Thermodynamic Formalism on the Skorokhod space: the continuous time Ruelle operator, entropy, pressure, entropy production and expansiveness

J. Knorst, A. O. Lopes, G. Muller, and A. Neumann

August 3, 2022

Abstract

Consider the semi-flow given by the continuous time shift $\Theta_t : \mathscr{D} \to \mathscr{D}$, $t \geq 0$, acting on the Skorokhod space $\mathscr{D}$ of *càdlàg* paths (right continuous with left limits) $w : [0,\infty) \to S^1$, where S^1 is the unitary circle (one can also take $[0,1]$ instead of S^1). We equip the space $\mathscr{D}$ with the Skorokhod metric and we show that the semi-flow is expanding. We also introduce a stochastic semi-group e^{tL}, $t \geq 0$, where L (the infinitesimal generator) acts linearly on continuous functions $f : S^1 \to \mathbb{R}$. This stochastic semi-group and an initial vector of probability π defines an associated stationary shift-invariant probability $\mathbb{P}$ on the Polish space $\mathscr{D}$. This probability $\mathbb{P}$ will play the role of an *a priori* probability. Given such $\mathbb{P}$ and a Hölder potential $V : S^1 \to \mathbb{R}$, we define a continuous time Ruelle operator, which is described by a family of linear operators $\mathbb{L}^t_V$, $t \geq 0$, acting on continuous functions $\varphi : S^1 \to \mathbb{R}$. More precisely, given any Hölder V and $t \geq 0$, the operator $\mathbb{L}^t_V$, is defined by

$$\varphi \to \psi(y) = \mathbb{L}^t_V(\varphi)(y) = \int_{w(t)=y} e^{\int_0^t V(w(s))\,ds}\, \varphi(w(0))\, d\mathbb{P}(w).$$

We show the existence of an eigenvalue λ_V and an associated Hölder eigenfunction $\varphi_V > 0$ for the semi-group $\mathbb{L}^t_V$, $t \geq 0$. After a coboundary procedure we obtain another stochastic semi-group, with infinitesimal generator L_V, and this will define a new probability $\mathbb{P}_V$ on $\mathscr{D}$, which we call the Gibbs (or, equilibrium) probability for the potential V. We define entropy, for some shift-invariant probabilities on $\mathscr{D}$, and we consider a variational problem of pressure. Finally, we define entropy production and we analyze its relation with time-reversal and symmetry of L. We wonder if the point of view described here provides a sketch (as an alternative for the Anosov one) for the chaotic hypothesis for a particle system held in a nonequilibrium stationary state, as delineated by Ruelle, Gallavotti, and Cohen.

1 Introduction

We consider the semi-flow given by the continuous time shift $\Theta_t : \mathscr{D} \to \mathscr{D}$, $t \geq 0$, acting on the Skorokhod space $\mathscr{D}$ of *càdlàg* paths (right continuous with left limits) $w : [0,\infty) \to S^1$, where S^1 is the unitary circle (one can take $[0,1]$ instead of S^1). We will prefer to state the results in $[0,1]$. The set $\mathscr{D}$ is equipped with the Skorokhod metric. The Skorokhod space $\mathscr{D}$ is a non compact Polish space. We will show that continuous time shift Θ_t, $t \geq 0$, is expanding (see Proposition 6.1).

Continuous time Stochastic Processes X_t, $t \geq 0$, taking values on $[0,1]$ are described by probabilities $\mathbb{P}$ on $\mathscr{D}$. To say that the process is stationary is equivalent to say that the associated probability $\mathbb{P}$ is invariant for the action of the shift Θ_t, $t \geq 0$.

The results presented in the initial part of our work are in some way related to [2], [20], [23] and [24]. Our main purpose here is to describe a version of Thermodynamic Formalism for semi-flows specified by infinitesimal generators. More precisely, in section 3 we follow the program of introducing a Ruelle operator from a potential and an *a priori* probability (in a similar fashion as in [3], [25], [2] and [20]).

We introduce a stochastic semi-group e^{tL}, $t \geq 0$, where L (the infinitesimal generator) acts on continuous functions $f : [0,1] \to \mathbb{R}$. This stochastic semi-group and an initial vector of probability π defines an associated stationary shift-invariant probability $\mathbb{P}$ on the $\mathscr{D}$ (see [19]). This probability $\mathbb{P}$ will play the role of an *a priori* probability (a continuos time version of the point of view of [25] and [3]).

Given the *a prori* probability $\mathbb{P}$ on $\mathscr{D}$ and a Hölder continuous potential $V : [0,1] \to \mathbb{R}$, we define the Ruelle operator $\mathbb{L}_V^t$, $t \geq 0$, in such way that for $\varphi : [0,1] \to \mathbb{R}$, we get $\mathbb{L}_V^t(\varphi) = \psi$, $t \geq 0$, when

$$\varphi \to \psi(y) = \mathbb{L}_V^t(\varphi)(y) = \int_{w(t)=y} e^{\int_0^t V(w(s))\,ds}\,\varphi(w(0))\,d\mathbb{P}(w). \tag{1}$$

The above expression can be recognized as in Feynman-Kac form if the infinitesimal generator is symmetric according to Lemma 1 in Section 3 and figure 1 (see also [23]).

