UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
DISSERTAÇÃO DE MESTRADO

CONSTRUCTION OF THE DAWSON-WATANABE PROCESS

EDUARDO SAMPAIO PIMENTA

Salvador-Bahia
Fevereiro de 2021

CONSTRUCTION OF THE DAWSON-WATANABE PROCESS

EDUARDO SAMPAIO PIMENTA

Dissertação de Mestrado apresentada ao Colegiado da Pós-Graduação em Matemática da Universidade Federal da Bahia como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Tertuliano Franco.

Salvador-Bahia
Março de 2021

Construção do Processo de Dawson-Watanabe

Eduardo Sampaio Pimenta

Dissertação apresentada ao Colegiado do Curso de Pós-graduação em Matemática da Universidade Federal da Bahia, como requisito parcial para obtenção do Título de Mestre em Matemática.

Banca examinadora

Tertuliano Franco

Prof. Dr. Tertuliano Franco Santos Franco (UFBA)

Prof. Dr. Dirk Erhard (UFBA)

Prof. Dr. Hubert Lacoin (IMPA)

À minha família, aos meus amigos e aos meus professores de Matemática.

Agradecimentos

Agradeço, inicialmente, aos meus pais Antônio Batista e Maristânea Alves, por sempre demonstrarem tamanho carinho e empatia, por serem minha primeira referência de amor e respeito.

Às minhas irmãs: Ana Clara, Daniela e Maria Alice por me apoiarem sempre, por todos estes anos me acompanhando, ajudando e por saber que sempre poderei contar com elas para tudo e ao meu sobrinho incrivelmente fofo e sapeca que alegra meus dias de uma forma quase mágica, Benício.

Aos meus demais familiares, tios e primos, por me motivarem, pelas festas em família, em particular aos meus primos, Alexandre, Matheus, Renan e Victor, por além de tudo, serem amigos inestimáveis.

Aos meus novos amigos que o mestrado em Salvador me trouxe: Bruna, Caio, Carol, Diego, Elivan, Enio, Glaene, Iago, Janara, Jonathas, Leandro Correia, Leandro Dorea, Leydiane, Rafael, Savio, Tamires, Valéria, Yure, Enathiele, Joedson, Moacyr e Pedro, a vocês, meu muito obrigado por esses meses de muito aprendizado e todo crescimento que tive para além do meio acadêmico. Ao meu grande amigo Juan, que por todos os sabados estudando e as poucas idas aos bares e praia, realmente um amigo admirável o qual respeito como pessoa e como matemático. Há também aqueles outros que não necessariamente estavam no programa, os quais gostaria de agradecer pelos bons momentos dentre eles Laércio, Edward e Jacqueline.

Aos meus amigos de Belo Horizonte Alan Nascimento, Amanda Lobo, Carla Ferreira, Júlia Drumond, Natália Moller, Pedro Rodrigues, Phillippe Samer, Rafael Souto e Sergio Martins que me apoiaram e se mantiveram presentes, mesmo com a distância física e meus horários difíceis. Ao Rodrigo Viégas, o motivo da bela imagem no começo do segundo capítulo da dissertação. Também ao Paulo Velloso, Marcos Costa, Lucas e Samuel Quirino pelas seções de rpg, mesmo que acontecessem uma vez por mês. Em particular, aos meus amigos mais antigos por ter crescido tanto com eles, apesar de alguns já não estarem mais tão próximos: Matheus Salles, Augusto, Gabriel, Humberto, João Pedro, João Gabriel, Michael e Warley e aos tantos outros amigos não citados.

A Taís Jesus de Brito e sua família, por me receber em sua casa durante a pandemia, um momento tão delicado e difícil e por todos os outros momentos de apoio incondicional, pelo carinho imenso que espero ter retribuído e por tornar as manhãs chuvosas em tardes ensolaradas.

Agradeço ao Prof. Dr. Hubert Lacoin por aceitar avaliar este trabalho. Ao Prof. Dr. Dirk Erhard, por além de estar na banca, me ensinar tanto durante este meu percurso no mestrado e ser sempre solícito a ajudar. Foi realmente um privilégio. Ao meu orientador prof. Dr. Tertuliano, pelo exemplo de professor e pesquisador, se desdobrando para nos orientar durante esta pandemia, mesmo com dois filhos muito fofos demandando uma energia tremenda, é uma referência como pessoa, professor e pesquisador. Foi uma experiência muito boa mesmo trabalhar com você durante o mestrado.

A todos professores doutores que me ensinaram durante o meu percurso no mestrado, por terem contribuído de tamanha forma para minha formação como mestre: Vilton Jeovan Viana Pinheiro, Georg Wilhelm Klein, Jaime Leonardo Orjuela Chamorro, Mathieu Molitor, Cristina Lizana Araneda e os já citados Tertuliano Francos Santos Franco e Dirk Erhard, e também aos professores que me ensinaram durante minha graduação, em particular, ao professor Dr. Marcelo Richard Hilário e Remy de Paiva Sanchis por fazerem despertar, cada um de sua forma, a minha paixão pela Probabilidade.

Por fim, agradeço ao Programa de Pós-Graduação em Matemática da UFBA pela oportunidade, e à CAPES, pelo apoio financeiro a este trabalho.

*“The success is the sum of small efforts,
repeated day-in and day-out”*

–Robert Collier

Resumo

Esta dissertação tem como objetivo a construção do processo de Dawson-Watanabe como limite de escala do movimento browniano ramificado, sendo que este último também qual é aqui relacionado à solução de uma equação do calor com fonte.

A existência do processo de Dawson-Watanabe (também chamado de superbrowniano ou superprocesso) é uma consequência de tal processo ser solução de um problema martingal obtido através do limite de escala do movimento browniano ramificado. A prova segue a estrutura clássica de rigidez e unicidade de pontos limite, sendo que a unicidade, neste caminho seguido, decorre da caracterização do processo de Dawson-Watanabe como dual da solução de uma certa equação diferencial parcial.

Palavras-chave: Processo de Dawson-Watanabe, Super Movimento Browniano.

Abstract

This dissertation aims at the construction of the Dawson-Watanabe process as a scaling limit of the branching Brownian motion, the latter being also related here to the solution of a heat equation with a source.

The existence of the Dawson-Watanabe process (also called super Brownian motion or superprocess) is a consequence that this process is a solution to a martingal problem obtained through the scale limit of the branching Brownian motion. The proof follows the classic structure of tightness and uniqueness of limit points, and the uniqueness, in this followed path, results from the characterization of the Dawson-Watanabe process as a dual solution of a certain partial differential equation.

Keywords: Super Brownian motion, Dawson-Watanabe superprocess.

Contents

Introduction

Much of the current research on probability theory is concerned with stochastic processes taking values in infinite-dimensional spaces. In 1951, Feller in [10] observed that the evolution of large populations could be study from a Galton-Watson branching process by re-scaling and passing to the limit, and such device was known as a Feller diffusion approximation. Superprocesses arise from this idea where we study not only the size of large populations, but also their spatial distribution.

The Dawson-Watanabe superprocess is one example of such superprocesses. In general terms, the Dawson-Watanabe process can be seen as a scaling limit of interacting particles systems, in which we give a distribution for the size of the population according to a reproduction law and we also give a spatial movement of each particle given by a Brownian motion.

The branching Brownian motion, rigorously constructed in [13, 14, 3], is a stochastic process with relevance in the theory of the superprocesses since from it we may construct the Dawson-Watanabe superprocess through a scaling limit. Intuitively, the branching Brownian motion is defined as follows: we give to a particle starting at $x \in \mathbb{R}^d$ a spatial movement given by a Brownian motion $\mathbb{P}_x$ and, at exponential random time, this particle dies. Then, according to some probability generating function, a random number of descendants of this ancestor is generated and each of the descendants performs a Brownian motion in $\mathbb{R}^d$.

Observe that such a process is not described by a continuous trajectory, since it may generate or destroy mass every time a particle in the system dies. We thus need to work with the space of càdlàg functions under a suitable topology: the Skorohod space [8].

As mentioned before, the Dawson-Watanabe superprocess will be constructed here as the scaling limit of the branching Brownian motion which, in a certain sense can be seen as a "discrete particle" counterpart. We will characterize the branching Brownian motion as a solution of the reaction-diffusion equation, first studied by Skorohod in [23], and posteriorly by H. McKean in [20]. Once the expectation of certain function of the branching Brownian motion was characterized as a solution of a reaction-diffusion equation, we may obtain its infinitesimal generator and consequently the Dynkin martingales. Taking the limit of such Dynkin martingales eventually serves us as a basis to show the existence and uniqueness of the superprocess.

The construction of the scaling limit process will be done by resizing the initial measure, we will no longer start with a single particle, but with many ones of small mass, wheres the total mass is of order of some constant. We will also give a shorter lifetime to each particle and a longer time for the spatial movement of Brownian motions to evolve.

The dissertation is divided into three chapters. The first one recalls some probabilistic background, basic concepts and stochastic processes, as well as the construction and understanding of the Skorohod space, an important tool to study superprocesses. The second chapter focuses on the construction of the branching Brownian motion as a the junction between a Galton-Watson process and a Brownian motion, its characterization via a solution to certain diffusion-reaction equation, and as a solution of a martingal problem. In the third chapter, we will start by making the scale limit of the branching Brownian motion, we will show certain important tools to show its uniqueness, here we can cite Prohorov's classic theorem, and the Aldous-Rebolledo criterion [21] criterion to guarantee the existence of the superprocess and we will also use the duality method to relate the Dawson-Watanabe superprocess to its deterministic dual process which we know is unique.

This dissertation aims to present a fluid and simple way to understand the book of Alison Etheridge [7] and generate an intuition to deal with more advanced problems, as perhaps analogous superprocess with boundary conditions such as reflection or absorption, or even a sticky branching Brownian motion.

Chapter 1

Preliminaries

1.1 Basic notions

let $(\Omega, \mathcal{U}, \mathbb{P})$ be a probability space, we denote $\mathcal{L}_1(\Omega, \mathcal{U}, \mathbb{P})$ ($\mathcal{L}_1(P)$) to be the space of measurable random variables $X : \Omega \to \mathbb{R}$ with $\mathbb{E}[|X|] < \infty$ and $L_1 := \mathcal{L}_1(P)/\sim$ where two random variables a in relation to each other, if they are equal almost everywhere.

Definition 1.1. *If $(S, \mathcal{S})$ is a measure space then a stochastic process with state space S is a collection $\{X_t : t \in I\}$ of random variables*

$$X_t : \Omega \to S.$$

More generally, we will consider processes with finite life-time. Here we add an extra point Δ to the state space and we endow $S_\Delta = S \cup \{\Delta\}$ with the σ-algebra $\mathcal{S}_\Delta = \{B \cup \Delta : B \in \mathcal{S}\}$. A stochastic process with state space S and life-time ξ is defined as a process

$$X_t : \Omega \to S_\Delta,$$

and $X_t(w) = \Delta$ if $t \geq \xi(w)$ where $\xi : \Omega \to [0, \infty]$ is a random variable.

Remark 1.1. We will assumes that space S is a Polish space, i.e., S is a complete separable metric space. Note that for example open sets in $\mathbb{R}^n$ are polish spaces, although they are not complete w.r.t. the Euclidean metric.

Definition 1.2. *A filtration on $(\Omega, \mathcal{U}, \mathbb{P})$ is an increasing collection $(\mathcal{F}_t)_{t \in I}$ of σ-algebras on $\mathcal{U}$.*

A stochastic process $(X_t : t \in I)$ is called adapted with respect to filtration $(\mathcal{F}_t)_{t \in I}$ if X_t is $\mathcal{F}_t$-measurable for any $t \in I$.

Given a process $X = \{X_t : t \in I\}$, we have that the X is always adapted to the σ-algebra generated by X,

$$\mathcal{F}_t^X = \sigma(X_s : s \leq t),$$

for each $t \in I$. $\mathcal{F}_t^X$ is called the natural σ-algebra of the process X.

1.2 Poisson Point Process

The Poisson point processes is an important class of stochastic process used to model the occurrence or arrival of events over a continuous interval. We will present three different characterizations.

Recall that a Poisson random variable X with parameter λ has the distribution function given by

$$\mathbb{P}[X = k] = \begin{cases} \frac{\exp(-\lambda)\lambda^k}{k!} & \text{if } k = 0,1,2,\dots \\ 0 & \text{otherwise.} \end{cases}$$

1.2.1 Counting Process

We define a counting process $\{N_t : t \geq 0\}$ as a collection of non-negative, integer-valued random variables such that if $0 \leq s \leq t$, then $N_s \leq N_t$.

A Poisson process is a particular case of counting process. Consider a collection of events that arrives at random times starting at $t = 0$, let N_t denote the number of arrivals that occur at the time t, that is, the number of events in $[0, t]$. Suppose that we would like to model the arrival of events that happen completely at random at a rate λ per unit time. Observe that, at time $t = 0$, we have no arrivals and therefore, $N_0 = 0$.

Now, we will divide the interval $[0, \infty)$ into tiny sub-intervals of length δ to some $\delta > 0$. Thus, we have that $[0, \infty) = \bigcup_{k=0}^{\infty} [k\delta, (k+1)\delta)$ is a partition of the half-line $\mathbb{R}^+$. Now, for each sub-interval $[k\delta, (k+1)\delta)$, assume that we toss a coin with probability $p = \lambda\delta$ of landing head. If it lands head, we say that we have an arrival in that sub-interval, otherwise, there is no arrival.

Let N_t the number of arrivals from time 0 to time t. Observe that there are, approximately $\frac{t}{\delta}$ time slots in the interval $(0, t]$ and N_t is the numbers of heads in n coin flips.

Thus, we conclude that N_t is a Binomial(n, p). Observe that

$$np = n\lambda\delta = \frac{t}{\delta}\lambda\delta = \lambda\delta.$$

We know that, if Y_n are Binomial(n, p) such that $np \to \lambda$ to some $\lambda > 0$ then Y_n converges to a random variable X which is a Poisson with parameter λ. [6, Theorem 3.6.1., page 126].

Definition 1.3. *A Poisson process with paramenter $\lambda > 0$ is a counting process $\{N_t : t \geq 0\}$ with the following properties:*

(i) $N_0 = 0$

(ii) For all $t > 0$, N_t has a Poisson distribuition with parameter λ.

(iii) (Stationary increments) For all $s, t > 0$, $N_{t+s} - N_s$ has the same distribution as N_t. That is

$$P(N_{t+s} - N_s = k) = P(N_t = k) = \frac{\exp(-\lambda t)(\lambda t)^k}{k!} \textit{ for } k = 0,1,2,\dots$$

(iv) (Independent increments) For $0 \leq q < r \leq s < t$, $N_t - N_s$ and $N_r - N_q$ are independent random variables.

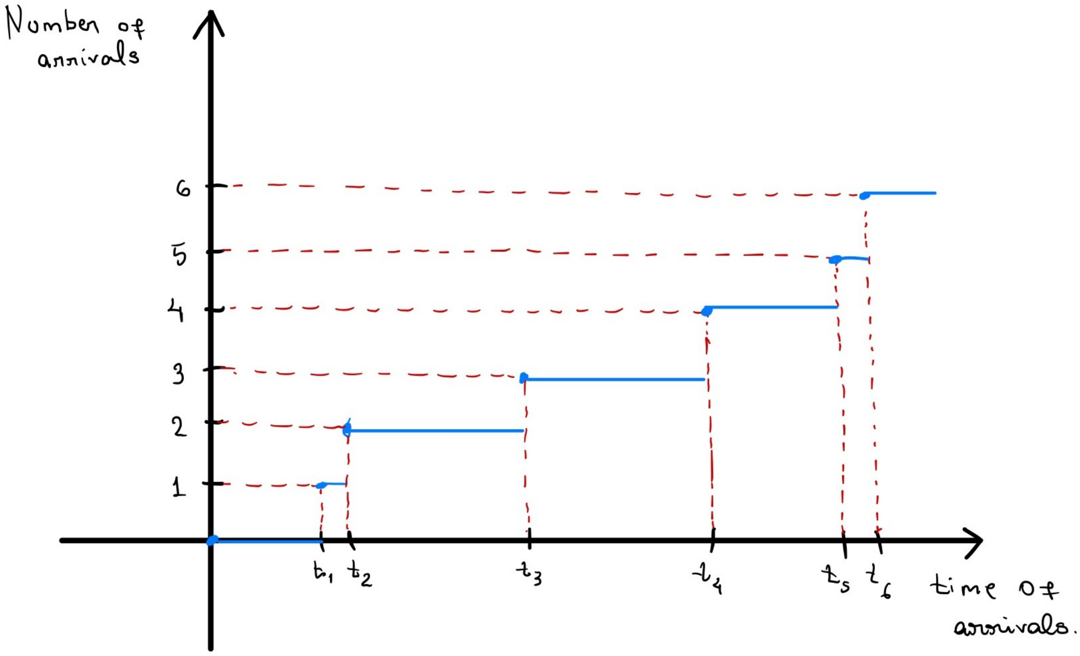


Figure 1.1: Counting process

Observe that the stationary increments property says that the distribution of the number of arrivals in an interval only depends on the length of the interval while the independent increments says that the number of arrivals on disjoint intervals are independent random variables.

The Figure 1.1 illustrate the path of counting process in which events occur in random times $t_1, \ldots, t_6$.

1.2.2 Arrival and interarrival times

Let N_t be a Poisson process with parameter λ and X_1 be the time of first arrival. Thus

$$\mathbb{P}[X_1 > t] = \mathbb{P}[\text{ there is no arrival in } (0,t]] = \exp(-\lambda t)$$

Therefore, we conclude that X is exponentially distributed with parameter λ. We will do a heuristic for this construction. Let X_2 the time between the first and the second arrival and consider $s > 0$ and $t > 0$. Observe that the intervals $(0, s]$ and $(s, s+t]$ are disjoint. Since the increments are independents, it follows that

$$\begin{aligned}
\mathbb{P}[X_2 > t \,|\, X_1 = s] &= \mathbb{P}[\text{ there is no arrival in } (s, s+t] \,|\, X_1 = s] \\
&= \mathbb{P}[\text{ there is no arrival in } (s, s+t]] \\
&= \mathbb{P}[N_{t+s} - N_s = 0] \\
&= \mathbb{P}[N_t = 0] = \exp(-\lambda t),
\end{aligned}$$

and we conclude that X_2 are exponentially distributed with parameter λ and X_1 and X_2 are independents. The random variables $X_1, X_2, \ldots$ are called the inter-arrival times of the counting process N_t. For a rigorous construction see [6, page 134]. Thus, another way to define the Poisson point process is:

Definition 1.4. *Let $X_1, X_2, \ldots$ be a sequence of i.i.d. exponential random variables with parameter $\lambda > 0$. For $t > 0$, let*

$$N_t = \max\{n : X_1 + \cdots + X_n \leq t\}$$

with $N_0 = 0$. Then $\{N_t : t \geq 0\}$ defines a Poisson process with parameter λ. Let

$$S_n = X_1 + \cdots + X_n, \text{ for } n = 1, 2, \ldots$$

We call $S_1, S_2 \ldots$ the arrival times of the process, where S_k is the time of the k-th arrival. Furthermore, let

$$X_k = S_k - S_{k-1}, \text{ for } k = 1, 2, \ldots$$

is the time between the $(k-1)$-th and k-th arrivals, with $S_0 = 0$.

Remember that if X is exponentially distributed with parameter λ, then X is memoryless random variable, i.e.,

$$\mathbb{P}[X > a + x \,|X > a] = \mathbb{P}[X > x]$$

for all $a, x \geq 0$. The memoryless of inter-arrival times is consistent with the independent increment property of the Poisson process.

To show the equivalence between these two definitions of the Poisson point process and consider X_i with $i = 0, 1, \ldots$ and N_t as in the definition of arrival and inter-arrival Poisson point process. Observe that for $k \geq 0$, $N_t = k$ if and only if $S_k \leq t < S_k + X_{k+1}$ and the densisty of S_k and X_{k+1} since they are independent is given by

$$f_{S_k, X_{k+1}}(s, x) = f_{S_k}(s) f_{X_{K+1}}(x) = \frac{\lambda^k s^{k-1} \exp(-\lambda s)}{(k-1)!} \lambda \exp(-\lambda x).$$

For $k \geq 0$

$$\begin{aligned}
\mathbb{P}[N_t = k] &= \mathbb{P}[S_k \leq t \leq S_k + X_{k+1}] \\
&= \mathbb{P}[S_k \leq t, X_{k+1} \geq t - S_k] \\
&= \int_0^t \int_0^\infty \frac{\lambda^k s^{k-1} \exp(-\lambda s)}{(k-1)!} \lambda \exp(-\lambda x)\, dx\, ds \\
&= \frac{\lambda^k}{(k-1)!} \int_0^t (s^{k-1} \exp(-\lambda(t-s))) ds \\
&= \frac{\exp(-\lambda t) \lambda^k}{(k-1)!} \int_0^t s^{k-1} ds = \frac{\exp(-\lambda t) \lambda^k}{k!}.
\end{aligned}$$

Which gives the distribution of the first definition and therefore, we have proved the following Proposition.

Proposition 1.1. *The definitions 1.3 and 1.4 are equivalent.*

1.2.3 Infinitesimal characterization

We can also define a Poisson point process via on an infinitesimal characterization. We write $f(h) = o(h)$ whenever

$$\lim_{h \to 0} \frac{f(h)}{h} = 0.$$

More generally, we say that f has the same asymptotic behavior of g, and write $f(h) = o(g(h))$, if

$$\lim_{h \to 0} \frac{f(h)}{g(h)} = 0.$$

Let N_t be a Poisson Process with rate λ. Consider a very short interval of length h. Then, the number of arrivals in this interval has the same distribution as N_h. Particularly, using the Taylor expansion we have that

$$\begin{aligned}\mathbb{P}[N_h = 0] &= \exp(-\lambda h) = 1 - \lambda h + \frac{\lambda^2}{2} h^2 - \dots \\ &= 1 - \lambda h + o(h).\end{aligned}$$

Analogously, it follows that

$$\begin{aligned}\mathbb{P}[N_h = 1] &= \exp(-\lambda h)(\lambda h) \\ &= (1 - \lambda h + o(h))(\exp(-\lambda h)) \\ &= \lambda h + o(h).\end{aligned}$$

and then, we have another way to define the Poisson point process. Rigorously,

Definition 1.5. *A Poisson process with parameter λ is a counting process $\{N_t : t \geq 0\}$ with the following properties:*

(i) $N_0 = 0$,

(ii) The process has stationary and independent increments,

(iii) $\mathbb{P}(N_h = 0) = 1 - \lambda h + o(h)$,

(iv) $\mathbb{P}(N_h = 1) = \lambda h + o(h)$.

Properties *(iii)* and *(iv)* essentially say that there is only be finitely many arrivals in a finite interval, and in an infinitesimal interval it may occur at most one arrival.

Example 1.1. Let $\{N_t : t \geq 0\}$ be a Poisson process. Then, by definition, we have that

$$\mathbb{P}(N_t = k) = \frac{\exp(-\lambda t)(\lambda t)^k}{k!}.$$

In particular $\mathbb{P}(N_t > 1) = 1 - \mathbb{P}(N_t = 0) - \mathbb{P}(N_t = 1)$. Thus,

$$\mathbb{P}(N_t > 1) = 1 - \exp(-\lambda t)(\lambda t)^0 - \exp(-\lambda t)(\lambda t).$$

Dividing $\mathbb{P}(N_t > 1)$ by t and taking the limit when t tends to 0

$$\begin{aligned}\lim_{t\to 0}\frac{\mathbb{P}(N_t>1)}{t} &= \lim_{t\to 0}\frac{1-\exp(-\lambda t)(\lambda t)^0-\exp(-\lambda t)(\lambda t)}{t}\\ &= \lim_{t\to 0}\frac{1-\exp(-\lambda t)}{t} - \lim_{t}\frac{\exp(-\lambda t)(\lambda t)}{t}\\ &= -(\exp(-\lambda t))'\big|_{t=0} - \lambda = 0.\end{aligned}$$

Therefore, $\mathbb{P}(N_t > 1) = o(t)$.

Assume that the definition 1.5 holds. We show that N_t has a Poisson distribution with parameter λt. Consider N_t the number of point in the interval $[0, t]$ and make a partition into n sub-intervals each of length $\frac{t}{n}$. For n sufficiently large, by the property *(iii)* and *(iv)* of the definition 1.5, the probability of each sub-interval has more than 1 arrivals are negligible.

Observe that each sub-interval has the form $(\frac{(k-1)t}{n}, \frac{kt}{n}]$ for $k = 0, 1, \ldots, n$ and by the stationary increments,

$$p_n := \mathbb{P}[N_{kt/n} - N_{(k-1)t/n} = 1] = \mathbb{P}[N_{t/n} = 1] = \frac{\lambda t}{n} + o\Big(\frac{t}{n}\Big).$$

Hence,

$$np_n = n\Big(\frac{\lambda t}{n} + o\Big(\frac{t}{n}\Big)\Big) = \lambda t + n\Big[o\Big(\frac{t}{n}\Big)\Big]$$

which converges to λt. By the theorem [6, Theorem 3.6.6. page 132], it follows that N_t is a Poisson with parameter λt. Therefore, we have that

Proposition 1.2. *The definitions 1.3 and 1.5 are equivalent.*

1.3 Markov Process and the Brownian Motion

The Markov process is a stochastic process that describes a sequence of possible events, in which the probability of the next state depends exclusively on the present state. Rigorously

Definition 1.6 (Markov Property). *A stochastic process $\{X_t : t \in I\}$ on $(\Omega, \mathcal{U}, \mathbb{P})$ with state space $(S, \mathcal{S})$ is called a $(\mathcal{F}_t)$-Markov process if and only if $\{X_t : t \in I\}$ is adapted w.r.t. the filtration $(\mathcal{F}_t)_{t\in I}$ and*

$$\mathbb{P}[X_t \in B|\mathcal{F}_s] = \mathbb{P}[X_t \in B|X_s] \tag{1.1}$$

$\mathbb{P}$-a.s. for any $B \in \mathcal{S}$ and $s, t \in I$ with $s \le t$.

Remark 1.2. A Markov process $\{X_t : t \in I\}$ is called a Markov chain when $I = \mathbb{N}$.

In addition of usual definition given by (1.1), we have other equivalents forms for the Markov Property. They are

$$\mathbb{P}[X_t \in B|\mathcal{F}_s] = P_{s,t}(X_s, B)$$

$\mathbb{P}$-a.s. for any $B \in \mathcal{S}$ with $s \le t$, and

$$\mathbb{E}\big[f\big(X_t\big)\big|\mathcal{F}_s\big] = (P_{s,t})f(X_s)$$

$\mathbb{P}$-a.s. for any measurable function and $s \le t$. In these definitions, $P_{s,t}(x, dy)$ is a regular version of the conditional probability distribution of X_t given $\mathcal{F}_s$ and

$$(P_{s,t}f)(x) = \int P_{s,t}(x, dy)f(y).$$

Moreover, by the item *(v)* of the Theorem A.3, the kernels $p_{s,t}$ satisfy the consistency condition

$$P_{s,u}(X_s, B) = \int P_{s,t}(X_s, dy)P_{t,u}(y, B),$$

$\mathbb{P}$-a.s. for any $B \in \mathcal{S}$ and $s \le t \le u$ and that is the same as

$$P_{s,u}f = P_{s,t}P_{t,u},$$

$\mathbb{P} \circ X_s^{-1}$-a.s. for all $0 \le s \le t \le u$.

Definition 1.7. *A probability kernel P on $(S, \mathcal{S})$ is a map $(x, B) \to P(x, B)$ from $S \times \mathcal{S}$ to $[0, 1]$ such that*

(i) for any $x \in S$, $P(x, \cdot)$ is a positive measure on $(S, \mathcal{S})$ with $P(x, S) = 1$,

(ii) for any $B \in \mathcal{S}$, $P(\cdot, B)$ is a measurable function on $(S, \mathcal{S})$.

A transition kernel is a collection of $P_{s,t}$ with $s, t \in I$ and $0 \le s \le t$, of probability kernels on $(S, \mathcal{S})$ satisfying

$$\begin{aligned} P_{t,t}(x, \cdot) &= \delta_x \text{ for any } x \in S \text{ and } t \in I \\ P_{s,t}P_{t,u} &= P_{s,u} \text{ for any } s \le t \le u \end{aligned} \tag{1.2}$$

and the composition of two probability kernels P and Q is the probability kernel PQ defined by

$$(PQ)(x, B) = \int P(x, dy)Q(y, B)$$

for any $x \in S$ and $B \in \mathcal{S}$. The equations in (1.2) are called the *Chapman-Kolmogorov equations*.

Definition 1.8 (Markov process with transition function)**.** *Let $P_{s,t}$ be a transition kernel. A stochastic process $\{X_t : t \in I\}$ on $(S, \mathcal{S}, \mathbb{P})$ is called an $(\mathcal{F}_t)$-Markov process with transition function $(P_{s,t})$ if, and only if, it is $(\mathcal{F}_t)$-adapted and*

$$\mathbb{P}[X_t \in B \,|\mathcal{F}_s] = P_{s,t}(X_s, B)$$

$\mathbb{P}$-a.s. for any $s \le t$ and $B \in \mathcal{S}$.

The process is called *time-homogeneous* if and only if there exist sub-probability kernels p_t such that $p_{s,t} = p_{t-s}$ for any $s \le t$.

