UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE MATEMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Thermodynamic formalism for jump processes and diffusions

Tese de Doutorado

Gustavo Henrique Müller

Porto Alegre, 3 de julho de 2023

Tese submetida por Gustavo Henrique Müller[1] como requisito parcial para a obtenção do grau de Doutor em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professora Orientadora:

Dra. Adriana Neumann de Oliveira

Banca Examinadora:

Dr. Artur Oscar Lopes (PPGMAT-UFRGS)
Dra. Susana Frómeta Fernández (PPGMAT-UFRGS)
Dr. Rodrigo Marinho de Souza (UFSM)
Dra. Josiane Stein (IFSul)

Data da apresentação: 12 de Junho de 2023

[1]Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

A todos aqueles que torcem pelo meu sucesso

Agradecimentos

Gostaria inicialmente de agradecer à todas as pessoas incríveis que tenho a sorte de ter na minha vida, especialmente aos meus pais, que sempre me apoiaram e incentivaram. Todos vocês sempre acreditaram em mim e, com certeza, fazem parte dessa conquista.

Agradeço aos meus professores por todos os conhecimentos adquiridos ao longo dos anos. Se hoje eu cheguei até aqui, devo isso a vocês. Agradeço, em especial, a Adriana Neumann de Oliveira, minha orientadora, por toda a paciência e atenção ao longo de todo esse processo.

Agradeço aos colaboradores Artur Oscar Lopes e Josué Knorst, com quem tive a honra de trabalhar em parte do trabalho desenvolvido. Nosso esforço, sempre em conjunto, foi o grande responsável por termos chegado a muitos dos resultados aqui apresentados.

Agradeço também aos demais membros da banca examinadora por todos os comentários e sugestões para melhorar o trabalho. Qualquer incorreção que permaneça é de minha inteira responsabilidade.

Agradeço aos amigos que fiz ao longo do caminho e com quem tive o prazer de compartilhar muitos momentos de aprendizagem. Todas as nossas discussões em grupo, geralmente na sala de estudos da pós, foram fundamentais para um melhor aprendizado teórico.

Agradeço aos professores e funcionários do Instituto de Matemática da UFGRS o suporte. Agradeço a CAPES pelo apoio financeiro.

Resumo:

Neste trabalho, generalizamos alguns conceitos de formalismo termodinâmico já conhecidos em casos mais simples, para dois tipos de processos de Markov a tempo contínuo: processos de salto e difusões, ambos com espaço de estados compacto. Para embasar esses estudos, foi necessário reorganizar e desenvolver alguns pontos da teoria de processos de Markov, o que fizemos no primeiro capítulo desta tese, com foco nos processos de salto. Para estes dois tipos de processos de Markov, utilizando um potencial V fixado, definimos o operador de Ruelle e o normalizamos, de modo a obter o processo de Gibbs e a respectiva probabilidade de Gibbs associada. Finalmente, fomos capazes de mostrar que o processo de Gibbs é o estado de equilíbrio que maximiza um problema variacional para a pressão.

Abstract:

In this work, we generalize some concepts of thermodynamic formalism already known for simpler cases, for two types of continuous-time Markov processes: jump processes and diffusions, both with compact state space. To support these studies, it was necessary to reorganize and develop some points of the Markov process theory, which we made in the first chapter of this thesis, focusing on jump processes. For this two types of Markov processes, using a fixed potential V, we define the Ruelle operator and normalize it, getting the Gibbs process and its respective Gibbs probability associated. Finally, we were able to show that the Gibbs process is the equilibrium state that maximizes a variational problem for the pressure.

Contents

Introduction

In this introduction, we start by providing a part[2] of the historical background on the thermodynamic formalism in order to present the main names in the area and see how this Ph.D. Thesis fits into this mathematical field. It also contains an overview of each of the three chapters of this work showing all the choices made during their constructions and how they are connected. In the end, some possible directions for future studies are also suggested, following paths that have already been taken in other contexts.

The origin of thermodynamic formalism goes back to statistical physics, where Josiah Willard Gibbs (1839 - 1903) may be considered the first one to include probability theory in his analysis and mint the term "statistical mechanics". In his book [19] of 1902, Gibbs treats his results with atypical mathematical rigor for the time, praised even by Albert Einstein, but which was criticized for not addressing the physical issues involved. Gibbs' contributions continue to echo to this day with many things named after him, as you can see throughout this work.

In the 1970s, thermodynamic formalism was introduced in the mathematical field of dynamical systems. David Ruelle (1935 -) is one of the first ones to do that, and his work [43] may be considered the field's first important book. Among all his contributions to the field, Ruelle has an operator named after him that you will see here.

From that time to the present day, many things have been developed in this field: as usual, it started with simpler discrete-time processes (see, for instance, [5, 29]), moving on to continuous-time processes with countable state space (see [6, 32]) and arriving at the present work, where the time and the state space are continuous. More specifically, the main goal of this work is to describe versions of thermodynamic formalism for semi-flows of two types of continuous-time Markov processes with continuous state space: jump processes and diffusions. We consider the semi-flow given by the continuous-time shift $\Theta_t : S \to S, t \geqslant 0$, acting on the trajectories space S. This continuous-time shift Θ_t is defined in such way that $(\Theta_t w)_s = w_{s+t}$.

Using a continuous-time Markov process $\{X_t, t \geqslant 0\}$ taking values on a state space E, we introduce a homogeneous Markov semigroup $P_t = e^{tL}$, $t \geqslant 0$, where L (the infinitesimal generator) acts on some type of functions $f : E \to \mathbb{R}$. This semigroup plays the role of a transition function for the continuous-time Markov process. The exact domain $\mathcal{D}(L)$ of the infinitesimal generator depends on the

[2]For a more complete historical context, we refer to [11]

characteristics of the process: in the case of jump processes we will consider this domain as the set $C_b(E)$ of all bounded functions, while for diffusions we will have $\mathcal{D}(L) = C^2(E)$, the set of all functions of class C^2 (see [15] for examples of generators of other types of Markov process and their respective domains).

Taking a measure ν on E as the initial measure of a continuous-time Markov process, one can induce a probability $\mathbb{P}$ on S. To say that the process is stationary (the distribution at any time t is equal to ν) is equivalent to saying that the associated probability $\mathbb{P}$ is invariant for the action of the shift Θ_t, $t \geqslant 0$. We say that a probability $\mathbb{P}$ on S is invariant for Θ_t, $t \geqslant 0$ if, for all measurable set $A \subset S$ and any $t \geqslant 0$, we have that $\mathbb{P}(A) = \mathbb{P}((\Theta_t)^{-1}(A))$.

Chapter 1 addresses the results of [36] and fits into this work as the fundamental theory on which the other chapters will be based. Its results have been specifically designed to be applicable to our settings, but we have tried to leave these results in the general form as possible including, for example, the possibility of time dependence. There exists a huge difference in bibliography between diffusion and jump processes since the Brownian Motion is a well-known process and so the diffusion theory is a lot more developed. Taking this into consideration, we opted to set the results of Chapter 1 in the form of jump processes even though they are valid on a more general case. General references for basic results on diffusions that we use here appear, for example, in [4, 8, 21, 22, 47]. In this chapter, we define multiple martingales from a continuous-time Markov process. Among them, two classic martingales stand out: the Dynkin martingale and the exponential martingale. We used stochastic calculus to get the quadratic variation of the Dynkin martingale, but it is important to notice that this is the only part of this work we needed such advanced technique.

Furthermore, from a bounded function $V : [0, \infty) \times E \to \mathbb{R}$, we disturb the homogeneous semigroup P_t to get a nonhomogeneous semigroup $P^V_{s,t}$ from which we can prove an important result, the Feynman-Kac formula, which gives us a solution of the partial differential equation $\frac{\partial u}{\partial t} - Lu - Vu = 0$.

Another important contribution of this chapter involves a Radon-Nikodym derivative: using L_s and $\overline{L}_s$, any two infinitesimal generators of jump processes that depend on time s, we were able to define their respective nonhomogeneous semigroups and get a formula for the Radon-Nikodym between them. This is an alternative approach to the one presented in Appendix A of [3].

About Chapters 2 and 3, both are strongly related and deal with a standard procedure in the field of thermodynamic formalism. In fact, we follow the exact same procedure used in the first part of [32] for continuous-time Markov chains with values on the Bernoulli space. Similar results to the ones presented in these chapters are also given by [5, 28, 29] on discrete cases, by [6] for continuos-time Markov chains with finite state space, and by [10] for quantum semigroups.

Unlike Chapter 1, where we consider a general state space E, in Chapters 2 and 3 we consider a compact state space. It is not very relevant whether we take this state space as the unitary circle S^1 or the interval $[0,1]$ since we can refer to S^1 as $[0,1]$ with the periodicity boundary condition $0 \equiv 1$.

The biggest difference between these chapters is the characteristic of the paths generated by the Markov process we are considering in each case. While we have continuous paths for the diffusions case and consider the trajectories space to be $C([0,T],E)$ with the usual supremum norm, in the case of jump process we need to consider the Skorohod space $D = D([0,+\infty),E)$ of *càdlàg* paths (right continuous with left limits) $\omega : [0,+\infty) \to E$ and equip it with a Skorohod metric (see [15] for more details on this metric). This Skorohod space is a noncompact Polish space. In both cases, one can take a shift-invariant probability $\mathbb{P}$ on the trajectories space induced by a Markov process with stationary probability π (see [26]), to play the role of an *a priori* probability (a continuous time version of the point of view of [5, 28]). Considering this probability and a continuous potential $V : E \to \mathbb{R}$, we define, for $t \geqslant 0$, the Ruelle operator $\mathbb{L}_V^t$ in such way that, for $\varphi : E \to \mathbb{R}$, we get

$$(\mathbb{L}_V^t \varphi)(x) = \mathbb{E}_x \left[e^{\int_0^t V(X_r)dr} \varphi(X_t) \right].$$

Notice that, by this expression, the Ruelle operator depends on L (because $\mathbb{P}_x$ is induced by L and the initial measure δ_x).

Under the right assumptions, we can normalize the non-Markovian semigroup (associated with the infinitesimal generator $L+V$) defined by this Ruelle operator in order to get a new Markovian semigroup. This can be done using the main eigenvalue of $L+V$, called λ_V, for which the respective eigenfunctions are positive. In the discrete-time analogous procedure, we get these via Perron-Frobenius Theorem (see [37]), but we still can not get a generalization of this theorem to our setting. The new associated stationary Markov process we get in this way will be called the Gibbs process associated with the perturbation V and the shift-invariant probability on the trajectories space obtained from this process will be called the Gibbs probability associated with the potential V (see also [6, 25, 32]).

From two different homogeneous Markov processes and a Hölder continuous potential V, we consider a variational problem in the continuous-time setting which is analogous to the pressure problem in the discrete-time setting. This was done via relative entropy, a negative value that represents the relation of two processes and depends on the Radon-Nikodym derivative between them. In both our settings, we were able to prove that the pressure is equal to λ_V as the supremum happens on the associated Gibbs process.

In Chapter 2, we also analyze the properties of the time-reversal process intending to generalize the concept of entropy production, a physical concept that can be used to quantify the amount of work dissipated by an irreversible system. The entropy rate is a positive value defined as the additive inverse of the relative entropy between a process and its time reversal. This entropy production rate is equal to zero if the process is reversible, which is the case for the Brownian motion, making this analysis unnecessary in Chapter 3.

1. General results for continuous-time Markov processes

In the course of this chapter, we consider a Polish space E as general as possible, for example, E can be uncountable. By doing this, the results are presented here in its integral form, but it is important to notice that the same results are valid for countable space E if we replace the integrals by summations. Although the space E is very general, the continuous-time Markov process we will consider are jump process and diffusions, both with bounded infinitesimal generator.

The results and definitions presented in the first three sections of this chapter are very similar to the ones presented in Appendix 1 of [24] for a countable E, but we prefer to restate them here in order to be clear that they can be extended to more general spaces even though some of the proofs will be exactly the same. However, on Section 1.4 we are considering a process that is not time homogeneous and therefore we have even more general results.

Before starting this chapter, we need to say that the main goal of the present chapter is to be a tool box for Chapters 2 and 3. Fortunatelly, it does not make this a boring part of the text, since it is a beautiful theory we present here.

1.1. Markov processes

In this section, we will introduce the Markov processes and a set of premilinary results that will be useful during this text. We start by providing the general definition of a Markov process and its relation with the semigroup who acts as transition probability. Then, we state some theory about the Markov jump process, a specific type of process with discontinuous (*càdlàg*) trajectories that are constant by intervals. For the classical construction of this type of process we refer to Section 2 on Appendix 1 of [24]. In the last subsection, we introduce diffusions, another classical type of Markov process, with continuous trajectories, because in Chapter 3 of this thesis we will handle with this process.

1.1.1. Basic definitions

We begin this subsection by introducing the concept of a Markov process in a general way.

Definition 1.1. *A collection of variables $\{X_t, t \geqslant 0\}$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and taking values in a state space E is a continuous time Markov* *process if, for every $s, t \geqslant 0$ and $y \in E$,*

$$\mathbb{P}\left[X_{s+t} = y \middle| \mathcal{F}_t\right] = \mathbb{P}\left[X_{s+t} = y \middle| X_t\right] ,$$

where $\mathcal{F}_t = \sigma\{X_r, r \leqslant t\}$, which is called the natural filtration. Furthermore, this Markov process is called homogeneous if

$$\mathbb{P}\left[X_{s+t} = y \middle| X_t\right] = \mathbb{P}_{X_t}\left[X_s = y\right],$$

where $\mathbb{P}_x$ denotes the probability on Ω defined by

$$\mathbb{P}_x[\,\cdot\,] := \mathbb{P}[\,\cdot\,|X_0 = x].$$

Let $(E, \mathcal{E})$ be a measurable space. We denote $C_b(E)$ by the space of all bounded measurable functions on $(E, \mathcal{E})$. In this space, we will consider the supremum norm, denoted by $||\cdot||_\infty$.

Definition 1.2. *A family of operators $P_{s,t}, 0 \leqslant s \leqslant t$, defined on $C_b(E)$ is called a Markov semigroup of operators if satisfies*

(i) (linearity) Each $P_{s,t} : C_b(E) \to C_b(E)$ is a linear operator.

(ii) (initial condition) For all $s \geqslant 0$, $P_{s,s} = I$, the identity operator.

(iii) (semigroup property). For every $0 \leqslant s \leqslant t \leqslant u$, we have

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

(iv) (right continuity property). For every $f \in C_b(E)$ and $s \geqslant 0$, the map $t \mapsto P_{s,t}f$ is right continuous.

(v) (positivity preserving). If $f \geqslant 0$, then $P_{s,t}f \geqslant 0$, $0 \leqslant s \leqslant t$.

(vi) (mass conservation). For all $0 \leqslant s \leqslant t$, $P_{s,t}(1) = 1$, where 1 is the constant function equal to 1.

If the family $P_{s,t}$ satisfies only the items $(i) - (iv)$, it is called just semigroup, see [16]. Furthermore, if the operators $P_{s,t}$ depends only on the difference $t - s$, we call it a homogeneous semigroup and write $P_{s,t}$ simply as P_{t-s}.

Now we are ready to set a relation between the two definitions above. This same relation is introduced by [39, Chapter III] as the most basic definition on Markov process theory.

Definition 1.3 (Relation of Markov process and semigroup)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)$ a filtration in this space. A process X with state space*

$(E, \mathcal{E})$ and adapted to $(\mathcal{F}_t)$ is called a Markov process with transition function given by the semigroup $P_{s,t}$ if, for all $f \in C_b(E)$ and all $s < t$, we have

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

The probability $X_0(\mathbb{P}) := \mathbb{P}X_0^{-1}$ is called starting distribution of X.

In the homogeneous case we have that

$$\mathbb{E}_x[f(X_t)|\mathcal{F}_s] = P_{t-s}f(X_s)\,. \tag{1.1}$$

1.1.2. Infinitesimal Generators

In this subsection, we will define the infinitesimal generators of two types of Markov process and use that to define their respective semigroups. We start by making a close analysis about the jump process and later we take a look around the diffusion process, this time without going into details. We choose this two processes to present here because they will be used in the Chapters 2 and 3. Although, as we said before, there is a huge bibliography about diffusion process, then we decided to put more effort in the jump process.

1.1.2.1 Markov jump process

In the space $C_b(E)$, we denote by L the operator

$$(Lf)(x) = \lambda(x) \int_E [f(y) - f(x)]\, P(x, dy)\,, \tag{1.2}$$

where λ is a nonnegative bounded function on $C_b(E)$ and P is a transition probability (see Section 1.C for the definition). The following lemma give us a important property of L in this context.

Lemma 1.4. *The operator L, defined on equation* (1.2)*, is a linear and bounded operator acting on $C_b(E)$.*

Proof. By definition, the function λ is bounded. Denote by λ^* the upper bound of λ on E, that is, $0 \leqslant \lambda(x) \leqslant \lambda^*$, for any $x \in E$. Then, we have

$$||Lf||_\infty = \sup_{x\in E} |Lf(x)| \leqslant \lambda^* \sup_{x\in E} \int_E 2||f||_\infty P(x, dy) = 2\lambda^* ||f||_\infty,$$

where the last equality is a consequence of P be a transition probability on E. We conclude that $||L|| \leqslant 2\lambda^*$. $\square$

Clearly, for this operator L we have that $Lf \in C_b(E)$, for all $f \in C_b(E)$.

Then, by induction, for every $n \in \mathbb{N}$, we have that $L^n f \in C_b(E)$. By this,

$$||L^j f||_\infty = ||L(L^{j-1}f)||_\infty \leqslant ||L||\,||L^{j-1}f||_\infty \leqslant \cdots \leqslant ||L||^j ||f||_\infty. \tag{1.3}$$

As consequence of this, we can set, for all $f \in C_b(E)$ and $x \in E$,

$$P_t f(x) := e^{tL} f(x) = \sum_{k=0}^{\infty} \frac{t^k}{k!}(L^k f)(x) = f(x) + tLf(x) + \frac{t^2}{2}L^2 f(x) + \cdots \tag{1.4}$$

where this operator is well defined because the bound over L^k given by equation (1.3) implies the convergence of this series.

Lemma 1.5. *This family of operators $\{P_t\}_{t\geqslant 0}$ is a semigroup of operators.*

Proof. In this proof, we need to show that this family satisfies the first four conditions of Definition 1.2 in the homogeneous case (we can think of P_t as $P_{0,t}$). For *(i)*, we just need to notice that, for every $t \geqslant 0$, $P_t f$ is a sum of linear operators. The initial condition is also valid: $P_0 f(x) = e^0 f(x) = f(x)$. The condition *(iii)* is an immediate consequence of the properties of exponentials. Finally, notice that, under the semigroup property, the fourth condition is equivalent to prove that $P_t f$ converges to f, when t goes to zero, and this is true because making t goes to zero on equation (1.4), every term of the sum vanishes except the first term $f(x)$. ☐

Lemma 1.6. *With L and P_t as defined above, as t goes to zero, we get*

$$\left|\left|\frac{P_t f - f}{t} - Lf\right|\right|_\infty \longrightarrow 0.$$

Proof. Notice that, for every $x \in E$,

$$\left(\frac{P_t f - f}{t} - Lf\right)(x) = \frac{1}{t}\sum_{k=1}^{\infty} \frac{t^k}{k!}(L^k f)(x) - Lf(x) = \sum_{k=2}^{\infty} \frac{t^{k-1}}{k!}(L^k f)(x),$$

whose the supremum norm vanishes when t goes to zero because of equation (1.3). ☐

The Kolmogorov equations, including Kolmogorov forward equation and Kolmogorov backward equation, characterize stochastic processes. In particular, they describe how the probability that a stochastic process is in a certain state changes over time.

Proposition 1.7 (Kolmogorov equations)**.** *The operators P_t, defined in equation* (1.4)*, are continuously differentiable in time and satisfy, for a function $f \in C_b(E)$, the following:*

(i) Kolmogorov forward equation[1]*:*

$$\partial_t(P_t f)(x) = P_t(Lf)(x)\,;$$

(ii) Kolmogorov backward equation[2]*:*

$$\partial_t(P_t f)(x) = L(P_t f)(x)\,.$$

Proof. For the forward equation, we compute

$$\frac{P_{t+h}f(x) - P_t f(x)}{h} = \frac{P_t(P_h f)(x) - P_t(P_0 f)(x)}{h} = P_t\left[\frac{P_h f - P_0 f}{h}\right](x)$$

because of the linearity of P_t. Making h goes to zero, the Lemma 1.6 and the continuity of P_t allow us to conclude that

$$\partial_t(P_t f)(x) = P_t(Lf)(x).$$

The backward equation is a direct consequence of the application of Lemma 1.6 for the function $P_t f$ because we can write

$$L(P_t f)(x) = \lim_{h\to 0}\frac{P_h(P_t f)(x) - P_0(P_t f)(x)}{h} = \partial_t(P_t f)(x).$$

$\square$

1.1.2.2 One-dimensional Markov diffusion

A time-homogeneous Markov process is a one-dimensional diffusion if its infinitesimal generator is the operator L who acts on functions $f \in D(L)$ as

$$Lf(x) = \frac{1}{2}a(x)\frac{d^2 f}{dx^2}(x) + b(x)\frac{df}{dx}(x), \tag{1.5}$$

where the functions $a \neq 0$ and b are measurable, non-negative and bounded. We call a the diffusion coeficient and b the drift of the process. A classical example of diffusion is the Brownian Motion, which is the case where $a \equiv 1$ and $b \equiv 0$. For more information, see [40]. A general study on d-dimensional diffusions can be found on [39, Chapter VII], where the infinitesimal generator is generalized using partial derivatives.

In the general case of Markov processes, given a homogeneous semigroup P_t, the domain of L, denoted by $D(L)$, is the set of all functions to whom exists the

[1]or Fokker-Planck equation, see [9].
[2]or parabolic equation, see [16], in the context of diffusions.

limit $\frac{P_t f - f}{t}$ when t decreses to zero, see [16, Chapter 7]. In this text we will consider that this infinitesimal generator acts on $C_b^2(E)$, the subset of $C_b(E)$ that contains all functions of class C^2.

Diffusion processes have a lot of nice properties that are well known. One of this properties is the fact it has continuous paths. We will not show the proprieties about this type of process here, because this work will be focused on Markov jump process, but the reader can find a deeper analysis on diffusions in [39, 40].

1.2. Martingales and Markov processes

In this section we show how to obtain martingales from Markov processes. The goal here is to use the most general Markov process as possible. The first result is the well know Dynkin Martingale, presented in the Theorem 1.8. The second construction of a martingale from a Markov process, presented in the Theorem 1.10, is also a powerful tool. For example, it is very important in the proof of Kolmogorov equations for the perturbed process on Section 1.3.

To be more general in our analysis, when necessary, we may consider functions $F : [0, \infty) \times E \to \mathbb{R}$ that depends on time (to simplify the notation, we write $F_s(x)$ for $F(s, x)$). In this case, we need to suppose that these functions satisfies the following assumption:

Assumption 1.1. *We assume that a bounded function $F : [0, \infty) \times E \to \mathbb{R}$ is smooth in the first coordinate uniformly over the second, i.e., for each $x \in E$, the function $F(\cdot, x)$ is twice continuously differentiable and there is a finite constant C_F such that, for $j = 0, 1, 2$,*

$$\sup_{(s,x)} \left|(\partial_s^j F_s)(x)\right| \leqslant C_F \,,$$

where $\partial_s^0 F_s(x)$ stands for $F_s(x)$.

Observe that, to apply it in the diffusion case, due to the nature of its infinitesimal generator (see equation (1.5)), we also need to suppose that the functions are of class C^2 on the spatial variable.

Theorem 1.8 (Dynkin martingale)**.** *Let $\{X_t,\, t \geqslant 0\}$ be a Markov process adapted to the filtration $\{\mathcal{F}_t; t \geqslant 0\}$. For each function F satisfying the Assumption 1.1, we define*

$$M^F(t) = F_t(X_t) - F_0(X_0) - \int_0^t (\partial_s + L) F_s(X_s)\, ds \,. \tag{1.6}$$

The process $M^F(t)$ is $\mathcal{F}_t$-martingale.