The Feynmann-Kac formula is the partial differential equation

$$\frac{\partial u}{\partial t} + Lu + V\,u = 0.$$

General results for continuous-time Markov chains that were specially designed to be applicable to our setting appear on [38].

Note that expression (1) depends also on L (because $\mathbb{P}$ depends on L).

From the *a priori* probability $\mathbb{P}$ on $\mathscr{D}$, in section 4 we are able to introduce the concepts of entropy, for a certain class of shift-invariant probabilities on $\mathscr{D}$, and pressure for a potential $V : [0,1] \to \mathbb{R}$ (see Definition 4.1 and expression (40)). For the existence of an eigenvalue and a positive eigenfunction for the Ruelle operator (see Theorem 1), an assumption on regularity of the potential V will be required (Hölder or Lipschitz continuous will be enough) as discussed in Proposition 3.4 in section 3. Example 3 presents explicit expressions for the eigenvalue and the eigenfunction $f : [0,1] \to \mathbb{R}$ solutions for a certain class of infinitesimal generators L and quadratic potentials V. In the Appendix Section 9 we show that this regularity assumption is necessary. From L and V, after a kind of coboundary procedure, we obtain another stochastic semi-group, with infinitesimal generator L_V, and this will define a new probability $\mathbb{P}_V$ on $\mathscr{D}$, which we call the equilibrium probability for the potential V (see Definitions 3.5, 3.9, 3.8, Lemma 2 and expressions (26) and (25)). The initial stationary vector of probability for such a process is given by Proposition 3.6. A nice formula related to the main eigenvalue is given by (41). Note that V is completely independ of the dynamics of the shift Θ_t, $t \geq 0$, and the *a priory* probabilty defined by L.

We define entropy production in section 5 and we discuss some properties related to time-reversal and the symmetry of the infinitesimal generator L (see Propositions 5.2 e 5.4). Related results for continuous time quantum channels (where the infinitesimal generator is a Lindbladian) appear in [6].

In [11], [12], [35], [1] and [36], the authors use an idea of Ruelle's as a guiding principle to describe nonequilibrium stationary states in general. The purpose 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 probabilities for Axiom A (or Anosov) systems and entropy production rate (see also [14], [27],[28] and [32]). 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.

The Appendix sections 7 and 8 are of technical nature and they have the purpose of analyzing some integral kernels which naturally appear in our reasoning.

In Appendix section 10 we present the details of the claims mentioned in Example 1 which describes in explicit form an interesting working case.

Some of our results are related to the ones in [7], [17], [18], [27], [28], [23], [14], [13], [21] and [29].

2 Motivation and Preliminaries

In order to motivate our reasoning, we will begin with a review of some basic and simple properties of Markov chains taking values on $[0,1]$ (in a similar fashion one consider the case where the process is taking values in S^1.

Consider $P(x,y) > 0$, $P : [0,1] \times [0,1] \to \mathbb{R}$ continuous such that for all $y \in [0,1]$

$$\int P(x,y)\, dx = 1. \tag{2}$$

Note that is not true that the supremum of P is smaller than 1.

Let $\theta : [0,1] \to \mathbb{R}$ be a strictly positive function such that

$$\iint P(x,y)\theta(y)\, dx\, dy = 1,$$

and also that for any x

$$\int P(x,y)\theta(y) dy = \theta(x). \tag{3}$$

The function θ is the initial invariant vector of probability for a stationary discrete-time Markov chain with values on $[0,1]$.

The above reasoning was just to explain what is a line sum 1 stochastic matrix with values on $[0,1]$. In analogy with Markov chains with finite state space, $P(x,y)$ should be seen as a matrix with entries in $[0,1] \times [0,1]$ where x is in the vertical axis and y is in the horizontal axis (see [22] for related results).

We define the infinitesimal generator L acting on the left in **periodic** functions $f : [0,1] \to \mathbb{R}$, by

$$L(f)(y) = \int f(x)P(x,y)dx - f(y),$$

which by (2) means

$$L(f)(y) = \int \big[f(x) - f(y)\big] P(x,y)\,dx \tag{4}$$

Note that $L(1) = 0$.

We call L the *a priori* infinitesimal generator.

Later we consider (see the expression (25)) the action of infinitesimal generators of the form

$$f \to \gamma(y) \int \big[f(x) - f(y)\big]\, P(x,y)dx, \tag{5}$$

where γ is positive, as in [20].

We will consider $L : \mathscr{L}^2(dx) \to \mathscr{L}^2(dx)$ and the dual $L^* : \mathscr{L}^2(dx) \to \mathscr{L}^2(dx)$, which acts on probability densities $g : [0,1] \to \mathbb{R}$ by

$$L^*(g)(x) = \int P(x,y)g(y)dy - g(x). \tag{6}$$

Now setting $\mu(dx) = \theta(x)dx$, the probability measure with density θ, we get $L^*(\theta) = 0$, by (3). In this context, this means that μ is invariant for the action of L^*.

L and L^* are bounded operators.

Note that for any f, g we have for $\mathscr{L}^2$ inner product

$$< g, L(f) > = < L^*(g), f > .$$

L acts on (the left) functions $f : [0,1] \to \mathbb{R}$ (on the variable y) and L^* acts on (the right) densities $g(x)dx$ (or on probabilities).

