2018-07-23

Fluid Limit and Stochastic Stability for a Genetic Mutation Model

Carlos Bajo Caraballo
University of Miami, gordonpimm@gmail.com

UNIVERSITY OF MIAMI

FLUID LIMIT AND STOCHASTIC STABILITY FOR A GENETIC MUTATION MODEL

By

Carlos M. Bajo Caraballo

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2018

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

FLUID LIMIT AND STOCHASTIC STABILITY FOR A GENETIC MUTATION MODEL

Carlos M. Bajo Caraballo

Approved:

Ilie Grigorescu, Ph.D.
Associate Professor of Mathematics

Robert Stephen Cantrell, Ph.D.
Professor of Mathematics

Subramanian Ramakrishnan, Ph.D.
Associate Professor of Mathematics

Guillermo Prado, Ph.D.
Dean of the Graduate School

Burton J. Rosenberg, Ph.D.
Associate Professor of Computer Science

BAJO CARABALLO, CARLOS M. (Ph.D., Mathematics)
Fluid Limit and Stochastic Stability for a Genetic Mutation Model (August 2018)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Associate Professor Ilie Grigorescu.
No. of pages of text. (86)

We develop the continuous time version of the particle model studied in [13]. The evolution is described by a pure jump, continuous time Markov process on the space of words of length L with a size N alphabet. Words change randomly in search of a *preferred* state, here the vector zero. In genome population models, this is the genome presenting selection advantage [21]; in cancer development, it is a state of a damaged gene by deleterious mutations, and in epidemiological models the number of infected individuals in the population. In the last two models, the characters in the preferred word have a probability γ of returning among ordinary states. It will turn out to be essential that γ depends on the configuration, leading to an *interacting particle system*. We investigate the scaling limit of the empirical measure, and study several types of random perturbations, together with applications. Chapter 1 presents the mathematical model. Chapter 2 proves the *Fluid Limit*, i.e. a Law of Large Numbers for a (random) dynamical system; Chapter 3 determines the *Fluctuation near Equilibria*, a fine scale (second order approximation) result; and Chapter 4, titled *Generalized Logistic Equation with Noise*, explores the relationship between quasi-stationarity and stable equilibria, a random perturbation question.

Contents

List of Frequently Used Notation and Symbols

$\mathcal{A}$	Infinitesimal generator of a continuous Markov process.
$C_b(a,b)$	Continuous and bounded functions over the (possibly infinite) interval (a,b).
$C^k[a,b]$	Functions with k continuous derivatives up to the boundary.
$C_c^k(a,b)$	Space of functions in $C^k(a,b)$ having compact support.
$\mathcal{D}([0,\infty),\mathbb{X})$	The Skorohod space of right-continuous with left-limit paths (rcll, cádlág) endowed with the J_1 topology.
$\mathbb{P}_x$	The probability law of a process starting at x.
S	The alphabet $\mathbb{Z}_N^L$, $L \in \mathbb{N}$.
$\mathbb{X}$	A Polish space denoting the state space of an stochastic process.
$\mathbf{Z}$	A word (vector) in the space S.
Z^j	The j^{th} Component of the vector $\mathbf{Z} = (Z_1, \ldots, Z_L)$.
$\mathbf{Z}_t,\ (Z_t^j)$	The process indexed by time $t \geq 0$ (its j^{th} component) with state space $S \ni \mathbf{Z}$.

$\mathbf{Z}^j(k)$	The vector $\mathbf{Z}$, as defined in (1.1), with the j^{th} component updated with the letter $k \in \mathbb{Z}_N$.
$\mathbb{Z}_N$	The set $\{0, 1, ..., N-1\}$.

Chapter 1

Introduction

1.1 Markov Processes in Genetics

We begin this introduction with a brief presentation of the theory of Markov processes used in connection with the study of genetics, population evolution and biological models.

We follow Norris [17] together with the modern mathematical population genetics monograph by Ewens [7], and the recent survey by Durrett [6] for a brief history of these ideas.

A *gene* represents a configuration of a chromosome, typically evolving under mutation, which is assumed random in various setups. The term *allele* means a single state in the set of possible mutations of a chromosome. The *Wright-Fisher model* was introduced in the thirties to study alleles inheritance. In every generation there are a fixed number of alleles of two types. The type of alleles in a given generation is obtained by selecting randomly among the types in the previous generation. Evolution follows a Markov chain in the set of alleles.

The *Moran model* proposes a birth-and-death model consisting of individuals of two types, in which, at time n, we choose randomly an individual from the population and add an individual of the same type, then we choose again an individual from the population and remove it. This process determines the population at time $n + 1$.

A very popular tool in the study of genetics evolution is the *Theory of branching processes*. The first model was introduced by Galton and Watson in the 1870s for the study of family names survival. (For a classical references about this subject we indicate the reader to Harris [11] and Athreya and Ney [1].) Many other models inspired by this theory have been developed throughout the years. For instance, Durrett et al. [5, 6], proposes multi-type branching processes with mutations as an alternative to the study of cancer development. In his set up, there are several types of cells. Cells of type i follow a birth-and-death process, and in addition, give birth to individuals of of type $i + 1$ at a certain rate, producing mutations. He examines the time of the first type k mutations.

We find appropriate to mention the excellent book by Ewens (2004) titled *Mathematical Population Genetics*, and invite the reader to research it as an introductory reference source in many contemporary areas of study in the field of genetics. Of particular interest to us is the use of the theory of diffusions as an approximation tool of Markov Chains; in particular, the Wright-Fisher and Moran models, initiated by Feller [8] with the celebrated *Feller diffusion*.

We follow a similar strategy and approximate continuous time Markov processes. However, our work is focussed on the scaling limit (fluid/hydrodynamic limit) of a generalized discrete logistic equation and its random perturbations.

Our study of evolutionary population models, with genetics and epidemic interpretations, is not based on branching processes theory. Instead, the problem of gene mutation is examined following [21] and expands on the ideas from [13], were evolution is modeled via a Markov chain on the space of alleles acting independently except at a special fixation site. Among other things, we work in continuous time. In the discrete setting the scaling limit may reach chaotic behavior [16], but that is absent in our setup.

1.2 Previous Work

An important theme in genetic models is scaling, which links the microscopic to the macroscopic dynamics and the transition between random and deterministic. A vast literature exists (see Ewens [7] for genetics and De Masi et al. [4] for scaling limits). We shall focus on a particular model we develop in the present work. In [21] the expected value of time needed for a certain gene mutation to occur is calculated explicitly with a combinatorial argument. The model assume an "in-parallel" evolutionary process. More precisely, evolution is described as the process of guessing a preferred word of length L with letters from an alphabet of size N. Here, the "in-parallel" assumption means that if, at any round of guessing, here done by mutation, we select some letters correctly, then we keep these and only update the remaining letters. To simplify, the preferred word will be the vector zero. The characters are independent and as soon as one reaches zero, fixation occurs, and that character never moves away. This is called the *ratchet effect* in genetics. The asymptotic value as $L \to \infty$ of the *time for evolution* τ is studied, this is defined to be the first time to reach a prescribed word of length L using an alphabet with N letters. If the *time* necessary to reach the preferred word is of large order in L, evolution is not realistic. However, here it amounts to *the maximum* of L independent geometric random variables. Naturally, it is the minimum that is trivial; the maximum proves to have expected time $E[\tau] \sim N \ln(L)$, which is "short" in L, and provides an elegant but simple argument for the feasibility of evolution.

An extension of the model in [21] is considered in [13], where fixation is not certain, characters having a probability $\gamma \geq 0$ to escape zero; the [21] case corresponds simply to $\gamma = 0$. It is at this point that the model becomes suitable to other interpretations, as the accumulation of deleterious effects leading to cancer and the epidemiological view based on contamination/recovery controlled by the same γ. Both are discussed in detail in Section 2.3.

We note that as long as $\gamma = constant$ (uniform model), the characters evolve independently. With configuration dependence we enter the realm of interacting particle systems. Now probability γ is allowed to be a function of the whole configuration vector, more precisely depending on the average of the configuration (interacting case, mean-field).

Theorem 3 in [13] shows that the empirical measure converges to the solution of a discrete logistic equation with possible nonzero steady state. We start our work with Theorem 2.5, which proves the analogue result in continuous time, with limit defined by a generalized logistic equation (deterministic).

1.3 Main Results

As we explained, the presence of escape probability γ is central to our problem and allows us to cover, besides evolutionary models (genetics), a cancer development model and the propagation of infectious diseases (epidemiology).

We develop the continuous time version of the model studied in [13] deriving its fluid limit, and studying several types of random perturbations, together with applications.

In our setting, the evolution is described by a pure jump, continuous time Markov process in which each ordinary letter waits an exponential time and then updates by selecting a new character in the alphabet uniformly. Those letters matching the characters in the preferred word have a probability γ of returning among ordinary states, and, conditional on that, distribute uniformly. The number of particles is L and the alphabet has size N, where $L \to \infty$ is the scaling factor. As before, it will turn out to be essential that γ depends on the configuration, leading to an *interacting particle system* (cf. De Masi et al. [4] and Landim [14]).

The thesis presents the mathematical model in Chapter 1 and pursues its study on three main topics divided into the other chapters: Chapter 2, regarding the *Fluid Limit*, i.e. a Law of Large Numbers for a (random) dynamical system; Chapter 3, presenting the *Fluctuation near Equilibria*, a fine scale (second order approximation) result; and Chapter 4, on the *Generalized Logistic Equation with Noise*, exploring the relationship between quasi-stationarity and stable equilibria, a random perturbation question.

1.3.1 Fluid Limit and Applications to Biology

Our first result, **Theorem 2.5**, is the continuous version of Theorem 3 in [13]. We show that the empirical measure converges to the solution of the deterministic generalized logistic equation (2.8).

Section 2.3, while not listed separately, illustrates the main applications, and could figure as a main result in in its own right. We believe it provides a simple pattern, yet rich enough to make sharp differences between regimes of recovery, define the intervention time, intensity of the disease and treatment. Equation (2.8) is analyzed in the light of population evolutionary models. We interpret the parameters c to be the intensity of treatment, and a to be the intensity of the disease. We let γ to follow, as before, the power law $\gamma(u) = cu^a$. This gives enough flexibility to study several regimes when $a < 1$ and $a \geq 1$ with one, respectively two stationary solutions, as well as the complete picture when there are two equilibrium points. A discussion about recovery, successful treatment, successful detection, among other topics is presented in **Section 2.3.8**, which summarizes these results.

1.3.2 Fluctuation Near Equilibrium

A different scaling is introduced to study behavior of the system (2.8) around its equilibrium points u'. It is shown that, under mild regularity conditions for the direction field function $H(u)$ introduced in (2.7), the system behaves as an Ornstein-Uhlenbeck diffusion (henceforth OU). This is our second main result, stated in **Theorem 3.1**. It is related to *slow-fast dynamical systems*, see Berglund et al. [2], p 9. We note that our peturbation is not obtained by superposing noise - it comes instead directly from the underlying Markov process before the scaling factor L goes to infinity.

The value of this approach rests in the fact that, in contrast with the fluid limit case (2.8), here it is possible to cross a stable point u' and reach high values *even when starting from below* u'. This is important in applications. Of course, the opposite is possible, namely to cross the singularity from above. For instance, it would be interesting to study the dependence on noise and other parameters of the probability to escape on the positive side when the initial value is negative. On the other hand, at stable points, the intensity is negative and the OU process is recurrent, with well known invariant measure, which allows many explicit calculations see [3].

1.3.3 Generalized Logistic Equation with Noise

In Section 4 we allow randomness in equation (2.8) in the form of a noise term. To make clear, the scaling from Subsection 1.3.2 was possible *only at equilibrium points*. To take advantage of noise at other points, we need to superpose it in the usual way.

The influence of such noise is determined by a coefficient σ, taken to be constant in this treatise but proposed to be depending on space and/or time in future research. This gives rise to the stochastic differential equation (4.2). We impose boundary conditions related with the biological interpretation suggested in Chapter 2.

When the noise is absent, a generalized logistic equation governs the limiting process. For a small noise, the process is a diffusion absorbed at zero. All equilibria of the ODE, with the exception of zero, are now absent from the picture. We conjecture that their presence is felt as points of higher mass for the *Quasi Stationary Distribution* (qsd) which is defined in Chapter 4. More specifically, when the noise tends to zero, the qsd, which depends on the noise (a function of the size of the diffusion coefficient), should approach a discrete distribution concentrated at stable points. Since we only have one interior stable point, according to Section 2.3.3, the qsd approaches the Dirac distribution at that point. This is our conjecture given in **Theorem 4.1**.

Three questions are posed: the question of extinction (reaching the particular state 0) of the solution process of (4.2); the existence and uniqueness of the quasi-stationary distribution (qsd); and the asymptotic perturbation result described in the preceding paragraph.

The question of extinction is answered in **Proposition 4.2**. It is shown that the solution process of (4.2) goes extinct with probability one. A formula for the time of extinction is included in Proposition 4.3.

To answer the second question we use the theory of Sturm-Liouville compact operators. **Theorem 4.4** proves the existence and uniqueness of the quasi-stationary distribution.

For the last question, numerical analysis is included in order to justify the conjecture from Theorem 4.1 in **Section 4.6.2**. Particular solutions of the Sturm-Liouville problem are computed for different values of σ: $a = 0$ (linear drift) in Figure 4.3, but we emphasize the non-linear case depicted in **Figure 4.6**.

1.4 Mathematical Model and Notations

This follows a genetic model introduced in [21] aimed at estimating the so called *time for evolution* as a function of the genome length. The case was to prove that random mutations can lock in a certain configuration in logarithmic, and thus, achievable time.

This simple model is extended in [13] by allowing a small probability of mutation after reaching the preferred evolutionary state, in [13] it is considered the case when all values are zero. In the present work, *the evolutionary biology aspect is not pursued*, instead we shall apply the same mathematical construction to analyze cell pathology and possible recovery, as well as the parallel epidemiological model.

Consider the set of alleles of length L with N possible types, which are determined by chromosome structures. Mathematically, this set can be described as the set of words of length L formed with letters from an alphabet of size N, both positive integers. We will represent the alphabet as $\mathbb{Z}_N = \{0, 1, ..., N-1\}$, as $N \geq 1$ and the set of words of length $L \geq 1$ using the alphabet by $S = \mathbb{Z}_N^L$. This will be the state space of a pure jump continuous time Markov process $(\mathbf{Z}_t^L)_{t\geq 0}$.

For convenience, we chose $0 \in \mathbb{Z}_N$ (zero) as the singular value of interest. The evolution depends on the number of letters equal to this special value.

We denote the vectors, or words, in S by bolded upper case letters (e.g. $\mathbf{Z}^L \in \mathbb{Z}_N^L$), while a component is denoted by regular upper case letters with super indexes (e.g. $Z^j \in \mathbb{Z}_N$). For a given vector $\mathbf{Z}^L = (Z^1, \ldots, Z^L) \in S$, the notation $\mathbf{Z}^j(k)$ represents the vector $\mathbf{Z}^L$ with the j^{th} component updated with the letter $k \in \mathbb{Z}_N$. i.e.

$$\mathbf{Z}^j(k) = (Z^1, \ldots, \underbrace{k}_{j^{th}}, \ldots, Z^L) \tag{1.1}$$

Let us denote by $Z_t^j \in \mathbb{Z}_N$ the j^{th} component, $1 \leq j \leq L$ of the configuration at time $t \geq 0$ of the vector $\mathbf{Z}_t^L = (Z_t^1, ..., Z_t^L)$.

In the model, the standard construction via exponential holding times (i.e. Poissonization)

Definition 1.1. *The evolution of the process is given by the transition matrix (1.2) - (1.4).*

- *Holding times between jumps are i.i.d. exponentials with intensity* $\lambda(\mathbf{Z}) \equiv L$*.*
- *At a jump time* τ*, one component* j*,* $1 \leq j \leq L$ *is chosen with probability* $1/L$ *(uniformly) and gets updated;*
- *conditional on the fact that component* $Z_{\tau-}^j \neq 0$*, it changes uniformly to* any value in the alphabet, including 0*, that is*

$$Z_\tau^j = k \quad \textit{with probability} \quad \frac{1}{N}, \qquad \textit{for all} \quad k \in \mathbb{Z}_N, \tag{1.2}$$

- *while, conditional on* $Z_{\tau-}^j = 0$*, it changes to*

$$Z_\tau^j = k \quad \textit{with probability} \quad \frac{G(Z_{\tau-})}{N-1}, \qquad \textit{when} \quad k \neq 0, \tag{1.3}$$

$$Z_\tau^j = k \quad \textit{with probability} \quad 1 - G(Z_{\tau-}), \qquad \textit{when} \quad k = 0. \tag{1.4}$$

Here $0 \leq G(\mathbf{V}) \leq 1$ (to be defined later) is a function depending on the configuration $\mathbf{V}^L \in \mathbb{Z}_N^L$, e.g. $\mathbf{V}^L = \mathbf{Z}_{\tau-}$.

In the next chapter we define the infinitesimal generator of the pure jump process (Proposition 2.7). It is standard to assume (cf. [22]) that the pure jump process $(\mathbf{Z}_t^L(\omega))_{t\geq 0}$

is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions. Moreover, with probability one, the paths of the process belong to the Skorokhod space $D([0,\infty), \mathbb{R}^L)$.

This process will be called the *microscopic particle system*. To simplify notation, we suppress the superscript L, as well as the random element $\omega \in \Omega$, henceforth denoting the process $\mathbf{Z}_t^L(\omega)$ as $\mathbf{Z}_t$.

Remark. We notice that if $G(\cdot)$ is a constant, the components are independent. In [21], the state equal to the zero vector $\mathbf{0} = (0, ..., 0) \in S$ models an ideal state of the random evolutionary model, the only one where fixation is possible, and the expected time to reach it is calculated with an exact asymptotic formula as $L \to \infty$. The nonzero value for G was introduced in [13].

We introduce the associated process $U_t = U(\mathbf{Z}_t) = \sum_{j=1}^{L} \mathbf{1}_{(Z_t^j \neq 0)}$. This is the process equal to the number of non-zero components of the vector $\mathbf{Z}_t \in S$. If, in addition, there exists $\gamma : [0,1] \to [0,1]$ continuous such that

$$G(\mathbf{Z}_t) = \gamma\left(\frac{U_t}{L}\right), \quad t \geq 0, \qquad \mathbf{Z}_t \in S, \quad 0 \leq U_t \leq L, \tag{1.5}$$

then U_t, $t \geq 0$, corresponding to $\mathbf{Z}_t$, is a Markov chain on the space $\{0, 1, ..., L\}$.

As established in [13], the generating function of the transition probabilities associated to the process U_τ, at a jump time τ is given by,

$$E\left(s^{U_\tau} | U_{\tau-} = U\right) = \left(\frac{1}{N} + \left(1 - \frac{1}{N}\right)s\right)^{U} \left(1 - \gamma(\frac{U}{L}) + \gamma(\frac{U}{L})s\right)^{L-U}. \tag{1.6}$$

In other words, after jump, U_τ is the sum of two independent binomials, one with

$U = U_{\tau-}$ trials and probability of success (to remain non-zero) $1 - \frac{1}{N}$, and one with $L - U$ trials corresponding to the zero components, with probability of success (i.e. to convert into a non-zero character) equal to $\gamma(\frac{U}{L})$. The presence of the factor L at the denominator indicates dependence on the empirical measure (relative frequency) of the state 0, as in

$$u_t^L = u^L(\mathbf{Z}_t) = \frac{U_t}{L} = \frac{1}{L}\sum_{j=1}^{L} \mathbf{1}_{(Z_t^j \neq 0)} \,. \tag{1.7}$$

This points out a *mean-field* dependence, leading to the natural scaling of a Law of Large Numbers. When in time-dependent setup and established for dependent particles, such a scaling limit is known as a *fluid limit*. The empirical measure of the zero states, here simply $\frac{U_t}{L}$, converges in probability to the deterministic solution of an ode (2.8).

This is our first result, Theorem (2.5). It is the continuous time analogue of Theorem (3) in [13].