Example 1.2 (Continuous time Markov chains)**.** If $\{X_t : t \in I\}$ is a time-homogeneous Markov chain on a polish space, and $\{N_t : t \geq 0\}$ is a Poisson process with parameter $\lambda > 0$ which is independent of $\{X_t : t \in I\}$, then the process

$$Y_t = X_{N_t}$$

is a time homogeneous Markov process in continuous-time with transition function

$$P_t(x, B) = \sum_{k \geq 0} \frac{\exp\{-\lambda t\}(\lambda t)^k}{k!} P_k(x, B).$$

Let the filtration $(\mathcal{F}_t)_{t \in I}$. Consider $T : \Omega \to [0, \infty)$ a map such that $\{w : T(w) \leq t\} \in \mathcal{F}_t$, for all $t \in [0, \infty)$ then T is called a $(\mathcal{F}_t)$-stopping time.

Definition 1.9 (Strong Markov Property)**.** *Let X be a càdlàg S-valued Markov process, $(\mathcal{F}_t : t \in I)$ the natural filtration and T a $(\mathcal{F}_t)$-stopping time. Then X is strong Markov at T if*

$$\mathbb{P}\Big[X(t+T) \in B \Big| \mathcal{F}_t\Big] = P_{t, X(t+T)}(X(t), B),$$

for all $t \geq 0$ and for all $B \in \mathscr{B}$.

We say that X is a strong Markov process if it has the strong Markov property for all $(\mathcal{F}_t)$-stopping times.

Definition 1.10. *A one-dimensional Brownian motion is a real-valued process $\{B_t : t \geq 0\}$ such that satisfies the following properties*

(i) The process $\{B(t) - B(s) : 0 \leq s \leq t\}$ are independent of $\{B(r) : r \leq s\}$ and $B(0) = 0$.

(ii) If $s, t \geq 0$ then

$$\mathbb{P}[B(s+t) - B(s) \in A] = \int_A \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{x^2}{2t}\Big) dx,$$

Note that for the d-dimensional Brownian motion, the transition function are given by

$$P_t(x, B) = \int_B \frac{1}{(2\pi t)^{d/2}} \exp\left(\frac{-\|x-y\|^2}{2t}\right) dy$$

where B is a Borel subset in $\mathbb{R}^d$.

(iii) With probability one, $t \to B_t$ is a continuous function.

Lemma 1.1 (Scaling invariance)**.** *Suppose that $\{B_t : t \geq 0\}$ is a standard Brownian motion and let $a > 0$. Then the process $X_t = \frac{1}{a} B_{a^2 t}$ is a standard Brownian motion.*

Proof. Consider the process $\{B(t) - B(s) : 0 \leq s \leq t\}$. Observe that this process have the distribution $\frac{1}{a^2}\mathcal{N}(0, a^2(t-s))$, i.e., has mean zero and variance $a^2(t-s)$. $\square$

Lemma 1.2 (Time inversion)**.** *Suppose that B_t is a standard Brownian motion. Then the process defined by*

$$X_t := \begin{cases} 0 & , \text{if } t = 0 \\ tB(\frac{1}{t}) & , \text{if } t > 0 \end{cases}$$

is a standard Brownian motion.

For a proof of above, see [6, page 302].

Definition 1.11. *if $B_1(t), \ldots, B_d(t)$ are all independent Brownian motions started in $x_1, \ldots, x_d$, then the randon process given by*

$$B_t = (B_1(t), \ldots, B_d(t)),$$

is called a d-dimensional Brownian motion started at $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

We define $\mathbb{P}_x$ as the probability measure which makes the random process $\{B_t : t \geq 0\}$ a d-dimensional Brownian motion started at $x \in \mathbb{R}^d$.

Theorem 1.1 (Markov property)**.** *Let $\{B(t) : t \geq 0\}$ is a Brownian motion started at $x \in \mathbb{R}^d$. Then the process $\{B(t+s) - B(s) : 0 \leq s \leq t\}$ is a Brownian motion started at origin and is independent of $\{B(t) : 0 \leq t \leq s\}$*

Definition 1.12. *The germ σ-algebra is defined as $\mathcal{F}^+(0)$ where*

$$\mathcal{F}_s^+ = \bigcap_{t>s} \mathcal{F}_t^0.$$

and $\mathcal{F}_s^0 := \sigma(B_r : r \leq s)$ is the natural σ-algebra generated by $\{B(t) : t \geq 0\}$.

Remark 1.3. Observe that the σ-algebra $\mathcal{F}_s^+$ give us more information than the natural σ-algebra also, note that $\mathcal{F}_s^+$ encodes information about a (infinitesimal) near future.

Let $x \in \mathbb{R}^d$ and $\mathbb{P}_x$ a probability measure on a measurable space $(\Omega, \mathcal{F})$, so that under $\mathbb{P}_x$, $B_t(w) = w(t)$ is a Brownian motion starting at x. In addition, for $s \geq 0$, define the *shift transformation* $\theta_s : \Omega \to \Omega$ such that $(\theta_s w)(t) = w(s+t)$.

Theorem 1.2 (Strong Markov Property)**.** *Let $(s, w) \to Y_s(w)$ be bounded and measurable. If S is a stopping time, then for all $x \in \mathbb{R}^d$*

$$\mathbb{E}_x[Y_S \circ \theta_S | \mathcal{F}_S] = \mathbb{E}_{B(S)}[Y_S]$$

where the right-hand side is the function $\phi(x, t) = \mathbb{E}_x Y_t$ evaluated at $x = B(S)$, $t = S$.

Proof. See [6, Theorem 8.3.7., page 314]. □

A particular case for the Brownian movement of Theorem 1.2 can be stated as

Theorem 1.3 (Strong Markov Property for Brownian motion)**.** *For every almost surely finite stopping time T, the process $\{W(T+t) - W(T) : t > 0\}$ is a standard Brownian motion independent of $\mathcal{F}_T^+$.*

Remark 1.4. The Donsker Theorem [4, page 68] is the classical construction of the Brownian motion which is one of the most important stochastic process in probability theory and statistical mechanics.

1.4 Martingales and infinitesimal generators

A martingale is a stochastic process for which, at a particular time, the conditional expectation of the next value in the sequence, given all prior values, is equal to the present value. Rigorously a stochastic process $X : \mathbb{R} \times \Omega \to \mathbb{R}$ is a continuous martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if

(i) $\{X_t : t \leq 0\}$ is adapted to $\mathcal{F}_t$,

(ii) for all t, $\mathbb{E}[|X_t|] < \infty$,

(iii) for all $s \leq t$, $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$.

Definition 1.13. *We will say that the process X_n is a local martingale if given a filtration $F_* = \{\mathcal{F}_t\}$ and $X : [0,\infty) \times \Omega \to S$ there exists a non-decreasing sequence $\{T_k\}$ of F_*-stopping times such that $T_k \uparrow \infty$ a.s. and, for every n, the stopped process M^{T_n} is a uniformly integrable martingale, i.e.,*

$$X_t^{T_k} := X_{\{\min\{t,T_k\}\}},$$

is an F_*-martingale for every k.

Definition 1.14. *A real valued process $X = \{X_t : t \geq 0\}$ is called a continuous semi-martingale if*

$$X_t = M_t + A_t,$$

where M is a continuous local martingale and A is a predictable process of locally bounded variation.

Proposition 1.3. *The semi-martingale X has unique decomposition in a continuous local martingale and a predictable process of locally bounded variation.*

Proof. Suppose that the semi-martingale can be decomposed in $X_t = M_t + A_t$ and $X_t = M_t' + A_t'$ where M_t and M_t' are continuous local martingales and A_t and A_t' are a predictable process of locally bounded variation. Thus

$$M_t + A_t = M_t' + A_t',$$

and therefore

$$M_t - M_t' = A_t - A_t'.$$

Now, observe that $A_t - A_t'$ is a continuous process of locally bounded variation and hence, $M_t - M_t'$ also is and by the theorem in [18, Theorem 4.8, page 78] it follows that $M_t = M_t'$ for all t almost surely. $\square$

The proposition above will be used in this dissertation only to evaluate some quadratic variation in the proof of Lemma 2.2.

1.4.1 Infinitesimal generator

Consider $X = \{X_t : t \geq 0\}$ be a Markov process. The infinitesimal generator $\mathcal{L}$ of X_t is defined as:

$$\mathcal{L}f(X_t) = \lim_{h\downarrow 0} \frac{1}{h}\mathbb{E}_\nu\big[f\big(X_{t+h}\big) - f(X_t)|X_t = x\big] = \lim_{h\downarrow 0} \frac{P_{t,t+h}f(x) - f(x)}{h},$$

given that this limit exists. Here, $P_tf(x) := \mathbb{E}_x\big[f\big(X_t\big)\big]$ denotes the transition semigroup of $\{X_t : t \geq 0\}$.

The set of functions $f : \mathbb{R}^d \to R$ such that the above limit exists at x is denoted by $\mathcal{D}(\mathcal{L}_x)$. Thus

$$\frac{d}{dt}P_tf(x) = \mathcal{L}f(x), \tag{1.3}$$

this means that the generator is the time derivative of the function $t \mapsto P_tf(x)$. Moreover, we can see the equation (1.3) as a partial differential equation. Indeed, let $u(t,x) = P_tf(x)$. Then, $u(t,x)$ solves

$$\begin{cases} \frac{d}{dt}u(t,x) = \mathcal{L}f(x) \\ u(0,x) = f(x). \end{cases}$$

Lemma 1.3. *For every bounded function f, the sequence $\{t^{-1}(P_tf - f) : t \geq 0\}$ converges uniformly to $\mathcal{L}f$ as $t\downarrow 0$.*

The proof of the lemma above can be found in [16, Lemma 3.1, page 323]

Proposition 1.4. *Let $\{X_t : t \geq 0\}$ be a Markov process. For every $f \in C_b^2(\mathbb{R}^d) \cap \mathcal{D}(\mathcal{L})$ and initial measure ν, the process*

$$M_t = f\big(X_t\big) - f\big(X_0\big) - \int_0^t \mathcal{L}f(X_s)ds,$$

is a $\mathbb{P}_\nu$-martingale, where $\mathcal{L}$ is the infinitesimal generator of X_t

Proof. Note that the process $\{M_t : t \geq 0\}$ is $\mathcal{F}_t$-adapted and

$$\begin{aligned}
\mathbb{E}_\nu[|M_t|] &= \mathbb{E}_\nu\left[\left|f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds\right|\right] \\
&\leq \mathbb{E}_\nu[|f(X_t)|] + \mathbb{E}_\nu[|f(X_0)|] + \mathbb{E}_\nu\left[\left|\int_0^t \mathcal{L}f(X_s)ds\right|\right] \\
&\leq \mathbb{E}_\nu[|f(X_t)|] + \mathbb{E}_\nu[|f(X_0)|] + \mathbb{E}_\nu\left[\int_0^t |\mathcal{L}f(X_s)|\, ds\right].
\end{aligned}$$

Since f is bounded, $\mathbb{E}_\nu[|M_t|] < \infty$. It remains to show that $\mathbb{E}_\nu[M_t|\mathcal{F}_s] = M_s$. By linearity of the conditional expectation, it follows that

$$\mathbb{E}_\nu\Big[M_t\,|\,\mathcal{F}_s\Big] = \mathbb{E}_\nu[f(X_t)\,|\,\mathcal{F}_s] - \mathbb{E}_\nu[f(X_0)\,|\,\mathcal{F}_s] - \mathbb{E}_\nu\left[\int_0^s \mathcal{L}f(X_u)\,du + \int_s^t \mathcal{L}f(X_u)\,du\,|\,\mathcal{F}_s\right].$$

Thus, since $\mathcal{L}$ is the infinitesimal generator of $\{X_t : t \geq 0\}$, by the theory of semigroup operators [8, Chapter 1, section 1-2],

$$\int_0^{t-s} \mathbb{E}_\nu[\mathcal{L}f(X_{u+s})]\, du \;=\; \mathbb{E}_\nu[f(X_t) \;-\; f(X_s)].$$

Now, consider $A \in \mathcal{F}_s$. By Markov property and the Fubini Theorem

$$\begin{aligned}
\mathbb{E}_\nu\left[\mathbb{E}_\nu\left[\int_s^t \mathcal{L}f(X_u)\, du \,|\mathcal{F}_s\right]\mathbb{1}_A\right] &= \mathbb{E}_\nu\left[\mathbb{E}_\nu\left[\theta_{\xi_s} \circ \int_0^{t-s} \mathcal{L}f(X_{u+s})\, du \,|\mathcal{F}_s\right]\mathbb{1}_A\right] \\
&= \mathbb{E}_\nu\left[\mathbb{E}_{\xi_s}\left[\int_0^{t-s} \mathcal{L}f\big(X_{u+s}\big)\, du\right]\mathbb{1}_A\right] \\
&= \mathbb{E}_\nu[\mathbb{E}_{\xi_s}[f(X_t) \;-\; f(X_s)]\mathbb{1}_A] \\
&= \mathbb{E}_\nu[\mathbb{E}_\nu[f(X_t) \;-\; f(X_s)]\mathbb{1}_A]
\end{aligned}$$

Since $f(X_0), f(X_s), \int_0^s \mathcal{L}f(X_u)du \in \mathcal{F}_s$, we conclude that M_t is a $\mathbb{P}_\nu$-martingale. □

Remark 1.5. The process M_t in the Proposition 1.4 is known as the *Dynkin's martingale*. The Dynkin's martingale has an analogous statement for the discrete case in which the process is given by

$$M_t \;=\; f(X_t) \;-\; \sum_{s=0}^{t-1}(\mathcal{L}f)(X_s)$$

and its generator is $(\mathcal{L}f)(X_t) \;:=\; \mathbb{E}[f(X_{t+1}) \;-\; f(X_t)\,|\,\mathcal{F}_t]$, $\mathbb{P}$-almost surely.

Lemma 1.4. *Let $\{X_t : t \;\geq\; 0\}$ be a stochastic process and*

$$M_t(X) \;=\; f(X_t) \;-\; f(X_0) \;-\; \int_0^t \mathcal{L}f(X_s)ds$$

be a martingale with respect to the natural σ-algebra $\mathcal{F}_t$ where $f \;\in\; C_b(S) \cap \mathcal{D}(\mathcal{L})$. Define

$$M_{t,t+h}(X) \;=\; f(X_{t+h}) \;-\; f(X_t) \;-\; \int_t^{t+h} \mathcal{L}f(X_s)ds.$$

Then, $\mathbb{E}\Big[M_{t,t+h}(X)|\mathcal{F}_t\Big] \;=\; 0$.

Proof. It is enough observe that $M_{t,t+h}(X) \;=\; M_{t+h}(X) \;-\; M_t(X)$. Thus

$$\mathbb{E}\Big[M_{t,t+h}(X)|\mathcal{F}_t\Big] \;=\; \mathbb{E}\Big[M_{t+h}(X) \;-\; M_t(X)|\mathcal{F}_t\Big] \;=\; \mathbb{E}\Big[M_{t+h}(X)|\mathcal{F}_t\Big] \;-\; M_t(X) \;=\; 0\,.$$

□

1.5 Skorohod space

1.5.1 The Space D

Let $C(\mathbb{R}^+, E)$ the space of continuous functions with the uniform topology, where (E, ρ) is a metric space. This space is not enough to describe certain classes of processes when they contain jumps, like a Poisson process for example.

For this reason, we need more suitable spaces to deal with such processes. Consider the metric q in the space E given by $q(x, y) = \min\{\rho(x, y), 1\}$. We will denote $D(\mathbb{R}^+, E)$ to be the space of the functions $x : [0, \infty) \to E$ that are right-continuous and have left-side limits.

Lemma 1.5. *If $x \in D(\mathbb{R}^+, E)$, then x have at most countable many points of discontinuity.*

Proof. Define $A_n = \{t > 0 : \rho(x(t), x(t-)) > \frac{1}{n}\}$. Since the x have the left-hand limits, the set A_n shall not have limit points, that is, A_n is countable. Thus, follows the result because $D(\mathbb{R}^+, E) = \bigcup_{n\in\mathbb{N}} A_n$. □

Now, we will equip the space with a metric such that will make our space well behaved. We will say that two functions $x, y \in D(\mathbb{R}^+, E)$ are near in the space D, if $x(t)$ can be carried onto the graph of the $y(t)$ by uniformly small perturbation in the range and a uniformly small perturbation on the scale of time, formally: Let Λ' the space of (strictly) increasing continuous function $\lambda : [0, \infty) \to [0, \infty)$ such that $\lambda(0) = 0$ and $\lambda(\infty) = \infty$ (this space is called reparametrization space). Let $\gamma : \Lambda' \to \mathbb{R}$ defined as

$$\gamma(\lambda) := \operatorname*{ess\,sup}_{t\geq 0} \left|\log \lambda'(t)\right| = \sup_{s>t\geq 0} \left|\log \frac{\lambda(s) - \lambda(t)}{s - t}\right|.$$

Recalling that the essential supremum of f is the smallest number $a \in \mathbb{R}$ for which f only exceeds a on a set of measure zero, i.e., if $U_f^a := \{x : f(x) \geq a\}$ for each $a \in \mathbb{R}$, define $U_0 := \{a : m(U_f^a) = 0\}$ where m is the Lebesgue measure. Then $\operatorname{ess\,sup}(f) = \inf U_0$.

Define $\Lambda := \{\lambda \in \Lambda' : \lambda \text{ is Lipschitz continuous and } \gamma(\lambda) < \infty\}$ and for $x, y \in D(\mathbb{R}^+, E)$, let

$$d(x, y) = \inf_{\lambda\in\Lambda} \left[\gamma(\lambda) \vee \int_0^\infty \exp\{-u\} d(x, y, \lambda, u)\, du\right], \tag{1.4}$$

where $d(\cdot, \cdot, \cdot, \cdot)$ represents the supremum of the distance until fixed time u for given λ, i.e.,

$$\begin{aligned} d(x, y, \lambda, u) &= \sup_{t\geq 0} q(x(t \wedge u), y(\lambda(t) \wedge u) \\ &= \sup_{t\geq 0}\{\min\{\rho(x(t \wedge u), y(\lambda(t) \wedge u)), 1\}\}. \end{aligned}$$

Note that it is adopted the exponential weight $\exp(-u)$ and the minimum with 1 in the definition of q in order to make the integral finite.

Example 1.3. Let $f(t) = \exp(\exp(-t))$ and $g(t) = 0$, see Figure 1.2. Since

$$d(f, g) = \inf_{\lambda\in\Lambda} \left[\gamma(\lambda) \vee \int_0^\infty \exp\{-u\} d(f, g, \lambda, u)\, du\right],$$

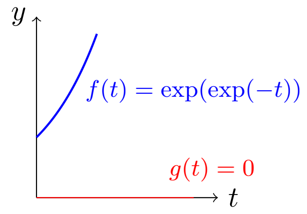


Figure 1.2: Illustration that d is a metric in the sense that d is finite.

we can take the λ to be the identity and so $\gamma(\lambda) = 0$. Thus, since $f(t) > e$ for all t, it follows that $d(f, g, \lambda, u) = 1$, hence $d(f, g) = 1$.

Remark 1.6. An equivalent notion of convergence between sequences in this space is the following: Consider $\{x_n\}, \{y_n\}$ sequences of $D(\mathbb{R}^+, E)$. Thus $\lim_{n\to\infty} d(x_n, y_n) = 0$ if and only if there exists $\{\lambda_n\}$ in Λ such that $\lim_{n\to\infty} \gamma(\lambda_n) = 0$, and for all $\varepsilon > 0$ and $u_0 > 0$

$$\lim_{n\to\infty} m(\{u \in [0, u_0] : d(x_n, y_n, \lambda_n, u) > \varepsilon\}) = 0.$$

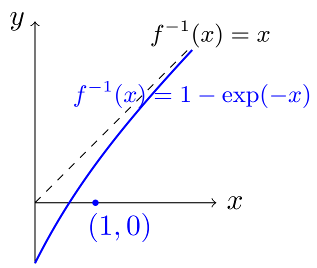
We claim that

$$\operatorname*{ess\,sup}_{t\geq 0} \left|\lambda'(t) - 1\right| \leq 1 - \exp\left(-\gamma(t)\right) \leq \gamma(\lambda), \tag{1.5}$$

for each $\lambda \in \Lambda$. Consider $f(x) = 1 - \exp(-x)$ which is an increasing function. Hence

$$\gamma(f(\lambda)) \leq f(\gamma(\lambda)).$$

Moreover $f(x) \leq x$ for all x.



Therefore, (1.5) is valid. Thus, if we have that $\gamma(\lambda_n) \to 0$, by (1.5) for each $T > 0$ it follows that

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} |\lambda_n(t) - t| = 0.$$

Proposition 1.5. *The function d as defined in (1.4) is a metric.*

Remark 1.7. The metric d induces the so-called J_1-Skorohod topology on $D(\mathbb{R}^+, E)$.

Proof. Let $x, y \in D(\mathbb{R}^+, E)$ such that $d(x, y) = 0$. So, as we have that x, y are right continuous function and $D(\mathbb{R}^+, E)$ has at most countable many points of discontinuity, $x(t) = y(t)$ for every t continuity point of y and follows $x = y$.

Since λ is strictly increasing and is defined on $[0, \infty)$, it follows that

$$\sup q(x(t \wedge u), y(\lambda(t) \wedge u)) \; = \; \sup q(x(\lambda^{-1}(t) \wedge u), y(t \wedge u)),$$

for each $\lambda \in \Lambda$ and $u > 0$ hence $d(x, y, \lambda, u) \; = \; d(x, y, \lambda^{-1}, u)$ and this implies that $d(x, y) \; = \; d(y, x)$.

It remains to show the triangular inequality. Let $x, y, z \in D(\mathbb{R}^+, E)$ and $\lambda_1, \lambda_2 \in \Lambda$,

$$\begin{aligned} \sup_{t \geq 0} q(x(t \wedge u), z(\lambda_2 \circ \lambda_1(t) \wedge u)) \; \leq \; & \sup_{t \geq 0} q(x(t \wedge u), y(\lambda_1(t) \wedge u)) \\ & + \sup_{t \geq 0} q(y(\lambda_1(t) \wedge u), z(\lambda_2 \circ \lambda_1(t) \wedge u)) \\ & \leq \sup_{t \geq 0} q(x(t \wedge u), y(\lambda_1(t) \wedge u)) \\ & + \sup_{t \geq 0} q(y(s \wedge u), z(\lambda_2(s) \wedge u)) \end{aligned} \tag{1.6}$$

that is, $d(x, y, \lambda_2 \circ \lambda_1, u) \; \leq \; d(x, y, \lambda_1, u) + d(x, y, \lambda_2, u)$. To assume the inequality $d(x, z) \; \leq \; d(x, y) \; + \; d(y, z)$, its enough to show that

$$\gamma(\lambda_2 \circ \lambda_1) \leq \gamma(\lambda_1) \; + \; \gamma(\lambda_2).$$

Since $\lambda_2 \circ \lambda_1 \in \Lambda$, applying the chain's rule gives us

$$\begin{aligned} \gamma(\lambda_2 \circ \lambda_1) \; &= \; \operatorname*{ess\,sup}_{t \geq 0} \left| \log(\lambda_2 \circ \lambda_1)'(t) \right| \\ &= \; \operatorname*{ess\,sup}_{t \geq 0} \left| \log \lambda_2'(\lambda_1(t))) \lambda_1'(t) \right| \\ &= \operatorname*{ess\,sup}_{t \geq 0} \left| \log \lambda_2'(\lambda_1(t))) \; + \; \log \lambda_1'(t) \right| \\ &\leq \operatorname*{ess\,sup}_{t \geq 0} \left| \log \lambda_2'(\lambda_1(t))) \right| + \; \operatorname*{ess\,sup}_{t \geq 0} \left| \log \lambda_1'(t) \right| \\ &= \; \gamma(\lambda_1) \; + \; \gamma(\lambda_2). \end{aligned}$$

Thus, d is a metric. □

The next example illustrates the difference between the uniform and the Skorohod metric.

Example 1.4. Consider the following functions,

$$x(t) \; = \; \begin{cases} 0, & \text{if } t \in [0, \frac{1}{2}), \\ 1, & \text{if } t \geq \frac{1}{2} \end{cases}$$

and

$$y(t) \; = \; \begin{cases} 0, & \text{if } t \in [0, \frac{1}{2} \; + \; \frac{\varepsilon}{2}), \\ 1, & \text{if } t \geq \frac{1}{2} \; + \; \frac{\varepsilon}{2}. \end{cases}$$

Under the uniform metric, we have that

$$\|x - y\|_\infty \; = \; \sup_{t \geq 0} |x(t) \; - \; y(t)| \; = \; 1.$$

On the other hand, defining a perturbation $\lambda \in \Lambda$ by

$$\lambda(t) = \begin{cases} (1+\varepsilon)t, \text{ if } t \in [0, \frac{1}{2}), \\ (1-\varepsilon)t + \varepsilon, \text{ if } t \geq \frac{1}{2}, \end{cases}$$

which is the candidate to minimize the distance between x and y.

We have two cases. Consider $t \in [0, \frac{1}{2})$, then $\lambda(t) = (1+\varepsilon)t \leq \frac{1}{2} + \frac{\varepsilon}{2}$. This implies that $x(t) = 0 = y(\lambda(t))$ for $t \in [0, \frac{1}{2})$, hence $d(x, y, \lambda, u) = 0$. Moreover $\gamma(\lambda) = |1+\epsilon|$ and consequently

$$d(x(t), y(\lambda(t))) = \inf_{\lambda \in \Lambda} \left[\gamma(\lambda) \vee \int_0^\infty \exp\{-u\} d(x, y, \lambda, u)\, du \right] \leq |\log(1+\varepsilon)|\,.$$

Now, for $t \geq \frac{1}{2}$, $\lambda(t) = (1-\varepsilon)t + \varepsilon \geq \frac{1}{2} + \frac{\varepsilon}{2}$. Hence, $x(t) = 1$ and $y(\lambda(t)) = 1$ for $t \geq \frac{1}{2}$, then $d(x, y, \lambda, u) = 0$. As the same way for the firs case, it follows that $d(x(t), y(\lambda(t))) \leq |\log(1-\varepsilon)|$.

Proposition 1.6. *If $\{x_n\}$ a sequence in $D(\mathbb{R}^+, E)$ and $x \in D(\mathbb{R}^+, E)$. the following statements are equivalent:*

(i) $\lim_{n\to\infty} d(x_n, x) = 0\,.$

(ii) There exists a sequence $\{\lambda_n\}$ in Λ such that $\gamma(\lambda_n) \to 0$, and

$$\lim_{n\to\infty} d(x_n, x, \lambda_n, u) = 0\,,$$

for all $u \in [0, \infty)$.

(iii) There exists a sequence $\{\lambda_n\}$ in Λ such that $\gamma(\lambda_n) \to 0$, and

$$\lim_{n\to\infty} d(x_n, x, \lambda_n, u) = 0,$$

for any continuity point u of x.

Proof. (*i*) $\Rightarrow$ *(ii)*

Let $u \in [0, \infty)$. By definition of convergence there exist sequences $\{\lambda_n\} \subset \Lambda$ and $\{u_n\} \subset [u, \infty)$ with $u_n \downarrow u$ such that for all $\varepsilon > 0$

$$\lim_{n\to\infty} m\{u \in [0, u_n] : d(x_n, x, \lambda_n, u) > \varepsilon\} = 0,$$

thus

$$\lim_{n\to\infty} d(x_n, x, \lambda_n, u_n) = 0, \tag{1.7}$$

m-almost everywhere in $t \wedge u_n$. By the triangular inequality

$$d(x_n, x, \lambda_n, u) \leq \sup_{t\geq 0} q(x_n(t\wedge u), x(\lambda_n(t\wedge u)\wedge u_n)) + \sup_{t\geq 0} q(x(\lambda_n(t\wedge u)\wedge u_n), x(\lambda_n(t)\wedge u)). \tag{1.8}$$

Note that the first term of (1.8) go to zero by (1.7). Now, observe that

$$\begin{aligned}\sup_{t\geq 0} q(x(\lambda_n(t\wedge u)\wedge u_n), x(\lambda_n(t)\wedge u)) &= \sup_{t\leq u} q(x(\lambda_n(t)\wedge u_n), x(\lambda_n(t)\wedge u)) \\ &+ \sup_{t>u} q(x(\lambda_n(u)\wedge u_n), x(\lambda_n(t)\wedge u)).\end{aligned}$$

By the monotonicity of each function λ_n, if $t \leq u$ then $\lambda_n(t) \leq \lambda_n(u)$ for all n, and similarly if $t > u$ it follows that $\lambda_n(t) > \lambda_n(u)$. The first term clearly tends to zero. In fact, it is enough to perceive that $\lambda_n(t)\wedge u \leq \lambda_n(t)\wedge u_n$ and $u_n \downarrow u$.

Since $d(x_n, x) \to 0$, by the notion of convergence in Skorohod space $\gamma(\lambda_n) \to 0$ and it follows that

$$|\lambda_n(u) - u| \leq \sup_{0\leq t\leq u_n} |\lambda_n(t) - t| \to 0\,.$$

The second term $\lambda_n(t)$ is smaller than u only for finite times because $t > u$. Then

$$q(x(\lambda_n(u)\wedge u_n), x(\lambda_n(t)\wedge u)) \to 0.$$

(ii) $\Rightarrow$ *(iii)* By the Lemma 1.5, the set of discontinuity points have zero Lebesgue measure.
(iii) $\Rightarrow$ *(i)* Since $d(x_n, x, \lambda_n, u) \leq 1$, $|\exp(-u)d(x_n, x, \lambda_n, u)| \leq \exp(-u)$ and $\exp(-u)$ is integrable, by the Dominated Convergence Theorem we can conclude. □

The previous proposition tell us that $\lim_{n\to\infty} d(x_n, x) = 0$ implies that

$$\lim_{n\to\infty} x_n(u) = \lim_{n\to\infty} x_n(u-) = x(u)$$

for all continuity points u of x.