A nice reference for continuous time processes with infinitesimal generator L is [7], where it is considered a strongly continuous semigroup mapping T^t, $t \geq 0$, acting on the set of continuous functions f on compact manifold, satisfying for all $t \geq 0$: $T^t(f)$ is a positive function, if f is positive and $T^t 1 = 1$.

The operator e^{tL}, for fixed $t \geq 0$, is an integral operator, that is, there exist a function $K_t : [0,1] \times [0,1] \to \mathbb{R}^+$ such that

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

The funcion $K_t : [0,1] \times [0,1] \to \mathbb{R}^+$ satisfies the following equations (see appendix 1):

$$\frac{d}{dt} K_t(x,y) = \int K_t(x,z)P(z,y)dz - K_t(x,y) + e^{-t} P(x,y), \tag{7}$$

$$\begin{aligned} \frac{d}{dt} K_t(x,y) &= L(K_t(x,\cdot))(y) + e^{-t} P(x,y) \\ &= L^*(K_t(\cdot,y))(x) + e^{-t} P(x,y), \end{aligned}$$

and

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

Example 1. *Take* $P(x,y) = \cos[(x-y)2\pi]/2 + 1$. *This* P *is symmetric and continuous on* $[0,1]$. *Since* $\int \cos[(x-y)2\pi]dx = 0$, *for any* $y \in [0,1]$ *we get that* $\int P(x,y)dx = 1$.

$K_t(x,y), t \geq 0$ *can be explicitly expressed by*

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

The Lebesgue probability dx is the unique invariant probability. The proofs of these claims are presented in Appendix section 10.

$\diamondsuit$

For each t fixed, $K_t(x,y)$ can be seen as a matrix with entries in $[0,1]\times[0,1]$, where x is in the vertical axis and y in the horizontal axis.

We denote by $\mathscr{D}$ the Skorokhod space of *càdlàg* paths (right continuous with left limits) $w:(0,\infty)\to[0,1]$ (see [9] for general properties)

The continuous time shift $\Theta_t:\mathscr{D}\to\mathscr{D}$, $t\geq 0$, is defined in such way that, $\Theta_t(w_a)=w_b$, if for all $s\geq 0$, we have $w_b(s)=w_a(s+t)$.

We say that a probability P on $\mathscr{D}$ is invariant for the semi-flow $\Theta_t:\mathscr{D}\to\mathscr{D}$, $t\geq 0$, if for all measurable set $C\subset\mathscr{D}$, and any $t\geq 0$, we have that $P(C)=P(\,(\Theta_t)^{-1}(C)\,)$.

The kernel K_t , $t\geq 0$, defines a Markov Process X_t, $t\geq 0$, with values on $[0,1]$. Given an initial density function φ_0 on $[0,1]$, this Markov Process determines a probability $\mathbb{P}$ on $\mathscr{D}$. For example, for the cylinder set $C=\{X_0\in(a_0,b_0),X_{t_1}\in(a_1,b_1),X_{t_2}\in(a_2,b_2),X_{t_3}\in(a_3,b_3)\}$ we get that

$$\mathbb{P}(C)=\int_{a_0}^{b_0}\int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3}K_{t_1}(x_0,x_1)\,K_{t_2-t_1}(x_1,x_2)\,K_{t_3-t_2}(x_2,x_3)\,\varphi_0(x_0)dx_3dx_2dx_1dx_0.$$

Given L (as in (4)) 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

$$\int L(f)(x)\theta(x)dx=0. \tag{8}$$

Moreover, we assume that L is such that the above-defined θ is unique.

Definition 2.1. *Given L (as in* (4)*) and an initial density* $\varphi_0=\theta$ *satisfying* (8)*, we get a continuous time stationary Markov process* X_t*,* $t\geq 0$*, with values on* $[0,1]$ *(see [4],[5] or [19]). This process defines a probability* $\mathbb{P}=\mathbb{P}_{L,\theta}$ *on the Skorokhod space* $\mathscr{D}$*. This probability* $\mathbb{P}$ *is invariant for the shift* Θ_t*,* $t\geq 0$.

Consider an infinitesimal generator L (where L is given by (4)) for the semigroup e^{tL}, $t\geq 0$. This semigroup satisfies $e^{tL}(1)=1$. Moreover, $e^{tL^*}(\theta)=\theta$, where L^* was given by (6) and θ satisfies (8).

Now we take a continuous function $V:[0,1]\to\mathbb{R}$, which will be called a potential. In Statistical Mechanics $H=-e^V$ should correspond in some sense to the Hamiltonian. For some results, we will assume that V is of Hölder class.

By definition, the operator $L+V:\mathscr{L}^2(dx)\to\mathscr{L}^2(dx)$ (acting on the left on functions on the variable y) is such that

$$(L+V)(f)(y)=g(y)=\int\big[f(x)-f(y)\big]P(x,y)\;dx+V(y)f(y).$$

The dual operator acts (on the right) on density functions g on the variable x

$$(L^*+V)(g)(x)=f(x)=\int P(x,y)g(y)dy-\;g(x)+V(x)\,g(x).$$

If $P(x,y)$ is symmetric the spectral properties of L and L^* is the same.