Chapter 2

Fluid Limit

2.1 Preliminaries

We begin by stating some general definitions and propositions which are necessary for the development of the theory presented below.

Definition 2.1. *Let* $(\Omega, \mathcal{F})$ *be a sample space. A family* $\mathcal{F}_t = \{\mathcal{F}_t \,|\, t \geq 0\}$ *of* σ*-algebras such that for all* $t \geq 0$, $\mathcal{F}_t \subset \mathcal{F}$ *and*

$$0 \leq s < t \implies \mathcal{F}_s \subset \mathcal{F}_t$$

is called a Filtration on $(\Omega, \mathcal{F})$.

In general every process (X_t) induces a filtration $\hat{\mathcal{F}}_t = \hat{\sigma}(X_s, 0 \leq s \leq t)$, the $\hat{\sigma}$*-algebra* induced by the truncated process $\{X_s,\ 0 \leq s \leq t\}$. The notion of a filtration $\mathcal{F}_t$ can be thought as information up to time t.

We prefer to work with processes that take into consideration this information and, in some cases, might not depend on it. The following definition presents an special type of processes that behave in the above mention manner.

Definition 2.2 (*Martingales*)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. An stochastic process* $(M_t)_{t\geq 0}$ *is a martingale with respect to a filtration* $\mathcal{F}_t$ *if*

(i) $\forall x \in \mathbb{R}$, $\{\omega : M_t(\omega) \leq x\} \in \mathcal{F}_t$ *(i.e: M_t is $\mathcal{F}_t$ - measurable for all $t \geq 0$);*

(ii) $\mathbb{E}(|M_t|) < \infty$ *for every $t \geq 0$;*

(iii) $\mathbb{E}(M_t|M_s) = M_s$ *for $s \leq t$.*

Let $\mathbb{X}$ be the state space. In the most general case $\mathbb{X}$ can be taken a Polish space, i.e. a complete separable metric space. In our applications, the state space is a subspace of $\mathbb{R}^d$, in some cases $d = L$ and in other cases $d = 1$. In this setup, we assume $\mathbb{X}$ has a norm, denoted by $||\cdot||$.

A pure jump Markov process on $\mathbb{X}$ is defined by its generator $\mathcal{A}$. Let $p(x, dy)$, with the property $p(x, \{x\}) = 0$, be a family of probability measures on the state space, measurable in x, in the sense that

$$x \to \int g(y)p(x, dy) \qquad \text{is measurable for any } g \in C_b(\mathbb{X})$$

and $\lambda(x) \geq 0$ be measurable functions, uniformly bounded by a constant $\sup_{x\in\mathbb{X}} \lambda(x) \leq ||\Lambda|| < \infty$. Then, for any $g \in C_b(\mathbb{X})$, let

$$\mathcal{A}g(x) = \lambda(x) \int (g(y) - g(x))p(x, dy)\,. \tag{2.1}$$

It is a standard construction to obtain $(Y_t)_{t\geq 0}$, on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$. filtration Conditional on being at state x, the process has holding time of intensity $\lambda(x)$ and jumps instantaneously to a random point Y with distribution $p(x, dy)$. We can assume the filtration satisfies the *usual conditions*.

Associated to the generator, two fundamental martingales emerge.

Proposition 2.3 (Appendix 1 in [14])**.** *For every g continuous and bounded, the processes*

$$M_t^g = g(Y_s) - g(Y_0) - \int_0^t \mathcal{A}g(Y_s)ds \tag{2.2}$$

$$N_t^g = (M_t^g)^2 - \int_0^t \left[\mathcal{A}g^2(Y_s) - 2g(Y_s)\mathcal{A}g(Y_s)\right] ds \tag{2.3}$$

are martingales w.r.t the filtration $(\mathcal{F}_t)_{t\geq 0}$.

This is true for more general Feller processes, but in this case the underlying process (Y_t) *is a* pure jump process*, then*

$$\langle M^g \rangle_t = \int_0^t \lambda(Y_s) \int_{\mathbb{X}} p(Y_s, dy)\,(g(y) - g(Y_s))^2\, ds. \tag{2.4}$$

Remark. The integrands in (2.2)-(2.4) depend on the process at time $s-$, not s; for example Y_s should be Y_{s-} throughout. However, from the Lebesgue-Stieltjes integration by parts formula, because ds, the Lebesgue measure, is continuous, the values coincide over the jump.

Pure jump processes, and a large class of Feller processes can be canonically constructed on the Skorokhod space $\mathcal{D}([0,\infty), \mathbb{X})$ of right-continuous with left-limit paths (rcll, also known as cádlág). Tightness is the notion of pre-compactness of probability laws defined by Prokhorov's theorem.

Definition 2.4. *A sequence of processes* $(Y^L_\cdot)_{L>0}$ *on a Polish space* $(\mathbb{X}, \|\cdot\|)$ *with right-continuous with left limits paths (in the Skorokhod space) is* C-tight*, if for any* $T \geq 0$

$$(i) \qquad \lim_{M\to\infty} \limsup_{L\to\infty} P\Big(\|Y^L_T\| > M\Big) = 0 \quad \textit{and} \tag{2.5}$$

$$(ii) \quad \forall\, \epsilon > 0 \qquad \lim_{\delta\to 0} \limsup_{L\to\infty} P\Big(\sup_{t,t'\in[0,T],|t'-t|<\delta} \|Y^L_{t'} - Y^L_t\| > \epsilon\Big) = 0\,. \tag{2.6}$$

C-tightness refers to the fact that the family of processes indexed by $N > 0$ is not only tight as a family on the Skorokhod space with the J_1 topology, but in addition, any limit point is continuous. This result is standard in proving convergence of pure jump processes to processes with continuous paths, as either diffusions or simply solutions of classical differential equations. See for instance the general treatise [14] and a similar derivation in [24].

2.2 Statement and Proof of Theorem 2.5 (Fluid Limit)

With $\gamma(u)$ defined in (1.5), let

$$H(u) := \left(-\frac{1}{N}\right)u + \gamma(u)(1-u)\,, \qquad 0 \leq u \leq 1\,. \tag{2.7}$$

Theorem 2.5. *Assume that* $u^L_0 = \frac{U_0}{L}$ *converges in probability, as* $L \to \infty$*, to the deterministic state* $\bar{u} \in [0,1]$ *and* $\gamma = \gamma(u)$*,* $0 \leq u \leq 1$ *from (1.5) is continuous. Then, as* $L \to \infty$*, the Markov process with state space* $[0,1]$ *equal to* $u^{(L)}_t = \frac{U_t}{L}$ *for all* $t \geq 0$*, converges in* *distribution to the deterministic process* $(u_t)_{t\geq 0}$ *on* $[0,1]$, *equal to the unique solution of*

$$\frac{du}{dt} = H(u)\,, \qquad u(0) = \bar{u}\,. \tag{2.8}$$

Remark. Notice that since $0 \leq \gamma(u) \leq 1$, we can see that $u(t) \leq \max\{\bar{u}, \frac{N}{N+1}\} \leq 1$. However, $u(t)$ may reach zero (extinction) in finite time (if zero is not an equilibrium) or approach it arbitrarily closely.

The proof of Theorem 2.5 will be done in multiple steps.

Step 1. We write the action of the generator $\mathcal{A}$, of the underlying pure jump process $(\mathbf{Z}_t)$ defined by (1.2)-(1.4), on the empirical average (u_t^L). I.e. we want to know how the generator acts on test functions of the form $g(\mathbf{Z}_t) = f(u^L(\mathbf{Z}_t))$. Since $u^L(\mathbf{Z}_t) \in [0,1]$ by construction, we may adopt functions $f \in C^2([0,1],\mathbb{R})$ and consider their extensions to the space of functions with compact support $C_c^2(\mathbb{R},\mathbb{R})$.

Step 2. We find uniform bounds for the formula of the generator $\mathcal{A}$ and for the quadratic variation of the martingale process $M_t^{f,L}$, associated with (u_t^L), given by formula (2.2).

Step 3. We prove that the process (u_t^L) is C-tight as in definition (2.4). This implies that the sequence of probability laws of the process $(u_\cdot^L)$ has a convergent subsequence to a limit probability law.

Step 4. In step four, we prove that any limit probability law of the sequence (u_t^L) must solve the initial value problem (2.8). Existence and uniqueness of solutions for ordinary differential equations gives that there is a unique accumulation point for (u_t^L). Thus, we conclude that the sequence (u_t^L) is convergent, to the unique solution of (2.8). The type of convergence is proven to be *in probability*, more precise

Definition 2.6. *A sequence of processes* $(Y^L_\cdot)_{L>0}$ *on a Polish space* $(\mathbb{X}, \|\cdot\|)$ converges in probability *to* $(Y_\cdot)$*, uniformly in finite time, if for any* $T > 0$*, the process* $t \to (Y^L_t)_{t\geq 0}$ *satisfies that, for every* $\epsilon > 0$

$$\lim_{L\to\infty} P\Big(\sup_{t\in[0,T]} \|Y^L_t - Y_t\| > \epsilon\Big) = 0\,. \tag{2.9}$$

Step one: Action of the generator

We now apply the general theory to the process $(\mathbf{Z}_t)$ from Definition 1.1. Let's denote by $P(\mathbf{Z}, \tilde{\mathbf{Z}})$ the transition probability from $\mathbf{Z}$ to $\tilde{\mathbf{Z}}$. Notice that $\lambda(x)$ are constant equal to one. The generator of the process $\mathbf{Z}_t$, applied to a function $F : S \to \mathbb{R}$, is given by

$$\begin{aligned}\mathcal{A}^L F(\mathbf{Z}) &= \sum_{\tilde{\mathbf{Z}}\in S} P(\mathbf{Z}, \tilde{\mathbf{Z}})\Big(F(\tilde{\mathbf{Z}}) - F(\mathbf{Z})\Big)\\ &= \sum_{j=1}^{L}\sum_{k=0}^{N-1} P(\mathbf{Z}, \mathbf{Z}^j(k))\Big(F(\mathbf{Z}^j(k)) - F(\mathbf{Z})\Big)\\ &= \sum_{j=1}^{L}\sum_{k=0}^{N-1} \frac{1}{N}\Big(F(\mathbf{Z}^j(k)) - F(\mathbf{Z})\Big)\mathbf{1}_{(Z^j\neq 0)} + \sum_{j=1}^{L}\sum_{k=1}^{N-1} \frac{G(\mathbf{Z})}{N-1}\Big(F(\mathbf{Z}^j(k)) - F(\mathbf{Z})\Big)\mathbf{1}_{(Z^j=0)} +\\ &\quad + \sum_{j=1}^{L}(1 - G(\mathbf{Z}))\Big(F(\mathbf{Z}^j(0)) - F(\mathbf{Z})\Big)\mathbf{1}_{(Z^j=0)}\,.\end{aligned} \tag{2.10}$$

Proposition 2.7. *The matrix*

$$Q(\mathbf{Z}, \tilde{\mathbf{Z}}) = \lambda(\mathbf{Z})P(\mathbf{Z}, \tilde{\mathbf{Z}})\,, \quad \textit{if} \quad \tilde{\mathbf{Z}} \neq \mathbf{Z} \qquad \textit{and} \quad -Q(\mathbf{Z}, \mathbf{Z}) = \sum_{\tilde{\mathbf{Z}}\neq\mathbf{Z}} P(\mathbf{Z}, \tilde{\mathbf{Z}})$$

is a Q-matrix in the sense of Definition A.6 in the Appendix.

Proof. The proposition is easy to verify since the state space is finite and the probabilities are non-negative and the intensities are all constant equal to L. □

We now turn to the process u_t^L is Markov with Q-matrix given by (1.6). We recall that at jump time, the evolution of u_t^L depends on the change in the components $Z_{\tau-}^j$. Each $Z_{\tau-}^j \neq 0$ could become 0 with probability $\frac{1}{N}$; while each $Z_{\tau-}^j = 0$ could become not zero with probability $\gamma(u_t^L)$. Hence, provided that there is a change in any component of the vector Z_τ, the scaled process u_t^L might remain unchanged or change by a factor of $\frac{1}{L}$. The precise description is that for a given state $\mathbf{Z}$,

given that $Z^j \neq 0$

$$u^L(\mathbf{Z}^j(k)) = u^L(\mathbf{Z}) \quad \text{if} \quad k \neq 0$$
$$u^L(\mathbf{Z}^j(k)) = u^L(\mathbf{Z}) - \tfrac{1}{L} \quad \text{if} \quad k = 0$$

and if $Z^j = 0$

$$u^L(\mathbf{Z}^j(k)) = u^L(\mathbf{Z}) + \tfrac{1}{L} \quad \text{if} \quad k \neq 0$$
$$u^L(\mathbf{Z}^j(k)) = u^L(\mathbf{Z}) \quad \text{if} \quad k = 0$$

Let $u^L = u^L(\mathbf{Z})$, and $\mathcal{T}$ be the family of functions F such that $F(\mathbf{Z}_t) = f(u^L(\mathbf{Z}_t))$ with f having compact support containing the closed interval $[0,1]$. Then it is the case that the generator acts over the family of test functions $\mathcal{T}$ as

$$\begin{aligned}\mathcal{A}^L f(u) &= \lambda \sum_{j=1}^{L} \left[\frac{1}{N} \left(f(u^L - \frac{1}{L}) - f(u^L) \right) \mathbf{1}_{(Z_t^j \neq 0)} + \gamma(u^L) \left(f(u^L + \frac{1}{L}) - f(u^L) \right) \mathbf{1}_{(Z_t^j = 0)} \right] \\ &= L\lambda \left[\frac{1}{N} \left(f(u^L - \frac{1}{L}) - f(u^L) \right) u^L + \gamma(u^L) \left(f(u^L + \frac{1}{L}) - f(u^L) \right) (1 - u^L) \right].\end{aligned} \tag{2.11}$$

Step two: Uniform bounds for $\mathcal{A}^L f(u_t^L)$ and $M_t^{L,f}$

We begin to work on the bound for (2.15). We begin by developing the function f

around the point u_t^L using the Taylor formula with remainder of order two.

$$f(u_t^L \pm \frac{1}{L}) - f(u_t^L) = \pm\frac{1}{L} f'(u_t^L) + R_1^{\pm}(u_t^L) \tag{2.12}$$

where $R_1^{\pm}(u_t^L) = \frac{1}{2}\left(\pm\frac{1}{L}\right)^2 \int_0^1 w f''(u_t^L + (1-w)(\pm\frac{1}{L}))\, dw$. We note that, since f has compact support, the integral in the remainder is uniformly bounded, thus $R_1^{\pm}(u_t^L) \in O(\frac{1}{L})$. Replacing into (2.15) and collecting all terms we have

$$\begin{aligned} \mathcal{A}^L f(u) &= L\lambda \left[\frac{1}{N}\left(-\frac{1}{L} f'(u_t^L) + R_1^{-}(u_t^L)\right) u_t^L + \gamma(u_t^L)\left(\frac{1}{L} f'(u_t^L) + R_1^{+}(u_t^L)\right)(1-u_t^L)\right] \\ &= \lambda H(u) f'(u) + c(f,u,L)\,, \qquad |c(f,u,L)| \le c(f) L^{-1}\,, \end{aligned} \tag{2.13}$$

where $c(f,u,L)$ is the error term, which collects all expressions involving the Taylor remainders $R_1^{\pm}(u_t^L)$, and $c(f) > 0$ is a constant depending only on f, obtained form the remainder of the Taylor formula $R_1^{\pm}(u_t^L)$ applied to f. We obtain the important uniform bound on $\mathcal{A}^L f(u)$,

$$\left|\mathcal{A}^L f(u)\right| \le \lambda c(f') + c(f) L^{-1} \tag{2.14}$$

We now focus on the uniform bound for the quadratic variation of the martingale process $M_t^{L,f} = f(u_t^L) - f(u_0^L) - \int_0^t \mathcal{A}^L f(u_s^L)\, ds$. Recall that for continuous jump processes the quadratic variation can be computed by formula (2.4). This formula, combined with (2.15) gives that the quadratic variation can be expressed by

$$\langle M^{L,f}\rangle_t = \lambda \sum_{j=1}^{L} \left[\frac{1}{N}\left(f(u_t^L - \frac{1}{L}) - f(u_t^L)\right)^2 \mathbf{1}_{(Z_t^j \neq 0)} + \gamma(u_t^L)\left(f(u_t^L + \frac{1}{L}) - f(u_t^L)\right)^2 \mathbf{1}_{(Z_t^j = 0)}\right]$$

A similar development in Taylor expansion for $\langle M^{L,f}\rangle_t$, as for the generator, gives that

$$\begin{aligned}\langle M^{L,f}\rangle_t &= L\lambda\left[\frac{1}{N}\left(f(u_t^L-\frac{1}{L})-f(u_t^L)\right)^2 u_t^L+\gamma(u_t^L)\left(f(u_t^L+\frac{1}{L})-f(u_t^L)\right)^2(1-u_t^L)\right]\\ &= L\lambda\left[\frac{1}{N}\left(\frac{1}{L}f'(u_t^L)+R_1^-(u_t^L)\right)^2 u_t^L+\gamma(u_t^L)\left(-\frac{1}{L}f'(u_t^L)+R_1^+(u_t^L)\right)^2(1-u_t^L)\right]\\ &= \frac{1}{L}\lambda\left[\frac{1}{N}\left(f'(u_t^L)+LR_1^-(u_t^L)\right)^2 u_t^L+\gamma(u_t^L)\left(-f'(u_t^L)+LR_1^+(u_t^L)\right)^2(1-u_t^L)\right].\end{aligned} \tag{2.15}$$

Where the expressions $R_1^-(u_t^L)$, $R_1^+(u_t^L)$ are the remainders in the Taylor formula (2.12), hence are of order $O(\frac{1}{L^2})$. Thus, the quadratic variation comprises only jumps of size $\frac{1}{L}$. We concluded that $|\langle M^{L,f}\rangle_t| < \frac{1}{L}C$, where the constant C depends only on the derivatives of f. Thus, the bound is uniform in ω and time, and $u_t \in [0,1]$.

Step three: Tightness of the process u_t^L

We now can show that for any test function $f \in \mathcal{T}$, the process $Y_t^L = f(u_t^L)$ is *C-tight*.

Condition (i) in (2.5) is trivial because $u_t^L \in [0,1]$ and f has compact support without further restrictions necessary.

Condition (ii) in (2.6) is implied by Markov's inequality and the fact that $\mathcal{A}^L f(u_s^L)$ is uniformly bounded by (2.14). So we let $\delta > 0$ and on the interval $[0,T]$, we choose $0 \le t < t' \le T$ such that $\delta > t' - t$. We have

$$\mathbb{P}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}}\left|\int_t^{t'}\mathcal{A}^L f(u_s^L)ds\right| > \epsilon\right) \le \frac{1}{\epsilon}\mathbb{E}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}}\int_t^{t'}|\mathcal{A}^L f(y_s^L)|\,ds\right) \le \frac{1}{\epsilon}\left(\lambda c(f')+c(f)L^{-1}\right)\delta.$$

Hence,

$$\lim_{\delta\to 0}\limsup_{L\to\infty}\mathbb{P}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}}\left|\int_t^{t'}\mathcal{A}^L f(u_s^L)ds\right| > \epsilon\right) = 0.$$

To obtain (2.6) for the martingale $M_t^{L,f}$, we apply the Doob's L^2-norm maximal inequality for the square norm on $[t,t']$, taking without loss of generality $0 \leq t \leq t' \leq T$. We use formula (2.3) for the martingale $M_{t'}^{f,L} - M_t^{f,L}$ and the uniform bound obtained in (2.2). More precisely

$$\mathbb{P}\left(\sup_{\substack{0\leq t<t'\leq T\\ t'-t<\delta}} \left|M_{t'}^{f,L} - M_t^{f,L}\right| > \epsilon\right) \leq \frac{1}{\epsilon^2}\mathbb{E}\left(\left|M_{t'}^{f,L} - M_t^{f,L}\right|^2\right) = \frac{1}{\epsilon^2}\mathbb{E}\left(\left|\langle M^f f, L\rangle_{t'} - \langle M^{f,L}\rangle_t\right|\right) \leq \frac{\delta}{\epsilon^2 L}C.$$

Hence

$$\lim_{\delta\to 0}\limsup_{L\to\infty} \mathbb{P}\left(\sup_{\substack{0\leq t<t'\leq T\\ t'-t<\delta}} \left|M_{t'}^{f,L} - M_t^{f,L}\right| > \epsilon\right) = 0.$$

Proposition 2.8. *The processes* $(u_t^L)_{t\geq 0}$*, indexed by* $L \in \mathbb{N}$*, are* C-tight.