Proposition 1.7. *Let $\{x_n\}$ and x elements of $D(\mathbb{R}^+, E)$. The following statements are equivalent:*

(i) $\lim_{n\to\infty} d(x_n, x) = 0$,

(ii) There exists $\{\lambda_n\} \subset \Lambda$ such that

$$\lim_{n\to\infty} \gamma(\lambda_n) = 0\,,$$

and

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} \rho(x_n(t), x(\lambda_n(t))) = 0,$$

for all $T > 0$.

(iii) For each $T > 0$, there exists $\{\lambda_n\} \subset \Lambda'$ such that

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} |\lambda_n(t) - t| = 0\,,$$

and

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} \rho(x_n(t), x(\lambda_n(t))) = 0\,.$$

Proof. Keep in mind that since $q(x,y) = \rho(x,y) \wedge 1$, if $\rho(x_n, x) \to 0$ we also have that $q(x_n, x) \to 0$.

(i) ⇒ *(ii)* Suppose that $\lim_{n\to\infty} d(x_n, x) = 0$. By the Proposition 1.6 there exists $\{\lambda_n\} \subset \Lambda$ and $u_n \in (0,\infty)$, $\{u_n\}$ sequence of continuity points of x with $u_n \to \infty$ such that $\gamma(\lambda_n) \to 0$, and $d(x_n, x, \lambda_n, u_n) \to 0$ when $n \to \infty$. Then, we can find $u_n \geq T := T(u_n) > 0$ such that $\lambda_n(T) \leq \lambda_n(u_n) \vee u_n$. Thus

$$\sup_{0\leq t\leq T} \rho(x_n(t), x(\lambda_n(t))) \leq \sup_{t\geq 0} \rho(x(t \wedge u_n), x(\lambda_n(t) \wedge u_n))$$

which goes to zero.

(ii) ⇒ *(i)* Let $\{\lambda_n\} \subset \Lambda$ satisfying the hypothesis. Fix any $T > 0$ and take $\{u_n\} \subset [0,T]$ such that $u_n \leq \lambda_n(T \wedge t) \vee T$. Thus

$$\sup_{t\geq 0} \rho(x_n(t \wedge u_n), x(\lambda_n(t) \wedge u_n) \leq \sup_{0\leq t\leq T} \rho(x_n(t), x(\lambda_n(t))).$$

Since the right side of above goes to zero for all $T > 0$, we can take any continuity point u of x and $\{u_n\} \subset [u,\infty)$ as in Proposition 1.6 and therefore we can conclude.

(ii) ⇒ *(iii)* We have already seen that $\gamma(\lambda_n) \to 0$ imply $|\lambda_n(t) - t| \to 0$ and that is enough to conclude because $\Lambda \subset \Lambda'$.

(iii) ⇒ *(ii)* The idea is to construct a family of polygonal paths $\{\lambda'_n\}$ from $\{\lambda_n\} \subset \Lambda'$ where λ'_n in Λ for all n such that λ'_n approaching to λ_n in a certain sense and their derivative is close to one.

Given $N \in \mathbb{N}$, choose $\{\lambda_n^N\} \subset \Lambda'$ that satisfies the item *(iii)* with $T = N$ such that for all $t \geq N$,

$$\lambda_n^N(t) := \lambda_n^N(N) + t - N,$$

for all $n \in \mathbb{N}$. Now, define $\tau_0 = 0$, and for each integer k,

$$\tau_k^N := \begin{cases} \inf\{t > \tau_{k-1}^N : \rho(x(t), x(\tau_{k-1}^N)) > \frac{1}{N}\}, & \text{if } \tau_{k-1}^N < \infty, \\ \infty, & \text{if } \tau_{k-1}^N = \infty. \end{cases}$$

Note that $\{\tau_k^N\}_k$ is a collection of times at which $x(\tau_k^N)$ are at least $\frac{1}{N}$ away from $x(\tau_{k-1}^N)$. The sequence $\{\tau_k^N\}_k$ is strictly increasing and since the left limits exist, τ_k^N goes to infinity as $t \to \infty$.

Now, we will create a sequence of strictly increasing Lipschitz functions which will be our candidate to be λ'_n. To do so, for each integer n, define the following sequence of times

$$u_{k,n}^N := (\lambda_n^N)^{-1}(\tau_k^N),$$

with $k \in \mathbb{N}$, where $(\lambda_n^N)^{-1}(\infty) = \infty$. Note that $\{u_{k,n}^N\}_n$ is a collection of increasing times that converges to τ_k^N. In fact, by the hypothesis *(iii)* we have that

$$\left|\tau_k^N - u_{k,n}^N\right| = \left|\lambda_n^N(u_{k,n}^N) - u_{k,n}^N\right| \leq \sup_{0\leq t\leq T} \left|\lambda_n^N(t) - t\right| \to 0. \tag{1.9}$$

For each k integer, let

$$\mu_n^N(t) := \begin{cases} \tau_k^N + (t - u_{k,n}^N)(u_{k+1,n}^N - u_{k,n}^N)^{-1}(\tau_{k+1}^N - \tau_k^N), & \text{if } t \in [u_{k,n}^N, u_{k+1,n}^N) \cap [0,N] \\ \mu_n^N(N) + t - N, & \text{if } t > N. \end{cases}$$

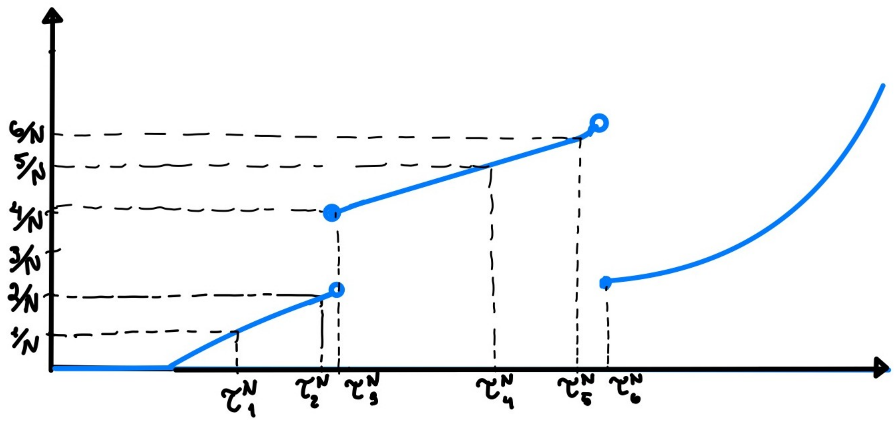


Figure 1.3: Illustration of the family $\{\tau_k^N\}$

and, by convention, define $\infty^{-1}\infty = 1$. The function μ_n^N as defined above is a polygonal path which is a strictly increasing Lipschitz function which carries $u_{k,n}^N$ to τ_k^N.

We claim that $\mu_n^N \in \Lambda$. First, observe that $\mu_n^N \in \Lambda'$. The continuity of $\mu_n^N(t)$ is guaranteed by the continuity of λ_n^N and the strictly increasing monotonicity property, because $t \in [u_{k,n}^N, u_{k+1,n}^N)$. Hence $\mu_n^N \in \Lambda$. It remains to show that μ_n^N will be Lipschitz and $\gamma(\mu_n^N) < \infty$ to ensure that $\mu_n^N \in \Lambda'$. Its enough to realize that

$$\frac{d}{dt}\mu_n^N = (u_{k+1,n}^N - u_{k,n}^N)^{-1}(\tau_{k+1}^N - \tau_k^N) < \infty.$$

Therefore, for each integer N and n we have that $\mu_n^N \in \Lambda$.

Now, by equation (1.9), when n goes to infinity it follows that $(\mu_n^N)' \to 1$ and therefore, $\gamma(\mu_n^N) \to 0$. Thus, for each $T = N$, put $\lambda_n' = \mu_n^N$ and observe that

$$\sup_{0\le t\le T} \rho(x_n(t), x(\mu_n^N(t))) \le \sup_{0\le t\le T} \rho(x_n(t), x(\lambda_n^N(t))) + \sup_{0\le t\le T} \rho(x(\lambda_n^N(t)), x(\mu_n^N(t))),$$

and

$$\sup_{0\le t\le T} \rho(x(\lambda_n^N(t)), x(\mu_n^N(t))) = \max_{0\le i\le n} \sup_{u_{i,n}^N\le t<u_{i+1,n}^N} \rho(x(\lambda_n^N(t)), x(\mu_n^N(t))) \le \frac{1}{N}.$$

By the second statement of hypothesis *(iii)*, we can conclude this implication which proves the proposition.

□

Remark 1.8. In items *(ii)* and *(iii)* of the proposition above in order to simplify some results, we can replace

$$\lim_{n\to\infty} \sup_{0\le t\le T} \rho(x_n(t), x(\lambda_n(t))) = 0,$$

by

$$\lim_{n\to\infty} \sup_{0\le t\le T} \rho(x_n(\lambda_n(t)), x(t)) = 0.$$

We can do that because the inverse of each λ_n inherit all properties of λ_n and therefore, λ_n^{-1} belongs to the same space that λ_n.

Theorem 1.4. *If E is separable, then $D(\mathbb{R}^+, E)$ is separable. If the metric space (E, ρ) is complete, then $D(\mathbb{R}^+, E)$ is complete.*

Proof. Let E separable and consider $x \in D(\mathbb{R}^+, E)$ arbitrary. Let $T > 0$ and set $\tau_0 = 0$ and for each k integer, define

$$\tau_k^T = \begin{cases} \inf\{t > \tau_{k-1}^T : \rho(x(t), x(\tau_{k-1}^T)) > \frac{1}{T}\}, & \text{if } \tau_{k-1}^T < \infty, \\ \infty, & \text{if } \tau_{k-1}^T = \infty. \end{cases}$$

Now, define the operator $J^T : D(\mathbb{R}^+, E) \times [0, \infty) \mapsto E$ by

$$J^T(x)(t) = \begin{cases} x(\tau_{i-1}^T), & \text{if } t \in [\tau_{i-1}^T, \tau_i^T) \cap [0, T] \\ x(T), & \text{if } t \ge T. \end{cases}$$

and observe that $\rho(x(t), J^T(x)(t)) < \frac{1}{T}$ for each $t \in [0, T]$ uniformly. Let $\{a_i\}_{i\in\mathbb{N}} \subset E$ to be the dense subset. Thus, for any $T > 0$ considering the family $\{\tau_i^T\}_{i\in\mathbb{N}}$ as defined, we can find $a_i \in B_\rho(x(\tau_{i-1}^T), \frac{1}{T})$.

Thus, we can approach the operator J^T by functions that assume values in a dense subset of E, and therefore, this function only have a countable distinct values. Define

$$y^T(t) = \begin{cases} a_i, & \text{if } t \in [\tau_{i-1}^T, \tau_i^T) \cap [0, T] \\ x(T), & \text{if } t \ge T, \end{cases}$$

and note that $\rho(x(t), y^T(t)) < \frac{1}{T}$. Observe that to get a dense subset of $D(\mathbb{R}^+, E)$, it is enough to show that we can approach any function that has a jump in irrational times by countable sequence of functions as defined as above.

Let $x \in D(\mathbb{R}^+, E)$ be a function that have at least one jump in an irrational time, let's say $t_i^* \in [\tau_{i-1}^T, \tau_i^T) \cap [0, T]$. Thus, let $\{q_n^i\}_{n\in\mathbb{N}} \subset \mathbb{Q}$ such that $q_n^i \to t_i^*$ as $n \to \infty$. Now, we can construct a sequence of $\{\lambda_n^i\}_{n\in\mathbb{N}} \subset \Lambda$ as in example 1.4. Since the set of discontinuity points is countable by Lemma 1.5 and making a linear interpolation of $\{\lambda_n^i\}_{i\in\mathbb{N}}$ then we conclude that $D(\mathbb{R}^+, E)$ is separable.

Let $\{x_n\} \subset D(\mathbb{R}^+, E)$ a Cauchy sequence. Thus, there exists $n_k \in \mathbb{N}$ large enough such that for all $m, n \ge n_k$,

$$d(x_n, x_m) \le \frac{1}{2^{k+1}} \exp(-k), \text{ for each } k \in \mathbb{N}.$$

Define $y_k = x_{n_k}$ to be a subsequence of $\{x_n\}$. Taking $u_k > k$ and $\{\lambda_k\} \subset \Lambda$ so that

$$\max\{\gamma(\lambda_k), d(x_n, x_m, \lambda_k, u_k)\} \le \frac{1}{2^k}$$

and that is true because $\{y_k\}$ is a subsequence of a Cauchy sequence in the Skorohod metric. Let

$$\mu_k := \lim_{n\to\infty} \lambda_{k+n} \circ \cdots \circ \lambda_k$$

We claim that μ_k exists uniformly on bounded intervals. Consider the interval $[0,T]$ for any $T > 0$ and consider the partial composition $\mu_k^n = \lambda_{k+n} \circ \cdots \circ \lambda_k$ for each integer n.

Let $m, n \in \mathbb{N}$ such that $k < n < m$, then

$$\sup_{0\le t\le T} |\mu_k^n - \mu_k^m| \le \gamma(\mu_k^n) + \gamma(\mu_k^m) < \frac{1}{2^{n-1}} + \frac{1}{2^{m-1}}$$

Therefore, $\{\mu_k^n\}$ is a Cauchy sequence and converge uniformly in $[0,T]$ as $n \to \infty$. Moreover since each $\lambda_i \in \Lambda$ for each i integer, we have that $\mu_k \in \Lambda$ for each k. Indeed

$$\begin{aligned}\gamma(\mu_k) &= \operatorname{ess\,sup} \left|\log \frac{d}{dt}\mu_k(t)\right| = \operatorname{ess\,sup} \left|\log \frac{d}{dt}(\lim_{n\to\infty} \lambda_{k+n} \circ \cdots \circ \lambda_k)\right| \\ &\le \sum_{j=k}^{\infty} \operatorname{ess\,sup} \left|\log \frac{d}{dt}\lambda_j(t)\right| \le \frac{1}{2^{k+1}},\end{aligned}$$

Moreover $\gamma(\mu_k)$ is finite and μ_k is composition of strictly increasing Lipschitz functions.

We claim that $\mu_{k+1}^{-1} = \lambda_k \circ \mu_k^{-1}$. Indeed

$$\begin{aligned}\lambda_k \circ \mu_k^{-1} &= \lambda_k \circ \big(\lim_{n\to\infty} \lambda_{k+n} \circ \ldots \circ \lambda_k\big)^{-1} \\ &= \lambda_k \circ \big(\lim_{n\to\infty} \lambda_k^{-1} \circ \ldots \circ \lambda_{(n-1)+(k+1)}^{-1}\big) \\ &= \lim_{n\to\infty} \lambda_k \circ \lambda_k^{-1} \circ \ldots \circ \lambda_{(n-1)+(k+1)}^{-1} \\ &= \big(\lim_{n\to\infty} \lambda_{(k+1)+(n-1)} \circ \ldots \circ \lambda_{k+1}\big)^{-1} = \mu_{k+1}^{-1}.\end{aligned}$$

Thus

$$\begin{aligned}\sup_{t\ge 0} q(y_k(\mu_k^{-1}(t) \wedge u_k), y_{k+1}(\mu_{k+1}^{-1}(t))) &= \sup_{t\ge 0} q(y_k(\mu_k^{-1}(t) \wedge u_k), y_{k+1}(\lambda_k \circ \mu_k^{-1}(t) \wedge u_k)) \\ &= \sup_{t\ge 0} q(y_k(s \wedge u_k), y_{k+1}(\lambda_k(s) \wedge u_k)) \\ &= d(y_k, y_{k+1}, \lambda_k, u_k) \le \frac{1}{2^k},\end{aligned}$$

for each k. From the completeness of (E, r), for each $T > 0$ set $z_k := y_k \circ \mu_k^{-1}$ and observe that $\{z_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence and then $\{z_k(t)\} \subset E$ is a Cauchy sequence. Thus for each $t_0 \in [0,T]$, $z_k(t_0)$ converges to some $a_0 \in E$ as $k \to \infty$. Define $y : [0,\infty) \to E$ to be $y(t) = \lim_{k\to\infty} z_k(t)$. Since $\{z_k\}_{k\in\mathbb{N}}$ by [19, Corollary 1, page 186], z_k converges uniformly to y on $[0,T]$. Observe that for each k we have that $z_k \in D(\mathbb{R}^+, E)$ and therefore $y \in D(\mathbb{R}^+, E)$.

To conclude it is enough to observe that $\gamma(\mu_k^{-1}) \to 0$ as $n \to \infty$,

$$\lim_{n'\to\infty} \sup_{0\le t\le T} \rho(y_k(\mu_k^{-1}(t)), y(t)) = 0$$

and by the the equivalence *(ii)*, *(i)* of the Proposition 1.7 follows that $D(\mathbb{R}^+, E)$ is complete. □

1.5.2 Compact sets in Skorohod space

Our next goal will be characterize compact sets in the space $D(\mathbb{R}^+, E)$ aiming to achieve tightness criterion for families of probability measures in the Skorohod space.

Consider $x \in D(\mathbb{R}^+, E)$. Define $s_0 = 0$ and for each $k \in \mathbb{N}$,

$$s_k(x) = \begin{cases} \inf\{t > s_{k-1}(x) : x(t) \neq x(t-)\}, & \text{if } s_{k-1}(x) \neq \infty, \\ \infty, & \text{if } s_{k-1}(x) = \infty, \end{cases}$$

the points where x is not continuous.

Let $\Gamma \subset E$ a compact set and given $\delta > 0$, define $A(\Gamma, \delta)$ to be the set of all step functions $x \in D(\mathbb{R}^+, E)$ with values $x(t)$ on Γ such that two consecutive jumps are at least a distance δ, that is

$$A(\Gamma, \delta) := \{x \in D(\mathbb{R}^+, E) : x([0,\infty)) \subset \Gamma, x \text{ is a step function and } s_k(x) - s_{k-1}(x) > \delta \text{ for each } k\}$$

Lemma 1.6. *The set $A(\Gamma, \delta)$ is relatively compact.*

Proof. Consider a sequence $\{x_n\} \subset D(\mathbb{R}^+, E)$ then either $s_k(x_{n'}) = \infty$ or $s_k(x_{n'}) < \infty$. For each $k \in \mathbb{N}$, by a diagonalization argument, we can extract a subsequence $\{x_{n'}\}_{n'} \subset \{x_n\}_n$ such that, $s_k(x_{n'})_{n'}$ is a convergent sequence.

If $s_k(x_{n'}) < \infty$ for each $n' \in \mathbb{N}$, it follows that $\lim_{n' \to \infty} s_k(x_{n'}) = t_k$ exists and since the set Γ is compact,

$$\lim_{n' \to \infty} x_{n'}(t_k) = a_k \in \Gamma$$

for each $n' \in \mathbb{N}$. Hence $\lim_{n' \to \infty} x_{n'} \in D_E[0,\infty]$ taking $\{\lambda_n\}_{n \in \mathbb{N}} \subset \Lambda$ as in example 1.4.

Let k the first moment that $s_k(x_{n'}) = \infty$. Observe that the $\{x_{n'}\} \subset A(\Gamma, \delta)$ and therefore $s_i(x_{n'}) - s_{i-1}(x_{n'}) > \delta$ for all $i \in \{1, \dots, k-1\}$ hence we are in the previous cases.

Now, define $y(t) = a_k$ whenever that $t \in [t_k, t_{k+'})$. The sequence $\{x_{n'}\}_{n' \in \mathbb{N}}$ converges uniformly to y. Then, $A(\Gamma, \delta)$ is relatively compact. $\square$

To establish the compactness criterion for subsets of $D(\mathbb{R}^+, E)$, let us define the modulus of continuity of a functions. For $x \in D(\mathbb{R}^+, E)$, $\delta > 0$ and $T > 0$, define

$$w'(x, \delta, T) := \inf_{\{t_i\}} \max_{1 \leq i \leq n} \sup_{s,t \in [t_{i-1}, t_i)} \rho(x(s), x(t)),$$

where $\{t_i\}_{i=1}^n$ is the form $0 = t_0 < \ldots < t_{n-1} < T \leq t_n$ and for any $n \geq 1$ we have that $\min_{1 \leq i \leq n} (t_i - t_{i-1}) > \delta$.

Note that, if $\delta_1 > \delta_2$, it follows that $\min_{1 \leq i \leq n} (t_i - t_{i-1}) > \delta_1 > \delta_2$ then

$$w'(x, \delta_1, T) \geq w'(x, \delta_2, T)$$

Similarly if $T_1 \geq T_2$, i.e., w' is non-decreasing in δ and in T.

We claim that,

$$w'(x, \delta, T) \leq w'(y, \delta, T) + 2 \sup_{0 \leq s < T+\delta} \rho(x(s), y(s)). \tag{1.10}$$

Indeed, by the triangular inequality,

$$\begin{aligned}\max_{0\le i\le n}\sup_{s,t\in[t_{i-1},t_i)}\rho(x(s),x(t)) &\le \max_{0\le i\le n}\sup_{s\in[t_{i-1},t_i)}\rho(x(s),y(s))\\ &+\max_{0\le i\le n}\sup_{s,t\in[t_{i-1},t_i)}\rho(y(s),y(t))+\max_{0\le i\le n}\sup_{t\in[t_{i-1},t_i)}\rho(y(t),x(t))\\ &\le \max_{0\le i\le n}\sup_{s,t\in[t_{i-1},t_i)}\rho(y(s),y(t))+2\sup_{s\in[0,T+\delta)}\rho(x(s),y(s))\end{aligned}$$

and since we take the infimum of $\{t_i\}_{i=1}^n$ with the form

$$0 = t_0 < \ldots < t_{n-1} < T \le t_n$$

such that $\min_{1\le i\le n}\{t_i - t_{i-1}\} > \delta$ on the both sides of equation, we can conclude.

Lemma 1.7. *(i) For each $x \in D(\mathbb{R}^+, E)$ and $T > 0$ we have that $w'(x,\delta,T)$ is right-continuous in δ and*

$$\lim_{\delta\downarrow 0} w'(x,\delta,T) = 0.$$

(ii) If $\{x_n\} \subset D(\mathbb{R}^+, E)$, $x \in D(\mathbb{R}^+, E)$ and $\lim_{n\to\infty} d(x_n, x) = 0$, follows that

$$\limsup_{n\to\infty} w'(x_n,\delta,T) \le w'(x,\delta,T+\varepsilon),$$

for all $\delta > 0, T > 0$ and $\varepsilon > 0$.

(iii) For each $\delta > 0$ and $T > 0$, $w'(x,\delta,T)$ is Borel measurable in x.

Proof. (i) To show the right-continuity of the modulus of continuity in δ, for each $x \in D(\mathbb{R}^+, E)$ and $T > 0$, let $\delta > 0$ and consider $w'(x,\delta,T)$. Then, we can find $\delta' > \delta$ such that $\min_{0\le i\le n}\{t_i - t_{i-1}\} > \delta'$ and since the modulus of continuity is non-decreasing in δ It follows that w' is right-continuous in δ.

To show the second statement, define $\tau_k^N := \min\{t > \tau_{k-1}^N : \rho(x(t), x(\tau_{k-1}^N)) > \frac{1}{N}\}$ and consider $0 < \delta < \min_{\tau_k^N < T}\{\tau_k^N - \tau_{k-1}^N\}$. Then

$$w'(x,\delta,T) \le \frac{1}{N}.$$

Thus whenever $N \to \infty$ it follows that $\delta\downarrow 0$ hence, $\lim_{\delta\to 0} w'(x,\delta,T) = 0$.

(ii) Let $\{x_n\} \subset D(\mathbb{R}^+, E)$ and $x \in D(\mathbb{R}^+, E)$ such that $\lim_{n\to\infty} d(x_n, x) = 0$. By the Proposition 1.7 item *(iii)* there exists $\{\lambda_n\} \subset \Lambda'$ such that, for each $T > 0$ we have that

$$\begin{aligned}&\lim_{n\to\infty}\sup_{0\le t\le T}|\lambda_n(t) - t| = 0\\ &\lim_{n\to\infty}\sup_{0\le t\le T}\rho(x_n(t), x(\lambda_n(t))) = 0\end{aligned}\tag{1.11}$$

Observe that, if we make a small perturbation δ in T, (1.11) still valid. Thus, for each $n \in \mathbb{N}$, set $\delta_n = \sup_{0\le t\le T}[\lambda_n(t+\delta) - \lambda_n(t)]$. We claim that $\delta_n \to \delta$. Indeed, since λ_n is non-decreasing and

$$\sup_{0\le t\le T}|\lambda_n(t+\delta) - \lambda_n(t) - ((t+\delta) - t)| \le \sup_{0\le t\le T}|\lambda_n(t+\delta) - (t-\delta)| + \sup_{0\le t\le T}|\lambda_n(t) - t|.$$

By (1.11) the above equation converges to zero. Now, by (1.10) and (1.11)

$$\begin{aligned}\limsup_{n\to\infty} w'(x_n,\delta,T) &\le \limsup_{n\to\infty} w'(x\circ\lambda_n,\delta,T) + \limsup_{0\le t\le T+\delta} \rho(x_n(t),x(\lambda_n(t)))\\ &\le \limsup_{n\to\infty}\lim_{n\to\infty} w'(x,\delta_n\vee\delta,\lambda_n(T)) = w'(x,\delta,T+\varepsilon)\end{aligned}$$

and such $\varepsilon>0$ is given by the condition (1.11).

(iii) By the previous result, we have that

$$\lim_{\varepsilon\downarrow 0} w'(x,\delta,T+\varepsilon) = w'(x,\delta,T+),$$

thus $w'(x,\delta,T)$ is upper semi-continuous function in T for all $x\in D(\mathbb{R}^+,E)$ and therefore, is Borel-measurable. □

Proposition 1.8. *Let (E,ρ) be a complete metric space. If $A\subset D(\mathbb{R}^+,E)$ is compact, then for each $T>0$ there exists a compact set $\Gamma_T\subset E$ such that $x(t)\in\Gamma_T$ for $0\le t\le T$ and for all $x\in A$.*

Proof. Fix $T>0$ and define $\Gamma_T := \{x(t) : x\in A \text{ and } t\in[0,T]\}\subset E$ and let $\{a_n\}_{n\in\mathbb{N}}\subset\Gamma_T$. By definition of Γ_T there exists $\{x_n\}_{n\in\mathbb{N}}\subset A$ such that $a_n = x_n(t_n)$ for some $t_n\in[0,T]$. By the compacity of $[0,T]$, we can extract a subsequence $\{t_{n'}\}_{n'\in\mathbb{N}}\subset\{t_n\}_{n\in\mathbb{N}}$ such that converges to some $t\in[0,T]$. Since A is compact then there exists a convergent subsequence $\{x_{n'}\}_{n'\in\mathbb{N}}\subset\{x_n\}_{n\in\mathbb{N}}$. Observe that this subsequence is a Cauchy sequence with the metric of Skorohod.

Then, given $\varepsilon>0$ there exist $\{\lambda^T_{n'}\}_{n'\in\mathbb{N}}\subset\Lambda$ with $\gamma(\lambda^T_{n'})\to 0$ as $n'\to\infty$, and $n_0\in\mathbb{N}$ such that, for all $m',n'\ge n_0$

$$\rho(x_{n'}\circ\lambda^T_{n'}(t_{n'}),x_{m'}\circ\lambda^T_{m'}(t_{m'})) \le \sup_{0\le s\le T}\rho(x_{n'}\circ\lambda^T_{n'}(t),x_{m'}\circ\lambda^T_{m'}(t)) < \varepsilon,$$

and therefore $\{x_{n'}(t_{n'})\}_{n'\in\mathbb{N}}$ is a Cauchy sequence in Γ_T. Hence since E is complete, for each $t\in[0,T]$ which is a limit of $\{t_{n'}\}_{n'\in\mathbb{N}}$ there exists $a_t\in E$ such that $x_{n'}(t_{n'})\to a_t$ as $n'\to\infty$. Moreover Γ_T is closed subset of E then $a_t\in\Gamma_T$ for each $t\in[0,T]$. So Γ_T is a compact subset of E. □

Theorem 1.5. *Let (E,ρ) be complete. The subset $A\subset D(\mathbb{R}^+,E)$ is relatively compact if and only if*

(i) For each rational $t\ge 0$, there exists a compact $\Gamma_t\subset E$ such that $x(t)\in\Gamma_t$ for all $x\in A$,

(ii) For each $T>0$,

$$\lim_{\delta\downarrow 0}\sup_{x\in A} w'(x,\delta,T) = 0.$$

Proof. Suppose that exists $A\subset D(\mathbb{R}^+,E)$ is relatively compact. Thus, taking $B = \operatorname{cl} A$ then by the Proposition 1.8, the first statement is true.

Since A is relatively compact, given $\{x_n\}_{n\in\mathbb{N}}\subset D(\mathbb{R}^+,E)$ there exists $\{x_{n'}\}_{n'\in\mathbb{N}}\subset\{x_n\}_{n\in\mathbb{N}}$ that convergent for some $x\in D(\mathbb{R}^+,E)$. Suppose there is $\eta>0$ and $T>0$ such that for all $n>0$

$$w'(x_{n'},\frac{1}{n'},T) > \eta. \tag{1.12}$$

By Lemma 1.7 *(i)* and *(ii)*, we have a contradiction with (1.12).