Example 2. $P(x,y) = \cos[(x-y)2\pi]/2+1$ *and* $V(y) = (y-1/2)^2$.
Then,

$$(L+V)(f)(y) = \int f(x)[\cos[(x-y)2\pi]/2+1]\,dx - f(y) + (y-1/2)^2 f(y).$$

$\diamondsuit$

Now consider the semigroup $e^{t\,(L+V)}$. By Feymann-Kac (see [33] or [4]), we can write

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

where X_t, $t \geq 0$, is the Markov process with infinitesimal generator L. This semigroup is not a stochastic semigroup.

This operator is an integral operator, that is, there exist $K_t^V : [0,1]\times[0,1] \to \mathbb{R}^+$, such that,

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

The function $K_t^V : [0,1]\times[0,1] \to \mathbb{R}^+$ satisfies the following equation (see (55) and (56) in the Appendix)

$$\begin{aligned}\frac{d}{dt}K_t^V(x,y) &= (L+V)(K_t^V(x,\cdot))(y) + e^{t(V(x)-1)}P(x,y)\\ &= (L+V)^*(K_t^V(\cdot,y))(x) + e^{t(V(y)-1)}P(x,y).\end{aligned}$$

3 Ruelle Operator

We denoted by $\mathscr{D} := \mathscr{D}\big([0,+\infty),[0,1]\big)$ the path space of *càdlàg* trajectories taking values in $[0,1]$ (see [19]). This space is endowed with the Skorokhod metric.

Remember that the flow $\Theta_t : \mathscr{D} \to \mathscr{D}$, $t \geq 0$, satisfies: given t we have that $\Theta_t(w_1) = w_2$, if for all $s \geq 0$, $w_2(s) = w_1(s+t)$.

We assume in this section that L is of the form (4).

Lemma 1. *If L is symmetric, then*

$$\int_{w(0)=y} e^{\int_0^t V(w(s))\,ds}\,\varphi(w(t))\,d\mathbb{P}_L(w) = \int_{w(t)=y} e^{\int_0^t V(w(s))\,ds}\,\varphi(w(0))\,d\mathbb{P}_L(w). \quad (9)$$

A more general version of the above result appears in the Appendix. This Lemma, in particular, is an immediate consequence of (58) for $T = t$.

Definition 3.1. *Given L and V, consider, for each fixed t, the continuous time Ruelle operator* $\mathbb{L}_V^t$, $t \geq 0$, *where* $\mathbb{L}_V^t : C^0[0,1] \to C^0[0,1]$, *associated to V (in a similar way to [2]): in this case for* $t \geq 0$ *we denote* $\mathbb{L}_V^t(\varphi) = \psi$, *when*

$$\varphi \to \psi(y) = \int_{w(0)=y} e^{\int_0^t V(w(s))\,ds}\,\varphi(w(t))\,d\mathbb{P}_L(w),$$

that is,

$$\varphi \to \psi(y) = \mathbb{L}_V^t(\varphi)(y) = e^{t(L+V)}(\varphi)(y), \quad (10)$$

where $\varphi, \psi : [0,1] \to \mathbb{R}$.

This operator (which was considered in similar cases in [2] and [20]) is the continuous time version of the classical Ruelle operator (discrete time case). Indeed, Lemma 1 confirms the claim and figure 1 schematically support this statement.

The left hand side of expression (9) is more suitable for the Feymann-Kac formula.

According to our notation the continuous time Ruelle operator $\mathbb{L}_V^t$, $t \geq 0$, is a family of linear operators indexed by t.

Definition 3.2. *Given L and $V : [0,1] \to \mathbb{R}$ we say that the familiy of Ruelle operators $\mathbb{L}_V^t$, $t \geq 0$, is normalized if $\mathbb{L}_V^t(1) = 1$, for all $t \geq 0$.*

Given L, in the case V is constantly equal to zero, the family of Ruelle operators $\mathbb{L}_0^t, t \geq 0$, is normalized.

In this case $\mathbb{L}_0^t(f) = g$, when

$$f \to g(y) = \int_{w(0)=y} f(w(t))\, d\mathbb{P}_L(w) = e^{tL}(f)(y) = \mathbb{L}_0^t(f)(y),$$

where $f, g : [0,1] \to \mathbb{R}$ are periodic.

Definition 3.3. *Given L and V we say that $f : [0,1] \to \mathbb{R}$ is an eigenfunction of the continuous time Ruelle operator $\mathbb{L}_V^t$, $t \geq 0$, associated to the eigenvalue $\lambda \in \mathbb{R}$, if for all $t \geq 0$,*

$$\mathbb{L}_V^t(f) = e^{\lambda t} f. \tag{11}$$

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

Assume that the positive function $f : [0,1] \to \mathbb{R}$ is such that

$$(L+V)(f) = \lambda f, \tag{12}$$

then, for all $t \geq 0$,

$$e^{t\,(L+V)}(f) = e^{t\,\lambda} f, \tag{13}$$

that is, $f : [0,1] \to \mathbb{R}^+$ is an eigenfunction (for the semigroup generated by the infinitesimal generator $L+V$) associated to the eigenvalue $\lambda \in \mathbb{R}$.

We say that such λ (which can be positive or negative) is the main eigenvalue for $L+V$.