Proof. We have shown that both the martingale $M_t^{f,L}$ and the integral term $\int_0^t \mathcal{A}^L f(u_s^L) ds$ are *C-tight*, hence $f(u_t^L)$ is *C-tight* for any f continuous and with compact support. Since the process $u_t^L \in [0,1]$, we can take any f such that $f(u)|_{[0,1]} \equiv id(u)|_{[0,1]}$, having compact support. We obtain that $(u_t^L)_{t\in[0,T]}$, indexed by L, is *C-tight*.

□

Step four: Limit distribution of the process u_t^L

Denote, for u_t, $t \in [0,T]$, written simply as $u_\cdot$, an rcll path in the Skorokhod space

$$\Psi(u_\cdot) := \left| f(u_t) - f(\bar{u}) - \int_0^t H(u_s) f'(u_s)\, ds \right| \tag{2.16}$$

In view of the bounds $|c(f, u_t^L, L)| \leq c(f)/L$ in (2.13), and $|\langle M^{L,f}\rangle_t| < \frac{1}{L}C$ from (2.15), toguether with the fact that the process N_t^g, defined in (2.3), is a martingale we obtain that

$$\mathbb{E}[\psi(u_\cdot^L)] = \mathbb{E}\left[\left|f(u_t^L) - f(\bar{u}) - \int_0^t H(u_s^L)f'(u_s)\,ds - \int_0^t c(f,u_s^L,L)ds + \int_0^t c(f,u_s^L,L)ds\right|\right]$$

$$\leq \mathbb{E}\left[|M_t^{f,L}|\right] + \mathbb{E}\left[\left|\int_0^t c(f,u_s^L,L)ds\right|\right] \leq \sqrt{\mathbb{E}\left[|M_t^{f,L}|^2\right]} + Tc(f)\frac{1}{L}$$

$$= \sqrt{\mathbb{E}\left[\langle M_t^{f,L}\rangle\right]} + Tc(f)\frac{1}{L} \leq \sqrt{C\frac{1}{L}} + Tc(f)\frac{1}{L}.$$

Thus we obtain that

$$\lim_{L\to\infty} E\Big[\Psi(u_\cdot^L)\Big] = 0\,. \tag{2.17}$$

Given that Ψ is a bounded continuous functional on the Skorokhod space (see [24]), let $u_\cdot$ be a limit point of the *C-tight* family $u_\cdot^L$. Then $E[\psi(u_\cdot)] = 0$, implying that, with probability one, the possibly random continuous process $u_\cdot$ satisfies

$$f(u(t)) - f(\bar{u}) - \int_0^t H(u_s)f'(u_s)\,ds = 0\,.$$

This identity is valid for any $f \in C_c^2(\mathbb{R})$; in particular taking $f(u)|_{[0,1]} \equiv id(u)|_{[0,1]}$ with compact support once again, shows that any possible continuous limiting path solves

$$u_t - \bar{u} - \int_0^t H(u_s)\,ds = 0\,,$$

which is exactly (2.8) in integral form. The function H is continuous, so is $u_\cdot$, thus the integrand is continuous, implying that u_t is differentiable in classical sense. We proved that any limit point solves the ode (2.8). Moreover, H is a Lipschitz function, having continuous derivative on $[0,1]$, proving uniqueness of the solution. We proved that any limit point *must be equal to the unique solution to the initial value problem (2.8)*. Finally, the convergence takes place in distribution, to a delta function concentrated on the unique solution of the ode. Convergence in distribution to a delta measure is equivalent to convergence in probability. Because all along the convergence was uniform in time over $[0, T]$, T fixed but arbitrary, we concluded the proof of Definition 2.6, and hence of the theorem.

2.3 Applications to Biological Models

Both the random process U_t (1.6) - microscopic - and its deterministic scaling u_t shown in (2.8) - macroscopic - can be used as population evolution models. Two main setups are proposed: (i) *the cancer development model* and (ii) *the epidemic model*. Then U_t, out of the total population L, is understood as the set of alleles in non-deleterious states in (i) and as the non-infected population in (ii). After scaling, u_t is the averaged value, consistent with a law of large numbers for empirical measures, calculated out of a normalized population of size one.

Equation (2.8) is valid for a general continuous function $\gamma(u)$. Since it is an autonomous equation, we are interested in stability about the equilibria u, given by the solutions of $H(u) = 0$. In both micro - and macroscopic models, we are still interested in the relation between the solution and the sensitive state 0 (zero). For this reason, we shall adopt models when $\gamma(0) = 0$, so that $u = 0$ is an equilibrium point. In general, since the initial value $\bar{u}$ is non-negative and the equation is autonomous, the solution u_t remains non-negative.

Evidently, a state of zero would be absorbing, as the intrinsic condition is that recovery depends with positive co-relation on the non-deleterious/infected population u_t. To satisfy that assumption, the mathematical model proposed, for both cancer and epidemic examples, will be a power law for the *probability of recovery*

$$\gamma(u) = cu^a , \quad \text{for} \quad 0 \le c \le 1 , \quad a \ge 0 . \tag{2.18}$$

The framework is as follows. In the natural state, the population follows (2.18) with $c = 0$ (i.e: $\gamma \equiv 0$). This is the pre-intervention level. It is assumed hereby that an exponential rate of aging/contamination (cancer and epidemic, respectively) drives the healthy population down, towards eventual extinction, *in the absence of treatment*, here represented by $\gamma(u)$. An empirical level $u = \alpha \in (0, 1)$ designates the *detection threshold*. It is only as soon as u drops below α that tests or symptoms make the disease detectable. At this point, an intervention takes place, with a specific *probability of recovery* $\gamma(u)$ (2.18) depending on the "healthy" proportion of the population u; the strength of the treatment c; and the intensity or virulence of the disease a.

A couple of observations are in order to motivate the definition of γ. First, notice that this function is increasing in u meaning that, for bigger u our probability of recovery is greater, consistent with u being the healthy proportion of the population. Also, if we fix $u > 0$ and consider $a \to \infty$, then the function $\gamma(u) = cu^a \to 0$, thus the probability of recovery reduces as the intensity of the disease increases.

Equation (2.8) depends on the *logistic* factor $\gamma(u)$. Since $G(\mathbf{Z})$ from eq. (1.5) belongs to $[0, 1]$ macroscopically (before letting $L \to \infty$), it is the case that $\gamma(u) = cu^a \in [0, 1]$ as well. Thus, to avoid technical complications, we adopt $c \in [0, 1]$ while $u \in [0, 1]$. As explained in the previous paragraph, for $a \ge 0$, the power function is increasing in u but decreasing in a, again consistent with the model interpretation. The particular cases $a = 0$ and $a = 1$ are studied in [13] in a discrete time setting. The continuous time case is briefly discussed in subsections 2.3.5 ($a = 0$) and 2.5.1, Proposition 2.11 ($a = 1$).

2.3.1 The Intrinsic Parameters

These are parameters intrinsic to a disease and its treatment. They are imbedded in the recovery probability function $\gamma(u)$, i.e. the intensity of the disease, respectively of the intervention, as well as N. In the cancer growth model setup, $1/N$ is the probability of a deleterious mutation, a feature of aging, and intrinsic to the cell; c is the effectiveness of the intervention (treatment); a is the cancer type, or aggressiveness. In the epidemic model, $1/N$ is the contagiousness of the disease (e.g. probability to contract the virus); c is the strength of a treatment of vaccine, and a is the virulence of the disease.

2.3.2 The Extrinsic Parameters

As opposed to the parameters defining the disease, these parameters are set independently. We postulate two values $\alpha, \beta \in [0, 1]$, where $u = \alpha$ is the *detectability* level and $u = \beta$ is the *quality of life*, a satisfactory health threshold, especially in the cancer setting. In the epidemic model it is the containment level, at which the population is considered out of an epidemic state. It is natural to consider $\alpha < \beta$. The case $\alpha \geq \beta$ is practically trivial and prophylactic care would prevail, as it allows a priori early detection.

2.3.3 The Equilibrium Values

All cases in the power law model have at least one stable equilibrium in the interval $[0, 1)$. If $c = 0$, then $u_0 = 0$ is the only equilibrium value (Subsection 2.3.4). If $c > 0$, then

(i) If $a = 0$, there exists only one equilibrium value u_2 (Newton's equation, subsection 2.3.5);

(ii) If $a > 1$, the number of equilibrium points of the system (2.8) will depend on the parameter a, varying from only one equilibria to either two or three. In this case the point $u_0 = 0$ is always a stable equilibrium. In the case of the presence of just another non zero-equilibrium point u_1, this will be half-stable. Finally, the case of three points $0 = u_0 < u_1 < u_2 < 1$, will exhibit u_1 as unstable and u_2 as stable - see Section 2.4.

(iii) If $0 < a \leq 1$, there exist two equilibrium values $0 = u_0 = u_1 < u_2 < 1$, where u_0 is unstable and u_2 is stable - see Section 2.5.

2.3.4 Natural State of the System $c = 0$

A natural state of the system is when there is no intervention, i.e. $c = 0$. In this case, the only dynamics is due to aging $u(t) = u(0)e^{-\frac{1}{N}t}$ determined by the exponential rate of decay $1/N$, with unique stable equilibrium at $u_0 = 0$ (see Section 1.1 in [9]).

Remark. The equilibrium value u_1 is a function of Nc and a, and thus ΔT is a function depending on the dynamical system (2.8) and not just on the trivial exponential decay. This time for detection successfully describes the interplay between the equilibrium point u_1 and the threshold α. Notice that if $u_1 > \alpha$ the time of intervention is negative, which means that the detection was late.

We now start analyzing the power law model (2.18) case by case.

2.3.5 Case $c > 0$, $a = 0$

This borderline case, when $a = 0$ and $c > 0$, the solution is the so called Newton's equation (usually of temperature) approaching its unique stable nonzero equilibrium $u_2 = (1 + \frac{1}{Nc})^{-1}$ exponentially fast. In our interpretation, this permits an intervention since $u_2 > 0$; yet, it is successful only if $u_2 \geq \beta$ as $c \uparrow 1$.

2.3.6 Case $c > 0$, $a > 0$

This is by far the most important case and is treated in detail in Sections 2.4 ($a > 1$), respectively 2.5 ($0 < a \leq 1$). We recall that the equilibrium points of the system (2.8) are given by the zeroes of the function in (2.7). With the power law,

$$H(u) = cuf(u)\,, \qquad f(u) := u^{a-1}(1-u) - \frac{1}{Nc}\,, \quad u \in [0,1]\,. \tag{2.19}$$

Let u be an equilibrium point of the dynamical system (2.8), then $H(u) = 0$ which implies that either $u = 0$ or, for the non-zero equilibrium points of (2.8), are the solutions $u > 0$ of $f(u) = 0$. It will be shown that there are at most two $0 < u_1 \leq u_2 < 1$ such solutions.

The analysis of the stability of the equilibrium points of (2.8) will require to determine the sign of H' (see, in Section 2.4, p 24 in [18]) or, because we are in dimension one, equivalently, just the sign of H. We write H and its derivative H' in terms of f, and focus essentially on the analysis of the function f. Any result that we obtain for the function f will easily imply the corresponding consequence for the functions H, H' via the formulas

$$H(u) = cuf(u)\,, \quad H'(u) = c\left(f(u) + uf'(u)\right)\,. \tag{2.20}$$

Without loss of generality, we assume $c > 0$ and $a > 0$, since the analysis regards the system after intervention ($c > 0$) and $a = 0$ is trivial because the recovery probability $\gamma(u)$ does not depend on u. Then $0 = u_0 \leq u_1 < u_2 < 1$.

2.3.7 Time Window for Detection

Let u_1 be the unstable equilibrium, where $u_1 > 0$ for $a > 1$ and $u_1 = u_0 = 0$ when $a \in (0, 1]$. The simple but important *intervention window* ΔT between reaching α and u_1 is defined by

$$\Delta T = T_{u_1} - T_\alpha = -N \ln(\frac{u_1}{\alpha}), \tag{2.21}$$

noticing that $\Delta T = +\infty$ when $u_1 = 0$. This is important because it prescribes the *time between tests or checkups*, obviously meaningful only when $a > 1$, which corresponds to a more aggressive disease.

2.3.8 Discussion and Interpretation of the Results

The most complex case, present only when $a > 1$, is when there are two nonzero stable points. This is characterized exactly in Proposition 2.9 eq. (2.22) in case (2). The following discussion can be applied to the other cases, with the corresponding simplifications.

It is the relation between the stable points (equilibria) and α, β that decides the outcome of the treatment c, applied to the disease, identified by the parameter a. From the outset, we see that $\alpha < u_1 < \beta$, detection will occur too late. When $u_2 < \beta$, one can never regain a satisfactory health level, even after detection. Thus the ideal configuration is $u_1 < \alpha < \beta < u_2$, as seen below.

• **Successful treatment.** When $u_1 < \alpha < \beta < u_2$, if detection occurs at a state $u \in (u_1, \alpha)$, then recovery is achieved as the solution evolves towards u_2. Detection occurs if testing is done at intervals not greater than $\Delta T = T_\alpha - T_{u_1} = N \ln(\frac{u_1}{\alpha})$.

• **Successful detection, insufficient treatment.** Here $u_1 < \alpha < u_2 < \beta$. Detection is successful but the treatment achieves a state u_2 that may be pathological/endemic.

• **Ineffective detection, unsuccessful prophylactic treatment.** If $\alpha < u_1 < \beta < u_2$. Detection would be too late, but prophylactic treatment would prevent the disease/epidemic.

• **Non intervention case, follow up.** If $u_1 < u_2 < \alpha < \beta$, detection is early and a follow up is required. No treatment should be necessary.

• **Ineffective detection and treatment.** If $\alpha < u_1 < u_2 < \beta$, detection is too late and treatment would be ineffective. The most pessimistic scenario.

• **Ineffective detection, successful early treatment.** If $\alpha < \beta < u_1 < u_2$, detection is too late and only early treatment would be effective.

While a modulates the aggressiveness of the disease, making γ smaller, c would push it up. For a given a, u_2 is increasing, and u_1 decreasing, in c. In the first application model introduced before, the action of increasing c is equivalent to improve the treatment. Since $\gamma(u)$ is a probability, $0 \leq c \leq 1$, thus we could improve treatment up to $c = 1$. It is of interest to see the optimal values of u_1, u_2 we can achieve, as shown, for example in case $a > 1$, in Proposition 2.10.

2.4 The case $a > 1$

Proposition 2.9. *If $\gamma(u) = cu^a$, $a > 1$, then the point $u_0 = 0$ is a stable equilibrium point of the system and, the number of equilibrium points in $[0, 1]$ of the dynamical system (2.8) is determined by the sign of the number q, defined as*

$$q = \frac{1}{a}\left(1 - \frac{1}{a}\right)^{a-1} - \frac{1}{Nc} \tag{2.22}$$

1. *if $q < 0$, the only equilibrium point is $u_0 = 0$.*

2. *if $q > 0$, there are 3 equilibrium points $0 = u_0 < u_1 < u_2 < 1$.*

The points u_0 (seen for $u \geq 0$) and u_2 are stable, and u_1 is unstable.

3. *if $q = 0$, there are two equilibrium points, namely*

$$u_0 = 0 \qquad \textit{and} \qquad u_1 = u_2 = u_m = \frac{a-1}{a},$$

where u_m is the maximizer point of f on $[0,1]$; u_0 is stable, and u_m is half-stable.

The above proposition describes a change in behavior of the system (2.8) as the parameter a ranges in $[0, 1]$. It is useful to have the phase of portrait of each situation in mind.

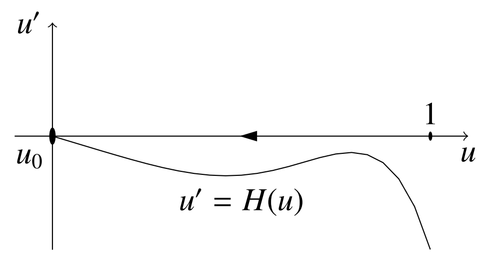

Figure 2.1: *Phase Portrait case $q < 0$, obtained by selecting $a = 5$, $N = 10$, $c = 1$. The unique equilibrium point u_0 is stable.*

Proof. Part 1. Number of equilibrium points. The function f defined in (2.19) satisfies $f(0) = f(1) = -\frac{1}{Nc} < 0$, thus it has an extreme value in the interval $(0, 1)$. The derivative of the function f is given by

$$f'(u) = au^{a-2}\left(\frac{a-1}{a} - u\right), \tag{2.23}$$

hence it has a unique zero at the point $u_m = \frac{a-1}{a}$ on the interval $(0, 1)$. This point is a global maximum over the interval $[0, 1]$. Hence the function f, over the same interval, is less or

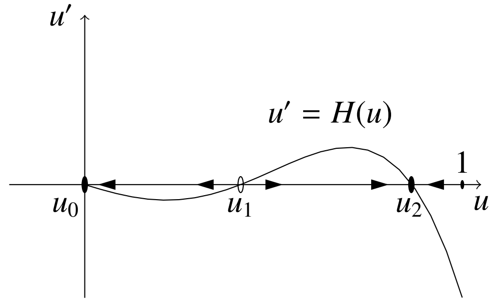

Figure 2.2: *Phase Portrait case $q > 0$, obtained by selecting $a = 3$, $N = 10$, $c = 1$. The equilibrium points u_0, u_2 are stable, and u_1 is unstable.*

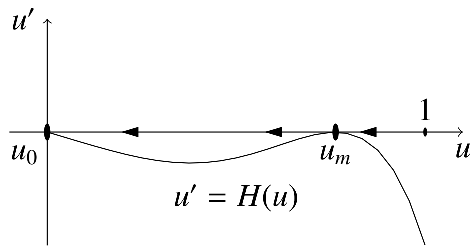

Figure 2.3: *Phase Portrait case $q = 0$, obtained by selecting $a = 4.2$, $N = 10$, $c = 1$. The equilibrium point u_0 is stable, and u_m is half-stable.*

equal to the value q given by

$$q = f\left(\frac{a-1}{a}\right) = \frac{1}{a}\left(1 - \frac{1}{a}\right)^{a-1} - \frac{1}{Nc}\,. \tag{2.24}$$

The function f is continuous and it is strictly monotone restricted to the intervals $[0, u_m]$ and $[u_m, 1]$; then f is injective on each of these intervals. If $q < 0$, the equation $f = 0$ has no solutions. If $q = 0$, a unique solution is obtained at u_m, and if $q > 0$, then there are two solutions that we denote by $0 < u_1 < u_2 < 1$.

Part 2. Stability. For the point $u_0 = 0$ we have that $H'(0) = -\frac{1}{N} < 0$, and thus it is always a stable point.

In the case $q > 0$, let's denote these points by $u_0 = 0$, u_1, and u_2, with $u_1 < u_2$ as before. We notice that $u_1 \in (0, u_m)$ and that $u_2 \in (u_m, 1)$. Hence, we have from 2.20 that, $H'(u_1) = cu_1 f'(u_1) > 0$, and hence u_1 is an unstable point. Similarly, we have that $H'(u_2) = cu_2 f'(u_2) < 0$ which means that u_2 is a stable point.

In case $q = 0$, we have that $f(u_m) = f'(u_m) = 0$. Since f attains global maximum at u_m we have that: $H(u) = cuf(u) < cuf(u_m) = 0, \forall u \in [0, 1]$. Hence, the point u_m is half-stable.

□

Proposition 2.10. *(i) The function $q = q(a)$ is decreasing for $a \geq 1$ with maximum value at $a = 1$ equal to $q(1) = 1 - \frac{1}{cN}$, equal to the limiting value of u_1 when $a \downarrow 1$ and minimum value $q(\infty) = -\frac{1}{cN} < 0$.*

(ii) If $q(1) \leq 0$, the only equilibrium point is $u_0 = 0$ and hence no recovery is possible.