Now, consider a subset $A \subset D(\mathbb{R}^+, E)$ that satisfies the hypothesis *(i)* and *(ii)*.For each $N > 0$, let $0 < \delta_N < 1$ such that

$$\sup_{x \in A} w'(x, \delta_N, N) \leq \frac{1}{N}.$$

Recalling that $\Gamma_T := \{x(t) : t \in [0, T] \,\text{and}\, x \in A\}$. Let m_N big enough such that $\frac{1}{m_N} < \delta_N$ and define

$$\Gamma^{(N)} := \bigcup_{i=1}^{N m_N} \Gamma_{\frac{i}{m_N}}.$$

Set $A_N = A(\Gamma^{(N)}, \delta_N)$ as in Lemma 1.6. Thus given $x \in A$, let l integer such that

$$0 = t_0 < t_1 < \ldots < t_{l-1} < N \leq t_l < N + 1 < t_{l+1}$$

with $\min_{i \leq n}(t_i - t_{i-1}) > \delta_N$ and

$$\max_{1 \leq i \leq l} \sup_{s,t \in [t_{i-1}, t_i)} \rho(x(s), x(t)) \leq \frac{2}{N}. \tag{1.13}$$

Define

$$x'(t) = \sum_{i=1}^{l} x(t_i) \mathbb{1}_{[t_i, t_{i+1})}(t).$$

We claim that $x' \in A_n$. Observe that, for any $k \in \{1, \ldots, l\}$ and for all $t \in [t_k, t_{k+1})$, $x'(t) \in \Gamma_{\frac{i}{m_N}}$ for all $\frac{i}{m_N} \geq t_k$ and it follows that $x'(t) \in \Gamma^{(N)}$, and therefore $x' \in A_n$. By the equation (1.13),

$$\sup_{1 \leq t \leq N} \rho(x'(t), x \circ \lambda^N(t)) = \max_{0 \leq i \leq l} \sup_{s,t \in [t_{i-1}, t_i)} \rho(x \circ \lambda^N(t), x'(t)) \leq \frac{2}{N},$$

and taking $\lambda(t) = t$

$$\begin{aligned} d(x', x) &\leq \int_0^\infty \exp(-u) \sup_{t \geq 0} q(x(\lambda(t) \wedge u), x'(t \wedge u))\, du \\ &\leq \int_0 \exp(-u) \sup_{0 \leq t \leq n} \rho(x(\lambda(t) \wedge u), x'(t \wedge u))\, du + \int_N^\infty \exp(-u)\, du \\ &\leq \int_0^N \exp(-u) \frac{2}{N}\, du + \int_N^\infty \exp(-u)\, du \leq \frac{2}{N} + \exp(-N) \leq \frac{3}{N}. \end{aligned} \tag{1.14}$$

Define $A_N^{3/N} := \{y \in A : d(y, A_N) < \frac{3}{N}\}$. By (1.14) we have that $x \in A_N^{3/N}$ and therefore $A \subset A_N^{3/N}$. Moreover, since N is arbitrary, it follows that $A \subset \bigcap_{N \geq 1} A_N^{3/N}$.

We claim that $\bigcap_{N \geq 1} A_N^{3/N}$ is totally bounded. Indeed by the Lemma 1.6 $\operatorname{cl} A_N$ is relatively compact then for each $x_\theta \in \operatorname{cl} A_N$ consider $B(x_\theta, \frac{3}{N}) = \{y \in D(\mathbb{R}^+, E) : d(x_\theta, y) <$

$\frac{3}{N}\}$. Thus, we can extract a finite sub-covering such that $\mathrm{cl}\, A_N \subset B_d(x_{\theta_1}, \frac{3}{N}) \cup \ldots \cup B(x_{\theta_j}, \frac{3}{N})$. Since $\bigcap_{N\geq 1} A_N^{3/N} \subset A_N^{3/N}$ and $A_N^{3/N} \subset B_d(x_{\theta_1}, \frac{3}{N}) \cup \ldots \cup B(x_{\theta_j}, \frac{3}{N})$ it follows the claim.

Now, since A is contained in a totally bounded set we conclude that A is relatively compact. □

Theorem 1.6. *Let (E,ρ) an arbitrary space, $\{x_n\} \subset D(\mathbb{R}^+, E)$ and $x \in D(\mathbb{R}^+, E)$. Thus, $\lim\limits_{n\to\infty} d(x_n, x) = 0$ if and only if whenever $\{t_n\} \subset [0,\infty)$ converges to $t \geq 0$, the following statements are satisfied*

(i)

$$\lim_{n\to\infty} \rho(x_n(t_n), x(t)) \wedge \rho(x_n(t_n), x(t-)) = 0,$$

(ii) If $\lim\limits_{n\to\infty} \rho(x_n(t_n), x(t)) = 0$, then for any sequence of times $\{s_n\} \subset [0,\infty)$ with $s_n \geq t_n$ and $\lim\limits_{n\to\infty} s_n = t$, we have that

$$\lim_{n\to\infty} \rho(x_n(s_n), x(t)) = 0,$$

(iii) If $\lim\limits_{n\to\infty} \rho(x_n(t_n), x(t-)) = 0$, then for any sequence of times $\{s_n\} \subset [0,\infty)$ with $s_n \leq t_n$ and $\lim\limits_{n\to\infty} s_n = t$, we have that

$$\lim_{n\to\infty} \rho(x_n(s_n), x(t-)) = 0,$$

Remark 1.9. The proposition 1.6 gives us a more intuitive idea for the convergence in Skorohod space, in the sense of emphasizing that we have at most two limit points for the sequence $\{x_n(t_n)\}$, the second condition tell us that the limit is right-continuous and the third condition the existence of left-limits. So, we characterize the limit as an element of $D(\mathbb{R}^+, E)$.

Moreover, this proposition gives us an equivalent result of sequential convergence in the uniform convergence in $C(\mathbb{R}^+, E)$.

Proof. Suppose that $\lim\limits_{n\to\infty} d(x,x) = 0$, and let $\{t_n\} \subset [0,\infty)$ and $t \geq 0$ with $t_n \to t$. Choose $T > 0$ such that $\{t_n\} \subset [0,T]$ and $0 \leq t \leq T$. By the Proposition 1.7 item *(iii)*, there exists $\{\lambda_n\} \subset \Lambda'$ such that

$$\sup_{0\leq t\leq T} |\lambda_n(t) - t| = 0 \tag{1.15}$$

and

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} \rho(x_n(t), x(\lambda_n(t))) = 0 \tag{1.16}$$

By triangular inequality, it follows that

$$\begin{aligned} \rho(x_n(t_n), x(t)) \wedge \rho(x_n(t_n), x(t-)) \leq & \sup_{0\leq t\leq t_n<T} \rho(x_n(t), x(\lambda_n(t))) \\ & + \rho(x \circ \lambda_n(t_n)), x(t) \wedge \rho(x \circ \lambda_n(t_n), x(t-)) \end{aligned} \tag{1.17}$$

Since we have that $t_n \to t$, by (1.15) it follows that $\lim_{n\to\infty} \lambda_n(t_n) = t$ and in addition with (1.16) we have that (1.17) tends to zero and therefore, we conclude *(i)*. Observe that maybe neither $\rho(x \circ \lambda_n(t_n)), x(t)$ nor $\rho(x \circ \lambda_n(t_n), x(t-))$ tends to zero, but the infimum between them, goes to zero.

Let $\{s_n\} \subset [0, \infty)$ such that $t_n \leq s_n \leq T$, for each $n \in \mathbb{N}$ where T is the same as choose previously, and $s_n \to t$ as $n \to \infty$. Observe that

$$\rho(x_n(s_n), x(t)) \leq \sup_{0 \leq u \leq s_n < T} \rho(x_n(u), x_n \circ \lambda_n(u)) + \rho(x_n \circ \lambda_n(s_n), x(t)), \tag{1.18}$$

and

$$\rho(x_n \circ \lambda_n(s_n), x(t)) \leq \sup_{0 < t_n \leq u \leq s_n < T} \rho(x_n \circ \lambda_n(u), x_n(u)) + \rho(x_n(t_n), x(t)). \tag{1.19}$$

By the hypothesis and (1.16) it follows that (1.19) tends to zero whenever $n \to \infty$. Again by the hypothesis and in addition with (1.19) it follows that (1.18) converges to zero as $n \to \infty$ and then we conclude the item *(ii)*. To show the last statement, it's enough to observe that

$$\rho(x_n(s_n), x(t-)) \leq \sup_{0 \leq u \leq s_n < T} \rho(x_n(u), x_n \circ \lambda_n(u)) + \rho(x_n \circ \lambda_n(s_n), x(t)),$$

and

$$\rho(x_n \circ \lambda_n(s_n), x(t)) \leq \sup_{0 < s_n \leq u \leq t_n < T} \rho(x_n \circ \lambda_n(u), x_n(u)) + \rho(x_n(t_n), x(t-)).$$

The idea of sufficiency of *(i)* to *(iii)* will be divided into three steps. The first, we will show that the set of points that do not satisfy the convergence is empty. The second step, we will create a sequence of $\{\lambda_n\}_{n\in\mathbb{N}} \subset \Lambda'$ which satisfies the item *(iii)* of Proposition 1.7. Finally, these $\{\lambda_n\}_{n\in\mathbb{N}}$ will be an upper (lower) bound for the sequence $\{s_n\}_{n\in\mathbb{N}}$ of this proposition.

Fix $T > 0$ and, for each $n \in \mathbb{N}$, define

$$\varepsilon_n := 2 \inf\{\varepsilon : \Gamma(t, n, \varepsilon) \neq \emptyset, \text{ for all } 0 \leq t \leq T\},$$

where

$$\Gamma(t, n, \varepsilon) := \{s \in (t - \varepsilon, t + \varepsilon) \cap [0, \infty) : \rho(x_n(s), x(t)) < \varepsilon, \rho(x_n(s-), x(t-)) < \varepsilon\}.$$

We claim that $\varepsilon_n \to 0$. Assume that the sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ does not converge to zero. Then there exists $\varepsilon > 0$, and sequences $\{t_k\}_{k\in\mathbb{N}} \subset [0, T]$ and $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ such that $\Gamma(t_k, n_k, \varepsilon) = \emptyset$ for all $k > 0$. Thus, without loss of generality, we can suppose, taking sub-sequence if necessary, that exists $t \in [0, T]$ such that, or $t_k \uparrow t$, or $t_k \downarrow t$ or $t_k = t$ for all k. For the first case, $t_k < t$ for all k and since the left limits exists it follows that

$$\lim_{k\to\infty} x(t_k) = \lim_{k\to\infty} x(t_k-) = x(t-). \tag{1.20}$$

In the second case, by the right-continuity of x, we have that

$$\lim_{k\to\infty} x(t_k) = \lim_{k\to\infty} x(t_k-) = x(t). \tag{1.21}$$

Thus, by item *(i)*, we have that for all $\varepsilon > 0$

$$\rho(x_n(s), x(s)) \wedge \rho(x_n(s), x(s-)) < \varepsilon, \tag{1.22}$$

for each continuity point s of x.

By the Lemma 1.5 and the Proposition 1.6, for the respectively cases there exist sequences $\{a_k\}_{k\in\mathbb{N}}$ and $\{b_k\}_{k\in\mathbb{N}}$ of continuity points of x such that $t_k \le a_k \le t$ or $t \le b_k \le t_k$ and therefore by (1.22), (1.20)

$$\rho(x_n(a_k-), x(t-)) < \rho(x_n(a_k-), x(a_k-)) + \rho(x(a_k-), x(t-)),$$

tends to zero and similarly by (1.22) and (1.21)

$$\rho(x_n(b_k), x(t)) < \rho(x_n(b_k), x(b_k)) + \rho(x(b_k), x(t))$$

also converges to zero as $k \to \infty$. Thus $a_k, b_k \in \Gamma(t_k, k, \varepsilon)$. It is a contradiction. It remains to show the case when $t_k = t$ for all k.

Now, we have two cases, the first is when $x(t) = x(t-)$. In this case, $t \in \Gamma(t_k, n_k, \varepsilon)$ for all $k \in \mathbb{N}$. For the second case, $x(t) \neq x(t-)$ then there exists a $\delta > 0$ such that $\rho(x(t), x(t-)) = \delta$. By the hypothesis *(ii), (iii)*, we can take a sequences $\{a_n\}, \{b_n\} \subset [0, T]$ with $a_n \downarrow t$ and $b_n \uparrow t$. Thus, there exists $n_0 \in \mathbb{N}$ such that, for all $n > n_0$,

$$\rho(x_n(a_n), x(t)) \vee \rho(x_n(b_n), x(t-)) < \frac{1}{2}(\varepsilon \wedge \delta), \tag{1.23}$$

and therefore, for $a_n, b_n \subset (t - \varepsilon, t + \varepsilon)$,

$$\sup_{a_n \le s \le b_n} \rho(x_n(s), x(t-)) \wedge \rho(x_n(s), x(t)) < \frac{1}{2}(\varepsilon \wedge \delta). \tag{1.24}$$

Now, define for each $n > n_0$,

$$s_n = \inf\{s > a_n : \rho(x_n(s), x(t)) < \frac{1}{2}(\varepsilon \wedge \delta)\}.$$

Note that by equation (1.23), $s_n \in (a_n, b_n] \subset (t - \varepsilon, t + \varepsilon)$ and therefore

$$\rho(x_n(s_n), x(t)) \le \frac{1}{2}(\varepsilon \wedge \delta),$$

and by definition of s_n

$$\rho(x_n(s_n-), x(t)) \ge \frac{1}{2}(\varepsilon \wedge \delta). \tag{1.25}$$

Thus, by the equations (1.24), (1.25) follows that

$$\rho(x_n(s_n-), x(t-)) < \rho(x(t-), x(t)) - \rho(x(t), x_n(s_n-)) < \frac{1}{2}(\varepsilon \wedge \delta),$$

and therefore $s_n \in \Gamma(t, n, \varepsilon)$ for all $n \ge n_0$ resulting a contradiction. This ensure that $\varepsilon_n \to 0$ as $n \to \infty$.

By the definition of modulus of continuity, fixed $T > 0$ and for each $n \in \mathbb{N}$, consider the partition

$$0 = t_0^n < \cdots < t_{m_n-1}^n < T \leq t_{m_n}^n,$$

with $\min_{0\leq i\leq m_n}(t_i^n - t_{i-1}^n) > 3\varepsilon_n$ such that

$$\max_{0\leq i\leq m_n} \sup_{s,t\in[t_i^n - t_{i-1}^n)} \rho(x(t), x(s)) \leq w'(x, 3\varepsilon_n, T) + \varepsilon_n. \tag{1.26}$$

Set $m_n^* = \max\{i \geq 0 : t_i^n \leq T\}$ and define $\lambda_n^T(0) := 0$ and $\lambda_n^T(t_i^n) := \inf \Gamma(t_i^n, n, \varepsilon_n)$, for $i \in \{1, \ldots, m_n^*\}$ and and interpolate linearly within the range $[0, t_{m_n^*}^n]$. For $t \geq t_{m_n^*}^n$, define

$$\lambda_n^T(t) = t - t_{m_n^*}^n + \lambda_n^T(t_{m_n^*}^n).$$

Easily follows that $\lambda_n^T \in \Lambda'$. Moreover

$$\sup_{0\leq t\leq T} \left|\lambda_n^T(t) - t\right| = \sup_{0\leq t\leq T} |\inf \Gamma(t, n, \varepsilon_n) - t| \leq \varepsilon_n$$

because $t \in [t_i^n, t_{i+1}^n)$ for some $i \in \{0, \ldots, m_n^*\}$. We claim that for each $T > 0$, we have that

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} \rho(x_n \circ \lambda_n^T(t), x(t)) = 0,$$

and, by the proposition 1.7, we will conclude that $d(x_n, x) \to 0$.

Let $\{t_n\} \subset [0, T]$ and $0 \leq t \leq T$ with $t_n \to t$ and let us separate in three cases. The first case consider $x(t) = x(t-)$. Then by item *(i)* of this proposition and taking $t_n = \lambda_n^T(t)$, follows that $d(x_n, x) \to 0$.

By item *(i)* of Lemma 1.7 and by the equation (1.26) then there exists $n \in \mathbb{N}$ such that there exist $i_n \in \{0, \ldots, m_n^*\}$ such that $t = t_{i_n}^n$. Thus, taking subsequence if necessary, we have that either $\{t_n\}_{n\in\mathbb{N}} \subset [0, t]$ or $\{t_n\}_{n\in\mathbb{N}} \subset [t, T]$.

For the first case, $\lambda_n^T(t_n) \uparrow \lambda_n^T(t_{i_n}^n) = \lambda_n^T(t)$. Thus, either

$$\rho(x_n \circ \lambda_n^T(t_{i_n}^n), x(t-)) < \varepsilon_n,$$

or

$$\rho(x_n \circ \lambda_n^T(t_{i_n}^n), x(t-)) < \varepsilon_n\,.$$

Hence there exists $\lambda_n^T(t_n) < s_n < \lambda_n^T(t_{i_n}^n)$ and therefore

$$\rho(x_n \circ \lambda_n^T(s_n), x(t)) < \varepsilon_n,$$

or

$$\rho(x_n \circ \lambda_n^T(s_n), x(t-)) < \varepsilon_n.$$

By the hypothesis *(iii)*, we have the desired convergence. For the second case it is enough to do the same argument and use the hypothesis *(ii)*.

$\square$

1.5.3 Tightness criterion in the space D

Let (S, r) a metric space.

Definition 1.15. *We say that a family of probability measures $\{\mathbb{P}_\alpha\}_{\alpha\in J}$ is tight if for all $\varepsilon > 0$ there exists a compact set $K = K(\varepsilon) \subset S$ such that $\mathbb{P}_\alpha(K) > 1 - \varepsilon$ for all $\alpha \in J$.*

In the case where E is complete and separable, then any singleton $\{\mu\}$ (where μ is a Borel probability) is tight.

Lemma 1.8. *Let (E, ρ) a complete and separable metric space. Then, all probability on E is tight.*

Proof. Let $\{x_i\}_{i\in\mathbb{N}}$ dense subset of E. Then, given $\varepsilon > 0$, we have that $E = \bigcup_{n\in\mathbb{N}} B(x_n, \frac{1}{k})$ for some $k \in \mathbb{N}$. Thus, we can find n_k such that, for all $n \geq n_k$

$$\mathbb{P}\Big[\bigcup_{n=0}^{n_k} B\Big(x_n, \frac{1}{k}\Big)\Big] > 1 - \frac{\varepsilon}{2^k}.$$

Define $R = \bigcap_{k\geq 1} \bigcup_{i=0}^{n_k} B\Big(x_i, \dfrac{1}{k}\Big)$. Since R is totally bounded set and E is complete, it follows that R is relatively compact. Thus

$$\mathbb{P}[R^{\complement}] = \mathbb{P}\Big[\bigcup_{k\geq 1}(\bigcap_{i=1}^{n_k} B(x_i, \frac{1}{k}))^{\complement}\Big] \leq \sum_{k=1}^{\infty} \mathbb{P}\Big[\Big(\bigcap_{i=1}^{n_k} B\big(x_i, \frac{1}{k}\big)\Big)^{\complement}\Big] = \sum_{k\geq 1} \frac{\varepsilon}{2^k}.$$

Taking $\operatorname{cl} R$ to be our compact set, the result follows. □

Let $\mathcal{M}(S)$ to denote a family of probability measures defined on S. We say that this family is tight if for all $\varepsilon > 0$ there exists a compact set $K = K(\varepsilon) \subset S$ such that $\mu(K) > 1 - \varepsilon$ for all $\mu \in \mathcal{M}(S)$.

Definition 1.16. *We will say that the family of probability measure $\mathcal{M}(S)$ is relatively compact if for all sequence $\{\mu_n\} \subset \mathcal{M}(S)$ we can find a subsequence $\{\mu_{n'}\} \subset \{\mu_n\}$ such that for all $\psi \in C_b(S)$ we have that $\langle \psi, \mu_{n'}\rangle \to \langle \psi, Q\rangle$ for some probability measure Q.*

The next theorem, due to Yuri Vasilyevich Prohorov [4, page 35, chapter 1], is an important tool that establishes a relation between probabilistic property about a family of measures and a topological property.

Theorem 1.7 (Prohorov's Theorem)**.** *Let S complete and separable space. The family $\mathcal{M}(S)$ is tight if and only if $\mathcal{M}(S)$ is relatively compact.*

Let $\{X_\alpha\}$ be a family of a stochastic process with sample path in $D(\mathbb{R}^+, E)$ and let $\{P_\alpha\}$ their associated probability distribution. We say that a sequence $\{X_n\}_{n\in\mathbb{N}}$ taking values in Skorohod space $D(\mathbb{R}^+, E)$ is tight if their distribution is tight. Now applying the Prohorov's Theorem on the Theorem 1.5, we have a criterion for $\{X_\alpha\}$ to be tight since E is a compact space.

Now we are interested in obtaining probabilistic conditions to characterize relatively compact sets the in $D(\mathbb{R}^+, E)$ space. Given a compact set $\Gamma \subset E$ and $\varepsilon > 0$, we define

$\Gamma^{\varepsilon} := \{a \in E : \inf_{x\in\Gamma} d(x,a) \leq \varepsilon\}$. In addition, given $\eta > 0$ and $T > 0$, we define $\Gamma_{\eta,T}$ as in the Proposition 1.8 such that

$$\mathbb{P}\{X_\alpha(t) \in \Gamma_{\eta,T} : 0 \leq t \leq T\} \geq 1 - \eta.$$

Now, we will give a result to characterize a topological property once knowing a probabilistic property.

Theorem 1.8. *Let (E,ρ) be a complete and separable metric space. Consider $\{X_\alpha\}$ a family of process with sample paths in $D(\mathbb{R}^+, E)$. Then $\{X_\alpha\}$ is relatively compact if and only if the following conditions hold.*

(i) For every $\eta > 0$ and a rational $t \geq 0$, there exists a compact set $\Gamma_{\eta,t} \subset E$ such that

$$\inf_\alpha \mathbb{P}\{X_\alpha(t) \in \Gamma^{\eta}_{\eta,t}\} \geq 1 - \eta.$$

(ii) For every $\eta > 0$ and $T > 0$ then there exists $\delta > 0$ such that

$$\sup_\alpha \mathbb{P}\{w'(X_\alpha, \delta, T) \geq \eta\} \leq \eta.$$

Proof. If $\{X_\alpha\}$ is relatively compact, by the theorems, 1.4, 1.5 and 1.7 it follows *(i)* and *(ii)*.

Conversely, let $\varepsilon > 0$ and $T > 0$ such that $\exp(-T) < \frac{\varepsilon}{2}$. Choose $\delta > 0$ such that the to second condition for $\eta = \frac{\varepsilon}{4}$. Let $m > \frac{1}{\delta}$ e define

$$\Gamma = \bigcup_{n=0}^{mT} \Gamma_{\frac{\varepsilon}{2^{i+2}}, \frac{i}{m}}.$$

Observe that

$$\begin{aligned}\mathbb{P}\big[X_\alpha(t) \in \Gamma^{\complement}\big] &= \mathbb{P}\Bigg[X_\alpha \in \Bigg(\bigcup_{n=0}^{mT} \Gamma_{\frac{\varepsilon}{2^{i+2}}, \frac{i}{m}}\Bigg)^{\complement}\Bigg] \\ &\leq \sum_{i=1}^{mT} \mathbb{P}\Big[X_\alpha \in \Gamma_{\frac{\varepsilon}{2^{i+2}}, \frac{i}{m}}\Big] = \frac{\varepsilon}{4}\sum_{i=1}^{mT}\frac{1}{2^i} < \frac{\varepsilon}{4}.\end{aligned}$$

Recalling that $\Gamma^{\varepsilon} := \{a \in E : \inf_{x\in\Gamma} d(x,a) \leq \varepsilon\}$, then by the hypothesis *(ii)* it follows that

$$\mathbb{P}\Big[X_\alpha\Big(\frac{i}{m}\Big) \in \Gamma^{\varepsilon/4} : i \in \{1, \ldots, mT\}\Big] \geq 1 - \frac{\varepsilon}{2}.$$

Now, with the notation of the Lemma 1.6, Define $A = A(\Gamma, \delta)$ which is relatively compact. Thus, given $x \in D(\mathbb{R}^+, E)$ with $w'(x, \delta, T) < \frac{\varepsilon}{4}$ and $x(\frac{i}{m}) \in \Gamma^{\varepsilon/4}$ for all $i \in \{0, \ldots, mT\}$ let $0 = t_0 < \ldots t_{n-1} < T < t_n$ with $\min_{0\leq i\leq n}\{t_i - t_{i-1}\} > \delta$ such that

$$\max_{0\leq i\leq n} \sup_{s,t\in[t_{i-1},t_i)} \rho(x(s), x(t)) < \frac{\varepsilon}{4}.$$

Taking $\{a_i\}_{i=1}^{mT} \subset \Gamma$ such that $\rho(x(\frac{i}{m}), a_i) < \frac{\varepsilon}{4}$ where $i \in \{0, \ldots, mT\}$.

Define $x' \in A$ as following:

$$x'(t) := \begin{cases} a_{\left[\frac{t_{i-1}}{m}\right]+1}, & \text{if } t \in [t_{i-1}, t_i) \\ a_{mT}, & \text{if } t \geq T, \end{cases}$$

and it follows that

$$\sup_{0 \leq t \leq T} \rho(x(t), x'(t)) \leq \frac{\varepsilon}{4}.$$

Thus, taking λ to be the identity

$$\begin{aligned} d(x, x') &< \int_0^\infty \exp(-u) d(x, x', \lambda, u)\, du \\ &= \int_0^T \exp(-u) \frac{\varepsilon}{4}\, du + \int_T^\infty \exp(-u) du \\ &\leq \frac{\varepsilon}{2} + \exp(-T) < \varepsilon. \end{aligned}$$

Hence, $x \in \{y : \inf_{x \in A} d(x, y) < \varepsilon\} =: A^\varepsilon$. It follows that $\inf_\alpha \mathbb{P}\big[X_\alpha \in A^\varepsilon\big] \geq 1 - \varepsilon$. Now, by the theorems 1.4 and 1.7, we conclude that $\{X_\alpha\}$ is relatively compact. $\square$

Definition 1.17. *We say that a family process $\{X_\alpha\}$ satisfies the compact containment condition if for every $\eta > 0$ and $T > 0$ there exist a compact set $\Gamma_{\eta,T} \subset E$ such that*

$$\inf_\alpha \mathbb{P}[X_\alpha(t) \in \Gamma_{\eta,T} : \textit{for all } 0 \leq t \leq T] \geq 1 - \eta.$$

Observe that, if $\{X_\alpha\}$ is relatively compact, then $\{X_\alpha\}$ satisfies the compact containment condition. The following Corollary is an important characterization for that processes to be relatively compacts. It follows by a little modification of the Theorem 1.8.

Corollary 1.1. *Let (E, ρ) be a complete and separable metric space and $\{X_n\}_{n \in \mathbb{N}}$ is a sequence of process with sample paths in $D(\mathbb{R}^+, E)$. Then, $\{X_n\}_{n \in \mathbb{N}}$ is relatively compact if and only if*

(i) For every $\eta > 0$ and rational $t \geq 0$, there exists a compact set $\Gamma_{\eta,t} \subset E$ such that

$$\liminf_{n \to \infty} \mathbb{P}[X_n(t) \in \Gamma^\eta_{\eta,t}] \geq 1 - \eta.$$

(ii) For every $\eta > 0$ and $T > 0$, there exists $\delta > 0$ such that

$$\limsup_{n \to \infty} \mathbb{P}[w'(X_n, \delta, T) \geq \eta] \leq \eta.$$

Remark 1.10. Note that the first condition of the Corollary 1.1 gives us the tight condition at each fixed time and the second hypothesis prevents large jumps to be close together.

Proof. Fix $\eta > 0$, rational $t \geq 0$ and $T > 0$. For each $n > 0$, since (E, ρ) is complete and separable metric space, every probability measure is tight by Lemma 1.8, there exists a compact set $\Gamma_n \subset E$ such that

$$\mathbb{P}\{X_n(t) \in \Gamma^\eta_n\} \geq 1 - \eta,$$

and by the Lemma 1.7 *(i)* there exists $\delta_n > 0$ such that

$$\mathbb{P}\{w'(X_n, \delta_n, T) \geq \eta\} \leq \eta.$$

Now, suppose *(i)* and *(ii)*. Then, there exist a compact set $\Gamma_0 \subset E$, δ_0 and a integer n_0 such that

$$\inf_{n \geq n_0} \mathbb{P}\{X_n(t) \in \Gamma_0^\eta\} \geq 1 - \eta,$$

and

$$\sup_{n \geq n_0} \mathbb{P}\{w'(X_n, \delta_n, T) \geq \eta\} \leq \eta.$$

Putting $\Gamma_0 = \bigcup_{n=0}^{n_0-1} \Gamma_n$ and $\delta_0 = \min_{0 \leq n \leq n_0-1} \delta_n$, we can take $n_0 = 1$ and, by the Theorem 1.8, we conclude. $\square$

Chapter 2

The Branching Brownian Motion

The Branching Brownian Motion (BBM) is a classical model in stochastic processes which can be briefly described as follows. Consider a single particle moving on $\mathbb{R}^d$ according to a Brownian motion. After a exponential random time of parameter V, this particles splits into a random number k of offsprings with probability generating function $\delta(s) = \sum_{k=0}^{\infty} p_k s^k$ where $\sum_{k=0}^{\infty} p_k = 1$ and are independent copies of itself which starts Brownian motion at the same location where she died. This process then repeats *ad infinitum*.