Example 3. *Consider the periodic function $g(x) = \frac{6}{7}(1 + x(1-x))$ and $P(x,y)$ defined for $x, y \in [0,1]$ (or, in $S^1 \times S^1$), by*

$$P(x,y) = g(x+y),\ \textit{if}\ (x+y) < 1,\ \textit{and}\ P(x,y) = g(x+y-1),\ \textit{if}\ (x+y) \geq 1. \tag{14}$$

One can show that the kernel P is symmetric and therefore the corresponding density θ satisfying (3) *is equal to* 1. *Consider now the function $V : S^1 = [0,1) \to \mathbb{R}$, given by $V(y) = 1 + \frac{1}{7} y(1-y)$. Taking*

$$\lambda = \frac{1}{70}(35 + \sqrt{1345})\ \textit{and}\ y \to f(y) = \frac{1}{10}(35 + \sqrt{1345}) + (1-y)y > 0,$$

we get for all $y \in S^1$

$$(L+V)(f)(y) = \int_0^1 P(x,y) f(x) dx - f(y) + V(y) f(y) = \lambda f(y), \tag{15}$$

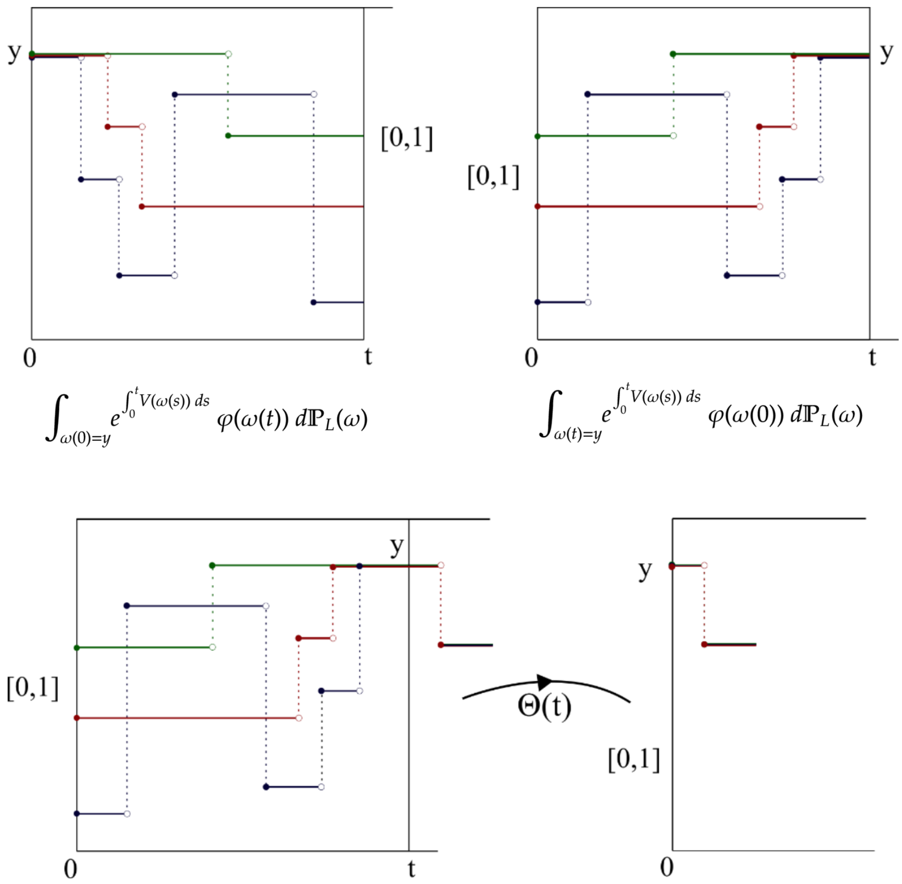


Figure 1: The point $y \in [0,1]$ is the value at time $t = 0$ of the path obtained as the image - by the continuous time shift Θ_t - of the set of paths described above.

and therefore, f and λ solve (12) *for such P and V.*

As P is symmetric the function $f : S^1 \to \mathbb{R}$ also solves

$$(L^* + V)(f) = \lambda f.$$

More generally, given V of the form $V(y) = r + s\,y(1-y)$, and the quadratic density $g(x) = \frac{d+cx(1-x)}{d+c/6}$ (defining P as in (14)*), one can find a density f of the form $f(y) = a + b\,y(1-y)$, such that, for some λ we get* (15)*.*

$\diamondsuit$

Given λ as in (12), if $g(y) > 0$ is such that

$$(L^* + V)(g) = \lambda g, \tag{16}$$

then, for all $t \geq 0$,

$$e^{t\,(L^*+V)}(g) = e^{t\,\lambda} g,$$

that is $g : [0,1] \to \mathbb{R}^+$ is an eigenfunction (for the semigroup generated by the infinitesimal generator $L^* + V$) associated to the eigenvalue $\lambda \in \mathbb{R}$. It is natural to assume the normalization condition $\int g(y)dy = 1$ so we can see g as a density.

Note that we ask for f and g to have the same eigenvalue λ.

A natural normalization assumption for f is to assume that

$$\int f(x)g(x)dx = 1. \tag{17}$$

In this case $\pi(x) = f(x)g(x)$ is a density on $[0,1]$, and this will be important later (see proposition 3.6).

Compare all this with pages 52-54 in [16] and around page 113 in [33].

We want to show that given L and V one can find a solution for (12). This will follow from Krein-Rutman Theorem (see [8] and [22]).