Proof. First, we will prove that $M^F(t)$ is a martingale. Fix $s \in [0,t)$. We need to show that

$$M^F(s) = \mathbb{E}_x \left[F_t(X_t) - F_0(X_0) - \int_0^t (\partial_r + L)F_r(X_r)\, dr \middle| \mathcal{F}_s \right] .$$

Rewriting the equality above, we have

$$\begin{aligned} \mathbb{E}_x[F_t(X_t)|\mathcal{F}_s] &= M^F(s) + F_0(X_0) + \int_0^t \mathbb{E}_x[(\partial_r + L)F_r(X_r)|\mathcal{F}_s]\, dr \\ &= F_s(X_s) + \int_s^t \mathbb{E}_x[(\partial_r + L)F_r(X_r)|\mathcal{F}_s]\, dr\, . \end{aligned} \tag{1.7}$$

Making a change of variables $r \mapsto r + s$ on the last integral above, we get

$$\int_0^{t-s} \mathbb{E}_x[(\partial_{r+s} + L)F_{r+s}(X_{r+s})|\mathcal{F}_s]\, dr\, .$$

By equation (1.1), the last integral becomes

$$\int_0^{t-s} \{P_r(\partial_{r+s}F_{r+s})(X_s) + P_r(LF_{r+s})(X_s)\}\ dr\, .$$

Thus, we conclude that the expression we need to prove is

$$(P_{t-s}F_t)(X_s) - F_s(X_s) = \int_0^{t-s} \{(P_r\partial_{r+s}F_{r+s})(X_s) + (P_rLF_{r+s})(X_s)\}\ dr\, .$$

Observe that, if $t = s$, this equality is trivial. Then, we just need to check that the time derivatives are equal on both sides:

$$\partial_t(P_{t-s}F_t)(x) = (P_{t-s}\partial_t F_t)(x) + (P_{t-s}LF_t)(x),$$

for any $x \in E$ and $s \in [0,t)$. To prove this, fix $h > 0$, add and subtract $\mathbb{E}_x[F_t(X_{t-s+h})]$ to rewrite $\frac{1}{h}\{(P_{t-s+h}F_{t+h})(x) - (P_{t-s}F_t)(x)\}$ as

$$\frac{1}{h}\mathbb{E}_x[F_{t+h}(X_{t-s+h}) - F_t(X_{t-s+h})] + \frac{1}{h}\mathbb{E}_x[F_t(X_{t-s+h}) - F_t(X_{t-s})].$$

Observe that, for any $u > 0$, we get $\partial_r\mathbb{E}_x[F_r(X_u)] = \mathbb{E}_x[\partial_r F_r(X_u)]$ and $\partial_r\mathbb{E}_x[F_u(X_r)] = \partial_r(P_rF_u)(x) = (P_rLF_u)(x)$, the last equality is a consequence of Kolmogorov forward equations. Then, we can rewrite the previous expression as

$$\frac{1}{h}\int_t^{t+h} \mathbb{E}_x[\partial_r F_r(X_{t-s+h})]\, dr + \frac{1}{h}\int_{t-s}^{t-s+h} (P_rLF_t)(x), dr\, .$$

Adding and subtracting appropriate terms, the expression above is equal to

$$\frac{1}{h}\int_t^{t+h} \mathbb{E}_x[\partial_r F_r(X_{t-s+h}) - \partial_t F_t(X_{t-s+h})]\,dr + \mathbb{E}_x[\partial_t F_t(X_{t-s+h}) - \partial_t F_t(X_{t-s})] + \mathbb{E}_x[\partial_t F_t(X_{t-s})] + \frac{1}{h}\int_{t-s}^{t-s+h} (P_r L F_t)(x)\,dr\,. \tag{1.8}$$

Now, let us analyse what happens with each part of the expression above when h goes to zero. We have that $(\partial F)(\cdot, x)$ is a Lipschitz function uniformly on x, then

$$\frac{1}{h}\int_t^{t+h} \mathbb{E}_x[|\partial_r F_r(X_{t-s+h}) - \partial_t F_t(X_{t-s+h})|]\,dr \leqslant C\mathbb{E}_x\left[\frac{1}{h}\int_0^h |r|\,dr\right].$$

By the Lebesgue Differentiation Theorem, this upper bound vanishes when h goes to zero. Thus the first term in equation (1.8) vanishes. The second term of equation (1.8) also vanishes, as $h \to 0$, because it can be rewritten as $P_{t-s+h}(\partial_t F_t)(x) - P_{t-s}(\partial_t F_t)(x)$ and the semigroup P_t is continuous. Using again the Lebesgue Differentiation Theorem, the last term in equation (1.8) converges to $(P_{t-s}LF_t)(x)$. With all that, we get

$$\partial_t(P_{t-s}F_t)(x) = (P_{t-s}\partial_t F_t)(x) + (P_{t-s}LF_t)(x),$$

and it shows that $M^F(t)$ is a martingale. □

Now, for each function F satisfying the Assumption 1.1, we define

$$N^F(t) = (M^F(t))^2 - \int_0^t \left[(\partial_s + L)F_s^2(X_s) - 2F_s(X_s)(\partial_s + L)F_s(X_s)\right] ds\,.$$

In the next result, we prove this is a martingale. Thus, we conclude that the integral part of $N^F(t)$ is the quadratic variation of $M^F(t)$.

Proposition 1.9. *The process $N^F(t)$ is $\mathcal{F}_t$-martingale.*

Proof. We start by analyse $(M^F(t))^2$. For simplicity, we denote $I_t = \int_0^t (\partial_s + L)F_s(X_s)\,ds$, then we can write

$$(M^F(t))^2 \quad = \quad (F_t(X_t))^2 + I_t^2 - 2F_t(X_t)I_t + M_1(t)\,,$$

where

$$M_1(t) := F_0(X_0)(-2F_t(X_t) + 2I_t + F_0(X_0)) = F_0(X_0)(-2M^F(t) - F_0(X_0))\,.$$

Since $M^F(t)$ is a martingale, $M_1(t)$ is a martingale. Now, we consider the Dynkin martingale for F^2,

$$M^{F^2}(t) = F_t^2(X_t) - F_0^2(X_0) - \int_0^t (\partial_s + L)F_s^2(X_s)\, ds,$$

and note that $M_0^{F^2}(t) := M^{F^2}(t) + F_0^2(X_0)$ is also a martingale. Using it, we get

$$(M^F(t))^2 = I_t^2 - 2F_t(X_t)I_t + M_1(t) + M_0^{F^2}(t) + \int_0^t (\partial_s + L)F_s^2(X_s)\, ds.$$

Denote the martingale $M^F(t) + F_0(X_0)$ by $M_0^F(t)$. And use $F_t(X_t) = M_0^F(t) + I_t$ to rewrite $(M^F(t))^2$ as

$$-2M_0^F(t)\, I_t - I_t^2 + M_1(t) + M_0^{F^2}(t) + \int_0^t (\partial_s + L)F_s^2(X_s)\, ds\,.$$

To handle with the multiplicative term $-2M_0^F(t)\, I_t$, we note that I is a predictable process (I is adapted and continuous) and F satisfies the Assumption 1.1 that implies $M_0^F(t)$ is a martingale, then we evoke Proposition 1.29 from Appendix 1.C , to say that

$$M_0^F(t)\, I_t = \int_0^t I_s\, dM_0^F(s) + \int_0^t M_0^F(s)\, dI_s,$$

where the first integral in the right hand-side above is a martingale. Besides that, the second term of the sum above can be rewritten as

$$\int_0^t F_s(X_s)(\partial_s + L)F_s(X_s)\, ds - \int_0^t I_s\, I_s'\; ds\,.$$

Using the classical integration by parts formula, the second term in the previous expression is equal to $\frac{1}{2}I_t^2$. Then,

$$\begin{aligned} (M^F(t))^2 \quad &= \quad M_1(t) + M_0^{F^2}(t) - 2\int_0^t I_s\, dM_0^F(s) \\ &\quad -2\int_0^t F_s(X_s)(\partial_s + L)F_s(X_s) ds + \int_0^t (\partial_s + L)F_s^2(X_s)\, ds\,. \end{aligned}$$

Note that $M_1(t) + M_0^{F^2}(t) - 2\int_0^t I_s\, dM_0^F(s)$ is a martingale. Then the result is proved if we denote this martingale by $N^F(t)$.

$\square$

Another important result that give us martingales from Markov process is the following:

Theorem 1.10. *Let $\{X_t,\, t \geqslant 0\}$ be a Markov process adapted to the filtration $\{\mathcal{F}_t; t \geqslant 0\}$. Let $F, V : [0,\infty) \times E \to \mathbb{R}$ two bounded functions such that F satisfies the Assumption 1.1. Then,*

$$F_t(X_t)\, e^{\int_0^t V_r(X_r)\, dr} - \int_0^t e^{\int_0^s V_r(X_r)\, dr} [F_s(X_s)V_s(X_s) + (\partial_s + L)F_s(X_s)]\, ds \tag{1.9}$$

is a martingale.

Proof. Let us introduce some notation to rewrite the equation (1.9). As usual we denote by $M_0^F(t) = F_t(X_t) - \int_0^t (\partial_s + L)F_s(X_s)\, ds$. Note that $M_0^F(t)$ is a martingale, due to the fact that $M_0^F(t)$ is the Dynkin Martingale $M^F(t)$ plus a random variable $F_0(X_0)$. Denote by $I_t = \int_0^t (\partial_s + L)F_s(X_s)\, ds$, thus $F_t(X_t) = M_0^F(t) + I_t$ and $dI_s = (\partial_s + L)F_s(X_s)\, ds$. Moreover, we write $Z_t = e^{\int_0^t V_s(X_s)\, ds}$, then $dZ_s = e^{\int_0^s V_r(X_r)\, dr} V_s(X_s)\, ds$[3]. Finally, using all the notations above we can rewrite equation (1.9) as

$$\begin{aligned} &[M_0^V(t) + I_t]\, Z_t - \int_0^t [M_0^F(s) + I_s]\, dZ_s - \int_0^t Z_s\, dI_s \\ =&\mathcal{N}_t + I_t\, Z_t - \int_0^t I_s\, dZ_s - \int_0^t Z_s\, dI_s\,, \end{aligned}$$

where $\mathcal{N}_t = M_0^F(t)\, Z_t - \int_0^t M_0^F(s)\, dZ_s$ is a martingale by the Theorem 1.31 of Appendix 1.C. Thus, to conclude the proof we need to observe that

$$I_t\, Z_t - \int_0^t I_s\, dZ_s - \int_0^t Z_s\, dI_s = I_0 Z_0 = 0\,,$$

because for a fixed trajectory the processes I and Z have bounded variation, then the result follows from Integration by parts formula, see Proposition 1.28 on Appendix 1.C.

□

As consequence of the result above, we can produce another classical martingale, called exponential martingale. This martingale has the nice property of being positive and it will be used to set a Radon-Nikodym derivative on Section 1.B.1.

Corollary 1.11 (Exponential martingale)**.** *Fix a function $F : \mathbb{R}_+ \times E \to \mathbb{R}$ satisfying Assumption 1.1. The expression*

$$\mathbb{M}^F(t) := \exp\left\{ F_t(X_t) - F_0(X_0) - \int_0^t e^{-F_s(X_s)} (\partial_s + L)\, e^{F_s(X_s)}\, ds \right\} \tag{1.10}$$

[3] Z has trajectories absolutely continuous by the hypothesis over V, see Lemma 1.32.

is a $\mathcal{F}_t-$martingale.

Proof. First of all, we rewrite the expression in the statement of this corollary as

$$e^{F(t,X_t)-F(0,X_0)}\exp\left\{-\int_0^t e^{-F(r,X_r)}(\partial_r+L)e^{F(r,X_0)}dr\right\}.$$

Define $\varphi(t,X_t)=e^{F(t,X_t)-F(0,X_0)}$ and $V(r,X_r)=-e^{-F(r,X_r)}(\partial_r+L)e^{F(r,X_0)}$. As F satisfies the Assumption 1.1, we get that φ satisfies this assumption too. By Theorem 1.10, if we consider the last expression minus

$$\int_0^t e^{\int_0^s V(r,X_r)dr}[\varphi(s,X_s)V(s,X_s)+(\partial_s+L)\varphi(s,X_s)]\,ds$$

we obtain a martingale. Then, we just need to prove that the last integral vanishes. For this, notice that, for any x and s,

$$\begin{aligned}\varphi(s,x)V(s,x) &= e^{F(s,x)}e^{-F(0,x_0)}[-e^{-F(s,x)}(\partial_s+L)e^{F(s,x)}]\\ &= -e^{-F(0,x_0)}(\partial_s+L)e^{F(s,x)}=-(\partial_s+L)\varphi(s,x),\end{aligned}$$

what proves that $\varphi(s,X_s)V(s,X_s)+(\partial_s+L)\varphi(s,X_s)=0$ and concludes the proof. $\square$

1.3. A perturbed process

In this section we will study a perturbation of the process X. It is important to analyse the consequences of this kind of perturbation as this will be used later to introduce the Ruelle Operator on Chapters 2 and 3. Fix a bounded function $V:\mathbb{R}_+\times E\to\mathbb{R}$, then, disturb the homogeneous semigroup P_t of the Markov process X using the function V in the following way:

$$P^V_{s,t}f(x)=\mathbb{E}_x\left[e^{\int_s^t V_r(X_r)\,dr}f(X_t)\right],$$

for $x\in E$, $s<t$ and $f\in C_b(E)$. Note that, in general, $P^V_{s,t}$ is a nonhomogeneous semigroup, but it becomes homogeneous if V is constant in time, i.e.,

$$P^V_{s,t}f(x)=\mathbb{E}_{X_s}\left[e^{\int_0^{t-s} V(X_u)\,du}f(X_{t-s})\right]=:P^V_{t-s}f(X_s)\,, \tag{1.11}$$

for $x\in E$, $s<t$ and $f\in C_b(E)$.

To be more general in our analysis, we consider the operator $P^V_{s,t}$ acting on functions $F:\mathbb{R}_+\times E\to\mathbb{R}$ satisfying the Assumption 1.1 instead of $f\in C_b(E)$, i.e.,

$$P^V_{s,t}F_t(x)=\mathbb{E}_x\left[e^{\int_s^t V_r(X_r)\,dr}F_t(X_t)\right].$$

Taking conditional expectation concerning $\mathcal{F}_s$ in the expression above and using the Markov Propriety, homogeneity and change of variables, we get that $P^V_{s,t}$ acts on functions $F : \mathbb{R}_+ \times E \to \mathbb{R}$ satisfying the Assumption 1.1, in this way:

$$P^V_{s,t}F_t(x) = \mathbb{E}_x\left[e^{\int_0^{t-s} V^s_u(X_u)\,du}F_t(X_{t-s})\right],$$

where $V^s_u(x) := V_{u+s}(x)$. From the definition of this family of operators, using the Markov property on the expression above, we get a semigroup property (or Chapman-Kolmogorov equation)

$$P^V_{s,t}(P^V_{t,u}F_u)(x) = P^V_{s,u}(F_u)(x)\,,$$

for all $s < t < u$ and $F : \mathbb{R}_+ \times E \to \mathbb{R}$ satisfying the Assumption 1.1.

1.3.1. Kolmogorov Equations for the perturbed process

To understand the complete evolution of this semigroup we present the Kolmogorov equations. In order to do it, we need to define the operator $L^V_t = L + V_t$, which acts in functions $f \in C_b(E)$ as

$$L^V_t f(x) = Lf(x) + V_t(x)f(x)\,.$$

Proposition 1.12 (Kolmogorov equations)**.** *For any $0 \leqslant s < t$, the operator $P^V_{s,t}$ are continuously differentiable in time and satisfy, for functions $F : \mathbb{R}_+ \times E \to \mathbb{R}$ satisfying the Assumption 1.1, the following:*

(i) Kolmogorov forward equation:

$$\partial_t\left(P^V_{s,t}F_t(x)\right) = P^V_{s,t}(L^V_t F_t)(x) + P^V_{s,t}(\partial_t F_t)(x)\,;$$

(ii) Kolmogorov backward equation:

$$\partial_s\left(P^V_{s,t}F_t(x)\right) = -L^V_s(P^V_{s,t}F_t)(x)\,.$$

Proof. We start by introducing the notations

$$A^s_t := e^{\int_0^t V^s_u(X_u)\,du}F_{t+s}(X_t)$$

and

$$B^s_t := \int_0^t \left\{(\partial_{r+s} + L)F_{r+s}(X_r) + V^s_r(X_r)F_{r+s}(X_r)\right\} e^{\int_0^r V^s_u(X_u)\,du}\,dr\,.$$

Recall that $V^s_u(x) = V_{u+s}(x)$. Note that $P^V_{s,t}F_t(x) = \mathbb{E}_x[A^s_{t-s}]$. Moreover, by the hypotheses over V and F, and from Theorem 1.10, we get that $\{A_t^s - B_t^s\}_{t\geqslant s}$ is a martingale, for $s \geqslant 0$ fixed. Since $\mathbb{E}_x[A_{t-s}^s - B_{t-s}^s] = \mathbb{E}_x[A_0^s - B_0^s] = F_s(x)$, for all $t \geqslant s \geqslant 0$, we can rewrite $P_{s,t}^V F_t(x) = F_s(x) + \mathbb{E}_x[B_{t-s}^s]$. Thus

$$\begin{aligned}
&\tfrac{1}{h}\left\{P_{s,t+h}^V F_{t+h}(x) - P_{s,t}^V F_t(x)\right\} = \tfrac{1}{h}\left\{\mathbb{E}_x[B_{t-s+h}^s] - \mathbb{E}_x[B_{t-s}^s]\right\} \\
&= \mathbb{E}_x\left[\frac{1}{h}\int_{t-s}^{t-s+h} \{(\partial_{r+s} + L)F_{r+s}(X_r) + V_r^s(X_r)F_{r+s}(X_r)\}\, e^{\int_0^r V_u^s(X_u)\,du}\, dr\right].
\end{aligned}$$

For a fixed s, the Lebesgue Differentiation Theorem shows that

$$\begin{aligned}
&\partial_t\left(P_{s,t}^V F_t(x)\right) = \lim_{h\to 0} \tfrac{1}{h}\left\{P_{s,t+h}^V F_{t+h}(x) - P_{s,t}^V F_t(x)\right\} \\
&= \mathbb{E}_x\left[e^{\int_0^{t-s} V_u^s(X_u)\,du}\, L_t^V F_t(X_{t-s})\right] + \mathbb{E}_x\left[e^{\int_0^{t-s} V_u^s(X_u)\,du}\, \partial_t F_t(X_{t-s})\right] \\
&= P_{s,t}^V(L_t^V F_t)(x) + P_{s,t}^V(\partial_t F_t)(x)\,.
\end{aligned}$$

For the backward Kolmogorov equation, we compute the following limit

$$\partial_s(P_{s,t}^V F_t(x)) = -\lim_{h\to 0} \frac{P_{s-h,t}^V F_t(x) - P_{s,t}^V F_t(x)}{h}\,,$$

for a fixed t. We start observing that

$$\begin{aligned}
P_{s-h,t}^V F_t(x) &= \mathbb{E}_x\left[e^{\int_{s-h}^t V_r(X_{r-(s-h)})\,dr} F_t(X_{t-(s-h)})\right] \\
&= \mathbb{E}_x\left[e^{\int_{s-h}^s V_r(X_{r-(s-h)})\,dr} e^{\int_s^t V_r(X_{r-s+h})\,dr} F_t(X_{t-s+h})\right]
\end{aligned}$$

and, using the Markov propriety and homogeneity of X, we obtain

$$\begin{aligned}
&P_h(P_{s,t}^V F_t)(x) = E_x[P_{s,t}^V F_t(X_h)] = \mathbb{E}_x\left[\mathbb{E}_{X_h}\left[e^{\int_0^{t-s} V_{u+s}(X_u)\,du} F_t(X_{t-s})\right]\right] \\
&= \mathbb{E}_x\left[e^{\int_0^{t-s} V_{u+s}(X_{u+h})\,du} F_t(X_{t-s+h})\right] = \mathbb{E}_x\left[e^{\int_s^t V_r(X_{r-s+h})\,dr} F_t(X_{t-s+h})\right].
\end{aligned}$$

Thus, $\frac{1}{h}[P_{s-h,t}^V F_t(x) - P_{s,t}^V F_t(x)]$ is equal to

$$\begin{aligned}
&\mathbb{E}_x\left[\left(\frac{e^{\int_{s-h}^s V_r(X_{r-(s-h)})\,dr} - 1}{h}\right) e^{\int_s^t V_r(X_{r-(s-h)})\,dr} F_t(X_{t-s+h})\right] \\
&\qquad + \frac{P_h(P_{s,t}^V F_t)(x) - P_{s,t}^V F_t(x)}{h}\,.
\end{aligned} \tag{1.12}$$

By the hypothesis over V and F, the first term in equation (1.12) is bounded.

Then the Dominated Convergence Theorem implies that it converges to

$$\mathbb{E}_x\left[V_s(X_0)\, e^{\int_s^t V_r(X_{r-s})\,dr} F_t(X_{t-s})\right] = V_s(x) P^V_{s,t} F_t(x)\,,$$

as $h \to 0$. The second term in equation (1.12) converges to $L(P^V_{s,t}F_t)(x)$, as $h \to 0$, because the backward Kolmogorov equation for the process X. Finally,

$$\partial_s \left(P^V_{s,t} F_t(x)\right) = -[LP^V_{s,t}F_t(x) + V_s(x)P^V_{s,t}F_t(x)] = -L^V_s\,(P^V_{s,t}F_t)(x)\,.$$

□

1.3.2. Feynman-Kac Formula

In this subsection, we present a result that links parabolic partial differential equations and stochastic processes. Another important fact about the Feynman-Kac Formula is that it proves rigorously the real case of Feynman's path integrals from Quantum Mechanics.

Proposition 1.13. *Fix $T > 0$. For all $V : \mathbb{R}_+ \times E \to \mathbb{R}$ satisfying the Assumption 1.1 and $f \in C_b(E)$. Fix $T > 0$ we define the function $u : [0,T]\times E \to \mathbb{R}$ as*

$$u_t(x) = \mathbb{E}_x\left[e^{\int_0^t V_{T-t+r}(X_r)\,dr} f(X_t)\right]\,. \tag{1.13}$$

Then u is a solution of the partial differential equation

$$\begin{cases} \partial_t u_t(x) = (Lu_t)(x) + V_{T-t}(x)u_t(x)\,, \quad x \in E\,, \quad t \in (0,T]\,, \\ u_0(x) = f(x)\,, \quad x \in E\,. \end{cases}$$

Proof. We start the proof by computing $\partial_t u_t(x) - Lu_t(x)$ using Lemma 1.6. This is equal to

$$\lim_{h\to 0} \frac{u_{t+h}(x) - u_t(x)}{h} - \lim_{h\to 0} \frac{P_h u_t(x) - u_t(x)}{h} = \lim_{h\to 0} \frac{u_{t+h}(x) - P_h u_t(x)}{h}\,. \tag{1.14}$$

In order to compute the limit above, we start by studding $u_{t+h}(x)$. By the definition of the function u, in equation (1.13), and taking the conditional expectation concerning $\mathcal{F}_h$, we get

$$\begin{aligned} u_{t+h}(x) &= \mathbb{E}_x\left[e^{\int_0^{t+h} V_{T-t-h+r}(X_r)\,dr} f(X_{t+h})\right] \\ &= \mathbb{E}_x\left[e^{\int_0^h V_{T-t-h+r}(X_r)\,dr}\, \mathbb{E}_x\left[e^{\int_h^{t+h} V_{T-t-h+r}(X_r)\,dr} f(X_{t+h})\Big|\mathcal{F}_h\right]\right]\,. \end{aligned}$$

Using the Markov property and homogeneity, the last expectation is equal to

$$\begin{aligned}
&\mathbb{E}_x\left[e^{\int_0^h V_{T-t-h+r}(X_r)\,dr}\mathbb{E}_{X_h}\left[e^{\int_0^t V_{T-t+w}(X_w)\,dw}f(X_t)\right]\right]\\
&= \mathbb{E}_x\left[e^{\int_0^h V_{T-t-h+r}(X_r)\,dr}u_t(X_h)\right].
\end{aligned}$$

Thus, the quotient $\frac{u_{t+h}(x)-P_h u_t(x)}{h}$, presented in the limit equation (1.14), can be rewrite as

$$\frac{1}{h}\left\{\mathbb{E}_x\left[e^{\int_0^h V_{T-t-h+r}(X_r)\,dr}u_t(X_h)\right]-\mathbb{E}_x[u_t(X_h)]\right\}.$$

By adding and subtracting the appropriate terms, we can rewrite the expression above as

$$\begin{aligned}
&\mathbb{E}_x\left[\left(\frac{e^{\int_0^h V_{T-t-h+r}(X_r)\,dr}-e^{\int_0^h V_{T-t+r}(X_r)\,dr}}{h}\right)u_t(X_h)\right]\\
&+\mathbb{E}_x\left[\left(\frac{e^{\int_0^h V_{T-t+r}(X_r)\,dr}-1}{h}-V_{T-t}(X_0)\right)u_t(X_h)\right]\\
&+V_{T-t}(x)\,\mathbb{E}_x[u_t(X_h)-u_t(X_0)]+V_{T-t}(x)\,u_t(x).
\end{aligned}\tag{1.15}$$

To conclude the result, it is enough proving that the three first terms in the sum above go to zero when $h\to 0$. For the first term, we write the expression inside the parenthesis as

$$e^{\int_0^h V_{T-t+r}(X_r)\,dr}\left(\frac{e^{\int_0^h [V_{T-t-h+r}(X_r)-V_{T-t+r}(X_r)]\,dr}-1}{h}\right).$$

Since the first term of this product is bounded, we just need to show that the second term goes to zero. Using that V is a Lipschitz function on time (as consequence of Assumption 1.1), the last term is bounded from above by

$$\frac{e^{\int_0^h C_V h\,dr}-1}{h}=\frac{e^{C_V h^2}-1}{h},$$

which converges to zero, as $h\to 0$. The second term of the sum equation (1.15) goes to zero, because the fraction part converges to $V_{T-t}(X_0)$. For the third term of the sum equation (1.15) we note that $\mathbb{E}_x[u_t(X_h)-u_t(X_0)]=P_h u_t(x)-u_t(x)$ and by the right continuity of the semigroup (see Definition 1.2), we also have that the third term of equation (1.15) goes to zero, as $h\to 0$. $\square$

Remark. *An important observation is that when a function V is constant in* *time, we have*

$$u_t(x) = \mathbb{E}_x\left[e^{\int_0^T V(X_r)\,dr} f(X_T)\right] = P_t^V f(x),$$

where $P_t^V f(x)$ *was defined in equation* (1.11)*. Moreover, by Proposition 1.13,* $P_t^V f(x)$ *is a solution of*

$$\begin{cases} \partial_t(P_t^V f)(x) = L^V(P_t^V f)(x)\,, & x \in E\,, \quad t > 0\,, \\ P_0^V f(x) = f(x)\,, & x \in E\,, \end{cases}$$

where $L^V = L + V$*.*

The next Lemma allow us to analyze a particular case when we start the process from an initial probability ν that may not be invariant. This result is originally stated on [2], but we prefer to repeat it here.