Naturally, we will have a lot of ask about this model but we will be more interested in the way this process spreads out and his behavior. A particular interest of BBM stems from the fact that it is closely linked to a non-linear partial differential equation introduced by Fischer and later by Kolmogorov, Petrovsky and Piscounov, the F-KPP equation.

The F-KPP equation is a differential partial equation that belongs to the class of reaction-diffusion equation.

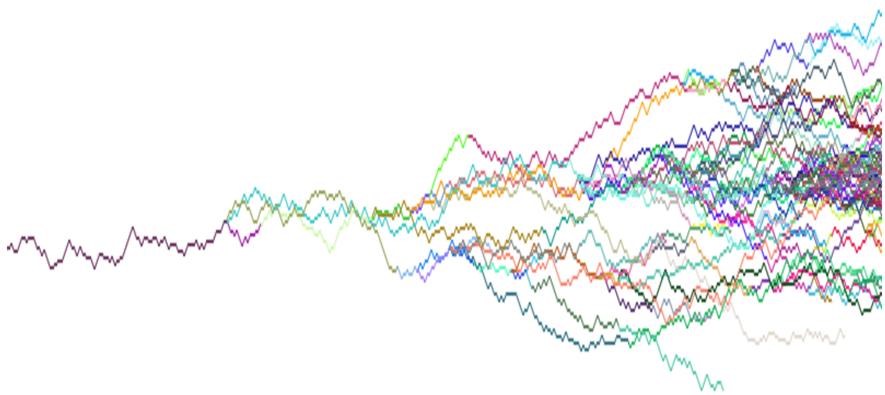


Figure 2.1: Illustration of the BBM in dimension one. Image courtesy of Rodrigo Viégas.

2.1 The Galton-Watson branching process

The branching process is a class of stochastic processes in which independent particles move around the space according to some (Markovian) process. This class is useful because these processes model the growth of a population.

At time t, denote $\mathcal{N}_t$ the set of all particles alive, and for $u \in \mathcal{N}_t$ let X_u^t be the position of each particle in the system. We are interested in rooted ordered trees so that we can identify and study the living population at time t. For that, set $\mathbb{N} = \{0, 1, 2, \dots\}$. Define

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n, \quad U = \mathcal{U} \cup \mathbb{N}^{\infty}.$$

Let $u \in \mathcal{U}$. Then u is a finite sequence of integers $u = (u_1, u_2, \dots, u_n)$ for some $n \in \mathbb{N}$. Denote $|u| = n$ to represents the generation of u. Define the map $p : \mathcal{U} \setminus \varnothing \to \mathcal{U}$ such that $p(u_1, \dots, u_n) = (u_1, \dots, u_{n-1})$. Note that this map carries information from the ancestors of the n-th generation of u. To illustrate what its mean, consider the following example

Example 2.1. Consider the $u \in \mathcal{U}$ such that $u = (2, 3, 4)$ then, u represents the 4-th child of the 3-th child of 2-th child of the root and, the function $p(u) = (2, 3)$ telling us that the descendants of u are the 3-th child of the 2-th child of the root. Moreover, we have that $|u| = 3$ which means that u belongs to the third generation. Figure 2.2 illustrate this example.

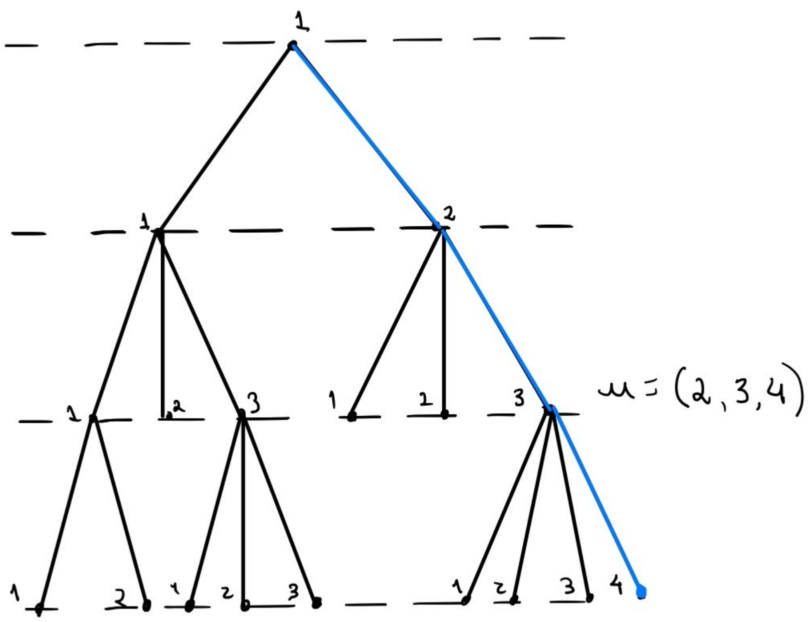


Figure 2.2: In blue, the 4-th child of the 3-th child of the 2-th child of the root.

Definition 2.1. *A (locally finite, rooted) plane tree τ is a subset of $\mathcal{U}$ such that* *(i)* $\varnothing \in \tau$,

(ii) $u \in \tau \setminus \varnothing$ *then* $p(u) \in \tau$,

(iii) *For every* $u = (u_1, \dots, u_n) \in \tau$ *there exists an integer* $A_u \geq 0$ *such that for every* $j \in \mathbb{N}$, $(u_1, \dots, u_j) \in \tau$ *if and only if* $j \leq A_u$.

Let us denote by $\mathcal{T}$ *the set of all plane trees.*

Consider the following collection of independent random variables $(A_u, u \in \mathcal{U})$ each of them with law $\mathbb{P}$. Denote by $\mathbf{T}$ the random subset of $\mathcal{U}$ defined by

$$\mathbf{T} = \{u \in \mathcal{U} : u_j \leq A_{p(u)}, \text{ for every } 1 \leq j \leq n\}.$$

Definition 2.2. *A Galton-Watson process is a Markov process in* $\mathbb{Z}^+$ *which is defined as*

$$Z_0 = 1, \textit{ and for all } n \geq 0, Z_{n+1} = \sum_{i=1}^{Z_n} X_i^{n+1}, \tag{2.1}$$

where $\{X_i^n : n, i \in \mathbb{N}\}$ *are i.i.d random variables such that* $\mathbb{P}[X_i^n = k] = p_k$ *and* $\sum_{k \geq 0} p_k = 1$.

Thus, we have that $\mathbf{T}$ is a $\mathbb{P}$-Galton-Watson tree and $(Z_n, n \in \mathbb{N})$ is the associated Galton-Watson processes where $Z_n = \#\{u \in \mathbf{T} : |\mathbf{u}| = \mathbf{n}\}$.

Remark 2.1. Observe that the branching process is a Markov chain. Indeed the size of next generation only depends on the size of previous generation, i.e., Z_{n+1} is independent of $Z_0, Z_1, \cdots, Z_{n-1}$.

Observe that the Galton-Watson branching process has probability generating function given by

$$G_{Z_n}(s) = \sum_{k=1}^{Z_{n-1}} s^{Z_n} \mathbb{P}[Z_n = k].$$

Therefore, if Z_n, W_n are branching processes, it follows that $G_{Z_n+W_n}(s) = G_{Z_n}(s) G_{W_n}(s)$ which it is enough to conclude that if p is the transition function

$$p(\cdot, Z_n + W_n) = p(\cdot, Z_n) * p(\cdot, W_n).$$

Proposition 2.1. *Let* $\{Z_n : n \in \mathbb{N}\}$ *be a Galton-Watson process as defined in 2.1 such that* $\mathbb{E}[X_1] < \infty$. *Then for each* $n \in \mathbb{N}$, $\mathbb{E}[Z_n] = (E[X_1])^n$.

Proof. Observe that

$$\begin{aligned}
\mathbb{E}[Z_n] &= \sum_{k\geq 0} \mathbb{E}[Z_n \mid Z_{n-1}=k]\mathbb{P}[Z_{n-1}=k] \\
&= \sum_{k\geq 0} \mathbb{E}\Big[\sum_{i=0}^{Z_{n-1}} X_i | Z_{n-1}=k\Big]\mathbb{P}[Z_{n-1}=k] \\
&= \sum_{k\geq 0} \mathbb{E}\Big[\sum_{i=0}^{k} X_i | Z_{n-1}=k\Big]\mathbb{P}[Z_{n-1}=k] \\
&= \sum_{k\geq 0} \mathbb{E}\Big(\sum_{i=0}^{k} X_i\Big)\mathbb{P}[Z_{n-1}=k] \\
&= \mathbb{E}[X_1]\sum_{k\geq 0} k\mathbb{P}[Z_{n-1}=k] \;=\; \mathbb{E}[X_1]\,\mathbb{E}[Z_{n-1}].
\end{aligned}$$

Iterating the result above with a basic recurrence it follows that $\mathbb{E}[Z_n] = (\mathbb{E}[X_1])^n$. □

Definition 2.3. *A branching process is said subcritical if* $\mathbb{E}[Z_n] < 1$*, critical if* $\mathbb{E}[Z_n] = 1$ *and supercritical if* $\mathbb{E}[Z_n] > 1$.

Proposition 2.2. *Let* $\{Z_n\}$ *a Galton-Watson process and* $\mathcal{F}_n = \sigma(X_i^m : i \geq 1, 0 \leq m \leq n)$*. If* $\mu = \mathbb{E}[X_i^n] < \infty$*, then the process* $\{\frac{Z_n}{\mu^n}\}$ *is a* $(\mathcal{F}_n)$*-martingale.*

Proof. Let us check the martingale properties. The process is clearly $\mathcal{F}_n$-adapted because it is sum of $\mathcal{F}_n$-measurable random variables. Moreover

$$\mathbb{E}\,|Z_n| = \mathbb{E}\left|\sum_{i=0}^{Z_{n-1}} X_i^n\right| \leq \sum_{i=0}^{Z_{n-1}} \mathbb{E}\,|X_i^n| = \sum_{i=0}^{Z_{n-1}} \mathbb{E}[X_i^n] < \infty.$$

Let us compute the conditional expectation. By the linearity of conditional expectation and the Monotone Convergence Theorem,

$$\begin{aligned}
\mathbb{E}[Z_{m+1} \mid \mathcal{F}_m] &= \mathbb{E}[Z_{m+1}\mathbb{1}_{\bigcup_{k\geq 0}[Z_m=k]} \mid \mathcal{F}_m] \\
&= \mathbb{E}[\lim_{k\to\infty}\sum_{i=0}^{k} Z_{m+1}\mathbb{1}_{[Z_m=k]} \mid \mathcal{F}_m] \\
&= \lim_{k\to\infty}\sum_{i=0}^{k} \mathbb{E}[Z_{m+1}\mathbb{1}_{[Z_m=k]} \mid \mathcal{F}_m].
\end{aligned} \tag{2.2}$$

Since $\mathbb{1}_{[Z_m=k]} \in \mathcal{F}_m$, $\{X_i^n\}$ are independent and identically distributed and $\{X_i^{m+1}\}$ are independent of $\mathcal{F}_m$, the equation (2.2) becomes

$$\lim_{k\to\infty}\sum_{i=0}^{k} \mathbb{1}_{[Z_m=k]}\mathbb{E}[\sum_{i=0}^{k} X_i^{m+1} \mid \mathcal{F}_m] = \lim_{k\to\infty}\sum_{i=0}^{k} k\mathbb{1}_{[Z_m=k]}\mu = \mu Z_n$$

Finally, dividing both sides by μ^{m+1}, the result follows. □

2.1.1 The branching Brownian motion

The spatial tree is an element $\tau \in \mathcal{T}$ with more structure. For each $u \in \tau$, we associate a life-time $\sigma_u \geq 0$ and that allows us to define the birth-time of u by $b_u = \sum_{v<u} \sigma_u$ and its death time by $d_u = b_u + \sigma_u$. Moreover, for each $u \in \tau$ and time t there is a map $X_t : \mathbb{R}^+ \to \mathcal{N}_t$, that is, a map that gives an "order" to all living individuals at time t. Once defined a spatial tree, we can define a marked tree.

Definition 2.4. *We define a marked tree as the triplet*

$$\mathbf{t} = (\tau, \sigma, \mathbf{X}) = (\tau, \{\sigma_{\mathbf{u}}, (\mathbf{X}_{\mathbf{u}}(\mathbf{s}), \mathbf{s} \geq \mathbf{0}), \mathbf{u} \in \tau\}).$$

Definition 2.5. *Let $\mathcal{N}_t \subset \mathcal{U}$ be the set of all particles alive at time t, that is*

$$\mathcal{N}_t = \{u \in \tau : b_u \leq t \leq d_u\}.$$

Remark 2.2. Note that, if $u \in \mathbb{R}^d$ is a particle alive at time t, we can define its position, inductively, as

$$Y_t^u := X_u(t - b_u) + Y_{d_u}^{p(u)},$$

where $Y_{d_u}^{p(u)}$ denotes the position where the "father" of u died.

Once defined the *Galton-Watson process* and the *spatial tree*, we can define rigorously a branching Brownian motion considering the following properties:

1. A branching mechanism such that $A_u = \Psi(u)$, for all $u \in \mathcal{U}$ where $\Psi(u)$ is the probability generating function of u,
2. The branching rate: For all $u \in \mathcal{U}$, σ_u are i.i.d. exponentially distributed with parameter V,
3. The spatial motion: The random variables Y^u are standard Brownian motions.

Definition 2.6. *Let $w = (\tau, (\sigma_u, Y^u)_{u\in\tau}, A_u)$ then we define*

$$Y_t = Y_t(w) = (Y_t^u, u \in \mathcal{N}_t),$$

to be the branching Brownian motion process.

Remark 2.3. For this process, the natural filtration is given by

$$\mathcal{F}_t = \sigma\{Y_s : s \leq t\} = \sigma\{Y_s, \mathcal{N}_t\}.$$

The distribution for this process is usually denoted by $\mathbb{P}$ or $\mathbb{P}_x$ when we want to highlight that $x \in \mathcal{N}_t$ is the initial state.

The branching Brownian motion process also can be viewed as a measure-valued process. The ξ_t will denote the measure representing the whole population at time t

$$\xi_t(\cdot) = \sum \delta_{Y_t^i}(\cdot),$$

which is also $\mathcal{F}_t$-adapted.

Remark 2.4. Observe that the following random variables are measurable with respect to $\mathcal{F}_t$.

- All the σ_v such that $v < u$ for some $u \in \mathcal{N}_t$.
- Any Y_s^u for $s \leq t$ and $u \in \mathcal{N}_t$.

Proposition 2.3. *The branching Brownian motion Y_t (or ξ_t) is strongly Markovian.*

Proof. By the memoryless property of exponentials and the Brownian motions is strongly Markovian, the result follows. □

Remark 2.5. If the initial population is the (purely atomic) measure ν, then we write $\mathbb{P}_\nu(\cdot)$ for the distribution of the process and $P_t(\cdot, \nu)$for the corresponding transition probability.

Since the branching Brownian motion is a branching processes, the following property holds

$$P_t(\cdot, \nu_1 + \nu_2) = P_t(\cdot, \nu_1) * P_t(\cdot, \nu_2),$$

where the symbol $*$ denotes the convolution between the measures ν_1 and ν_2.

2.2 F-KPP equation

The Fischer, Kolmogorov, Petrovski, Piscounov (F-KPP) equation was first considered in 1937 by R. A. Fisher in [11], which was proposed and studied in the context of population dynamics to describe the spatial spread of an advantageous allele, analyzing its traveling wave solutions. Kolmogorov, Petrovsky, and Piskunov [17] studied simultaneously this equation. Over the years, these reaction-diffusion equations have been studied by many authors, among them: Kolmogorov, Fisher, Skorohod, McKean, Dawson, Le Gall, and Perkins to name just some of them.

The equation describes the evolution of an invasion front from a stable phase into an unstable phase. It is a semi-linear equation of the form

$$u_t = \frac{1}{2} u_{xx} + f(u) \tag{2.3}$$

where the forcing term f is assumed to be in $C[0, 1]$ and satisfies the conditions

$$f(0) = f(1) = 0, \, f(u) \geq 0 \, , u \in (0, 1]$$

and

$$f'(0) > 0, \, f'(u) \leq f'(0), \, u \in (0, 1].$$

This equation is ubiquitous in the study of reaction-diffusion phenomena and front propagation. It appears in models related to diverse fields as ecology, population genetics, combustion, epidemiology, etc. It is one of the simplest examples of a semi-linear parabolic equation, which admits traveling wave solutions. Briefly, a mathematical wave is a function of the form $u(x, t) = f(x - ct)$, with $c > 0$ by convention, and such wave is solution of the constant-coefficient transport equation

$$u_t + cu_x = 0.$$

At $t = 0$, the wave has the form $f(x)$, which is the boundary condition for the problem, so the wave $f(x - ct)$ represents the wave profile translated to the right by ct spatial units.

2.3 McKean representation

The next Theorem, due to H. McKean [20], gives a representation of solutions to the F-KPP equation in term of the BBM which will allow us to study branching Brownian motion from the heat equation. We can understand the heat in the discrete case, as follows.

Consider the finite set $A \in \mathbb{Z}^d$ with boundary ∂A. We will define the temperature at the border as zero, and the initial condition set the temperature at $x \in A$ to be $p_t(x)$ with $t \in \mathbb{N}$ the time unit. For each integer t, the heat in x at time t is spreading among its $2d$ first neighbors. If one of those neighbors is a boundary point, then the heat that goes to that site is lost forever.

In a more probabilistic way, we can understand as a very large amount of "heat particles" which perform a random walk in A until the moment they touch the boundary. The temperature at x at time t, $p_t(x)$ is given by density of particles $x \in A \subset \mathbb{Z}^d$.

Remark 2.6. We use to denote the domain of the Laplacian $\mathcal{D}(\Delta)$. In general, $\mathcal{D}(\mathcal{L})$ will be the domain of the linear operator $\mathcal{L}$. Subscripts will indicate a time variable, *not a derivative*.

The theorem is often called the *McKean representation*. It says that the solution of the F-KPP equation can be viewed as an expectation of spatial distribution of the braching Brownian motion. It is, essentially, a type of Feynman-Kac result according with [16, page 334].

Theorem 2.1. *The distribution of the branching Brownian motion can be characterized as follows: for $\psi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$ with $0 \leq \psi(x) \leq 1$ for $x \in \mathbb{R}^d$,*

$$\mathbb{E}_{\delta_x}\left[\prod_{u \in \mathcal{N}_t} \psi(Y_t^u)\right] = v(t,x)\,, \tag{2.4}$$

and v solves

$$\begin{cases} \dfrac{\partial v}{\partial t} = \dfrac{1}{2}\Delta v + (\Psi(v) - v)V \\ v(0,x) = \psi(x)\,. \end{cases} \tag{2.5}$$

where $\Psi(v)$ is the generating probability function and V is the branching rate.

Before proving the theorem, let us discuss its meaning. Once the theorem is proved, we will have that the expected value for the position of the particles in a branching Brownian motion is described by the heat equation, i.e., it is described by the phenomenon of diffusion. These are particular cases of so-called transport processes, in which there is a directed transfer of internal energy, mass, linear moment, etc. in supposedly continuous environments.

Such problems can be approached in a relatively same way in the sense of considering each of these measures as "fluids" such that is modeled by a certain conservation law. In general, all these cases obey a continuity equation, which can be found in its physical-mathematical construction elsewhere [12].

The heat equation is a particular case of the diffusion process. The heat conduction process is determined by the temperature $u(t,x)$. As well known, heat flows from the highest temperature points to the lowest temperature points, and in the case of the presence of heat sources or sinks, the forcing term in (2.3) will represent the amount added or removed.

With this in mind, we can understand that the expected value of the particle positions in $\mathbb{R}^d$ will play the role of temperature, diffusing through space according to the heating process.

Remark 2.7. The hypothesis $\psi \in [0,1]$ ensures that function $v(t,x)$ defined in the Theorem 2.5 is finite.

Proof. Initially let $\psi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$ defined as

$$\prod_{i\in\mathcal{N}_t} \psi(Y_t^i) \overset{\text{def}}{=} \begin{cases} 1, & \text{if } \mathcal{N}_t = \varnothing \\ \prod_{i\in\mathcal{N}_t} \psi(Y_t^i), & \text{otherwise.} \end{cases}$$

Let $v(t,x) \overset{\text{def}}{=} \mathbb{E}_{\delta_x}\Big[\prod \psi(Y_t^i)\Big]$. It is clear that $v(0,x) = \psi(x)$. In fact, at time $t = 0$, we have only one particle, which we will call the original ancestor and therefore follows the equality $Y_0^i = x$.

We assume that at time t, $v(t,x)$ is twice differentiable with respect to the space variable x and by [9, Theorem 8, page 69] $v(t,x)$ is smooth as a function of x at all later times.

At time t, a certain number of individuals are alive. Each one of them will be called of *ancestor* and recall that they are all independent. We are interested in the behaviour between 0 and t for each of these individuals. To clarify ideas, fix one of the original ancestors, which we call I_1. Define the hitting time

$$T = \inf\{t \geq 0 : \text{ the original ancestor dies before time } t\ \}$$

and consider the space partition

$$K = \{\text{ after the original ancestor dies, he leaves exactly } k \text{ offsprings }\}.$$

Then, we can writing our solution as

$$\begin{aligned} \mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t} \psi(Y_t^i)\Big] &= \mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t} \psi(Y_t^i)\mathbb{1}_{[T\leq t]}\Big] + \mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t} \psi(Y_t^i)\mathbb{1}_{[T>t]}\Big] \\ &= \sum_{k\geq 0} \mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t} \psi(Y_t^i)\mathbb{1}_{[T\leq t]}\mathbb{1}_K\Big] + \mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t} \psi(Y_t^i)\mathbb{1}_{[T>t]}\Big]. \end{aligned}$$

Now we will handle each term above. We claim that:

- $\mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t} \psi(Y_t^i)\mathbb{1}_{[T\leq t]}\mathbb{1}_K\Big] = \int_0^t ds\, V \exp\{-Vs\}\, \mathbb{E}_{\delta x}\big[p_k v^k\big(t-s, Y_t\big)\big],$
- $\mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t} \psi(Y_t^i)\mathbb{1}_{[T>t]}\Big] = \mathbb{E}_x[v(t-T, B_t)] \int_t^\infty ds\, V \exp\{-Vs\},$

where $\mathbb{E}_x$ denotes expectation for a Brownian motion started at the point x at time zero and B_t represents the Brownian that survived.

Since the process is Markovian and T have exponential distribution with parameter V, it follows that

$$\begin{aligned}
\mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t}\psi(Y_t^i)\mathbb{1}_{[T\le t]}\mathbb{1}_K\Big] &= \mathbb{E}_{\delta_x}\Big[\mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t}\psi(Y_t^i)|[T\le t]\Big]\mathbb{1}_{[T\le t]}\mathbb{1}_K\Big]\\
&= \mathbb{E}_{\delta_x}\Big[\mathbb{E}_{\delta_x}\Big[\prod_{j=1}^{k}\prod_{u\in\mathcal{N}_t}\psi(Y_{T-t}^i)\circ\theta_T|[T\le t]\Big]\mathbb{1}_{[T\le t]}\mathbb{1}_K\Big]\\
&= \mathbb{E}_{\delta_x}\Big[\mathbb{E}_{Y_T}\Big[\prod_{j=1}^{k}\prod_{u\in\mathcal{N}_t}\psi(Y_{T-t}^i)\Big]\mathbb{1}_{[T\le t]}\mathbb{1}_K\Big]\\
&= \int_0^t ds\, p_k V\exp\{-Vs\}\,\mathbb{E}_{\delta x}\Big[\mathbb{E}_{Y_T}\Big[\prod_{j=1}^{k}\prod_{u\in\mathcal{N}_t}\psi(Y_{T-t}^i)\Big]\Big].
\end{aligned}$$

Now, given that the original ancestor died at position $Y_{T_k}^1\in\mathbb{R}^d$, the particles generated are independent of each other and each particle are a Brownian motion. Therefore the product over all alive particles become k copies of a Brownian motion started at $Y_{T_k}^1\in\mathbb{R}^d$. Thus, by the linearity of expectation and definition of $v(t,x)$

$$\begin{aligned}
\mathbb{E}_{\delta_x}\Big[\psi(Y_t^0)\mathbb{1}_{[T\le t]}\mathbb{1}_K\Big] &= \int_0^t ds\, p_k V\exp\{-Vs\}\,\mathbb{E}_{\delta x}\Big[\mathbb{E}_{Y_T}\Big[\Big(\prod_{u\in\mathcal{N}_t}\psi(Y_{t-T}^i)\Big)^k\Big]\Big]\\
&= \int_0^t ds\, p_k V\exp\{-Vs\}\,\mathbb{E}_{\delta x}\Big[\mathbb{E}_{Y_T}\Big[\prod_{u\in\mathcal{N}_t}\psi\Big(Y_{t-T}^i\Big)\Big]^k\Big]\\
&= \int_0^t ds\, V\exp\{-Vs\}\,\mathbb{E}_{\delta x}\big[p_k v^k\big(t-s,Y_t\big)\big].
\end{aligned}$$

To show the second claim note that the Brownian motion has survived and is independent of the hitting time T. Then

$$\begin{aligned}
\mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t}\psi(Y_t^i)\mathbb{1}_{[T>t]}\Big] &= \mathbb{E}_{\delta_x}\Big[\mathbb{E}\Big[\psi(Y_t^0)|T>t\Big]\mathbb{1}_{[T>t]}\Big]\\
&= \mathbb{E}_{\delta_x}\Big[\psi(Y_t^1)\mathbb{1}_{[T>t]}\Big]\\
&= \mathbb{P}[T>t]\mathbb{E}\big[\psi(B_t)\big],
\end{aligned}$$

where B_t denotes the position of the Brownian motion which is alive at time $T-t$. Now, we will show that expectation solves the P.D.E. (2.5) and we have two cases whether $\mathcal{N}_t$ is empty or not. For the first case, by definition, the product over all particles is one.

Trivially for this case, $v(t,x)$ solves (2.5). Suppose now that $\mathcal{N}_t$ is not empty. Hence

$$\begin{aligned}
\mathbb{E}_{\delta x}\Big[\prod_{u\in\mathcal{N}_t}\psi(Y_t^i)\Big] &= \mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t}\psi(Y_t^i)\mathbb{1}_{[T\le t]}\Big] + \mathbb{E}_{\delta_x}\Big[\prod_{u\in\mathcal{N}_t}\psi(Y_t^i)\mathbb{1}_{[T>t]}\Big]\\
&= \exp\{-Vt\}\mathbb{E}_{\delta x}\big[\psi(B_t)\big] + \int_0^t ds\, V\exp\{-Vs\}\,\mathbb{E}_{\delta x}\Big[\sum_{k\ge 0}p_k v^k\big(t-s,Y_t\big)\Big]\\
&= \exp\{-Vt\}\mathbb{E}_{\delta x}\big[\psi(B_t)\big] + \int_0^t ds\, V\exp\{-Vs\}\,\mathbb{E}_{\delta x}\Big[\Psi\Big(v\big(t-s,Y_t\big)\Big)\Big].
\end{aligned}$$

Since $v(t,\cdot)$ is continuously differentiable with respect to the space variable x, making the substitution $t-s\to s'$ in the integral and by [15, Lemma 2.3, page 171] we get

$$\begin{aligned}
\partial_t v(t,x) &= -V\exp\{-Vt\}\mathbb{E}_{\delta x}\big[\psi(B_t)\big] + \exp\{-Vt\}\frac{1}{2}\Delta\mathbb{E}_{\delta x}\big[\psi(B_t)\big]\\
&+ \partial_t\Big[\int_0^t ds\; V\exp\{-V(t-s')\}\mathbb{E}_{\delta x}\Big[\Psi\Big(v\big(t-s,Y_t\big)\Big)\Big]\Big]\\
&= -V\exp\{-Vt\}\mathbb{E}_{\delta x}\big[\psi(B_t)\big] + \exp\{-Vt\}\frac{1}{2}\Delta\mathbb{E}_{\delta x}\big[\psi(B_t)\big]\\
&+ V\sum_{k\ge 0}p_k v^k(t,x) + \int_0^t ds'\,\Big[-V^2\exp\{-V(t-s')\}\mathbb{E}_{\delta x}\big[v^k\big(t-s,Y_t\big)\big]\\
&+ V\exp\{-V(t-s')\}\partial_t\mathbb{E}_{\delta x}\big[v^k\big(t-s,Y_t\big)\big]\Big]\\
&= -V\mathbb{E}_{\delta_x}\Big[\prod_{i\in\mathcal{N}_t}\psi\big(Y_t^i\big)\mathbb{1}_{[T>t]}\Big] + \frac{1}{2}\Delta\mathbb{E}_{\delta_x}\Big[\prod_{i\in\mathcal{N}_t}\psi\big(Y_t^i\big)\mathbb{1}_{[T>t]}\Big]\\
&+ V\Psi(v(t,x)) - V\mathbb{E}_{\delta_x}\Big[\prod_{i\in\mathcal{N}_t}\psi\big(Y_t^i\big)\mathbb{1}_{[T\le t]}\Big] + \frac{1}{2}\Delta\mathbb{E}_{\delta_x}\Big[\prod_{i\in\mathcal{N}_t}\psi\big(Y_t^i\big)\mathbb{1}_{[T\le t]}\Big]\\
&= \frac{1}{2}\Delta v(t,x) + V\Big(\Psi(v(t,x)) - v(t,x)\Big).
\end{aligned}$$

Once this process is established for the original ancestor, we can reproduce for each of the descendants, considering them as an original ancestor, and this proves the theorem.