Theorem 1. *Assume that $V : [0,1] \to \mathbb{R}$ is Hölder , then, there exist $\lambda \in \mathbb{R}$, $\ell : [0,1] \to \mathbb{R}^+$ and $r : [0,1] \to \mathbb{R}^+$, where ℓ, r are also Hölder (in particular $\ell, r \in \mathscr{L}^2(dx)$), such that,*

$$\ell(L+V) = \lambda \ell$$

and

$$(L+V)r = \lambda r.$$

Proof. The action of $\ell \to \ell(L+V)$ (acting on the left side) can be seen as the action on the dual, i.e., $\ell \to (L^* + V)\ell$ (acting on the right side). Consider $\mathscr{H}_\alpha$ the Banach space of Hölder continuous real functions on $[0,1]$ with constant α and the norm

$$\|h\|_\alpha = \|h\| + \sup_{x \neq y} \frac{|h(y) - h(x)|}{|y-x|^\alpha}.$$

The above supremmum is denoted as $Höl_h$. Let $K \subset \mathscr{H}_\alpha$ be the cone of positive α-Hölder functions in $[0,1]$. The interior of K, denoted K^o, is the set of strictly positive α-Hölder functions. We will use item (a) in Theorem 19.3 in [8].

Indeed, consider $z = \|V\| + 1$. We claim that $(L+V+zI)$ is a strongly positive operator (take non-null positive functions to strictly positive functions). Then, from Krein-Rutman Theorem (see Theorem 19.3 page 228 in [8]) there exists a unique eigenfunction r for $(L+V+zI)$ in the set K^o. The same is true for the left eigenfunction ℓ.

We now check the assumptions of Krein-Rutman theorem. Notice that $-\|V\| \leq V(x)$ and therefore $V(x) + z - 1 \geq 0$ for all $x \in [0,1]$. It follows that

$$(L+V+zI)f(x) = \int f(y)P(y,x)dy + (V(x)+z-1)f(x) \geq \int f(y)P(y,x)dy.$$

We started with $f(y) \geq 0$ and $P(y,x) > 0$, for every $x, y \in [0,1]$. For $f \neq 0$, by continuity there exists an open set in which f is strictly positive and then $(L+V+zI)f(x) > 0$. This means $(L+V+zI)(K \backslash \{0\}) \subset K^o$.

Now to see that the operator is compact, consider $f \in B_\alpha(1)$, the unitary ball of $\mathscr{H}_\alpha$. Then $\|f\| \leq 1$, $Höl_f \leq 1$ and

$$|(L+V+zI)f(y) - (L+V+zI)f(x)| \leq |Lf(y) - Lf(x)|$$

$$+|V(y)f(y) - V(x)f(x)| + z|f(y) - f(x)|$$

$$\leq \int |f(w)||P(w,y)-P(w,x)|\,dw+(z+1)|f(y)-f(x)|+$$

$$|V(y)||f(y)-f(x)|+|f(x)||V(y)-V(x)|.$$

Since $P:[0,1]^2\to\mathbb{R},\ V:[0,1]\to\mathbb{R}$ are continuous and therefore uniformly continuous, given $\varepsilon>0$, there exists $\delta>0$ such that $|x-y|<\delta\Rightarrow|P(w,y)-P(w,x)|<\frac{\varepsilon}{3}$ for every $w\in[0,1]$ and $|V(y)-V(x)|<\frac{\varepsilon}{3}$. If we take δ small enough to have $\delta^\alpha<\frac{\varepsilon}{3(z+1+\|V\|)}$ also, we get

$$|(L+V+zI)f(y)-(L+V+zI)f(x)|\leq\frac{2\varepsilon}{3}+(z+1+\|V\|)|f(y)-f(x)|$$

$$\leq\frac{2\varepsilon}{3}+(z+1+\|V\|)|y-x|^\alpha<\varepsilon.$$

This means that $(L+V+zI)(B_\alpha(1))$ is equicontinuous. It is bounded also, since $L+V+zI$ is bounded by $1+2\|V\|$, and finally the operator is compact. The same analysis can be done for $L+V+zI$ acting on the right, which is equivalent to the action of L^*+V+zI.

From the above we get the existence of $\ell,r:[0,1]\to(0,\infty)$ and a λ which satisfy

$$(L+V-\lambda I)\,r=0, \tag{18}$$

and

$$\ell(L+V-\lambda I)=0, \tag{19}$$

by Krein-Rutman (as in [1]).

□

We denote by ℓ_V, r_V and $\lambda(V)$ the solutions of the above equations.
We consider normalization conditions: $\int r_V(y)dy=1$ and (see (17))

$$\int r_V(x)\ell_V(x)dx=1. \tag{20}$$

The equation for the above left eigenfunction $\ell=\ell_V$ is

$$\int \ell(z)P(z,y)\,dz-(\,1+\lambda(V)\,-\,V(y))\,\ell(y)=0. \tag{21}$$

for any y.
The equation for the above right eigenfunction $r=r_V$ is

$$\int P(z,x)r(x)\,dx-(\,1+\lambda(V)\,-V(x)\,)\,r(x)=0, \tag{22}$$

for any x.

