+ \longrightarrow \mathbb{R}_+$ be two continuous functions such that $g'(\cdot,\cdot)$ is also continuous, where the prime denotes differentiation with respect to the first argument. Assume also that $f(x_0,t) = g(x_0,t) = 0$ for all $t \geq 0$ and $g'(x_0,0) \neq 0$ for some $x_0 \in \mathbb{R}$. Then the Itô SDE*

$$dX_t = f(X_t,t)\,dt + \sqrt{2g(X_t,t)}\,dW_t, \qquad X_0 = x_0,$$

admits the trivial solution $X_0 = x_0$, but its formal Stratonovich dual

$$dX_t = \left[f(X_t,t) - \frac{1}{2}g'(X_t,t)\right]dt + \sqrt{2g(X_t,t)}\circ dW_t, \qquad X_0 = x_0,$$

does not admit such a solution in the interval $[0,T]$ for any $T > 0$. Correspondingly the Stratonovich SDE

$$dX_t = f(X_t,t)\,dt + \sqrt{2g(X_t,t)}\circ dW_t, \qquad X_0 = x_0,$$

admits the trivial solution $X_0 = x_0$, but its formal Itô dual

$$dX_t = \left[f(X_t,t) + \frac{1}{2}g'(X_t,t)\right]dt + \sqrt{2g(X_t,t)}\,dW_t, \qquad X_0 = x_0,$$

does not admit such a solution in the interval $[0,T]$ for any $T > 0$.

Proof. We will proof explicitly the first affirmation as the proof of the second follows identically. The SDE

$$dX_t = f(X_t,t)\,dt + \sqrt{2g(X_t,t)}\,dW_t, \qquad X_0 = x_0,$$

actually means

$$X_t = x_0 + \int_0^t f(X_s, s)\, ds + \int_0^t \sqrt{2g(X_s, s)}\, dW_s, \qquad t \in [0, T].$$

Substituting $X_t = x_0$ we find

$$x_0 = x_0 + \int_0^t f(x_0, s)\, ds + \int_0^t \sqrt{2g(x_0, s)}\, dW_s = x_0, \qquad t \in [0, T],$$

so it is obviously a solution. On the other hand the SDE

$$dX_t = \left[f(X_t, t) - \frac{1}{2} g'(X_t, t) \right] dt + \sqrt{2g(X_t, t)} \circ dW_t, \qquad X_0 = x_0,$$

actually means

$$X_t = x_0 + \int_0^t \left[f(X_s, s) - \frac{1}{2} g'(X_s, s) \right] ds + \int_0^t \sqrt{2g(X_s, s)} \circ dW_s, \qquad t \in [0, T],$$

so substituting $X_t = x_0$ yields

$$x_0 = x_0 - \frac{1}{2} \int_0^t g'(x_0, s)\, ds \iff \int_0^t g'(x_0, s)\, ds = 0,$$

for every $t \in [0, T]$. Note that this integral is well-defined by continuity of g' and assume without loss of generality that $g'(x_0, 0) > 0$. Again by continuity we know that $g'(x_0, t) > 0$ for $t \in [0, \delta)$ for some $\delta > 0$ sufficiently small; then

$$\int_0^t g'(x_0, s)\, ds > 0 \quad \text{for every } t \in (0, \delta),$$

hence x_0 is not a solution in [0,T] for any $T > 0$. □

5.2 Dynamic Behaviour

In the previous section, we have seen that some specific constant solutions solve for an Itô SDE but not its Stratonovich counterpart (and the other way around), but we have not yet proved that the equations at hand are well-defined and exists at all times, that is the purpose of this section. Not only that, we calculate with precision when solutions to these SDEs converge to the constant solutions calculated before. We prove as well that the problematic behavior of these equations has to do with

the boundaries of the interval of definition.

For that, we first need to introduce some terminology.

Definition 5.1. *A boundary point in the interval of the definition of an SDE is called accessible if the probability of reaching that point in a finite time is positive.*

Definition 5.2. *An accessible boundary point $\{a\}$ in the interval of the definition of an SDE is called instantaneously reflecting if and only if the Lebesgue measure of the set $\{t : X_t = a\}$ is zero almost surely, is called slowly reflecting if and only if the Lebesgue measure of the set $\{t : X_t = a\}$ is finite almost surely, and is called reflecting if it is either instantaneously or slowly reflecting.*

We will be focusing on equations of the type (5.1) with initial conditions

$$dX_t = f(X_t)\,dt + \sqrt{2g(X_t)}\,dW_t, \qquad X_0 = \gamma_0 \in \mathbb{R},$$

that is, on time-homogeneous diffusions, but prior to analyzing their dynamical behavior, we need a result that precisely states their well-posedness.

Definition 5.3. *Let $h : [a, \infty) \longrightarrow U \subseteq \mathbb{R}$, with $a \in \mathbb{R}$. We say that $h \in C^2[a, \infty)$ whenever there exist an $\epsilon > 0$ and a function $h_\epsilon \in C^2(a - \epsilon, \infty)$ such that $h_\epsilon = h$ in $[a, \infty)$. Moreover we write $h \in BC^2[a, \infty)$ if the function h along with its first and second derivatives are bounded in its domain of definition, with the understanding that $h'(a) := h'_\epsilon(a)$ and $h''(a) := h''_\epsilon(a)$.*

Remark. *Although we assume that the drift and diffusion terms of the SDEs under consideration belong to $BC^2[a, \infty)$, this can be relaxed to assuming that their first and second derivatives are bounded, but the functions themselves just obey the linear growth condition (i.e., they are bounded by an affine function). All the results in this work still hold under this milder requirement, as it already forbids finite time blow-ups, see for instance [35].*

Lemma 5.1. *Let $f : [a, \infty) \longrightarrow \mathbb{R}$ and $g : [a, \infty) \longrightarrow \mathbb{R}_+$, with $a \in \mathbb{R}$, and $f, g \in BC^2[a, \infty)$. Assume that $g(x) = 0$ if and only if $x = a$, with $g'(a) \neq 0$, and that there exist constants $C, \delta > 0$ such that $g(x) \geq C$ for $x \geq a + \delta$. Additionally assume that the function f satisfies the compatibility condition $f(a) \geq 0$. Then, the Itô SDE*

$$dX_t = f(X_t)\,dt + \sqrt{2g(X_t)}\,dW_t, \qquad X_0 = x_0 \in [a, \infty), \tag{5.7}$$

possesses a unique strong solution for all $t \geq 0$.

Proof. First observe that $f, g \in BC^2[a, \infty)$ implies they are globally Lipschitz and satisfy the linear growth condition. This last condition is also satisfied by $\sqrt{g}$, and moreover, if $x, y \geq a + \delta$ then

$$\left|\sqrt{g(x)} - \sqrt{g(y)}\right| \leq \sup_{z \geq a+\delta} \left\{ \frac{g'(z)}{2\sqrt{g(z)}} \right\} |x - y| .$$

This inequality implies that $\sqrt{g}$ is globally Lipschitz for $x \geq a + \delta$. So the lapse of existence of any local in time solution, if it exists, can be arbitrarily extended, see for instance [35].

Now, the string of inequalities

$$\begin{aligned} \left|\sqrt{x} - \sqrt{y}\right|^2 &= x + y - 2\sqrt{xy} \\ &\leq x + y - 2\min\{x, y\} \\ &\leq \max\{x, y\} - \min\{x, y\} \\ &= |x - y| \end{aligned}$$

implies

$$\left|\sqrt{g(x)} - \sqrt{g(y)}\right| \leq \sqrt{|g(x) - g(y)|}.$$

These results, together with the classical theorem of Watanabe and Yamada [60, 11, 35, 52, 53], imply existence and uniqueness of a strong solution to the SDE (5.7), which is defined for all times $t > 0$. □

Remark. *We are implicitly assuming all over this work that $g \geq 0$ so all the Itô diffusions we consider are real-valued. In particular, this implies $g'(a) > 0$.*

Remark. *An analogous result holds mutatis mutandis if the domain of definition of f and g is shifted to either $(-\infty, b]$ or $[a, b]$. For instance, the compatibility condition for the right boundary point should be $f(b) \leq 0$, i.e., the reflecting properties of the boundary should remain. From now on (unless explicitly indicated) we state our results just for the original case of domain $[a, \infty)$, but with the understanding that they still hold for any of the other two after simple modifications.*

The following theorem characterizes the boundary behavior of the solution to an Itô SDE of type (5.7) and compares it with the properties of its formal Stratonovich counterpart.

Theorem 5.3. *Let f and g be functions satisfying the same assumptions as in Proposition 5.1. The*

boundary $\{a\}$ is accessible for the Itô SDE

$$dX_t = f(X_t)\,dt + \sqrt{2g(X_t)}\,dW_t \tag{5.8}$$

if and only if $f(a) < g'(a)$, and absorbing if and only if $f(a) = 0$. On the other hand, its formal Stratonovich dual

$$dX_t = \left[f(X_t) - \frac{1}{2}g'(X_t)\right]dt + \sqrt{2g(X_t)} \circ dW_t \tag{5.9}$$

only admits the constant solution $X_t = a$ if $f(a) = \frac{1}{2}g'(a) < g'(a)$.

Remark. *This theorem implies that SDE* (5.9) *admits the constant solution $X_t = a$ only when SDE* (5.8) *has a reflecting boundary at a; respectively, when this boundary is absorbent, SDE* (5.9) *does not admit this constant solution.*

Remark. *We adopt the convention of denoting all constants by the letter C, independently of their specific value.*

Proof. The boundary behavior is studied using the method of speed/scale measures [60, 11, 61]. According to it, the boundary $\{a\}$ is accessible for equation (5.8) if and only if

$$\int_a^{a+\delta}\int_x^{a+\delta} \frac{1}{g(y)}\, e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx < \infty,$$

for some constant $\delta > 0$. We consider the cases $f(a) = 0$ (for which the inequality $f(a) < g'(a)$ is automatically guaranteed) and $f(a) > 0$ separately.

First we assume that $f(a) = 0$. Then by L'Hôpital rule we find

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)} < \infty.$$

So the integral $\int_x^y \frac{f(s)}{g(s)}\,ds$ is finite for any fixed δ since $y \le a + \delta$. Moreover we have the estimate

$$\begin{aligned}
\int_a^{a+\delta}\int_x^{a+\delta} \frac{1}{g(y)}\, e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx &< C\int_a^{a+\delta}\int_x^{a+\delta} \frac{dy\,dx}{g(y)} \\
&= C\int_a^{a+\delta}\int_a^{y} \frac{dx\,dy}{g(y)} \\
&= C\int_a^{a+\delta} \frac{y-a}{g(y)}\,dy < \infty,
\end{aligned}$$

where the first inequality comes from the boundedness of $\int_x^y \frac{f(s)}{g(s)}\,ds$ and the last one from the second application of L'Hôpital rule $\lim_{y\to a^+}\frac{y-a}{g(y)} = \frac{1}{g'(a)}$. Thus $\{a\}$ is an accessible boundary for $f(a) = 0$. Since $X_t = a$ is a solution to (5.8) in this case, then by uniqueness of solution we conclude this boundary is absorbing.

Now we turn to examining the finiteness of the integral

$$\int_a^{a+\delta}\int_x^{a+\delta}\frac{1}{g(y)}\,e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx$$

in the case $f(a) > 0$; for a small enough $\epsilon > 0$ (clearly it should be sufficiently smaller than δ) we find

$$\begin{aligned}
\int_{a+\epsilon}^{a+\delta}\int_x^{a+\delta}\frac{1}{g(y)}\,e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx &= \int_{a+\epsilon}^{a+\delta}\int_x^{a+\delta}\frac{1}{f(y)}\,\frac{f(y)}{g(y)}\,e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx\\
&\geq C\int_{a+\epsilon}^{a+\delta}\int_x^{a+\delta}\frac{f(y)}{g(y)}\,e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx\\
&= C\int_{a+\epsilon}^{a+\delta}\int_x^{a+\delta}\frac{d}{dy}\,e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx\\
&= C\int_{a+\epsilon}^{a+\delta}\left(e^{\int_x^{a+\delta} \frac{f(s)}{g(s)}\,ds} - 1\right)dx\\
&= C\int_{a+\epsilon}^{a+\delta}\left(e^{\int_x^{a+\delta} \frac{(s-a)f(s)}{g(s)}\frac{ds}{s-a}}\right)dx - \delta\\
&= C\int_{a+\epsilon}^{a+\delta}\left(e^{\int_x^{a+\delta} \frac{(s-a)f(a)}{g(s)}\frac{ds}{s-a}+\int_x^{a+\delta}\frac{(s-a)[f(s)-f(a)]}{g(s)}\frac{ds}{s-a}}\right)dx - \delta\\
&\geq C\int_{a+\epsilon}^{a+\delta}\left(e^{\int_x^{a+\delta} \frac{(s-a)f(a)}{g(s)}\frac{ds}{s-a}}\right)dx - \delta,
\end{aligned}$$

where the first inequality comes from the sign of $f(a)$ by choosing δ small enough, and the second comes from the boundedness of $\int_x^{a+\delta}\frac{f(s)-f(a)}{g(s)}\,ds$ in the range $x \in [a, a+\delta]$ (i.e. the bound is uniform in ϵ) as can be seen by means of the application of L'Hôpital rule $\lim_{s\to a^+}\frac{f(s)-f(a)}{g(s)} = \frac{f'(a)}{g'(a)}$. Now use the Taylor expansion $g(s) = g'(a)(s-a) + \frac{1}{2}g''(\bar{a})(s-a)^2$, where $\bar{a} \in [a, a+\delta]$, and choose a δ small enough to find

$$\begin{aligned}
\int_{a+\epsilon}^{a+\delta}\left(e^{\int_x^{a+\delta} \frac{(s-a)f(a)}{g(s)}\frac{ds}{s-a}}\right)dx - \delta &\geq C\int_{a+\epsilon}^{a+\delta}\left(e^{\int_x^{a+\delta} \frac{f(a)}{g'(a)}\frac{ds}{s-a}}\right)dx - \delta\\
&= C\int_{a+\epsilon}^{a+\delta}\left(\frac{\delta}{x-a}\right)^{\frac{f(a)}{g'(a)}}dx - \delta;
\end{aligned}$$

so by taking the limit $\epsilon \to 0$ we establish

$$\begin{aligned}\int_a^{a+\delta}\int_x^{a+\delta} \frac{1}{g(y)} e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx &\geq C\int_a^{a+\delta}\left(\frac{\delta}{x-a}\right)^{\frac{f(a)}{g'(a)}}dx-\delta \\ &= \infty \iff f(a)\geq g'(a).\end{aligned}$$

Arguing analogously one also establishes

$$\begin{aligned}\int_a^{a+\delta}\int_x^{a+\delta} \frac{1}{g(y)} e^{\int_x^y \frac{f(s)}{g(s)}\,ds}\,dy\,dx &\leq C\int_a^{a+\delta}\left(\frac{\delta}{x-a}\right)^{\frac{f(a)}{g'(a)}}dx-\delta \\ &< \infty \iff f(a)< g'(a).\end{aligned}$$

Combining these two results we conclude that $\{a\}$ is accessible if and only if $f(a) < g'(a)$. Moreover, together with the compatibility condition $f(a) > 0$ and uniqueness of solution we find that the boundary is instantaneously reflecting in this case [62].

Finally, the part of the statement that corresponds to the Stratonovich SDE is evident. □

To further exemplify this theorem we reconsider equation (5.1). For any initial condition $x_0 \in [0, 1]$, according to Theorem 5.3 (see also Remark 5.2), the boundary points $\{0\}$ and $\{1\}$ are both accessible and absorbing. Therefore, by Markovianity, we conclude that $\lim_{t\to\infty} X_t \in \{0, 1\}$ a.s. On the other hand, the formal Stratonovich counterpart of this equation, i.e. equation (5.3), does not possess either of these solutions.

5.3 Mean Time of Absortion

From the previous section not only the Itô formulation possesses constant solutions that the formal 'equivalent' Stratonovich does not possess, it also shows us that all solutions converge to these constant solutions over time, but it may be that the average time to absorption is infinite and so that the real practical importance of the result is not so significant.

What we prove here is that is not the case, i.e., the average time of absorption to these constant solutions is finite. Lets take again equation (5.2)

$$dX_t = \sqrt{\lambda X_t(1-X_t)}\,dW_t, \qquad X_0 = x_0,$$

which falls under the assumptions of Theorems 5.2 and 5.3. Denote by $T(x_0)$ the average time to absorption; it obeys the partial differential equation

$$\frac{\lambda}{2}x_0(1-x_0)\,\frac{\partial^2 T}{\partial x_0^2} = -1$$

subject to the boundary conditions $T(0) = T(1) = 0$ [61]. The explicit solution to this boundary value problem is

$$T(x_0) = -\frac{2}{\lambda}\left[(1-x_0)\log(1-x_0) + x_0\log(x_0)\right],$$

which is clearly bounded uniformly in $x_0 \in [0,1]$. Therefore absorption has a well-defined time scale and consequently the asymptotic behavior mentioned in the previous section will be reached at intermediate times.