□

For now on we will use $\langle\cdot,\cdot\rangle$ to denote the integral, i.e., $\langle\phi,\mu\rangle = \int\phi\,d\mu$.

Corollary 2.1. *We can define the branching Brownian motion as following: assuming that the initial state of the population is represented by the purely atomic measure ν, its state at time t will be the measure ξ_t determined by*

$$\mathbb{E}_\nu\big[\exp\left(\langle\log\psi,\xi_t\rangle\right)\big] = \exp\big(\langle\log v(t,\cdot),\nu\rangle\big). \tag{2.6}$$

Proof. Let $\mathcal{N}_t$ the number of alive individuals at time t. Thus, $\xi_t = \{\delta_{Y_t^1},\dots,\delta_{Y_t^{\mathcal{N}_t}}\}$ and therefore

$$\langle\log\psi,\xi_t\rangle = \Big\langle\log\psi,\sum_{i=1}^{\mathcal{N}_t}\delta_{Y_t^i}\Big\rangle = \log\psi(Y_t^1) + \ldots + \log\psi(Y_t^{\mathcal{N}_t}).$$

Hence

$$\exp\left(\langle \log\psi, \xi_t\rangle\right) \;=\; \prod_{i=1}^{\mathcal{N}_t} \psi(Y_t^i)$$

and Theorem 2.1 allows to conclude. □

Remark 2.8. The expression (2.6) is the so-called Laplace functional of the Branching Brownian motion.

2.4 A martingale characterisation

Towards to characterization of branching Brownian motion as a martingale, we will need to deal with the functionals of the branching Brownian motion which can be expressed in terms of functions of the position of each particle alive in time t and then we would like to be able to determine when these functionals is a martingale, and it is easy when we know the infinitesimal generator of this processes.

Proposition 2.4. *Let ξ_t be the branching Brownian motion with branching rate V. Then its generator is given by*

$$\mathcal{L}_1 F(\eta) \;=\; \left\langle \frac{\frac{1}{2}\Delta\psi \;-\; V\big(\Psi(\psi) - \psi\big)}{\psi}, \eta \right\rangle \exp\left\{\langle \log(\psi), \eta\rangle\right\}. \tag{2.7}$$

Proof. Let v as defined in (2.4). By the Corollary 2.1 we know that

$$\mathbb{E}_\eta\big[\exp\left(\langle \log\psi, \xi_t\rangle\right)\big] \;=\; \exp\left\{\langle \log v(t,\cdot), \eta\rangle\right\}$$

By the above and Theorem 2.1, it follows that

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}[F(\xi_t)] \;&=\; \frac{\partial}{\partial t}\exp\left(\langle \log v(t,\cdot), \eta\rangle\right) \\
&=\; \exp\left(\langle \log v(t,\cdot), \eta\rangle\right)\frac{\partial}{\partial t}\langle \log v(t,\cdot), \eta\rangle \\
&=\; \exp\left(\langle \log v(t,\cdot), \eta\rangle\right)\left\langle \frac{\frac{1}{2}\Delta v \;-\; V\big(\Psi(v) - v\big)}{v}, \eta \right\rangle.
\end{aligned}$$

Since

$$\mathcal{L}_1 F(\eta) \;=\; \frac{d}{dt}\mathbb{E}_\eta[F(\xi_t)]\Big|_{t=0},$$

we conclude the proof. □

Observe that, by the Proposition A.2 *(ii)*, if $\mathcal{L}_1$ is the infinitesimal generator of the branching Brownian motion with Laplace functional f, we have that $\frac{d}{dr}P_t f \;=\; P_t\mathcal{L}_1 f$. Thus

$$\frac{d}{dr}\mathbb{E}_\nu[F(\xi_r)] \;=\; \mathbb{E}_\nu[\mathcal{L}_1 F(\xi_r)]$$

By the Lebesgue Differentiation theorem under the sign of Integral and the Fubini's Theorem, integrating both side of the equation above on the interval $[s, s+t]$ it follows that

$$\mathbb{E}_\nu\big[F\big(\xi_{s+u}\big) \;-\; F\big(\xi_s\big)\big] \;=\; \mathbb{E}_\nu\bigg[\int_s^{s+u} \mathcal{L}_1 F\big(\xi_y\big)\, dy\bigg], \tag{2.8}$$

Proposition 2.5. *The process*

$$M(\xi_t) = F(\xi_t) - F(\xi_0) - \int_0^t \mathcal{L}_1 F(\xi_s)\,ds \tag{2.9}$$

is a mean zero $\mathbb{P}_\nu$-martingale for all $F \in \mathcal{D}(\mathcal{L}_1)$ where $\mathcal{L}_1$ is given by equation (2.7).

Remark 2.9. The proof of the Proposition 2.5 is slightly similar to the proof of Proposition 1.4 which is more general. Here we will use an explicit formula for the infinitesimal generator of the branching Brownian motion while in Proposition 1.4 we used the theory of semi-groups argument for a generic generator which did not done in this dissertation and therefore, we will keep this demonstration although are quite similar.

Proof. Note that F and $\mathcal{L}_1$ are continuous and since the integral is a continuous operator it follows that the process is $\mathcal{F}_t$-adapted. To show that the process have first moment finite, it is enough to note that ϕ and $\mathcal{L}_1$ are bounded. By the linearity of conditional expectation

$$\mathbb{E}_\nu[M(\xi_t)|\,\mathcal{F}_s] = \mathbb{E}_\nu[F(\xi_t)\,|\,\mathcal{F}_s] - \mathbb{E}_\nu[F(\xi_0)\,|\,\mathcal{F}_s] - \mathbb{E}_\nu\Big[\int_0^s \mathcal{L}F(\xi_u)\,du + \int_s^t \mathcal{L}F(\xi_u)\,du\,|\,\mathcal{F}_s\Big].$$

Let u such that $s+u=t$ and $A \in \mathcal{F}_s$, then by the Markov property

$$\begin{aligned}
\mathbb{E}_\nu\Big[\mathbb{E}_\nu\Big[\int_s^t \mathcal{L}F(\xi_u)\,du\,|\mathcal{F}_s\Big]\mathbb{1}_A\Big] &= \mathbb{E}_\nu\Big[\mathbb{E}_\nu\Big[\theta_{\xi_s}\circ\int_0^u \mathcal{L}F(\xi_{u+s})\,du\,|\mathcal{F}_s\Big]\mathbb{1}_A\Big] \\
&= \mathbb{E}_\nu\Big[\mathbb{E}_{\xi_s}\Big[\int_0^u \mathcal{L}F\big(\xi_{u+s}\big)\,du\Big]\mathbb{1}_A\Big] \\
&= \mathbb{E}_\nu[\mathbb{E}_{\xi_s}[F(\xi_t) - F(\xi_s)]\mathbb{1}_A] \\
&= \mathbb{E}_\nu[\mathbb{E}_\nu[F(\xi_t) - F(\xi_s)|\mathcal{F}_s]\mathbb{1}_A].
\end{aligned}$$

Thus

$$\mathbb{E}\Big[F(\xi_t) - F(\xi_0) - \int_0^t \mathcal{L}_1 F(\xi_s)\,ds\,|\,\mathcal{F}_s\Big] = F(\xi_s) - F(\xi_0) - \int_0^s \mathcal{L}_1 F(\xi_y)\,dy,$$

and therefore $M(\xi_t)$ is a mean zero $\mathbb{P}_\nu$-martingale. □

Definition 2.7. *We say that the measurable random variable ξ, or equivalently its distribution $\mathbb{P}_\nu$, solves the $(\mathcal{L},\nu)$-martingale problem if*

$$\mathbb{P}_\nu\big[\xi_0 = \nu\big] = 1$$

and

$$F\big(\xi_t\big) - F\big(\xi_0\big) - \int_0^t \mathcal{L}F\big(\xi_s\big)ds$$

is a mean-zero $\mathbb{P}_\nu$-martingale for all $F \in \mathcal{D}(\mathcal{L})$.

A classical theorem in the theory of probability due to Stroock and Varadhan [8, Theorem 4.2, Chapter 4, page 184] tell us that if there exist two solutions of a given martingale problem which has the same one-dimensional distributions then this solution is unique, therefore is important understand this characterization for the branching Brownian motion ξ_t to get convergence when $t \to \infty$.

Theorem 2.2 (Characterization via a martingale problem)**.** *The distribution $\mathbb{P}_\nu$ of branching Brownian motion with initial value ν is the unique solution to the $(\mathcal{L}_1, \nu)$ martingale problem in $(\Omega_1, \mathcal{F})$.*

The demonstration of this theorem is an easy modification of the general theorem which proof the uniqueness of Dawson-Watanabe superprocess and therefore we will omit it.

Lemma 2.1. *Let $\{M_t^\theta\}_{\theta \in (0,1)}$ be a family of $(\mathcal{F}_t)$-martingale where $\mathcal{F}_t$ denote the natural σ-algebra. Consider that $\frac{\partial}{\partial\theta} M_t^\theta$ exists almost surely and*

$$\left\| \frac{\partial}{\partial\theta} M_t^\theta \right\|_\infty < \infty. \tag{2.10}$$

Then, $\frac{\partial}{\partial\theta} M_t^\theta|_{\theta=\theta_0}$ is a $\mathcal{F}_t$-martingale for all $\theta_0 \in (0,1)$.

Proof. The process is adapted because it is a limit of adapted processes. The condition (2.10) guarantees that the process is in L_1. It remains to show that for all $s \leq t$, $\mathbb{E}\big[\frac{\partial}{\partial\theta} M_t^\theta|_{\theta=\theta_0}|\mathcal{F}_s\big] = \frac{\partial}{\partial\theta} M_s^\theta|_{\theta=\theta_0}$. Let $A \in \mathcal{F}_s$. Since M_t^θ is a martingale,

$$\int_A \mathbb{E}\big[M_t^\theta|\mathcal{F}_s\big] d\mathbb{P} = \int_A M_t^\theta d\mathbb{P}.$$

differentiating both sides of the above equation with respect to θ, it follows that

$$\frac{\partial}{\partial\theta} \int_A M_s^\theta d\mathbb{P} = \frac{\partial}{\partial\theta} \int_A M_t^\theta d\mathbb{P},$$

and by the Lebesgue Differentiation theorem under the sign of integral [2, Corollary 5.9, page 46] it follows that

$$\int_A \frac{\partial}{\partial\theta} M_s^\theta d\mathbb{P} = \int_A \frac{\partial}{\partial\theta} M_t^\theta d\mathbb{P}.$$

Since the result holds for all θ, particularly it is true for $\theta = \theta_0$. □

We characterize the branching Brownian motion as solution of the martingale problem as follows.

Lemma 2.2. *If ξ solves the $(\mathcal{L}_1, \nu)$-martingale problem, then for each $\phi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$, $\langle \phi, \xi_s \rangle$ is a semimartingale. Moreover,*

$$M_t(\phi) := \langle \phi, \xi_t \rangle - \langle \phi, \xi_0 \rangle - \int_0^t \langle \frac{1}{2}\Delta\phi, \xi_s \rangle \, ds - \int_0^t \langle V\big(\Psi'(1) - 1\big)\phi, \xi_s \rangle \, ds$$

is a mean-zero $\mathbb{P}_\nu$-martingale with quadratic variation

$$\big[M(\phi)\big]_t = \int_0^t \Big\langle 2\nabla\phi.\nabla\phi + 2V\left[\frac{\Psi(\exp(-\phi)) - \exp(-\phi)}{\exp(-\phi)} + \big(\Psi'(1) - 1\big)\phi\right], \xi_s \Big\rangle \, ds\,.$$

Proof. Note that by the Theorem 2.2, the process $\langle \psi, \xi_t \rangle$ is a $\mathbb{P}_\nu$-semimartingale. Indeed,

$$M_t(\psi) \;:=\; F(\xi_t) \;-\; F(\xi_0) \;-\; \int_0^t \mathcal{L}_1 F(\xi_s)\, ds$$

is a mean zero $\mathbb{P}_\nu$-martingale as seen previously and since $F(\xi_0) \;+\; \int_0^t \mathcal{L}_1 F(\xi_s)\, ds$ is a predictable process of bounded variation, it follows that $F(\xi_t)$ is a semimartingale.

Let $F(\xi) \;=\; \exp\left(\langle \log \psi, \xi \rangle\right)$ and take $\psi \;=\; \exp\left(-\theta\phi\right)$ then it follows that $F_\theta(\xi) \;=\; \exp\left(-\langle \theta\phi, \xi \rangle\right)$, and since

$$\begin{aligned} \Delta \exp\left(-\theta\phi\right) \;&=\; \nabla(\nabla \exp\left(-\theta\phi\right)) \;=\; \nabla(\exp\left(-\theta\phi\right)(-\theta\nabla\phi)) \\ &=\; \exp\left(-\theta\phi\right)(-\theta\nabla\phi)^2 \;+\; \exp\left(\theta\phi\right)(-\theta\Delta\phi)\,, \end{aligned}$$

we have that

$$\begin{aligned} \mathcal{L}_1 F_\theta(\xi_s) \;&=\; \left\langle \frac{\frac{1}{2}\Delta\psi \;-\; V\big(\Psi(\psi) - \psi\big)}{\psi}, \xi_s \right\rangle \exp\left(\langle \log(\psi), \xi_s \rangle\right) \\ &= \left\langle \frac{1}{2}\big(- \theta\Delta\phi + \theta^2\nabla\phi.\nabla\phi\big) \;+\; V\frac{\big[\Psi(\exp\left(-\theta\phi\right)) - \exp\left(-\theta\phi\right)\big]}{\exp\left(-\theta\phi\right)}, \xi_s \right\rangle \exp\left(\langle -\theta\phi, \xi_s \rangle\right). \end{aligned} \tag{2.11}$$

Since ξ solves the $(\mathcal{L}_1, \nu)$-martingale problem, it follows that

$$\mathbb{E}_\nu\left[F_\theta(\xi_{t+u}) - F_\theta(\xi_t) - \int_t^{t+u} \mathcal{L}_1 F_\theta(\xi_s)\, ds\right] = 0.$$

Observe that

$$\begin{aligned} &\frac{d}{d\theta}\left\langle \frac{1}{2}\big(- \theta\Delta\phi + \theta^2\nabla\phi.\nabla\phi\big) \;+\; V\frac{\big[\Psi(\exp\left(-\theta\phi\right)) - \exp\left(-\theta\phi\right)\big]}{\exp\left(-\theta\phi\right)}, \xi_s \right\rangle \exp\left(\langle -\theta\phi, \xi_s \rangle\right) \\ &= \left\langle \frac{1}{2}\big(- \Delta\phi + 2\theta\nabla\phi.\nabla\phi\big) \;+\; V\phi\frac{\big[- \Psi'(\exp\left(-\theta\phi\right)) \;+\; \exp\left(-\theta\phi\right)\big]}{\exp\left(-\theta\phi\right)}, \xi_s \right\rangle \exp\left(\langle -\theta\phi, \xi_s \rangle\right) \\ &- \left\langle \frac{1}{2}\big(- \theta\Delta\phi + \theta^2\nabla\phi.\nabla\phi\big) \;+\; V\frac{\big[\Psi(\exp\left(-\theta\phi\right)) - \exp\left(-\theta\phi\right)\big]}{\exp\left(-\theta\phi\right)}, \xi_s \right\rangle \exp\left(\langle -\theta\phi, \xi_s \rangle\right)\langle \phi, \xi_s \rangle. \end{aligned}$$

Thus differentiating the martingale and evaluating it in $\theta \;=\; 0$, since $\Psi(1) \;=\; 1$ it follows that

$$\begin{aligned} \frac{d}{d\theta}\left(F_\theta(\xi_{t+u}) \;-\; F_\theta(\xi_t) \;-\; \int_t^{t+u} \mathcal{L}_1 F_\theta(\xi_s)\, ds\right)\Bigg|_{\theta=0} &= \langle \phi, \xi_t \rangle \;-\; \langle \phi, \xi_0 \rangle \\ &- \int_t^{t+u} \langle \frac{1}{2}\Delta\phi \;-\; V(\Psi'(1) \;-\; 1)\phi, \xi_s \rangle ds. \end{aligned}$$

Thus, by the Lebesgue Differentiation theorem under the sign of integral

$$\begin{aligned} 0 &= \frac{d}{d\theta}\mathbb{E}_\nu\Big[F_\theta(\xi_{t+u}) - F_\theta(\xi_t) - \int_t^{t+u} \mathcal{L}_1 F_\theta(\xi_s)\,ds\Big]\Big|_{\theta=0} \\ &= \mathbb{E}_\nu\Big[\frac{d}{d\theta}\Big(F_\theta(\xi_{t+u}) - F_\theta(\xi_t) - \int_t^{t+u} \mathcal{L}_1 F_\theta(\xi_s)\,ds\Big)\Big|_{\theta=0}\Big] \\ &= \mathbb{E}_\nu\Big[\langle\phi,\xi_t\rangle - \langle\phi,\xi_0\rangle - \int_0^u \langle\frac{1}{2}\Delta\phi - V\big(\Psi'(1) - 1\big)\phi,\xi_s\rangle\,ds\Big]. \end{aligned}$$

This equation is enough to ensure that $M_t(\phi)$ is a mean-zero $\mathbb{P}_\nu$-martingale. Observe that $\frac{d}{d\theta}F_\theta(\xi_t)|_{\theta=0}$ is a semimartingale as previously seen which proves the first statement. Now since the process $\langle\phi,\xi_t\rangle$ is a semimartingale by the Itô's formula [5, Theorem 14, page 135],

$$F(\langle\phi,\xi_t\rangle) = F(\langle\phi,\xi_0\rangle) + \int_0^t F'(\langle\phi,\xi_{s-}\rangle)d\langle\phi,\xi_s^-\rangle + \frac{1}{2}\int_0^t F''(\langle\phi,\xi_{s-}\rangle)d\big[\langle\phi,\xi_s^-\rangle\big]_{s^-}, \tag{2.12}$$

which is a semimartingale for any F twice continuously differentiable. Observe that we can write the semi-martingale as

$$\langle\phi,\xi_t\rangle = M_t(\phi) + \langle\phi,\xi_0\rangle + \int_0^t \langle\frac{1}{2}\Delta\phi - V(\Psi'(1) - 1)\phi,\xi_s\rangle ds,$$

therefore (2.12) becomes

$$\begin{aligned} F(\langle\phi,\xi_t\rangle) &= F(\langle\phi,\xi_0\rangle) \\ &+ \int_0^t F'(\langle\phi,\xi_{s-}\rangle)d\Big(M_t(\phi) + \langle\phi,\xi_0\rangle + \int_0^t \langle\frac{1}{2}\Delta\phi - V(\Psi'(1) - 1)\phi,\xi_s\rangle ds\Big) \\ &+ \frac{1}{2}\int_0^t F''(\langle\phi,\xi_{s-}\rangle)d\Big[M_t(\phi) + \langle\phi,\xi_0\rangle + \int_0^t \langle\frac{1}{2}\Delta\phi - V(\Psi'(1) - 1)\phi,\xi_s\rangle ds\Big]_{s^-}, \end{aligned}$$

and once that $d\langle\phi,\xi_0\rangle = 0$ and $d\Big[\int_0^t\langle\frac{1}{2}\Delta\phi - V(\Psi'(1) - 1)\phi,\xi_s\rangle\Big]_s = 0$, it follows that

$$\begin{aligned} F(\langle\phi,\xi_t\rangle) &= F(\langle\phi,\xi_0\rangle) \\ &+ \int_0^t F'(\langle\phi,\xi_{s-}\rangle)dM_{s^-}(\phi) + \int_0^t F'(\langle\phi,\xi_{s-}\rangle)\langle\frac{1}{2}\Delta\phi - V(\Psi'(1) - 1)\phi,\xi_s\rangle ds \\ &+ \frac{1}{2}\int_0^t F''(\langle\phi,\xi_{s-}\rangle)d[M(\phi)]_{s^-}, \end{aligned}$$

where $\int_0^t F'(\langle \phi, \xi_{s^-}\rangle) dM_{s^-}(\phi)$ is a martingale [4, Theorem 5.4, chapter 5, page 101]. Thus, taking $F(x) = \exp(-x)$, the function F satisfies the Itô's Formula hypothesis, therefore

$$\int_0^t \exp(-\langle \phi, \xi_{s^-}\rangle) dM_{s^-}(\phi) = \exp(-\langle \phi, \xi_t\rangle) - \exp(-\langle \phi, \xi_0\rangle)$$
$$+ \int_0^t \langle \frac{1}{2}\Delta\phi + V(\Psi'(1) - 1)\phi, \xi_s\rangle \exp(-\langle \phi, \xi_s\rangle) ds - \frac{1}{2}\int_0^t \exp(-\langle \phi, \xi_s\rangle) d\big[M(\phi)\big]_s, \tag{2.13}$$

is $\mathbb{P}_\nu$-martingale. Consider $F_\theta(\xi)$ and its generator (2.11). By the Dynkin's formula, the process (2.9) define a martingale $N_t(\phi)$ with $\theta = 1$. Then

$$N_t(\phi) = \exp(-\langle \phi, \xi_t\rangle) - \exp(-\langle \phi, \xi_0\rangle)$$
$$- \int_0^t \left\langle \frac{1}{2}(2\nabla\phi\nabla\phi - \Delta\phi) + V\frac{\Psi'(\exp(-\phi)) - \exp(\phi)}{\exp(-\phi)}, \xi_s \right\rangle \exp(-\langle \phi, \xi_s\rangle) ds, \tag{2.14}$$

which is a mean-zero $\mathbb{P}_\nu$-martingale. Since the decomposition of semimartingale is unique by the Proposition 1.3, equating (2.13) with (2.14) we have that

$$\int_0^t \exp(-\langle \phi, \xi_s\rangle) d\big[M(\phi)\big]_s$$
$$= \int_0^t \exp(-\langle \phi, \xi_s\rangle) 2\left\langle \nabla\phi.\nabla\phi + V\left[\frac{\Psi(\exp(-\phi)) - \exp(-\phi)}{\exp(-\phi)} + (\Psi'(1) - 1)\phi\right], \xi_s \right\rangle ds,$$

which gives the second statement. □

The above Lemma uses a slightly different definition for the quadratic variation than found in classical references like [18] and in this case, they are equivalent, see [24].

Remark 2.10. The Lemma 2.2 is an important piece to show the existence of the Dawson-Watanabe superprocess.

Chapter 3

The Dawson-Watanabe superprocess

In this chapter, we will construct the so-called Dawson-Watanabe superprocess as a limit scale of branching Brownian motion.

3.1 Re-scaling and Tightness

The Dawson-Watanabe superprocess is construct as a re-scaling limit of branching Brownian motion as follows:

Firstly, at the 1-th stage, the re-scaled process is the branching Brownian motion. For the n-th stage of re-scaling, we give to the system $O(n)$ amount of individuals in the following way, if the re-scaled process has size x, then there are nx particles alive. Thus, in the next generation, the number of individuals alive is given by the sum of nx independent Poisson random variables each of them with parameter one, in which each particle has $\frac{1}{n}$ mass. It corresponds to the initial measure ν at time zero.

Each particle of the system has an independent exponential lifetime with parameter nV with a spatial position given by a Brownian motion. Also, observe that at time zero, it is impossible to distinguish if the system is in the n-th or m-th re-scaling stage, for $m \neq n$. Since we normalized the mass of all particles alive at time zero, and each of them behaves according to a Brownian motion, there is no method to distinguish which stage of re-scaling it is. Besides, note that once the lifetime of each particle has parameter nV, each of them dies faster than the usual branching Brownian motion.

As previously seen for the branching Brownian motion, whenever each particle dies, it leaves behind at the same position with a random number of descendants given by the probability generating function Φ. Also, the re-scaling process inherits this property. Finally, for any time t of the re-scaled process, we give nt time to the branching Brownian motion evolve. As previously seen, whenever each particle dies, it leaves behind at the same position with a random number of descendants given by the probability generating function Φ and the re-scaling process inherits this property.

Rigorously, we define the re-scaled process in terms of branching Brownian motion $\xi_t = \sum_{i \in \mathcal{N}_t} \delta_{Y_t^i}$, as following:

Let $\{X^{(n)}\}_{n\geq 1}$ to be the process such that

$$X_0^{(n)} = \frac{1}{n}\xi_0^{(n)}, \tag{3.1}$$

and for any time t,

$$X_t^n = \frac{1}{n}\xi_{nt}^{(n)}. \tag{3.2}$$

Observe that the process at time zero converges to X_0^∞, by the weak law of large numbers [6, Theorem 2.2.9., page 54]. Indeed, since

$$X_0^{(n)} = \frac{1}{n}\sum_{i\in\mathbb{N}_0}\delta_{x_i},$$

and $\{x_i : i \in \mathbb{N}_0\}$ is a Poisson point process which is independent and identically distributed, thus $\{\delta_{x_i} : i \in \mathbb{N}_0\}$ also it is. Moreover $\mathbb{E}[|\delta_{x_1}|] < \infty$, and hence the result follows.

In addition, in agreement with Roelly-Copoleta [22, page 60], we assume that the branching mechanism is critical, i.e., $\Phi'(1) = 1$. Denote the variance by σ^2 which is given by $\Phi''(1)$.

Proposition 3.1. *Suppose that the branching process is critical. Then* $\{\langle 1, X_t^{(n)}\rangle\}_{n\geq 1}$ *is a* $\mathbb{P}^{(n)}$*-martingale.*

Remark 3.1. Note that this result is a kind of variation of the Proposition 2.2, once we suppose that the process is critical, i.e., $\mu_n = 1$ for all $n \in \mathbb{N}$.

Proof. Let denote $Y_t^{(n)} := \langle 1, X_t^{(n)}\rangle$ to be the total mass of continuous-time branching process. Clearly the process is adapted to the natural σ-algebra $\mathcal{F}_{t,n} = \sigma(\langle 1, X_s^n\rangle : 0 \leq s \leq t, n \in \mathbb{N})$ and observe that

$$\mathbb{E}^{(n)}[\left|Y_t^{(n)}\right|] = \mathbb{E}^{(n)}\left[\left|\langle 1, X_t^{(n)}\rangle\right|\right] = \mathbb{E}^{(n)}\left[\langle 1, X_t^{(n)}\rangle\right] < \infty,$$

once the process is critical.

Let $s < t$. Since the branching process is Markovian and Y_t is the total mass at time t,

$$\mathbb{E}^{(n)}[Y_t^{(n)} \mid \mathcal{F}_s] = \mathbb{E}^{(n)}[\theta_s \circ Y_{t-s}^{(n)} \mid \mathcal{F}_s] = \mathbb{E}_{X_s}^{(n)}[Y_{t-s}^{(n)}] = Y_s\mathbb{E}_1^{(n)}[Y_{t-s}^{(n)}] \tag{3.3}$$

Let $B = \{$occur a branching up to time t$\}$, $R = \{$number of particles after branching$\}$. Without loss of generality, take $n = 1$. Thus

$$\begin{aligned} f(t) := \mathbb{E}_1^{(n)}[Y_t^{(n)}] &= \mathbb{E}_1^{(n)}[Y_t^{(n)}\mathbb{1}_{[B>t]} + Y_t^{(n)}\mathbb{1}_{[B\leq t]}\mathbb{1}_{\bigcup_{k\geq 0} R=k]} \\ &= \mathbb{E}_1^{(n)}[1\mathbb{1}_{[B>t]}] + \sum_{k\geq 0}\mathbb{E}_1^{(n)}[Y_t^{(n)}\mathbb{1}_{[B\leq t]}\mathbb{1}_{R=k}] \\ &= \exp(-Vt) + \sum_{k\geq 0} k\mathbb{P}_1[R=k]\int_0^t V\exp(-Vs)f(t-s)\,ds\,. \end{aligned}$$

Now, since the process $\{Y_t^{(n)}\}_{t\geq 0}$ is critical $\sum_{k\geq 0} k\mathbb{P}_1[R = k] = 1$, so we have the following O.D.E.

$$\begin{cases} \frac{d}{dt}f(t) = -V\exp(-Vt) + V\exp(-Vt)f(0) \\ f(0) = 1. \end{cases}$$

Observe that $\frac{d}{dt}f(t) = 0$ for all $t \geq 0$ and since $f(0) = 1$ it follows that $f(t) \equiv 1$. By (3.3), we have that $\{Y_t\}_{t\geq 1}$ is a $\mathbb{P}^{(n)}$-martingale. □

Although the Corollary 1.1 gives us a strong condition for tightness, in real processes, it is not malleable. David Aldous in 1978 [1] has been constructed a helpful criterion to determine if a process is tight. In 1980, Rebolledo et al. [21], make a criterion using semi-martingales to determines if a process is tight or not, and once we had characterized the quadratic variation and the predictable process of finite variation process of the semi-martingale given by the branching Brownian motion, we will use this criterion.

Theorem 3.1 (Aldous-Rebolledo Criterion)**.** *Let $\{Y^{(n)}\}_{n\geq 1}$be a sequence of real valued semimartingales with càdlàg paths. Consider $V^{(n)}$ for the corresponding predictable finite variation process and $[M^{(n)}]_t = \langle M^{(n)}, M^{(n)}\rangle_t$ the quadratic variation of the martingale part of $Y^{(n)}$. Suppose that the following conditions are satisfied.*

(i) For each fixed t, $\{Y_t^{(n)}\}_{n\geq 1}$ is tight.