It follows from the existence of ℓ satisfying (19) and the above that:

Proposition 3.4. *Given L and the Hölder continuous function* $V:[0,1]\to\mathbb{R}$ *there exists* f *and* λ*, such that,*

$$\mathbb{L}^t_V(f)=e^{\lambda t}f. \tag{23}$$

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

$$\gamma_V(y) = 1 + \lambda(V) - V(y), \qquad Q_V(z,y) = \frac{\ell(z)P(z,y)}{\ell(y)\gamma_V(y)}, \tag{24}$$

$$\mathscr{L}_V(f)(y) = \gamma_V(y)\int [f(z) - f(y)]Q_V(z,y)\,dz \tag{25}$$

and

$$\mathscr{P}_t^V(f)(y) = \frac{e^{t(L+V)}(\ell f)(y)}{e^{\lambda(V)t}\ell(y)}. \tag{26}$$

Lemma 2. *The operator $\mathscr{P}_t^V$, defined in* (26)*, is the semi-group associated to the infinitesimal generator$\mathscr{L}_V$, defined in* (25)*, that is,*

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

Proof. Since

$$\gamma_V(y) = \int \ell(z)\tfrac{P(z,y)}{\ell(y)}\,dz,$$

we have

$$\int Q_V(z,y)\,dz = \int \frac{\ell(z)P(z,y)}{\ell(y)\gamma_V(y)}\,dz = 1. \tag{27}$$

We can rewrite $\frac{\mathscr{P}_t^V(f)(y)-f(y)}{t}$ as

$$\frac{1}{e^{\lambda_V t}\ell(y)}\Big(\frac{e^{t(L+V)}(\ell f)(y) - \ell(y)f(y)}{t}\Big) + f(y)\Big(\frac{e^{-\lambda(V)t} - 1}{t}\Big).$$

Taking limit as $t \to 0$, we get $\frac{1}{\ell(y)}(L+V)(\ell f)(y) - \lambda(V)f(y)$. Using (24), the last expression becomes

$$\int [f(z) - f(y)]\,\frac{\ell(z)}{\ell(y)}\,P(z,y)\,dz = \mathscr{L}^V f(y).$$

□

The semigroup $e^{t\mathscr{L}_V}, t \geq 0$, is normalized.

Definition 3.5. *The continuous Markov chain process with values on $[0,1]$ and infinitesimal generator $\mathscr{L}_V$ has an initial stationary positive density $\pi_V : [0,1] \to \mathbb{R}$. We denote the associated stationary continuous time Markov chain by X_t^V, $t \geq 0$. We call such Process the* **Gibbs Markov Process for the potential** *V (see section in 3 in [20]).*

In Proposition 3.6 we show that $\pi_V = \ell_V\, r_V$, where the normalization conditions (20) are assumed to be satisfied.

Note that from (27) we get that $Q_V(x,y)$, $x,y \in [0,1]$, defines a continuos time Markov chain with a generator L of the form (4) where we replace P by Q_V.

Then, multiplying (22) by $\ell(y)$ and integrating over y we get

$$\iint r(z)P(z,y)\ell(y)dzdy + \int V(y)r(y)\ell(y)dy = \lambda + 1. \tag{28}$$

Note that when P is symmetric we have from (28) that $\ell = r$ and

$$\iint P(x,y)\ell(y)\ell(x)dxdy + \int V(y)\ell^2(y)dy = \lambda + 1. \tag{29}$$

The above reasoning is similar to section 5 in [20].

Lemma 3. *The dual of the operator $\mathscr{L}_V$, defined in* (25) *is the operator*

$$g \to \mathscr{L}_V^*(g)(z) = \int \gamma_V(y) g(y) Q_V(z,y) dy - \gamma_V(z)\, g(z). \tag{30}$$

Proof. Given the functions f, g we get

$$\int \mathscr{L}_V(f)(y)\, g(y)\, dy = \int [\gamma_V(y) \int [f(z) - f(y)] Q_V(z,y)\, dz]\, g(y) dy =$$

$$\iint \gamma_V(y) f(z) g(y)\, Q_V(z,y)\, dz\, dy - \int \gamma_V(y) f(y) g(y)\, [\int Q_V(z,y)\, dz]\, dy =$$

$$\int f(z)\, [\int \gamma_V(y)\, g(y)\, Q_V(z,y)\, dy]\, dz - \int f(z)\, \gamma_V(z)\, g(z)\, dz =$$

$$\int f(z)\, [\int \gamma_V(y)\, g(y)\, Q_V(z,y)\, dy - \gamma_V(z)\, g(z)\,] dz = \int f(z)\, \mathscr{L}_V^*(g)(z)\, dz.$$

□

Given the the Markov Process X_t^V, $t \geq 0$, with infinitesimal generator $\mathscr{L}_V$, we ask: how to get the stationary initial probability $\pi_V = \pi : [0,1] \to \mathbb{R}$.

We assume that $\int r_V(z) dz = 1$ and, moreover, that $\int \ell_V(z) r_V(z) dz = 1$.