Lemma 1.14 (Feynman-Kac's lemma without invariant measure)**.** *Let* ν *be a probability measure in* E *and* V *be a bounded function. Define*

$$\Gamma_t = \sup_{\|f\|_2=1}\Big\{\langle V_t, f^2\rangle_\nu + \langle Lf, f\rangle_\nu\Big\} = \sup_{\|f\|_2=1}\Big\{\langle L_t^V f, f\rangle_\nu\Big\},\ \textit{for all } t \geqslant 0, \tag{1.16}$$

where $\langle\cdot,\cdot\rangle_\nu$ *denotes the inner product in* $\mathscr{L}^2(\nu)$ *and* $\|\cdot\|_2 = \langle\cdot,\cdot\rangle_\nu^{1/2}$*. Then*

$$\mathbb{E}_\nu\Big[e^{\int_0^t V_r(X_r)\,dr}\Big] \leqslant \exp\Big\{\int_0^t \Gamma_s\,ds\Big\}.$$

Proof. For a function V, define the nonhomogeneous semigroup

$$(P_{s,t}^V f)(x) = \mathbb{E}_x\left[e^{\int_0^{t-s} V_{s+r}(X_r)\,dr} f(X_{t-s})\right], \quad \text{for all } t \geqslant s \geqslant 0.$$

Then, $\mathbb{E}_\nu\left[e^{\int_0^t V(r,X_r)\,dr}\right] = \langle P_{0,t}^V 1, 1\rangle_\nu$. To bound $\langle P_{0,t}^V 1, 1\rangle_\nu$, we start with the Cauchy-Schwarz inequality

$$\big\langle P_{0,t}^V 1, 1\big\rangle_\nu \leqslant \big\langle P_{0,t}^V 1, P_{0,t}^V 1\big\rangle_\nu^{1/2}.$$

In the remaining of the proof we will look at $\big\langle P_{s,t}^V 1, P_{s,t}^V 1\big\rangle_\nu$ as a function of s and apply Gronwall's inequality. First of all, notice that

$$\partial_s\big\langle P_{s,t}^V 1, P_{s,t}^V 1\big\rangle_\nu = -2\,\big\langle L_s^V P_{s,t}^V 1, P_{s,t}^V 1\big\rangle_\nu. \tag{1.17}$$

To show this, we differentiate under the integral sign on

$$\left\langle P_{s,t}^{V}1, P_{s,t}^{V}1\right\rangle_{\nu} = \int_E (P_{s,t}^{V}1)^2(x)\, d\nu(x)$$

and use that $\partial_s (P_{s,t}^{V}1)^2(x) = 2(P_{s,t}^{V}1)(x)\partial_s(P_{s,t}^{V}1)(x)$ to conclude that via backward Kolmogorov equation.

Since $g(x) = (P_{s,t}^{V}1)(x)/\|P_{s,t}^{V}1\|_2$ is such that $\|g\|_2 = 1$, by equation (1.16), we have

$$2\Gamma_s \geqslant \frac{2\left\langle L_s^V P_{s,t}^{V}1, P_{s,t}^{V}1\right\rangle_{\nu}}{\left\langle P_{s,t}^{V}1, P_{s,t}^{V}1\right\rangle_{\nu}}.$$

Plugging into equation (1.17),

$$\partial_s\left\langle P_{s,t}^{V}1, P_{s,t}^{V}1\right\rangle_{\nu} \geqslant (-2\Gamma_s)\left\langle P_{s,t}^{V}1, P_{s,t}^{V}1\right\rangle_{\nu}.$$

Applying Gronwall's inequality, we get

$$\left\langle P_{0,t}^{V}1, P_{0,t}^{V}1\right\rangle_{\nu} \leqslant \left\langle P_{t,t}^{V}1, P_{t,t}^{V}1\right\rangle_{\nu} \exp\left\{\int_0^t 2\Gamma_s\, ds\right\} = \exp\left\{\int_0^t 2\Gamma_s\, ds\right\},$$

where the last equality follows from the fact that $P_{t,t}^{V}1(x) = 1$ and it finishes the proof. $\square$

1.4. Radon-Nikodyn derivative between two continuous-time Markov jump process

In this section, we will state the Radon-Nikodym derivative between two nonhomogeneous jump process X and $\overline{X}$ with the same state space E. The main result of this section, Proposition 1.16, can be found on Appendix A of [3], but here we are making a new analyse from a whole different perspective. We start by looking at properties of X, but the same is valid for $\overline{X}$. Let the infinitesimal generator of this process be

$$(L_s f)(x) = \lambda(s,x)\int_E [f(y) - f(x)] P_s(x, dy),$$

where $\lambda : [0, +\infty) \times E \to \mathbb{R}$ is assumed to be nonnegative and bounded, as usual.

Remark. *In terms of the construction of this generalized jump process from a skeleton Markov chain ξ_n (see section 2 on appendix 1 of [24] for the classical construction), we consider that τ_{n+1} is distributed according to an exponential law of parameter $\lambda(T_n + t, \xi_n)$, where $T_n = \tau_1 + \cdots + \tau_n$. This means that the* *density function of each τ_{n+1} is*

$$f(t) = \begin{cases} \lambda(T_n + t, \xi_n) e^{-\int_0^t \lambda(T_n+s,\xi_n)ds}, & t \geqslant 0 \\ 0, & t < 0 \end{cases},$$

which is a probability density if we assume $\int_a^\infty \lambda(s,x) = \infty$, for all $a \in \mathbb{R}$ and $x \in E$.

From these time-changing infinitesimal generator, we want to set a nonhomogeneous Markov semigroup. The following result brings a natural extension of the semigroup defined in the case where L does not depend on time, since in this case we can see the expoent tL as the integral of a constant L on a time interval of size t.

Proposition 1.15. *The nonhomogeneos semigroup associated with the infinitesimal generator L_s is*

$$P_{s,t}(f) = \left(e^{\int_s^t L_r dr}\right)(f).$$

Furthermore, this semigroup satisfies the Kolmogorov equations:

(i) Kolmogorov forward equation:

$$\partial_t P_{s,t}(f) = P_{s,t}(L_t f);$$

(ii) Kolmogorov backward equation:

$$\partial_s P_{s,t}(f) = -L_s(P_{s,t} f).$$

Proof. First, notice that $P_{s,t}$ satisfies the semigroup property

$$(P_{s,t}P_{t,u})(f) = \left(e^{\int_s^t L_r dr} e^{\int_t^u L_r dr}\right)(f) = P_{s,u} f.$$

Morover, the positivity of the exponential along with the properties of integration implies that $P_{s,t}$ is a Markov semigroup (see Definition 1.2).

Now, observe that

$$\lim_{h \downarrow 0} \frac{P_{t,t+h} f - f}{h} = \lim_{h \downarrow 0} \left(\frac{e^{\int_t^{t+h} L_r dr} - 1}{h}\right)(f) = L_t f,$$

what allow us to prove the Kolmogorov equations:

$$\partial_t P_{s,t}(f) = \lim_{h \downarrow 0} \frac{P_{s,t+h} f - P_{s,t} f}{h} = \lim_{h \downarrow 0} \frac{(P_{s,t}P_{t,t+h}) f - P_{s,t} f}{h}$$

$$= \lim_{h\downarrow 0} P_{s,t}\left(\frac{P_{t,t+h}f - f}{h}\right) = P_{s,t}(L_t f)$$

and

$$\begin{aligned}
\partial_s P_{s,t}(f) &= \lim_{h\downarrow 0} \frac{P_{s+h,t}f - P_{s,t}f}{h} = \lim_{h\downarrow 0} \frac{P_{s+h,t}f - (P_{s,s+h}P_{s+h,t})f}{h} \\
&= \lim_{h\downarrow 0} \left(\frac{1 - P_{s,s+h}}{h}\right)(P_{s+h,t}f) = -L_s(P_{s,t}f).
\end{aligned}$$

□

Now, we consider the infinitesimal generator

$$(\overline{L}_s f)(x) = \overline{\lambda}(s,x)\int_E [f(y) - f(x)]\overline{P}_s(x,dy)$$

of $\overline{X}$ satisfying the same properties above. We would like to compute the Radon-Nikodym derivative of $\mathbb{P}$ with respect to $\overline{\mathbb{P}}$ restricted to the σ-field $\mathcal{F}_t$. For this, on the following result, we consider k instants of time $0 \leqslant t_1 < \cdots < t_k \leqslant t$ of the interval $[0,t]$, a function $F : \mathbb{R}^k \to \mathbb{R}$ and we get the Radon-Nikodym derivative by the equality

$$\mathbb{E}_x[F(X_{t_1},\ldots,X_{t_k})] = \overline{\mathbb{E}}_x\left[\left(\left.\frac{d\mathbb{P}}{d\overline{\mathbb{P}}}\right|_{\mathcal{F}_t}\right)F(X_{t_1},\ldots,X_{t_k})\right], \tag{1.18}$$

because the function $F(X_{t_1},\ldots,X_{t_k})$ is $\mathcal{F}_t$-measurable.

Proposition 1.16. *This Radon-Nikodym derivative is given by*

$$\left.\frac{d\mathbb{P}}{d\overline{\mathbb{P}}}\right|_{\mathcal{F}_t} = \exp\left\{\int_0^t [\overline{\lambda}(s,X_s) - \lambda(s,X_s)]ds + \sum_{s\leqslant t}\log\left(\frac{\lambda(s,X_{s-})}{\overline{\lambda}(s,X_{s-})}\frac{dP_s}{d\overline{P}_s}(X_{s-},X_s)\right)\right\},$$

where, for a fixed x, $\frac{dP_s}{d\overline{P}_s}(x,y)$ is the Radon-Nikodym derivative of $P_s(x,dy)$ with respect to $\overline{P}_s(x,dy)$.

Remark. *Observe that the above sum is well defined, because it is not equal to zero just on the jumps, which are almost surely finite.*

Proof. Let T_n be the time of the jump n of the process. Partitioning with respect to the number of jumps until the time t, we get

$$\begin{aligned}
\mathbb{E}_x[F(X_{t_1},\ldots,X_{t_k})] &= \sum_{n=0}^{\infty}\mathbb{E}_x\left[F(X_{t_1},\ldots,X_{t_k})\mathbb{1}_{[T_n\leqslant t<T_{n+1}]}\right] \\
&= \sum_{n=0}^{\infty}\mathbb{E}_x\left[F_n(\xi_1,T_1,\ldots,\xi_n,T_n)\mathbb{1}_{[T_n\leqslant t<T_{n+1}]}\right],
\end{aligned}$$

because on $[T_n \leqslant t < T_{n+1}]$ the function F could be writen as a function of $(\xi_1, T_1, \ldots, \xi_n, T_n)$, where $\{\xi_k\}_{k\in\mathbb{N}}$ is the discrete-time skeleton chain. Remember that the pair $\{(\xi_n, T_n)\}_n$ is markovian, then by the proprieties of conditional expectation,

$$
\begin{aligned}
& \mathbb{E}_x[F(X_{t_1}, \ldots, X_{t_k})] \\
= & \sum_{n=0}^{\infty} \mathbb{E}_x \left[\mathbb{E}\left[F_n(\xi_1, T_1, \ldots, \xi_n, T_n) \mathbb{1}_{[T_n \leqslant T < T_{n+1}]} \middle| \sigma(\xi_1, T_1, \ldots, \xi_n, T_n) \right] \right] \\
= & \sum_{n=0}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{[T_n \leqslant T]} F_n(\xi_1, T_1, \ldots, \xi_n, T_n) \mathbb{E}\left[\mathbb{1}_{[T < T_{n+1}]} \middle| \sigma(\xi_n, T_n) \right] \right]
\end{aligned}
$$

Denoting by G_n the function on the expectation above, we can compute

$$
\begin{aligned}
& \mathbb{E}_x[G_n(\xi_1, T_1, \ldots, \xi_n, T_n)] \\
= & \int_E \cdots \int_E \int_0^\infty \cdots \int_0^\infty \prod_{j=0}^{n-1} \left(\frac{\lambda(T_j + t_{j+1}, x_j)}{e^{\int_0^{t_{j+1}} \lambda(T_j+s, x_j) ds}} P_{T_j + t_{j+1}}(x_j, dx_{j+1}) dt_{j+1} \right) G_n \\
= & \int_E \cdots \int_E \int_0^\infty \cdots \int_0^\infty \prod_{j=0}^{n-1} \left(\frac{\overline{\lambda}(T_{j+1}, x_j)}{e^{\int_0^{t_{j+1}} \overline{\lambda}(T_j+s, x_j) ds}} \overline{P}_{T_{j+1}}(x_j, dx_{j+1}) dt_{j+1} \right) \\
& G_n \prod_{k=0}^{n-1} \left(\frac{\lambda(T_{k+1}, x_k)}{\overline{\lambda}(T_{k+1}, x_k)} e^{\int_{T_k}^{T_{k+1}} [\overline{\lambda}(u, x_k) - \lambda(u, x_k)] du} \frac{dP_{T_{k+1}}}{d\overline{P}_{T_{k+1}}}(x_k, dx_{k+1}) \right) \\
= & \overline{\mathbb{E}}_x \left[G_n \prod_{k=0}^{n-1} \left(\frac{\lambda(T_{k+1}, \xi_k)}{\overline{\lambda}(T_{k+1}, \xi_k)} e^{\int_{T_k}^{T_{k+1}} [\overline{\lambda}(u, \xi_k) - \lambda(u, \xi_k)] du} \frac{dP_{T_{k+1}}}{d\overline{P}_{T_{k+1}}}(\xi_k, d\xi_{k+1}) \right) \right].
\end{aligned}
$$

Now, we can use this expression to do the same computations, in the inverse order, to come back to F and get, by equation (1.18), the expected expression for the Radon-Nikodym derivative. $\square$

As consequence of this, we get the Proposition 2.6 of Appendix 1 of [24]:

Corollary 1.17. *For a function λ that does not depend on time, we get*

$$
\left.\frac{d\mathbb{P}}{d\overline{\mathbb{P}}}\right|_{\mathcal{F}_t} = \exp\left\{ \int_0^t [\overline{\lambda}(X_s) - \lambda(X_s)] ds + \sum_{s \leqslant t} \log\left(\frac{\lambda(X_{s-})}{\overline{\lambda}(X_{s-})} \frac{dP}{d\overline{P}}(X_{s-}, X_s) \right) \right\}.
$$

Appendices

We will finish this chapter with some appendix sections related to it. These results are presented here in order to obtain a more complete chapter, so we recommend that the reader skips this part in a first reading. The tools in this part are properly called in the text if needed.

1.A. Some results for functions without time dependence

This section is dedicated to explore the case where the functions F_s does not depend on time s. Thus, we replace the function $F : \mathbb{R}_+ \times E \to \mathbb{R}$ satisfying the Assumption 1.1 by $f \in C_b(E)$. The proofs of some results will be omitted as they are just an adaptation of the results we already proved.

We start by stating the form of the Dynkin martingale in this context. The following result is a version of Theorem 1.8:

Theorem 1.18. *For a function $f \in C_b(E)$, the Dynkin martingale*

$$M^f(t) = f(X_t) - f(X_0) - \int_0^t Lf(X_s)\, ds \tag{1.19}$$

is a $\mathcal{F}_t$-martingale.

Conversely, the next result shows that the infinitesimal generator L is the only operator who turns the equation (1.19) into a martingale.

Proposition 1.19. *If $f \in C_b(E)$ and exists a function $g \in C_b(E)$ such that*

$$f(X_t) - f(X_0) - \int_0^t g(X_s) ds$$

is a $\mathcal{F}_t$-martingale, then $Lf = g$.

Proof. Notice that, for every g, this martingale has expectation equals to zero at time $t = 0$. Consequently, by martingales properties, this expectation is equal to zero at every time t. Then, for every x and t,

$$\mathbb{E}_x \left[f(X_t) - f(X_0) - \int_0^t g(X_u) du \right] = 0.$$

Using the linearity of the expetcation along with Fubini's theorem, we have

$$\mathbb{E}_x[f(X_t)] - \mathbb{E}_x[f(X_0)] - \int_0^t \mathbb{E}_x[g(X_s)] ds = 0.$$

This equality can be rewrited as

$$P_t f(x) - f(x) - \int_0^t P_s g(x) ds = 0.$$

We finish the proof by showing that, as consequence of the expression above, g is equal to the derivative of $P_t f$ at time $t = 0$. In fact,

$$\left|\left| \frac{1}{t}(P_t f - f) - g \right|\right| = \left|\left| \frac{1}{t} \int_0^t (P_s g - g) ds \right|\right| \leqslant \frac{1}{t} \int_0^t ||P_s g - g|| ds,$$

which vanishes when $t \to 0$. $\square$

Going on with the adaptations, now we show a version of Proposition 1.9. To keep it simple, we will enunciate this result in the case of Markov jump process, where the time derivative term vanishes as time dependence on X_s has no influence.

Proposition 1.20. *In the case of Markov jump process, if $f \in C_b(E)$, then*

$$N^f(t) = (M^f(t))^2 - \int_0^t \Gamma(f,f)(X_s)\, ds\,,$$

where Γ is the Carré du champ operator, which is a bilinear map

$$\Gamma(f,g)(x) = L(fg)(x) - f(x)Lg(x) - g(x)Lf(x)\,, \quad x \in E\,,$$

defined for every $(f,g) \in C_b(E) \times C_b(E)$.

As consequence of this, we know that, in this context, the quadratic variation of M^f is given by integration of the Carré du champ operator on the pair (f,f). Futhermore, we can also extend this result for the quadratic covariation by

Lemma 1.21. *If X is a Markov jump process and $f, g \in C_b(E)$, the quadratic covariation of the Dynkin's martingales M^f and M^g is given by*

$$\langle M^f, M^g \rangle_t = \int_0^t \Gamma(f,g)(X_s) ds.$$

Proof. The quadratic covariation of two martingales M, N is calculated by

$$\langle M, N \rangle_t = \frac{1}{4}\left(\langle M+N \rangle_t - \langle M-N \rangle_t\right),$$

where $\langle \cdot \rangle_t$ denotes the quadratic variation. If we notice that $M^f_t \pm M^g_t = M^{f\pm g}_t$, then we can write

$$\langle M^f, M^g \rangle_t = \frac{1}{4}\left(\langle M^{f+g}_t \rangle_t - \langle M^{f-g}_t \rangle_t\right).$$

We know the quadratic variation of both this processes, then

$$\langle M^f, M^g \rangle_t = \frac{1}{4}\int_0^t \left(\Gamma(f+g, f+g) - \Gamma(f-g, f-g)\right)(X_s) ds.$$

Using the definition of the Carré du champ operator and the linearity of L, we can compute the expression above. After some cancelations, we get

$$\langle M^f, M^g \rangle_t = \frac{1}{4}\int_0^t \left(4L(fg) - 4fLg - 4gLf\right)(X_s) ds = \int_0^t \Gamma(f,g)(X_s) ds.$$

□

Another result that we can easily adapt to this case is Theorem 1.10. To prove this version, we just need to follow the same steps, but changing the Dynkin martingale by the one presented in Theorem 1.18. The result we ended up with is

Theorem 1.22. *Let f be a function on $C_b(E)$. Then,*

$$f(X_t)\, e^{\int_0^t V_r(X_r)\, dr} - \int_0^t e^{\int_0^s V_r(X_r)\, dr}[f(X_s)V_s(X_s) + Lf(X_s)]\, ds$$

is a martingale.

The Kolmogorov equations can also be stated in this case. In the same way we used Theorem 1.10 in the proof of Proposition 1.12, here we can use the above theorem to get

Proposition 1.23. *If we consider a function $f \in C_b(E)$, the Kolmogorov equations, for all $t > s \geqslant 0$ and $x \in E$, becomes*

(i) Kolmogorov forward equation:

$$\partial_t \left(P^V_{s,t} f(x)\right) = P^V_{s,t} L^V_t f(x)\,;$$

(ii) Kolmogorov backward equation:

$$\partial_s \left(P^V_{s,t} f(x)\right) = -L^V_s P^V_{s,t} f(x)\,.$$

1.B. Extra results for other perturbations

In this section, we will expose some extra theory about other types of perturbations of Markov process that are somehow related to the ones presented before. For instance, the first subsection allow us to get the exponential martingale, defined in Section 1.2, as a formula for the Radon-Nikodym derivative between the probabilities induced on Skorohod space by two different Markov jump processes. The second subsection give us the perturbation we need to do in order to get a process whose infinitesimal generator differs from the original by the Carré du Champ operator. This is a consequence of the results without time dependence, presented on Section 1.A.

1.B.1 One perturbed process

Define the following perturbation of process X_t by its infinitesimal generator

$$(\mathcal{L}^{F_t} f)(x) = \lambda(x) \int_E e^{F_t(y)-F_t(x)}[f(y) - f(x)]\, P(x, dy)\,, \tag{1.20}$$

with $f \in C_b(E)$, $t \geqslant 0$. Denote by $\mathbb{P}^F$ the probability induced in the Skorohod space D by the process with generator L^{F_t}.

In the present section we present how the Kolmogorov equations for the perturbed process (presented in Section 1.3.1) can be used to get an explicit formula for the Radon-Nikodym derivative of $\mathbb{P}^F$ with respect to $\mathbb{P}$ restricted to the σ-field $\mathcal{F}_t$:

$$\left.\frac{d\mathbb{P}^F}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left\{F_t(X_t) - F_0(X_0) - \int_0^t e^{-F_s(X_s)}(\partial_s + L)e^{F_s(X_s)}\, ds\right\}, \quad (1.21)$$

where the expression in the right-hand side above is the exponential martingale $\mathbb{M}^F(t)$, defined in equation (1.10).

To prove it we define explicitly the probability $\mathbb{P}^F$. For a fix time $T > 0$ and, for each $x_0 \in E$, we define on $\mathcal{F}_T$ the probability measure $\mathbb{P}^F_{x_0}$ by

$$\mathbb{E}^F_{x_0}[G] = \mathbb{E}_{x_0}[G\mathbb{M}^F(T)]\,,$$

for all bounded function G $\mathcal{F}_T$-measurable.

Then we prove that the process which induce $\mathbb{P}^F$ has infinitesimal generator $\mathcal{L}^{F_t}$, defined in equation (1.20). It follows from the next lemma.