(iii) In case $q(1) > 0$, for sufficiently small $a > 1$, since $q(a) > 0$ the system is in case (1) of Proposition 2.9.

(iv) Additionally, for such values of a, as the treatment *c satisfies $c \uparrow 1$, the values of $u_1 \downarrow u_1(1)$ and $u_2 \uparrow u_2(1)$, showing that the test of optimality (at $c = 1$, most efficient treatment) is if the double inequality $0 \leq u_1(1) \leq \alpha < \beta \leq u_2(1) < 1$ is satisfied.*

(v) The critical values are a_- such that $u_1(a_-) = \alpha$ and $a_+ = u_2(a_+) = \beta$ solving

$$a_- = 1 - \frac{\ln(N(1-\alpha))}{\ln \alpha}, \qquad a_+ = 1 - \frac{\ln(N(1-\beta))}{\ln \beta}. \tag{2.25}$$

Let $a_* = a_*(cN)$ be the solution of $q(a) = 0$, always satisfying $a_* < N$. In the interval $a \in (1, a_*)$, the upper critical value u_2 sweeps decreasingly the interval $(1 - \frac{1}{a_*}), 1 - \frac{1}{cN})$. For a given β in this interval, there exists a maximal $a = a_+$ given by (2.25).

Remark. The restriction that $c \leq 1$ is intrinsic to the power law model $\gamma(u) = cu^a$,

$a \geq 0$, since $\gamma(\frac{U}{L})$ must be a probability, before scaling. In principle, it appears meaningful to have monotonicity in c (increasing) for γ, and other models may be considered, including $\gamma(u) = cu^a \wedge 1$ or a consistent mollification.

Proof. The function $a \to q(a)$ will be decreasing if and only if the function $a \to p(a) = \frac{1}{a}\left(1 - \frac{1}{a}\right)^{(a-1)}$ is decreasing since q is a translation of p. An elementary calculation shows that the derivative of p is negative when $a > 1$, which implies $1 - \frac{1}{a} \in (0, 1)$. In detail,

$$\ln(p(a)) = -\ln(a) + (a-1)(\ln(a-1) - \ln(a)) \,, \quad \frac{d}{da}\ln(p(a)) = \ln\left(1 - \frac{1}{a}\right) < 0 \,.$$

The monotonicity in c of the solutions u_1, u_2 of $f(u) = 0$ (equilibria for $H(u)$) stems from the monotonicity of f (2.19) on each side of the value u_m (maximizer). □

2.5 Case $a \in (0, 1]$

2.5.1 The case $a = 1$

We begin with the analysis for the case $a = 1$. The following proposition is the continuous version discussed in [13]. We remind the reader that in Proposition 2.10 was introduced the quantity $q(1) = 1 - \frac{1}{Nc}$.

Proposition 2.11. *Under the conditions of Theorem 2.5, if $a = 1$, equivalently $\gamma = cu$, then $u(t)$ solves the standard logistic equation*

$$\frac{du}{dt} = cu(u_2 - u)\,, \qquad u(0) = \bar{u}\,. \tag{2.26}$$

with carrying capacity $u_2 = q(1)$. For $q(1) \leq 0$ the solution converges to zero, and for

$q(1) > 0$, the solution converges to the unique nonzero stable stationary state u_2. There is no unstable equilibrium u_1.

Proof. The classical solution of the logistic equation (see Section 1.2 in [9])

$$u(t) = u_2\left(1 + (\frac{u_2}{\bar{u}} - 1)e^{-cu_2 t}\right)^{-1}, \qquad \bar{u} \neq 0 \tag{2.27}$$

and $u(t) \equiv 0$ when $\bar{u} = 0$, proves the results, considering that the initial value $\bar{u} \in [0, 1]$ by construction. □

2.5.2 The case $a \in (0, 1)$

Proposition 2.12. *If $\gamma(u) = cu^a$, $0 < a < 1$, then there are exactly two equilibrium points $u_0 = 0$, and u_2 in $(0, 1)$ of the dynamical system 2.8. The point u_0 is unstable, and u_2 is stable.*

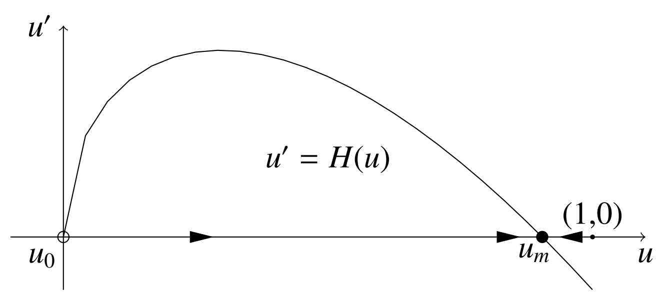

Figure 2.4: *Phase Portrait obtained by selecting $a = \frac{1}{2}$, $N = 10$, $c = 1$. The equilibrium point u_0 is unstable, and u_2 is stable.*

Proof. Part 1. Number of equilibrium points. As before $u_0 = 0$ is an equilibrium point.

Let $u > 0$, we have that $\lim_{u\to 0^+} f(u) = +\infty$ and $f(1) = -\frac{1}{Nc}$. The continuity of f on $(0, 1]$ guarantees that $f = 0$ has a solution on the same interval. Thus, we will have at least one more equilibrium point on $(0, 1)$.

To determine the precise number of solutions we recall that $f'(u) = au^{a-2}\left(\frac{a-1}{a} - u\right)$. Since $a < 1$, then $f'(u) < 0$, thus f' is strictly decreasing and therefore f injective. Hence, we obtain only one equilibrium point u_2 in $(0, 1)$.

Part 2. Stability. For the point $u_0 = 0$ we notice that, since $\lim_{u\to 0^+} f(u) = +\infty$, then for points greater but close enough to zero the expression $H(u) = cuf(u)$ is positive. We conclude then that $u_0 = 0$ is unstable in this case.

For the analysis of the point u_2 we begin by noticing that since $f' < 0$, then in particular $f'(u_2) < 0$. Thus, $H'(u_2) = cu_2 f'(u_2) < 0$, so u_2 is a stable point.

□

Chapter 3

Fluctuation Near Equilibrium

In this chapter we will analyze behavior near equilibrium points. We will show that if we use a different scaling near equilibria, the process u_t behaves as a diffusion. The study of the stability of the system is then approached by the investigation of the resulting stochastic process. We will follow closely Section (2.15) in which the process $u_t^L = \frac{1}{L}U = \frac{1}{L}\sum_j \mathbf{1}_{(Z_t^j \neq 0)}$ was introduced and studied. We recall that the function H, defined in (2.7), is

$$H(u) := \left(-\frac{1}{N}\right)u + \gamma(u)(1-u)\,, \qquad 0 \leq u \leq 1\,.$$

In this chapter, the function γ is not required to follow the power law $\gamma(u) = cu^a$. However, we require γ to be continuous. If, in addition, we want to keep the biological interpretations then, we will impose the conditions $\gamma(0) = 0$ and γ increasing in u, as explained in Section 2.3. This freedom in the selection of the function γ adds generality to our conclusions, however, it is interesting to keep the power law model in mind as a reference, since we have great information about the stability of the system (2.8) in such case.

3.1 Introduction to Scaling near Equilibrium and Initial Values

To describe fluctuations around an equilibrium point u' of the function H, we introduce a new process y_t^L of the form

$$y_t^L = \sqrt{L}(u_t^L - u'), \qquad H(u') = 0. \tag{3.1}$$

Throughout we assume the initial condition at time zero

$$\sqrt{L}(u_0^L - u') \xrightarrow{p} y_0 \in \mathbb{R}, \tag{3.2}$$

i.e. y_0^L converges in probability to a real value y_0. This amounts to studying the fluctuation process (3.1) when we start from values within $o(L^{-1})$.

We shall prove that, with the initial condition (3.2), the process y_t^L converges in distribution to y_t, where (y_t) is an Ornsetin-Uhlenbeck process with rate $r = H(u')$. We note the convergence is not only for marginals at given time t, but as a process, which is seen as a random variable on the space of RCLL paths.

3.2 Statement of Theorem 3.1 (Perturbation near Equilibrium)

We begin with the statement of the theorem and a brief description of the proof, which will be broken down in the next subsections. Precisely, we have

Theorem 3.1 (Perturbation near Equilibrium)**.** *Let H be defined as in (2.7), with u' an equilibrium point of H. With the additional hypothesis that H' satisfies a Lipschitz condition, and provided the initial condition (3.2) then, the process (y_t^L) converges in distribution, as*

$L \to \infty$, to the linear diffusion with generator

$$\mathcal{U}g(x) = -rxg'(x) + \frac{1}{2}\sigma^2 g''(x)\,, \qquad r = -H'(u')\,, \quad \sigma^2 = \frac{2u'}{N}\,. \tag{3.3}$$

Thus, the limit is a one-dimensional Ornstein - Uhlenbeck process.

3.3 Proof of Theorem 3.1

The proof of the theorem follows the following strategy. We will show that the scaled process y_t^L is C-tight. Recall that, by Prokhorov's theorem, tightness is pre-compactness in the space of probability measures. Denote by $\mathcal{L}^L$ the probability measures of the processes $y_\cdot^L$, indexed by $L \in \mathbb{Z}_+$, with values in the Skorokhod space $\mathcal{D}([0,\infty),\mathbb{R})$ of RCLL paths. We investigate a limit point $\mathcal{L}$ of such probability laws. More specifically, let $\mathcal{L}^{L_j} \Rightarrow \mathcal{L}$ as $j \to \infty$. We denote by y_t the process with probability law $\mathcal{L}$, and we show in Section 3.3.4 that $\mathcal{L}$ solves the martingale problem associated with the generator (3.3). This is, for any $g \in C_c^\infty(\mathbb{R})$, the process

$$M_t^g = g(y_t) - g(y_0) - \int_0^t \mathcal{U}g(y_s)\,ds$$

is a continuous $(\mathcal{F}_t)$ - martingale under the probability measure $\mathcal{L}$. Then, since the coefficients in $\mathcal{U}$ are bounded and continuous, the theory of martingale problem of Stroock and Varadhan guarantees the uniqueness of the process y_t. We identify y_t to be the Ornstein - Uhlenbeck process described in theorem 3.1.

Each step in the proof requires several considerations; that is why the proof of the theorem is divided into subsections. In Section 3.3.1 we prove that the process y_t^L is uniformly bounded in L^2. An explicit bound is provided in Proposition (3.2). In Section 3.3.3 we show that both, the generator $\mathcal{A}^L g(y_s^L)$ (see Proposition 3.3), and the quadratic variation of the martingale $M_t^{g,L}$ (see Proposition 3.4), are uniformly bounded in L^2.

Next, we prove tightness. Part (ii) of tightness (2.6) is harder than in Theorem 2.5 because the process (y_t^L) is no longer bounded naturally by the interval $[0,1]$. The fluctuations from the limit live on a different scale and can be arbitrarily large, both positive or negative. In fact, we show they are well behaved, but still unbounded. Due to the uniform bound obtained in Proposition 3.2, it is sufficient to prove the modulus of continuity (in essence, an Arzela-Ascoli equicontinuity condition in probability) on compact sets. More precisely, let $g \in C_c^\infty(\mathbb{R})$ be an infinitely differentiable function with compact support. The plan consists in evaluating bounds of the time integral, respectively of the martingale part of the differential formula (2.2) applied to $g(y_t^L)$. With this in mind, in section (3.3.3) we prove that the processes $\mathcal{A}^L g(y_s^L)$ and $M_t^{g,L}$ are C-tight (see proposition (3.5)). We obtain that the process y_t^L is C-tight (see (3.7)).

Finally, we prove in section 3.3.4 that, for every $g \in C_0^\infty(\mathbb{R})$, the process M_t^g is a continuous $(\mathcal{F}_t)$-martingale under the probability measure $\mathcal{L}$. We then invoke Theorem 4.28 (due to Stroock & Varadhan (1969)) and Corollary 4.29 in [10]. This concludes our proof.

3.3.1 Uniform L^2 Bound

In an effort to prove that the process y_t^L is C-tight, we meet with the problem of proving that y_t^L is square integrable. We will encounter such situation in (3.3), when trying to prove that the generator $\mathcal{A}^L g(y_t^L)$ is also square integrable. For that reason, we devote this section to the proof of Proposition (3.2) which, on one hand, gives a uniform bound for $\mathbb{E}\left[\left(y_t^L\right)^2\right]$, while on the other, introduces the proof's technique for the next section.

Remark: We make an important observation now that will be used through the entire chapter. Since the process u_t^L lives in the interval $[0,1]$, for each L the process y_t^L lives in the compact interval $[-\sqrt{L}u', \sqrt{L}(1-u')]$. We wish to apply the differential formula (2.15) obtained in Section 2.5. Thus, we must work with a restricted set of functions, namely continuous with compact support. At this point we take the bold step to denote by *id* any function which coincides with the identity function on the interval $[-\sqrt{L}u', \sqrt{L}(1-u')]$ and that is continuous with compact support. The selection of the function *id* is not important for the conclusion of any of the following results, and thus, we feel the reader will appreciate the simplification.

Due to the initial condition (3.2), the initial value has uniformly bounded second moment. We prove a slightly stronger result.

Proposition 3.2. *Assume that the initial values are random variables and there exists a positive C, independent of L, such that $E[(y_0^L)^2] \le C$. Let $T > 0$ be arbitrary but fixed. Then the process y_t^L is square integrable. Moreover, we have the bounds*

$$\mathbb{E}\left[\left(y_t^L\right)^2\right] \le K_1 \exp\left(K_2\, t\right) \qquad 0 \le t \le T. \tag{3.4}$$

Where K_1, K_2 are constants independent of t, and are given explicitly by $K_1 = 3C + 3\left(1+\frac{2}{N}\right)T$ and $K_2 = 3T(c_1(H))^2$, where $c_1(H)$ is bound of the function H'.

Proof. Apply the differential formula (2.15) to the function *id*, and expand H around the point $u_s^L = u' + \frac{y_s^L}{\sqrt{L}}$ using Taylor formula with remainder in integral form (see (B.1)). We get the following bound for the generator, in s, $0 \leq s \leq t \leq T$

$$\begin{aligned}
\mathcal{A}^L y_s^L &= \sum_{j=1}^{L} \left[\frac{1}{N} \left((y_s^L - \frac{1}{\sqrt{L}}) - y_s^L \right) \mathbf{1}(Z_s^j \neq 0) + \gamma(u_s^L) \left((y_s^L + \frac{1}{\sqrt{L}}) - y_s^L \right) \mathbf{1}(Z_s^j = 0) \right] \\
&= \frac{1}{N} \left(-\frac{1}{\sqrt{L}} \right) L u_s^L + \gamma(u_s^L) \left(\frac{1}{\sqrt{L}} \right) L(1 - u_s^L) \\
&= H(u_s^L) \sqrt{L} \\
&= y_s^L \int_0^1 H'(u' + (1-w) \frac{y_s^L}{\sqrt{L}}) dw.
\end{aligned} \tag{3.5}$$

We obtain that

$$|\mathcal{A}^L y_s^L| \leq c_1(H) \, |y_s^L| \tag{3.6}$$

where $c_1(H)$ is the supremum of H' on $[0, 1]$ which is independent of s and L.

A similar calculation for the quadratic variation of the martingale $M_t^{id,L}$, given by formula (2.4), shows

$$\begin{aligned}
\langle M^{id,L} \rangle_t &= \int_0^t \left(\sum_{j=1}^{L} \left[\frac{1}{N} \left((y_s^L - \frac{1}{\sqrt{L}}) - y_s^L \right)^2 \mathbf{1}(Z_s^j \neq 0) \; + \; \gamma(u_s^L) \left((y_s^L + \frac{1}{\sqrt{L}}) - y_s^L \right)^2 \mathbf{1}(Z_s^j = 0) \right] \right) ds \\
&= \int_0^t \left(\frac{1}{N} \left(-\frac{1}{\sqrt{L}} \right)^2 L u_s^L + \gamma(u_s^L) \left(\frac{1}{\sqrt{L}} \right)^2 L(1 - u_s^L) \right) ds \\
&= \int_0^t \left(H(u_s^L) + \frac{2}{N} u_s^L \right) ds.
\end{aligned}$$

Since $H(u) \leq 1$ in $[0, 1]$, then

$$\mathbb{E}\left[\langle M^{id,L} \rangle_t \right] \leq \left(1 + \frac{2}{N} \right) T \qquad 0 \leq t \leq T. \tag{3.7}$$

Recall that by formula (2.2) we can write

$$y_t^L = y_0^L + \int_0^t \mathcal{A}^L y_s^L \, ds + M_t^{id,L} \tag{3.8}$$

where $\mathcal{A}_t^L y_s^L$ is given by the differential formula (3.5). Thus, we obtain that

$$(y_t^L)^2 \leq 3\left((y_0^L)^2 + \left(\int_0^t \left|\mathcal{A}^L y_s^L\right| \, ds\right)^2 + \left(M_t^{id,L}\right)^2\right).$$

Remember also, that

$$N_t^{id} = \left(M_t^{id}\right)^2 - \langle M^{id,L}\rangle_t \tag{3.9}$$

is a martingale. Thus, by Cauchy-Schwarz, Fubini's Theorem, the initial condition on $\mathbb{E}[(y_0^L)^2]$, and relations (3.6) and (3.7) we obtain that

$$\begin{aligned}
\mathbb{E}\left[\left(y_t^L\right)^2\right] &\leq 3\left(\mathbb{E}\left[\left(y_0^L\right)^2\right] + \mathbb{E}\left[T\int_0^t \left(\left|\mathcal{A}^L y_s^L\right|\right)^2 \, ds\right] + \mathbb{E}\left[\left(M_t^{id,L}\right)^2\right]\right) \\
&\leq 3\mathbb{E}\left[\left(y_0^L\right)^2\right] + 3(c_1(H))^2\, T\int_0^t \mathbb{E}\left[\left(y_t^L\right)^2\right] ds + 3\mathbb{E}\left[\langle M^{id,L}\rangle_t\right] \quad (3.6) \\
&\leq 3C + 3(c_1(H))^2\, T\int_0^t \mathbb{E}\left[\left(y_t^L\right)^2\right] ds + 3\left(1 + \frac{2}{N}\right)T. \qquad (2.3), (3.7)
\end{aligned}$$

Hence

$$\mathbb{E}\left[\left(y_t^L\right)^2\right] \leq K_1 + K_2\int_0^t \mathbb{E}\left[\left(y_s^L\right)^2\right] ds$$

where $K_1 = 3C + 3\left(1 + \frac{2}{N}\right)T$, $K_2 = 3T(c_1(H))^2$.

By Gronwall's inequality

$$\mathbb{E}\left[\left(y_t^L\right)^2\right] \leq K_1 \exp\left(K_2\, t\right) \qquad 0 \leq t \leq T. \tag{3.10}$$

□

3.3.2 Modulus of Continuity

Right from the onset we clarify that all functions considered in the remainder of the proof belong to the family of test functions of the form $g(y_s^L) = g\left(\sqrt{L}(u_t^L - u')\right)$ where $g \in C_c^3(\mathbb{R})$. We will apply formula (2.15) to this family in order to obtain an uniform bound for $\mathcal{A}^L g(y_s^L)$. We recall from (2.2) and (2.4) that the process

$$M_t^{L,g} = g(y_s^L) - g(y_0^L) - \int_0^t \mathcal{A}g(y_s^L)ds$$

is a martingale and that its quadratic variation is given by the expression

$$\langle M^{g,L}\rangle_t = \int_0^t \sum_{y\in S} p(y_s^L, y)\left(g(y) - g(y_s^L)\right)^2 ds.$$

We will use this formula to obtain an uniform bound for the quadratic variation. These results are stated in the next two propositions.

Proposition 3.3. *For every function $g \in C_c^3(\mathbb{R})$ and $H \in C^2(\mathbb{R})$ the generator of the process y_t^L is in L^2. Moreover, we have the estimate*

$$\mathbb{E}\left(\left|\mathcal{A}^L g(y_s^L)\right|\right)^2 \leq C_1 \mathbb{E}\left(|y_s^L|\right)^2 + C_2$$

where C_1 and C_2 are constants depending on the functions g, H and their derivative but not depending on L, t.