(ii) Given a sequence of stopping times τ_n, bounded by T, for each $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that

$$\sup_{n\geq n_0}\sup_{\theta\in[0,\delta]} \mathbb{P}\Big[\Big|V^{(n)}(\tau_n + \theta) - V^{(n)}(\tau_n)\Big| > \varepsilon\Big] \leq \varepsilon,$$

and

$$\sup_{n\geq n_0}\sup_{\theta\in[0,\delta]} \mathbb{P}\Big[\Big|[M^{(n)}]_{\tau_n+\theta} - [M^{(n)}]_{\tau_n}\Big| > \varepsilon\Big] \leq \varepsilon.$$

Then, the sequence $\{Y^{(n)}\}_{n\geq 1}$ is tight.

The next theorem, [8, chapter 3, Theorem 9.1, page 142], state a relationship between the tight condition of a process $\{X^{(n)}\}_{n\geq 1}$ and their image $\{f(X^{(n)})\}_{n\geq 1}$ where f belongs to a dense subset Θ of the bounded continuous function which will be useful for working with a relatively simpler class of process.

Theorem 3.2. *Let (E, ρ) be a complete and separable metric space. Consider $\{X^{(n)}\}_{n\geq 1}$ be a family of process with sample path in $D(\mathbb{R}^+, E)$. Suppose that the compact containment condition holds. Let Θ be a dense subset of bounded continuous function with the topology of uniform convergence on compact sets. Then $\{X^{(n)}\}_{n\geq 1}$is relatively compact if and only if $\{f(X^{(n)})\}_{n\geq 1}$ is relatively compact as a family of process in $D_{\mathbb{R}}[0,\infty)$ for each $f \in \Theta$.*

Let $\mathcal{M}_F(S)$ the space of all finite measure defined on S, i.e., $\mathcal{M}_F(\mathbb{R}^d) := \{\mu \in \mathcal{M}_1(\mathbb{R}^d) : |\langle 1, \mu\rangle| < K, \text{ for } K < \infty\}$ is not a compact set in $\mathcal{M}_1(\mathbb{R}^d)$ with the weak topology. Indeed, it is enough take the measure $\mu_n = \delta_n$. We have that $\langle 1, \mu_n\rangle = 1$ for all n. Note that as $n \to \infty$, the mass escape to the infinity and therefore, we cannot extract convergent sub-sequence in $\mathcal{M}_F(\mathbb{R}^d)$.

To deal with this problem, we consider the space $\hat{\mathbb{R}}^d$ which is the one point compactification of $\mathbb{R}^d$. We equip $\hat{\mathbb{R}}^d$ with the usual topology with addition to the infinity point, i.e., we consider the sets of the form $V \cup \{\infty\}$ where $V^{\complement}$ is compact set in $\mathbb{R}^d$.

Theorem 3.3. *Let $\{X^{(n)}\}_{n\geq 1}$ the sequence of re-scaled process defined as* (3.1) *and* (3.2) *such that the branching mechanism is critical. Then $\{X^{(n)}\}_{n\geq 1}$ is tight in $\mathcal{M}_F(\hat{\mathbb{R}^d})$.*

Remark 3.2. Dividing the proof into two parts: Firstly, we will check the compact containment condition 1.17 to achieve the tight condition of $\{X_t^{(n)}\}_{n\geq 1}$ for each t by the theorem 3.2. After that, we will show the hypothesis of Aldous-Rebolledo Criterion 3.1 to conclude the proof.

Proof. Define $Y_t^{(n)} = \langle \phi, X_t^{(n)} \rangle$ for any ϕ in a dense subset of $C_b^+(\hat{\mathbb{R}}^d)$. Observe that to show the compact containment condition for $Y_t^{(n)}$ it is equivalent to show for $\langle 1, X_t^{(n)} \rangle$ once $\|\phi\|_\infty < \infty$. Then, for each fixed $T > 0$, we want to show that given $\varepsilon > 0$ there exists $K > 0$ such that

$$\mathbb{P}\Big[\sup_{0\leq t\leq T} \langle 1, X_t^{(n)} \rangle \leq K\Big] \geq 1 - \varepsilon.$$

By the Proposition 3.1 the process $\{\langle 1, X_0^{(n)} \rangle\}_{n\geq 1}$ is a $\mathbb{P}^{(n)}$-martingale and by the Doob's inequality we have that

$$\mathbb{P}^{(n)}\Big[\sup_{0\leq t\leq T} \langle 1, X_t^{(n)} \rangle > K\Big] \leq \frac{1}{K}\mathbb{E}^{(n)}\big[\langle 1, X_t^{(n)} \rangle\big] = \frac{1}{K}\mathbb{E}^{(n)}[\langle 1, X_0^{(n)} \rangle].$$

Now, since $\{X_0^{(n)}\}_{n\geq 1}$ converges almost surely, $\langle 1, X_0^{(n)} \rangle < \infty$ for all $n \in \mathbb{N}$. Taking $K \to \infty$, the right side of the above equation tends to zero uniformly in n. We show that the hypothesis of the theorem 3.2 is satisfied, moreover, we show that the process $\{\langle \phi, X_t^{(n)} \rangle\}_{n\geq 1}$ is tight for each t fixed.

Let $\{\tau_n\}_{n\geq 1}$ be a sequence of stopping times bounded by n. By the Lemma 2.2, since $\{\langle \psi, \xi_t \rangle\}_{t\in\mathbb{R}}$ is a $\mathbb{P}_\nu$-semimartingale for all $\psi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$. Observe that

$$\langle \psi, X_t^{(n)} \rangle = \langle \psi, \frac{1}{n}\xi_{nt}^{(n)} \rangle = \langle \frac{1}{n}\psi, \xi_{nt}^{(n)} \rangle$$

and it follows that the process $\{\langle \psi, X_t^{(n)} \rangle\}_{t\in\mathbb{R}}$ is a $\mathbb{P}_\nu$-semimartingale. Using that $\Phi'(1) = 1$, its predictable process is given by

$$V^{(n)}(t) = \langle \psi, X_0^{(n)} \rangle - \int_0^t \langle \frac{1}{2}\Delta\psi, X_s^{(n)} \rangle \, ds,$$

and, since

$$\begin{aligned}\Delta(\psi^2) - 2\psi\Delta\psi &= \nabla(2\psi\nabla\psi) - 2\psi\Delta\psi \\ &= 2\nabla\psi\nabla\psi + 2\psi\Delta\psi - 2\psi\Delta\psi = 2\nabla\psi\nabla\psi\,,\end{aligned}$$

the process has the following quadratic variation $[M(\psi)]_t$ of the martingale part

$$[M(\psi)]_t = \int_0^t \langle \Delta(\psi^2) - 2\psi\Delta\psi + 2V\left(\frac{\Psi(\exp(-\psi)) - \exp(-\psi)}{\exp(-\psi)}\right), X_s^{(n)} \rangle \, ds.$$

Thus, given $\varepsilon > 0$ let $\delta = \frac{2\varepsilon^2}{\|\Delta\psi\|_\infty \langle 1, X_0^{(\infty)}\rangle}$ where $X_0^{(\infty)}$ will denotes the limit process of $\{X_0^{(n)}\}_{n\geq 1}$. Hence

$$\begin{aligned}
\sup_{0\leq\theta\leq\delta} \mathbb{P}^{(n)}\left[\left|V^{(n)}(\tau_n+\theta) - V^{(n)}(\tau_n)\right| > \varepsilon\right] &= \sup_{0\leq\theta\leq\delta} \mathbb{P}^{(n)}\left[\left|\int_{\tau_n}^{\tau_n+\theta} \langle \frac{1}{2}\Delta\psi, X_s^{(n)}\rangle\, ds\right| > \varepsilon\right] \\
&\leq \sup_{0\leq\theta\leq\delta} \mathbb{P}^{(n)}\left[\int_{\tau_n}^{\tau_n+\theta} \left|\langle \frac{1}{2}\Delta\psi, X_s^{(n)}\rangle\right| ds > \varepsilon\right] \\
&\leq \sup_{0\leq\theta\leq\delta} \mathbb{P}^{(n)}\left[\frac{1}{2}\|\Delta\psi\|_\infty \int_{\tau_n}^{\tau_n+\theta} \langle 1, X_s^{(n)}\rangle ds > \varepsilon\right] \\
&\leq \mathbb{P}^{(n)}\left[\frac{1}{2}\|\Delta\psi\|_\infty\, \delta \sup_{0\leq t\leq n+\delta} \langle 1, X_t^{(n)}\rangle > \varepsilon\right].
\end{aligned} \tag{3.4}$$

Since $\langle 1, X_t^{(n)}\rangle$ is a $\mathbb{P}^{(n)}$-martingale, we apply the Doob's inequality and (3.4) is bounded by $\frac{\|\Delta\psi\|_\infty \delta\langle 1, X_0^{(\infty)}\rangle}{2\varepsilon}$, so the predictable process has variation in probability lesser than ε. Now, we do the same argument to the quadratic variation of the martingale part it follows that

$$\begin{aligned}
\sup_{0\leq\theta\leq\delta} \mathbb{P}^{(n)}\left[\left|[M^{(n)}]_{\tau_n+\theta} - [M^{(n)}]_{\tau_n}\right| > \varepsilon\right] = \sup_{0\leq\theta\leq\delta} \mathbb{P}^{(n)}\Bigg[\Bigg|\int_0^t \langle \Delta(\psi^2) - 2\psi\Delta\psi \\
+ 2V\left(\frac{\Psi(\exp(-\psi)) - \exp(-\psi)}{\exp(-\psi)}\right), X_s^{(n)}\rangle\, ds\Bigg| > \varepsilon\Bigg].
\end{aligned}$$

Denote $B = \|\Delta\psi^2\|_\infty + \|2\psi\Delta\psi\|_\infty + \left\|2V\frac{\Psi(\exp(-\psi)) - \exp(-\psi)}{\exp(-\psi)}\right\|_\infty$ which is bounded once that $\psi \in C_b^+(\mathbb{R}^d)\cap \mathcal{D}(\Delta)$.

So as the same way that we did to the predictable process it follows that the quadratic variation of the martingale part is bounded,

$$\sup_{0\leq\theta\leq\delta} \mathbb{P}^{(n)}\left[\left|[M^{(n)}]_{\tau_n+\theta} - [M^{(n)}]_{\tau_n}\right| > \varepsilon\right] \leq \mathbb{P}^{(n)}\left[B\delta \sup_{0\leq\theta\leq\delta} \langle 1, X_t^{(n)}\rangle > \varepsilon\right] \leq \frac{B\delta}{\varepsilon}\langle 1, X_0^{(n)}\rangle.$$

Taking $\delta = \frac{\varepsilon^2}{B\langle 1, X_0^{(\infty)}\rangle}$ we conclude by the Aldous-Rebolledo criterion 3.1 that the process $\{\langle\psi, X^{(n)}\rangle\}$ is tight. □

Observe that the re-scaled process is relatively compact with the weak topology in $\mathcal{M}_F(\hat{\mathbb{R}}^d)$. Indeed, by the theorem 3.3 $\{X^{(n)}\}_{n\geq 1}$ is tight and, since the space $\mathcal{M}_F(\hat{\mathbb{R}}^d)$ is Polish, by the Prohorov's Theorem 1.7, it follows that $\{X^{(n)}\}_{n\geq 1}$ is relatively compact in its space. Now we will identify the limit points of the sequence as solutions to a martingale problem.

Note that $nX^{(n)}$ is a branching Brownian motion with branch rate nV hence we can apply the Lemma 2.2 to characterize the predictable process and quadratic variation of the martingale part given by the Theorem 2.2 to some function $F \in \mathcal{D}(\mathcal{L}_1)$.

Define the test functions $f_n = 1 - \frac{\phi}{n}$ to some $\phi \in C_b^+(\mathbb{R}^d)\cap\mathcal{D}(\Delta)$ and observe that $f_n \in C_b^+(\mathbb{R}^d)\cap\mathcal{D}(\Delta)$. For each $n \in \mathbb{N}$ define $F_n(\cdot) = \exp(\langle \log f_n, \cdot\rangle)$. We know that for each $n \in \mathbb{N}$ the process

$$\begin{aligned}&\exp\left(\langle \log f_n, nX_t^{(n)}\rangle\right) - \exp\left(\langle \log f_n, nX_0^{(n)}\rangle\right)\\&- \int_0^t \left\langle \frac{\frac{1}{2}\Delta f_n + nV(\Psi(f_n) - f_n)}{f_n}, nX_s^{(n)}\right\rangle \exp\left(\langle \log f_n, nX_s^{(n)}\rangle\right) ds\end{aligned} \tag{3.5}$$

is a mean-zero $\mathbb{P}^{(n)}$-martingale. We know that $\Psi(1) = \sum_{k=0}^{\infty} p_k(1)^k = 1$. Making the Taylor's expansion for the probability generating function centering at 1 it follows that

$$\begin{aligned}\Psi(f_n) = \Psi\Big(1 - \frac{\phi}{n}\Big) &= \Psi(1) - \frac{\phi}{n}\Psi'(1) + \frac{\phi}{2n^2}\Psi''(1) + o\Big(\frac{1}{n^2}\Big)\\&= 1 - \frac{\phi}{n} + \frac{\phi^2}{2n^2}\sigma^2 + o\Big(\frac{1}{n^2}\Big).\end{aligned} \tag{3.6}$$

Thus, the $\mathbb{P}^{(n)}$-martingale (3.5) becomes

$$\begin{aligned}&\exp\left(\left\langle n\log\left(1 - \frac{\phi}{n}\right), X_t^{(n)}\right\rangle\right) - \exp\left(\left\langle n\log\left(1 - \frac{\phi}{n}\right), X_0^{(n)}\right\rangle\right)\\&- \int_0^t \left\langle \frac{n\frac{1}{2}\Delta(1 - \frac{\phi}{n}) + n^2V\Big(1 - \frac{\phi}{n} + \frac{\phi^2}{2n^2}\sigma^2 - (1 - \frac{\phi}{n})\Big)}{(1 - \frac{\phi}{n})}, X_s^{(n)}\right\rangle\\&\times \exp\left(\langle n\log\left(1 - \frac{\phi}{n}\right), X_s^{(n)}\rangle\right) ds\end{aligned}$$

and since $\left(1 - \frac{\psi}{n}\right)^n \to \exp(-\psi)$, as $n \to \infty$ it follows that (3.5) converges, and its limit is given by

$$\exp\left(-\langle\phi, X_t\rangle\right) - \exp\left(-\langle\phi, X_0\rangle\right) - \int_0^t \left\langle -\frac{1}{2}\Delta\phi + \frac{1}{2}V\sigma^2\phi^2, X_s\right\rangle \exp\left(-\langle\phi, X_s\rangle\right) ds\,. \tag{3.7}$$

We claim that the expression (3.7) is a $\mathbb{P}^{(n)}$-martingale. Observe that, if A is the limit of the equation (3.5) and X is the process (3.7) it follows that

$$\mathbb{E}^{(n)}[|A - X|] = \lim_{n\to\infty} o\Big(\frac{1}{n^2}\Big) = 0, \tag{3.8}$$

Therefore we have a well approximation for martingale of the limit of rescaled branching Brownian motion.

The next theorem (c.f. [8, Theorem 4.8.10, page 234]) will guarantee that the limit of (3.7) converges to some process in $\mathcal{M}_F(\hat{\mathbb{R}}^d)$ which is the solution of martingale problem in $D(\mathbb{R}^+, \mathcal{M}_F(\hat{\mathbb{R}}^d))$, in other words, the martingale property is preserved under passage to the limit.

Theorem 3.4. *Let (E,ρ) be complete and separable metric space. Let $A \subset C_b(E) \times C_b(E)$ and $\nu \in \mathcal{P}(E)$ and suppose that the $D(\mathbb{R}^+, E)$ martingale problem for (A,ν) has at most one solution. Suppose that X_n is $\mathcal{F}_t^n$-adapted process with sample path in $D(\mathbb{R}^+, E)$, $\{X_n\}_{n\in\mathbb{N}}$ is relatively compact, $\mathbb{P}X_n(0)^{-1} \Rightarrow \nu$, and $M \subset C_b(E)$ is separating. Then, the following statements are equivalent:*

(i) There exists a solution X of the (A,ν)-martingale problem in $D(\mathbb{R}^+, E)$, and $X_n \Rightarrow X$.

(ii) There exists a countable set $\Gamma \subset [0,\infty)$ *such that for each* $(f,g) \in A$,

$$\lim_{n\to\infty} \mathbb{E}\Big[\Big(f(X_n(t+s)) - f(X_n(t)) - \int_t^{t+s} g(X_n(u)du)\Big)\prod_{j=1}^{k} h_j(X_n(t_j))\Big] = 0$$

for all $k \geq 0$, $0 \leq t_1 < \ldots < t_k \leq t < t+s$ *with* $t_i, t, t+s \notin \Gamma$ *and* $h_1, \ldots h_k \in M$.

Observe that the re-scaled branching Brownian motion $\{X^{(n)}\}_{n\in\mathbb{N}}$ satisfies the hypothesis of Theorem 3.4. Let us check that $\{X^{(n)}\}_{n\in\mathbb{N}}$ satisfies the second hypothesis for test function f_n as previously defined.

Let $0 < t_1 < \ldots < t_k \leq t < t+s$ as in theorem and $h_1, \ldots, h_k \in C_b(\mathcal{M}_F(\hat{\mathbb{R}}^d))$. Thus, the second condition is translated to

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\Big[\Big(&\exp\Big(\Big\langle n\log\Big(1-\frac{\phi}{n}\Big), X_{t+s}^{(n)}\Big\rangle\Big) - \exp\Big(\Big\langle n\log\Big(1-\frac{\phi}{n}\Big), X_t^{(n)}\Big\rangle\Big) \\ &- \int_t^{t+s}\Big\langle \frac{n\frac{1}{2}\Delta(1-\frac{\phi}{n}) + n^2V\Big(1-\frac{\phi}{n}+\frac{\phi}{2n^2}\sigma^2-(1-\frac{\phi}{n})\Big)}{(1-\frac{\phi}{n})}, X_s^{(n)}\Big\rangle \\ &\times \exp\big(\langle n\log(1-\frac{\phi}{n}), X_s^{(n)}\rangle\big)\,ds\Big)\prod_{j=1}^{k} h_j\Big(X_{t_j}^{(n)}\Big)\Big].\end{aligned} \tag{3.9}$$

Since $h_j(X_{t_j}^{(n)}) \in \mathcal{F}_t^n$ and the process (3.9) of re-scaled branching Brownian motion inside the expectation is expressed in terms of Lemma 1.4, it follows that the expectation is equal to zero because the process is, for each n, a mean-zero $\mathbb{P}^{(n)}$-martingale.

Now, observe that each h_j is bounded, and since the process inside of conditional expectation is also bounded, by the Dominated Convergence Theorem, we conclude that the limit of the expectation is zero. Using the Theorem 3.4 it follows that exists a solution to the martingale problem for the Dawson-Watanabe measure X such that $X^{(n)} \Rightarrow X$ as $n \to \infty$.

Lemma 3.1. *Let* $T_t\psi$ *the heat semi-group for* $\psi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$ *and* X *the Dawson-Watanabe measure. Then* $\mathbb{E}[\langle\psi, X_t\rangle] = \mathbb{E}[\langle T_t\psi, X_0\rangle]$.

Proof. Let $F(\cdot) = \exp(\langle\log\phi, \cdot\rangle)$ to $\phi \in C_b^+(\mathbb{R}^d)\cap\mathcal{D}(\Delta)$. Since the infinitesimal generator of the branching Brownian motion is the same as the Dawson-Watanabe superprocess it follows that

$$\frac{d}{dt}\mathbb{E}[F(X_t)] = \mathbb{E}\Big[\exp(\langle\log\phi, X_t\rangle)\Big\langle\frac{-\frac{1}{2}\Delta\phi + V[\Psi(\phi)-\phi]}{\phi}, X_t\Big\rangle\Big]. \tag{3.10}$$

Let $\phi = \exp(-\theta\psi)$ for some $\psi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$. Thus

$$\begin{aligned}\frac{d}{d\theta}\exp\left(\langle \log\phi(\theta), X_t\rangle\right)\Big\langle \frac{-\frac{1}{2}\Delta\phi(\theta) + V[\Psi(\phi(\theta)) - \phi(\theta)]}{\phi(\theta)}, X_t\Big\rangle &= \frac{d}{d\theta}\exp\left(\langle -\theta\psi, X_t\rangle\right) \\ &\times \Big\langle \frac{-\frac{1}{2}\Delta\exp(-\theta\psi) + V[\Psi(\phi(\theta)) - \phi(\theta)]}{\exp(-\theta\psi)}, X_t\Big\rangle \\ = -\langle\psi, X_t\rangle \exp\left(-\langle\theta\psi, X_t\rangle\right)\Big\langle \frac{-\frac{1}{2}\Delta\exp(-\theta\psi) + V[\Psi(\exp(-\theta\psi)) - \exp(-\theta\psi)]}{\exp(-\theta\psi)}, X_t\Big\rangle \\ + \exp\left(-\langle\theta\psi, X_t\rangle\right)\Big\langle -\frac{1}{2}(2\theta\nabla\psi\nabla\psi - \Delta\psi) + V\phi\frac{\left[-\Psi'(\exp(-\theta\phi)) + \exp(-\theta\phi)\right]}{\exp(-\theta\phi)}, X_t\Big\rangle.\end{aligned} \tag{3.11}$$

Evaluating (3.11) at $\theta = 0$ and since $\Psi'(1) = 1$ it follows by the Lebesgue Differentiation Theorem under the sign of Integral that

$$\mathbb{E}\Big[\langle\frac{1}{2}\Delta\psi, X_t\rangle\Big] = \frac{d}{d\theta}\frac{d}{ds}\mathbb{E}\Big[F_\theta(X_s)\Big]\Big|_{\theta=0} = \frac{d}{ds}\mathbb{E}\Big[\frac{d}{d\theta}F_\theta(X_s)\Big|_{\theta=0}\Big].$$

Observe that, again by the semi-group theory argument and by the Fubini's Theorem, integrating the time s over $[0,t]$, it follows that

$$\mathbb{E}\Big[\int_0^t \langle\frac{1}{2}\Delta\psi, X_s\rangle ds\Big] = \mathbb{E}\Big[\frac{d}{d\theta}F_\theta(X_t)\Big|_{\theta=0} - \frac{d}{d\theta}F_\theta(X_0)\Big|_{\theta=0}\Big] = \mathbb{E}[\langle\psi, X_t\rangle - \langle\psi, X_0\rangle].$$

Therefore

$$\mathbb{E}[\langle\psi, X_t\rangle] = \mathbb{E}\Big[\langle\psi, X_0\rangle + \int_0^t \langle\frac{1}{2}\Delta\psi, X_s\rangle\, ds\Big].$$

□

Now we will show that the process takes it is values in the smaller space. Consider $\phi = \mathbb{1}_{\{\|x\| > R\}}$ for some $R \in \mathbb{R}^+$. Thus, if X denotes the Dawson-Watanabe measure, by the Markov inequality,

$$\mathbb{P}[X_t\{\|x\| > R\} > \varepsilon] = \mathbb{P}[\langle\phi, X_t\rangle > \varepsilon] \le \frac{1}{\varepsilon}\mathbb{E}[\langle\phi, X_t\rangle]. \tag{3.12}$$

So, by the Lemma 3.1 the equation 3.12 is bounded by $\frac{1}{\varepsilon}\mathbb{E}[\langle T_t\mathbb{1}_{\{\|x\| > R\}}, X_0\rangle]$, which converges to zero as $R \to \infty$. We give an intuition about that: Observe that at time zero, since the process is a Poisson point process, there is no infinite cluster. Thus whenever R tends to infinity, the indicator function ϕ is almost surely equal to zero and, only the particles that are to infinity will be considered.

Remark 3.3. A strong result shows that the Dawson-Watanabe measure has compact support [7, Corollary 6.8, page 110].

Now, we can conclude that the Dawson-Watanabe superprocess is a measure defined in $\mathcal{M}_F(\mathbb{R}^d)$ equipped with the weak topology.

3.2 Uniqueness of solution of the martingale problem

To show the uniqueness of the Dawson-Watanabe superprocess, we will characterize the Dawson-Watanabe martingale problem as was done to the branching Brownian motion, we will do the same kind of logic and then, we will show that they has a dual deterministic process which have unique solution.

Proposition 3.2. *The measure $\mathbb{P}_{X_0} \in \mathcal{M}_1(\mathcal{M}_F(\hat{\mathbb{R}}^d))$ solves the $(\mathcal{L}, X_0)$-martingale problem, then for each $\phi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$, $\langle \phi, X_t \rangle$ is a semimartingale. Moreover*

$$N_t(\phi) = \langle \phi, X_t \rangle - \langle \phi, X_0 \rangle - \int_0^t \langle \frac{1}{2}\Delta\phi, X_s \rangle ds$$

is $\mathbb{P}_{X_0}$-martingale and its quadratic variation is given by

$$[N(\phi)]_t = V\sigma^2 \int_0^t \langle \phi^2, X_s \rangle ds.$$

Remark 3.4. The proof of this proposition is quite similar to the demonstration of Lemma 2.2. Moreover, the infinitesimal generator $\mathcal{L}$ is $\mathcal{L}_1$ with the Taylor expansion of the probability generating function (3.6), critical branching mechanism and finite variance $\Psi''(1) = \sigma^2$.

Proof. Let ψ be the function taken in the test function to be equal to $\psi = \exp(-\theta\phi)$ for $\theta \in (0, 1)$ and $\phi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$. Consider $F_\theta(\cdot) = \exp(\langle \log\psi(\theta), \cdot \rangle)$. Since $\mathbb{P}_{X_0}$ solves the $(\mathcal{L}, X_0)$-martingale problem, it follows that

$$\mathbb{E}_{X_0}\big[F_\theta(X_t) - F_\theta(X_0) - \int_0^t \mathcal{L}F_\theta(X_s)ds\big] = 0\,.$$

We know that the derivative with respect to θ of above equation is also a martingale thus differentiate and make $\theta = 0$ it follows that $N_t(\phi)$ is a $\mathbb{P}_{X_0}$-martingale and consequently $\frac{d}{d\theta}F_\theta(X_t)|_{\theta=0}$ is a semimartingale. Since

$$\langle \phi, X_t \rangle = N_t(\phi) + \langle \phi, X_0 \rangle + \int_0^t \langle \frac{1}{2}\Delta\phi, X_s \rangle ds$$

and $F(x) = \exp(-x)$ by Itô's formula we have that

$$\begin{aligned} \int_0^t \exp(-\langle \phi, X_{s^-} \rangle) dN_{s^-}(\phi) &= \exp(-\langle \phi, X_t \rangle) - \exp(-\langle \phi, X_0 \rangle) \\ &+ \int_0^t \langle \frac{1}{2}\Delta\phi, \xi_s \rangle \exp(-\langle \phi, \xi_s \rangle) ds - \frac{1}{2}\int_0^t \exp(-\langle \phi, \xi_s \rangle) d\big[N(\phi)\big]_s \end{aligned} \tag{3.13}$$

is $\mathbb{P}_{X_0}$-martingale. By the uniqueness of decomposition of semimartingale and equating (3.13) with (3.7) it follows that

$$\frac{1}{2}\int_0^t \exp(-\langle \phi, \xi_s \rangle) d\big[N(\phi)\big]_s = \int_0^t \exp(-\langle \phi, X_s \rangle) \langle \frac{1}{2}V\sigma^2\phi^2, X_s \rangle ds$$

which proves the second statement. ∎

An interesting assumption that we can make is asymptotically approximating the expected number of descendants. Consider $f \in C_b(\mathbb{R}^d)$ which depends of the spatial position of individual, and make the following approximation in the expectation. $\Phi'(1) = 1 + \frac{f}{n}$. So the equation (3.7) will convert to

$$\exp(-\langle\phi, X_t\rangle) - \exp(-\langle\phi, X_0\rangle) - \int_0^t \Big\langle -\frac{1}{2}\Delta\phi + \frac{1}{2}Vf\sigma^2\phi, X_s\Big\rangle \exp(-\langle\phi, X_s\rangle)ds\,.$$

Following the same reasoning, we will have that the sequence will be rigid and any limit point solves the martingale problem

$$M_t(\phi) = \langle\phi, X_t\rangle - \langle\phi, X_0\rangle - \int_0^t \langle\frac{1}{2}\Delta\phi + fV\phi, X_s\rangle ds.$$

with same quadratic variation given in Proposition 3.2.

Let us show that the process is unique using the method of duality which relates two processes. Both of them will have the same evolution in time, and if one process has a unique solution, the other also will have.

Theorem 3.5 (The method of duality). *Let E_1, E_2 metric spaces and suppose that $\mathbb{P}_1$ and $\mathbb{P}_2$ (equivalently X and Y are stochastic processes) are distribution on the Skorohod space $D(\mathbb{R}^+, E_1)$ and $D(\mathbb{R}^+, E_2)$ respectively. Let f, g bounded functions defined on $E_1 \times E_2$ that satisfies*

(i) For each $y \in E_2$, $f(\cdot, y)$ and $g(\cdot, y)$ are continuous functions on E_1.

(ii) For each $x \in E_2$, $f(x, \cdot)$ and $g(x, \cdot)$ are continuous functions on E_2.