We can even prove a much more general statement.

Theorem 5.4. *Let f and g be functions as in Proposition 5.1, Remark 5.2, and Theorem 5.3. Then the solution to the Itô SDE*

$$dX_t = f(X_t)\,dt + \sqrt{2g(X_t)}\,dW_t, \qquad x_0 \in [a,b],$$

subject to a finite state space $[a,b]$ in which both boundary points are accessible, and at least one is absorbing and the other is either absorbing or reflecting, is absorbed in finite mean time.

Proof. By Proposition 5.1 we know there exists a unique and global strong solution X_t to this equation. Denote by $\mathcal{A}$ the set of absorbing boundary points that, according to our assumptions, could be either $\{a\}$, $\{b\}$, or $\{a,b\}$. For every initial condition $\gamma_0 \in [a,b]$ define

$$\epsilon_T(\gamma_0) := P_{\gamma_0}(X_T \in \mathcal{A}),$$

so $\epsilon_T(\gamma_0) > 0$ for every large enough $T > 0$ by assumption. Now assume $\{b\}$ is absorbing to find

$$\epsilon_T(\gamma_0) \geq P_{\gamma_0}(X_T = b) \geq P_{x_0}(X_T = b)$$

for any $x_0 \leq \gamma_0$ by the $t-$continuity and Markovianity of X_t [35, 10]. Analogously if $\{a\}$ is absorbing then

$$\epsilon_T(\gamma_0) \geq P_{\gamma_0}(X_T = a) \geq P_{x_0}(X_T = a)$$

for any $x_0 \geq \gamma_0$. From now on we consider a T large enough so $\min\{P_{x_0}(X_T = b), P_{x_0}(X_T = a)\} > 0$ for every given initial condition x_0, which is always possible by the assumption on the accessibility of both boundary points.

For the time being let us assume that $\mathcal{A} \equiv \{a, b\}$. Now fix some c, $a < c < b$, to find

$$\begin{aligned}
\inf_{x_0 \in [a,b]} P_{x_0}(X_T \in \mathcal{A}) &= \min\left\{ \inf_{x_0 \in [a,c]} P_{x_0}(X_T \in \mathcal{A}), \inf_{x_0 \in [c,b]} P_{x_0}(X_T \in \mathcal{A}) \right\} \\
&\geq \min\left\{ \inf_{x_0 \in [a,c]} P_{x_0}(X_T = a), \inf_{x_0 \in [c,b]} P_{x_0}(X_T = b) \right\} \\
&\geq \min\left\{ P_c(X_T = a), P_c(X_T = b) \right\} > 0,
\end{aligned}$$

where we have employed, as in the previous paragraph, the Markovianity and $t-$continuity of X_t, and where a larger enough T has been selected, in case that had been necessary. Therefore we can define

$$\varepsilon_T := \inf_{\gamma_0 \in [a,b]} \epsilon_T(\gamma_0)$$

as a positive quantity. Note that the same conclusion arises analogously in the cases $\mathcal{A} \equiv \{a\}$ and $\mathcal{A} \equiv \{b\}$.

From now on we fix the value of T; then compute

$$\begin{aligned}
P\left(X_{(n+1)T} \in \mathcal{A}\right) &= P\left(X_{(n+1)T} \in \mathcal{A} \cap X_{nT} \in \mathcal{A}\right) + P\left(X_{(n+1)T} \in \mathcal{A} \cap X_{nT} \notin \mathcal{A}\right) \\
&= P\left(X_{(n+1)T} \in \mathcal{A} \middle| X_{nT} \in \mathcal{A}\right) P\left(X_{nT} \in \mathcal{A}\right) \\
&\quad + P\left(X_{(n+1)T} \in \mathcal{A} \middle| X_{nT} \notin \mathcal{A}\right) P\left(X_{nT} \notin \mathcal{A}\right) \\
&= P\left(X_{nT} \in \mathcal{A}\right) + P\left(X_{(n+1)T} \in \mathcal{A} \middle| X_{nT} \notin \mathcal{A}\right) \left[1 - P\left(X_{nT} \in \mathcal{A}\right)\right] \\
&\geq P\left(X_{nT} \in \mathcal{A}\right) + \varepsilon_T \left[1 - P\left(X_{nT} \in \mathcal{A}\right)\right]
\end{aligned}$$

for $n = 1, 2, \cdots$, where the inequality follows from the Markov property of X_t, and obviously $P\left(X_T \in \mathcal{A}\right) \geq \varepsilon_T$. Compare this inequality with the recursion relation

$$\begin{aligned}
U(n+1) &= U(n) + \varepsilon_T[1 - U(n)] \\
U(1) &= \varepsilon_T
\end{aligned}$$

that can be solved to yield

$$U(n) = 1 - (1 - \varepsilon_T)^n$$

for all $n \geq 1$. Clearly $P(X_T \in \mathcal{A}) \geq U(1)$ and by induction

$$\begin{aligned} P\left(X_{(n+1)T} \in \mathcal{A}\right) &\geq P(X_{nT} \in \mathcal{A}) + \varepsilon_T \left[1 - P(X_{nT} \in \mathcal{A})\right] \\ &= \varepsilon_T + P(X_{nT} \in \mathcal{A})\left[1 - \varepsilon_T\right] \\ &\geq \varepsilon_T + U(n)\left[1 - \varepsilon_T\right] \\ &= U(n) + \varepsilon_T \left[1 - U(n)\right] \\ &= U(n+1), \end{aligned}$$

therefore

$$P(X_{nT} \in \mathcal{A}) \geq 1 - (1 - \varepsilon_T)^n \longrightarrow 1$$

when $n \to \infty$, so adsorption happens almost surely in the long time limit.

Now define

$$\tau_{x_0}(\omega) := \inf\left\{t \geq 0 \,:\, (X_t | X_0 = x_0) \in \mathcal{A}\right\},$$

that is, the random variable $\tau_{x_0}(\omega)$ is the exit time from the set $\mathcal{A}^c$ or, in other words, the absorption time. We compute

$$\begin{aligned} \mathbb{E}(\tau_{x_0}) &= \int_0^\infty P(\tau_{x_0} > t)\, dt \\ &= \sum_{n=0}^{\infty} \int_{nT}^{(n+1)T} P(\tau_{x_0} > t)\, dt \\ &\leq T \sum_{n=0}^{\infty} P(\tau_{x_0} > nT) \\ &= T \sum_{n=0}^{\infty} P(X_{nT} \notin \mathcal{A}) \\ &= T \sum_{n=0}^{\infty} \left[1 - P(X_{nT} \in \mathcal{A})\right] \\ &\leq T \sum_{n=0}^{\infty} (1 - \varepsilon_T)^n \\ &= \frac{T}{\varepsilon_T} < \infty \end{aligned}$$

for each $x_0 \in [a, b]$, so the statement follows. □

From this, it is clear that the interchangeability between the Itô and Stratonovich formulations is not valid in all cases. Moreover, it is not valid for several of the SDEs arising from the study of Doi-Peliti field theory and duality applied to chemical reactions, as the ones studied in previous chapters.

This kind of equations also appears in another kind of fields where the results proved here still applies.

The spatially extended version (i.e., the stochastic partial differential equation version) of the equation

$$dX_t = \sqrt{X_t}\, dW_t,$$

i.e. of the Feller branching diffusion, has been used to model the growth and motion of plankton populations [63]; for related works (and in turn related to the critical Galton-Watson process) see [64, 65]. This equation has also been used in mathematical finance [60]; for an extension of this model see [62]. The equation

$$dX_t = (\alpha - \beta X_t)\, dt + \sqrt{X_t(1 - X_t)}\, dW_t,$$

with $\beta \geq \alpha \geq 0$ has been considered in population genetics [66]; for the case $\alpha = \beta = 0$ one can see [67]. The SDE

$$dX_t = \gamma X_t(1 - X_t)\, dt + \sqrt{X_t(1 - X_t)}\, dW_t,$$

with $\gamma > 0$, has been used to describe reaction processes, as well as its partial differential version has been used to describe reaction-diffusion processes [68]; related equations, both in the ordinary and partial stochastic differential setting, can be found in [5]. It is also remarkable that some spatially extended versions of this equation appear in the field of high energy physics [69, 70].

Chapter 6

Coalescence: How to Tackle a Complex SDE

In this chapter, we are going to study the coalescence reaction and its relation with complex noise in the context of SDEs.

As we saw in Section (3.3.3), the coalescence reaction is defined, using the notation given by abstract chemical reactions, as

$$A + A \xrightarrow{\lambda} \emptyset. \tag{6.1}$$

As a quick reminder, the generating function representation of the previous Markov process can be written as

$$\frac{\partial G}{\partial t} = \frac{\lambda}{2}\left(1 - x^2\right)\frac{\partial^2 G}{\partial x^2}. \tag{6.2}$$

The most striking fact about this reaction, is that the coherent transform $\Psi(\phi, t)$ possesses negative 'diffusion' [1]

$$\frac{1}{\lambda}\frac{\partial \Psi}{\partial t} = \frac{\partial}{\partial \phi}(\phi^2 \Psi) - \frac{1}{2}\frac{\partial^2}{\partial \phi^2}(\phi^2 \Psi), \tag{6.3}$$

where $\Psi(\phi, t)$ is defined as

$$G(x, t) = \int_0^\infty e^{\phi(x-1)}\Psi(\phi, t)\, d\phi. \tag{6.4}$$

[1] Technically speaking, a diffusion coefficient must the positive by definition.

Which means that, formally, the SDE [2]

$$d\phi = -\phi^2 dt + i\phi dW, \tag{6.5}$$

is 'equivalent' to the Markov process defined via (6.1).

This chapter aims to provide some answer to the following questions:

1. What is the meaning of reaction (6.5) and how does it relate to the PDE (6.3)?
2. What is the precise meaning of equivalence, if any, between reactions (6.1) and (6.5)?

For a deeper study on the topics of this chapter see [71]

6.1 Equivalence of moments

The first fact concerning the equivalence between (6.1) and (6.5) is that, at least formally, they both possess the same system of ODEs for the moments. This can be seen making use of Itô's rule, which can be written very easily in our context.

In particular, the Itô SDE (6.5) is equivalent to the two-real dimensional system of SDEs

$$\begin{pmatrix} d\phi_1 \\ d\phi_2 \end{pmatrix} = -\begin{pmatrix} \phi_1^2 - \phi_2^2 \\ 2\phi_1\phi_2 \end{pmatrix} dt + \begin{pmatrix} -\phi_2 & 0 \\ \phi_1 & 0 \end{pmatrix} \begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix}, \tag{6.6}$$

and, in this context the Itô's rule is clear. We claim that

Theorem 6.1. *Applying Itô's rule for an holomorphic function $f = f_1 + if_2 = (f_1, f_2)$ in* (6.5) *formally and in* (6.6) *in the usual $\mathbb{R}^2$ Itô's rule, coincide.*

Proof. It's important to remember that:

$$\begin{cases} \dfrac{\partial}{\partial\phi} = \dfrac{1}{2}\left(\dfrac{\partial}{\partial\phi_1} - i\dfrac{\partial}{\partial\phi_2}\right) \\ \dfrac{\partial}{\partial\overline{\phi}} = \dfrac{1}{2}\left(\dfrac{\partial}{\partial\phi_1} + i\dfrac{\partial}{\partial\phi_2}\right) \end{cases}$$

[2] In this section, we are always re-scaling time so that λ will not appear explicitly in our equations

The fact that f is holomorphic, which implies $\frac{\partial f}{\partial \overline{\phi}} = 0$, along with the definitions of $\frac{\partial f}{\partial \phi}$ and $\frac{\partial f}{\partial \overline{\phi}}$, yield the following identities

$$\begin{aligned} \frac{\partial f}{\partial \phi} &= \frac{\partial f}{\partial \phi_1} = -i\frac{\partial f}{\partial \phi_2}, \\ \frac{\partial^2 f}{\partial \phi^2} &= \frac{\partial^2 f}{\partial \phi_1^2} = -\frac{\partial^2 f}{\partial \phi_2^2} = -i\frac{\partial^2 f}{\partial \phi_1 \partial \phi_2}. \end{aligned}$$

On the other hand, the two dimensional Itô rule yields

$$\begin{aligned} df &= \left[-(\phi_1^2 - \phi_2^2)\frac{\partial f}{\partial \phi_1} - 2\phi_1\phi_2\frac{\partial f}{\partial \phi_2} + \frac{1}{2}\phi_2^2\frac{\partial^2 f}{\partial \phi_1^2} + \frac{1}{2}\phi_1^2\frac{\partial^2 f}{\partial \phi_2^2} - \phi_1\phi_2\frac{\partial^2 f}{\partial \phi_1 \phi_2}\right] dt \\ &\quad + \left[\phi_1\frac{\partial f}{\partial \phi_2} - \phi_2\frac{\partial f}{\partial \phi_1}\right] dW_t. \end{aligned}$$

Now, by substituting the previous expressions, we find

$$\begin{aligned} df &= \left[-(\phi_1^2 - \phi_2^2)\frac{\partial f}{\partial \phi} - i2\phi_1\phi_2\frac{\partial f}{\partial \phi} + \frac{1}{2}\phi_2^2\frac{\partial^2 f}{\partial \phi^2} - \frac{1}{2}\phi_1^2\frac{\partial^2 f}{\partial \phi^2} - i\phi_1\phi_2\frac{\partial^2 f}{\partial \phi^2}\right] dt \\ &\quad + \left[i\phi_1\frac{\partial f}{\partial \phi} - \phi_2\frac{\partial f}{\partial \phi}\right] dW_t \\ &= \left[-\phi^2\frac{\partial f}{\partial \phi} - \frac{1}{2}\phi^2\frac{\partial^2 f}{\partial \phi^2}\right] dt + i\phi\frac{\partial f}{\partial \phi} dW_t \\ &= -\phi^2\left[\frac{\partial f}{\partial \phi} + \frac{1}{2}\frac{\partial^2 f}{\partial \phi^2}\right] dt + i\phi\frac{\partial f}{\partial \phi} dW_t. \end{aligned}$$

□

Theorem 6.2. *The moments for the Markov process* (6.1) *and the complex SDE* (6.5) *coincide, in particular, both follow the system of ODEs*

$$\frac{1}{\lambda}\frac{d}{dt}\mathcal{M}_m = -\frac{m(m-1)}{2}\mathcal{M}_m - m\mathcal{M}_{m+1} \tag{6.7}$$

Proof. For the case of the Markov process (6.1) we know from (3.6) that its moments follow the sytem (6.7).

For the case of the complex SDE (6.5) we just need to apply formally Itô's rule to the function $f(\phi) = \phi^m$ to end up with

$$\frac{d}{dt}E\left[\phi^m\right] = -\frac{m(m-1)}{2}E\left[\phi^m\right] - mE\left[\phi^{m+1}\right], \tag{6.8}$$

which is the same system of ODEs as (6.7) □

6.2 Imaginary Noise is Not Unreal

Guided by the equivalence between moments the of the Markov process (6.1) and the complex noise Itô's equation (6.5), we are going to explore the relation between complex noise Itô's equations and their formally equivalent Fokker-Plank equations.

We see that such a relation exist, and we do this using the Cauchy transform that is a representation of certain space of distributions to a space of holomorphic functions.

The complex Itô SDE is not but a $\mathbb{R}^2$ Stochastic process that has its correspondent and well defined already, Fokker-Planck or Kolmogorov PDE associated with the usual properties. One could also formally talk about the 1-dimensional negative diffusion 'Fokker-Plank' equation, but this connection has not been proved yet.

They are three ingredients of the same underlying mathematical concept, and these formulations are equivalent to each other in a way that we precise in this Section.

The purpose of this Section is to analyze the particular reaction (6.5) and by no means, we try to develop a general theory, but we believe the results provided here can be generalized in a way that covers a great part of negative diffusion PDEs. In particular, we write the results with a notation as general as possible, but certain very restricting conditions are inspired by the particular properties of (6.5).

We will use the following notation for a general 1-dimensional negative diffusion PDE, that in particular covers our main case (6.3):

$$\frac{\partial \Psi}{\partial t} = -\frac{\partial}{\partial \phi}\left\{A(\phi)\Psi\right\} - \frac{1}{2}\frac{\partial^2}{\partial \phi^2}\left\{D^2(\phi)\Psi\right\}, \tag{6.9}$$

where $A(\phi)$ and $D(\phi)$ will be polynomials w.r.t. ϕ.

The first observation is that one should not pose such an equation in any classical space of differentiable functions, as the tendency of these PDEs is to concentrate probability in small regions giving rise to explosions and singularities, so one should restrict to spaces of distributions for posing this type of PDEs.

One convenient way to work with these type of equations involving distributions is to study their equivalent complex analytic representations [72].