Proposition 3.4. *For every function $g \in C_c^3(\mathbb{R})$ and every $t \in [0, T]$, the quadratic variation of $M_t^{g,L}$ is bounded by*

$$\left|\langle M^{g,L}\rangle_t\right| \leq \left(M_1 + \frac{1}{\sqrt{L}}M_2\right)t \tag{3.11}$$

where M_1, M_2 are constants depending on g, H and their derivatives but not depending on L, t.

Proof. **(Proof of proposition (3.3))** We choose a test function $g \in C_c^3(\mathbb{R})$, and apply the differential formula (2.15). We get the formula

$$\begin{aligned}\mathcal{A}^L g(y_s^L) &= \sum_{j=1}^{L}\left[\frac{1}{N}\left(g(y_s^L - \frac{1}{\sqrt{L}}) - g(y_s^L)\right)\mathbf{1}(Z_s^j \neq 0) + \gamma(u_s^L)\left(g(y_s^L + \frac{1}{\sqrt{L}}) - g(y_s^L)\right)\mathbf{1}(Z_s^j = 0)\right] \\ &= \frac{1}{N}\left(g(y_s^L - \frac{1}{\sqrt{L}}) - g(y_s^L)\right)Lu_s^L + \gamma(u_s^L)\left(g(y_s^L + \frac{1}{\sqrt{L}}) - g(y_s^L)\right)L(1-u_s^L).\end{aligned} \tag{3.12}$$

We develop each difference in the integrand using Taylor expansion of order two around the point y_s^L with remainder in the integral form

$$\begin{aligned}g(y_s^L \pm \frac{1}{\sqrt{L}}) - g(y_s^L) &= g'(y_s^L)\left(\pm\frac{1}{\sqrt{L}}\right) + \frac{1}{2}g''(y_s^L)\left(\pm\frac{1}{\sqrt{L}}\right)^2 \\ &\quad + \frac{1}{2}\left(\pm\frac{1}{\sqrt{L}}\right)^3 \int_0^1 w^2 g'''\left(y_s^L + (\pm\frac{1}{\sqrt{L}})(1-w)\right)^2 dw.\end{aligned} \tag{3.13}$$

Replacing these expressions back in (3.12) we obtain that the generator takes the form

$$\begin{aligned}\mathcal{A}^L g(y_s^L) = &\ \frac{1}{N}\left(g'(y_s^L)\left(-\frac{1}{\sqrt{L}}\right) + \frac{1}{2}g''(y_s^L)\left(-\frac{1}{\sqrt{L}}\right)^2\right)Lu_s^L \\ &+ \gamma(u_s^L)\left(g'(y_s^L)\left(\frac{1}{\sqrt{L}}\right) + \frac{1}{2}g''(y_s^L)\left(\frac{1}{\sqrt{L}}\right)^2\right)L(1-u_s^L) + R\left(\omega, g, L, s\right).\end{aligned}$$

Which simplifies to

$$\mathcal{A}^L g(y_s^L) = \sqrt{L}\, g'(y_s^L)H(u' + \frac{y_s^L}{\sqrt{L}}) + \frac{1}{2}g''(y_s^L)\left(H(u_s^L) + \frac{2}{N}u_s^L\right) + R\left(\omega, g, L, s\right). \tag{3.14}$$

Where $R(\omega, g, L, s)$ is a function obtained by collecting all expressions involving the remainder expressions in the Taylor formulas (3.13).

Since the function g has compact support, it is possible to estimate the remainders formulas in (3.13). In any case we have that, there is a constant $c_3(g)$, depending on the function g''', such that the remainders are bounded by

$$\left|\frac{1}{2}\left(\pm\frac{1}{\sqrt{L}}\right)^3 \int_0^1 w^2 g'''\left(y_s^L + (\pm\frac{1}{\sqrt{L}})(1-w)\right)^2 dw.\right| \leq \frac{1}{6}\left(\frac{1}{\sqrt{L}}\right)^3 c_3(g). \tag{3.15}$$

Now, bound (3.15) together with $u_t^L \in [0,1]$ imply that the function $R(\omega, g, L, s)$ can be estimated by

$$|R(\omega, g, L, s)| \leq \frac{1}{6}\left(\frac{1}{\sqrt{L}}\right)\left(H(u_s^L) + \frac{2}{N}u_s^L\right)c_3(g) \leq \frac{1}{6}\frac{1}{\sqrt{L}}\left(1 + \frac{2}{N}\right)c_3(g). \tag{3.16}$$

At this point, we take a close look at expression (3.14) and observe that all summands involve are bounded except the first one. The last summand is bounded by (3.16) and the second summand is bounded since g has compact support, while $H(u_s^L)$ and u_s^L belong to $[0, 1]$. However, in the first summand the factor $\sqrt{L}$ gives the impression the this term is uncontrolled. To sort this misperception, we develop further the function H by Taylor formula of order one, and we use the fact that $H(u') = 0$ at the equilibrium point u', to obtain

$$H(u' + \frac{y_s^L}{\sqrt{L}}) = \underbrace{H(u')}_{=0} + \frac{y_s^L}{\sqrt{L}}\int_0^1 H'(u' + w\,\frac{y_s^L}{\sqrt{L}})dw. \tag{3.17}$$

Plugging back into (3.14) we obtain that the generator is

$$\mathcal{A}^L g(y_s^L) = g'(y_s^L)y_s^L \int_0^1 H'(u' + w\,\frac{y_s^L}{\sqrt{L}})dw + \frac{1}{2}g''(y_s^L)\left(H(u_s^L) + \frac{2}{N}u_s^L\right) + R(\omega, g, L, s)\,. \tag{3.18}$$

Since H' is continuous, the integral remainder in (3.17) can be bounded by a constant $c_1(H)$ depending on the function H'. Thus, expression (3.18) presents a clean formula for the generator in which each term in now, by proposition (3.2) and the comments above, clearly bounded, at least in L^2 sense. Denote by $c_1(g)$, $c_2(g)$ uniform bounds for the function g' and g'' respectively. More precisely, we obtain the L^2 bound

$$\begin{aligned}\mathbb{E}\left[\left(\mathcal{A}^L g(y_s^L)\right)^2\right] &\leq \mathbb{E}\left[3\left(\left(g'(y_s^L)c_1(H)y_s^L\right)^2 + \frac{1}{4}\left(g''(y_s^L)\left(H(u_s^L) + \frac{2}{N}u_s^L\right)\right)^2 + \left(R\left(\omega, g''', L, s\right)\right)^2\right)\right] \\ &\leq \mathbb{E}\left[3c_1^2(g)c_1^2(H)\left(y_s^L\right)^2 + 3\left(1 + \frac{2}{N}\right)^2\left(\frac{1}{4}c_2^2(g) + c_3^2(g)\frac{1}{L}\right)\right].\end{aligned}$$

Since L and N can be taken bigger than 1 we obtain that,

$$\mathbb{E}\left[\left(\mathcal{A}^L g(y_s^L)\right)^2\right] \leq C_1 \mathbb{E}\left[\left(y_s^L\right)^2\right] + C_2$$

where $C_1 = 3c_1^2(g)c_1^2(H)$ and $C_2 = 27\left(\frac{1}{4}c_2^2(g) + c_3^2(g)\right)$. □

Proof. **(Proof of proposition (3.4))** We apply formula (2.4) and obtain that the quadratic variation can be written as

$$\langle M^{g,L}\rangle_t = \int_0^t \left(\sum_{j=1}^{L} \frac{1}{N}\left(g(y_{s-}^L - \frac{1}{\sqrt{L}}) - g(y_{s-}^L)\right)^2 \mathbf{1}(Z_{s-}^j \neq 0) + \gamma(u_{s-}^L)\left(g(y_{s-}^L + \frac{1}{\sqrt{L}}) - g(y_{s-}^L)\right)^2 \mathbf{1}(Z_{s-}^j = 0)\right) ds$$

Which simplifies to the expression

$$\langle M^{g,L}\rangle_t = \int_0^t \left(\frac{1}{N}\left(g(y_{s-}^L - \frac{1}{\sqrt{L}}) - g(y_{s-}^L)\right)^2 L\, u_t^L + \gamma(u_{s-}^L)\left(g(y_{s-}^L + \frac{1}{\sqrt{L}}) - g(y_{s-}^L)\right)^2 L(1 - u_t^L)\right) ds. \tag{3.19}$$

We develop the differences in the integrand by Taylor's formula

$$g(y_s^L \pm \frac{1}{\sqrt{L}}) - g(y_s^L) = g'(y_s^L)\left(\pm\frac{1}{\sqrt{L}}\right) + r_\pm(\omega, g, L, s)$$

where $r_\pm(\omega, g, L, s)$ are the remainders of the Taylor expansion of order one in integral form

$$r_\pm(\omega, g, L, s) = \left(\pm\frac{1}{\sqrt{L}}\right)^2 \int_0^1 w\, g''\left(y_s^L + (\pm\frac{1}{\sqrt{L}})(1-w)\right) dw$$

These expressions are bounded uniformly in time t and ω by

$$|r_\pm(\omega, g, L, s)| \;\leq\; \frac{1}{2}\left(\frac{1}{\sqrt{L}}\right)^2 c_2(g) \tag{3.20}$$

Replacing the Taylor expansion in (3.19) we obtain that, the quadratic variation of the martingale is

$$\begin{aligned}\langle M^{g,L}\rangle_t &= \int_0^t \left(\frac{1}{N}\left(-\frac{1}{\sqrt{L}}g'(y_s^L) + r_-\right)^2 u_t^L\, L + \gamma(u_{s-}^L)\left(\frac{1}{\sqrt{L}}g'(y_s^L) + r_+\right)^2 (1-u_t^L)\, L\right) ds \\ &= \int_0^t \left(\left(g'(y_s^L)\right)^2 \left(H(u_s^L) + \frac{2}{N}u_s^L\right) + \bar{R}\,(g''', L, s)\right) ds\end{aligned} \tag{3.21}$$

where $\bar{R}\,(g''', L, s)$ collects all terms involving the the Taylor remainders. We do not need to worry to much with the exact form of $\bar{R}\,(g''', L, s)$, it is only important to know that $\bar{R}\,(g''', L, s) \in O(\frac{1}{\sqrt{L}})$, i.e: there is a constant M_2 such that $\bar{R}\,(g''', L, s) \leq \frac{1}{\sqrt{L}}M_2$. Thus, we obtain that

$$\left|\langle M^{g,L}\rangle_t\right| \leq \left(c_1^2(g)\left(1 + \frac{2}{N}\right) + \left|\bar{R}\,(g''', L, s)\right|\right) t \leq \left(c_1^2(g)\left(1 + \frac{2}{N}\right) + \frac{1}{\sqrt{L}}M_2\right) t.$$

Hence, by denoting $M_1 = c_1^2(g)\left(1+\frac{2}{N}\right)$ we obtain

$$\left|\langle M^{g,L}\rangle_t\right| \leq \left(M_1 + \frac{1}{\sqrt{L}}M_2\right)t.$$

□

3.3.3 Tightness of the Process y_t^L

We want now to prove that the process $(y_t^L)_{t\geq 0}$ is C-tight. We already used this procedure in the proof of Theorem 2.5. We show first that the processes $\int_0^t \mathcal{A}^L g(y_s^L)$ and $M_t^{g,L}$ are C-tight. Since the process

$$g(y_t^L) = g(y_0^L) + \int_0^t \mathcal{A}^L g(y_s^L) + M_t^{g,L}$$

we immediately obtain that, for any test function g, the process $g(y_t^L)$ is C-tight. A simple localization argument then gives that y_t^L is C-tight. So we will prove the following three propositions:

Proposition 3.5. *The processes $d_t^{g,L} := \int_0^t \mathcal{A}^L g(y_s^L)\,ds$ is C-tight.*

Proposition 3.6. *The process $M_t^{g,L}$ is C-tight.*

Proposition 3.7. *The process y_t^L is C-tight.*

The two key steps in order to prove (3.5) and (3.6) are to verify (i) and (ii) from Definition 2.4.

Proof. **(Proof of Proposition 3.5)** Condition (i) in definition (2.4) is easily satisfied by $d_t^{g,L}$ since g has compact support. In order to prove condition (ii) for $d_t^{g,L}$ we let $\delta > 0$. On the interval $[0,T]$, we choose $0 \le t < t' \le T$ such that $t' - t < \delta$. For the process $d_t^{g,L}$ we have

$$\begin{aligned}
\mathbb{P}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}} \left|d_{t'}^{g,L} - d_t^{g,L}\right| > \epsilon\right) &\le \frac{1}{\epsilon^2}\mathbb{E}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}} \left(\int_t^{t'} \mathcal{A}^L g(y_s^L)\, ds\right)^2\right) && \text{by Chebyshev's inequality}\\
&\le \frac{1}{\epsilon^2}\mathbb{E}\left(\delta \sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}} \int_t^{t'} \left(\mathcal{A}^L g(y_s^L)\right)^2\, ds\right) && \text{by Cauchy-Schwarz inequality}\\
&\le \frac{\delta}{\epsilon^2}\mathbb{E}\left(\int_0^T \left(\mathcal{A}^L g(y_s^L)\right)^2\, ds\right) && \text{by positivity of the integrand}\\
&= \frac{\delta}{\epsilon^2}\int_0^T \mathbb{E}\left[\left(\mathcal{A}^L g(y_s^L)\right)^2\right] ds && \text{by Fubini's Theorem}\\
&\le \frac{3\delta}{\epsilon^2}\int_0^T \left(C_1\mathbb{E}\left[\left(y_s^L\right)^2\right] + C_2\right) ds && \text{by (3.3)}\\
&\le \frac{3\delta}{\epsilon^2}\int_0^T \left[C_1 K_1 \exp\left(K_2\, s\right) + C_2\right] ds && \text{by (3.2)}\\
&= \frac{3\delta}{\epsilon^2}\left[C_1 K_1 \frac{\exp\left(K_2\, T\right) - 1}{K_2} + C_2 T\right].
\end{aligned}$$

hence,

$$\lim_{\delta\to 0} \limsup_{L\to\infty} \mathbb{P}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}} \left|d_{t'}^{g,L} - d_t^{g,L}\right| > \epsilon\right) = 0.$$

□

Proof. **(Proof of Proposition 3.6)** As before, condition (i) in definition (2.4) is easily satisfied by $M_t^{g,L}$ since g has compact support. Again, we let $\delta > 0$. On the interval $[0,T]$, we choose $0 \le t < t' \le T$ such that $t' - t < \delta$. For the process $M_t^{g,L}$ we have

$$\begin{aligned}
\mathbb{P}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}} \left|M_{t'}^{g,L} - M_t^{g,L}\right| > \epsilon\right) &\le \frac{1}{\epsilon^2}\mathbb{E}\left(\left|M_{t'}^{g,L} - M_t^{g,L}\right|^2\right) && \text{by Doob's } L^2\text{-norm maximal inequality}\\
&= \frac{1}{\epsilon^2}\mathbb{E}\left(\left|\langle M^{g,L}\rangle_{t'} - \langle M^{g,L}\rangle_t\right|\right) && \text{by (2.3)}\\
&\le \frac{\delta}{\epsilon^2}\left(M_1 + \frac{1}{\sqrt{L}}M_2\right) && \text{by (3.4)}
\end{aligned}$$

hence

$$\lim_{\delta\to 0}\limsup_{L\to\infty}\mathbb{P}\left(\sup_{\substack{0\le t<t'\le T\\ t'-t<\delta}}\left|M^{g,L}_{t'}-M^{g,L}_{t}\right|>\epsilon\right)=0.$$

□

Proof. **(Proof of Proposition 3.7)** From (2.2)

$$g(y^L_t)-g(y^L_0)=M^{g,L}_t+d^{g,L}_t \tag{3.22}$$

both processes $M^{g,L}_t$ and $d^{g,L}_t$ are *C-tight*. Take $g(u)=id(u)$. □

3.3.4 Solution of the Martingale Problem

In this section we want to prove that any limit point (y_t) of the sequence $(y^L_t)_{L\in\mathbb{Z}_+}$, seeing as a set in the Skorokhod space with the J_1 topology, satisfies the stochastic differential equation

$$dy_t=-H'(u')y_t dt+\frac{2u'}{N}dB_t \tag{3.23}$$

In order to do this, we will show that any limit point solves the martingale problem associated with the generator of the solution process of equation (3.23).

With this in mind, let $(y_t)_{t\ge 0}$ a limit point, seen as a random variable on $C([0,T],\mathbb{R})$, of the processes $(y^L_t)_{t\ge 0}$, indexed by L. We begin evaluating an useful bound.

Proposition 3.8. *There exist D_1, D_2 and D_3 independent of s and L (but dependent on T) such that*

$$|\mathcal{A}^L g(y^L_s)-\mathcal{U}g(y^L_s)|\le\left((D_1+D_2|y^L_s|)^2+D_3\right)L^{-\frac{1}{2}}\,. \tag{3.24}$$

Remark. The estimate above, combined with (3.2) shows that

$$\mathbb{E}\left[|\mathcal{A}^L g(y_s^L) - \mathcal{U}(y_s^L)|\right] \leq DL^{-\frac{1}{2}}.$$

Where D is a constant that does not depend on s or L, but depends on T.