Lemma 1.24. *For the probability measure defined above we have that, for all $G \in \mathcal{F}_T$, the conditional expectation is*

$$\mathbb{E}^F_{x_0}[G|\mathcal{F}_s] = \frac{\mathbb{E}_{x_0}[G\mathbb{M}^F(T)|\mathcal{F}_s]}{\mathbb{M}^F(s)}.$$

Proof. We denote $Y = \mathbb{E}^F_{x_0}[G|\mathcal{F}_s]$. Then, by the definition of conditional expectation, for all $\Gamma \in \mathcal{F}_s$,

$$\mathbb{E}^F_{x_0}[Y\mathbb{1}_\Gamma] = \mathbb{E}^F_{x_0}[G\mathbb{1}_\Gamma].$$

As $Y \in \mathcal{F}_s$ and $\mathbb{1}_\Gamma \in \mathcal{F}_s$, we have $Y\mathbb{1}_\Gamma \in \mathcal{F}_s$ and the left hand-side of the equality above is $\mathbb{E}_{x_0}[Y\mathbb{1}_\Gamma\mathbb{M}^F(s)]$. On the other hand, $G \in \mathcal{F}_T$ and $\mathbb{1}_\Gamma \in \mathcal{F}_s \subseteq \mathcal{F}_T$ implies that $G\mathbb{1}_\Gamma \in \mathcal{F}_T$ which allow us to rewrite the right side as $\mathbb{E}_{x_0}[G\mathbb{1}_\Gamma\mathbb{M}^F(T)]$. Then, for all $\Gamma \in \mathcal{F}_s$,

$$\mathbb{E}_{x_0}[Y\mathbb{M}^F(s)\mathbb{1}_\Gamma] = \mathbb{E}_{x_0}[G\mathbb{M}^F(T)\mathbb{1}_\Gamma].$$

This concludes the proof because

$$Y\mathbb{M}^F(s) = \mathbb{E}_{x_0}[Y\mathbb{M}^F(s)|\mathcal{F}_s] = \mathbb{E}_{x_0}[G\mathbb{M}^F(T)|\mathcal{F}_s].$$

□

As consequence of the previous Lemma and Markov property, we have

$$\mathbb{E}^F_{x_0}[f(X_t)|\mathcal{F}_s] = \mathbb{E}_{x_0}\left[f(X_t)\frac{\mathbb{M}^F_t}{\mathbb{M}^F(s)}\middle| \mathcal{F}_s\right] = \mathbb{E}_{x_0}\left[f(X_t)\frac{\mathbb{M}^F_t}{\mathbb{M}^F(s)}\middle| X_s\right]. \quad (1.22)$$

Lemma 1.25. *We get the Markov property for the process associated with the probability* $\mathbb{P}^F_{x_0}$*, that is,*

$$\mathbb{E}^F_{x_0}[f(X_t)|\mathcal{F}_s] = \mathbb{E}^F_{x_0}[f(X_t)|X_s]\,.$$

Proof. Define $Y = \mathbb{E}^F_{x_0}[f(X_t)|\mathcal{F}_s]$. Notice that, by equation (1.22), $Y \in \sigma(X_s)$, then $Y = \mathbb{E}^F_{x_0}[Y|X_s]$ and we just need to prove that, for all $\Gamma \in \sigma(X_s)$,

$$\mathbb{E}^F_{x_0}[f(X_t)\mathbb{1}_\Gamma] = \mathbb{E}^F_{x_0}[Y\mathbb{1}_\Gamma].$$

We start by computing the right hand-side above. By definition,

$$\mathbb{E}^F_{x_0}[Y\mathbb{1}_\Gamma] = \mathbb{E}_{x_0}[Y\mathbb{1}_\Gamma\mathbb{M}^F(s)] = \mathbb{E}_{x_0}[E^F_{x_0}[f(X_t)|\mathcal{F}_s]\mathbb{1}_\Gamma\mathbb{M}^F(s)]\,.$$

Using Lemma 1.24 and the fact of $\mathbb{1}_\Gamma \in \mathcal{F}_s$, we get

$$\begin{aligned}\mathbb{E}^F_{x_0}[Y\mathbb{1}_\Gamma] &= \mathbb{E}_{x_0}[\mathbb{E}_{x_0}[f(X_t)\mathbb{M}^F(t)|\mathcal{F}_s]\mathbb{1}_\Gamma] \\ &= \mathbb{E}_{x_0}[f(X_t)\mathbb{1}_\Gamma\mathbb{M}^F(t)] \\ &= \mathbb{E}^F_{x_0}[f(X_t)\mathbb{1}_\Gamma]\,,\end{aligned}$$

because $f(X_t)\mathbb{1}_\Gamma \in \mathcal{F}_t$. $\square$

Now, with this Markov property, we can define a family of operators

$$Q^F_{s,t}f(X_s) = \mathbb{E}^F_{x_0}[f(X_t)|X_s].$$

To characterize this operators, we compute

$$\begin{aligned}\mathbb{E}^F_{x_0}[f(X_t)|X_s] &= \mathbb{E}^F_{x_0}[f(X_t)|\mathcal{F}_s] = \mathbb{E}_{x_0}\left[f(X_t)\frac{\mathbb{M}^F_t}{\mathbb{M}^F_s}\middle| X_s\right] \\ &= \mathbb{E}_{x_0}\left[f(X_t)e^{F_t(X_t)-F_s(X_s)}e^{-\int_s^t e^{-F_r(X_r)}(\partial_r+L)e^{F_r(X_r)}dr}\middle| X_s\right].\end{aligned}$$

If we denote $V_r(z) = -e^{-F_r}(z)(\partial_r + L)e^{F_r(z)}$, the homogeneity of the non-perturbed Markov process and a change of variables give us that

$$\mathbb{E}^F_{x_0}[f(X_t)|X_s] = \mathbb{E}_{X_s}\left[f(X_{t-s})e^{F_t(X_{t-s})-F_s(X_0)}e^{\int_s^t V_r(X_{r-s})dr}\right]$$

$$
\begin{aligned}
&= e^{-F_s(X_s)}\mathbb{E}_{X_s}\left[e^{\int_0^{t-s} V_{u+s}(X_u)du} e^{F_t(X_{t-s})} f(X_{t-s})\right] \\
&= e^{-F_s(X_s)} P^V_{s,t}(fe^{F_t})(X_s).
\end{aligned}
$$

Then, for all $x \in E$, we have this characterization:

$$
\begin{aligned}
Q^F_{s,t}f(x) &= e^{-F_s(x)} P^V_{s,t}(fe^{F_t})(x) \\
&= \mathbb{E}_x\left[e^{\int_0^{t-s} V_{u+s}(X_u)du} e^{F_t(X_{t-s})-F_s(X_0)} f(X_{t-s})\right].
\end{aligned}
$$

We will now prove that the operators $Q_{s,t}$ forms a semigroup. Taking $0 < s < t < u$, we have

$$
Q^F_{s,t}Q^F_{t,u}(x) = e^{-F_s(x)}\mathbb{E}_x\left[e^{\int_0^{t-s} V_{v+s}(X_v)dv} e^{F_t(X_{t-s})} Q^F_{t,u}f(X_{t-s})\right],
$$

which the expectation part can be rewrite, by the Markov property, as

$$
\mathbb{E}_x\left[e^{\int_0^{t-s} V_{v+s}(X_v)dv}\mathbb{E}_x\left[e^{\int_0^{u-t} V_{v+t}(X_{v+t-s})dv} e^{F_u(X_{u-t+t-s})} f(X_{u-t+t-s})\middle| \mathcal{F}_{t-s}\right]\right].
$$

Using the properties of the conditional expectation and making a change of variables, we prove that

$$
Q^F_{s,t}Q^F_{t,u}(x) = e^{-F_s(x)}\mathbb{E}_x\left[e^{\int_0^{u-s} V_{r+s}(X_r)dr} e^{F_u(X_{u-s})} f(X_{u-s})\right] = Q_{s,u}f(x).
$$

Finally, we will study the Kolmogorov equation for $Q^F_{s,t}$:

$$
\begin{aligned}
\partial_t(Q^F_{s,t}f)(x) &= \partial_t\left(e^{F_s(x)} P^V_{s,t}(fe^{F_t})(x)\right) = e^{-F_s(x)}\partial_t P^V_{s,t}(fe^{F_t})(x) \\
&= e^{-F_s(x)}\left\{P^V_{s,t}(L^V_t(fe^{F_t}))(x) + P^V_{s,t}(\partial_t(fe^{F_t}))\right\}.
\end{aligned}
$$

Notice that, by the definition of V,

$$
L^V_t(fe^{F_t})(z) = L(fe^{F_t})(z) + V_t(z)f(z)e^{F_t(z)} = L(fe^{F_t})(z) - (L+\partial_t)e^{F_t}(z),
$$

then the previous equation is

$$
\partial_t(Q^F_{s,t}f)(x) = e^{-F_s(x)}\left\{P^V_{s,t}(L(fe^{F_t}) - fLe^{F_t})\right\}.
$$

By the definitions of $P^V_{s,t}$ and L, we can compute the bracket part as

$$
\mathbb{E}_x\left[e^{\int_s^t V_r(X_r)dr}\lambda(X_t)\int_E e^{F_t(y)}(f(y) - f(X_t))P(X_t, dy)\right]
$$

$$
\begin{aligned}
&= \mathbb{E}_x\left[e^{\int_s^t V_r(X_r)dr} e^{F_t(X_t)} \lambda(X_t) \int_E e^{F_t(y)-F_t(X_t)} (f(y)-f(X_t)) P(X_t, dy)\right] \\
&= P^V_{s,t}(e^{F_t} \mathcal{L}^{F_t}(f))(x).
\end{aligned}
$$

Finally, by the definition of $Q^F_{s,t}$, we conclude that

$$
\partial_t(Q^F_{s,t} f)(x) = Q^F_{s,t}(\mathcal{L}^{F_t}(f))(x).
$$

1.B.2 Another pertubated processes

Let X_t be a Markov process on $(\Omega, \mathcal{F}_t, \mathbb{P}_x)$, where $\Omega = C(\mathbb{R}_+, \mathbb{R}^d)$. Given a pair of functions (f, F), suppose that $D_t = e^{f(X_t)-f(X_0)-\int_0^t F(X_s)dx}$ is a continuous martingale for every x. We can define a new probability $\mathbb{P}^f_x$ on $\mathcal{F}_\infty$ by $\mathbb{P}^f_x = D_t \cdot \mathbb{P}_x$ on $\mathcal{F}_t$, see [39].

We have a general method to find such pairs (f, F). For each f, denote by X^f_t the pertubated process whose law is given by the probability $\mathbb{P}^f_x$ as defined in the beginning of the section, with $F = Lf + \frac{1}{2}\Gamma(f, f)$. In this setting, we get

Proposition 1.26. *If L is the infinitesimal generator of X, the infinitesimal generator of X^f is equal to $L + \Gamma(f, \cdot)$.*

Proof. Given a function $g \in C_b(E)$, we denote M^g_t by the Dynkin martingale. As consequence of the Girsanov's theorem, see Theorem 1.7 in [39, Chapter VIII], $M^g_t - \langle M^f, M^g \rangle_t$ is a $\mathbb{P}^f$-martingale. By Lemma 1.21, we conclude that

$$
g(X_t) - g(X_0) - \int_0^t \left(Lg + \Gamma(f, g)\right)(X_s)ds
$$

is a $\mathbb{P}^f$-martingale. This ends the proof by Proposition 1.19. □

1.C. Basic definitions and results

We start this section with the definition of a kernel. As you can see, this is the transition probability in this context (similar to the one of the discrete-time case) and it give us a measure to integrate on E.

Definition 1.27. *A kernel N on E is a map from $E \times \mathcal{E}$ into $[0, \infty) \cup \{+\infty\}$ such that*

(i) for each $x \in E$, the map $A \to N(x, A)$ is a measure on $\mathcal{E}$;

(ii) for each $A \in \mathcal{E}$, the map $x \to N(x, A)$ is $\mathcal{E}$-measurable.

Furthemore, this kernel is called a transition probability if $N(x, E) = 1$, for all $x \in E$.

If f is a positive measurable function and N is a kernel, we define

$$Nf(x) = \int_E N(x, dy)f(y).$$

Notice that Nf is also a positive measurable function. Besides that, if we have two kernels M and N, then

$$MN(x, A) := \int_E M(x, dy)N(y, A)$$

is also a kernel.

Now, we will expose some versions of the integration by parts formula. The first one will be the case of two functions of bounded variation, given by Proposition 4.5 on [39, Chapter 0].

Proposition 1.28. *Suppose that A and B are two functions of bounded variation defined on $[0, +\infty)$. For any time $t \geqslant 0$,*

$$A_t B_t = A_0 B_0 + \int_0^t A_s dB_s + \int_0^t B_{s-} dA_s,$$

where B_{s-} denote the limit of a sequence of times increasing to s.

Proof. Let μ and ν denote measures associated with A and B, respectively. Notice that both sides of this equality are equal to $\mu \otimes \nu$ on the square $[0, t]^2$. If we divide this square in two triangles by the diagonal, each integral in the right side is the measure of one triangle while the first term represents the origin. The B_{s-} is used to exclude the diagonal on one half to not be counted twice. □

The second version of the integration by parts formula we state here will be a stochastic version, between two semimartingales, specifically in the case where one of them have bounded variation. This result is a immediate consequence of the classical integration by parts formula of stochastic calculus, see Proposition 3.1 of [39, Chapter IV].

Proposition 1.29. *Let X and Y be two continuous semimartingales. If one of this martingales is of bounded variation, then*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s.$$

Proof. By Proposition 3.1 of [39, Chapter IV], the classical stochastic version of the integration by parts formula is given by

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t.$$

To finish the proof, we just need to notice that, when one of this processes have bounded variation, the covariation term above vanishes. □

We want to prove another result related to integration by parts on martingales theory, but first, we need the following lemma:

Lemma 1.30. *Let $\{M_t,\, t \geqslant 0\}$ be a càdlàg process, which is bounded in the following sense: for each $t > 0$ there exists a constant $C_t > 0$ such that $|M_r| \leqslant C_t$, for any $r \in [0,t]$. And, let $\{Z_t,\, t \geqslant 0\}$ be a bounded variation process, with $dZ_r = Z'_r dr$ and $|Z'_r| \leqslant A_t$, for any $r \in [0,t]$, for some $A_t > 0$. Then, for all $0 \leqslant s \leqslant t$ there exists, $\{t_0, \dots, t_N\}$, a partition of the $[s,t]$ such that*

$$\sum_{j=1}^{N} M_{t_j}(Z_{t_j} - Z_{t_{j-1}}) \xrightarrow{L_1} \int_s^t M_r\, dZ_r, \tag{1.23}$$

as $N \to \infty$.

Proof. Since

$$\left| \sum_{j=1}^{N} M_{t_j}(Z_{t_j} - Z_{t_{j-1}}) \right| \leqslant C_t \sum_{j=1}^{N} |Z_{t_j} - Z_{t_{j-1}}| \leqslant C_t\, A_t\, (t-s)$$

and by the Dominated Convergence Theorem, it is enough to prove that convergence in equation (1.23) is almost surely. In order to do this, we note that

$$M_r \mathbb{1}_{(s,t]}(r) = \sum_{j=1}^{N} (M_r - M_{t_j}) \mathbb{1}_{(t_{j-1},t_j]}(r) + \sum_{j=1}^{N} M_{t_j} \mathbb{1}_{(t_{j-1},t_j]}(r).$$

Then, integrating by dZ_r in the interval $[s,t]$, we just need to prove that, when $N \to \infty$,

$$\sum_{j=1}^{N} \int_{t_{j-1}}^{t_j} (M_r - M_{t_j})\, dZ_r \xrightarrow{a.s.} 0. \tag{1.24}$$

By the hypothesis $\{M_t,\, t \geqslant 0\}$ has càdlàg trajectories, then a trajectory of $\{M_t,\, t \geqslant 0\}$ has a finite number of jumps in a compact interval. Thus, to handle with the sum above, for a fixed trajectory, we split it in two parts: the first one with terms where the trajectory of $\{M_t,\, t \geqslant 0\}$ is uniformly continuous and the other one includes the remaining terms. Note that the quantity of terms where the trajectory of $\{M_t,\, t \geqslant 0\}$ has jumps is bounded by the number of jumps of it in the time interval $[s,t]$. Then it is possible to control the number of remaining terms.

In order to write with precision this idea, we need to introduce some notation. For all below, we work ever with the same fixed trajectory of $\{M_t,\, t \geqslant 0\}$.

Denote by $s_1, \ldots, s_m$ the times where the trajectory of $\{M_t,\, t \geqslant 0\}$ jumps in the time interval $[s,t]$. And let us consider the following partition of $[s,t]$:

$$t_j^n = s + (t-s)\tfrac{j}{2^n}, \quad j \in \{0,1,\ldots,2^n\}\,. \tag{1.25}$$

Let $\varepsilon > 0$ and take n_1 such that $\frac{m(t-s)}{2^{n_1}} < \varepsilon$. Define the index set

$$\Lambda^1_{n_1} = \left\{ j \in D_{n_1};\ \text{ there is some } i \in \{1,\ldots,m\} \text{ such that } s_i \in (t^{n_1}_{j-1}, t^{n_1}_j] \right\},$$

where $D_{n_1} = \{1,\ldots,2^{n_1}\}$, and the compact subset of $[s,t]$

$$\mathcal{K} = [s,t] \backslash \bigcup_{\ell \in \Lambda^1_{n_1}} (t^{n_1}_{\ell-1}, t^{n_1}_{\ell+1})\,.$$

Since $r \mapsto M_r$ is uniformly continuous in $\mathcal{K}$, there exists $\delta_0 > 0$ such that for every $r, s \in \mathcal{K}$ with $|r-s| < \delta_0$, we have $|M_r - M_s| < \varepsilon$. Then, let us choose $n_2 \geqslant n_1$ such that $\frac{t-s}{2^{n_2}} < \delta_0$. For all $n \geqslant n_2$, define the index set

$$\Lambda^2_n = \left\{ j \in D_n;\ \text{ there exists some } \ell \in \Lambda^1_{n_1} \text{ such that } (t^n_{j-1}, t^n_j] \subset (t^{n_1}_{\ell-1}, t^{n_1}_\ell] \right\}.$$

Finally, using Λ^2_n we can split the sum

$$\sum_{j=1}^{2^n} \int_{t^n_{j-1}}^{t^n_j} (M_r - M_{t^n_j})\, dZ_r \tag{1.26}$$

in two parts. The first one is the sum in $\{1,\ldots,2^n\}\backslash\Lambda^2_n$, and for study it we observe that for all $j \in \{1,\ldots,2^n\}\backslash\Lambda^2_n$ the interval $(t^n_{j-1}, t^n_j] \subset \mathcal{K}$, then $|M_r - M_{t^n_j}| < \varepsilon$, for all $r \in (t^n_{j-1}, t^n_j]$. Thus, by the hypothesis over Z, we have

$$\Big| \sum_{\substack{j=1 \\ j \notin \Lambda^2_n}}^{2^n} \int_{t^n_{j-1}}^{t^n_j} (M_r - M_{t^n_j})\, dZ_r \Big| \leqslant \varepsilon \sum_{\substack{j=1 \\ j \notin \Lambda^2_n}}^{2^n} \int_{t^n_{j-1}}^{t^n_j} |Z'_r|\, dr \leqslant \varepsilon A_t\, (t-s)\,. \tag{1.27}$$

By the hypotheses over M and Z, the second part of the sum equation (1.26), the summation over Λ^2_n, satisfies

$$\Big| \sum_{j \in \Lambda^2_n} \int_{t^n_{j-1}}^{t^n_j} (M_r - M_{t^n_j})\, dZ_r \Big| \leqslant 2\, C_t \sum_{j \in \Lambda^2_n} \int_{t^n_{j-1}}^{t^n_j} |Z'_r|\, dr \leqslant 2\, C_t\, A_t \sum_{j \in \Lambda^2_n} (t^n_j - t^n_{j-1})\,. \tag{1.28}$$

Using the definitions of $\Lambda^1_{n_1}$ and Λ^2_n, we have

$$\sum_{j\in\Lambda^2_n}(t^n_j - t^n_{j-1}) = \sum_{\ell\in\Lambda^1_{n_1}}(t^{n_1}_\ell - t^{n_1}_{\ell-1})\,.$$

The last sum is bounded from above by $m\frac{t-s}{2^{n_1}} < \varepsilon$, because of the definition of the partition, see equation (1.25), the fact that $\Lambda^1_{n_1}$ has at most m elements and the choice of n_1. Putting it in equation (1.28), we get

$$\Big|\sum_{j\in\Lambda^2_n}\int_{t^n_{j-1}}^{t^n_j}(M_r - M_{t^n_j})\,dZ_r\Big| \leqslant 2\,C_t\,A_t\,\varepsilon\,. \tag{1.29}$$

From equations (1.27) and (1.29), we obtain

$$\Big|\sum_{j=1}^{2^n}\int_{t^n_{j-1}}^{t^n_j}(M_r - M_{t^n_j})\,dZ_r\Big| \leqslant \varepsilon\,A_t\,(t-s+2C_t)\,,\quad \text{for all } n\geqslant n_2.$$

Then, we can conclude equation (1.24) and it finishes the proof. □

Now, we have what we need to prove the next theorem. This result is an extension, for cádlág martingales, of the continuous version presented in Theorem 1.2.8 on [46], which should be viewed as the integration by parts formula for martingale theory.

Theorem 1.31. *Let $\{M_t,\, t\geqslant 0\}$ be a càdlàg martingale with respect to the filtration $\{\mathcal{F}_t,\, t\geqslant 0\}$, which is bounded in the following sense for each $t>0$ there exists a constant $C_t>0$ such that $|M_r|\leqslant C_t$, for any $r\in[0,t]$. And, let $\{Z_t,\, t\geqslant 0\}$ be a bounded variation process and adapted to the filtration $\{\mathcal{F}_t,\, t\geqslant 0\}$, with $dZ_r = Z'_r dr$ and $|Z'_r|\leqslant A_t$, for any $r\in[0,t]$, for some $A_t>0$. Then,*

$$M_tZ_t - \int_0^t M_r\,dZ_r$$

is a martingale with respect to the filtration $\{\mathcal{F}_t,\, t\geqslant 0\}$.

Proof. It is enough to prove that

$$\mathbb{E}[M_tZ_t|\mathcal{F}_s] - M_sZ_s = \mathbb{E}\Big[\int_s^t M_r\,dZ_r\,\Big|\mathcal{F}_s\Big]\,.$$

In order to get the equality above, we start by study the integral $\int_s^t M_r\,dZ_r$. Now, use Lemma 1.30, which says that there exists, $\{t_0,\dots,t_N\}$, a partition of the $[s,t]$ such that

$$\sum_{j=1}^{N} M_{t_j}(Z_{t_j} - Z_{t_{j-1}}) \xrightarrow{L_1} \int_s^t M_r\, dZ_r,$$

as $N \to \infty$. Then, we can write

$$\mathbb{E}\Big[\int_s^t M_r\, dZ_r \,\Big|\mathcal{F}_s\Big] = \lim_{N\to\infty} \sum_{j=1}^{N} \mathbb{E}\big[M_{t_j}(Z_{t_j} - Z_{t_{j-1}})\,\big|\mathcal{F}_s\big]\,,$$

where the $\{t_0, \ldots, t_N\}$ is a partition of $[s,t]$. Taking the conditional expectation concerning to the filtration $\mathcal{F}_{t_{j-1}}$, we can write

$$\mathbb{E}\big[M_{t_j}(Z_{t_j} - Z_{t_{j-1}})\,\big|\mathcal{F}_s\big] = \mathbb{E}\big[M_{t_j}Z_{t_j}\,\big|\mathcal{F}_s\big] - \mathbb{E}\Big[Z_{t_{j-1}}\mathbb{E}\big[M_{t_j}\,\big|\,\mathcal{F}_{t_{j-1}}\big]\,\Big|\,\mathcal{F}_s\Big]\,.$$

Using that $\{M_t,\, t \geqslant 0\}$ is a martingale, we obtain the last limit is equal to

$$\lim_{N\to\infty} \sum_{j=1}^{N} \Big(\mathbb{E}[M_{t_j}Z_{t_j}\,|\mathcal{F}_s] - \mathbb{E}[M_{t_{j-1}}Z_{t_{j-1}}\,|\mathcal{F}_s]\Big) = \mathbb{E}[M_tZ_t\,|\mathcal{F}_s] - M_sZ_s\,.$$

$\square$

Finally, we end this appendix with a last result that we need for Theorem 1.10.

Lemma 1.32. *Let $\psi : [0,\infty) \to \mathbb{R}$ be a bounded measurable function. Then*

$$\frac{d}{dt}\Big(e^{\int_0^t \psi(r)\,dr}\Big) = \psi(t)\, e^{\int_0^t \psi(r)\,dr}\,.$$

Proof. By Taylor expansion and the fact that ψ is bounded, we get

$$\tfrac{1}{h}\Big(e^{\int_0^{t+h} \psi(r)\,dr} - e^{\int_0^t \psi(r)\,dr}\Big) = e^{\int_0^t \psi(r)\,dr}\Big(\tfrac{1}{h}\int_t^{t+h} \psi(r)\,dr + O_\psi(h)\Big)\,,$$

for all $h > 0$. The result follows from Lebesgue Differentiation Theorem. $\square$

2. Thermodynamic formalism for jump processes

This chapter is a joint work with Josué Knorst, Artur Lopes and Adriana Neumann. In [25], we deal with thermodynamic formalism for processes whose infinitesimal generator L is in the form of equation (1.2). We consider the state space a compact set that can be the unitary circle S^1 or the interval $[0,1]$. Notice that we can consider S^1 as the interval $[0,1]$ with $0 \equiv 1$, then to fix ideas our state space will be the interval $[0,1]$ and when we want to refer to S^1 we will consider periodic boundary conditions, that is, identifying 0 and 1 ($0 \equiv 1$). The integrals presented in this chapter are assumed to be on the whole state space unless we say otherwise.

We introduce a Ruelle operator from a continuous potential $V : [0,1] \to \mathbb{R}$ and an *a priori* probability $\mathbb{P}$ (induced by the infinitesimal generator L and a initial measure) on the Skorohod space $D = D([0,+\infty),[0,1])$ of *càdlàg* paths $\omega : [0,+\infty) \to [0,1]$. Assuming a Hölder regularity of the potential V, we can prove the existence of an eigenvalue and a positive eigenfunction for the Ruelle operator. After a kind of normalization procedure, we obtain another process, called the Gibbs Markov process, that induces a probability on D, called Gibbs (or equilibrium) probability. From this, we were able to introduce the concepts of relative entropy and pressure. Lately, we define entropy production by discussing some properties related to time-reversal and symmetry of infinitesimal generators.

2.1. The Model

Consider an infinitesimal generator L of a Markov jump process, on the form of equation (1.2), with jump rate function $\lambda \equiv 1$ and a kernel $P(x,dy)$ that can be decomposed as $P(x,y)dy$, where the function $P : [0,1]^2 \to [0,1]$ later will be asked to satisfy equation (2.3). This operator acts on periodic functions $f : [0,1] \to \mathbb{R}$ by

$$(Lf)(x) = \int \big[f(y) - f(x)\big] P(x,y)dy. \tag{2.1}$$

Notice that $L(1) = 0$. We call L the *a priori* infinitesimal generator.

We will denote by L^* the dual of L in $\mathscr{L}^2(dx)$, which acts on functions $g : [0,1] \to \mathbb{R}$ by

$$(L^*g)(x) = \int P(y,x)g(y)dy - g(x). \tag{2.2}$$

Let θ be the invariant vector for P on the left. In the subsection entitled "Markov Chains with values on S^1" of [28, Section 3] it is shown that, under Hölder assumption, there exists a unique θ. Define $\mu(dx) = \theta(x)dx$ the probability measure with density θ. This means that it satisfies

$$\int \theta(y)P(y,x)dy = \theta(x). \tag{2.3}$$

By this, we get $L^*(\theta) = 0$, what means that μ is invariant for the action of L^*.