Let us denote by $\mathcal{H}(\cdot)$ the class of holomorphic functions on a given domain. Then the following result holds (see Theorem 2.2.10 in [72])

Theorem 6.3 (Analytic representation of distributions)**.** *For every $\Psi \in C_c^\infty(\mathbb{R})'$ there exists a $\{\Psi\}_a \in \mathcal{H}(\mathbb{C}\backslash\mathbb{R})$ such that for all $f \in C_c^\infty(\mathbb{R})$,*

$$\langle \Psi | f \rangle = \lim_{\phi_2 \to 0^+} \int_{-\infty}^{\infty} \left(\{\Psi\}_a (\phi_1 + i\phi_2) - \{\Psi\}_a (\phi_1 - i\phi_2) \right) f(\phi_1) d\phi_1,$$

where $\langle \cdot | \cdot \rangle$ represents the duality product between $C_c^\infty(\mathbb{R})$ and $C_c^\infty(\mathbb{R})'$.

Note that this representation is not unique as any $\{\Psi\}_a \in \mathcal{H}(\mathbb{C})$ leads to the trivial distribution [73]. We focus now on the following analytical representation.

Definition 6.1 (Cauchy representation)**.** *For every $\Psi \in C_c^\infty(\mathbb{R})'$ we define its Cauchy representation as*

$$\{\Psi\}_a (\phi) := \frac{1}{2\pi i} \left\langle \Psi(\cdot) \middle| \frac{-1}{\phi - \cdot} \right\rangle,$$

where $\phi \in \mathbb{C}$.

Remark. *It is possible to prove that $\{\Psi\}_a(\phi)$ is always well-defined in $\mathcal{H}(\mathbb{C}\backslash\mathbb{R})$ [72]. Whenever $\Psi \in C_c(\mathbb{R})$ its Cauchy representation can be written as the integral*

$$\{\Psi\}_a (\phi) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{\Psi(s)}{s - \phi} ds.$$

From now on we will employ this integral notation with the understanding that the actual meaning is the duality product in Definition 6.1.

As an example of this transformation, it is easy to see from the definition that the Cauchy representation of the $n-$th derivative of the Dirac delta is:

$$\text{If} \quad \Psi = \delta_0^{(n)} \quad \text{then} \quad \{\Psi\}_a (\phi) = \frac{1}{2\pi i} \frac{(-1)^{n+1} n!}{\phi^{n+1}}.$$

Lemma 6.1. *Let $\{\Psi\}_a \in \mathcal{H}(\mathbb{C}\backslash\mathbb{R})$ be the Cauchy representation of a (compactly-supported) distribution $\Psi \in C_c^\infty(\mathbb{R})'$. Then:*

- $\{\mathcal{P}(s)\Psi(s)\}_a = \mathcal{P}(\phi)\{\Psi\}_a(\phi)$ *in* $\mathcal{H}(\mathbb{C}\backslash\mathbb{R})/\mathcal{H}(\mathbb{C})$ *and*
- $\left\{\dfrac{\partial^m \Psi(s)}{\partial s^m}\right\}_a = \dfrac{\partial^m \{\Psi(\phi)\}_a}{\partial \phi^m}$ *in* $\mathcal{H}(\mathbb{C}\backslash\mathbb{R})/\mathcal{H}(\mathbb{C})$,

where $\phi \in \mathbb{C}\backslash\mathbb{R}$, $s \in \mathbb{R}$, $\mathcal{P}$ *is an arbitrary polynomial, and* m *an arbitrary positive integer.*

Proof. It is clear that all expressions in the statement are well-defined. Now, to prove the first property note that it is enough to show that it is true for $\mathcal{P}$ being an arbitrary monomial, say ϕ^n. The case $n = 0$ is trivial, for $n = 1$ compute

$$\begin{aligned}
\phi\{\Psi\}_a(\phi) &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{\phi\Psi(s)}{s-\phi} ds \\
&= \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{(s-(s-\phi))\Psi(s)}{s-\phi} ds \\
&= \{s\Psi(s)\}_a - \underbrace{\frac{1}{2\pi i}\int_{-\infty}^{+\infty} \Psi(s) ds}_{\in\mathcal{H}(\mathbb{C})}.
\end{aligned}$$

The case $n > 1$ follows from the computation:

$$\begin{aligned}
\phi^n\{\Psi\}_a(\phi) &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{\phi^n\Psi(s)}{s-\phi} ds \\
&= \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{(s^n-(s^n-\phi^n))\Psi(s)}{s-\phi} ds \\
&= \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{s^n\Psi(s)}{s-\phi} ds - \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{s^n-\phi^n}{s-\phi}\,\Psi(s)\, ds \\
&= \{s^n\Psi(s)\}_a - \underbrace{\frac{1}{2\pi i}\sum_{m=0}^{n-1} \phi^{n-1-m}\int_{-\infty}^{+\infty} s^m\,\Psi(s)\, ds}_{\in\mathcal{H}(\mathbb{C})}.
\end{aligned}$$

To prove the second property we proceed by induction, commencing with the case $m = 1$:

$$\begin{aligned}
\frac{\partial}{\partial\phi}\{\Psi\}_a(\phi) &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \Psi(s)\frac{\partial}{\partial\phi}\frac{1}{s-\phi} ds \\
&= \frac{-1}{2\pi i}\int_{-\infty}^{+\infty} \Psi(s)\frac{\partial}{\partial s}\frac{1}{s-\phi} ds \\
&= \frac{1}{2\pi i}\int_{-\infty}^{+\infty} \frac{\Psi'(s)}{s-\phi} ds \\
&= \left\{\frac{\partial\Psi(s)}{\partial s}\right\}_a.
\end{aligned}$$

For the general case we use the induction hypothesis to find

$$\begin{aligned}\frac{\partial^{m+1}}{\partial\phi^{m+1}}\{\Psi\}_a(\phi) &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{\partial^m\Psi(s)}{\partial s^m}\frac{\partial}{\partial\phi}\frac{1}{s-\phi}ds \\ &= \frac{-1}{2\pi i}\int_{-\infty}^{+\infty}\frac{\partial^m\Psi(s)}{\partial s^m}\frac{\partial}{\partial s}\frac{1}{s-\phi}ds \\ &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{\partial^{m+1}\Psi(s)}{\partial s^{m+1}}\frac{1}{s-\phi}ds \\ &= \left\{\frac{\partial^{m+1}\Psi(s)}{\partial s^{m+1}}\right\}_a.\end{aligned}$$

□

Corollary 6.1. *Let $A(\cdot)$ and $D(\cdot)$ be polynomials and let Ψ be a $C^1([0,T],C_c^\infty(\mathbb{R})')$ solution, for some $T>0$, to*

$$\begin{cases}\dfrac{\partial\Psi}{\partial t} = -\dfrac{\partial}{\partial\phi}\left[A(\phi)\Psi\right] - \dfrac{1}{2}\dfrac{\partial^2}{\partial\phi^2}\left[D^2(\phi)\Psi\right] \\ \Psi(\phi,0) = \Psi_0(\phi), \quad \phi\in\mathbb{R};\end{cases}$$

then its Cauchy representation $\{\Psi\}_a$ is a $C^1([0,T],\mathcal{H}(\mathbb{C}\backslash\mathbb{R})/\mathcal{H}(\mathbb{C}))$ solution to

$$\begin{cases}\dfrac{\partial\{\Psi\}_a}{\partial t} = -\dfrac{\partial}{\partial\phi}\left[A(\phi)\{\Psi\}_a\right] - \dfrac{1}{2}\dfrac{\partial^2}{\partial\phi^2}\left[D^2(\phi)\{\Psi\}_a\right] \\ \{\Psi\}_a(\phi,0) = \{\Psi_0\}_a(\phi), \quad \phi\in\mathbb{C}.\end{cases} \tag{6.10}$$

Proof. The proof of this Corollary is a direct consequence of Lemma 6.1. □

This last Corollary, allows us to see a negative diffusion PDE in two equivalent ways, that cannot be clarified just given the PDE itself, as their formulations are the same. With this point of view at hand, we will see the connection of the complex SDE with the complex PDE formulation instead of the original attempt of relating it with the distribution posed PDE.

The imaginary-noise SDE formally associated to problem (6.10) is

$$dz = A(z)dt + iD(z)dW. \tag{6.11}$$

And this SDE is in turn associated with the real two-dimensional Fokker-Planck equation

$$\frac{\partial}{\partial t}P = -\frac{\partial}{\partial z_1}\left[A_1P\right] - \frac{\partial}{\partial z_2}\left[A_2P\right] + \frac{1}{2}\frac{\partial^2}{\partial z_1^2}\left[D_2^2P\right] + \frac{1}{2}\frac{\partial^2}{\partial z_2^2}\left[D_1^2P\right] - \frac{\partial^2}{\partial z_1\partial z_2}\left[D_1D_2P\right]. \tag{6.12}$$

Our main result, connect (6.10) with (6.11) through (6.16) and (6.15).

Theorem 6.4. *Let Ψ be a $C^1([0,T], C_c^\infty(\mathbb{R})')$ solution to*

$$\begin{cases} \dfrac{\partial \Psi}{\partial t} = -\dfrac{\partial}{\partial \phi}\left[A(\phi)\Psi\right] - \dfrac{1}{2}\dfrac{\partial^2}{\partial \phi^2}\left[D^2(\phi)\Psi\right] \\ \Psi(\phi,0) = \Psi_0(\phi), \quad \phi \in \mathbb{R}. \end{cases} \tag{6.13}$$

Its Cauchy representation can be expressed as

$$\{\Psi\}_a(\phi,t) = \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P(z_1,z_2,t)\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0, \tag{6.14}$$

if there exists a compactly-supported $C^1([0,T], \mathcal{M}(\mathbb{R}^2))$ solution P to

$$\begin{cases} \dfrac{\partial}{\partial t}P = -\dfrac{\partial}{\partial z_1}\left[A_1 P\right] - \dfrac{\partial}{\partial z_2}\left[A_2 P\right] + \dfrac{1}{2}\dfrac{\partial^2}{\partial z_1^2}\left[D_2^2 P\right] + \dfrac{1}{2}\dfrac{\partial^2}{\partial z_2^2}\left[D_1^2 P\right] - \dfrac{\partial^2}{\partial z_1 \partial z_2}\left[D_1 D_2 P\right] \\ P(z_1,z_2,0) = \delta(z_1-x_0)\delta(z_2), \end{cases} \tag{6.15}$$

for $\phi \in \overline{supp(P)\cup\mathbb{R}}^c$, where $A_1(z_1,z_2) = \Re[A(z_1+iz_2)]$, $A_2(z_1,z_2) = \Im[A(z_1+iz_2)]$, $D_1(z_1,z_2) = \Re[D(z_1+iz_2)]$, and $D_2(z_1,z_2) = \Im[D(z_1+iz_2)]$.

Proof. We start with the initial condition:

$$\begin{aligned} \Psi(\phi,0) &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P(z_1,z_2,0)\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0 \\ &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \delta(z_1-x_0)\delta(z_2)\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0 \\ &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{\Psi_0(x_0)}{x_0-\phi}dx_0. \end{aligned}$$

Using the relations

$$\frac{\partial}{\partial z_1}\frac{1}{z_1+iz_2-\phi} = -\frac{\partial}{\partial \phi}\frac{1}{z_1+iz_2-\phi}$$

and

$$i\frac{\partial}{\partial z_2}\frac{1}{z_1+iz_2-\phi} = \frac{\partial}{\partial \phi}\frac{1}{z_1+iz_2-\phi}$$

we find

$$
\begin{aligned}
\frac{\partial}{\partial t}\Psi(\phi,t) &= \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\partial_t P(z_1,z_2,t)\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0,\\
&= -\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\partial}{\partial z_1}\left[A_1P\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad -\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\partial}{\partial z_2}\left[A_2P\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad +\frac{1}{4\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\partial^2}{\partial z_1^2}\left[D_2^2P\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad +\frac{1}{4\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\partial^2}{\partial z_2^2}\left[D_1^2P\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad -\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\partial^2}{\partial z_1\partial z_2}\left[D_1D_2P\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&= -\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[A_1P\right]\frac{\partial}{\partial\phi}\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad -\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[iA_2P\right]\frac{\partial}{\partial\phi}\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad +\frac{1}{4\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[D_2^2P\right]\frac{\partial^2}{\partial\phi^2}\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad +\frac{1}{4\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[-D_1^2P\right]\frac{\partial^2}{\partial\phi^2}\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad -\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[iD_1D_2P\right]\frac{\partial^2}{\partial\phi^2}\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&= -\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[AP\right]\frac{\partial}{\partial\phi}\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad -\frac{1}{4\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[D^2P\right]\frac{\partial^2}{\partial\phi^2}\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&= -\frac{\partial}{\partial\phi}\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[AP\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
&\quad -\frac{1}{2}\frac{\partial^2}{\partial\phi^2}\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[D^2P\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0.
\end{aligned}
$$

We conclude noticing that by a similar argument to that in the proof of Proposition 6.1 we have

$$
\begin{aligned}
&\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[AP\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
-\ &\frac{A}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}P\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0\\
\in\ &\mathcal{H}(\mathbb{C})
\end{aligned}
$$

and

$$\begin{aligned} & \frac{1}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[D^2P\right]\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0 \\ - & \frac{D^2}{2\pi i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}P\frac{\Psi_0(x_0)}{z_1+iz_2-\phi}dz_1dz_2dx_0 \\ \in & \mathcal{H}(\mathbb{C}). \end{aligned}$$

□

6.2.1 Compactly Supported Initial Conditions

Given Theorem (6.4), one can see now that the complex SDE can be related to its formally complex PDE through the Cauchy transform, which is an interesting result by itself, but we started our study of equation (6.5) because of its connection with (6.1), in particular, we realize about the moment connection (6.2).

The purpose of this Section is to close this connection with moments to see why the generating function formulation was natural in this context.

We restrict ourselves to the case of compactly supported initial conditions for the Markov process (6.1), i. e. we assume that $P_n = 0$ for all $n > N$, where $N \in \mathbb{N}$ is arbitrarily large but fixed.

We present now the analog of Theorem 6.4 in the present context.

Theorem 6.5. *Let Ψ be a $C^1([0,T], C_c^\infty(\mathbb{R})')$ solution to*

$$\begin{cases} \dfrac{\partial\Psi}{\partial t} = -\dfrac{\partial}{\partial\phi}\left[A(\phi)\Psi\right] - \dfrac{1}{2}\dfrac{\partial^2}{\partial\phi^2}\left[D^2(\phi)\Psi\right] \\ \Psi(\phi,0) = \Psi_0(\phi), \quad \phi\in\mathbb{R}. \end{cases} \tag{6.16}$$

Its Cauchy representation can be expressed as

$$\{\Psi\}_a(\phi,t) = \frac{1}{2\pi i}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{\phi^{n+1}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(z_1+iz_2)^nP(z_1,z_2,t)\Psi_0(x_0)dz_1dz_2dx_0 \tag{6.17}$$

for $\phi \in \mathbb{C} \setminus \{0\}$. *Where* $P \in C^1([0,T], \mathcal{M}(\mathbb{R}^2))$ *is solution to*

$$\begin{cases} \dfrac{\partial}{\partial t}P = -\dfrac{\partial}{\partial z_1}\left[A_1 P\right] - \dfrac{\partial}{\partial z_2}\left[A_2 P\right] + \dfrac{1}{2}\dfrac{\partial^2}{\partial z_1^2}\left[D_2^2 P\right] + \dfrac{1}{2}\dfrac{\partial^2}{\partial z_2^2}\left[D_1^2 P\right] - \dfrac{\partial^2}{\partial z_1 \partial z_2}\left[D_1 D_2 P\right] \\ P(z_1, z_2, 0) = \delta(z_1 - x_0)\delta(z_2), \end{cases}$$

for $\phi \in \overline{supp(P) \cup \mathbb{R}}^c$, *where* $A_1(z_1, z_2) = \Re[A(z_1 + iz_2)]$, $A_2(z_1, z_2) = \Im[A(z_1 + iz_2)]$, $D_1(z_1, z_2) = \Re[D(z_1 + iz_2)]$, $D_2(z_1, z_2) = \Im[D(z_1 + iz_2)]$

Proof. Note that

$$\begin{aligned} & \frac{1}{2\pi i}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{\phi^{n+1}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(z_1 + iz_2)^n P(z_1, z_2, 0)\Psi_0(x_0)dz_1 dz_2 dx_0 \\ = & \frac{1}{2\pi i}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{\phi^{n+1}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(z_1 + iz_2)^n \delta(z_1 - x_0)\delta(z_2)\Psi_0(x_0)dz_1 dz_2 dx_0 \\ = & \frac{1}{2\pi i}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{\phi^{n+1}}\int_{-\infty}^{+\infty} x_0^n\,\Psi_0(x_0)dx_0 \\ = & \frac{1}{2\pi i}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{\phi^{n+1}}\mathcal{M}_n(0), \end{aligned}$$

where $\mathcal{M}_n(0)$ denote the moments of the distribution $\Psi_0(x_0)$. Consider now the (infinite-dimensional) vector

$$\begin{aligned} & \left\{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(z_1 + iz_2)^n P(z_1, z_2, t)\Psi_0(x_0)dz_1 dz_2 dx_0\right\}_{n=1}^{\infty} \\ = & \{\mathcal{D}_{nn}\}_{\mathbb{Z}\times\mathbb{Z}}\left\{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(z_1 + iz_2)^n P(z_1, z_2, 0)\Psi_0(x_0)dz_1 dz_2 dx_0\right\}_{n=1}^{\infty} \\ = & \{\mathcal{D}_{nn}\}_{\mathbb{Z}\times\mathbb{Z}}\left\{\int_{-\infty}^{+\infty} x_0^n\,\Psi_0(x_0)dx_0\right\}_{n=1}^{\infty} \\ = & \mathcal{M}_n(t), \end{aligned}$$

where $\{\mathcal{D}_{nn}\}_{\mathbb{Z}\times\mathbb{Z}}$ denotes the infinite-dimensional operator that describes the time evolution of the moment system of ODEs (6.7). Therefore

$$\{\Psi\}_a(\phi, t) = \frac{1}{2\pi i}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{\phi^{n+1}}\mathcal{M}_n(t), \tag{6.18}$$

for all $t \geq 0$, and thus

$$\Psi(\phi, t) = \sum_{n=0}^{\infty}\frac{\mathcal{M}_n(t)}{n!}\,\delta^{(n)}(\phi).$$

Now, by the definition of coherent state, conclude by computing $G(x,t)$ which is the generating function of (6.1)

$$\begin{aligned} G(x,t) &= \int_0^\infty \Psi(\phi,t)\, e^{\phi(x-1)} d\phi \\ &= \sum_{n=0}^\infty \frac{\mathcal{M}_n}{n!}\,(1-x)^n, \end{aligned}$$

so we recover our original generating function Taylor-expanded at $x=1$ instead of $x=0$ (note that, in the present case of compactly supported initial conditions, the generating function is simply a polynomial). □

Remark. *The formal connection between Theorems 6.4 and 6.5 comes from the identity*

$$\frac{1}{z-\phi} = \sum_{n=0}^\infty \frac{(-1)^{n+1}}{\phi^{n+1}}\, z^n,$$

which is valid for $|z| < |\phi|$.