Proof. From (3.18) and (3.3) we have that

$$\begin{aligned}\mathcal{A}^L g(y_s^L) - \mathcal{U}g(y_s^L) = {} & \underbrace{g'(y_s^L)y_s^L \int_0^1 \left(H'(u' + w\,\frac{y_s^L}{\sqrt{L}}) - H'(u')\right)dw}_{(I)} \; + \\ & + \underbrace{\frac{1}{2}g''(y_s^L)\left(H(u_s^L) + \frac{2}{N}u_s^L - \frac{2}{N}u'\right)}_{(II)} \; + \\ & + \underbrace{R\left(\omega, g''', L, s\right).}_{(III)}\end{aligned}$$

Since H' satisfies a Lipschitz condition, we obtain in (I) that

$$\left|H'(u' + w\,\frac{y_s^L}{\sqrt{L}}) - H'(u')\right| \leq \left|w\frac{y_s^L}{\sqrt{L}}\right|$$

Hence, if we let $c_1(g)$ be a positive bound for $\frac{1}{2}g'$, then expression (I) is bounded by

$$(I) \leq \frac{1}{2\sqrt{L}}\,|g'(y_s^L)|\,(|y_s^L|)^2 \leq \frac{1}{\sqrt{L}}\,c_1(g)\,(|y_s^L|)^2 \tag{3.25}$$

In expression (II), we write $u_t^L = u' + \frac{y_s^L}{\sqrt{L}}$ and develop the function H around the point u' using Taylor formula with remainder, in integral form, of order one. Thus, obtain that

$$H(u_s^L) = \frac{y_s^L}{\sqrt{L}}\int_0^1 H'\left(u' + (1-w)\frac{y_s^L}{\sqrt{L}}\right)dw$$

$$(II) = \frac{1}{2}g''(y_s^L)\left(\frac{y_s^L}{\sqrt{L}}\int_0^1 H'\left(u' + (1-w)\frac{y_s^L}{\sqrt{L}}\right)dw + \frac{2}{N}(u' + \frac{y_s^L}{\sqrt{L}}) - \frac{2}{N}u'\right)$$
$$= \frac{1}{2}g''(y_s^L)\frac{y_s^L}{\sqrt{L}}\left(\int_0^1 H'\left(u' + (1-w)\frac{y_s^L}{\sqrt{L}}\right)dw + \frac{2}{N}\right)$$

Since H' satisfies a Lipschitz condition then it is bounded, let $c_1(H)$ be a positive bound for H', and $c_2(g)$ a positive bound for $\frac{1}{2}g''$. Then

$$|(II)| \leq c_2(g)\frac{|y_s^L|}{\sqrt{L}}\left(c_1(H) + \frac{2}{N}\right) \tag{3.26}$$

Finally, (3.16) gives

$$|(III)| = \left|R\left(\omega, g^{(3)}, L, s\right)\right| \leq \frac{1}{\sqrt{L}}\left(1 + \frac{2}{N}\right)c_3(g). \tag{3.27}$$

Where $c_3(g)$ is a positive bound for g'''. Now bounds (3.25), (3.26) and (3.27) give the desire bound for the difference

$$|\mathcal{A}^L g(y_s^L) - \mathcal{U}g(y_s^L)| \leq \frac{1}{\sqrt{L}}\left(c_1(g)(|y_s^L|)^2 + c_2(g)|y_s^L|\left(c_1(H) + \frac{2}{N}\right) + \left(1 + \frac{2}{N}\right)c_3(g)\right) \tag{3.28}$$

Thus, by completing the square, we can write (3.28) as

$$|\mathcal{A}^L g(y_s^L) - \mathcal{U}g(y_s^L)| \leq \left((D_1 + D_2|y_s^L|)^2 + D_3\right)L^{-\frac{1}{2}}.$$

where

$$D_1 = \sqrt{c_1(g)},\ D_2 = \frac{c_2(g)\left(c_1(H) + \frac{2}{N}\right)}{2\sqrt{c_1(g)}},\ D_3 = \left(1 + \frac{2}{N}\right)c_3(g) - \frac{c_2^2(g)\left(c_1(H) + \frac{2}{N}\right)^2}{c_1(g)}.$$

□

Proposition 3.9. *For any* $g \in C_c^3(\mathbb{R})$

$$g(y_t) - g(y_0) - \int_0^t \mathcal{U} g(y_s)\, ds \qquad \text{is a } (\mathcal{F}_t) \text{ - martingale} \tag{3.29}$$

where $\mathcal{U}$ *is as in (3.3).*

Proof. Define $M_{s,t}^{g,\eta}$, $g \in C_c^3(\mathbb{R})$ the difference of expressions from (3.3) taken at two times $0 \le s \le t \le T$.

$$M_{s,t}^{g,\eta} = g(\eta_t) - g(\eta_s) - \int_s^t \mathcal{U} g(\eta_{s'})\, ds' \qquad \eta \in \mathcal{D}([0,T],\mathbb{R})\,. \tag{3.30}$$

Let $\psi(\eta_\cdot)$ a $\mathcal{F}_s$ - measurable function, equal to a finite product of continuous bounded functions applied at a finite number of times s', $0 \le s' \le s$. Define $\Psi : \mathcal{D}([0,T],\mathbb{R}) \to \mathbb{R}$

$$\Psi(\eta_\cdot) = M_{s,t}^{g,\eta} \psi(\eta_\cdot)\,. \tag{3.31}$$

Ψ is a bounded, continuous functional, (see [24]). It only remains to prove that $E[\Psi(y_\cdot^L)] = 0$. So we have

$$\begin{aligned}
E[\Psi(y_\cdot^L)] &= \mathbb{E}\left[\left(g(y_t^L) - g(y_s^L) - \int_s^t \mathcal{U} g(y_{s'}^L)\, ds'\right)\psi(y_\cdot^L)\right] \\
&= \mathbb{E}\left[\left(g(y_t^L) - g(y_s^L) - \int_s^t \mathcal{A}^L g(y_{s'}^L)\, ds' + \int_s^t (\mathcal{A}^L - \mathcal{U}) g(y_{s'}^L)\, ds'\right)\psi(y_\cdot^L)\right] \\
&= \mathbb{E}\left[(M_t^{g,L} - M_s^{g,L})\psi(y_\cdot^L)\right] + \mathbb{E}\left[\left(\int_s^t (\mathcal{A}^L - \mathcal{U}) g(y_{s'}^L)\, ds'\right)\psi(y_\cdot^L)\right] \\
&= \mathbb{E}\left[\left(\int_s^t (\mathcal{A}^L - \mathcal{U}) g(y_{s'}^L)\, ds'\right)\psi(y_\cdot^L)\right]
\end{aligned}$$

where the last equality follows from the fact that $M_t^{g,L}$ is a martingale. From (3.24) we have that the error term

$$\left|E[\Psi(y_\cdot^L)]\right| \leq \left|\mathbb{E}\left[\left(\int_s^t (\mathcal{A}^L - \mathcal{U})g(y_{s'})\,ds'\right)\psi(y_\cdot^L)\right]\right| \leq TDL^{-\frac{1}{2}} \to 0 \quad \text{as } L \to \infty.$$

We remind the reader that $y^L(\cdot) \Rightarrow y(\cdot)$ (converges in distribution), as a consequence of tightness. Again Portmanteau Theorem implies that $\lim_{L\to\infty} E[\Psi(y_\cdot^L)] = E[\Psi(y_\cdot)]$. We have shown the right-hand side limit is zero. Thus, $\Psi(y_\cdot) = 0$ implying that

$$g(y_t) - g(y_0) - \int_0^t \mathcal{U}g(y_s)\,ds$$

is a continuous $(\mathcal{F}_t)$ - martingale, thus the measure induced by the the process y_t satisfies the martingale problem associated with the generator

$$\mathcal{U}g(x) = -rxg'(x) + \frac{1}{2}\sigma^2 g''(x)\,, \qquad r = -H'(u')\,, \quad \sigma^2 = \frac{2u'}{N}\,.$$

The coefficients of r and σ^2 are bounded and continuous. By Theorem 4.28 and Corollary 4.29 in Karatzas et al. [10] (due to Stroock & Varadhan (1969) [23]), the martingale problem is well posed and there is at most one solution to the time-homogeneous martingale problem. We just proved that any possible limit point $\mathcal{L}$ of the laws $\mathcal{L}^L$, $L > 0$ *solves the martingale problem for* $\mathcal{U}$. In other words, any limiting process (y_t) has probability law $\mathcal{L}$.

The solution to the martingale problem is unique. We conclude that there is only one limit point. Thus, the sequence of processes y_t^L converge in distribution, as $L \to \infty$, to the the Ornstein-Uhlenbeck process described in Theorem 3.1. This concludes the proof.

□

Chapter 4

Generalized Logistic Equation with Noise

4.1 Overview

We have shown in Section 1.4, Theorem 2.5 that the deterministic limit (u_t) is the solution of the generalized logistic equation

$$du_t = H(u_t)dt\,, \qquad u_0 = \bar{u}\,, \tag{4.1}$$

where $H(u)$ is defined in (2.7). In Section 2.3, it is suggested that the process (u_t) can be used as a population evolutionary model. On the scale (1.5) of Theorem 2.5, the noise, represented by a martingale term, vanishes. This is natural in a law of large numbers.

However, it might be valuable to consider a perturbation of the process u_t, with small noise σdW_t, where W_t is a Wiener process. This *mesoscale* allows considering the root of our question (escape from a stable state), under the continuum assumption of a SDE. The process (u_t) is not perfectly determined by (4.1), but subject to some random biological effect. Such influence could be captured by the noise term.

With this in mind, in this chapter we propose the study of the equation

$$du_t = H(u_t)dt + \sigma dW_t\,, \qquad u_0 = \bar{u}\,, \tag{4.2}$$

where W_t is a standard Brownian motion and σ is a constant. The magnitude of σ depends on how much impact we consider the noise has over the deterministic model. We will denote the solution of equation (4.2) by u_t, since this section is self contained and u_t cannot be confused with the solution of the ODE (2.8). In order to avoid technical difficulties we impose the condition that the function H satisfies a Lipschitz condition. In the case that $\gamma(u) = cu^a$, then the regularity of H is guaranteed by choosing

$$a \in \{0\} \cup [1, \infty)\,, \quad \text{in eq.} \quad (2.18)\,. \tag{4.3}$$

Hence, we have existence and uniqueness [10] of the solution of (4.2) and the solution u_t is a diffusion.

In Section 2.3, we considered the deterministic process u_t naturally living in $[0, 1]$ and having 0 as an absorbing state for the system. In that spirit, we let the process, defined as the solution of equation (4.2), be absorbing at 0 and reflecting at 1, an assumption we regard as biologically consistent.

For such problem we begin to question the probability of reaching 0 starting at a given value in the interval from 0 to 1 (probability of extinction), and the time to reach 0, or time of extinction τ defined in (4.4), also lifetime of the process, also known as time to ruin in classical probability theory. These questions are answered in Propositions 4.2 and 4.3, respectively.

As $\tau < \infty$ a.s. we have a dissipative process. It is natural to investigate the existence of a quasi-stationary distribution (qsd), denoted by ν in (4.20), for the solution process of (4.2). In section 4.5 it is shown, via the theory of Sturm-Liouville operators, that the quasi-invariant distribution exists and it is unique. In section 4.6 some explicit calculations for the quasi-stationary distribution are presented.

We finalize this chapter with an important result, stated as a conjecture (in that evidence is provided numerically, and only in special cases). We find it of great interest because of its random perturbation theory implications.

Theorem 4.1. *(Conjecture) Let $q_\epsilon(x)$ be the density function of the qsd corresponding to $\sigma^2 = \epsilon$. Then, as $\epsilon \to 0$, the measure $q_\epsilon(x)dx$ converges weakly to a discrete measure with atoms at the equilibrium points of H. Moreover, only the stable equilibria have nonzero mass.*

We point the reader to Section 4.6.2 and specifically to the results for $a = 1$ from Figure 4.6, when $H(u)$ has a square nonlinearity, essentially a perturbation of the logistic equation. In this we illustrate how the graph of the qsd peaks like a bump function about the value of the non-zero stable equilibrium.

4.2 Extinction with Probability One

We study now the time of extinction of the diffusion u_t, defined as the first time in which the process reaches the value 0. Formally,

$$\tau = \inf\left\{t \geq 0 \,|\, u_t = 0\right\} \tag{4.4}$$

For $\alpha > 0$, define $g : [0,1] \to \mathbb{R}$ as

$$g(u) := \mathbb{E}_u\left[e^{-\alpha\tau}\mathbb{1}_{(\tau<\infty)}\right] \qquad \text{note that we simplified notation} \quad g(u) = g(u,\alpha)\,. \tag{4.5}$$

Thus, for u fixed, the function $g(u)$ is the Laplace transform of the time of extinction τ. The detail to keep in mind is that we can obtain the moments of τ by successive derivations with respect to the parameter α. For instance, we are interested in

$$p^0(u) := \lim_{\alpha\to 0^+} g(u) = \mathbb{E}_u\left[\lim_{\alpha\to 0^+} e^{-\alpha\tau}\mathbb{1}_{(\tau<\infty)}\right] = \mathbb{P}_u[\mathbb{1}_{(\tau<\infty)}] \quad \text{(Probability of Extinction)} \tag{4.6}$$

$$e^0(u) := \lim_{\alpha\to 0^+} -\frac{\partial g}{\partial \alpha}(u) = \mathbb{E}_u\left[\lim_{\alpha\to 0^+} \tau e^{-\alpha\tau}\mathbb{1}_{(\tau<\infty)}\right] = \mathbb{E}_u[\tau] \quad \text{(Expected time of Extinction).} \tag{4.7}$$

The study of g can be examined in the light of the Feynman-Kac connection. This will be our method of choice. Since we will need it, we take a moment to notice that, by the Itô formula, the generator of the process u_t is given by

$$\mathcal{A}f(u) = H(u)f'(u) + \frac{\sigma^2}{2}f''(u)\,, \qquad f \in C^2[0,1] \qquad f(0) = 0\,, \quad f'(1) = 0\,. \tag{4.8}$$

We are ready for our first proposition of the chapter. Due to the smoothness of $H(u)$, it is classic but since it is essential for the rest of the chapter, we give a direct proof.

Proposition 4.2. *The probability of extinction (hitting* 0*), of the process* u_t *defined as the solution of (4.2), with boundary conditions as in (4.8), is equal to one, for any initial point* $0 \le u \le 1$.

Proof. By the Feynman-Kac Formula we have that the function g satisfies the initial value problem

$$\begin{aligned} \mathcal{A}g(u) - \alpha g(u) &= 0, \\ g(0) = 1, \quad \frac{\partial g}{\partial u}(1) &= 0. \end{aligned} \tag{4.9}$$

where the boundary conditions are chosen to exhibit that the process is absorbing at 0 and reflecting at 1. By the formula (4.8), equation (4.9) is equivalent to

$$\begin{aligned} H(u)g'(u) + \frac{\sigma^2}{2}g''(u) - \alpha g(u) = 0, \\ g(0) = 1, \qquad \frac{\partial g}{\partial u}(1) = 0. \end{aligned} \tag{4.10}$$

If, in equation (4.9), we take limit as $\alpha \to 0^+$, keeping in mind that g is bounded by one, we obtain

$$\mathcal{A}\left[\lim_{\alpha\to 0^+} g(u)\right] - \underbrace{\lim_{\alpha\to 0^+} (\alpha g(u))}_{=0} = 0 \tag{4.11}$$

Thus, the function $p^0(u)$ satisfies the Cauchy problem

$$\begin{aligned} H(u)\frac{d}{du}[p^0(u)] + \frac{\sigma^2}{2}\frac{d^2}{du^2}[p^0(u)] = 0, \\ p^0(0) = 1, \qquad \frac{d}{du}[p^0(1)] = 0. \end{aligned} \tag{4.12}$$

Since the coefficients in (4.12) are $C^1[0,1]$, there is a unique solution to this boundary problem. We can check easily that $p^0 \equiv 1$ is the solution, hence $\mathbb{P}_u(\tau < \infty) = 1$. □

We finish this discussion with the general formula for the solution of (4.12) since we feel that other models, with different boundary conditions, can be studied later. For instance, we could explore the possibility that the process dies, instead of reflecting, at the boundary point $u = 1$. The general solution of (4.12) is

$$p^0(u) = K_1 \int_0^u \left[\exp\left(-\frac{2}{\sigma^2}\int_0^x H(y)dy\right)dx\right] + K_0.$$

It is immediate that the current case corresponds to $K_1 = 0$ and $K_0 = 1$.

4.3 Expected Time of Extinction

Proposition 4.3. *The expected time of extinction, of the process u_t defined as the solution of (4.2) which is absorbing at 0 and reflecting at 1, is given by the formula*

$$\mathbb{E}_u[\tau] = \frac{2}{\sigma^2}\int_0^u \mathrm{e}^{-\frac{2}{\sigma^2}\int_0^x H(s)ds}\left(\int_x^1 \mathrm{e}^{\frac{2}{\sigma^2}\int_0^y H(s)ds} dy\right)dx. \tag{4.13}$$

for every initial point $0 \le u \le 1$.

Proof. Differentiating equation (4.9) with respect to α we obtain

$$g(u) + \alpha\frac{\partial g}{\partial\alpha}(u) - \mathcal{A}\frac{\partial g}{\partial\alpha}(u) = 0 \tag{4.14}$$

Taking limit as $\alpha \to 0^+$ in equation (4.14), we get

$$\lim_{\alpha\to 0^+} g(u) + \underbrace{\lim_{\alpha\to 0^+}\left(\alpha\frac{\partial g}{\partial\alpha}(u)\right)}_{=0} + \mathcal{A}\left[-\lim_{\alpha\to 0^+}\frac{\partial g}{\partial\alpha}(u)\right] = 0. \tag{4.15}$$

By proposition (4.2), $\lim_{\alpha\to 0^+} g(u) = \mathbb{P}_u(\tau < \infty) = 1$. Thus we obtain that $e^0(u)$ satisfies the initial value problem

$$
\begin{aligned}
H(u)\frac{d}{du}[e^0(u)] + \frac{\sigma^2}{2}\frac{d^2}{du^2}[e^0(u)] &= -1, \\
e^0(0) = 0, \qquad \frac{d}{du}[e^0(1)] &= 0
\end{aligned}
\tag{4.16}
$$

To solve (4.16), we multiply by the integrating factor $\frac{2}{\sigma^2}\,\mathrm{e}^{\frac{2}{\sigma^2}\int_0^u H(s)ds}$ to obtain the equation

$$
\frac{d}{du}\left[\mathrm{e}^{\frac{2}{\sigma^2}\int_0^u H(s)ds}\frac{d}{du}[e^0(u)]\right] = -\frac{2}{\sigma^2}\,\mathrm{e}^{\frac{2}{\sigma^2}\int_0^u H(s)ds} \tag{4.17}
$$

We integrate and apply the boundary condition $w'(1) = 0$ and obtain

$$
\frac{d}{du}[e^0(u)] = \frac{2}{\sigma^2}\,\mathrm{e}^{-\frac{2}{\sigma^2}\int_0^u H(s)ds}\int_u^1 \mathrm{e}^{\frac{2}{\sigma^2}\int_0^x H(s)ds}dx \tag{4.18}
$$

The second boundary condition, $w(0) = 0$, implies that the particular solution is

$$
\mathbb{E}_u[\tau] = e^0(u) = \frac{2}{\sigma^2}\int_0^u \mathrm{e}^{-\frac{2}{\sigma^2}\int_0^x H(s)ds}\left(\int_x^1 \mathrm{e}^{\frac{2}{\sigma^2}\int_0^y H(s)ds}dy\right)dx. \tag{4.19}
$$

□

Remark.

We recall that

$$
H(u) = -\frac{1}{N}u + cu^a(1-u)\,, \qquad \int_0^x H(u)du = -\frac{1}{2N}x^2 + \frac{c}{a+1}x^{a+1} - \frac{c}{a+2}x^{a+2}\,.
$$

Even the case $a = 0$ (a linear diffusion) is nontrivial, even though classical; but an value $a > 0$ gives a highly nonlinear equation. The integrals in expression (4.19) can be quite complicated for a general function H. For instance, if $H(u)$ has a a positive integer, the solutions are related with various special functions (Airy, Error, Kummer, etc).

4.4 Quasi-Stationary Distribution

The diffusion defined (4.2) possesses a transition kernel $p(t, u, v)dv$ on the state space $S = [0, 1]$, which generates a (strongly) Feller semigroup S_t on $C(S)$, the Banach space of continuous functions with the uniform norm. More precisely, for f continuous on S,

$$S_t f(u) = \int_S p(t, u, v) f(v) dv \, .$$

A left-hand side positive eigenfunction (here, a Radon measure) of the semigroup is said a quasi-stationary distribution or qsd for short. There are several ways to identify the qsd's. Note that we allow, in general, the possibility of more than one (or, in fact, none) qsd. However, in the present case uniqueness will be a consequence of the fact that the semigroup is compact. We adopt the Sturm-Liouville approach to the problem, which is justified by the smooth kernel, which shows that the semigroup can be extended to $L^2(S, w(u)du)$, where $w(u) > 0$, continuous on S. Then, the existence of a largest eigenvalue, with its unique positive eigenfunction, is a classical result in the theory of second order differential equations with smooth coefficients. An alternative approach is to notice that the Krein-Rutman theorem applies, but we shall not pursue this approach due to the natural setting of Sturm-Liouville formalism in L^2 universally available in dimension one.

The semigroup is strongly continuous, contractive (C_0 - semigroup) with an infinitesimal generator $(\mathcal{A}, D(\mathcal{A}))$, where $\mathcal{A}$ is defined in (4.8). The domain includes smooth functions with continuous derivatives up to the boundary satisfying the boundary conditions we imposed

$$D(\mathcal{A}) \supseteq \{ f \in C^2(S) \, | \, f'(1) = 0 \quad \text{and} \quad f(0) = 0 \} \, .$$

Then, $\nu \in M_1(S)$ (the space of probability measures) is a qsd for S_t if there exists $\lambda > 0$ such that $\nu\mathcal{A} = \lambda\nu$. Since, the domain (4.4) includes a defining class for probability measures, this is the same as (4.20).

We begin the study of the quasi-stationary distribution ν, directly, as a left eigenvalue of the generator $\mathcal{A}$. Thus we have to find a positive radon measure ν on S and a number $\lambda < 0$ such that

$$< \lambda\nu(u) - \nu(u)\mathcal{A}, f(u) >= 0 \tag{4.20}$$

for f a twice continuously differentiable function satisfying $f(0) = 0$, $f'(1) = 0$. If ν exists, due to the absolute continuity of the transition probability and the smoothness of $p(t, u, v)$, the qsd is absolutely continuous as well as we see in the proof of Theorem 4.4.