Notice that L and L^* are bounded operators. Then, by equation (1.4), we can define the semigroup e^{tL}. For fixed $t \geqslant 0$, this semigroup is an integral operator, that is, there exists a kernel function $K_t : [0,1] \times [0,1] \to \mathbb{R}^+$ such that

$$(e^{tL}f)(x) = \int K_t(x,y)f(y)dy + e^{-t}f(x). \tag{2.4}$$

The existence of this function K_t is presented in Section 2.A along with some properties that it satisfies.

For a continuous-time Markov process $\{X_t, t \geqslant 0\}$, the kernel K_t plays the same role that the transition function has on the discrete-time case. Given an initial density function φ_0 and the probability $\mathbb{P}$ induced on D by this process, we can measure a cylinder set $\mathcal{C} = \{X_0 \in (a_0, b_0), X_{t_1} \in (a_1, b_1), X_{t_2} \in (a_2, b_2)\}$ by

$$\mathbb{P}(\mathcal{C}) = \int_{a_0}^{b_0} \int_{a_1}^{b_1} \int_{a_2}^{b_2} \varphi_0(x_0) K_{t_1}(x_0, x_1)\, K_{t_2 - t_1}(x_1, x_2)\, dx_2 dx_1 dx_0.$$

Now, let us see how we can compute this K_t with an example:

Example 2.1. *Take $P(x,y) = \cos[(x-y)2\pi]/2 + 1$. This P is symmetric and continuous on $[0,1]$. Since $\int \cos[(x-y)2\pi]dy = 0$, for any $x \in [0,1]$, we get that $\int P(x,y)dy = 1$. In this case, the kernel function $K_t(x,y), t \geqslant 0$ can be explicitly expressed by*

$$K_t(x,y) = 2\cos[2\pi(x-y)](e^{-3t/4} - e^{-t}) + (1 - e^{-t})$$

and the Lebesgue probability dx is the unique invariant probability.

First, we will calculate the K_t expression. Note that

$$\int \cos(2\pi(x-z)) \cdot \cos(2\pi(z-y))\, dz = \frac{1}{2}\cos(2\pi(x-y)).$$

Using induction, we show that $P^n(x,y) = \frac{\cos(2\pi(x-y))}{2^{2n-1}} + 1$:

$$\begin{aligned} P^{n+1}(x,y) &= \int P^n(x,z)P(z,y)dz \\ &= \int \left(\frac{\cos[2\pi(x-z)]}{2^{2n-1}} + 1 \right) \left(\frac{\cos[2\pi(z-y)]}{2} + 1 \right) dz \\ &= \frac{1}{2^{2n}} \int \cos[2\pi(x-z)] \cdot \cos[2\pi(z-y)]\, dz + 1 \end{aligned}$$

$$= \frac{\cos[2\pi(x-y)]}{2^{2n+1}} + 1.$$

By the general case, see Section 2.A, we know that

$$K_t(x,y) = \sum_{k=1}^{\infty} \frac{t^k}{k!} Q_k(x,y),$$

where

$$\begin{aligned} Q_k(x,y) &:= \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} P^j(x,y) \\ &= \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \left(\frac{\cos[2\pi(x-y)]}{2^{2j-1}} + 1 \right) \\ &= 2\cos[2\pi(x-y)] \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \frac{1}{2^{2j}} + \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \\ &= 2\cos[2\pi(x-y)] \left[(-1)^{k+1} + \left(-\frac{3}{4} \right)^k \right] + (-1)^{k+1}. \end{aligned}$$

This gives us exactly the formula we want for $K_t(x,y)$.

Now, we turn ourselves to the second claim. The fact that dx is invariant is an immediate consequence of symmetry: the function 1 satisfies $L^*(1) = L(1) = 0$. We need to go further to get unicity.

A continuous function $f : [0,1] \to \mathbb{R}$ can be seen as a periodic function $f : \mathbb{R} \to \mathbb{R}$ with period 1 so that we can employ Fourier Series. Write

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n x) + \sum_{n=1}^{\infty} b_n \sin(2\pi n x),$$

with $\frac{a_0}{2} = \int f(x)dx$, $a_n = 2\int f(x)\cos(2\pi n x)dx$ and $b_n = 2\int f(x)\sin(2\pi n x)dx$. Notice that $\cos(2\pi(x-y)) = \cos(2\pi x)\cos(2\pi y) + \sin(2\pi x)\sin(2\pi y)$. Then

$$\begin{aligned} (Lf)(x) &= \int f(y)dy + \frac{1}{2}\int f(y)\cos[2\pi(x-y)]dy - f(x) \\ &= \frac{a_0}{2} + \frac{1}{2}\cos(2\pi x)\frac{a_1}{2} + \frac{1}{2}\sin(2\pi x)\frac{b_1}{2} - f(x). \end{aligned}$$

Therefore, $Lf = 0$ if, and only if,

$$f(y) = \frac{a_0}{2} + \cos(2\pi y)\frac{a_1}{4} + \sin(2\pi y)\frac{b_1}{4}$$

and consequently $a_1 = a_1/4$, $b_1 = b_1/4$ and $a_n = b_n = 0$, $\forall n \geq 2$. We conclude that $L^* f = Lf = 0 \Leftrightarrow f \equiv \frac{a_0}{2}$, constant. This means that the only eigendensity of the operator e^{tL} is that of Lebesgue measure dx. $\diamondsuit$

Consider L as defined on equation (2.1), let us assume that there exists a positive continuous density function $\theta : [0,1] \to \mathbb{R}$, such that, for any continuous function $f : [0,1] \to \mathbb{R}$, we get

$$\int (Lf)(x)\theta(x)dx = 0. \tag{2.5}$$

Moreover, as a consequence of the relation above, valid for any f, it is easy to see that θ is also a solution of equation (2.3), which is unique under the Hölder assumption. Thereat, we can assume that L is such that the above defined θ is unique.

Definition 2.2. *Given L defined on equation* (2.1) *and an initial density θ satisfying equation* (2.5)*, we get a continuous-time stationary Markov process $\{X_t, t \geqslant 0\}$, with values on* $[0,1]$ *(see [4, 8, 26]). This process defines a probability $\mathbb{P}$ on the Skorohod space D. This probability $\mathbb{P}$ is invariant for the shift $\{\Theta_t, t \geqslant 0\}$, which acts on $\omega \in D$ as $(\Theta_t w)_s = w_{s+t}$*

For this infinitesimal generator, the associated semigroup satisfies $e^{t\,L}(1) = 1$. Moreover, $e^{t\,L^*}(\theta) = \theta$, where L^* was given by equation (2.2) and θ satisfies equation (2.5).

Now we take the continuous potential $V : [0,1] \to \mathbb{R}$ and introduce the operator $L + V$ as a particular case, with a function V that does not depend on time, of the one presented in Section 1.3. In the same way, we define $L^* + V$. If $P(x,y)$ is symmetric, the spectral properties of both are the same. Section 1.3 also give us a formula for the homogeneous semigroup

$$\left(e^{t\,(L+V)} f\right)(x) = \mathbb{E}_x \left[e^{\int_0^t V(X_r)dr} f(X_t) \right], \tag{2.6}$$

where $\{X_t,\, t \geqslant 0\}$, is the Markov process with infinitesimal generator L. Notice that this semigroup is not Markovian.

Similarly to e^{tL}, this semigroup is also an integral operator, that is, there exists a kernel function $K_t^V : [0,1] \times [0,1] \to \mathbb{R}^+$ such that

$$\left(e^{t\,(L+V)} f\right)(x) = \int K_t^V(x,y) f(y)dy + e^{-t} e^{tV(x)} f(x). \tag{2.7}$$

The properties of K_t^V are the ones presented in Section 2.B if we consider $\lambda \equiv 1$.

2.2. Ruelle Operator and the Gibbs Markov process

In this section, we will introduce the Ruelle operator (which was considered in similar cases in [6, 32]) and use a kind of normalization procedure to get the

Gibbs Markov process and its induced Gibbs probability on D.

Definition 2.3 (Ruelle Operator)**.** *Consider, for a fixed $t \geqslant 0$, the continuous-time Ruelle operator $\mathbb{L}_V^t$, associated with V, that acts on continuous functions $\varphi : [0,1] \to \mathbb{R}$ as*

$$(\mathbb{L}_V^t \varphi)(x) = \mathbb{E}_x \left[e^{\int_0^t V(X_r)dr} \varphi(X_t) \right] = \left(e^{t(L+V)} \varphi \right)(x).$$

For a symmetrical L, as the Feynman-Kac formulas for natural and reverse time processes coincide, this Ruelle operator can be seen as the continuous-time version of the classical Ruelle operator (discrete case). Figure 2.1 depicts this statement. The left approach is more suitable for the Feymann-Kac formula while the right one can be easily related to the classical discrete-time Ruelle operator for the n-coordinate shift $\sigma_n : [0,1]^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$, with y being the initial value of the shifted path.

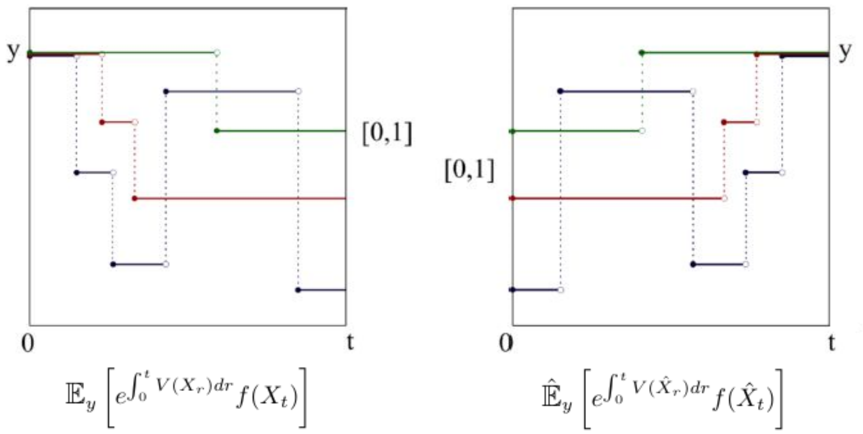


Figure 2.1: *If L is symmetric, we can use the Ruelle Operator at natural or reversal time.*

According to our notation, this continuous-time Ruelle operators $\mathbb{L}_V^t$, $t \geqslant 0$, are a family of linear operators indexed by t.

Definition 2.4. *Fix $V : [0,1] \to \mathbb{R}$. We say that the family of Ruelle operators $\mathbb{L}_V^t, t \geqslant 0$, is normalized if $\mathbb{L}_V^t 1 = 1$, for all $t \geqslant 0$.*

If the potential $V \equiv 0$, for any $t \geqslant 0$, the Ruelle operator is $\mathbb{L}_0^t = e^{tL}$. In this case, the family of Ruelle operators is normalized. From now to the end of this section, we will study non-normalized Ruelle operators in order to associate them with a normalized Gibbs operator.

Definition 2.5. *We say that $f : [0,1] \to \mathbb{R}$ is an eigenfunction on the right of the Ruelle operator $\mathbb{L}_V^t$, $t \geqslant 0$, associated with the eigenvalue $\lambda \in \mathbb{R}$, if for all $t \geqslant 0$,*

$$\mathbb{L}_V^t f = e^{\lambda\, t} f.$$

Similarly, we say that h is an eigenfunction on the left if

$$h\, \mathbb{L}_V^t = e^{\lambda\, t} h.$$

In order to find these eigenfunctions, we have to analyze the properties of the operator $L + V$ and $L^* + V$.

Assume that positive functions f, h are such that

$$(L + V)(f) = \lambda f \quad \text{and} \quad (L^* + V)(h) = \lambda h, \tag{2.8}$$

that is, $f, h : [0,1] \to \mathbb{R}^+$ are eigenfunctions of $L + V$ and $L^* + V$ associated with the same eigenvalue $\lambda \in \mathbb{R}$. Then, $e^{t(L+V)} f = e^{\lambda t} f$, what makes f an eigenfunction for the Ruelle operator associated. In addition, $e^{t(L^*+V)} h = e^{\lambda t} h$. We say that such λ (which can be positive or negative) is the main eigenvalue.

Notice that, by linearity, we have a whole class of functions that satisfies equation (2.8). It is natural to assume the normalization condition $\int h(x)dx = 1$, so we can see h as a density. Let us take the specific f that satisfies $\int f(x)h(x)dx = 1$. In this case, $\pi(x) = f(x)h(x)$ is a density on $[0,1]$.

In the following, we will assume that exists a solution for equation (2.8). Comparing with [45, pages 106 and 111], we can see that, in the discrete case, we simply get that via Perron-Frobenius theory.

Assumption 2.1. *Assume that there exists an eigenvalue $\lambda \in \mathbb{R}$ and two functions $\ell : [0,1] \to \mathbb{R}^+$ and $r : [0,1] \to \mathbb{R}^+$ of Hölder class, such that,*

$$(L + V)r = \lambda r \quad \text{and} \quad \ell\,(L + V) = \lambda \ell.$$

There are plenty of examples of pairs L and V for such ones the above condition is satisfied. To exemplify that, we will use a continuous function $g : [0,1] \to \mathbb{R}^+$, satisfying $\int g(x)dx = 1$, to define $P(x,y)$, for $x, y \in [0,1]$, by

$$P(x,y) = \begin{cases} g(x+y), & \text{if } (x+y) < 1, \\ g(x+y-1), & \text{if } (x+y) \geqslant 1. \end{cases}$$

This P is a symmetric kernel and the corresponding invariant density θ, satisfying equation (2.3), is equal to 1. Therefore, the g we choose defines L via P. Notice that $L^* = L$, then we just need to find f such that $(L + V)(f) = \lambda f$, because, in this case, we have $(L^* + V)h = \lambda h$ for $h = f$.

Example 2.6. *Consider g as a restriction of a polynomial of degree 2 to $[0,1]$, with $g(0) = g(1) \geqslant 0$. Assume L is defined via P using this polynomial g, as we mentioned above. Defining, for any $b \in \mathbb{R}$, a polynomial $V(x) = b + [1 - g(0)]x(1-x)$, there exists $\lambda \in \mathbb{R}$ and f a polynomial of degree 2, with $f(0) = f(1)$, satisfying*

$$(L+V)(f)(x) = \int [f(y) - f(x)]P(x,y)dy + V(x)f(x) = \lambda f(x). \tag{2.9}$$

Moreover, there is a solution f that is positive on $[0,1]$.

Write $g(x) = a_0 + a_1 x + a_2 x^2$, for some $a_0, a_1, a_2 \in \mathbb{R}$ with $a_0 > 0$ and $a_2 \neq 0$. The restriction $g(0) = g(1)$ imply that $a_2 = -a_1$, while the integral condition give us that $a_1 = 6(1 - a_0)$. Considering both, we have $g(x) = a_0 + 6(1-a_0)x(1-x)$ with $a_0 > 0$ and $a_0 \neq 1$.

In the same way, the polynomial f is of the form $f(x) = c_0 + c_1 x(1-x)$, for some $c_0, c_1 \in \mathbb{R}$. Using the definition of P, we can compute the integral term of equation (2.9) as

$$\begin{aligned} p(x) &:= \int_0^{1-x} [f(y) - f(x)]g(x+y)dy + \int_{1-x}^1 [f(y) - f(x)]g(x+y-1)dy \\ &= \frac{1}{30}(6-a_0)c_1 - c_1 x + a_0 c_1 x^2 + 2(1-a_0)c_1 x^3 - (1-a_0)c_1 x^4. \end{aligned}$$

Considering $g(0) = a_0$ on the definition of V, the expression $p(x) + V(x)f(x)$ of the left side of equation (2.9) turns to be

$$\left[bc_0 + \frac{c_1}{5} - \frac{a_0 c_1}{30}\right] + \left[(1-a_0)c_0 - (1-b)c_1\right]x + \left[(-1+a_0)c_0 + (1-b)c_1\right]x^2.$$

This expression needs to be equal to $\lambda f(x)$, also a polynomial of degree 2, for some $\lambda \in \mathbb{R}$. This means that both polynomials should have the same coefficients, which gives us three equations:

$$\begin{cases} bc_0 + \frac{c_1}{5} - \frac{a_0 c_1}{30} = \lambda c_0, \\ (1-a_0)c_0 - (1-b)c_1 = \lambda c_1, \\ (-1+a_0)c_0 + (1-b)c_1 = -\lambda c_1. \end{cases}$$

First of all, notice that the last two equations give us the same condition. As $a_0 \neq 1$, they give us that $c_0 = \left(\frac{b-1-\lambda}{a_0-1}\right)c_1$. Substituting on the first one, we get

$$c_1\left(\left[\frac{30b^2 - 30b - 6 + 7a_0 - a_0^2}{30(a_0-1)}\right] - \left[\frac{2b-1}{a_0-1}\right]\lambda + \left[\frac{1}{a_0-1}\right]\lambda^2\right) = 0.$$

The general solution of this is $\lambda = \frac{1}{30}\left(-15 \pm \sqrt{405 - 210a_0 + 30a_0^2 + 30b}\right)$, what means

$$f(x) = c_1 \left(\frac{b - 1 - \lambda}{a_0 - 1} + x(1 - x) \right). \tag{2.10}$$

The general eigenfunction f presented above can, sometimes, assume negative values on the interval $[0, 1]$. The Assumption 2.1 asks positivity, since we need that in our reasoning (see Equation (2.14)), but this is not a problem in this case because we have positivity for $0 < a_0 < 1$ if we take $c_1 > 0$ and for $a_0 \geqslant 1$ if $c_1 < 0$. $\diamondsuit$

Remark. *The same may not be true in other cases. For instance, considering* g *and* V *as polynomials of degree 3, we were able to get a positive polynomial* $f : [0, 1] \to \mathbb{R}^+$ *of degree 3, but the next example shows that its impossible to get a polynomial solution on* S^1*, considering the periodic boundary condition* $0 \equiv 1$*. In the search for suitables* V *and* f*, we are not able to get solutions satisfying the property that* g *is strictly positive. This is not very intuitive, since for larger degrees we have more free variables to work. In fact, in the degree* n *case, we have* $3n + 4$ *coefficients (*$n + 1$ *from each polynomial and one from* λ*) to cancel* $2n + 1$ *coefficients of* $p(x) + (V(x) - \lambda)f(x)$*, the constraints, which means there are left three coefficients for periodicity, one for* $\int g(x)dx = 1$ *and at least* $n - 1$ *to adjust positivity.*

Example 2.7. Consider periodic polynomials $f, g, V : [0, 1] \to \mathbb{R}$ of degree 3 as

$$g(x) = a_0 + a_1 x - 3(-4 + 4a_0 + a_1)x^2 + (-a_1 + 3(-4 + 4a_0 + a_1))x^3,$$

$$V(x) = b_0 + b_1 x + b_2 x^2 + (-b_1 - b_2)x^3,$$

$$f(x) = c_0 + c_1 x + c_2 x^2 + (-c_1 - c_2)x^3$$

and define a polynomial of degree 6 by

$$K(x) := \int_0^1 P(x, y)f(y)dy + V(x)f(x) - (1 - \lambda)f(x).$$

Suppose that $c_1, c_2 \in \mathbb{R}$ are such that:

1) $c_1 \neq 0$;

2) $c_1 + c_2 \neq 0$;

3) $2c_1 + c_2 \neq 0$;

4) $3c_1 + c_2 \neq 0$;

5) $54c_1^2 + 39c_1c_2 + 8c_2^2 \neq 0$;

6) $91368c_1^6+186948c_1^5c_2+159318c_1^4c_2^2+71367c_1^3c_2^3+17228c_1^2c_2^4+1984c_1c_2^5+64c_2^6 \neq 0$.

In this case, solving $K(x) = 0$, we find expressions[1] for $b_2, b_1, a_1, c_0, \lambda, a_0$ as functions of b_0, c_1, c_2. Thus, we get a whole class of polynomials f, g, V that solves $K(x) = 0$, for all x, but none of these solutions satisfies that g is strictly positive on $[0, 1]$. In order to see that, let us analyze the properties of a generic polynomial $G : \mathbb{R} \to \mathbb{R}$ of degree 3 satisfying $G(0) = G(1)$ and $\int_0^1 G(x)dx = 1$.

If we write G as

$$G(x) = \frac{1}{12}(12 + 2g_2 + 3g_3) + (-g_2 - g_3)x + g_2x^2 + g_3x^3,$$

the expressions for the two critical points of G are

$$X_1 = \frac{g_2 - \sqrt{g_2^2 + 3g_2g_3 + 3g_3^2}}{3g_3} \quad \text{and} \quad X_2 = \frac{g_2 + \sqrt{g_2^2 + 3g_2g_3 + 3g_3^2}}{3g_3}.$$

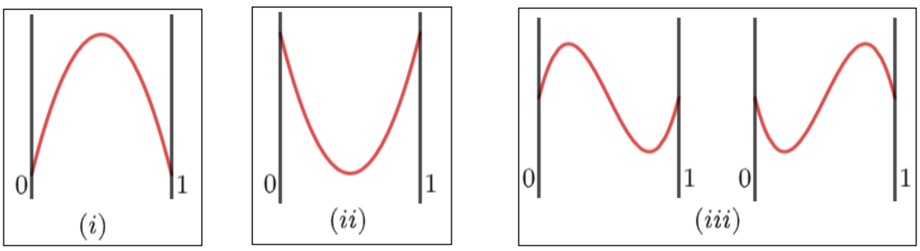
Looking closely at these critical points, we see that X_1 is a local maximum and X_2 is a local minimum for G. Furthermore, since $G(0) = G(1)$, there is at least one critical point of G on $[0, 1]$. Then, we can divide the analysis into three cases:

(i) $X_1 \in [0, 1]$ and $X_2 \notin [0, 1]$;

(ii) $X_2 \in [0, 1]$ and $X_1 \notin [0, 1]$;

(iii) $X_1, X_2 \in [0, 1]$.

Visually, we can see these cases as



Notice that these conditions are unically defined by the sign of the derivatives of G on $x = 0$ and $x = 1$. Moreover, the positivity is given by the absolute minimum on $[0, 1]$, which is $G(0) = G(1)$ on *(i)* and $G(X_2)$ on *(ii)* and *(iii)*.

Finally, we consider g_2 and g_3 as a_2 and a_3, the respective coefficients of g that solves $K(0) = 0$. Using the free variables $c_1, c_2 \in \mathbb{R}$, in all three cases, we

[1] We are not showing the exact expressions here due to its complexity, but it is possible to get them using the *Mathematica* software by nullifying each coefficient of K from the highest exponent to the smallest one.

can reduce $g(x) \geqslant 0$, for all $x \in [0,1]$, into

$$\begin{cases} c_2 \neq 0, & \text{if } c_1 = 0, \\ c_2 = -2c_1, & \text{if } c_1 \neq 0. \end{cases}$$

In both cases, there is no solution, since we have a clash with conditions 1 and 3 that we initially set for c_1, c_2. $\diamondsuit$

Observe that a function f obtained from Assumption 2.1 defines a whole set of functions $\{\alpha f, \alpha \in \mathbb{R}^+\}$ that also satisfies the same condition. In Example 2.6, this is very clear when we look to equation (2.10). One can use this subspace to get specific functions satisfying some conditions. For instance, for a fixed V, we take ℓ_V, r_V and λ_V the ones from Assumption 2.1 that satisfies the normalization conditions

$$\int \ell_V(x)dx = 1 \quad \text{and} \quad \int r_V(x)\ell_V(x)dx = 1. \tag{2.11}$$

An equation for the right eigenfunction r_V is

$$\int P(x,z)r_V(z)dz - (1 + \lambda_V - V(x))r_V(x) = 0, \tag{2.12}$$

for any x. On the other hand, the left eigenfunction ℓ_V satisfies

$$\int \ell_V(z)P(z,x)dz - (1 + \lambda_V - V(x))\ell_V(x) = 0. \tag{2.13}$$

For all $x, y \in [0,1]$, $t \geqslant 0$ and $f \in C_b([0,1])$, define

$$\gamma_V(x) = 1 + \lambda_V - V(x), \qquad Q_V(x,y) = \frac{P(x,y)r_V(y)}{r_V(x)\gamma_V(x)}, \tag{2.14}$$

$$(\mathcal{L}_V f)(x) = \gamma_V(x) \int [f(y) - f(x)] Q_V(x,y)\, dy$$

and

$$(\mathcal{P}_t^V f)(x) = \frac{e^{t(L+V)}(r_V f)(x)}{e^{\lambda_V t} r_V(x)}.$$

Remark. *One can also write*

$$(\mathcal{P}_t^V f)(x) = \frac{1}{e^{\lambda_V t} r_V(x)} \mathbb{E}_x \left[e^{\int_0^t V(X_s)\, ds} r_V(X_t) f(X_t) \right].$$

Lemma 2.8. *The operator $\mathcal{P}_t^V$ is the semigroup associated with the infinitesimal generator $\mathcal{L}_V$, that is,*

$$\lim_{t \to 0} \frac{(\mathcal{P}_t^V f)(x) - f(x)}{t} = (\mathcal{L}_V f)(x).$$

Proof. We can rewrite

$$\begin{aligned}\frac{(\mathcal{P}_t^V f)(x) - f(x)}{t} &= \frac{1}{e^{\lambda_V t} r_V(x)} \left(\frac{e^{t(L+V)}(r_V f)(x) - r_V(x) f(x)}{t} \right) \\ &\quad + f(x) \left(\frac{e^{-\lambda_V t} - 1}{t} \right).\end{aligned}$$

Taking limit as $t \to 0$, we get

$$\begin{aligned}& \frac{1}{r_V(x)} (L+V)(r_V f)(x) - \lambda_V f(x) \\ = \; & \frac{1}{r_V(x)} \left[\int P(x,y) r_V(y) f(y) dy + (V(x) - 1) r_V(x) f(x) \right] - \lambda_V f(x) \\ = \; & \frac{1}{r_V(x)} \int P(x,y) r_V(y) f(y) dy - \gamma_V(x) f(x) \\ = \; & \gamma_V(x) \int \left(\frac{P(x,y) r_V(y)}{r_V(x) \gamma_V(x)} \right) f(y) dy - \gamma_V(x) f(x).\end{aligned}$$

Using equation (2.12), we have

$$\int Q_V(x,y) dy = \int \frac{P(x,y) r_V(y)}{r_V(x) \gamma_V(x)} dy = 1. \tag{2.15}$$

Then,

$$\lim_{t \to 0} \frac{(\mathcal{P}_t^V f)(x) - f(x)}{t} = \gamma_V(x) \int [f(y) - f(x)] Q_V(x,y) dy = (\mathcal{L}_V f)(x).$$

□

Notice that the semigroup $\mathcal{P}_t^V$, $t \geqslant 0$, is normalized. Furthermore, from equation (2.15), it defines a jump process (with the generator in the form of equation (1.2), where we replace P by Q_V and consider the jump rate function γ_V).