Chapter 7

An Application to Biology: Stochastic Activation in a Genetic Switch Model

7.1 Introduction

In this chapter, we are going to apply the machinery of Martin-Siggia-Rose field theory to a stochastic system arising from cell biology, in which different chemicals act as part of the same genetic pathway.

As we studied in Section (3.4), Martin-Siggia-Rose is an alternative method to Doi-Peliti in the context of field-theoretic methods to model discrete stochastic systems. They are similar as there is a canonical transformation between the two; being specific to each use case when to apply each method. In this case, as we have a non-linear coupling that does not precisely fit in the typical combinatorial structure of Doi-Peliti, the Hamiltonian for the case of Martin-Siggia-Rose is much simpler to analyze and give us useful insights to important temporal scale information of the system.

The two chemical components that we are going to analyze are a protein and an mRNA, which production is influenced by a positive feedback loop. In particular, the concentration of proteins grows under the presence of mRNA; also, mRNA concentration grows under the presence of proteins, creating a positive feedback loop.

Typically, the influence of these chemicals is sigmoidal, that is, the rate of creation of mRNA grows in a non-linear sigmoidal way in the presence of protein, while protein concentration grows linearly relative to the concentration of mRNA. The concentration of each molecule can also degrade due to

complex mechanisms that occur in the cell.

Intuitively, this situation makes possible two static behaviors in the long run, either both concentrations are small, and the low rate of creation plus the biological degradation maintains this situation or both chemicals are in high concentration, and the high enhancement of chemical creation maintains the situation. We find out very quickly that this is the case, the system protein/mRNA is a bi-stable system.

Nevertheless, this system possesses an even more interesting component that in combination with the previous one makes the study of such an interaction not trivial at all: the existence of noise. The cell environment is a complex system in which both creation and degradation of molecules are non-deterministic. Additionally, the complex interactions with other parts of the system, that we do not study makes necessary the introduction of noise as an essential component of the evolution of such a system.

This stochastic influence translates, among other things, in the fact that random transitions between the two locally stable solutions exist, in which rare events influence the dynamics dramatically. How long does it take to have such a transition between these two locally stable situations? The answer to this question could have significant consequences in the behavior of a cell, and so this time scale can give us insights about cell behavior. A more extensive study of this situation can be found in [74]. For similar applications to biology some references are [75, 76, 77, 78, 79, 80, 81, 82, 83]

7.2 The Model

In our minimal stochastic model, we consider the transcription of a gene such that the production of mRNA is in a positive feedback loop with its own translated protein.

For that we define a Markov process for the probability

$$P(n, m, t) \equiv P_{n,m}(t) := \text{Probability of having n mRNAs and m proteins at time t.}$$

Let's define the rates of this Markov process:

- mRNA is created from the presence of protein molecules with a rate equal to the Hill function

$$g(y) = a + g_0 \frac{y^h}{y^h + K^h}, \tag{7.1}$$

  where h, a, g_0 are considered as free parameters. On the other hand, the degradation rate of mRNA molecule is constantly γ.

- Protein molecules get created with a rate linear to the number of mRNA molecules, with a scale factor c. The degradation rate is constantly 1, that is, we measure time in units of the protein degradation rate.

From the previous modeling assumptions the Master equation of the Markov process is

$$\begin{aligned} \frac{dP_{n,m}}{dt} &= \gamma(n+1)P_{n+1,m} - \gamma n P_{n,m} \\ &+ g(m)P_{n-1,m} - g(m)P_{n,m} \\ &+ cnP_{n,m-1} - cnP_{n,m} \\ &+ (m+1)P_{n,m+1} - mP_{n,m} \end{aligned} \tag{7.2}$$

Using now Martin-Siggia-Rose Field theory we can identify $P(n, m, t) = P(x, y, t)$ and use formally x, y as continuous variables with their correspondent momentum (derivatives) variables. Following the conventions of Section (3.4) we can write the Master equation as the linear 'PDE'

$$\frac{\partial P(x, y, t)}{\partial t} = H(x, p, y, q)P(x, y, t), \tag{7.3}$$

where

$$H(x, p, y, q) = (e^p - 1)\,\gamma x + \left(e^{-p} - 1\right) g(y) + (e^q - 1)\,y + \left(e^{-q} - 1\right) cx. \tag{7.4}$$

The reason for using Martin-Siggia-Rose Theory in this case, and not the previously studied Doi-Peliti Theory, is the convenience of writing non-linear rates like $g(y)$.

7.2.1 Deterministic Limit

We can consider the previous PDE as a stochastic perturbation around a deterministic model, that is, the linear approximation in p, q

$$H(x,p,y,q) = p(\gamma x - g(y)) + O(p^2) + q(y - cx) + O(q^2). \tag{7.5}$$

From which we derive that the deterministic PDE is equivalent to the system of ODEs

$$\dot{x} = g(y) - \gamma x, \tag{7.6}$$

$$\dot{y} = cx - y. \tag{7.7}$$

The stationary points of this ODE follow the system of equations

$$0 = g(y) - \gamma x, \tag{7.8}$$

$$0 = cx - y, \tag{7.9}$$

which can be re-written as

$$\begin{aligned} \frac{c}{\gamma} g(y) &= y, \\ x &= c^{-1} y. \end{aligned} \tag{7.10}$$

At this stage we define the ratio

$$\frac{c}{\gamma} = b, \tag{7.11}$$

which has a clear biological meaning, as b is the average number of proteins produced per mRNA lifetime. We will assume that the coefficients of our Hill function a, g_0 scale with b in such a way that the same solutions remain while varying b.

If we assume that $h > 1$, and a is big enough, there are exactly three solutions to the system (7.10) that we denote as $y_0 < y_1 < y_2$ with y_0, y_2 being the two attractors of the deterministic system. Analogously, we define $x_0 < x_1 < x_2$ solving for x in (7.10)

Nevertheless, the stochastic perturbation of our system makes that the solution transit between these two locally stable solutions if $1 \ll b \ll |y_{1,2} - y_0|$, analogously to the two-dimensional Kramers scape problem. Our goal is to study the rate of those transitions.

7.3 Bursting Limit

As a simplification of the problem at hand, and because of its biological relevance, we study the bursting limit of the system, that in our case consist on the limit

$$\gamma \longrightarrow \infty,$$

with b fixed. That is, the timescale of mRNA is 'accelerated' from the point of view of the lifetime of a protein, meaning that proteins are produced in bursts of b proteins (in average). This limit allows us to reduce our problem to an effective one-dimensional system in which only the protein concentration is relevant.

Theorem 7.1 (Bursting limit)**.** *In the bursting limit $\gamma \to \infty$ and b fixed, we can reduce the two dimensional Master equation* (7.2) *to the effective one dimensional Master equation*

$$\frac{dP_m(t)}{dt} = (m+1)P_{m+1} - mP_m - g(m)P_m + \frac{1}{1+b}\sum_{n=0}^{m}\left(\frac{b}{1+b}\right)^n g(m-n)P_{m-n}, \tag{7.12}$$

where P_m is the probability of having m proteins.

Remark. *That is, in the bursting limit we can see protein numbers as influencing themselves with a degradation rate of one and an exponential creation rate where more than one protein molecule can be created at once. This are the so-called exponential bursts.*

Proof. Taking our system of Master equations (7.2)

$$\begin{aligned}
\frac{dP_{n,m}}{dt} &= \gamma(n+1)P_{n+1,m} - \gamma nP_{n,m} \\
&+ g(m)P_{n-1,m} - g(m)P_{n,m} \\
&+ \gamma bnP_{n,m-1} - \gamma bnP_{n,m} \\
&+ (m+1)P_{n,m+1} - mP_{n,m}
\end{aligned}$$

we consider the asymptotic expansion

$$P_{n,m} = F_{n,m} + \gamma^{-1}S_{n,m} + O(\gamma^{-2}). \tag{7.13}$$

Which lead us to

$$\begin{aligned}\frac{dF_{n,m}}{dt} &= \gamma\left((n+1)F_{n+1,m} - nF_{n,m}\right) &+ &\left((n+1)S_{n+1,m} - nS_{n,m}\right)\\ &+ \quad g(m)\left(F_{n-1,m} - F_{n,m}\right)\\ &+ \quad \gamma bn\left(F_{n,m-1} - F_{n,m}\right) &+ &\; bn\left(S_{n,m-1} - S_{n,m}\right)\\ &+ \left((m+1)F_{n,m+1} - mF_{n,m}\right) &+ &\; O(\gamma^{-1}).\end{aligned} \tag{7.14}$$

The equation have to be fulfilled for each order separately. For order $O(\gamma)$ we find

$$F_{n+1,m} = (1+b)\frac{n}{n+1}F_{n,m} - b\frac{n}{n+1}F_{n,m-1} \tag{7.15}$$

for each $n, m \geq 0$, which implies $F_{n,m} = 0\ \forall n \geq 1$. We then simplify the notation to $F_{0,m} := P_m$.

Now for order $O(1)$ we can extract several equations, for $n = 0$

$$\frac{dP_m}{dt} = S_{1,m} - g(m)P_m + (m+1)P_{m+1} - mP_m. \tag{7.16}$$

For $n = 1$

$$S_{2,m} = \frac{1+b}{2}S_{1,m} - \frac{b}{2}S_{1,m-1} - \frac{g(m)}{2}P_m. \tag{7.17}$$

And, for $n \geq 2$

$$S_{n+1,m} = \frac{(1+b)n}{n+1}S_{n,m} - \frac{bn}{n+1}S_{n,m-1} \tag{7.18}$$

Summing for all m equations (7.16) and (7.17) we find

$$\sum_m S_{1,m} = \sum_m g(m)P_m, \tag{7.19}$$

and so

$$\sum_m S_{2,m} = 0. \tag{7.20}$$

As $F_{n,m} = 0\ \forall n \geq 1$ this implies $S_{n,m} \geq 0\ \forall n \geq 1$ as we are working with probabilities. Such a condition joined with the previous equality implies $S_{2,m} = 0\ \forall m$ which joined with (7.18) implies $S_{n,m} = 0\ \forall n \geq 2$ and (7.16) can be simplified to

$$S_{1,m} = \frac{b}{1+b}S_{1,m-1} + \frac{g(m)}{1+b}P_m, \tag{7.21}$$

which solution is

$$S_{1,m} = \frac{1}{1+b}\sum_{n=0}^{m}\left(\frac{b}{1+b}\right)^n g(m-n)P_{m-n}. \tag{7.22}$$

□

Using again Martin-Siggia-Rose Field Theory, the previous Master equation can be written as

$$\frac{dP(y,t)}{dt} = H_B(y,q)P(y,t), \tag{7.23}$$

where the bursting Hamiltonian is

$$H_B(y,q) = (e^{-q}-1)\left(\frac{bg(y)}{1+b(1-e^{-q})} - ye^{q}\right). \tag{7.24}$$

Having the Hamiltonian of the system, we can use the special structure that it has to calculate the stationary probability distribution easily

Lemma 7.1. *The system* (7.24) *has an unique stationary probability distribution given by*

$$P_\infty(y) = \frac{P_\infty(0)}{1+b}\prod_{k=0}^{y-1}\frac{b(k+g(k))}{(k+1)(1+b)}, \tag{7.25}$$

where $P_\infty(0)$ *is a (finite) constant that makes the sum of the series sum up to 1.*

Proof. By definition, the stationary distribution follows

$$H(y,q)P_\infty(y) = 0, \tag{7.26}$$

this implies

$$(1+b(1-e^{-q}))H(y,q)P_\infty(y) = (e^{-q}-1)\left[bg(y) - (1+b)e^{q}y + by\right]P_\infty(y) = 0, \tag{7.27}$$

which is the same as the following equation

$$bg(y)P_\infty(y) - (1+b)(y+1)P_\infty(y+1) + byP_\infty(y) = C \tag{7.28}$$

with C a constant. It is clear, making the sum in all $y's$ on both sides of the equation that the constant

should be $C = 0$ as otherwise we would have a finite constant (the stationary distribution has at least an exponential tail, so it has finite mean) at the left of the equation and infinity to the right. Giving this simplifications one can now easily derive the relation

$$\frac{P_\infty(y+1)}{P_\infty(y)} = \frac{b(g(y)+y)}{(1+b)(y+1)} \tag{7.29}$$

from where

$$P_\infty(y) = P_\infty(0)\left(\frac{b}{1+b}\right)^y \frac{g(0)(1+g(1))\cdots(y-1+g(y-1))}{y!} \tag{7.30}$$

and $P_\infty(0)$ should be the constant that makes the probabilities to sum up to 1. □

Remark. *As we are assuming that the equation $bg(y) = y$ has exactly three solutions it is clear that the stationary probability distribution has two local maximums and a local minimum giving by the closest integer solutions to the previous equation (for that we assume there are three well-separated natural numbers with that property). Which is something we already knew from the previous study of the deterministic system.*

7.4 WKB Approximation and Escape Rate

We are going to calculate the rate of escape for this system from the left meta-stable point y_0 to the right one y_2. The way we are going to calculate this rate is by using a WKB (Wentzel–Kramers–Brillouin) approximation for solving the quasi-stationary probability distribution with constant flux. The WKB approximation is tightly related to the theory of large deviations where one assumes there is some scale of the system in which the probability distribution behaves as an exponential dependence of a big parameter. One can use the Laplace method, and in that way, the symplectic nature of this kind of solutions appears naturally giving us one strong reason for treating the differential operator $H(y, q)$ as a Hamiltonian.

The goal is now to calculate the rate of escape η from y_0 to y_1 once we know that the stationary probability distribution is bi-stable.

The method we are using can be found in [84] and [76].

The Master equation (7.12) can be written in abstract as

$$\frac{\partial P(n,t)}{\partial t} = \sum_{k=-\infty}^{\infty} \left\{ \omega_{n-k,k} P(k,t) - \omega_{k-n,n} P(n,t) \right\}, \tag{7.31}$$

where, in our particular case,

$$\omega_{k,n} = \begin{cases} 0 & k \leq -2 \\ n & k = -1 \\ g(n)(1-r)r^k & k \geq 1 \end{cases} \tag{7.32}$$

are the transition rates, with $r = \dfrac{b}{1+b}$.

If we assume that y_0, y_1, y_2 are at least of order 10^2 to 10^3, as it is usual in biological situations, we can introduce a large parameter N of that order and define a reduced concentration variable $y = n/N$ of $O(1)$. The introduction of this factor is quite convenient in our case because we would like to derive an asymptotic expansion from applying the WKB method.