(iii) For each $y \in E_2$,

$$M_y(X) := f(X(t), y) - \int_0^t g(X(s), y)ds$$

is a $\mathbb{P}_1$-martingale.

(iv) For each $x \in E_1$

$$M_x(Y) := f(x, Y(s)) - \int_0^t g(x, Y(s))ds$$

is a $\mathbb{P}_1$-martingale.

Then

$$\mathbb{E}^{\mathbb{P}_1}_{X(0)}\big[f(X(t), Y(0))\big] = \mathbb{E}^{\mathbb{P}_2}_{Y(0)}\big[f(X(0), Y(t))\big]\,. \tag{3.14}$$

Proof. Define $h : \mathbb{R}^2 \to \mathbb{R}$ such that $h(t_1, t_2) = \mathbb{E}^{\mathbb{P}_1 \times \mathbb{P}_2}[f(X(t_1), Y(t_2))]$ and for each $t \in \mathbb{R}$ let $g_t : [0, t] \to [0, t]^2$ defined as $g_t(s) = (s, t - s)$. Observe that, since both functions h and g_t are continuous differentiable, it follows by the chain rule that

$$\frac{d}{ds} h \circ g_t(s) = \frac{d}{ds}\mathbb{E}^{\mathbb{P}_2}[f(X(s), Y(t-s))] - \frac{d}{ds}\mathbb{E}^{\mathbb{P}_1}[f(X(s), Y(t-s))]\,. \tag{3.15}$$

Observe that for each $y \in E_2$ fixed, we have that

$$
\begin{aligned}
\frac{d}{ds}\mathbb{E}^{\mathbb{P}_1}[f(X(t),y)] &= \frac{d}{ds}\mathbb{E}^{\mathbb{P}_1}\Big[M_y(X) + \int_0^t g(X(s),y)ds\Big] \\
&= \frac{d}{ds}\mathbb{E}^{\mathbb{P}_1}[M_y(X)] + \frac{d}{ds}\mathbb{E}^{\mathbb{P}_1}\Big[\int_0^t g(X(s),y)ds\Big] \\
&= \mathbb{E}^{\mathbb{P}_1}[g(X(t),y)]
\end{aligned}
$$

where we have used Fubini's Theorem in the second equality and the fact that the expectation of a martingale is constant to obtain the identity above. In the same way,

$$
\frac{d}{ds}E^{\mathbb{P}_2}[f(x,Y(s))] = \mathbb{E}^{\mathbb{P}_2}[g(x,Y(s))]
$$

for each $x \in E_1$ fixed. Therefore (3.15) equal to zero. Thus

$$
\frac{d}{ds}\mathbb{E}^{\mathbb{P}_1}[f(X(t),y)] = \frac{d}{ds}E^{\mathbb{P}_2}[f(x,Y(s))]. \tag{3.16}
$$

Integrating (3.16) with respect to s over $[0,t]$ it follows that

$$
\int_0^t \frac{d}{ds}\mathbb{E}^{\mathbb{P}_1}[f(X(s),Y(t-s))]ds = \int_0^t \frac{d}{ds}\mathbb{E}^{\mathbb{P}_2}[f(X(s),Y(t-s))]ds \tag{3.17}
$$

and this equality is enough to conclude. □

Proposition 3.3. *The equation* (3.14) *is sufficient to guarantee uniqueness of solutions of martingale problem.*

Proof. Let X be a solution of a martingale problem and Y it is the dual process of X such that both satisfies the hypothesis of the Theorem 3.5. Moreover, suppose that X' is another solution to the martingale problem, which also satisfies the hypothesis of the theorem it follows that $X(0) = X'(0)$, and for all $t \geq 0$

$$
\mathbb{E}^{\mathbb{P}_1}[f(X(t),Y(0))] = \mathbb{E}^{\mathbb{P}_2}[f(X(0),Y(t))] = \mathbb{E}^{\mathbb{P}_2}[f(X'(0),Y(t))] = \mathbb{E}^{\mathbb{P}_1}[f(X'(t),Y(0))].
$$

Thus, by the Portmanteau Theorem (see [4, page 24]), in terms of the Theorem, making the sequence identically to X it follows that the distribution of X converges weakly to X', i.e., they have the same distribution, and hence the solution of the martingale problem is unique. □

Remains to show that the Dawson-Watanabe have a dual deterministic process which satisfies, with the solution of the martingale problem, the equation (3.14).

Corollary 3.1. *The Dawson-Watanabe superprocess has a deterministic dual process prescribed as follows: For $\phi \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$,*

$$
\mathbb{E}[\exp(-\langle \phi, X_t\rangle)] = \exp(-\langle u(t,\cdot), X_0\rangle),
$$

where u solves

$$
\begin{cases}
\frac{\partial}{\partial t}u = \frac{1}{2}\Delta u - \frac{1}{2}V\sigma^2 u^2 \\
u(0,x) = \phi(x).
\end{cases} \tag{3.18}
$$

Consequently the solution of the Dawson-Watanabe martingale problem is unique.

Remark 3.5. (i) The equation (3.18) is often called the evolution equation for the Dawson-Watanabe superprocess.

(ii) Observe that the Theorem 2.2 is a particular case of the Corollary 3.1 and therefore, it follows that the solution of the martingale problem of branching Brownian motion is also unique.

Proof. Let Y to be the solution of the partial differential equation (3.18) and X to be the solution of Dawson-Watanabe martingale process. Consider $f(x,y) = \exp(-\langle y,x\rangle)$ and $g(x,y) = \langle \frac{\partial}{\partial t}y, x\rangle \exp(-\langle y,x\rangle)$. Observe that for each $x \in \mathcal{M}_1(\mathcal{M}_F(\mathbb{R}^d))$ and for each $y \in C_b^+(\mathbb{R}^d) \cap \mathcal{D}(\Delta)$, $f(y,\cdot)$, $g(y,\cdot)$ and $f(\cdot,x)$, $g(\cdot,x)$ are continuous.

Define

$$M_t(X) := f(X(t),Y(0)) - \int_0^t g(X(t),Y(0))ds,$$

and

$$N_t(Y) := f(X(0),Y(t)) - \int_0^t g(X(0),Y(s))ds.$$

Let us show that $M_t(X)$ and $N_t(Y)$ are martingales with their respectively distributions. Observe that $M_t(X)$ is given by (3.7) which is a $P^{(n)}$-martingale. For show that $N_t(Y)$ is a martingale, consider the natural σ-algebra $\mathcal{F}_t$. Then $N_t(Y)$ clearly is adapted. Since f and g are continuous bounded function the process $N_t(Y)$ is L_1. Define $h(t) := N_t(Y)$ and observe that Y is deterministic. Then, is enough to show that $h \equiv 0$. Observe that

$$h(t) = \exp(-\langle u_t, X_0\rangle) - \exp(-\langle \phi, X_0\rangle) - \int_0^t \langle -\frac{1}{2}\Delta u_s + \frac{1}{2}V\sigma^2 u_s^2, X_0\rangle \exp(-\langle u_s, X_0\rangle)ds.$$

We have the following O.D.E.,

$$\begin{cases} \frac{d}{dt}h(t) = -\exp(-\langle u_t, X_0\rangle)\frac{\partial}{\partial t}\langle u_t, X_0\rangle + \langle \frac{1}{2}\Delta u_t - \frac{1}{2}V\sigma^2 u_t^2, X_0\rangle \exp(-\langle u_t, X_0\rangle), \\ h(0) = 0. \end{cases}$$

By equation (3.18) it follows that $\frac{d}{dt}h \equiv 0$. Using the initial condition, we have that $h \equiv 0$ and since u is a deterministic process $N_t(Y)$ is a martingale. By the Theorem 3.5 we finish the prove. □

Appendix A

Some extra standard tools

A.1 Generating function

The probability generating function is a useful tool for dealing with discrete random variables and, therefore, a discrete process. In general, it is difficult to find the distribution of a sum using the traditional probability function. The probability generating function transforms a sum into a product which makes it easier to deal with such a problem.

Definition A.1. *Let X be a random variable taking values in the non-negative integer $\{0, 1, \ldots\}$. The probability generating function (PGF) of X is $G_X(s) = \mathbb{E}[s^X]$, for all $s \in \mathbb{R}$ for which the sum converges.*

Theorem A.1. *Let X be a discrete random variable with PGF $G_X(s)$. Then*

(i) $\mathbb{E}[X] = G'_X(s)$,

(ii) $\mathbb{E}[X(X-1)(X-2)\cdots(X-k+1)] = \frac{d^k}{ds^k} G_X(s)|_{s=0}$.

Theorem A.2. *Let $X_1, X_2, \ldots, X_k$ independent random variables and let $Y = X_1 + \ldots, X_k$. Then*

$$G_Y(s) = \prod_{i=1}^{k} G_{X_i}(s).$$

A.2 Conditional expectation

Definition A.2 (conditional expectation)**.** *Let $X \in \mathcal{L}_1(\Omega, \mathcal{U}, \mathbb{P})$ (or non-negative) and $\mathcal{F} \subset \mathcal{U}$. A random variable $Z \in \mathcal{L}_1(\Omega, \mathcal{U}, \mathbb{P})$ is called conditional expectation of X given $\mathcal{F}$ and written $Z = \mathbb{E}[X|\mathcal{F}]$, if*

1. *Z is $\mathcal{F}$-measurable;*
2. *For all $B \in \mathcal{F}$,*
$$\mathbb{E}[Z \cdot \mathbb{1}_B] = \mathbb{E}[X \cdot \mathbb{1}_B].$$

The random variable $\mathbb{E}[X|\mathcal{F}]$ is $\mathbb{P}$-unique. For a measurable space $(S, \mathcal{S})$ and an arbitrary random variable $Y : \Omega \to S$ we define $\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$

Theorem A.3. *let X, Y and X_n be non-negative or integrable random variables on $(\Omega, \mathcal{U}, \mathbb{P})$ and let $\mathcal{F}, \mathcal{G} \subset \mathcal{U}$ be σ-algebras. The following statements hold:*

(i) Linearity: $\mathbb{E}[\lambda X + \mu Y|\mathcal{F}] = \mathbb{E}[\lambda X|\mathcal{F}] + \mathbb{E}[\mu Y|\mathcal{F}]$ *$\mathbb{P}$-a.s. for all* $\lambda, \mu \in \mathbb{R}$,

(ii) Monotonicity: $X \geq 0$ *$\mathbb{P}$-a.s. implies* $\mathbb{E}[X|\mathcal{F}] \geq 0$ *$\mathbb{P}$-a.s.,*

(iii) $X = Y$ *$\mathbb{P}$-a.s. implies that* $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[Y|\mathcal{F}]$, *$\mathbb{P}$-a.s.,*

(iv) Monotone convergence: if $\{X_n : n \in \mathbb{N}\}$ *is increasing with* $X_1 \geq 0$, *then*

$$\mathbb{E}\big[\lim X_n|\mathcal{F}\big] = \lim \mathbb{E}\big[X_n|\mathcal{F}\big],$$

(v) Tower property: If $\mathcal{G} \subset \mathcal{F}$, *then*

$$\mathbb{E}\big[\mathbb{E}\big[X|\mathcal{F}\big]|\mathcal{G}\big] = \mathbb{E}\big[X|\mathcal{G}\big],$$

(vi) Let Y be $\mathcal{F}$-measurable and $Y, Y \cdot X \in \mathcal{L}_1$. Then

$$\mathbb{E}[XY|\mathcal{F}] = Y\mathbb{E}[X|\mathcal{F}],$$

(vii) Independence: If X is independent of $\mathcal{F}$, then $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$,

(viii) Doob's Inequality: Let $\{X_t : t \geq 0\}$ be a right-continuous sub-martingale. Then,

$$\mathbb{P}[\sup_{s \leq t} X_s \geq K] \leq \frac{1}{K}\mathbb{E}[X]$$

for every $K > 0$.

A.3 Operator semigroups

Semi-group operators is an algebraic structure with an associative binary operation and are fundamental for the existence of theorems for a particular class of Markov processes.

We will say that a family of bounded linear operators $\{T(t);\ t \geq 0\}$ on a Banach space L with norm $\|\cdot\|$ is a semigroup if $T(0) = 1$ and $T(s+t) = T(s) + T(t)$ for all $s, t \geq 0$. Furthermore, if

$$\lim_{t \downarrow 0} T(t)f = f$$

for all $f \in L$ we will say that the semi-group is strongly continuous. Also it is said to be a contraction semi-group if $\|T(t)\| \leq 1$ for all $t \geq 0$.

Definition A.3 (Feller semi-group). *A strongly continuous semigroup on L is called a Feller semigroup if*

(i) $T(t)1 = 1$ *and*

(ii) $T(t)f \geq f$ *for all nonnegative* $f \in L$.

Example A.1. Consider a bounded linear operator $B \in L$. Define

$$\exp(tB) \;=\; \sum_{k=0}^{\infty} \frac{1}{k!} t^k B^k.$$

Note that $\exp((t+s)B) \;=\; \exp(tB)\exp(sB)$ for all $s, t \geq 0$. Indeed,

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{1}{k!}(t+s)^k B^k \;&=\; \sum_{k=0}^{\infty} \frac{1}{k!} B^k \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} t^j s^{k-j} \\ &=\; \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{j!} t^j B^j \frac{1}{(k-j)!} s^{k-j} B^{k-j}, \end{aligned}$$

which is enough to conclude. Hence, $\{\exp(tB)\}$ is a semigroup. Also $\{\exp(tB)\}$ is strongly continuous. Observe that, for all $f \in L$

$$\lim_{t \downarrow 0} \exp(tB) f \;=\; \lim_{t \downarrow 0} \sum_{k \geq 0} \frac{1}{k!} t^k B^k f \;=\; f \;+\; \lim_{t \downarrow 0} \sum_{k \geq 1} \frac{1}{k!} t^k B^k \;=\; f.$$

Furthermore,

$$\|\exp(tB)\| \;=\; \left\| \sum_{k=0}^{\infty} \frac{1}{k!} t^k B^k \right\| \;\leq\; \sum_{k=0}^{\infty} t^k \left\| B^k \right\| \;=\; \exp(t\,\|B\|).$$

An inequality of this type holds in general for strongly continuous semi-groups.

Proposition A.1. *Let $\{T(t)\}_{t \geq 0}$ be a strongly countinuous semigroup in L. Then there exists constants $M \geq 1$ and $w \geq 0$ such that*

$$\|T(t)\| \;\leq\; M \exp(wt)$$

Proof. There exists constants $M \geq 1$ and $t_0 \geq 0$ such that $\|T(t)\| \leq M$ for all $0 \leq t \leq t_0$. Indeed, if this constant does not exists then we could find a sequence (t_n) of positive numbers tend to zero such that $\|T(t_n)\| \to \infty$, but by the uniform boundedness principle this would imply that $\sup \|T(t_n) f\| \;=\; \infty$ for some $f \in L$, but $\{T(t)\}_{t \geq 0}$ is strongly continuous which gives us a contradiction.

Let $t_0 \;=\; w \log M$ for some $w \;\geq\; 0$. Hence, given $t \geq 0$, write $t \;=\; kt_0 \;+\; s$ for some k nonnegative integer and $0 \leq s \leq t_0$, follows

$$\begin{aligned} \|T(t)\| \;&=\; \|T(kt_0 \;+\; s)\| \;=\; \|T(kt_0)T(s)\| \\ &=\; \left\| T(t_o)^k T(s) \right\| \;\leq \|T(s)\|\,\|T(t_0)\|^k \\ &\leq M M^k \;\leq\; M M^{\frac{t}{t_0}} \;=\; M \exp(wt) \end{aligned}$$

if $0 \leq h \leq t$. □

Corollary A.1. *let $\{T(t)\}_{t \geq 0}$ be a strongly continuous semigroup on L. Then, for each $f \in L$, the map $t \mapsto T(t)f$ is a continuous function from $[0, \infty)$ into L*

Proof. Let $f \in L$. By proposition A.1, of $t \geq 0$ and $h \geq 0$

$$\begin{aligned}\|T(t+h)f - T(t)f\| &= \|T(t)T(h)f - T(t)f\| \\ &= \big\|T(t)\big[T(h)f - f\big]\big\| \\ &\leq M\exp(wt)\,\|T(h)f - f\|\,.\end{aligned}$$

If $0 \leq h \leq t$

$$\begin{aligned}\|T(t-h)f - T(f)f\| &= \|T(t-h) - T(t-h+h)\| \\ &= \|T(t-h)f - T(t-h)T(h)f\| \\ &= \big\|T(t-h)\big[T(h)f - f\big]\big\| \\ &\leq M\exp(wt)\,\|T(h)f - f\|\,.\end{aligned}$$

As $T(t)$ is strongly continuous operator, the result follows. □

Remark A.1. Let $\{T(t)\}_{t\geq 0}$ be a strongly continuous semigroup on Banach space L and define $S(t) = \exp(-wt)T(t)$ for each $t \geq 0$. Then $\{S(t)\}$ is strongly continuous semigroup in L such that

$$\|S(t)\| \leq \|\exp(-wt)T(t)\| \leq M\exp(wt)\exp(-wt) \leq M, \tag{A.1}$$

for $t \geq 0$. In particular, if $M = 1$, then $\{S(t)\}$ is strongly continuous contraction semigroup on L.

Consider $\{S(t)\}_{t\geq 0}$ to be a strongly continuous semigroup on L such that (A.1) holds. Define the norm $\|\cdot\|_1$ on L by

$$\|f\|_1 = \sup_{t\geq 0}\|S(t)f\|\,.$$

Note that $\|f\| \leq \|f\|_1 \leq M\|f\|$ for each $f \in L$, hence the norm $\|\cdot\|_1$ is equivalent to the original $\|\cdot\|$. Moreover, S with the norm $\|\cdot\|_1$ is strongly continuous contraction semigroup on L

Definition A.4. *The (infinitesimal) generator of a semigroup $\{T(t)\}$ on L is the linear operator A defined by*

$$Af = \lim_{t\downarrow 0}\frac{1}{t}\big[T(t)f - f\big].$$

The domain $\mathcal{D}(A)$ of A is the subspace of all $f \in L$ for which this limit exists.

Let $[a,b]$ a closed interval of $(-\infty,\infty)$ for some $a < b$ and denote $C([a,b],L)$ the space of continuous functions $u : [a,b] \to L$ and $C^1([a,b],L)$ to be the space of continuously differentiable functions $u : [a,b] \to L$.

The function $u : [a,b] \to L$ is said Riemann integrable over $[a,b]$, if

$$\lim_{\delta\downarrow 0}\sum_{k=1}^{n} u(s_k)(t_k - t_{k-1})$$

exists, where $\{t_k\}_{k=1}^{n}$ is a partition of $[a,b]$ and $\max\{t_k - t_{k-1}\} = \delta$.

Lemma A.1. *(i) If* $u \in C([a,b], L)$ *and* $\int_{[a,b]} \|u(t)\| \ dt < \infty$ *then* u *is Riemann integrable and*

$$\left\| \int_{[a,b]} u(t) \ dt \right\| \ \leq \ \int_{[a,b]} \|u(t)\| \ dt.$$

(ii) Let B *a closed linear operator in* L*. Suppose that* $u \in C([a,b], L))$*,* $u(t) \in \mathcal{D}(B)$ *for all* $t \in [a,b]$*,* $Bu \in C([a,b], L)$*, an both* u *and* Bu *are integrables over* $[a,b]$*. Then* $\int_{[a,b]} u(t) \ dt \in \mathcal{D}(B)$ *and*

$$B \int_{[a,b]} u(t) \ dt \ = \ \int_{[a,b]} Bu(t) \ dt.$$

(iii) if $u \in C_L[a,b]$ *then*

$$\int_a^b \frac{d}{dt} u(t) \ dt \ = \ u(b) \ - \ u(a).$$

Proof. Let $\{t_k\}_{k=1}^n$ be a partition of $[a,b]$ with $\max\{t_k - t_{k-1}\} \ = \ \delta$ then

$$\int_{[a,b]} u(t) \ dt \ = \ \lim_{\delta \downarrow 0} \sum_{k=1}^{\infty} u(s_k)(t_k - t_{k-1}),$$

where $s_k \in (t_{k-1}, t_k)$. Therefore,

$$\begin{aligned} \left\| \int_{[a,b]} u(t) \ dt \right\| \ &= \ \left\| \lim_{\delta \downarrow 0} \sum_{k=1}^{\infty} u(s_k)(t_k - t_{k-1}) \right\| \\ &\leq \ \lim_{\delta \downarrow 0} \sum_{k=1}^{\infty} \|u(s_k)(t_k - t_{k-1})\| \\ &= \ \lim_{\delta \downarrow 0} \sum_{k=1}^{\infty} \|u(s_k)\| \, (t_k - t_{k-1}) \\ &= \ \int_{[a,b]} \|u(t)\| \ dt. \end{aligned}$$

Since $\int_{[a,b]} \|u(t)\| \ dt < \infty$, we have that u is integrable which proves the first statement.

Consider $\{T(t)\}$ a semigroup on L which defines the linear operator B, i.e.,

$$Bf \ = \ \lim_{s \downarrow 0} \frac{1}{s} \big[T(s)f - f \big].$$

Thus

$$
\begin{aligned}
B \int\limits_{[a,b]} u(t)\, dt &= \lim_{s\downarrow 0} \frac{1}{s}\left[T(s) \int\limits_{\Delta} u(t)\, dt \;-\; \int\limits_{[a,b]} u(t)\, dt\right] \\
&= \lim_{s\downarrow 0} \frac{1}{s}\left[\; \int\limits_{[a,b]} T(s)u(t)\, dt \;-\; \int\limits_{[a,b]} u(t)\, dt\right] \\
&= \lim_{s\downarrow 0} \frac{1}{s} \int\limits_{[a,b]} \big[T(s)u(t) - u(t)\big]\, dt \\
&= \int\limits_{[a,b]} \lim_{s\downarrow 0} \frac{1}{s}\big[T(s)u(t) - u(t)\big]\, dt.
\end{aligned}
$$

As $u(t) \in \mathcal{D}(B)$ it follows that $\int_{[a,b]} u(t)\, dt \in \mathcal{D}(B)$ and

$$
B \int\limits_{[a,b]} u(t)\, dt \;=\; \int\limits_{[a,b]} Bu(t)\, dt.
$$

Applying the Theorem of differentiate under the signal of the integral we conclude the proof. $\square$

Proposition A.2. *Let $\{T(t)\}$ be a strongly continuous semigroup on L with generator A.*

(i) If $f \in L$ and $t \geq 0$, then $\int_0^t T(s)f\, ds \in \mathcal{D}(A)$ and

$$
T(t)f \;-\; f \;=\; A \int\limits_0^t T(s)f\, ds.
$$

(ii) If $f \in \mathcal{D}(A)$ and $t \geq 0$, then $T(t)f \in \mathcal{D}(A)$ and

$$
\frac{d}{dt} T(t)f \;=\; AT(t)f \;=\; T(t)Af.
$$

(iii) If $f \in \mathcal{D}(A)$ and $t \geq 0$, then

$$
T(t)f \;-\; f \;=\; \int\limits_0^t AT(s)f\, ds \;=\; \int\limits_0^t T(s)Af\, ds.
$$

Proof. for all $h \geq 0$, define $A_h := \frac{1}{h}\big[T(h) - I\big]$. We have that

$$\begin{aligned} A_h \int_0^t T(s)f\,ds &= \frac{1}{h}\big[T(h) - I\big] \int_0^t T(s)f\,ds \\ &= \frac{1}{h}\int_0^t [T(h)T(s)f - T(h)f]\,ds \\ &= \frac{1}{h}\int_0^t [T(h+s)f - T(h)f]\,ds \\ &= \frac{1}{h}\left[\int_h^{t+h} T(t)ff\,ds - \int_0^t T(s)f\,ds\right] \\ &= \frac{1}{h}\int_t^{t+h} T(s)f\,ds - \frac{1}{h}\int_0^h T(s)f\,ds. \end{aligned}$$

Taking $h{\downarrow}0$, the first statement follows. For the second item, for all $h \geq 0$

$$\begin{aligned} A_h T(t)f &= \frac{1}{h}\big[T(h) - I\big]T(t)f \\ &= \frac{1}{h}\big[T(h+t)f - T(t)f\big] \\ &= T(t)\frac{1}{h}\big[T(h) - I\big]f \\ &= T(t)A_h f, \end{aligned}$$

therefore $T(t)Af \in \mathcal{D}(A)$. Its remain to show that $\frac{d^+}{dt}T(t)f = \frac{d^-}{dt}T(t)f$ for all $0 \leq h \leq t$. Note that

$$\begin{aligned} \frac{1}{h}\big[T(t)f - T(t-h)f\big] - T(t)Af &= T(t-h)\big[A_h - A\big]f \\ &\quad + \big[T(t+h) - T(t)\big]Af. \end{aligned}$$

Thus

$$\begin{aligned} \left\|\frac{T(t)f - T(t-h)f}{h} - T(t)Af\right\| &\leq \left\|T(t-h)\left(\frac{T(h)f - f}{h} - AF\right)\right\| \\ &\quad + \big\|T(t-h)\big(Af - T(h)Af\big)\big\| \\ &\leq \|T(t-h)\| \left[\left\|\frac{T(h)f - f}{h} - Af\right\| + \|Af - T(h)Af\|\right]. \end{aligned}$$

Hence, since we have $\|T(t-h)\| \leq 1$ and taking $h{\downarrow}0$, the result follows. Finally, to last statement apply A.1 (3). □

Corollary A.2. *Let A be a generator of a strongly continuous semi-group $\{T(t)\}_{t\geq 0}$ on L, then $\mathcal{D}(A)$ is dense in L and A is closed.*

Proof. Let $f \in L$ arbitrary. Since $\{T(t)\}_{t\geq 0}$ is a strongly continuous semi-group, we have that $\lim_{t\downarrow 0} T(t)f = f$ and hence it follows that

$$\lim_{t\downarrow 0} \frac{1}{t}\int_o^t T(t)f\,ds = f$$

By the item *(i)* of the proposition A.2, we have the density of $\mathcal{D}(A)$ in L. To show that A is closed, let $\{f_n\}_{n\geq 1}$ satisfying $f_n \to f$ and $Af_n \to g$ to some $f, g \in L$. Note that, for each each $t \geq 0$ it follows that

$$T(t)f_n - f_n = \int_0^t T(s)Af_n\,ds,$$

and, as $n \to \infty$ we have that $T(t)f - f = \int_0^t T(s)g\,ds$. Thus

$$Af = \lim_{t\downarrow 0} \frac{1}{h}(T(t)f - f) = \lim_{t\downarrow 0} \frac{1}{t}\int_0^t T(s)g\,ds = g.$$

which proves that $g \in \mathcal{D}(A)$. $\square$

Bibliography

[1] D. Aldous et al. Stopping times and tightness. *The Annals of Probability*, 6(2):335–340, 1978.

[2] Robert G. Bartle. *The elements of integration and Lebesgue measure*, volume 27. Wiley Online Library, 1995.

[3] Julien Berestycki. Topics on branching Brownian motion. *Lecture notes available at: http://www. stats. ox. ac. uk/˜ berestyc/articles. html*, 2015.

[4] P. Billingsley. Convergence of probability measures. *New York*, 1968.

[5] Claude Dellacherie. Un survol de la théorie de l'intégrale stochastique. In *Measure Theory Oberwolfach 1979*, pages 365–395. Springer, 1980.

[6] R. Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010.

[7] Alison Etheridge. *An introduction to superprocesses*. Number 20. American Mathematical Soc., 2000.

[8] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*. Wiley interscience, 1986.

[9] L. C. Evans. *Partial Differential Equations.* American Mathematical Society, 1997.

[10] W. et al Feller. Diffusion processes in genetics. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability*. The Regents of the University of California, 1951.

[11] R. A. Fisher. The wave of advance of advantageous genes. *Annals of eugenics*, 7(4):355–369, 1937.

[12] J. L. Gondar and R. Cipolatti. *Iniciação à Física Matemática*. 2011.

[13] N. Ikeda, M. Nagasawa, S. Watanabe, et al. Branching Markov processes i. *Journal of Mathematics of Kyoto University*, 8(2):233–278, 1968.

[14] N. Ikeda, M. Nagasawa, S. Watanabe, et al. Branching Markov processes ii. *Journal of Mathematics of Kyoto University*, 8(3):365–410, 1968.

[15] V. M. Iório. *EDP. Um curso de graduação, 3ed.* IMPA, 2012.

[16] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*, volume 320. Springer Science & Business Media, 1998.

[17] A. N. Kolmogorov. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Bull. Univ. Moskow, Ser. Internat., Sec. A*, 1:1–25, 1937.

[18] Jean-François Le Gall. *Brownian motion, martingales, and stochastic calculus*, volume 274. Springer, 2016.

[19] E. L. Lima. *Espaços métricos*, volume 5. Instituto de Matemática Pura e Aplicada, CNPq Rio de Janeiro, 2015.

[20] H. P. McKean. Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Communications on pure and applied mathematics*, 28(3):323–331, 1975.

[21] R. Rebolledo et al. Sur l'existence de solutions à certains problèmes de semi-martingales. 1980.

[22] Sylvie Roelly-Coppoletta. A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics: An International Journal of Probability and Stochastic Processes*, 17(1-2):43–65, 1986.

[23] Anatolii Volodimirovich Skorokhod. Branching diffusion processes. *Theory of Probability & Its Applications*, 9(3):445–449, 1964.

[24] Jan H Van Schuppen, Eugene Wong, et al. Transformation of local martingales under a change of law. *The Annals of Probability*, 2(5):879–888, 1974.