Let the density function be denoted φ, thus $\nu(u) = \varphi(u)du$, and equation (4.20) takes the form

$$\int_0^1 \left(\lambda\varphi(u)f(u) - \varphi(u)H(u)f'(u) - \frac{\sigma^2}{2}f''(u)\varphi(u)\right)du = 0.$$

We apply integration by parts and obtain an equation for the formal adjoint operator

$$\int_0^1 \left[\lambda\varphi(u) + (H(u)\varphi(u))' - \frac{\sigma^2}{2}\varphi(u)''\right] f(u)du \\ + \frac{\sigma^2}{2}\varphi(0)f'(0) + \left[\frac{\sigma^2}{2}\varphi'(1) - H(1)\varphi(1)\right] f(1) = 0\,. \tag{4.21}$$

Hence, in order to find ν, we must solve the second order differential equation with boundary conditions

$$\begin{aligned} &\lambda\varphi(u) + (H(u)\varphi(u))' - \frac{\sigma^2}{2}\varphi''(u) = 0 \\ &\varphi(0) = 0\,, \qquad \frac{\sigma^2}{2}\varphi'(1) + \frac{1}{N}\varphi(1) = 0 \end{aligned} \tag{4.22}$$

First, we should recast equation (4.22) in Sturm-Liouville form (self-adjoint form), whose general theory is summarized in the upcoming Section 4.5. Expanding $(H(u)\varphi(u))' = H'(u)\varphi(u) + H(u)\varphi'(u)$ and introducing the integrating factor

$$\rho(u) = \exp\left(-\frac{2}{\sigma^2}\int_0^u H(x)dx\right) \tag{4.23}$$

equation (4.22) becomes

$$\begin{gathered} -(\rho(u)\varphi'(u))' + \left[\frac{2}{\sigma^2}\rho(u)H'(u) - \lambda\left(-\frac{2}{\sigma^2}\rho(u)\right)\right]\varphi(u) = 0 \\ \varphi(0) = 0\,, \qquad \frac{\sigma^2}{2}\varphi'(1) + \frac{1}{N}\varphi(1) = 0 \end{gathered} \tag{4.24}$$

We introduce the notation, consistent to the Sturm-Liouville form given in the next section in eq. (4.31)

$$q(u) = \frac{2}{\sigma^2}\rho(u)H'(u)\,, \qquad \tilde{\lambda} = -\frac{2}{\sigma^2}\lambda\,, \qquad w(u) = \rho(u) \tag{4.25}$$

and, corresponding to the boundary conditions (4.30).

$$\alpha = 1, \quad \alpha' = 0\,, \qquad \beta = \frac{1}{N}\,, \quad \beta' = \frac{\sigma^2}{2}\,. \tag{4.26}$$

With this notation, equation (4.24) is exactly (4.31).

Theorem 4.4. *The process defined by (4.8) has a unique qsd ν having a continuous on* $[0,1]$*, twice differentiable, positive density on* $(0,1)$.

Proof. Existence. Once we have established (4.24), we see from the explicit formulas for its coefficients that all conditions of Theorem 4.6 are satisfied, i.e. the coefficients are continuous and $w(u) > 0$ on $[0,1]$. Moreover, from (4.22) is equivalent to (4.24), so the left eigenvalue $\tilde{\lambda}_0$, respective eigenfunction ψ_0 from Theorem 4.6 is a $C^2[0,1]$ function, positive on $(0,1)$. The positivity comes from the fact that the operator has a compact resolvent (Theorem 7.5.4 in [15]); this is directly stated in Theorem 4.6, being the case for all Sturm-Liouville operators. The function ψ_0 is continuous, thus integrable on $[0,1]$, which means $\nu(u) = C_0^{-1}\psi_0(u)du$ is a probability distribution, where $C_0 > 0$ is the normalizing constant.

Uniqueness. Let $\nu(du)$ be a qsd for $\mathcal{A}$ defined in (4.8), i.e. a positive, finite measure on $[0,1]$ satisfying the weak equation (4.20). Since the coefficients of $\mathcal{A}$ are $C^2[0,1]$ and the boundary conditions are classical, one dimensional, the solution $p(T-t,u,v)$, $t \in [0,T]$, for any $T > 0$ is jointly continuous up to the boundary and of class C^2 in (u,v) for $s > 0$. This implies that a left-side eigenfunction has a density. To see that, for any $\phi \in \mathcal{D}(\mathcal{A})$,

$$e^{\lambda_0 s}\int_0^1 \nu(dv)\phi(v) = \int_0^1\int_0^1 \nu(du)p(s,u,v)\phi(v)dv$$

implying that

$$e^{\lambda_0 s}\nu(dv) = \int_0^1 \nu(du)p(s,u,v)\,,$$

which is a C^2 function of $v \in [0,1]$. Notice that (4.20) implies the exponentiation relation for the semigroup shown in the equations from above. We've proved that ν is an $L^2([0,1], w(du))$ positive eigenfunction of (4.29). The only positive eigenfunction of (4.29) is, modulo a constant, the eigenfunction ψ_0. This is also unique, since the eigenspace of $\tilde{\lambda}_0$ is one dimensional. It cannot be equal to another eigenfunction, because it would be orthogonal; two positive orthogonal functions would be zero almost everywhere. Since both are also continuous, uniqueness is proven. □

Thus the system (4.24) can be written in the classical context of the *Sturm-Liouville* theory as:

$$
\begin{aligned}
&-(\rho(u)\varphi'(u))' + \left(q(u) - \rho(u)\tilde{\lambda}\right)\varphi(u) = 0 \\
&\varphi(0) = 0 \qquad \frac{\sigma^2}{2}\varphi'(1) + \frac{1}{N}\varphi(1) = 0
\end{aligned}
\tag{4.27}
$$

The generality of equation (4.27) makes the process of solving it explicit very difficult. However, this is a *Regular Sturm-Liouville Operator*. By the theory developed is section 4.5, we get very nice properties for the Sturm-Liouville operator associated with (4.27), such as compactness, self-adjointness and Fourier expansions of the solution in terms of orthogonal eigenfunctions. But most important, we get the existence of the quasi-stationary distribution by Corollary 4.7. Next, we give some numerical computation of the qsd.

4.5 Regular Sturm-Liouville Operators

In this section we review the theory of *Regular Sturm-Liouville Operators*, following very closely the corresponding chapter of [15] and [19], the latter especially for the particular form of Theorem 4.6. Such operators appear in the investigation of second-order differential equations, in dimension one, defined on a closed interval $a \leq u \leq b$, of the form

$$
-(\rho\varphi') + \left(q - \tilde{\lambda}w\right)\varphi = 0 \tag{4.28}
$$

together with special boundary conditions. For sufficiently smooth coefficients, any one dimensional second order time-homogeneous pde can be put in this form. We ask for $\rho \in C^1[a,b]$, $q, w \in C[a,b]$ and $\tilde{\lambda}$ constant. In addition, we impose $\rho(u) > 0$ and $w(u) > 0$ for every $a \leq u \leq b$.

We shall define the class of *Regular Sturm-Liouville operators* on the complex Hilbert space $\mathcal{H} = L^2([a,b], wdu)$

$$\mathcal{H} = \left\{ x(u) : \int_a^b |x(u)|^2 w(u) du < \infty \right\}$$

with inner product

$$< x, y >= \int_a^b x(u) y(\bar{u}) w(u) du \,.$$

Definition 4.5. *An operator defined on the subspace $\mathcal{D}$ of C^2-function $\varphi \in \mathcal{H}$ of the form*

$$L\varphi = \frac{1}{w} \left[-(\rho\varphi')' + q\varphi \right] \tag{4.29}$$

is called a Regular Sturm-Liouville Operator if for every function $\varphi \in \mathcal{D}$, satisfies the boundary conditions

$$\begin{aligned} (BC_1) \qquad & \alpha\varphi(a) + \alpha'\varphi'(a) = 0 \\ (BC_2) \qquad & \beta\varphi(b) + \beta'\varphi'(b) = 0 \,, \end{aligned} \tag{4.30}$$

where α, α', β, β' are real numbers such that $|\alpha| + |\alpha'| > 0$, $|\beta| + |\beta'| > 0$.

Remark. The definition implies that the domain of a Regular Sturm-Liouville operator L includes $\mathcal{D}$. Also, since we trivially have the inclusions $C_c^\infty[a,b] = C^\infty[a,b] \subset \mathcal{D}$, the operator L is densely defined.

Given an Sturm-Liouville operator L and a real number $\tilde{\lambda}$, the operator

$$L\varphi = \frac{1}{w} \left[-(\rho\varphi')' + (q - \tilde{\lambda} w)\varphi \right] \tag{4.31}$$

is a Regular Sturm-Liouville operator if it satisfies the boundary conditions BC_1 and BC_2.

A Regular Sturm-Liouville operator is symmetric [15]. Furthermore, for any real number $\tilde{\lambda}$ in the resolvent of the operator (4.29), the operator $(\tilde{\lambda}I - L)$ is invertible, and $(\tilde{\lambda}I - L)^{-1}$ is compact and self-adjoint [15]. It follows that since $\mathcal{H}$ is a Hilbert space, any Regular Sturm-Liouville operator L has at most a countable number of eigenvalues, and that these must be real. The following theorem provides a better description of the behavior of the eigenvalues and the corresponding eigenfunctions.

Theorem 4.6. *(Theorem 5, 10.8 in [19]) Any regular Sturm-Liouville system has an infinite sequence of real eigenvalues* $\tilde{\lambda}_0 < \tilde{\lambda}_1 < \tilde{\lambda}_2 < \cdots$ *with* $\lim_{n\to\infty} \tilde{\lambda}_n = \infty$*. The eigenfunction* $\psi_n(u)$ *belonging to the eigenvalue* $\tilde{\lambda}_n$ *has exactly n zeros in the interval* $a < u < b$ *and it is uniquely determined up to a constant factor.*

Corollary 4.7. *A regular Sturm-Liouville operator has a unique quasi-invariant distribution, which is the normalized eigenfunction associated with the smallest eigenvalue* $\tilde{\lambda}_0$.

Proof. By Theorem 4.6, the eigenfunction ψ_0 does not vanish over the compact interval $[a, b]$, and it is continuous since it belongs to $\mathcal{D}$. Thus, ψ_0 is integrable and does not change sign in $[a, b]$. Hence, we can define a probability measure by normalizing ψ_0 by its integral over $[a, b]$, this give us the quasi-invariant distribution. □

4.6 Conjecture 4.1 and Numerical Analysis

In this section we explore the application of the software Maple 2017 in order to find the quasi-stationary distribution. However, without any help, Maple is not able to solve the problem for us. A procedure is needed in order to guide the software through the process. The strategy is:

First, to find a fundamental system for the Sturm-Liouville problem. This is a system of two linearly independent solutions. Then, we argue as follows, given ψ_1, ψ_2 two linearly independent solutions of the Sturm-Liouville problem, a real number λ is an eigenvalue of of the operator L if and only if the determinant

$$Det(\lambda) = \begin{vmatrix} BC_1(\psi_1) & BC_1(\psi_2) \\ BC_2(\psi_1) & BC_2(\psi_2) \end{vmatrix} = 0. \tag{4.32}$$

This is because every solution ψ is a linear combination $\psi = c_1\psi_1 + c_2\psi_2$. Thus, applying the boundary conditions we obtain the system of equations

$$\begin{aligned} c_1 BC_1(\psi_1) + c_2 BC_1(\psi_2) &= 0 \\ c_1 BC_2(\psi_1) + c_2 BC_2(\psi_2) &= 0. \end{aligned} \tag{4.33}$$

If λ is an eigenvalue, then $Det(\lambda) = 0$ since we can find a nontrivial solution of (4.33). Conversely, if $Det(\lambda) = 0$, then there are constants c_1, c_2 that solve (4.33) such that $c_1c_2 \neq 0$. Thus, $\psi = c_1\psi_1 + c_2\psi_2$ is an eigenfunction corresponding to the eigenvalue λ.

We must solve for the smallest solution of equation (4.32). This is the first eigenvalue λ_0. We recall that, by (4.6), the set of solutions is countable and bounded below, and that the corresponding eigenfunction is positive in the interval $(0, 1)$.

Once we obtain λ_0, we can evaluate any of the equations in (4.33) and solve for one of the variables, say c_2 in terms of c_1. Finally, we normalize the eigenfunction $\psi = c_1\psi_1 + c_2(c_1)\psi_2$ to obtain the quasi-invariant distribution.

In order to make numerical computations less complicated, it is necessary to simplify (4.27) as much as possible. In order to eliminate the first derivative term we introduce the change of variables:

$$\begin{aligned}
&\underline{\psi(u) = \varphi(u)\sqrt{\rho(u)}}\\
&\psi'(u) = \varphi'(u)\sqrt{\rho(u)} + \varphi(u)\frac{\rho'(u)}{2\sqrt{\rho(u)}}\\
&\psi''(u) = \varphi''(u)\sqrt{\rho(u)} + \varphi'(u)\frac{\rho'(u)}{\sqrt{\rho(u)}} + \varphi(u)(\sqrt{\rho(u)})''.
\end{aligned} \tag{4.34}$$

We obtain the new regular problem for $\psi(u)$,

$$\begin{aligned}
\psi''(u) + \tilde{\lambda}\,\psi(u) &= Q(u)\,\psi(u)\\
\psi(0) = 0, \quad \sigma^2\psi'(1) + \frac{1}{N}\psi(1) &= 0
\end{aligned} \tag{4.35}$$

where

$$Q(u) = \frac{q(u) + \sqrt{\rho(u)}(\sqrt{\rho(u)})''}{\rho(u)} = \frac{1}{\sigma^2}H'(u) + \left(\frac{1}{\sigma^2}H(u)\right)^2 \tag{4.36}$$

We define the operator

$$L := -\psi''(u) + Q(u)\psi(u) \tag{4.37}$$

toguether with the boundary operators

$$BC_1(\psi) = \psi(0) \qquad BC_2(\psi) = \sigma^2\psi'(1) + \frac{1}{N}\psi(1). \tag{4.38}$$

We will illustrate how to obtain the quasi-invariant distribution for the system (4.35) in the particular cases, $a = 0$ and $a = 1$. The initialization for Maple is:

> *with(LinearAlgebra):*

> *with(Student[Calculus1]):*

> *with(RootFinding):*

> *with(MathematicalFunctions):*

> *with(PDEtools):*

>*_EnvAllSolutions := true*

4.6.1 Special Case, Explicit Solution $a = 0$

We analyze the special case in which the parameter $a = 0$. In this case the function H takes the simpler form $H(u) = -\frac{1}{N}u + c(1-u)$. We write equation (4.35) and the boundary conditions.

> $H(u) := -\frac{1}{N}u + c(1-u)$

> $Q(u) := \frac{1}{\sigma^2}H'(u) + \left(\frac{1}{\sigma^2}H(u)\right)^2$

> $DE := \psi''(u) + \tilde{\lambda}\psi(u) = Q(u)\psi(u)$

The general solution of the equation DE is

$$\_C2 \exp\left(\frac{u(c(u-2)N+u)}{2\sigma^2 N}\right)(c(u-1)N+u)\,\text{hypergeom}\left(\frac{2+(-\sigma^2\lambda+2c)N}{4cN+4}, \frac{3}{2}, \frac{(c(u-1)N+u)^2}{(cN+1)\sigma^2 N}\right) + \\ + \ \_C1\,\text{hypergeom}\left(\frac{\sigma^2 N\lambda}{4cN+4}, \frac{1}{2}, \frac{(c(u-1)N+u)^2}{(cN+1)\sigma^2 N}\right).$$

Where hypergeom (n, d, z) is the generalized *hypergeometric function* $F(n, d, z)$. This was obtain by

> *sol(u) := rhs(dsolve(DE))*

> $eq1 := sol(u)\Big|_{u=0} = 0$

> $eq2 := \sigma^2 sol'(u)\Big|_{u=1} + \frac{1}{N}sol(u)\Big|_{u=1} = 0$

We write the system of equations $\{eq1, eq2\}$ as a matrix, and impose the condition that the determinant of such matrix is zero. This will provide us with a nontrivial solution for the constants $\_C1$, $\_C2$.

> *M := GenerateMatrix([eq1, eq2], [_C1, _C2])[1]*

> *EQ := Determinant(M) = 0*

> *H := (lambda, N, sigma, c) → lhs(EQ)*

We take this moment to evaluate the parameters λ, N, σ, and c by some particular values for computational purposes.

> *f := simplify(eval(H(lambda, N, sigma, c), [c = 0.02, N = 100, sigma = 1]))*

We plot the determinant to check that our results are consistent with the theory.

> *plot(f, lambda = 0 .. 500, color = [red], thickness = 2)*

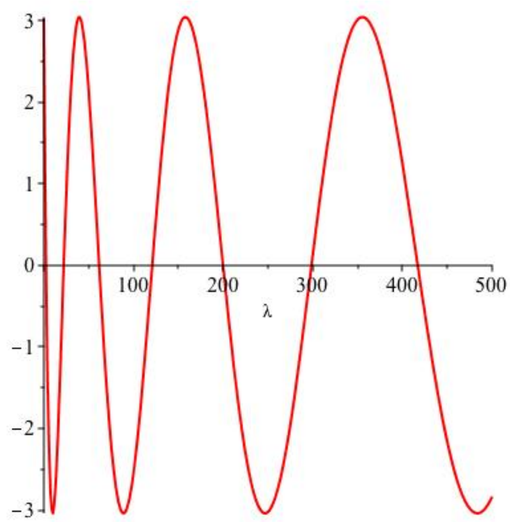

Figure 4.1: *Graphs of the determinant of the matrix M as a function of* λ *for* $N = 100$, $c = 0.02$, $\sigma = 1$.

We are interested in the smallest eigenvalue. The command *fsolve* finds all real roots of the function f on some interval. From graph (4.1) we can see that the first root lies in $(0, 10)$.

> *ev := fsolve(f, lambda = 0 .. 10)*

Thus we obtain that the first root is 2.457399838. This is the eigenvalue λ_0 corresponding with the *quasi-stationary distribution*. For such value λ_0, equations $eq1$ and $eq2$ are redundant. We could solve for $\_C2$ in $eq1$.

> *s := eval(solve(eq1, _C2), [lambda = ev, N = 100, sigma = 1, c = 0.02, _C1 = 1])*

We obtain $\_C2 = .3005649041$. The corresponding eigenfunction is, up to a constant, obtained by

> g := x → eval(sol(x), [lambda = ev, N = 100, sigma = 0.02, c = 1, _C1 = 1, _C2 = s]) We finally graph the eigenfunction corresponding to the first eigenvalue λ_0.

> plot(g(x), x = 0 .. 1)

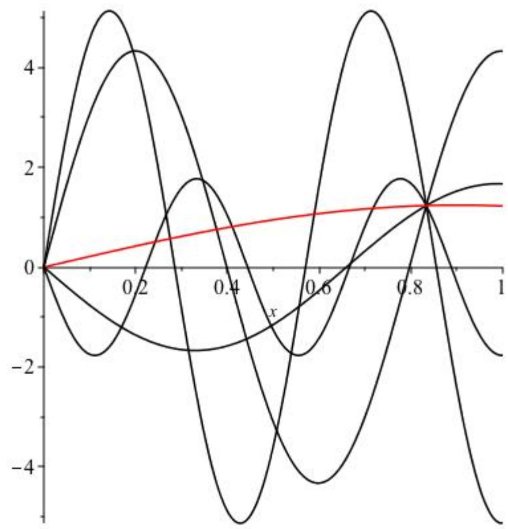

Figure 4.2: *Graphs of the quasi-stationary distribution (in red), and the next four eigenfunctions for* $a = 0$, $N = 100$, $c = 0.02$, $\sigma = 1$

Conjecture 4.1 - Analysis when $\sigma \to 0$

Finally, to obtain the *quasi-invariant distribution* we need to normalize the eigenfunction. We compute the area under the curve by

> area := int(ef1(x), x = 0 .. 1, numeric = true).