Rewriting our equation (7.31) in the new variable $y = n/N$, we have

$$\frac{\partial \Pi}{\partial t}(y,t) = \sum_{j=-\infty}^{\infty} N \left\{ \Omega_{-j}\left(y + \frac{j}{N}\right) \Pi\left(y + \frac{j}{N}\right) - \Omega_j(y)\,\Pi(y) \right\}, \tag{7.33}$$

with $\Pi(y,t) = P(yN,t)$ and

$$\Omega_k(y) = \begin{cases} 0 & k \leq -2 \\ y & k = -1 \\ N^{-1} g(yN)(1-r)r^k & k \geq 1. \end{cases} \tag{7.34}$$

We use a well-known technique called asymptotic matching that works in the following way:

- First, we approximate the quasi-stationary distribution near the unstable fixed point y_1 by solving the Fokker-Planck equation in this region, assuming a constant flux $J > 0$.
- Second, we use the WKB method between y_0 and y_1 in the region where the Fokker-Planck approximation is not valid.
- We then match these two solutions in the region where both are valid, enabling us to obtain a

formula for the rate of escape.

The smallest eigenvalue of the master equation is zero, and the next-smallest is exponentially small with respect to N. This second eigenvalue is the sum of the rates of the rare transitions from y_0 to y_2 and back. Under our present assumptions, the backward rate is exponentially weaker than the forward one, so this sum of rates is essentially the rate of escape from y_0.

We define the quasi-stationary probability distribution $\Pi(y)$ as the eigenfunction that corresponds to the second eigenvalue r_-, i.e. the rate of escape. The function $\Pi(y)e^{-r_- t}$ solves the Fokker-Planck approximation of (7.31) around y_1. We assume that the constant flux J through y_1 is re-injected into the mestastable well around y_0, so $\Pi(y)$ remains stationary. We then expand

$$0 = \sum_{j=-\infty}^{\infty} N \left\{ \Omega_{-j}\left(y + \frac{j}{N}\right) \Pi\left(y + \frac{j}{N}\right) - \Omega_j(y)\,\Pi(y) \right\} = \tag{7.35}$$

$$= \sum_{j=-\infty}^{\infty} N \left\{ \Omega_{j}\left(y - \frac{j}{N}\right) \Pi\left(y - \frac{j}{N}\right) - \Omega_j(y)\,\Pi(y) \right\}, \tag{7.36}$$

to second order in $1/N$:

$$0 = \sum_{k=-\infty}^{+\infty} N \left\{ \left(\Omega_j(y) - \frac{j}{N}\Omega_j'(y) + \frac{j^2}{2N^2}\Omega_j''(y) \right) \left(\Pi(y) - \frac{j}{N}\Pi'(y) + \frac{j^2}{2N^2}\Pi''(y) \right) - \Pi(y)\Omega_j(y) \right\}. \tag{7.37}$$

We collect the terms proportional to powers of $1/N$:

$$0 = -\sum_{k=-\infty}^{\infty} j \left\{ \Omega_j'(y)\Pi(y) + \Omega_j(y)\Pi'(y) \right\} + \frac{1}{2N} \sum_{k=-\infty}^{\infty} j^2 \left\{ \Omega_j''(y)\Pi(y) + 2\Omega_j'(y)\Pi'(y) + \Omega_j(y)\Pi''(y) \right\}. \tag{7.38}$$

This has the form of a stationary Fokker-Planck equation

$$0 = \frac{\partial}{\partial y}\left\{ A(y)\Pi(y) \right\} - \frac{1}{2N}\frac{\partial^2}{\partial y^2}\left\{ B(y)\Pi(y) \right\} \tag{7.39}$$

with

$$A(y) = \sum_{j=-\infty}^{\infty} j\Omega_j = bN^{-1}g(Ny) - y \tag{7.40}$$

$$B(y) = \sum_{j=-\infty}^{\infty} j^2\Omega_j = (1+2b)bN^{-1}g(Ny) + y. \tag{7.41}$$

The flux is

$$J = A(y)\Pi(y) - \frac{1}{2N}\frac{\partial}{\partial y}(B(y)\Pi(y)). \tag{7.42}$$

This can be solved for Π:

$$\Pi(y) = \frac{2JN}{B(y)} e^{2N\int_{y_1}^{y}\frac{A(z)}{B(z)}dz}\int_y^\infty e^{-2N\int_{y_1}^{z}\frac{A(\theta)}{B(\theta)}d\theta}dz \simeq \frac{JN}{y_1(1+b)}e^{\frac{(y-y_1)^2}{2\sigma^2}}\int_y^\infty e^{-\frac{(z-y_1)^2}{2\sigma^2}}dz, \tag{7.43}$$

with

$$\sigma^2 = \frac{y_1(1+b)}{N(bg'(Ny_1)-1)}. \tag{7.44}$$

We can simplify Π in the regime $y \ll y_1 - \sigma$

$$\Pi(y) = \frac{JN\sigma\sqrt{2\pi}}{y_1(1+b)}e^{\frac{(y-y_1)^2}{2\sigma^2}}. \tag{7.45}$$

Now we use the WKB method to obtain a solution valid in the range $y_0 < y < y_1 - \sigma$ that we will match with (7.45) in the region $y \ll y_1 - \sigma$, which will enable us to obtain J. The WKB *ansatz* is $\Pi = K(y)e^{-N\mathcal{W}(y)}$. With it we can expand:

$$\Pi\left(y \pm \frac{j}{N}\right) \simeq K\left(y \pm \frac{j}{N}\right)e^{-N\mathcal{W}\left(y\pm\frac{j}{N}\right)} \simeq \left\{K(y) \pm \frac{j}{N}K'(y)\right\}e^{-N\left[\mathcal{W}(y)\pm\frac{j}{N}\mathcal{W}'(y)+\frac{j^2}{2N^2}\mathcal{W}''(y)\right]}. \tag{7.46}$$

Now, expanding $e^{-N^{-1}\mathcal{W}''}$ in N^{-1}, we obtain, consistently to order $1/N$,

$$\Pi\left(y \pm \frac{j}{N}\right) \simeq \left\{K(y) \pm \frac{j}{N}K'(y)\right\}e^{-N\mathcal{W}(y)\mp j\mathcal{W}'(y)}\left\{1 - \frac{j^2}{2N}\mathcal{W}''(y)\right\}. \tag{7.47}$$

Also expanding Ω and using equation (7.35) we obtain

$$0 = \sum_{k=-\infty}^{\infty}\left\{\left(\Omega_j - \frac{j}{N}\Omega'\right)\left(K - \frac{j}{N}K'\right)\left(1 - \frac{j^2}{2N}\mathcal{W}''\right)e^{-N\mathcal{W}+j\mathcal{W}'} - \Omega_j K e^{-N\mathcal{W}}\right\}. \tag{7.48}$$

From the terms of $O(1)$ we have

$$0 = \sum_{j=-\infty}^{\infty}\Omega_j(y)\left(e^{j\mathcal{W}'(y)} - 1\right), \tag{7.49}$$

which can be interpreted as a stationary Hamilton-Jacobi equation for the Hamiltonian

$$H(y,q) = \sum_{j=-\infty}^{\infty} \Omega_j(y) \left(e^{jq} - 1\right) = (e^q - 1) \left(\frac{bN^{-1} g(Ny)}{1 - b(e^q - 1)} - y e^{-q} \right), \tag{7.50}$$

which coincides with (7.24). The $O(N^{-1})$ terms yield

$$0 = \frac{K'(y)}{K(y)} \sum_{j=-\infty}^{\infty} \left\{ j\Omega_j e^{j\mathcal{W}'} \right\} + \sum_{j=-\infty}^{\infty} \left\{ j\Omega_j' e^{j\mathcal{W}'} \right\} + \sum_{j=-\infty}^{\infty} \frac{j^2}{2} \left\{ \Omega_j \mathcal{W}'' e^{j\mathcal{W}'} \right\}. \tag{7.51}$$

The differential equation for K can be written in terms of derivatives of the Hamiltonian as

$$\frac{K'(y)}{K(y)} H_q(y, q(y)) = -\frac{1}{2} q'(y) H_{qq}(y, q(y)) - H_{qy}(y, q(y)), \tag{7.52}$$

with $q(y) = \mathcal{W}'(y)$, which solves $H(y, q(y)) = 0$. Thus, our WKB solution has the form

$$\Pi(y) = K(y) e^{-N\mathcal{W}(y)}, \tag{7.53}$$

where the action $\mathcal{W}$ is defined as $\mathcal{W} = \int_{y_0}^{y_1} q(y) dy$ and $K(y)$ solves(7.52).

To obtain K we just have to realize that $H(y, q(y)) = 0$ and take derivatives

$$H_q q' + H_y = 0 \tag{7.54}$$

and

$$H_{qq}(q')^2 + 2H_{qy}q' + H_q q'' + H_{yy} = 0, \tag{7.55}$$

from which

$$-q'(y) H_{qq}(y, q(y)) - 2H_{qy}(y, q(y)) = H_q \left(-\frac{H_{yy}}{H_y} + \frac{q''}{q'} \right). \tag{7.56}$$

Thus, the equation for the prefactor can be rewritten

$$2\frac{K'(y)}{K(y)} = \left(-\frac{H_{yy}}{H_y} + \frac{q''}{q'} \right). \tag{7.57}$$

The solution of this equation is

$$K(y) = \frac{C}{\sqrt{y(N^2 y + g(Ny))}}. \tag{7.58}$$

We also know that the flux is the escape rate times the population in the metastable well:

$$J = \eta \int_{-\infty}^{y_1} \Pi(y) dy. \tag{7.59}$$

Using equation (7.59) and the WKB solution, we can employ the Laplace method to obtain

$$\eta = \frac{J}{K(y_0)} \sqrt{\frac{N\mathcal{W}''(y_0)}{2\pi}} e^{N\mathcal{W}(y_0)}. \tag{7.60}$$

Approximating $K(y)\exp[-N\mathcal{W}(y)]$ by a second order Taylor expansion around y_1 and matching with (7.45), we obtain

$$\sigma^2 = \frac{1}{N\,|\mathcal{W}''(y_1)|} \tag{7.61}$$

$$J = \frac{K(y_1) y_1 (1+b) e^{-N\mathcal{W}(y_1)}}{\sigma\sqrt{2\pi}}. \tag{7.62}$$

Substituting this flux in equation (7.60) yields the rate

$$\eta = \frac{K(y_1)}{2\pi K(y_0)} y_1 (1+b) \sqrt{|\mathcal{W}''(y_0)\mathcal{W}''(y_1)|} e^{-N(\mathcal{W}(y_1)-\mathcal{W}(y_0))}. \tag{7.63}$$

From (7.58), it can be seen that $K(y_1)/K(y_0) = y_0/y_1 = y_0/y_1$. Also, because $\mathcal{W}'' = q'$,

$$\mathcal{W}''(y_{0,1}) = \frac{1}{1+b} \frac{1 - bg'(Ny_{0,1})}{y_{0,1}}, \tag{7.64}$$

so we can rewrite (7.63) as

$$\eta = \frac{1}{2\pi} \sqrt{\frac{y_0}{y_1}} \sqrt{(1 - bg'(y_0))\,|1 - bg'(y_1)|} e^{-N\Delta\mathcal{W}}, \tag{7.65}$$

with

$$\Delta W = \int_{y_0}^{y_1} q(y) dy = \int_{y_0}^{y_1} \log \frac{b}{1+b} \frac{N^{-1} g(Ny) + y}{y} dy \tag{7.66}$$

where

$$bN^{-1} g(Ny_{0,1}) = y_{0,1}. \tag{7.67}$$

Note that the calculation of the action ΔW can be done through the solution of the system of Hamilton

equations of the Hamiltonian (7.50) with the additional condition $H(y(t), q(t)) = 0$ giving

$$\Delta W = \int_{-\infty}^{\infty} q(t)\dot{y}(t)dt = \int_{y_0}^{y_1} q(y)dy. \tag{7.68}$$

This can be seen because the derivative W' follows the Hamilton-Jacobi equation. This analogy will be crucial to extend our methods for more than one dimension, as we will do in the next section.

7.4.1 Simulations

Using the Gillespie algorithm, we have simulated the bursting-limit model (7.12), measuring the average time $\overline{\tau}$ to reach the unstable point y_1. We take $1/(2\overline{\tau})$ as the empirical escape rate Γ_{sim}, where the factor of $1/2$ comes from the fact that a system at the unstable point has a probability of $1/2$ to leave it in either direction. We have done this for 13 values of b, from $b = 15$ to $b = 75$ in steps of $\Delta b = 5$. For each value of b we have simulated $10,000$ escape events. We find that the escape times are exponentially distributed for this range of b, giving an empirical error for the escape rate of 2% with confidence of 95%.

Throughout all our simulations, we are going to fix the values of the parameters on the Hill function as: $a = 13.33$, $g_0 = 116.6667$ and $h = 4$ and $K = 850$, for $b = 15$, thus preserving the fixed points $y_0 = 206.0185$ and $y_1 = 653.8648$. Whenever we vary the parameter b, we do it in such a way that preserve $bg(y)$, that is, we define $a = 13.33 \cdot 15/b$ and $g_0 = 116.6667 \cdot 15/b$ for an arbitrary b.

The results of our simulations (Fig. (7.1)) confirm the theoretical rate to a large extent. The empirical escape rate is slightly smaller (by a few percent) than the theoretical value. We attribute the discrepancy to the fact that our theory is only exact in the limit of infinitesimal escape rates. However, the simulation times necessary to confirm this quantitatively are prohibitively long.

7.5 The Full Problem: mRNA and Protein Dynamics

In this section we will tackle the study of the full system (7.3) with Hamiltonian (7.4). As we know from Section (7.2.1), this system possesses three stationary points (the solutions of (7.10)), two of them, (x_0, y_0) and (x_2, y_2), will be stable, and (x_1, y_1) will be unstable, with $x_0 < x_1 < x_2$ and $y_0 < y_1 < y_2$.

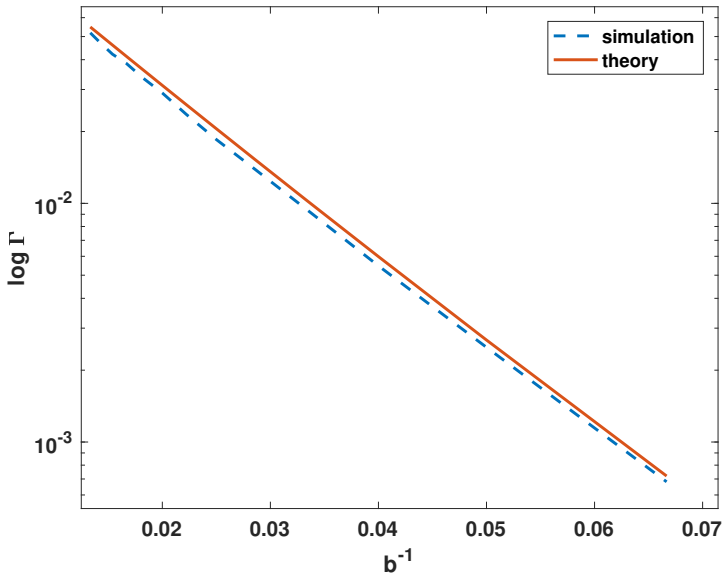


Figure 7.1: Rate of escape vs. b^{-1} for the bursting limit model: theory (red solid line) and measured in simulations (blue dashed line).

Reprinted figure with permission from [J.Hertz, J.Tyrcha, A.Correales,Stochastic activation in a genetic switch model, 98, 052403, 2018.] Copyright (2018) by the American Physical Society

Our goal is to calculate the escape rate η, from (x_0, y_0) to (x_2, y_2). For that, we use a similar procedure to the one used in the previous section when calculating the rate for the bursting limit case.

The calculation of the solution for this case is obviously much more involved, and we use the following assumption that is correct to a large extent: the rate of escape can be written as a perturbation around the parameter γ of the escape rate given by the bursting limit.

It is intuitive (though not totally trivial) that this should be the case when one is close to the limit $\gamma \to \infty$ but we see, through computational simulations, that it is the case even for values of γ of order $O(1)$.

We follow similar lines to the arguments used in the proof of the bursting limit (7.1) where we employed an asymptotic expansion around γ. Moreover, we do that in two separate stages, one when dealing with the exponent of the rate of escape and, in a later step, we do this for the prefactor.

7.5.1 Computing the Extremal Action

Analogously to the proof for the bursting limit case in section (7.3), one can use an argument involving the WKB approximation to see that the exponent of the exponential rate of escape is the solution to

the Hamilton-Jacobi equation

$$H(x, \frac{\partial S}{\partial x}, y, \frac{\partial S}{\partial y}) = 0, \tag{7.69}$$

which is,

$$S = \int_{-\infty}^{\infty} [p(t)\dot{x}(t) + q(t)\dot{y}(t)]dt, \tag{7.70}$$

where x, p, y, and q solve the Hamilton equations for the Hamiltonian of the system (7.4)

$$\dot{x} = \frac{\partial H}{\partial p} = g(y)\mathrm{e}^{p} - \gamma x \mathrm{e}^{-p} \tag{7.71}$$

$$\dot{p} = -\frac{\partial H}{\partial x} = \gamma(1 - \mathrm{e}^{-p}) - c(\mathrm{e}^{q} - 1), \tag{7.72}$$

$$\dot{y} = \frac{\partial H}{\partial q} = c x \mathrm{e}^{q} - y\mathrm{e}^{-q} \tag{7.73}$$

$$\dot{q} = -\frac{\partial H}{\partial y} = -g'(y)(\mathrm{e}^{p} - 1) + 1 - \mathrm{e}^{-q}, \tag{7.74}$$

with the condition that $H = 0$ along the extremal path and that the extremes of the path have to be stationary points for the Hamiltonian flow (stationary probability).