Definition 2.9 (Gibbs Markov process)**.** *We call Gibbs Markov process associated with the potential V (and the a priori infinitesimal generator L) the continuous-time Markov jump process generated by $\mathcal{L}_V$.*

Now, we want to prove that the invariant density for the Gibbs process is $\pi_V = \ell_V r_V$, where the normalization conditions given by equation (2.11) are assumed to be satisfied. To do this, we need to use the dual operator $\mathcal{L}_V^*$.

Lemma 2.10. *The dual of the operator $\mathcal{L}_V$ is the operator*

$$(\mathcal{L}_V^* g)(x) = \int \gamma_V(y) g(y) Q_V(y,x) dy - \gamma_V(x) g(x). \tag{2.16}$$

Proof. Given the functions f, g we get

$$\begin{aligned}
& \int (\mathcal{L}_V f)(x) g(x) dx \\
= & \int \gamma_V(x) g(x) \int [f(y) - f(x)] Q_V(x,y) dy dx \\
= & \int \int \gamma_V(x) g(x) Q_V(x,y) f(y) dy dx - \int \gamma_V(x) f(x) g(x) \left[\int Q_V(x,y) dy \right] dx \\
= & \int f(y) \int \gamma_V(x) g(x) Q_V(x,y) dx dy - \int \gamma_V(x) f(x) g(x) dx \\
= & \int f(z) \left[\int \gamma_V(x) g(x) Q_V(x,z) dx - \gamma_V(z) g(z) \right] dz = \int f(z) (\mathcal{L}_V^* g)(z) dz.
\end{aligned}$$

□

Next, we show that π_V is the stationary density.

Proposition 2.11. *The density π_V satisfies $\mathcal{L}_V^*(\pi_V) = 0$.*

Proof. From equations (2.13) and (2.14) we get that, for any point x,

$$\begin{aligned}
(\mathcal{L}_V^* \pi_V)(x) &= \int \gamma_V(y) \ell_V(y) r_V(y) Q_V(y,x) dy - \gamma_V(x) \ell_V(x) r_V(x) \\
&= \int \gamma_V(y) \ell_V(y) r_V(y) \left(\frac{P(y,x) r_V(x)}{r_V(y) \gamma_V(y)} \right) dy - \gamma_V(x) \ell_V(x) r_V(x) \\
&= r_V(x) \int \ell_V(y) P(y,x) dy - \gamma_V(x) \ell_V(x) r_V(x) \\
&= r_V(x) \left[\int \ell_V(y) P(y,x) dy - \gamma_V(x) \ell_V(x) \right] = 0.
\end{aligned}$$

□

From the above, we get that, for any x,

$$\int \gamma_V(y) \pi(y) Q_V(y,x) dy = \gamma_V(x) \pi_V(x).$$

Remark. *We have*

$$(\mathcal{P}_t^V 1) = e^{t\,\mathcal{L}_V}(1) = 1$$

and

$$(\mathcal{P}_t^V)^*(\pi_V) = (e^{t\,\mathcal{L}_V^*} \pi_V) = \pi_V.$$

Definition 2.12 (Gibbs probability)**.** *The probability $\mathbb{P}^V$ induced on D by the Gibbs Markov process (with infinitesimal generator $\mathcal{L}_V$ and stationary probability π_V) will be called the Gibbs probability for the potential V (and the a priori infinitesimal generator L). This $\mathbb{P}^V$ is invariant for the shift $\{\Theta_s, s \geqslant 0\}$.*

In the case $V \equiv 0$, $\mathbb{P}^V$ is the a priori probability $\mathbb{P}$ of Definition 2.2.

2.3. Relative Entropy, Pressure and the equilibrium state for V

In this section, we will consider a variational problem in the continuous-time setting which is analogous to the pressure problem in the discrete-time setting. This requires a meaning for entropy, so we will define the relative entropy. A continuous-time stationary Markov process that maximizes our variational problem is called *continuous-time equilibrium state for V*. The results of this section in some sense are similar to the ones in [32].

Consider the infinitesimal generator $\tilde{\mathcal{L}}$, which acts on bounded measurable functions $f : [0,1] \to \mathbb{R}$ as

$$(\tilde{\mathcal{L}}f)(x) = \int \big[f(y) - f(x)\big] \tfrac{\varphi(y)}{\varphi(x)} P(x,y)dy,$$

where $\varphi \in C_b([0,1])$. To rewrite the operator above on the form of equation (1.2) we consider

$$\tilde{\gamma}(x) := \frac{1}{\varphi(x)} \int \varphi(y)P(x,y)dy \quad \text{and} \quad \tilde{Q}(x,y) := \frac{\varphi(y)}{\varphi(x)\tilde{\gamma}(x)} P(x,y).$$

Then

$$(\tilde{\mathcal{L}}f)(x) = \tilde{\gamma}(x) \int \big[f(y) - f(x)\big] \tilde{Q}(x,y)\, dy\,.$$

Proposition 2.13. *The invariant probability for $\tilde{\mathcal{L}}$ is*

$$\tilde{\mu}(dx) = \frac{\varphi(x)\tilde{\ell}_\varphi(x)}{\|\varphi\|\|\tilde{\ell}_\varphi\|} dy,$$

where $\tilde{\ell}_\varphi$ satisfies

$$\frac{1}{\tilde{\ell}_\varphi(x)} \int \tilde{\ell}_\varphi(y)P(y,x)dy = \tilde{\gamma}(x).$$

Proof. Repeating the computation we did on the proof of Lemma 2.10, we can show that, for any density g,

$$\begin{aligned}(\tilde{\mathcal{L}}^* g)(x) &= \int \tilde{\gamma}(y)g(y)\tilde{Q}(y,x)dy - \tilde{\gamma}(x)g(x) \\ &= \int g(y)\frac{\varphi(x)}{\varphi(y)} P(y,x)dy - \tilde{\gamma}(x)g(x).\end{aligned}$$

In particular, for $g = \frac{\varphi\tilde{\ell}_\varphi}{\|\varphi\|\|\tilde{\ell}_\varphi\|}$, we have

$$(\tilde{\mathcal{L}}^* g)(x) = \frac{\varphi(x)}{\|\varphi\|\|\tilde{\ell}_\varphi\|} \left(\int \tilde{\ell}_\varphi(y)P(y,x)dy - \tilde{\gamma}(x)\tilde{\ell}_\varphi(x) \right) = 0.$$

$\square$

Definition 2.14. *The probability $\tilde{\mathbb{P}}_{\tilde{\mu}}$ on D is called admissible if it is induced by the continuous-time Markov chain with infinitesimal generator $\tilde{\mathcal{L}}$ and initial measure $\tilde{\mu}$.*

For $\tilde{\mathbb{P}}_{\tilde{\mu}}$ admissible and $\mathbb{P}_{\tilde{\mu}}$ the probability induced by the original continuous-time Markov chain with infinitesimal generator L, defined in equation (2.1), and initial probability $\tilde{\mu}$, define for $T > 0$,

$$H_T(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) = -\int_D \log\left(\frac{d\tilde{\mathbb{P}}_{\tilde{\mu}}}{d\mathbb{P}_{\tilde{\mu}}}\bigg|_{\mathcal{F}_T}\right)(\omega)d\tilde{\mathbb{P}}_{\tilde{\mu}}(\omega).$$

Notice that we are using the same initial measure $\tilde{\mu}$ for both processes, so the probabilities are absolutely continuous with respect to each other.

Using this H_T above defined, we introduce a meaning for the relative entropy similar to the one presented on [32].

Definition 2.15 (Relative entropy)**.** *For a fixed initial probability $\tilde{\mu}$, the limit*

$$H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) = \lim_{T\to\infty}\frac{1}{T}H_T(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}})$$

is called the relative entropy of $\tilde{\mathbb{P}}_{\tilde{\mu}}$ concerning $\mathbb{P}_{\tilde{\mu}}$.

Since L and $\tilde{\mathcal{L}}$ are both in the form of equation (1.2), they generate two Markov jump process and Corollary 1.17 implies that

$$\begin{aligned}
\log\left(\frac{d\tilde{\mathbb{P}}_{\tilde{\mu}}}{d\mathbb{P}_{\tilde{\mu}}}\bigg|_{\mathcal{F}_T}\right)(\omega) &= \int_0^T [1-\tilde{\gamma}(\omega_s)]ds + \sum_{s\leqslant T}\log\left(\tilde{\gamma}(\omega_{s-})\frac{\varphi(\omega_s)}{\varphi(\omega_{s-})\tilde{\gamma}(\omega_{s-})}\right)\\
&= \int_0^T [1-\tilde{\gamma}(\omega_s)]ds + \sum_{s\leqslant T}\{\log(\varphi(\omega_s)) - \log(\varphi(\omega_{s-}))\}\\
&= \int_0^T [1-\tilde{\gamma}(\omega_s)]ds + \log(\varphi(\omega_T)) - \log(\varphi(\omega_0)).
\end{aligned}$$

Then,

$$H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) = \int [\tilde{\gamma}(x)-1]d\tilde{\mu}(x). \tag{2.17}$$

For a Hölder class potential V, the probability $\mathbb{P}^V_{\pi_V}$ is admissible. Then,

$$H(\mathbb{P}^V_{\pi_V}|\mathbb{P}_{\pi_V}) = \int [\gamma_V(x)-1]d\pi_V(x) = \lambda_V - \int V(x)d\mu_V(x). \tag{2.18}$$

Definition 2.16 (Pressure)**.** *We denote the Pressure (or Free Energy) of V as the value*

$$\boldsymbol{P}(V) := \sup_{\substack{\tilde{\mathbb{P}}_{\tilde{\mu}}\\ admissible}}\left\{H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) + \int V(x)d\tilde{\mu}(x)\right\}.$$

Using equation (2.17), the pressure can be written as

$$\mathbf{P}(V) = \sup_{\substack{\tilde{\mathbb{P}}_{\tilde{\mu}} \\ \text{admissible}}} \int [\tilde{\gamma}(x) - 1 + V(x)]\, d\tilde{\mu}(x).$$

Recalling the expressions of $\tilde{\gamma}$ and $\tilde{\mu}$, we have

$$\mathbf{P}(V) = \sup_{\varphi>0} \int \left(\frac{\tilde{\ell}_\varphi}{\|\tilde{\ell}_\varphi\|} \right)(x)(L+V)\left(\frac{\varphi}{\|\varphi\|} \right)(x)dx \ = \lambda_V.$$

By equation (2.18), this means that the Gibbs probability is the one that maximizes the pressure. In some sense, similar results are true for other settings, see [5, 13, 23, 29, 32].

2.4. Time-reversal process and entropy production

In this section, we consider that the time parameter is bounded, $t \in [0, T]$ for a fixed $T > 0$, in order to explore the time-reversal process. We will show that this time-reversal process is the jump process generated by the dual operator of L in $\mathscr{L}^2(\mu)$, where μ is the invariant measure for L^* defined on Section 2.1. Later, we study the properties of the entropy production rate, that can be used to describe the amount of work dissipated by a irreversible system. Related results can be found in [20, 27, 34, 35, 38].

Remember that the invariant measure satisfies $\mu(dx) = \theta(x)dx$ and that $L^*(\theta) = 0$, where L^* acts on $\mathscr{L}^2(dx)$. The substantial change from $\mathscr{L}^2(dx)$ to $\mathscr{L}^2(\mu)$ is that our reference measure, which was simply Lebesgue measure dx, becomes now $\theta(x)dx$. Taking that into account, the inner product in this new space is given by

$$< f, g >_\mu = \int f(x)g(x)\mu(dx) = \int f(x)g(x)\theta(x)dx.$$

Proposition 2.17. *The dual operator of L over $\mathscr{L}^2(\mu)$ is*

$$(\mathfrak{L}^* g)(x) = \int [g(y) - g(x)] \frac{\theta(y)}{\theta(x)} P(y, x)dy.$$

Proof. To verify this, just compute

$$\begin{aligned} \langle Lf, g\rangle_\mu &= \int (Lf)(x)g(x)\theta(x)dx \\ &= \int\int g(x)\theta(x)P(x,y)f(y)dydx - \int f(x)g(x)\theta(x)dx \\ &= \int f(y) \int g(x)\theta(x)P(x,y)dxdy - \int f(x)g(x)\theta(x)dx \end{aligned}$$

$$
\begin{aligned}
&= \int f(z)\left(\int g(x)\frac{\theta(x)}{\theta(z)}P(x,z)dx\right)\theta(z)dz - \int f(z)g(z)\theta(z)dz \\
&= \int f(z)\left(\int [g(x)-g(z)]\frac{\theta(x)}{\theta(z)}P(x,z)dx\right)\theta(z)dz \\
&= \int f(z)(\mathfrak{L}^* g)(z)\theta(z)dz = \langle f, \mathfrak{L}^* g\rangle_\mu .
\end{aligned}
$$

In this computation, we use that $\int \frac{\theta(x)}{\theta(z)}P(x,z)dy = 1$, what follows directly from equation (2.3). □

Having discussed that, we turn now to defining the time-reversal process, associated with the stationary Markov process (X_t, μ) and an interval of time $[0,T]$. The new process, denoted by $(\hat{X}_t)$, satisfies

$$\mathbb{E}_\mu[\, g(\hat{X}_0)f(\hat{X}_t)] := \mathbb{E}_\mu[\, g(X_T)f(X_{T-t})].$$

Proposition 2.18. *The time-reversal process $\hat{X}_t$ has transition family equal to $P_t^* = e^{t\mathfrak{L}^*}$, the dual operator of P_t over $\mathscr{L}^2(\mu)$.*

Proof. Let $\hat{P}_t$ denote the transition family of $\hat{X}_t$. Using the Markov property and stationarity of the chain X_t, notice that this transition family satisfies, for all $f, g \in \mathscr{L}^2(\mu)$,

$$
\begin{aligned}
\langle \hat{P}_t f, g\rangle_\mu &= \int (\hat{P}_t f)(x) g(x)\ d\mu(x) = \mathbb{E}_\mu[f(\hat{X}_t)g(\hat{X}_0)] \\
&= \mathbb{E}_\mu[f(X_{T-t})g(X_T)] = \mathbb{E}_\mu\left[f(X_{T-t})\mathbb{E}_\mu[g(X_T)|\mathcal{F}_{T-t}]\right] \\
&= \mathbb{E}_\mu[f(X_0)\mathbb{E}_{X_0}[g(X_t)]] = \int f(x)(P_t g)(x)d\mu(x) \\
&= \langle f, P_t g\rangle_\mu .
\end{aligned}
$$

Since this is true for all $f, g \in \mathcal{L}^2(\mu)$, we get that $\hat{P}_t = P_t^*$. This also means that $\hat{L} = \mathfrak{L}^*$, where $\hat{L}$ is the infinitesimal generator of the semigroup $\hat{P}_t$. □

For a fixed $T > 0$, we are interested in the relative entropy of $\hat{\mathbb{P}}_\mu$ concerning $\mathbb{P}_\mu$, where $\hat{\mathbb{P}}_\mu$ is the probability induced on D by the time-reversal process with initial measure μ. Notice that, by definition,

$$H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) = -\int_D \log\left(\left.\frac{d\mathbb{P}_\mu}{d\hat{\mathbb{P}}_\mu}\right|_{\mathcal{F}_T}\right)(\omega)d\mathbb{P}_\mu(\omega).$$

Since, for the processes we are considering, we have $\lambda(x) = \hat{\lambda}(x) = 1$ and

$\hat{P}(x,dy) = \frac{\theta(y)}{\theta(x)} P(y,x)dy$, Corollary 1.17 implies that

$$\log\left(\left.\frac{d\mathbb{P}_\mu}{d\hat{\mathbb{P}}_\mu}\right|_{\mathcal{F}_T}\right) = \sum_{s\leqslant T}\log\left(\frac{P(X_{s-},X_s)\theta(X_{s-})}{P(X_s,X_{s-})\theta(X_s)}\right)$$

and, consequently,

$$\begin{aligned} -H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) &= \mathbb{E}_\mu\left[\sum_{s\leqslant T}\left\{\log\left(\frac{P(X_{s-},X_s)}{P(X_s,X_{s-})}\right) + \log(\theta(X_{s^-})) - \log(\theta(X_s))\right\}\right] \\ &= \mathbb{E}_\mu\left[\sum_{s\leqslant T}\log\left(\frac{P(X_{s-},X_s)}{P(X_s,X_{s-})}\right)\right], \end{aligned}$$

because, for μ invariant, the telescopic summation

$$\mathbb{E}_\mu\left[\sum_{s\leqslant T}\{\log(\theta(X_{s^-})) - \log(\theta(X_s))\}\right] = \mathbb{E}_\mu\left[\log(\theta(X_0)) - \log(\theta(X_T))\right] = 0.$$

In order to analyze the remaining term of this expression, we use the structure of the Markov process. Denoting by $0 = T_0 < T_1 < \cdots$ the jump times of this process and by ξ_n the value of the process on the interval $[T_{n-1}, T_n)$, we have

$$\begin{aligned} -H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) &= \sum_{n=1}^{\infty}\mathbb{E}_\mu\left[\sum_{s\leqslant T}\log\left(\frac{P(X_{s-},X_s)}{P(X_s,X_{s-})}\right)\mathbb{1}_{[T_n\leqslant T<T_{n+1}]}\right] \\ &= \sum_{n=1}^{\infty}\mathbb{E}_\mu\left[\sum_{k=0}^{n-1}\log\left(\frac{P(\xi_k,\xi_{k+1})}{P(\xi_{k+1},\xi_k)}\right)\mathbb{1}_{[T_n\leqslant T<T_{n+1}]}\right] \end{aligned}$$

For simplicity, denote $\psi(x,y) := \log\left(\frac{P(x,y)}{P(y,x)}\right)$. Then,

$$\begin{aligned} -H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) &= \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\mathbb{E}_\mu\left[\psi(\xi_k,\xi_{k+1})\mathbb{1}_{[T_n\leqslant T<T_{n+1}]}\right] \\ &= \sum_{n=1}^{\infty}\left(\mathbb{E}_\mu[\mathbb{1}_{[T_n\leqslant T<T_{n+1}]}]\sum_{k=0}^{n-1}\mathbb{E}_\mu[\psi(\xi_k,\xi_{k+1})]\right). \end{aligned}$$

In this computation, we use that the time variables T_n (defined as the sum of n independent exponential variables τ_k with parameter 1) are independent of the spatial variables ξ_k. It is important to notice that this is not always true. Actually, in the general case, see [24], each τ_k is distributed according to an exponential law of parameter $\lambda(\xi_k)$.

Now, we will analyze separately each expected value on the last expression.

The first one is

$$\begin{aligned}
\mathbb{E}_\mu[\mathbb{1}_{[T_n \leqslant T < T_{n+1}]}] &= \int_0^\infty ds_0 e^{-s_0} \cdots \int_0^\infty ds_n e^{-s_n} \left(\mathbb{1}_{[0 \leqslant T - \sum_{i=0}^{n-1} s_i < s_n]} \right) \\
&= e^{-T} \int_0^\infty \cdots \int_0^\infty ds_0 \dots ds_{n-1} \left(\mathbb{1}_{[\sum_{i=0}^{n-1} s_i \leqslant T]} \right) \\
&= e^{-T} \frac{T^n}{n!},
\end{aligned}$$

since the integrals can be recognized as a fraction (exactly $\frac{1}{2^n}$) of the volume of the ball in the $\mathbb{R}^n$ with 1-norm and radius T. For the second expected value, we use that μ is invariant for the chain to rewrite

$$\mathbb{E}_\mu[\psi(\xi_k, \xi_{k+1})] = \mathbb{E}_\mu[\psi(\xi_0, \xi_1)] = \int \mu(dx_0) \int P(x_0, x_1) \psi(x_0, x_1) dx_1,$$

which makes every term of the second sum equal. Then,

$$\begin{aligned}
-H_T(\mathbb{P}_\mu | \hat{\mathbb{P}}_\mu) &= \sum_{n=1}^{\infty} e^{-T} \frac{T^n}{n!} \left(n \int \mu(dx_0) \int P(x_0, x_1) \psi(x_0, x_1) dx_1 \right) \\
&= T e^{-T} \sum_{n=1}^{\infty} \frac{T^{n-1}}{(n-1)!} \int \mu(dx_0) \int P(x_0, x_1) \psi(x_0, x_1) dx_1 \\
&= T \int \mu(dx_0) \int P(x_0, x_1) \psi(x_0, x_1) dx_1
\end{aligned}$$

Using the tools explored above, we can now give meaning to the entropy production rate. This formulation, however, is not universal and depends on the physical system and its dynamical laws. Different formulations for entropy production are explored on [33], where the authors made a review of the progress of these formulations. The point of view presented here relates to the one presented on [7].

Definition 2.19. *The entropy production rate is defined as*

$$ep := -H(\mathbb{P}_\mu | \hat{\mathbb{P}}_\mu) = -\lim_{T \to \infty} \frac{1}{T} \; H_T(\mathbb{P}_\mu | \hat{\mathbb{P}}_\mu).$$

Using the computations we made before, is possible to write the entropy production rate as

$$ep = \int \int \log \left(\frac{P(x,y)}{P(y,x)} \right) P(x,y) dy d\mu(x).$$

Notice that, if we try to apply the concept of entropy production to a reversible process, satisfying $P(x,y) = P(y,x)$, we ended up with $ep = 0$.

Proposition 2.20. *For all transition functions $P(x,y) > 0$, we have $ep \geqslant 0$.*

Proof. Since $\mathfrak{L}^*(1) = 0$, we have that $\int (Lf)(x)\, d\mu(x) = 0$ for every continuous function f. For $f = -\log \circ \theta$, we have that

$$\int\int [\log(\theta(x)) - \log(\theta(y))] P(x, y) dy d\mu(x) = 0.$$

Therefore, we can add this term to the entropy production rate without changing its value:

$$\begin{aligned} ep &= \int\int \log\left(\frac{\theta(x)P(x,y)}{\theta(y)P(y,x)}\right) P(x,y) dy d\mu(x) \\ &= \int\int \left[\frac{\theta(x)P(x,y)}{\theta(y)P(y,x)}\right] \log\left(\frac{\theta(x)P(x,y)}{\theta(y)P(y,x)}\right) \frac{\theta(y)}{\theta(x)} P(y,x) dy d\mu(x). \end{aligned}$$

Since $\int\int \frac{\theta(y)}{\theta(x)} P(y,x) dy d\mu(x) = 1$, we can use this as a probability measure in order to apply the Jensen inequality for the convex function $\psi(z) = z \log z$. In this way,

$$\begin{aligned} ep &= \int\int \psi\left(\frac{\theta(x)P(x,y)}{\theta(y)P(y,x)}\right) \frac{\theta(y)}{\theta(x)} P(y,x) dy d\mu(x) \\ &\geqslant \psi\left(\int\int \left[\frac{\theta(x)P(x,y)}{\theta(y)P(y,x)}\right] \frac{\theta(y)}{\theta(x)} P(y,x) dy d\mu(x)\right) \\ &= \psi\left(\int\int P(x,y) dy d\mu(x)\right) = \psi(1) = 0. \end{aligned}$$

□

The idea of this proof was similar to the one in Lemma 3.3 in [37].