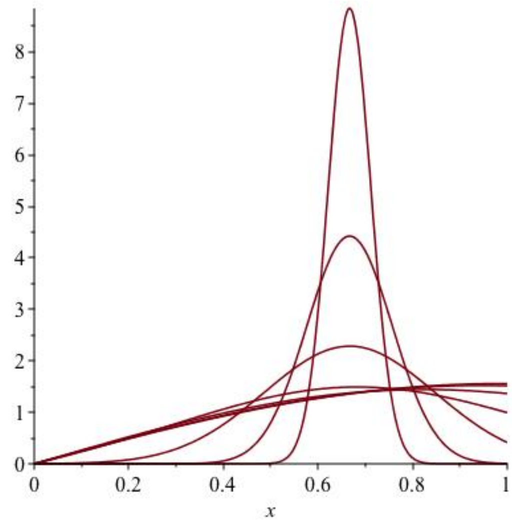

Figure 4.3: *Graphs of the quasi-stacionary distribution as* $\sigma \to 0$*, for* $a = 0$*,* $N = 100$*,* $c = 0.02$*,* $\sigma = \frac{1}{2^i},\ 1 \le i \le 7$.

We can see the quasi-stationary distribution converges to the unique equilibrium value. For $N = 100$, $c = 0.02$ and $a = 0$, by (2.3.5), the stable point is given by

$$u_2 = (1 + \frac{1}{Nc})^{-1} = .66666\bar{6}.$$

4.6.2 Special Case, Explicit Solution $a = 1$

We analyze the special case in which the parameter $a = 1$. In this case the function H takes the simpler form $H(u) = -\frac{1}{N}u + cu(1 - u)$. We write equation (4.35) and the boundary conditions and obtain that it has general solution given by

$$\_C1 \exp\left(\frac{x^2\left(\frac{3}{2} + N(x - \frac{3}{2})c\right)}{3\sigma^2 N}\right) \mathrm{HeunT}\left(\frac{(18\sigma^4)^{1/3}\lambda}{2c^{2/3}}, 3, -\frac{18^{1/3}(cN-1)^2}{4(c\sigma)^{4/3}}, \frac{18^{1/3}\left(\frac{1}{2} + c(x - \frac{1}{2})N\right)}{3c^2\sigma^{2/3}N}\right) +$$
$$+\ \_C2 \exp\left(\frac{x^2\left(\frac{3}{2} + N(x - \frac{3}{2})c\right)}{3\sigma^2 N}\right) \mathrm{HeunT}\left(\frac{(18\sigma^4)^{1/3}\lambda}{2c^{2/3}}, -3, -\frac{18^{1/3}(cN-1)^2}{4(c\sigma)^{4/3}}, -\frac{18^{1/3}\left(\frac{1}{2} + c(x - \frac{1}{2})N\right)}{3c^2\sigma^{2/3}N}\right).$$

Where HeunT $(\alpha, \beta, \gamma, z)$ is the solution of the *Heun Triconfluent Equation*. This is, again, obtained by

> $> sol(u) := rhs(dsolve(DE))$

> $> eq1 := sol(u)\Big|_{u=0} = 0$

> $> eq2 := \sigma^2 sol'(u)\Big|_{u=1} + \frac{1}{N} sol(u)\Big|_{u=1} = 0$

As before we write the system of equations $\{eq1,\ eq2\}$ as a matrix, and impose the condition that the determinant of such matrix is zero. This will provide us with a nontrivial solution for the constants $\_C1$, $\_C2$. Plotting the determinant as a function of λ we obtain

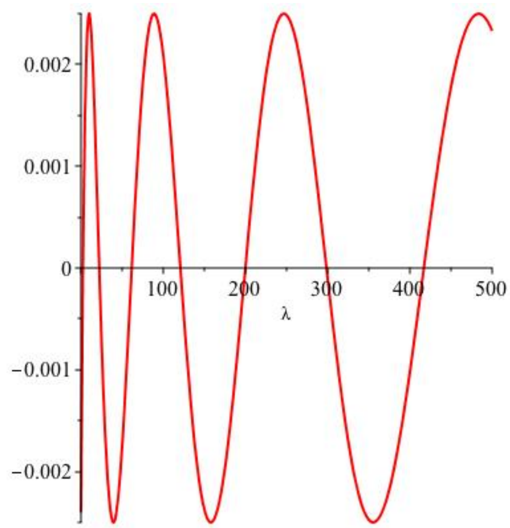

Figure 4.4: *Graphs of the determinant of the matrix M as a function of λ for $N = 100$, $c = 0.02$, $\sigma = 1$.*

We apply *fsolve* again to find all real roots of the function f on $(0, 10)$ and we obtain that the first root is 2.469293165. This is the eigenvalue λ_1 corresponding with the *quasi-stationary distribution*. For such value λ_1 we obtain $\_C2 = -.9997542569$. The corresponding eigenfunction is, up to a constant, obtained by

> *> g := x → eval(sol(x), [lambda = ev, N = 10000, sigma = 1, c = 1, _C1 = 1, _C2 = s])*

We finally graph the *quasi-invariant measure*.

> plot(g(x), x = 0 .. 1)

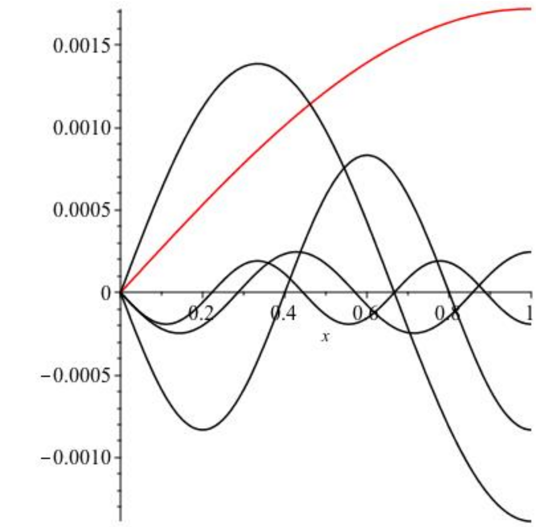

Figure 4.5: *Graphs of the quasi-stacionary distribution (in red), and the next four eigenfunctions for* $a = 1$, $N = 100$, $c = 0.02$, $\sigma = 1$.

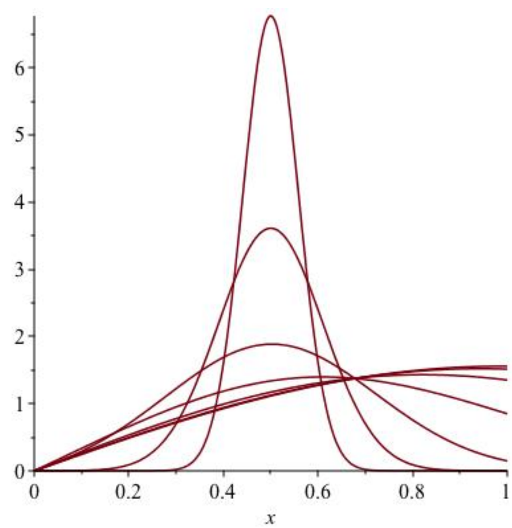

Figure 4.6: *Graphs of the quasi-stacionary distribution as* $\sigma \to 0$, *for* $a = 1$, $N = 100$, $c = 0.02$, $\sigma = \frac{1}{2^i}$, $1 \leq i \leq 7$.

This is the behavior of the sequence of points, at which the quasi-stationary distribution attains maximum, converging to the equilibrium value $q(1) = 1 - \frac{1}{Nc} = 0.5$. (see (2.5.1)).

Chapter 5

Future Research

- Refine the numerical calculations using MatLab.
- Prove Conjecture 4.1 under appropriate conditions.
- Investigate both perturbation results in higher dimensions and generalize to state dependent noise.
- Modify the nature of the escape probability from $\gamma(u) = cu^a$ (power law) to another monotone law, with some different tail behavior.
- Investigate the problem of random fixation/escape for a different mutation mechanism. Here we essentially have only two types of characters (zero, for fixation, and non-zero, for regular) while we could have a random walk on the alphabet $\{0, 1, \ldots, N - 1\}$ with the same fixation mechanism.

Appendix A

Discrete and continuous time Markov chains

A.1 Markov Chains

In this subsection we briefly present the general definitions of Markov Chains. The theory between the discrete and continuous cases is very close and the main object driving these toguether is the jump process Y_n introduced before.

Let $I \subseteq \mathbb{R}^n$ (the state space) and $X : \Omega \to I$ a random variable. The *distribution* of X is the vector μ with components given by

$$\mu_i = \mathbb{P}(\{\omega : X(\omega) = i\}) \quad i \in I.$$

A matrix $P = (p_{i,j}),\ i, j \in I$ is called an *Stochastic Matrix* if $0 \leq p_{i,j} \leq \infty$ and

$$\sum_{i \in I} p_{i,j} = 1.$$

Definition A.1 (*Discrete Markov Chain*). *Consider a countable family of random variables $X_n \to I$. Let $P = (p_{i,j})$ be an stochastic matrix with values on I. The family X_n is a Discrete*

Markov Chain with initial distribution μ *and transition matrix* P *if the random variable* X_0 *has distribution* λ *and*

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, ..., X_0 = i_0) = p_{i_n, i_{n+1}}.$$

The entries $p_{i_n, i_{n+1}}$ are called *transition probabilities* and we write X_n is *Markov*(μ, P) for a discrete Markov chain X_n defined as before. We mention that, in order to define a discrete Markov chain X_n, it's only needed to prescribe the initial distribution and the transition matrix P. The process is then completely described since, given that X_n is at state i, we know what stages it could reach and with what probability. Similarly, for a continuous process X_t the main object needed to describe its behavior is the *generator matrix* or *Q-matrix*, described in Definition A.6.

A.2 Pure Jump Processes

Markov chain theory is very vast, even when limited to the pure-jump case. We present concepts, definitions and a few martingale theory results that will be important for our work. We follow closely Norris [17] for notation and definitions. For a deeper understanding of the theory we point to the monograph by Roger and Williams [22].

For the rest of the chapter let I be a countable set. We will refer to the elements of the set I as *states*, and by I as *state space*.

Definition A.2 (*Pure Jump Continuous Random Processes*)**.** *By a Pure Jump Continuous Random Process*

$$(X_t)_{t\geq 0} = \{X(t, \omega) : 0 \leq t < \infty,\ \omega \in \Omega\}$$

on a stage space I, we mean a family of random variables $X_t : \Omega \to I$ *satisfying that:*

for all $\omega \in \Omega$, $t \geq 0$*, there is an* $\epsilon > 0$ *such that*

$$X_s = X_t \quad t \leq s \leq t + \epsilon.$$

The above definition implies that, a pure jump process must remain constant in a particular state for some period of time, until it *"jumps"* to a new one. The time the process spends between two consecutive stages is called a *holding time*, and the instant in which the process changes stages is referred as a *jump time*. The sum of all the holding times is referred as *explosion time*. We formalize these notions by giving an explicit relation with the process $(X_t)_{t\geq 0}$.

Definition A.3 (*Jump Times*)**.**

$$J_0 = 0, \quad J_{n+1} = \inf\{t \geq J_n : X_t \neq X_{J_n}\} \quad ,n = 0, 1, 2, ...$$

Definition A.4 (*Holding Times*)**.**

$$S_n = \begin{cases} J_n - J_{n-1} & ,J_{n-1} < \infty \\ \infty & ,J_{n-1} = \infty \end{cases} \qquad n = 0, 1, 2, ...$$

Definition A.5 (*Explosion Time*)**.**

$$\zeta = \sum_{n=1}^{\infty} S_n .$$

If we add an special state $\{\infty\}$ to the set I then, a process X_t is called a *minimal process* iff $X_t = \infty$, $\forall t \geq \zeta$. We must introduce another useful process associated with $(X_t)_{t\geq 0}$, the discrete process $Y_n = X_{J_n}$, $n = 0, 1, 2, ...$ is called the *jump process*. It turns out that the jump process is an essential tool in the study of the process X_t.

Definition A.6. *A Q-matrix on the state space I is a matrix* $Q = (q_{i,j}), i, j \in I$ *satisfying that:*

- $0 \leq q_{i,j} \quad i \neq j;$
- $0 \leq -q_{i,i} < \infty \quad i \in I;$
- $\sum_{j \in I} q_{i,j} = 0$

The entry $-q_{i,i}$ is denoted for short by $q_i = q(i)$.

Out of any Q-matrix Q we can obtain a transition matrix P, called the *jump matrix* Π, as follows:

$$\pi_{i,j} = \begin{cases} \frac{q_{i,j}}{q_i} & i \neq j, \ q_i \neq 0 \\ 0 & i \neq j, \ q_i = 0 \end{cases}$$

$$\pi_{i,i} = \begin{cases} 0 & q_i \neq 0 \\ 1 & q_i = 0 \end{cases}$$

We are now in position to define a continuous Markov process.

Definition A.7 (*Continuous Markov Pure Jump Process*)**.** *A minimal pure jump continuous processes* X_t *on I is a Markov Chain with initial distribution* μ *and matrix Q (Q - matrix) if its jump chain* Y_n *is* $Markov(\lambda, \Pi)$ *and if for each* $n \geq 1$*, conditional on* $Y_0, ..., Y_{n-1}$*, its holding times* $S_1, ..., S_n$ *are independent exponential random variables of parameters* $q(Y_0)$*, . . . ,* $q(Y_{n-1})$ *respectively.*

We omit the details of the construction of such processes but we refer the reader to [17] and [24] in case of interest. At this moment it is only important for us to note that a continuous time Markov pure jump process X_t on the state space I is completely determined if we describe the intensities of the holding times and the behavior of X_t at jump times. The generator matrix of the process is an special case of a more general notion, the *generator of a Markov process*. We should discuss a little bit further such theory next since both presentations will be used later.

Definition A.8 (*Generator of a pure jump Markov Process*)**.** *Let X_t be a continuous time pure jump Markov process on $I \subseteq \mathbb{R}^n$. Let $f \in C(\mathbb{R}^n, \mathbb{R})$. The (infinitesimal) generator $\mathcal{A}$ of X_t is defined by*

$$\mathcal{A}f(i) = \lim_{t \to 0} \frac{\mathbb{E}_u[f(X_t)] - f(i)}{t}; \qquad i \in I\,,$$

whenever it exists. The set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the limit exists is the domain of the generator and is denoted by $\mathcal{D}_{\mathcal{A}}$.

Remark. We note that for a finite state continuous Markov chain

$$\mathcal{A}f(u) = \sum_{j \neq i} q_{i,j}(f(j) - f(i))\,.$$

Martingales Associated to a Continuous Time Process

In general a Markov process does not need to be a martingale, not a martingale is always a Markov process. However, associated with any Markov process we can define a pair of martingales that are very useful.

$$M_t^g = g(X_t) - g(X_0) - \int_0^t \mathcal{A}g(X_s)ds \tag{A.1}$$

$$N_t^g = (M_t^g)^2 - \int_0^t \left[\mathcal{A}g^2(X_s) - 2g(X_s)\mathcal{A}g(X_s)\right]ds \tag{A.2}$$

Lemma A.9. *(Simplified version of Lemma 5.1 Appendix 1 in [14]) Let* $(X_t)_{t\geq 0}$ *be a continuous time, pure jump Markov process on a probability space* $(\Omega, \mathcal{F}, P)$ *adapted to the filtration* $(\mathcal{F}_t)_{t\geq 0}$*. For any test function* $g(t,x) \in C_0^2([0,\infty)\times\mathbb{R},\mathbb{R})$*, the processes* M_t^g *and* N_t^g *are* $\mathcal{F}_t$*-martingales.*

Appendix B

A Useful Formula Used in Chapter 3

In Chapter 3, we used several times an estimate of the following type. If $f \in C_c(\mathbb{R})$ such that f^{n+1} exists and is continuous on an open interval containing $[u', u' + \frac{y_s^L}{\sqrt{L}}]$, then by Taylor's formula with remainder in the integral form

$$f(u' + \frac{y_s^L}{\sqrt{L}}) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(u') \left(\frac{y_s^L}{\sqrt{L}}\right)^k + R_n(u' + \frac{y_s^L}{\sqrt{L}}) \tag{B.1}$$

where

$$R_n(u' + \frac{y_s^L}{\sqrt{L}}) = \frac{1}{n!} \left(\frac{y_s^L}{\sqrt{L}}\right)^{n+1} \int_0^1 w^n f^{(n+1)} \left(\left(u' + (1-w)\frac{y_s^L}{\sqrt{L}}\right)^k\right) dw. \tag{B.2}$$

The remainder is a process and since f^{n+1} is bounded on $[u', u' + \frac{y_s^L}{\sqrt{L}}]$, by some constant $c(f)$, then R_n satisfies

$$R_n(u' + \frac{y_s^L}{\sqrt{L}}) \leq \frac{1}{(n+1)!} \left(\frac{y_s^L}{\sqrt{L}}\right)^{n+1} c(f). \tag{B.3}$$

Bibliography

[1] Athreya, K. B. and Ney, P. E. (1972). Branching Processes. Springer, New York.

[2] Nils Berglund and Barbara Gentz *Noise-Induced Phenomena in Slow-Fast Dynamical Systems - A Sample Path Approach*, Springer 2006.

[3] L. Alili, P. Patie and J.L. Pedersen. *Representations of the First Hitting Time Density of an Ornstein-Uhlenbeck Process*, `http://www.math.ku.dk/~jesper/research/denou.pdf`

[4] Anna De Masi, Errico Presutti *Mathematical Methods for Hydrodynamic Limits*, Springer 1991.

[5] Rick Durrett, Jasmine Foo, Kevin Leder, John Mayberry, and Franziska Michor. *Evolutionary dynamics of tumor progression with random fitness values Theor Popul Biol.* 2010 Aug; 78(1): 54-66.

[6] Durrett, Richard *Branching process models of cancer.* Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems, 1.1. Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015.

[7] Warren J. Ewens *Mathematical Population Genetics. I - Theoretical Introduction*, Springer (2004).

[8] Feller, William. *The birth and death processes as diffusion processes.* J. Math. Pures Appl. **38**, 301-345 (1959).

[9] Hirsch, M, W., Smale, S. and Devaney, R. L. (2004). *Differential Equations, Dynamical Systems, and an Introduction to Chaos.* Elsevier, USA.

[10] Karatzas, I.; Shreve, S.E. *Brownian Motion and Stochastic Calculus* Springer-Verlag, Berlin, 1988

[11] Harris, T. E.; *The Theory of Branching Processes* Dover, New York, 1989.

[12] May, R. M. *Simple mathematical models with very complicated dynamics.* Nature (1976); 261: 459 - 67.

[13] Grigorescu, Ilie; Kang, Min *Fixation time for an evolutionary model.* Stoch. Models 29 (2013), no. 3, 328 - 340.

[14] C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems*, Springer-Verlag, New York, 1999.

[15] Naylor, Arch; Sell, George. *Linear Operator Theory in Engineering and Science.* Springer-Verlag, New York, 2000.

[16] Lopez-Ruiz, R.; Fournier-Prunaret, D. *Periodic and chaotic events in a discrete model of logistic type for the competitive interaction of two species.* Chaos Solitons Fractals 41 (2009), no. 1, 334 - 347.

[17] Norris, J. R. *Markov Chains.* Cambridge University Press, 1997.

[18] Strogatz, Steven H. *Nonlinear Dynamics and Chaos.* Perseus Books, 1994.

[19] Garrett Birkhoff, Gian-Carlo Rota. *Ordinary Differential Equations.* John Wiley & Sons, 1989.

[20] Cristian Tomasetti and Bert Vogelstein *Variation in cancer risk among tissues can be explained by the number of stem cell divisions Science*, 02 Jan 2015: Vol. 347, Issue 6217, pp. 78 - 81.

[21] Herbert S. Wilf and Warren J. Ewens *There's plenty of time for evolution* PNAS 2010 107: 22454-22456. `www.pnas.org/cgi/doi/10.1073/pnas.1016207107`

[22] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales*, Vol. 1, 2nd ed., Cambridge University Press 2000.

[23] Daniel W. Stroock and S.R.S. Varadhan *Multidimensional Diffusion Processes*, Springer, 1997 Edition.

[24] Zhe Zhang, *Scaling Limit of a Generalized Polya Urn Model* (2015). Open Access Dissertations. Paper 1490.