However, the condition $H = 0$ is insufficient in this 4-dimensional case to determine the x- and y-dependence of p and q. The only recourse is to integrate the Hamilton equations to find $x(t)$, $p(t)$, $y(t)$, and $q(t)$ and then use these in (7.70). Both steps have to be done numerically. We do this using the "relaxation" method described in Numerical Recipes [85].

Figs. (7.2) and (7.3) shows typical results for the solutions $x(t)$, $p(t)$, $y(t)$, and $q(t)$. The calculations here are for $\gamma = 4$, but the qualitative shape of the curves is the same for all γ. The mRNA and protein concentrations $x(t)$ and $y(t)$ are sigmoidal and the conjugate momenta $p(t)$ and $q(t)$ have single "bumps", beginning and ending at 0 for $t \to \pm\infty$. The protein curves lag the mRNA ones, as can be expected.

We have calculated the extremal action S for values of the mRNA degradation rate γ from 1 to 128, using a fixed burst-size parameter $b = 15$, so the translation rate $c = b\gamma$. The values of the parameters in the Hill function (7.1) were to be those used for $b = 15$ in the bursting-limit calculations described in subsection (7.4.1). As explained above, in the limit $\gamma \to \infty$ we recover the bursting model of the preceding section.

We examined the γ-dependence of the two terms in the extremal action (7.70). The qualitative features are quite simple: The protein action is relatively insensitive to γ, and its large-γ limit is consistent

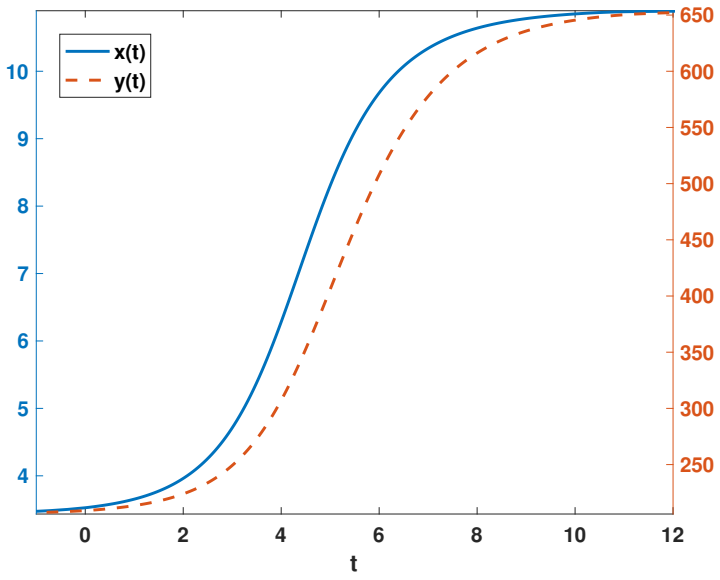


Figure 7.2: Typical concentration trajectories calculated using the relaxation method: mRNA concentration $x(t)$ (blue solid line) and protein concentration $y(t)$ (red dashed line).
Reprinted figure with permission from [J.Hertz, J.Tyrcha, A.Correales,Stochastic activation in a genetic switch model, 98, 052403, 2018.] Copyright (2018) by the American Physical Society

with the analytically calculated bursting-limit value. The mRNA action, on the other hand, decreases rapidly with increasing γ. At $\gamma = 1$ the two actions are about the same size, while at $\gamma = 128$ the mRNA term is two orders of magnitude smaller. The computed values are consistent with a $1/\gamma$ dependence over most of the range of γ for which the calculations were done.

7.5.2 Expanding Around the Bursting Limit

Using the results (A.15) and (A.18) from the Appendix (A), in which we make a complete calculation of the problem at hand, we realize that, in the limit of big γ, all the variables can be written in terms of $y(t)$, at least at order $O(\gamma^{-1})$. Meaning we can calculate an approximation to order $O(\gamma^{-1})$ of the action S

$$S \approx \int_{y_0}^{y_1} q(y)dy + \int_{y_0}^{y_1} p(y)\frac{dx}{dy}dy \tag{7.75}$$

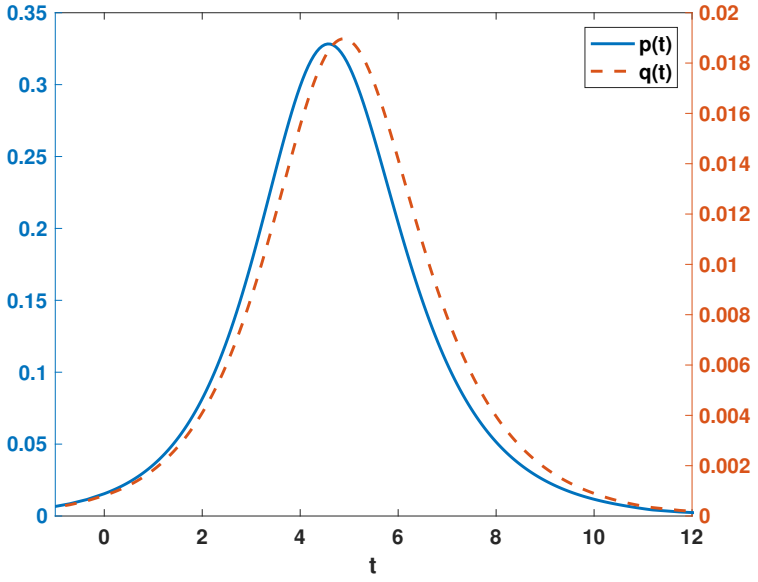


Figure 7.3: Typical momentum trajectories calculated using the relaxation method: mRNA momentum $p(t)$ (blue solid line) and protein momentum $q(t)$ red dashed line).

Reprinted figure with permission from [J.Hertz, J.Tyrcha, A.Correales,Stochastic activation in a genetic switch model, 98, 052403, 2018.] Copyright (2018) by the American Physical Society

For the first term, it is clear we obtain the limiting bursting limit rate [1]. For the second, and using (A.15) and (A.18)

$$\int_{y_0}^{y_1} p(y)\frac{dx}{dy}dy = \frac{1}{\gamma(1+b)^2}\int_{y_0}^{y_1} \log\left(\frac{1}{1+b}\frac{y+g(y)}{g(y)}\right)\frac{d}{dy}\frac{(y+g(y))^2}{g(y)}dy \tag{7.76}$$

We can even simplified a bit more the previous equation, taking integral by parts. Given the definition of $y_{0,1}$

$$bg(y_{0,1}) = y_{0,1}, \tag{7.77}$$

the boundary terms get eliminated and we can write

$$\int_{y_0}^{y_1} p(y)\frac{dx}{dy}dy = \frac{1}{\gamma(1+b)^2}\int_{y_0}^{y_1} \frac{(y+g(y))(g(y)-yg'(y))}{g(y)^2} \tag{7.78}$$

which is the $O(\gamma^{-1})$ error make when approaching the bursting limit at $\gamma \to \infty$. Our relaxation calculations for the numerical action can be now compared with this theoretical estimate and, quite remarkable, this term describes pretty well the behavior of the numerical case even for orders of γ of $O(1)$.

[1] It is clear to observe, given the set of relations, (A.15) and (A.18) that there is no order γ^{-1} terms appearing from this part of the integral, in particular the terms with presence of y_2 cancel each other out.

7.5.3 Prefactors

Like that of the action, the calculation of the prefactor for the full models is non-trivial and must be done numerically. A procedure based on a WKB scheme was set out by Maier and Stein [86] for Gaussian-noise models. It was extended to birth-death processes by Roma et al [87] for a model with mutual competition between two proteins and treated more generally by Bressloff [76].

In the appendix (A) we make the calculations more explicitly than here following the WKB procedure mentioned above. For the escape rate we get the formula

$$\eta = \frac{\lambda_+}{2\pi}\sqrt{\frac{\det(Z_0)}{|\det(Z_1)|}}\frac{K_1}{K_0}, \tag{7.79}$$

differing from the classical Eyring formula [88] in the presence of the last factor. Here,

$$\lambda_+ = \frac{-(1+\gamma)+\sqrt{(1+\gamma)^2+4\gamma(bg'(x_1)-1)}}{2} \tag{7.80}$$

is the positive eigenvalue of the rate equation matrix at the unstable fixed point, $K_{0,1}$ are the limits as $t \to \pm\infty$ of the solutions of the differential equation

$$\frac{\mathrm{d}\log K(t)}{\mathrm{d}t} = H_{xp} + H_{yq} + \tfrac{1}{2}(Z_{xx}H_{pp} + Z_{yy}H_{qq}), \tag{7.81}$$

and $Z_{0,1}$ are the corresponding limits of the symmetic 2×2 matrix function of t

$$Z = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yz} & Z_{yy} \end{pmatrix} = \begin{pmatrix} \frac{\partial p}{\partial x} & \frac{\partial p}{\partial y} \\ \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} \end{pmatrix}, \tag{7.82}$$

which solves

$$-\dot{Z} = ZBZ + ZA + A^T T + C. \tag{7.83}$$

The elements of the matrices A, B, and C are second derivatives of H:

$$A = \begin{pmatrix} H_{px} & H_{py} \\ H_{qx} & H_{qy} \end{pmatrix}, \tag{7.84}$$

$$B = \begin{pmatrix} H_{pp} & H_{pq} \\ H_{qp} & H_{qq} \end{pmatrix}, \tag{7.85}$$

and

$$C = \begin{pmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{pmatrix}. \tag{7.86}$$

These matrices are functions of t through their dependence on the solutions x, y, p and q of the Hamilton equations of motion. Explicitly,

$$A = \begin{pmatrix} -\gamma \mathrm{e}^{-p} & cg'(cy)\mathrm{e}^{p} \\ \mathrm{e}^{q/c} & -\mathrm{e}^{-q/c} \end{pmatrix}, \tag{7.87}$$

$$B = \begin{pmatrix} g(cy)\mathrm{e}^{p} + \gamma x \mathrm{e}^{-p} & 0 \\ 0 & (x\mathrm{e}^{q/c} + y\mathrm{e}^{-q/c})/c \end{pmatrix}, \tag{7.88}$$

and

$$C = \begin{pmatrix} 0 & 0 \\ 0 & c^2 g''(cy)(\mathrm{e}^{p} - 1) \end{pmatrix}. \tag{7.89}$$

Here, and in the calculation described below, we have rescaled the protein concentration y by a factor $1/c$, so that $y = x$ at the fixed points, and, correspondingly, rescaled q by a factor c.

We have solved (7.83) numerically for Z, using the same relaxation method employed above in finding the optimal path $x(t), y(t), p(t), q(t)$. (As in previous calculations, we use a burst size $b = 15$.) Putting the elements $Z_{xx}(t)$ and $Z_{zz}(t)$ into (7.81) and integrating, we evaluate K_1/K_0 and thus the prefactor η from (7.79). The product $\sqrt{\det(Z_0)/|\det(Z_1)|}\cdot K_1/K_0$ is almost independent of γ; the γ-dependence of η is accounted for entirely (within our numerical accuracy) by that of the rate equation eigenvalue λ_+. We are tempted to conjecture that this is exact, but we have not been able to prove it.

There is a cloud hanging over these results: The theoretical treatment rests on the assumption that the probability density $P(x, y; t)$ everywhere along the optimal path (the solutions of the Hamilton equations) is Gaussian, with inverse correlation matrix $Z(t)$. However, in a later paper, Maier and Stein [89] showed that, except for a special class of models for which the flow near the saddle point is derivable from a potential, the final ($t \to \infty$) density near the unstable point (x_1, y_1) is not Gaussian because that point is not accessible from all directions in the local flow. Our model is not in this special class, so our calculation of the escape rate cannot be correct.

Maier and Stein investigated the general case and showed, in an elaborate calculation, how one can find moments of the true exit distribution. However, they did not go so far as to give an explicit result for the value of the prefactor, and we do not attempt this calculation here. Lacking such a result, we have performed numerical simulations to find out how big an error we have made in our naive calculation above.

We have simulated the model using the Gillespie algorithm, estimating the mean first passage time to reach the protein concentration y_1 by averaging over $10,000$ trials, for $b = 15$ and $\gamma = 8$, 16 and 32. We define an empirical prefactor by multiplying the transition rate measured in simulations by $\exp(S_0)$, where S_0 is the barrier calculated in the preceding subsection. They show clearly that the naive theoretical calculation is wrong, as Maier and Stein would have anticipated. For $\gamma = 8$ and 16, the theoretical prefactors are about 20% higher than those found in the simulations; for $\gamma = 32$ it is about 15% higher. We expect that as $\gamma \to \infty$ both the theoretical and the empirical values should approach the bursting-limit prefactor, so the discrepancy should disappear.

Chapter 8

Conclusion

8.1 Summary of Thesis Achievements

Through this thesis, we have studied several topics in the intersection between probability, applied mathematics, and theoretical physics. The main focus has been in the study of field-theoretic techniques applied to probability. We summarize our main conclusions in order by chapter

1. In Chapter 3, the main focus has been the study of field-theoretic methods applied to the field of abstract chemical reactions, a subclass of Markov processes in discrete state space that appear naturally in different models of applied mathematics, physics, and probability theory.

   We have presented a mathematical approach to the main concepts that appear in this field: state spaces, annihilation and creation operators, the Master equation evolution (3.1), and coherent representations, among others.

   One of the main Theorems (3.4) shows that we can not expect a one to one relation between continuous probability distributions in the interval $(0, \infty)$ and sequences of probabilities in $\mathbb{N}$, as one could naively expect from the coherent transform. In particular, for fully recovering the space of sequences of probabilities, one would need to go, at least, to a space of distributions (with derivatives of Dirac delta functions). Thanks to this observation, it is not strange to see that there are coherent PDEs with negative diffusion coefficient (3.45) that give rise, formally, to complex noise SDEs.

   We have as well seen in this chapter that the theory of generating functions is a particular

representation of the more abstract Doi-Peliti theory where one takes as a base the sequence of polynomials $\{x^n\}$.

Finally, we have made explicit the relation between Doi-Peliti and Martin-Siggia-Rose field theories due to a symplectic transformation between them.

2. In Chapter 4, we have made a step further into the investigations of the previous chapter, and we have made two main observations.

   The first of those observations is that the generating function representation of Doi-Peliti field theory gives rise, in some instances, to stochastic dualities between discrete state-space Markov processes and Itô SDEs.

   The key ingredient to that connection comes from the application of the Feynman-Kac theorem (4.3). We have proven a slightly generic version of the classical Feynman-Kac theorem, in which we include the possibility of degeneracy in the diffusion coefficient. This diffusion coefficient usually is null in the boundary of its interval of definition, which gives rise to absorbing states for the equivalent SDE.

   The second main observation is that any Sheffer sequence can play the role of monomials in the generating function representation. Following that idea, we have found Hermite representations and have studied some examples in which we can apply the duality concepts to this representation as well.

3. In Chapter 5, we use as motivation the previously studied absorbing SDEs to studying a classic problem on the theory of Stochastic Differential Equations. In particular, we have realized that the formally equivalent Stratonovich counterparts of those equations lack some of the solutions of the Itô SDEs (5.2). Itô and Stratonovich formalism are not equivalent to each other for this particular class of stochastic phenomena. To our surprise, these equations are widespread in the literature in applications in probability, applied mathematics, and physics, so we believe the impact of this result can have consequences on the way these equations are analyzed.

4. In Chapter 6, we return to a question that already appeared in Chapter 3 when dealing with the coherent transformation of the coalescence reaction. The main question that we try to answer in this chapter is whether or not one can use the formally equivalent complex SDE to study the negative diffusion Kolmogorov equation and the related coalescence process.

   We provide an affirmative answer to this question in Theorem (6.4), at least for the case of finite

particle initial conditions. We have seen this relating the formally equivalent complex SDE and the negative diffusion Fokker-Plank equation through the Cauchy transform.

5. In Chapter 7, we model a system of Protein/mRNA interactions inside the cell as a coupled stochastic system that presents bi-stability. The main question we want to answer is whether one can obtain a formula for the mean transition time between these two locally stable solutions. The methods applied in this chapter go from the application of asymptotic analysis to reduce the problem first to a one-dimensional case, in which one can study the question with the help of a WKB approach and Hamilton-Jacobi equations theoretically. For the general problem, we proceed by making asymptotic expansions and, with the help of numerical methods, we can see our method gives a reasonable answer, given the complexity of the problem.

We believe the methods developed in this thesis can serve to future studies in several directions, in particular, all of our lines of investigation open more questions than they answer, future lines of investigation from this thesis can be

1. Extend duality relations to other types of Sheffer sequences.
2. Study duality relations for the case of abstract chemical reactions with more than one species.
3. Further implications of the broken Itô Stratonovich relation in the context of absorbing SDEs, in particular in more than one dimension.
4. Develop further consequences of the theory of complex noise SDEs in connection with Cauchy transformations.