Proposition 2.21. *The entropy production rate of the time reversal process is the same as the original process:*

$$ep^* := -H(\hat{\mathbb{P}}_\mu | \mathbb{P}_\mu) = ep.$$

Proof. Since $L(1) = 0$, we have that $\int (\mathfrak{L}^* g)(x)\, d\mu(x) = 0$ for every continuous function g, For $g = \log \circ \theta^2$, we have that

$$\int\int [\log(\theta^2(y)) - \log(\theta^2(x))] P^*(x, y) dy d\mu(x) = 0,$$

where $P^*(x,y) = \frac{\theta(y)}{\theta(x)} P(y,x)$. One can show that

$$\begin{aligned} ep^* &= \int\int \log\left(\frac{P^*(x,y)}{P^*(y,x)}\right) P^*(x,y) dy d\mu(x) \\ &= \int\int \log\left(\frac{\theta^2(y)P(y,x)}{\theta^2(x)P(x,y)}\right) P^*(x,y) dy d\mu(x) \end{aligned}$$

$$
\begin{aligned}
&= \int\int \log\left(\frac{P(y,x)}{P(x,y)}\right) P^*(x,y)dyd\mu(x) \\
&= \int\int \log\left(\frac{P(y,x)}{P(x,y)}\right) \frac{\theta(y)}{\theta(x)} P(y,x)dy[\theta(x)dx] \\
&= \int\int \log\left(\frac{P(y,x)}{P(x,y)}\right) P(y,x)d\mu(y)dx \\
&= \int\int \log\left(\frac{P(x,y)}{P(y,x)}\right) P(x,y)dyd\mu(x) = ep.
\end{aligned}
$$

□

2.5. Expansiveness of the semi-flow Θ_t on D

In this section, we consider an extended Skorohod space $\hat{D}$ of the *càdlàg* paths $w : \mathbb{R} \to [0,1]$ and $\hat{\Theta}_t$, $t \in \mathbb{R}$, the bidirectional flow on $\hat{D}$, acting on w by translation to the left: $(\hat{\Theta}_t w)(s) = w(s+t)$. One can show that two paths on D that coincide up to time t have a distance between them, using the Skorohod metric defined below, limited by e^{-t}. This means that, given two paths of such type, is possible to increase the distance by applying Θ_t.

Let Λ be the set of continuous functions f such that

$$\gamma(\lambda) := \operatorname*{ess\,sup}_{t\geq 0} |\log \lambda'(t)| < \infty$$

and recall the definition of the Skorohod distance (see [15]):

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

Let D^* be the set of paths $w : [0,+\infty) \to [0,1]$ continuous at left and with a limit at right. We can denote a typical path w in $\hat{D}$ as

$$w(s) = \langle w_1 | w_2 \rangle (s) = \begin{cases} w_1(-s), & \text{for } s < 0, \\ w_2(s), & \text{for } s \geqslant 0, \end{cases}$$

where $\omega_1 \in D^*$ and $w_2 \in D$. In this way, we can identify $\hat{D} \mapsto D^* \times D$ and define the projections $\Pi_1(w) = w_1$ and $\Pi_2(w) = w_2$. By convention, we will always use the time $t = 0$ to set this.

From two paths $w_1 \in D^*$ and $w_2 \in D$, we can go to $\hat{D}$ by $\langle w_1 | w_2 \rangle$, then apply $\hat{\Theta}_{-t}$ and go back to D using Π_2. By doing this, we ended up with

$$\Pi_2(\hat{\Theta}_{-t}\langle w_1|w_2\rangle)(s) = (w_1|_t\, w_2)(s) = \begin{cases} w_1(t-s), & \text{for } s < t, \\ w_2(s-t), & \text{for } s \geqslant t, \end{cases}$$

defined for $s \geqslant 0$, as Figure 2.2 shows.

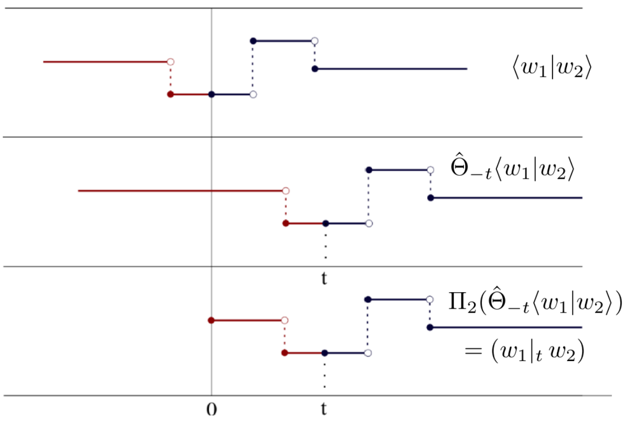


Figure 2.2: *The bilateral shift and the projection* Π_2

Proposition 2.22. *The continuous-time shift* Θ_t*, acting on the Skorohod space* D*, is expanding: given paths* $w_1 \in D^*$ *and* $w_2, w_2' \in D$*, for all* $t \geqslant 0$*,*

$$d\left((w_1|_t\, w_2), (w_1|_t\, w_2')\right) \leqslant \int_t^\infty e^{-u} du = e^{-t}. \tag{2.19}$$

Proof. Fix I as the identity function. Then, $\gamma(I) = 0$ and

$$\begin{aligned} d((w_1|_t\, w_2), (w_1|_t\, w_2')) &\leq \int_0^\infty e^{-u} d((w_1|_t\, w_2), (w_1|_t\, w_2'), I, u) du \\ &= \int_0^\infty e^{-u} \sup_{s \geqslant 0} q\left((w_1|_t\, w_2)(s \wedge u), (w_1|_t\, w_2')(s \wedge u)\right) du, \end{aligned}$$

where $q = r \wedge 1$ with r denoting the (Lebesgue) metric on the state space $[0, 1]$.

For $u < t$, the distance q above is $q(w_1(t - s \wedge u), w_1(t - s \wedge u)) = 0$. Otherwise, the distance q is upper bounded by 1. Then,

$$d\left((w_1|_t\, w_2), (w_1|_t\, w_2')\right) \leqslant \int_t^\infty e^{-u} du = e^{-t}.$$

□

Appendices

We will finish this chapter with some appendix sections related to it. These results are presented here in order to obtain a more complete chapter, so we recommend that the reader skips this part in a first reading. The tools in this part are properly called in the text if needed.

2.A. Existence of $K_t(x,y)$

In this section, we will show explicitly the existence of a function $K_t(x,y)$ which has a relation with the semigroup e^{tL} given by equation (2.4). We can write $L = \mathcal{L} - I$, where $\mathcal{L}$ is acting on functions as $(\mathcal{L}f)(x) = \int f(y)P(x,y)dy$. One can write down the action of the powers L^k which appear in e^{tL} in a simple way using the Newton binomial, since $\mathcal{L}$ and $-I$ commute:

$$L^k = (\mathcal{L} - I)^k = \sum_{j=0}^{k} \binom{k}{j} \mathcal{L}^j (-I)^{k-j} = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \mathcal{L}^j,$$

where $\mathcal{L}^0(f) = I(f) = f$. To go further, we need to consider the following transition functions: for all $k \geq 2$,

$$P^k(x,y) := \int \cdots \int P(x,z_1)P(z_1,z_2)\cdots P(z_{k-1},y)dz_1dz_2...dz_{k-1}.$$

Of course, $P^1(x,y) = P(x,y)$ and $P^{k+1}(x,y) = \int P^k(x,z)P(z,y)dz$. Now, we state that

$$(\mathcal{L}^k f)(x) = \int f(y)P^k(x,y)dy,$$

for every $k \geq 1$. To verify this, one can use induction:

$$\begin{aligned}
(\mathcal{L}^{k+1} f)(x) &= \mathcal{L}^k(\mathcal{L}f)(x) \\
&= \int (\mathcal{L}f)(y)P^k(x,y)dy \\
&= \int \int f(z)P(y,z)dz\, P^k(x,y)dy \\
&= \int f(z) \int P^k(x,y)P(y,z)dy\ dz \\
&= \int f(z)P^{k+1}(x,z)dz.
\end{aligned}$$

Above, to change the order of integration, we use the continuity of P and f over the compact state space or the continuity of P and the boundedness of f to assure that the integral is finite.

Now, we can compute L^k:

$$\begin{aligned}
(L^k f)(x) &= \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} (\mathcal{L}^j f)(x) \\
&= (-1)^k f(x) + \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \int f(y)P^j(x,y)dy.
\end{aligned}$$

Changing the order of terms, we get

$$\begin{aligned}
(L^k f)(x) &= (-1)^k f(x) + \int f(y) \left[\sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} P^j(x,y) \right] dy \\
&= (-1)^k f(x) + \int f(y)\; Q_k(x,y) dy,
\end{aligned}$$

where $Q_k(x,y)$ is the expression inside the bracket. Notice that

$$K_t(x,y) = \sum_{k=1}^{\infty} \frac{t^k}{k!}\; Q_k(x,y)$$

is our desired function, because

$$\begin{aligned}
(e^{tL} f)(x) &= f(x) + \sum_{k=1}^{\infty} \frac{t^k}{k!} (L^k f)(x) \\
&= f(x) + \sum_{k=1}^{\infty} \frac{t^k}{k!} \left[(-1)^k f(x) + \int f(y) Q_k(x,y) dy \right] \\
&= f(x) \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} + \int f(y)\; \sum_{k=1}^{\infty} \frac{t^k}{k!}\; Q_k(x,y)\; dy \\
&= f(x) e^{-t} + \int f(y) K_t(x,y)\; dy.
\end{aligned}$$

Considering the dynamics involved, the first term, which cannot be merged into $K_t(x,y)$, corresponds to the probability of not observing any jump in the interval $[0,t]$.

2.A.1 Properties of $K_t(x,y)$

We denote by $P_t = e^{tL}$. Then, we calculate

$$\partial_t (P_t f)(x) = -e^{-t} f(x) + \int f(y) (\partial_t K_t)(x,y) dy$$

and

$$\begin{aligned}
L(P_t f)(x) &= \int (P_t f)(y) P(x,y) dy - (P_t f)(x) \\
&= \int e^{-t} f(y) P(x,y) dy + \int \int f(z) K_t(y,z) dz P(x,y) dy - (P_t f)(x) \\
&= \int e^{-t} f(y) P(x,y) dy + \int f(z) \left(\int P(x,y) K_t(y,z) dy \right) dz \\
&\quad - e^{-t} f(x) - \int f(y) K_t(x,y) dy.
\end{aligned}$$

Reordering the terms, we conclude that $L(P_t f)(x)$ is equal to

$$-e^{-t}f(x) + \int f(y)\left(-K_t(x,y) + e^{-t}P(x,y) + \int P(x,z)K_t(z,y)dz\right)dy.$$

As P_t is the homogeneous semigroup generated by the infinitesimal generator L, the Kolmogorov equations imply that $L(P_t f) = \partial_t P_t(f) = P_t(Lf)$. From this, we conclude the equality of these two expressions, for every f. Then,

$$\partial_t K_t(x,y) = -K_t(x,y) + e^{-t}P(x,y) + \int P(x,z)K_t(z,y)dz.$$

The above is equal to

$$\partial_t K_t(x,y) = L(K_t(\cdot,y))(x) + e^{-t}P(x,y)$$

and, if we write down the other equation $\partial_t P_t f = P_t(Lf)$, the only change is the last integral for $\int P(z,y)K_t(x,z)dz$, which results in

$$\partial_t K_t(x,y) = L^*(K_t(x,\cdot))(y) + e^{-t}P(x,y).$$

Another way to explore $K_t(x,y)$ is looking to the property of semigroup: $P_s \circ P_t = P_{s+t}$. This leads us to

$$(P_{s+t}f)(x) = (e^{(s+t)L}f)(x) = e^{-(s+t)}f(x) + \int f(y)K_{s+t}(x,y)dy,$$

while $P_t(P_s f)(x)$ is equal to

$$\begin{aligned}
& e^{tL}(e^{sL}f)(x) \\
= \;& e^{-t}(e^{sL}f)(x) + \int (e^{sL}f)(y)K_t(x,y)dy \\
= \;& e^{-t}\left[e^{-s}f(x) + \int f(y)K_s(x,y)dy\right] \\
& + \int \left[e^{-s}f(y) + \int f(z)K_s(y,z)dz\right]K_t(x,y)dy \\
= \;& e^{-(t+s)}f(x) + \int f(y)\left(e^{-t}K_s(x,y) + e^{-s}K_t(x,y)\right)dy \\
& + \int\int f(z)K_s(y,z)K_t(x,y)dzdy \\
= \;& \int f(y)\left(e^{-t}K_s(x,y) + e^{-s}K_t(x,y) + \int K_t(x,z)K_s(z,y)dz\right)dy \\
& + e^{-(t+s)}f(x).
\end{aligned}$$

This means

$$K_{s+t}(x,y) = e^{-t}K_s(x,y) + e^{-s}K_t(x,y) + \int K_t(x,z)K_s(z,y)dz.$$

Notice that the last equation is the expression (1.3.1) in [4] for our transition function $p_t(y,dx) = K_t(x,y)dx + e^{-t}\delta_y(dx)$.

2.B. Existence of K_t^V

In this section, we will show explicitly the existence of a function $K_t^V(x,y)$ which has a relation with the semigroup $e^{t(L+V)}$ given by equation (2.7). Here we are considering a general L acting on functions according to equation (1.2). We will analyze the equation (2.6) in terms of the graphic construction of the jump process (X_t). This means that we will use that the trajectories are piece-wise constants:

$$\mathbb{E}_x\left[e^{\int_0^t V(X_r)dr} f(X_t)\right] = \sum_{n=0}^{\infty} \mathbb{E}_x\left[e^{\int_0^t V(X_r)dr} f(X_{T_n})\mathbb{1}_{[T_n \leqslant t < T_{n+1}]}\right],$$

where $0 = T_0 < T_1 < T_2 < \cdots$ are the times that X_t jumps.

The $n = 0$ term of this sum represents the time before the first jump. In this case, we have $s < T_1$ and the process $X_s \equiv x$. Then, this first term is equal to

$$e^{tV(x)}f(x)\mathbb{P}_x[\tau_0 > t] = e^{tV(x)}f(x)e^{-t\lambda(x)},$$

where τ_0 is a random variable with exponential distribution of parameter $\lambda(x)$.

For the terms $n \geqslant 1$, we need further analysis. For each k, set $x_k = X_{T_k}$ and let τ_k be a exponential random variable with parameter $\lambda(x_k)$. By this, under $\mathbb{1}_{[T_n \leqslant t < T_{n+1}]}$, we have

$$\int_0^t V(X_r)dr = \sum_{i=0}^{n-1} \tau_i V(x_i) + \left(t - \sum_{i=0}^{n-1} \tau_i\right) V(x_n).$$

Now, define

$$\varphi_t^{n,V}(x_0, ..., x_n) = \exp\left[\sum_{i=0}^{n-1} \tau_i V(x_i) + \left(t - \sum_{i=0}^{n-1} \tau_i\right) V(x_n)\mathbb{1}_{[T_n \leq t < T_{n+1}]}\right].$$

Notice that, for a fixed t, all functions $\varphi_t^{n,V}$ are null except for the one whose n is equal to the number of jumps until time t. In this way, using the kernel $P(x,dy) = P(x,y)dy$, the nth term of the summation becomes

$$\int \cdots \int \varphi_t^{n,V}(x_0, ..., x_n) f(x_n) P(x_0, x_1)dx_1 \cdots P(x_{n-1}, x_n)dx_n.$$

This expression is equal to $\int Q_t^{n,V}(x,x_n)f(x_n)dx_n$ if we define

$$Q_t^{n,V}(x,x_n)=\int\cdots\int\varphi_t^{n,V}(x_0,...,x_n)P(x_0,x_1)\cdots P(x_{n-1},x_n)dx_1\cdots dx_{n-1}.$$

Finally,

$$\begin{aligned}\left(e^{t(L+V)}f\right)(x) &= e^{tV(x)}f(x)e^{-t\lambda(x)}+\sum_{n=1}^{\infty}\int Q_t^{n,V}(x,x_n)f(x_n)dx_n\\ &= e^{t(V(x)-\lambda(x))}f(x)+\int\sum_{n=1}^{\infty}Q_t^{n,V}(x,y)f(y)dy\\ &= e^{t(V(x)-\lambda(x))}f(x)+\int K_t^V(x,y)f(y)dy,\end{aligned}$$

where $K_t^V(x,y)=\sum_{n=1}^{\infty}Q_t^{n,V}(x,y)$. Notice that $Q_t^{n,V}(x,y)\geqslant 0$ for all t. Furthermore, it is strictly positive when n is equal to the number of jumps until time t. Then, $K_t^V(x,y)>0$, for every $x,y\in[0,1]$.

2.B.1 Properties of K_t^V

Now, we proceed in the same way that we have done with K_t, looking for a differential equation that K_t^V satisfies, in the case of $\lambda\equiv 1$. For the semigroup $P_t^V=e^{t(L+V)}$, we have $(L+V)(P_t^Vf)=\partial_tP_t^V(f)=P_t^V((L+V)f)$. The middle term opens as

$$\partial_tP_t^V(f)(x)=(V(x)-1)e^{tV(x)-t}f(x)+\int\partial_tK_t^V(x,y)f(y)dy$$

while the last term is

$$\begin{aligned}&P_t^V((L+V)f)(x)\\ =\;& e^{tV(x)-t}(L+V)(f)(x)+\int K_t^V(x,y)(L+V)(f)(y)dy\\ =\;& e^{tV(x)-t}\left[\int P(x,y)f(y)dy+(V(x)-1)f(x)\right]\\ &+\int K_t^V(x,y)(L+V)(f)(y)dy.\end{aligned}$$

We get that, for every f,

$$\int\partial_tK_t^V(x,y)f(y)dy=e^{tV(x)-t}\int P(x,y)f(y)dy+\int K_t^V(x,y)(L+V)(f)(y)dy.$$

Using the definition of $L+V$ we can make a computation to rewrite the right-hand side of the above equation as

$$\int \left[e^{tV(x)-t} P(x,y) + \int K_t^V(x,z) P(z,y) dz + K_t^V(x,y)(V(y)-1) \right] f(y) dy,$$

which means that

$$\begin{aligned} \partial_t K_t^V(x,y) &= e^{tV(x)-t} P(x,y) + \int K_t^V(x,z) P(z,y) dz + K_t^V(x,y)(V(y)-1) \\ &= e^{tV(x)-t} P(x,y) + (L^* + V)(K_t^V(x,\cdot))(y). \end{aligned}$$

Similarly, if we open the other equation $(L+V)(P_t^V f) = \partial_t P_t^V(f)$, we conclude

$$\partial_t K_t^V(x,y) = e^{tV(x)-t} P(x,y) + (L+V)(K_t^V(\cdot,y))(x).$$

2.C. Another look of Feynman-Kac formula for symmetrical L

Consider X_t a continuous-time process with state space $[0,1]$ and infinitesimal generator L. Let f and V be two functions on $[0,1]$ taking values on $\mathbb{R}$. For any fixed $T > 0$, we denote by $\hat{X}_s = X_{T-s}$ the time-reversal process and by $\hat{L}$ its generator. For this process $\hat{X}$, we have that, by Feynman-Kac (see Proposition 1.13), the function

$$u_t(x) = \hat{\mathbb{E}}_x \left[e^{\int_0^t V(\hat{X}_s) ds} f(\hat{X}_t) \right]$$

is the solution of the partial differential equation

$$\begin{cases} \partial_t u_t(x) = \hat{L} u_t(x) + V(x) u_t(x)\,, & t \in (0,T] \\ u_0(x) = f(x) \end{cases}.$$

If L is symmetric, i.e., $\hat{L} = L$, this partial differential equation is the same for the original process X, whose known solution, by Feynman-Kac, is

$$v_t(x) = \mathbb{E}_x \left[e^{\int_0^t V(X_s) ds} f(X_t) \right].$$

Then, for any $t \in (0,T]$, we have that $v_t = u_t$. Looking at the paths, we get

$$\int_{w(0)=x} e^{\int_0^t V(w(s)) ds} f(w(t)) d\mathbb{P}(w) = \int_{w(T)=x} e^{\int_0^t V(w(T-s)) ds} f(w(T-t)) d\mathbb{P}(w).$$

Making a change of variables, we can rewrite this expression as

$$\int_{w(0)=x} e^{\int_0^t V(w(s))ds} f(w(t)) d\mathbb{P}(w) = \int_{w(T)=x} e^{\int_{T-t}^{T} V(w(s))ds} f(w(T-t)) d\mathbb{P}(w). \tag{2.20}$$

3. Thermodynamic formalism for diffusions

In this chapter, we consider X_t the Brownian Motion whose state space is a Riemannian compact manifold M. In order to simplify the notation we will assume that $M = S^1$, the same state space we considered on Chapter 2. For the general case, similar results can be obtained, but then we would get more cumbersome expressions. The results presented here are based on [30], a joint work with Artur Lopes and Adriana Neumann.

In the same way we did in the previous chapter, we use a Hölder potential V and an *a priori* probability on the trajectories space, in this case the space $C = C([0, T], M)$ of continuous functions, to introduce a Ruelle operator and get, due to a normalization procedure, a Gibbs Markov process and a Gibbs probability on C. From this, again we were able to introduce the concepts of relative entropy and pressure.

3.1. The Model

The Brownian Motion is a Markov process whose infinitesimal generator $L = \frac{1}{2}\Delta$ is on the form of equation (1.5), where $\Delta = \frac{\partial^2}{\partial x^2}$ denotes the Laplacian on the Riemannian manifold $M = S^1$. This operator L is self-adjoint (see [44]) and acts on functions $f \in C^2(M)$. The trajectories of the process are on C, the space of continuous functions, and induce, in this set, a probability $\mathbb{P}_\mu$, where μ denotes the initial probability.

Let μ be the invariant probability. The associated Markov process is stationary for the flow Θ_s, $s \geqslant 0$, and the probability $\mathbb{P} = \mathbb{P}_\mu$ obtained in this way will play the role of the *a priori* probability (in a similar way as in [6, 25, 32]).

Let $V : M \to \mathbb{R}$ a Hölder continuous function and consider the operator $L + V$, which acts on functions $f \in C^2(M)$ by the expression

$$(L + V)(f)(x) = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x) + V(x)f(x),$$

for all $x \in M$. There exists a positive differentiable eigenfunction $F : M \to \mathbb{R}$ associated with an eigenvalue λ_V for the above operator (see [**?**, 47]).

For $t \geqslant 0$, we consider the Ruelle operator

$$(P_t^V f)(x) := \mathbb{E}_x \left[e^{\int_0^t V(X_r)\, dr} f(X_t) \right],$$

for all continuous function $f : M \to \mathbb{R}$ and $x \in M$. By Feynman-Kac, P_t^V defines a non markovian semigroup associated with the infinitesimal operator $L + V$ (see

Section 1.3.2). Using self-adjointness, we get the same relation from the jump process case for a symmetric L. The Figure 3.1 visually supports this statement. Again, the left-hand side is more suitable for the Feymann-Kac formula while the right one is the natural generalization of the classical Ruelle operator from the discrete-time setting.

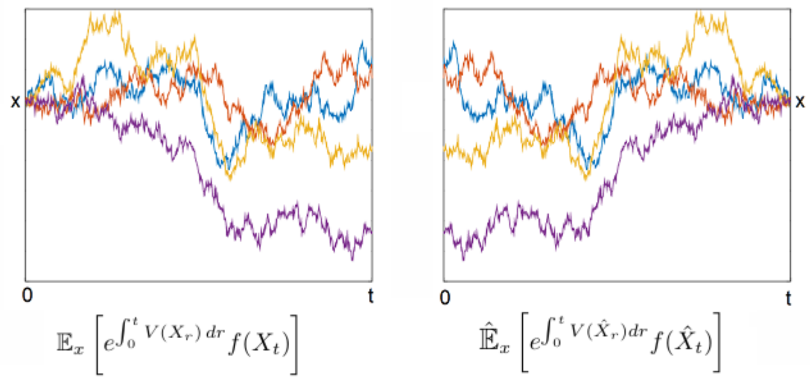


Figure 3.1: *The Ruelle operator at natural and reversal time. Created using simulations of [12].*

3.2. On the continuous time Gibbs state for the potential V

Let λ_V be the main eigenvalue of $L+V$ and F_V the strictly positive differentiable eigenfunction associated with λ_V (for the existence theorems see [4, 14, 47]). To make simply the notation we will denote F_V by F.

Using these λ_V and F, define

$$(\mathcal{P}_t^V f)(x) \;= \mathbb{E}_x\left[e^{\int_0^t V(X_r)dr}\frac{F(X_t)}{e^{\lambda_V t}F(x)}f(X_t)\right] = \frac{(P_t^V Ff)(x)}{e^{\lambda_V t}F(x)}.$$

Then $(\mathcal{P}_t^V 1)(x) = 1$, $\forall x \in M$. This defines a Markov semigroup, which is what we were looking for.

We define the operator $\mathcal{L}_V$ acting on $f \in C^2(M)$ as

$$\begin{aligned}
(\mathcal{L}_V f)(x) \quad &= \quad \frac{1}{F(x)}(L+V)(Ff)(x) - f(x)\lambda_V \\
&= \quad \frac{1}{F(x)}\left[\frac{1}{2}\Delta(Ff)(x) + V(x)F(x)f(x)\right] - f(x)\lambda_V \\
&= \quad \frac{1}{2}\Delta f(x) + \frac{1}{F(x)}\frac{\partial F}{\partial x}(x)\frac{\partial f}{\partial x}(x) + \frac{(L+V)(F)(x)f(x)}{F(x)} - \lambda_V f(x) \\
&= \quad \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x) + \frac{\partial}{\partial x}\log(F(x))\frac{\partial f}{\partial x}(x).
\end{aligned}$$

Proposition 3.1. *The operator $\mathcal{L}_V$ is the infinitesimal generator associated with the semigroup $\mathcal{P}_t^V$.*

Notice that a process induced by this kind of infinitesimal generator corresponds to a Brownian Motion with nonhomogeneous drift $\frac{\partial}{\partial x}\log(F(x))$.