8.2 Conclusiones de la tesis

A lo largo de esta tesis, hemos estudiado temas relacionados con teoría de la Probabilidad, Matemática Aplicada y Física Teórica. En particular, el tronco central alrededor del cual pueden verse los temas de esta tesis es la aplicación de métodos de teoría de campos a la teoría de probabilidad. Vamos a ver un resumen de los principales temas vistos en esta tesis capítulo por capítulo

1. En el capítulo (3), hemos aplicado la teoría de campos a las ecuaciones químicas abstractas, que son un ejemplo de procesos de Markov que aparecen en diversos modelos de la matemática aplicada, la teoría de probabilidad y la física.

Primero hemos estudiado rigurosamente los conceptos matemáticos que aparecen en estas formulaciones: espacios de estados, operadores de creación y destrucción, la dinámica de la ecuación maestra (3.1), y la transformada coherente.

En uno de los principales teoremas de este capítulo, hemos visto que uno no puede representar las funciones de densidad en el intervalo $(0, \infty)$ con el espacio de las secuencias de probabilidad en $\mathbb{N}$, al menos no a través de la transformada coherente, que es lo que uno esperaría a priori. Una de la principales consecuencias de este resultado es que, para poder representar el espacio de secuencias de probabilidad, uno debe estudiar transformadas coherentes de distribuciones (en particular, de derivadas de deltas de Dirac). A la vista de este resultado, no es de extrañar que aparezcan ecuaciones con ruido negativo en el estudio de ciertas transformadas coherentes (3.45).

Además, hemos visto que la teoría de funciones generatrices en probabilidad, puede verse como una representación de la teoría de Doi-Peliti en la que la base del espacio es representada por monomios $\{x^n\}$.

Finalmente, hemos relacionado la teoría de Doi-Peliti con la teoría de Martin-Siggia-Rose, que es otra teoría de campos aplicada comúnmente al caso de procesos de Markov con espacio discreto. Esta relación se basa en la existencia de una transformación simpléctica entre ambas teorías.

2. En el capítulo (4), continuamos el estudio iniciado en el capítulo previo y podemos resumir nuestros principales resultados en dos observaciones.

   La primera observación es que la aplicación de funciones generatrices a nuestras ecuaciones químicas abstractas dan lugar, en ciertos casos, a una dualidad estocástica donde un proceso estocástico con espacio discreto es dual a una ecuación estocástica de Itô (EDS).

   El ingrediente fundamental de esta dualidad es la aplicación del teorema de Feynman-Kac (4.3). De hecho, para ello necesitamos ampliar el caso clásico del teorema de Feynman-Kac al caso en el que tenemos ecuaciones de difusión con coeficientes de difusión degenerados en la frontera. Estos coeficientes de difusión se anulan en la frontera dando lugar a una representación en EDS con estados absorventes en la frontera.

   La segunda observación de este capítulo, es que uno puede considerar representaciones de la teoría de Doi-Peliti usando secuencias polinomiales de Sheffer, en particular, usamos la secuencia de Hermite para estudiar casos novedosos de dualidad que no eran posibles para el caso de funciones generatrices.

3. En el capítulo (5), y motivados por las EDS con estados absorventes de capítulos anteriores, nos centramos en el estudio de un problema clásico en el estudio de ecuaciones estocásticas, el dilema de Itô vs Stratonovich. Esta motivación viene del hecho de que los representantes formales en ecuación de Stratonovich de los ejemplos estudiados con anterioridad, carecen de dichas soluciones absorventes (5.2). Esto prueba que la dualidad Itô vs Stratonovich no es cierta en el caso de estas ecuaciones estocásticas. Sorprendentemente, este tipo de ecuaciones son usadas frecuentemente a la hora de modelar fenómenos en los campos de Probabilidad, Matemática aplicada y física, por lo que creemos que este resultado puede tener un cierto impacto a la hora de estudiar estos modelos.

4. En el capítulo (6), volvemos a retomar una ecuación que apareció por primera vez en el capítulo (3) cuando introdujimos la transformada coherente de la reacción de coalescencia. La cuestión que intentamos resolver es, hasta que punto se puede usar la ecuación con ruido imaginario obtenida formalmente como representante coherente del proceso de Markov de coalescencia para estudiar el proceso anterior y, más teóricamente, si existe alguna relación entre la ecuación estocástica con ruido imaginario y su correspondiente ecuación de Kolmogorov con difusión negativa.

   La respuesta a estas preguntas es positiva en ambos casos, ver teorema (6.4), al menos para el caso de condiciones iniciales con un número finito de partículas. Hemos obtenido este resultado relacionando la solución de la ecuación de diffusión negativa con la ecuación estocástica con ruido imaginario a través de la transformada de Cauchy, que permite representar distribuciones como funciones holomorfas en el plano complejo salvo la recta real.

5. En el capítulo (7), nos centramos en el estudio de un modelo estocástico para la evolución de la concentración de proteínas y RNAm donde soluciones biestables aparecen. Debido al componente estocástico de este modelo, existen transiciones entre las soluciones localmente estables y la pregunta que intentamos resolver es si existe un método para calcular el tiempo medio de transición entre estas dos soluciones, que será un evento raro. El método empleado para resolver esta cuestión es la aplicación de una expansión asintótica en la que reducimos primero las dimensiones de este problema a una dimensión, donde podemos aplicar técnicas como WKB y ecuaciones de Hamilton-Jacobi para responder a la pregunta en este límite concreto. Para el caso general, usamos otra expansión asintótica y métodos numéricos para comprobar que nuestra fórmula, si bien no es exacta, es hasta cierto punto precisa dada la complejidad del problema.

Creemos que los métodos desarrollados en esta tesis pueden servir para el estudio de líneas de investi-

gación futuras, en particular, en cada uno de los temas que estudiamos aquí se abren más preguntas con los resultados obtenidos, algunas de estas futuras líneas de investigación podrían ser

1. Estudiar nuevos tipos de dualidad con polinomios de Sheffer que no hemos tratado aquí.
2. Estudiar problemas de dualidad en el contexto de ecuaciones químicas abstractas con más de un componente químico
3. Estudiar más profundamente las implicaciones de la no equivalencia de las formulaciones de Itô y Stratonovich para ecuaciones estocásticas con estados absorventes.
4. Estudiar más profundamente las relaciones entre ecuaciones estocásticas con ruido imaginario y las ecuaciones de difusión negativas.

Appendix A

Calculations from Stochastic Activation in Switch Analysis

In this Appendix we present a detailed calculation of the asymptotic expansion on $O(\gamma^{-1})$ of the solution of the 4-dimensional Hamiltonian problem (7.5) with additional calculations for the asymptotic prefactor behavior in the limit of big γ. We begin writing the 4-dimensional Hamiltonian involved

$$H(x,p,y,q) = g(y)(e^p - 1) - \gamma x(1 - e^{-p}) + b\gamma x(e^q - 1) - y(1 - e^{-q}) \tag{A.1}$$

from which we extract the Hamiltonian flow follows the ODE's

$$\begin{aligned}
\dot{x} &= \frac{\partial H}{\partial p} = g(y)e^p - \gamma x e^{-p} \\
\dot{p} &= -\frac{\partial H}{\partial x} = \gamma(1 - e^{-p}) - \gamma b(e^q - 1) \\
\dot{y} &= \frac{\partial H}{\partial q} = b\gamma x e^q - y e^{-q} \\
\dot{q} &= -\frac{\partial H}{\partial y} = -g'(y)(e^p - 1) + 1 - e^{-q}
\end{aligned}$$

Now we are going to simplify things written everything in terms of

$$P = e^p \qquad\qquad Q = e^q \qquad\qquad X = \gamma x$$

So that the Hamiltonian is written as

$$H = g(y)(P-1) - X(1-P^{-1}) + bX(Q-1) - y(1-Q^{-1}) \tag{A.2}$$

and also

$$\gamma^{-1}\dot{X} = g(y)P - XP^{-1} \tag{A.3}$$

$$\gamma^{-1}\dot{P} = (P-1) - bP(Q-1) \tag{A.4}$$

$$\dot{y} = bXQ - yQ^{-1} \tag{A.5}$$

$$\dot{Q} = -g'(y)Q(P-1) + Q - 1 \tag{A.6}$$

A.1 $O(\gamma^{-1})$ for the Variables

Now we make the assumption that, in the large γ limit, everything could be written in terms of $O(\gamma^{-1})$. We use the following notation

$$\begin{aligned} y &= y + \gamma^{-1}y_2 \\ P &= P_1 + \gamma^{-1}P_2 \\ X &= X_1 + \gamma^{-1}X_2 \\ Q &= Q_1 + \gamma^{-1}Q_2 \end{aligned}$$

Actually, in the first equation there is abuse of notation as the y in the right hand side of the equality should be y_∞ which is the limit of y for big γ. We hope this does not induce much confussion to the reader, as far as we are consistent everything works fine.

Writting the hamiltonian at order γ^{-1}

$$\begin{aligned} &(g(y) + \gamma^{-1}g'(y)y_2)(P_1 - 1 + \gamma^{-1}P_2) - (X_1 + \gamma^{-1}X_2)(1 - P_1^{-1} + \gamma^{-1}P_1^{-2}P_2) + \\ &+b(X_1 + \gamma^{-1}X_2)(Q_1 - 1 + \gamma^{-1}Q_2) - (y + \gamma^{-1}y_2)(1 - Q_1^{-1} + \gamma^{-1}Q_1^{-2}Q_2) \end{aligned}$$

Which order $O(1)$ we suppose will be exactly zero as we want that condition independent of γ

$$g(y)(P_1 - 1) - X_1(1 - P_1^{-1}) + bX_1(Q_1 - 1) - y(1 - Q_1^{-1}) = 0 \tag{A.7}$$

Additionally, we want the differential equations (A.3) to be satisfied at all orders of γ. So, from the first two at order $O(1)$ we obtain the following equalities

$$0 = g(y)P_1 - X_1 P_1^{-1}$$
$$0 = (P_1 - 1) - bP_1(Q_1 - 1)$$

Which can be rewritten as

$$X_1 = g(y)P_1^2$$
$$P_1 = \frac{1}{1 - b(Q_1 - 1)}$$

and joining with the Hamiltonian for that same order we can write everything in terms of y_∞ $(= y)$

$$X_1 = \frac{1}{(1+b)^2}\frac{(y + g(y))^2}{g(y)} \tag{A.8}$$
$$P_1 = \frac{1}{1+b}\frac{y + g(y)}{g(y)} \tag{A.9}$$
$$Q_1 = \frac{1+b}{b}\frac{y}{y + g(y)} \tag{A.10}$$

which for $\gamma = 16$ seems pretty accurate, as assuming $y \equiv y_\infty$ the rest should be written in its terms. The following figures shows how different are the actual values of each variable with the values they should have at first order $O(1)$.

Now we proceed to extract the $O(\gamma^{-1})$ terms. For that, lets write the $O(\gamma^{-1})$ of the hamiltonian, which should evaluate exactly zero

$$P_2\left[g(y) - X_1 P_1^{-2}\right] + X_2\left[P_1^{-1} - 1 + b(Q_1 - 1)\right] + \tag{A.11}$$
$$+Q_2\left[bX_1 - yQ_1^{-2}\right] + y_2\left[g'(y)(P_1 - 1) + Q_1^{-1} - 1\right] = 0 \tag{A.12}$$

which, evaluating what we already know lead us to

$$\frac{b(y + g(y))^2(y - bg(y))}{(1+b)^2 y g(y)}Q_2 - \frac{(y - bg(y))(g(y) - yg'(y))}{(1+b)yg(y)}y_2 = 0 \tag{A.13}$$

and so

$$Q_2 = \frac{1+b}{b}\frac{g(y) - yg'(y)}{(y + g(y))^2}y_2 \tag{A.14}$$

Proceeding similarly, we could obtain the following identities from the first two differential equations written in order $O(\gamma^{-1})$

$$Q_2 = \frac{1+b}{b}\frac{g(y)-yg'(y)}{(y+g(y))^2}y_2 \tag{A.15}$$

$$P_2 = \frac{(y+g(y))^2(y-bg(y))(g(y)-yg'(y))}{(1+b)^3g(y)^4} + \frac{g(y)-yg'(y)}{(1+b)g(y)^2}y_2 \tag{A.16}$$

$$X_2 = -\frac{(y+g(y))^4g'(y)(y-bg(y))}{(1+b)^4g(y)^4} + \frac{(y+g(y))(g(y)g'(y)+2g(y)-yg'(y))}{(1+b)^2g(y)^2}y_2 \tag{A.17}$$

Theoretically, it would be possible to obtain y_2 from one of the two differential equations at $O(\gamma^{-1})$, which actually coincide. But it is very tedious and, more importantly, the actual form of y_2 will not be important whenever we write the equation for K as the coefficient in y_2 for that equation is exactly 0 which simplifies our lifes at to the point that we can suppose that $y_2 = 0$ for making our future calculations (as our goal is just to obtain the equation for K). The only thing that we are going to mention about y_2 is that $y_2 \equiv 0$ at each of the boundary points, so all our identities will be exact (at that order of γ) at the boundaries.

With that in mind, we actually can write all our variables at order γ^{-1} with the "axiom" that $y_2 = 0$

$$X = \frac{1}{(1+b)^2}\frac{(y+g(y))^2}{g(y)} - \gamma^{-1}\frac{(y+g(y))^4g'(y)(y-bg(y))}{(1+b)^4g(y)^4} \tag{A.18}$$

$$Q = \frac{1+b}{b}\frac{y}{y+g(y)} \tag{A.19}$$

$$y = y \tag{A.20}$$

$$P = \frac{1}{1+b}\frac{y+g(y)}{g(y)} + \gamma^{-1}\frac{(y+g(y))^2(y-bg(y))(g(y)-yg'(y))}{(1+b)^3g(y)^4} \tag{A.21}$$

As we are taking time derivatives, and now everything depends on y it is useful to calculate know $\dot{y}$ at order $O(1)$

$$\dot{y} = \frac{(y+g(y))(y-bg(y))}{(1+b)g(y)} \tag{A.22}$$

A.2 $O(\gamma^{-1})$ for Z Matrix and K Equation

First we are going to write the equation for K that can be seen in our references

$$-\frac{\dot{K}}{K} = \frac{\partial^2 H}{\partial x \partial p} + \frac{\partial^2 H}{\partial y \partial q} + \frac{1}{2}\frac{\partial p}{\partial x}\frac{\partial^2 H}{\partial p^2} + \frac{1}{2}\frac{\partial q}{\partial y}\frac{\partial^2 H}{\partial q^2} \tag{A.23}$$

from which we can extract the term $O(\gamma)$ and, as we suppose K should not depend on this order (in other case, K will either go to zero or to ∞ for big γ), we extract the order γ for $\frac{\partial p}{\partial x}$ which is

$$\frac{\partial p}{\partial x} = \frac{(1+b)^2 g(y)}{(y+g(y))^2}\gamma + c(y) \tag{A.24}$$

Now from the identity

$$\dot{p} = \frac{\partial p}{\partial x}\dot{x} + \frac{\partial p}{\partial y}\dot{y} \tag{A.25}$$

at order $O(1)$ we obtain, in this order,

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x} = -\frac{1+g'(y)}{y+g(y)} \tag{A.26}$$

and from the equation

$$\dot{q} = \frac{\partial q}{\partial x}\dot{x} + \frac{\partial q}{\partial y}\dot{y} \tag{A.27}$$

at order $O(1)$ we obtain

$$\frac{\partial q}{\partial y} = \frac{g(y) - y g'(y)}{y(y+g(y))} \tag{A.28}$$

We would need to know the function $c(y)$ for calculating the equation for K this comes from plugging all the information we have into the differential matrix equation for Z, where [1]

$$Z = \begin{pmatrix} \frac{\partial p}{\partial x} & \frac{\partial p}{\partial y} \\ \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} \end{pmatrix}$$

which is

$$-\dot{Z} = ZBZ + ZA + A^T Z + C \tag{A.29}$$

with

$$B = \begin{pmatrix} \frac{\partial^2 H}{\partial p^2} & \frac{\partial^2 H}{\partial p \partial q} \\ \frac{\partial^2 H}{\partial p \partial q} & \frac{\partial^2 H}{\partial q^2} \end{pmatrix} = \begin{pmatrix} g(y)e^p + \gamma x e^{-p} & 0 \\ 0 & b\gamma x e^q + y e^{-q} \end{pmatrix}$$

[1] $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$

$$A = \begin{pmatrix} \frac{\partial^2 H}{\partial p \partial x} & \frac{\partial^2 H}{\partial p \partial y} \\ \frac{\partial^2 H}{\partial q \partial x} & \frac{\partial^2 H}{\partial q \partial y} \end{pmatrix} = \begin{pmatrix} -\gamma e^{-p} & g'(y)e^{p} \\ \gamma b e^{q} & -e^{-q} \end{pmatrix}$$

$$C = \begin{pmatrix} \frac{\partial^2 H}{\partial x^2} & \frac{\partial^2 H}{\partial x \partial y} \\ \frac{\partial^2 H}{\partial x \partial y} & \frac{\partial^2 H}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & g''(y)(e^{p}-1) \end{pmatrix}$$

it is easy to see that the term in order γ^2 dissappears naturally and impossing that also the term at order $O(\gamma)$ should disappear (this hypothesis comes because that's what happen in the boundary, much more easy to calculate) we obtain the function $c(y)$ which is

$$c(y) = \frac{yg(y) + g'(y)\left[y^2 - g(y)(bg(y) + (b-2)y)\right]}{g(y)^2(y + g(y))} \tag{A.30}$$

So we can resume the information about the coefficients of the matrix Z as

$$\begin{aligned} \frac{\partial p}{\partial x} =& \frac{(1+b)^2 g(y)}{(y+g(y))^2}\gamma + \frac{yg(y) + g'(y)\left[y^2 - g(y)(bg(y) + (b-2)y)\right]}{g(y)^2(y+g(y))} \\ \frac{\partial q}{\partial y} =& \frac{g(y) - yg'(y)}{y(y+g(y))} \\ \frac{\partial p}{\partial y} =& \frac{\partial q}{\partial x} = -\frac{1+g'(y)}{y+g(y)} \end{aligned}$$

And now we can substitute in the equation for K and obtain

$$-\frac{\dot{K}}{K} = \frac{(y - bg(y))\left[g(y)(4y + g(y)) - yg'(y)(y - 2g(y))\right]}{2(1+b)yg(y)^2} \tag{A.31}$$

Which written in terms of derivatives w.r.t. y (y_∞) is

$$\begin{aligned} \frac{d}{dy}\log(K^{-1}) =& \dot{y}^{-1}\frac{(y - bg(y))\left[g(y)(4y + g(y)) - yg'(y)(y - 2g(y))\right]}{2(1+b)yg(y)^2} = \\ =& \frac{g(y)(4y + g(y)) - yg'(y)(y - 2g(y))}{2yg(y)(y + g(y))} = \frac{d}{dy}\log\sqrt{\frac{y(y+g(y))^3}{g(y)}} \end{aligned}$$

So, the ratio

$$\frac{K_1}{K_0} = \sqrt{\frac{y_0^3}{y_1^3}} \tag{A.32}$$

which is not the term that give the same ratio in the context of 1-d discrete.

A.3 Prefactor for $\gamma \to \infty$

Using the information we have, we can take the limit in the "without wedge" prefactor formula from Maier-Stein

$$\eta = \frac{\lambda_+}{\pi}\sqrt{\frac{\det(Z_0)}{\det(Z_1)}}\frac{K_1}{K_0} \tag{A.33}$$

where λ_+ is the positive (unstable) eigenvalue of the matrix

$$A_1 = \begin{pmatrix} -\gamma & g_1' \\ \\ \gamma b & -1 \end{pmatrix}$$

which is

$$\lambda_+ = \frac{-1-\gamma+\sqrt{1-2\gamma+4bg_1'\gamma+\gamma^2}}{2} \tag{A.34}$$

and the ratio

$$\frac{\det(Z_0)}{\det(Z_1)} = \frac{y_1^2\left[-\gamma+bg_0'(1+\gamma+g_0')\right]}{y_0^2\left[-\gamma+bg_1'(1+\gamma+g_1')\right]} \tag{A.35}$$

which in the limit of big γ is

$$\lim_{\gamma\to\infty}\frac{\det(Z_0)}{\det(Z_1)} = \frac{|1-bg_0'|\,y_1^2}{|1-bg_1'|\,y_0^2} \tag{A.36}$$

Additionally, the limit of λ_+ is

$$\lim_{\gamma\to\infty}\lambda_+ = \left|1-bg_1'\right| \tag{A.37}$$

So, using the formula for prefactor (which is not clear to be correct in the context of wedge, as we have)

$$\eta = \frac{\lambda_+}{\pi}\sqrt{\frac{|1-bg_0'|\,y_1^2}{|1-bg_1'|\,y_0^2}}\sqrt{\frac{y_0^3}{y_1^3}} = \frac{\sqrt{|1-bg_0'|\,|1-bg_1'|}}{\pi}\sqrt{\frac{y_0}{y_1}} \tag{A.38}$$

which is the same formula one obtains for the 1-d discrete case. This tell us that the formula "without wedge" from Maier-Stein is true at least in the limit.

Additionally, we have some information of order $1/\gamma$ for the case of finite gamma, that can not be obtained from the 1-d formula alone, which allows us to explain the behavior of the prefactor in the 2d case for finite γ partially.

We can not assure, with the information we have, that the prefactor would be independent of b for the case of finite γ. At least, this is true for the limit of big γ.

Bibliography

[1] M. Doi. Second quantization representation for classical many-particle system. *Journal of Physics A: Mathematical and General*, 9(9):1465, 1976.

[2] L. Peliti. Path integral approach to birth–death processes on a lattice. *Journal de Physique*, 46(9):1469–1483, 1985.

[3] J. Cardy. *Scaling and Renormalization in Statistical Physics*, volume 5. Cambridge University Press, 1996.

[4] M.A. Muñoz. Nature of different types of absorbing states. *Physical Review E*, 57:1377–1383, Feb 1998.

[5] F. Benitez, C. Duclut, H. Chaté, B. Delamotte, I. Dornic, and M.A. Muñoz. Langevin equations for reaction-diffusion processes. *Physical Review Letters*, 117(10):100601, 2016.

[6] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Berlin: Springer, 1985.

[7] B. Øksendal. Stochastic differential equations. In *Stochastic Differential Equations*, pages 65–84. Springer, 2003.

[8] L. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*, volume 2. Cambridge University Press, 1994.

[9] L.C. Evans. *An Introduction To Stochastic Differential Equations*, volume 82. American Mathematical Society, 2012.

[10] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293. Springer Science & Business Media, 2013.

[11] A. Borodin and P. Salminen. *Handbook of Brownian Motion-Facts and Formulae*. Birkhäuser, 2012.

[12] P. Mörters and Y. Peres. *Brownian Motion*, volume 30. Cambridge University Press, 2010.

[13] S.M. Roman and G.C. Rota. The umbral calculus. *Advances in Mathematics*, 27(2):95–188, 1978.

[14] M.I. Freidlin and A.D. Wentzell. Random perturbations. In *Random Perturbations of Dynamical Systems*, pages 15–43. Springer, 1998.

[15] A. Erdelyi. Asymptotic expansions. *Technical Report 3- Office of Naval Research*, 1955.

[16] P. Grassberger and A. De La Torre. Reggeon field theory (schlögl's first model) on a lattice: Monte carlo calculations of critical behaviour. *Annals of Physics*, 122(2):373–396, 1979.

[17] L. Peliti. Renormalisation of fluctuation effects in the a+ a — a reaction. *Journal of Physics A: Mathematical and General*, 19(6):L365, 1986.

[18] Masao Doi. Stochastic theory of diffusion–controlled reaction. *Journal of Physics A: Mathematical and General*, 9(9):1479, 1976.

[19] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry*, volume 1. Elsevier, 1992.

[20] C.W. Gardiner. *Stochastic Methods: a Handbook for the Natural and Social Sciences*. Springer, 2009.

[21] K. Wiese. Coherent-state path integral versus coarse-grained effective stochastic equation of motion: From reaction–diffusion to stochastic sandpiles. *Physical Review E*, 93:042117, Apr 2016.

[22] M. Assaf and B. Meerson. Spectral theory of metastability and extinction in birth-death systems. *Physical Review Letters*, 97(20):200602, 2006.

[23] M. Assaf and B. Meerson. Spectral theory of metastability and extinction in a branching-annihilation reaction. *Physical Review E*, 75(3):031122, 2007.

[24] D.A. McQuarrie. Kinetics of small systems. i. *The Journal of Chemical Physics*, 38(2):433–436, 1963.

[25] D.A. McQuarrie, C.J. Jachimowski, and M.E. Russell. Kinetics of small systems. ii. *The Journal of Chemical Physics*, 40(10):2914–2921, 1964.

[26] J. Baez and J. Biamonte. Quantum techniques for stochastic mechanics. *ArXiv:1209.3632*, 2012.

[27] M. Droz and A. McKane. Equivalence between poisson representation and fock space formalism for birth-death processes. *Journal of Physics. A, Mathematical and General*, 27(13):L467–L474, 1994.

[28] V. Elgart and A. Kamenev. Rare event statistics in reaction–diffusion systems. *Physical Review E*, 70(4):041106, 2004.

[29] V. Elgart and A. Kamenev. Classification of phase transitions in reaction–diffusion models. *Physical Review E*, 74(4):041101, 2006.

[30] J. Zinn-Justin. Quantum field theory and critical phenomena. 2002.

[31] M. Weber and E. Frey. Master equations and the theory of stochastic path integrals. *Reports on Progress in Physics*, 80(4):046601, 2017.

[32] Uwe C Täuber. *Critical Dynamics: a Field Theory Approach to Equilibrium and Non–Equilibrium Scaling Behavior*. Cambridge University Press, 2014.

[33] S. Asmussen. *Applied Probability and Queues*, volume 51. Springer Science & Business Media, 2008.

[34] W. Feller. *An Introduction to Probability Theory and its Applications: Volume I*, volume 3. John Wiley & Sons New York, 1968.

[35] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets*. Springer Science & Business Media, 2009.

[36] J. Cardy. Renormalisation group approach to reaction-diffusion problems. *ArXiv Preprint cond-mat/9607163*, 1996.

[37] M.J. Howard and U.C. Täuber. 'real' versus 'imaginary' noise in diffusion–limited reactions. *Journal of Physics A: Mathematical and General*, 30(22):7721, 1997.

[38] O. Deloubriere, L. Frachebourg, H.J. Hilhorst, and K. Kitahara. Imaginary noise and parity conservation in the reaction a+a — 0. *Physica A: Statistical Mechanics and its Applications*, 308(1):135–147, 2002.

[39] D. Gredat, I. Dornic, and J.M. Luck. On an imaginary exponential functional of brownian motion. *Journal of Physics A: Mathematical and Theoretical*, 44(17):175003, 2011.

[40] U.C. Täuber. Field theoretic methods. In *Computational Complexity*, pages 1080–1093. Springer, 2012.

[41] A. Kamenev. Keldysh and doi–peliti techniques for out–of–equilibrium systems. In *Strongly Correlated Fermions and Bosons in Low-Dimensional Disordered Systems*, pages 313–340. Springer, 2002.

[42] J. Cardy. Field theory and non-equilibrium statistical mechanics. *URL http://www-thphys. physics. ox. ac. uk/users/JohnCardy/notes. ps.*

[43] J. Cardy. Reaction-diffusion processes. 100:26, 2006.

[44] C. Perret. The stability of numerical simulations of complex stochastic differential equations. 2010.

[45] L. Chen and D.W. Stroock. The fundamental solution to the wright–fisher equation. *SIAM Journal on Mathematical Analysis*, 42(2):539–567, 2010.

[46] C.L. Epstein and R. Mazzeo. Wright–fisher diffusion in one dimension. *SIAM Journal on Mathematical Analysis*, 42(2):568–608, 2010.

[47] C.L. Epstein and C.A. Pop. The feynman–kac formula and harnack inequality for degenerate diffusions. *The Annals of Probability*, 45(5):3336–3384, 2017.

[48] P. Martin, E.D. Siggia, and H.A. Rose. Statistical dynamics of classical systems. *Physical Review A*, 8(1):423, 1973.

[49] P. Lévy and M. Loève. *Processus Stochastiques et Mouvement Brownien.* Gauthier–Villars Paris, 1965.

[50] S. Karlin and J. McGregor. The classification of birth and death processes. *Transactions of the American Mathematical Society*, 86(2):366–400, 1957.

[51] D.V. Lindley. The theory of queues with a single server. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 48, pages 277–289. Cambridge University Press, 1952.

[52] D. Stroock and SR. Varadhan. *Multidimensional Diffusion Processes.* Springer, 2007.

[53] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1):155–167, 1971.

[54] S. Jansen and N. Kurt. On the notion (s) of duality for markov processes. *Probability Surveys*, 11:59–120, 2014.

[55] R. Mannella and P. McClintock. Itô versus stratonovich: 30 years later. *Fluctuation and Noise Letters*, 11(01):1240010, 2012.

[56] N.G. Van Kampen. Itô versus stratonovich. *Journal of Statistical Physics*, 24(1):175–187, 1981.

[57] A. Correales and C. Escudero. Absence of itô/stratonovich duality in physical theories with absorbing states. *Journal of Mathematical Physics*, 2019.

[58] R.L. Stratonovich. A new representation for stochastic integrals and equations. *SIAM Journal on Control*, 4(2):362–371, 1966.

[59] HH. Kuo. *Introduction to Stochastic Integration*. Springer, New York, NY, 2006.

[60] L. Andersen and J. Andreasen. Volatility skews and extensions of the libor market model. *Applied Mathematical Finance*, 7(1):1–32, 2000.

[61] I.S. Helland. One–dimensional diffusion processes and their boundaries. 1996.

[62] L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50, 2007.

[63] R. Adler. Superprocesses and plankton dynamics. Technical report, North Carolina University at Chapel Hill Dept. Statistics, 1997.

[64] B. Houchmandzadeh. Clustering of diffusing organisms. *Physical Review E*, 66(5):052902, 2002.

[65] W. Young, A.J. Roberts, and G. Stuhne. Reproductive pair correlations and the clustering of organisms. *Nature*, 412(6844):328, 2001.

[66] W. Feller. Diffusion processes in genetics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. The Regents of the University of California, 1951.

[67] T. Tran, J. Hofrichter, and J. Jost. An introduction to the mathematical structure of the wright–fisher model of population genetics. *Theory in Biosciences*, 132(2):73–82, 2013.

[68] C.R. Doering, C. Mueller, and P. Smereka. Interacting particles, the stochastic fisher–kolmogorov–petrovsky–piscounov equation, and duality. *Physica A: Statistical Mechanics and its Applications*, 325(1-2):243–259, 2003.

[69] R. Peschanski. Traveling wave solution of the reggeon field theory. *Physical Review D*, 79(10):105014, 2009.

[70] N.V. Prikhod'ko. Homogeneous balitsky–kovchegov hierarchy and reggeon field theory. In *XXI International Baldin Seminar on High Energy Physics Problems*, volume 173, page 048. SISSA Medialab, 2013.

[71] A. Correales, C. Escudero, and M. Ptashnyk. Chemical kinetics, markov chains, and the imaginary itô interpretation. *Communications in Mathematical Physics*, 2019.

[72] A. Debrouwere. *Analytic Representations of Distributions and Ultradistributions*. PhD thesis, Ghent University, 2014.

[73] G. Köthe. Die randverteilungen analytischer funktionen. *Mathematische Zeitschrift*, 57(1):13–33, 1952.

[74] J. Hertz, J. Tyrcha, and A. Correales. Stochastic activation in a genetic switch model. *Physical Review E*, 98:052403, 2018.

[75] P. Bressloff. *Stochastic Processes in Cell Biology*, volume 41. Springer, 2014.

[76] P. Bressloff. Metastable states and quasicycles in a stochastic wilson-cowan model of neuronal population dynamics. *Physical Review E*, 82(5):051903, 2010.

[77] P. Bressloff. Stochastic switching in biology: from genotype to phenotype. *J Phys A*, 50:133001, 2017.

[78] C.C. Chow and M.A. Buice. Path integral methods for stochastic differential equations. *The Journal of Mathematical Neuroscience*, 5(1):8, 2015.

[79] N. Friedman, L. Cai, and X.S. Xie. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Physical Review Letters*, 97(16):168302, 2006.

[80] C. Long, N. Friedman, and X. Xie. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440(7082):358, 2006.

[81] V. Shahrezaei and P.S. Swain. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*, 105(45):17256–17261, 2008.

[82] A. Walczak, A. Mugler, and C.H. Wiggins. Analytic methods for modeling stochastic regulatory networks. *Computational Modeling of Signaling Networks*, pages 273–322, 2012.

[83] J. Yu, J. Xiao, X. Ren, K. Lao, and X. Xie. Probing gene expression in live cells, one protein molecule at a time. *Science*, 311(5767):1600–1603, 2006.

[84] C. Escudero and A. Kamenev. Switching rates of multistep reactions. *Physical Review E*, 79(4):041149, 2009.

[85] W. Press et al. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

[86] R.S. Maier and D.L. Stein. Escape problem for irreversible systems. *Physical Review E*, 48:931–938, 1993.

[87] D. Roma, R. O'Flanagan, A. Ruckenstein, A. Sengupta, and R. Mukhopadhyay. Optimal path to epigenetic switching. *Physical Review E*, 71:011902, 2005.

[88] H. Eyring. The activated complex in chemical reactions. *J Chem Phys*, 3:107–115, 1935.

[89] R.S. Maier and D. Stein. Limiting exit location distributions in the stochastic exit problem. *SIAM Journal on Applied Mathematics*, 57(3):752–790., 1997.