Proof. To prove this association, we need to observe that

$$\frac{(\mathcal{P}_t^V f)(x) - f(x)}{t} = \frac{1}{e^{\lambda_V t} F(x)} \left(\frac{(P_t^V Ff)(x) - (Ff)(x)}{t} \right) + f(x) \left(\frac{e^{-\lambda_V t} - 1}{t} \right).$$

Taking the limit as t goes to zero the expression turns into

$$\partial_t(\mathcal{P}_t^V f)(x) = \frac{1}{F(x)} \partial_t (P_t^V Ff)(x) - f(x)\lambda_V = (\mathcal{L}_V f)(x).$$

□

From now on, we will elaborate on the properties of initial invariant probability μ_V for the operator $\mathcal{L}_V$. In other words, μ_V is a probability in M such that, for any $f \in \mathcal{C}^2(M)$ and $t \geqslant 0$, we have

$$\int (\mathcal{P}_t^V f) d\mu_V = \int f d\mu_V \quad \text{or equivalently} \quad \int (\mathcal{L}_V f) d\mu_V = 0.$$

The following lemma will give us this invariant measure.

Lemma 3.2. *Let $G \in C^1(M)$ and define an operator $A : C^2(M) \to \mathbb{R}$ as*

$$Af = \frac{1}{2}\frac{\partial^2 f}{\partial x^2} + \frac{\partial G}{\partial x}\frac{\partial f}{\partial x},$$

for all $f \in C^2(M)$. Then, a measure μ such that $\frac{d\mu}{dx} = e^{2G}$ satisfies

$$\int Af \, d\mu = 0.$$

Proof. This proof follows from the Radon-Nikodym theorem and integration by parts. □

Thus, taking $G = \log F$, we get that $\tilde{\mu}_V$ satisfies $\frac{d\tilde{\mu}_V}{dx} = F^2$ is the invariant measure for $\mathcal{L}_V$. This measure maybe is not a probability, then we will consider the normalized measure

$$d\mu_V(x) = \frac{F^2(x)}{\gamma_V} dx,$$

where $\gamma_V = \int_M F^2(x)\, dx$.

Remark. *There is another way to find an invariant measure for $\mathcal{L}_V$. Following the reasoning of Section 2.2 one can find an eigenprobability ν_V of $L+V$ associated with eigenvalue λ_V. Then, consider*

$$\mu_V(dx) = F(x)\nu_V(dx),$$

where F is the eigenfunction associated with eigenvalue λ_V. We have

$$\int (\mathcal{L}_V f) d\mu_V = \int ((L+V)(Ff) - Ff\lambda_V)\, d\nu_V = 0.$$

Definition 3.3. *Given a Hölder function $V : M \to \mathbb{R}$, we define a continuous-time Markov process $\{Y_t^V, t \geqslant 0\}$ with state-space M whose infinitesimal generator is $\mathcal{L}_V$ and the initial stationary probability is μ_V. We call this process the continuous time Gibbs state for the potential V. This process induced a probability $\mathbb{P}^V_{\mu_V}$ on the space C, which we call the Gibbs probability for the potential V.*

Remark. *Suppose V is of class C^∞ and has a finite number of points with derivative zero. Let λ be the main eigenvalue of $L+V$ and F be the eigenfunction associated with λ. One can show an interesting property relating oscillations of V and the oscillations of the main eigenfunction F: If $V : [0,1] \to \mathbb{R}$ has only two points with derivative zero (V has a unique point of maximum and a unique point of minimum), then the eigenfunction F has less than four points with derivative zero. Given a value c there exist at most three values x such that $V(x) = c$. Suppose F has many values with derivative zero. Then, between each two of these points, there exists another one x_1 with $F''(x_1) = 0$. From $\frac{1}{2}F''(x_1) + V(x_1)F(x_1) = \lambda F(x_1)$ we get that $V(x_1) = \lambda$. By hypothesis, we can get at most three of these intervals, that means, four points of F with derivative zero. One can generalize this for V with more oscillations in a similar way. The analogous property for potentials and eigenfunctions in the setting where the state space has no differentiable structure is not so clear how to get it.*

3.3. Relative Entropy, Pressure and the equilibrium state for V

In this section, we will repeat for this process the same we did on Section 2.3 for the jump process. First of all, in order to define the relative entropy, we will analyze the Radon-Nikodym derivative of $\mathbb{P}^V_x$ with respect to the measure $\mathbb{P}_x$, induced by the Brownian Motion, with initial probability δ_x. Remember that the Radon-Nikodym derivative must satisfy

$$\mathbb{E}^V_x[G(w_{T_1}, w_{T_2}, \ldots, , w_{T_k})] = \mathbb{E}_x\left[G(w_{T_1}, w_{T_2}, \ldots, , w_{T_k})\, \frac{d\mathbb{P}^V_x}{d\mathbb{P}_x}\bigg|_{\mathcal{F}_t}\right],$$

for all $k \in \mathbb{N}$, $0 = T_0 < T_1 < \cdots < T_k = t < T$ and $G : M^k \to \mathbb{R}$. For this, it is enough to consider, for any $k \in \mathbb{N}$, functions $f_i : M \to \mathbb{R}$, $i \in \{1, \ldots, k\}$, a time partition as above and study the following:

$$\begin{aligned}
& \mathbb{E}_x^V[f_1(X_{T_1})f_2(X_{T_2})\ldots f_k(X_{T_k})] \\
= & \int_M P_{T_1}^V(x,dx_1)f_1(x_1)\cdots\int_M P_{T_k-T_{k-1}}^V(x_{k-1},dx_k)f_k(x_k) \\
= & \int_M P_{T_1}^V(x,dx_1)f_1(x_1)\cdots\int_M P_{D_{k-1}}^V(x_{k-2},dx_{k-1})(f_{k-1}\mathcal{P}_{D_k}^V f_k)(x_{k-1}) \\
= & \cdots \\
= & \mathcal{P}_{T_1}^V(f_1\ldots\mathcal{P}_{T_k-T_{k-1}}^V(f_k))(x),
\end{aligned}$$

where $P_t^V(x,dy)$ is the kernel of $\mathcal{P}_t^V$ and D_k denotes the k-th difference $T_k - T_{k-1}$.

To fix ideas consider $k = 2$ and analyze

$$\begin{aligned}
& \mathbb{E}_x^V[f_1(X_{T_1})f_2(X_{T_2})] \\
= & \mathcal{P}_{T_1}^V(f_1\mathcal{P}_{T_2-T_1}^V(f_2))(x) \\
= & \mathbb{E}_x\left[e^{\int_0^{T_1}V(X_r)dr}\frac{F(X_{T_1})f_1(X_{T_1})}{e^{\lambda_V T_1}F(x)}(\mathcal{P}_{T_2-T_1}^V f_2)(X_{T_1})\right] \\
= & \mathbb{E}_x\left[e^{\int_0^{T_1}V(X_r)dr}\frac{f_1(X_{T_1})}{e^{\lambda_V T_2}F(x)}\mathbb{E}_{X_{T_1}}\left[e^{\int_0^{T_2-T_1}V(X_r)dr}F(X_{T_2-T_1})f_2(X_{T_2-T_1})\right]\right] \\
= & \mathbb{E}_x\left[e^{\int_0^{T_1}V(X_r)dr}\frac{f_1(X_{T_1})}{e^{\lambda_V T_2}F(x)}e^{\int_{T_1}^{T_2}V(X_r)dr}F(X_{T_2})f_2(X_{T_2})\right] \\
= & \mathbb{E}_x\left[f_1(X_{T_1})f_2(X_{T_2})\left(\frac{e^{\int_0^{T_2}V(X_r)dr}}{e^{\lambda_V T_2}}\frac{F(X_{T_2})}{F(x)}\right)\right].
\end{aligned}$$

We can do an analogous computation for any k. Remembering that $T_k = t$, we have

$$\left.\frac{d\mathbb{P}_x^V}{d\mathbb{P}_x}\right|_{\mathcal{F}_t} = \exp\left\{\log F(X_t) - \log F(X_0) - \int_0^t(\lambda_V - V(X_r))dr\right\}.$$

Notice that, if we denote $\log F = g$, this function satisfies

$$\frac{1}{2}\left[\frac{\partial^2 g}{\partial x^2} + \left(\frac{\partial g}{\partial x}\right)^2\right] = \frac{1}{2}\left[\frac{\partial}{\partial x}\left(\frac{\frac{\partial F}{\partial x}}{F}\right) + \left(\frac{\frac{\partial F}{\partial x}}{F}\right)^2\right] = \frac{\frac{1}{2}\frac{\partial^2 F}{\partial x^2}}{F} = \frac{LF}{F} = \lambda_V - V. \tag{3.1}$$

The last equality is due to $(L+V)F = \lambda_V F$. Then, we have the following definition, according to that the probability $\mathbb{P}_x^V$, induced by the Gibbs Markov process with the initial probability δ_x, is admissible.

Definition 3.4. *The probability $\tilde{\mathbb{P}}_\mu$ on C is called admissible if exists a function*

$g \in C^2(M)$ *such that, for all* $t \geqslant 0$,

$$\left.\frac{d\tilde{\mathbb{P}}_x}{d\mathbb{P}_x}\right|_{\mathcal{F}_t} = \exp\left\{ g(X_T) - g(X_0) - \frac{1}{2}\int_0^t \left[\frac{\partial^2 g}{\partial x^2}(X_r) + \left(\frac{\partial g}{\partial x}\right)^2 (X_r)\right] dr \right\}.$$

Remark. *Denote by* $\tilde{X}_t$ *the process which the law in* C *is the probability* $\tilde{\mathbb{P}}_x$*. Notice that*

$$\Gamma(g,g) = \frac{1}{2}\frac{\partial^2}{\partial x^2}(g^2) - 2g\left(\frac{1}{2}\frac{\partial^2 g}{\partial x^2}\right) = \left(\frac{\partial g}{\partial x}\right)^2,$$

where Γ *denotes the Carré du Champ operator. If we define* $G = Lg + \frac{1}{2}\Gamma(g,g)$*, the pair* (g, G) *allow us to write the process* $\tilde{X}$ *as* X^g*, using the notation of Section 1.B.2. By Proposition 1.26, we conclude* $\tilde{L} = L + \Gamma(g, \cdot)$*. Then, the admissible process* $\tilde{X}_t$ *is a Brownian Motion with drift* $\frac{\partial g}{\partial x}$*.*

Following in the same way as in Section 2.3, we take the invariant probability for $\tilde{L}$, which we will denote by $\tilde{\mu}$. By Lemma 3.2, this probability is such that $d\tilde{\mu}(x) = \frac{e^{2g(x)}}{\tilde{\gamma}}dx$, where $\tilde{\gamma} = \int_M e^{2g(x)}\, dx$. Then, to define the relative entropy of the $\tilde{\mathbb{P}}_{\tilde{\mu}}$ with respect to $\mathbb{P}_{\tilde{\mu}}$ we set

$$H_T(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) = -\int_M\int_C \log\left(\left.\frac{d\tilde{\mathbb{P}}_x}{d\mathbb{P}_x}\right|_{\mathcal{F}_T}\right)(\omega)d\tilde{\mathbb{P}}_x(\omega)d\tilde{\mu}(x).$$

Remark. *Using Jensen's inequality, one can show that* $H_T(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) \leqslant 0$*. Negative entropy appears naturally when one analyzes a dynamical system with the property that each point has an uncountable number of preimages (see [28, 29]).*

Using the expression of the Radon-Nikodym derivative, we get

$$\begin{aligned} & H_T(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) \\ = \; & \mathbb{E}_{\tilde{\mu}}\left[g(X_0) - g(X_T) + \frac{1}{2}\int_0^T \left[\frac{\partial^2 g}{\partial x^2}(X_r) + \left(\frac{\partial g}{\partial x}\right)^2 (X_r)\right] dr\right] \\ = \; & \int_M \left\{(\tilde{P}_0 g)(x) - (\tilde{P}_T g)(x) + \frac{1}{2}\int_0^T \tilde{P}_r\left[\frac{\partial^2 g}{\partial x^2} + \left(\frac{\partial g}{\partial x}\right)^2\right](x)dr\right\} d\tilde{\mu}(x), \end{aligned}$$

where $\tilde{P}_t$ is the semigroup associated with $\tilde{L}$. By the Definition 2.15 and the Ergodic Theorem, the relative entropy is

$$H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) = \frac{1}{2}\int_M \left[\frac{\partial^2 g}{\partial x^2} + \left(\frac{\partial g}{\partial x}\right)^2\right] d\tilde{\mu}.$$

Finally, we can state the main result of this section:

Proposition 3.5. *The pressure of the potential V is given by*

$$\boldsymbol{P}(V) = H(\mathbb{P}^V_{\mu_V}|\mathbb{P}_{\mu_V}) + \int_M V d\mu_V = \lambda_V.$$

Proof. The second equality in the statement of the theorem comes from

$$\begin{aligned} H(\mathbb{P}^V_{\mu_V}|\mathbb{P}_{\mu_V}) + \int_M V d\mu_V &= \int_M \frac{1}{2}\left[\frac{\partial^2}{\partial x^2}\log F + \left(\frac{\partial}{\partial_x}\log F\right)^2] + V\right] d\mu_V \\ &= \int_M \frac{LF}{F} + V d\mu_V = \lambda_V, \end{aligned}$$

by equation (3.1) and $(L+V)F = \lambda_V F$.

In order to finish the proof, we need to analyze the variational formula for the pressure (see Definition 2.16) to show that $\mathbf{P}(V) \leqslant \lambda_V$. Notice that

$$\begin{aligned} H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) + \int_M V d\tilde{\mu} &= \frac{1}{\tilde{\gamma}}\frac{1}{2}\int_M \left[\frac{\partial^2 g}{\partial x^2} + \left(\frac{\partial g}{\partial x}\right)^2\right] e^{2g}dx + \frac{1}{\tilde{\gamma}}\int_M V e^{2g} dx \\ &= \frac{1}{\tilde{\gamma}}\int_M \left[\left(\frac{1}{2}\frac{\partial^2 g}{\partial x^2} + \left(\frac{\partial g}{\partial x}\right)^2\right) - \frac{1}{2}\left(\frac{\partial g}{\partial x}\right)^2 + V\right] e^{2g}dx \\ &= \frac{1}{\tilde{\gamma}}\int_M \left[V - \frac{1}{2}\left(\frac{\partial g}{\partial x}\right)^2\right] e^{2g}dx. \end{aligned}$$

The last equality follows from Lemma 3.2 for $f = g$ and $G = g$. Using that $V = \lambda_V - \frac{LF}{F}$, we can rewrite

$$H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) + \int_M V d\tilde{\mu} = \lambda_V + \frac{1}{\tilde{\gamma}}\frac{1}{2}\int_M \left[-\frac{\frac{\partial^2 F}{\partial x^2}}{F} - \left(\frac{\partial g}{\partial x}\right)^2\right] e^{2g}dx.$$

Applying integration by parts, the above expression becomes

$$\begin{aligned} & \lambda_V + \frac{1}{\tilde{\gamma}}\frac{1}{2}\left\{\int_M \left[\frac{\partial F}{\partial x}\frac{\partial}{\partial x}\left(\frac{e^{2g}}{F}\right)\right] dx - \int_M \left(\frac{\partial g}{\partial x}\right)^2 e^{2g}dx\right\} \\ = \; & \lambda_V - \frac{1}{\tilde{\gamma}}\frac{1}{2}\int_M \left(\frac{\partial}{\partial x}\log F - \frac{\partial g}{\partial x}\right)^2 e^{2g}dx, \end{aligned}$$

what is less or equal than λ_V. □

The immediate consequence of this result is the fact that just like in Chapter 2, the Gibbs probability is the one that maximizes the pressure.

Final Considerations

At the end of this thesis, we want to summarize that we contributed to the theory of continuous-time thermodynamic formalism with compact state space. Although, it is important to point out some of the questions that are still open for future studies in these settings. Although we were able to get, for continuous-time Markov processes with non-countable state space, most of the results that we would like to extend from the simplest cases, it will still take some work to obtain a generalization of the Perron-Frobenius Theorem that works in our setting, in order to replace Assumption 1.1.

Furthermore, one natural question to ask is about some type of large deviation principle (for the unperturbated process), following a similar way to what was done in [31] or [32]. Another thing we can try to extend from the discrete-time setting is to consider an extra parameter β, which is a multiple of the inverse of the temperature, on potential V and study what happens when the temperature goes to zero by making β increase (see [5]). One can also try to extend the potential V to a more general case whose domain is the set of trajectories, then the V presented here can be seen as a particular case that depends only on the value at time zero.

Another possibility for future studies is to try to replicate, in our setting, what was done in [1, 17, 18, 41, 42], where the authors use an idea of the Ruelle operator as a guiding principle to describe nonequilibrium stationary states in general. The purpose of this study is a better understanding of a model for the chaotic hypothesis for a single (moving) particle system held in a nonequilibrium stationary state. This model is described by properties of SBR (Sinai-Bowen-Ruelle) probabilities for Axiom A (or Anosov) systems and entropy production rate. In this case, the potential is fixed as the Lyapunov exponent. The reason for such interest is that the real physical problem behaves, in many respects, as if they were Anosov systems as far as their properties of physical interest are concerned. We wonder if our setting, where V is general, also provides a sketch (as an alternative for the Anosov one) for the chaotic hypothesis.

Bibliography

[1] P. Adamo, R. Belousov and L. Rondoni. Fluctuation-Dissipation and Fluctuation Relations: From Equilibrium to Nonequilibrium and Back. *Large Deviations in Physics*. Lecture Notes in Physics. Vol. 885, pp. 92-133, Springer Verlag (2014).

[2] R. Baldasso, O. Menezes, A. Neumann and R. Souza. Exclusion Process with Slow Boundary. *Journal of Statistical Physics*. Vol. 167, no. 5, pp. 1112-1142 (2017).

[3] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim. Macroscopic fluctuation theory for stationary non-equilibrium states. *Journal of Statistical Physics*. Vol. 107, no. 3, pp. 635-675 (2002).

[4] D. Bakry, I. Gentil and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. Springer (2014).

[5] A. Baraviera, L. M. Cioletti, A. O. Lopes, J. Mohr and R. R. Souza. On the general XY Model: positive and zero temperature, selection and non-selection. *Reviews in Mathematical Physics*. Vol. 23, n. 10, pp. 1063-1113 (2011).

[6] A. Baraviera, R. Exel and A. Lopes. A Ruelle Operator for continuous time Markov chains. *São Paulo Journal of Mathematical Sciences*. Vol. 4, n. 1, pp. 1-16 (2010).

[7] T. Benoist, V. Jakšic, Y. Pautrat and C-A. Pillet. On entropy production of repeated quantum measurements I. General Theory. *Communications in Mathematical Physics* Vol. 357, pp. 77-123 (2018).

[8] A. Bobrowski. *Functional Analysis for Probability and Stochastic Processes*. Cambridge University Press (2005).

[9] V. Bogachev, N. Krylov, M. Röckner and S. Shaposhnikov. Fokker-Planck-Kolmogorov Equations. *Mathematical Surveys and Monographs*. Vol. 207, American Mathematical Society (2015).

[10] J. E. Brasil, J. Knorst and A. O. Lopes. Thermodynamic formalism for continuous-time quantum Markov semigroups: the detailed balance condition, entropy, pressure and equilibrium quantum processes. *arXiv preprint* (2022).

[11] H. Bruin. *Notes on Thermodynamic Formalism*. Accessed on 12/02/2022. Avaliable on <https://www.mat.univie.ac.at/ bruin/TF.pdf>.

[12] F. Burgaud, C. Frank and D. Mavris. An Aircraft Development Methodology Aligning Design and Strategy to Support Key Decision Making. *57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (2016).

[13] M. D. Donsker and S. R. S. Varadhan. On a Variational Formula for the Principal Eigenvalue for Operators with Maximum Principle. *Proceedings of the National Academy of Sciences*. Vol. 72, no. 3, pp. 780-783 (1975).

[14] Y. V. Egorov and M. A. Shubin. *Foundations of the classical theory of partial differential equations*. Vol. 1, Springer Science & Business Media (1998).

[15] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons (2009).

[16] L. C. Evans. *Partial differential equations*. American Mathematical Sociery (2010).

[17] G. Gallavotti. Fluctuation patterns and conditional reversibility in nonequilibrium systems. *Annales de l'I.H.P. Physique théorique*. Vol. 70, no. 4, pp. 429-443 (1999).

[18] G. Gallavotti and E. G. D. Cohen. Dynamical Ensembles in Nonequilibrium Statistical Mechanics. *Physical review letters*. Vol. 74, no. 14, p. 2694 (1995).

[19] J. W. Gibbs. *Elementary Principles of Statistical Mechanics*. Charles Scribner's Sons (1902).

[20] D. Q. Jiang and D. Jiang. *Mathematical theory of nonequilibrium steady states: on the frontier of probability and dynamical systems*. Lectures Notes in Math. Springer Science & Business Media (2004).

[21] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*. Vol. 113, Springer Science & Business Media (2012).

[22] S. Karlin and H. E. Taylor. *A second course in stochastic processes*. Elsevier (1981).

[23] Y. Kifer. Principal eigenvalues, topological pressure, and stochastic stabilty of equilibrium states. *Israel Journal of Mathematics*. Vol. 70, no. 1, pp. 1-47 (1990).

[24] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems*. Springer Science & Business Media (1998).

[25] J. Knorst, A. O. Lopes, G. Muller and A. Neumann. Thermodynamic Formalism on the Skorokhod space: the continuous time Ruelle operator, entropy, pressure, entropy production and expansiveness. *arXiv preprint* (2022).

[26] T. M. Liggett. *Continuous Time Markov Processes*. American Mathematical Society (2010).

[27] A. O. Lopes and J. K. Mengue. On information gain, Kullback-Leibler divergence, entropy production and the involution kernel. *Discrete and Continuous Dynamical Systems - Series A*. Vol. 42, no. 7, pp. 3593-3627 (2022).

[28] A. O. Lopes, J. K. Mengue, J. Mohr and R. R. Souza. Entropy and Variational Principle for one-dimensional Lattice Systems with a general a-priori probability: positive and zero temperature. *Ergodic Theory and Dynamical Systems*. Vol. 35, no. 6, pp. 1925-1961 (2015).

[29] A. O. Lopes, J. Mohr, R. Souza and P. Thieullen. Negative Entropy, Zero temperature and stationary Markov chains on the interval. *Bulletin of the Brazilian Mathematical Society*. Vol. 40, no. 1, pp. 1-52 (2009).

[30] A. O. Lopes, G. Muller and A. Neumann. Diffusion Processes: entropy, Gibbs states and the continuous time Ruelle operator. *arXiv preprint* (2022).

[31] A. O. Lopes and A. Neumann. Large Deviations for stationary probabilities of a family of continuous time Markov chains via Aubry-Mather theory. *Journal of Statistical Physics*. Vol. 159, no. 4, pp. 797-822 (2015).

[32] A. O. Lopes, A. Neumann and P. Thieullen. A thermodynamic formalism for continuous time Markov chains with values on the Bernoulli Space: entropy, pressure and large deviations. *Journal of Statistical Physics*. Vol. 152, no. 5, pp. 894-933 (2013).

[33] G. T. Landi and M. Paternostro. Irreversible entropy production: From classical to quantum. *Reviews of Modern Physics*. Vol. 93, no. 3 (2021).

[34] C. Maes and K. Netocny, Time-Reversal and Entropy. *Journal of Statistical Physics*. Vol. 110, No. 1, pp. 269-310 (2003).

[35] C. Maes, K. Netocny, and B. Shergelashvili. A selection of nonequilibrium issues. *Methods of Contemporary Mathematical Statistical Physics*. pp. 247-306 (2009).

[36] G. Muller and A. Neumann. Some general results for continuous-time Markov chain. *preprint* (2023).

[37] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*. Vol. 187-188, pp. 1-268 (1990).

[38] H. Qian and Y. Wang. Mathematical Representation of Clausius' and Kelvin's Statements of the Second Law and Irreversibility. *Journal of Statistical Physics*. Vol. 179, no. 3, pp. 808-837 (2020).

[39] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Vol. 293, Springer Science & Business Media (2013).

[40] L. C. Rogers and D. Williams. *Diffusions, markov processes, and martingales: Volume 1, foundations*. Vol. 1, Cambridge university press (2000).

[41] D. Ruelle. A generalized detailed balance relation. *Journal of Statistical Physics*. Vol. 164, no. 3, pp. 463-471 (2016).

[42] D. Ruelle. Positivity of entropy production in nonequilibrium statistical mechanics. *Journal of Statistical Physics*. Vol. 85, no. 1, pp. 1-23 (1996).

[43] D. Ruelle. *Thermodynamic Formalism*. Addison-Wesley Publishing Company (1978).

[44] R. S. Strichartz. Analysis of the Laplacian on the complete Riemannian manifold. *Journal of functional analysis* Vol. 52, no. 1, pp. 48-79 (1983).

[45] W. Strock. *An introduction to the theory of large deviations*. Springer (2012).

[46] D. W. Stroock and S. S. Varadhan. *Multidimensional diffusion processes*. Springer Science & Business Media (1997).

[47] M. Taylor. *Partial differential equations II: Qualitative studies of linear equations*. Vol. 116, Springer Science & Business Media (2013).