Proposition 3.6. *The density $\pi_V(z) = \ell_V(z)\, r_V(z)$ satisfies $\mathscr{L}_V^*(\pi_V) = 0$.*

Proof. From (24) and (21) we get for any point z

$$\mathscr{L}_V^*(\pi_V)(z) = \int \gamma_V(y) \ell_V(y) r_V(y) Q_V(z,y) dy - \gamma_V(z)\, \ell_V(z) r_V(z) =$$

$$\int \gamma_V(y) \ell_V(y) r_V(y) \frac{\ell_V(z)\, P(z,y)}{\ell_V(y) \gamma_V(y)} dy - \gamma_V(z)\, \ell_V(z) r_V(z) =$$

$$\int r_V(y)\, \ell_V(z)\, P(z,y) dy - \gamma_V(z)\, \ell_V(z) r_V(z) =$$

$$[\int r_V(y)\, P(z,y) dy - \gamma_V(z)\, r_V(z)\,]\, \ell_V(z) = 0\, \ell_V(z) = 0.$$

□

From the above we get for any z

$$\int \frac{\gamma_V(y)}{\gamma_V(z)} \ell_V(y) r_V(y) Q(z,y) dy = \ell_V(z) r_V(z) \tag{31}$$

Corollary 3.7. *Given V we get for any $t \geq 0$ and continuous function $\varphi : [0,1] \to \mathbb{R}$:*

$$\mathscr{P}_t^V(\varphi)(x) = \mathbb{E}_x\, [e^{\int_0^t V(X_s)\, ds}\, \frac{\ell_V(X_t)}{e^{t\, \lambda(V)}\, \ell_V(x)} \varphi(X_t)\,] \tag{32}$$

Moreover,

$$\mathscr{P}_t^V(1) = e^{t\, \mathscr{L}_V}(1) = 1 \tag{33}$$

and

$$(\mathscr{P}_t^V)^*(\pi_V) = e^{t\, \mathscr{L}_V^*}(\pi_V) = \pi_V. \tag{34}$$

The expected value calculated above is relative to the *a priori* probability P on the Skorokhod space.

Definition 3.8. *The Markov Process X_t^V, $t \geq 0$ will be called the* **Gibbs stochastic process** *associated with $V : [0,1] \to \mathbb{R}$ (where the a priori P on $\mathscr{S}$ was given via the infinitesimal generator L which was fixed).*

Definition 3.9. *Given L and $V : [0,1] \to \mathbb{R}$, the associated probability $\mathbb{P}_V$ on the space $\mathscr{D}$ obtained from the* **Gibbs Markov Process** *X_t^V, $t \geq 0$ (with infinitesimal generator $\mathscr{L}_V$ and the stationary probability π_V) will be called the* **Gibbs probability for the interaction** *V (and the a priori infinitesimal generator L). $\mathbb{P}_V$ is invariant for the shift $\Theta_s, s \geq 0$.*

In the case $V = 0$ (and the Ruelle operator is normalized) $\mathbb{P}_0$ is the probability $\mathbb{P} = \mathbb{P}_L$ of Definition 2.1.

From [7],

$$\lim_{t\to\infty} \frac{1}{t} \log \sup_x \mathbb{E}_x \, [e^{\int_0^t V(w(s))\, ds}] = \lambda_V .$$

4 Relative Entropy, Pressure and the equilibrium state for V

The results of this section in some sense are similar to the ones in [20].

Given the inifinitesimal generators L_1 and L_2 (of the form (4)) consider the corresponding stationary probabilities μ_1 and μ_2 on $[0,1]$.

We denote by $\mathbb{P}_{L_1}$ and $\mathbb{P}_{L_2}$ the corresponding associated Θ_t-invariant probabilities on the Skorokhod space, $t \geq 0$.

Definition 4.1. *The relative entropy (or Kullback-Leibler divergence) of $\mathbb{P}_{L_1}$ and $\mathbb{P}_{L_2}$ is the value*

$$\frac{1}{T} H_T(\mathbb{P}_{L_2} | \mathbb{P}_{L_1}) = -\frac{1}{T} \int_{\mathscr{D}} \log \left(\frac{d\mathbb{P}_{L_2}}{d\mathbb{P}_{L_1}} \Big|_{\mathscr{F}_T} \right) (\omega) \, d\mathbb{P}_{L_1}(\omega) \tag{35}$$

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

$$\tilde{\mathscr{L}}(f)(x) = \int \left[f(y) - f(x) \right] \tfrac{\phi(y)}{\phi(x)} P(y,x)\, dy .$$

To rewrite the operator above we consider

$$\tilde{\gamma}(x) := \frac{1}{\phi(x)} \int \phi(y) P(y,x) dy \tag{36}$$

and $\tilde{Q}(y,x) := \frac{\phi(y)}{\phi(x)\tilde{\gamma}(x)} P(y,x)$. Then

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

The invariant probability for $\tilde{\mathscr{L}}$ is

$$\tilde{\mu}(dy) = \frac{\phi(y)\tilde{r}_\phi(y)}{\|\phi\|_2\|r_\phi\|_2}dy \tag{37}$$

where $\tilde{r}_\phi$ satisfies

$$\frac{1}{\tilde{r}_\phi(x)}\int P(x,z)\tilde{r}_\phi(z)dz = \tilde{\gamma}(x) = \frac{1}{\phi(x)}\int \phi(y)P(y,x)dy.$$

The probability $\tilde{\mathbb{P}}_{\tilde{\mu}}$ on $\mathscr{D}$ is called admissible, if it is induced by the continuous time Markov chain with infinitesimal generator $\tilde{\mathscr{L}}$ and the initial measure $\tilde{\mu}$. We point out that $\tilde{\mu}$ is invariant for this chain.

Given a Lipschitz function $V:[0,1]\to\mathbb{R}$, note that $\mathbb{P}^V_{\mu_V}$ is induced by the continuous time Markov chain with infinitesimal generator $\mathscr{L}^V$ and invariant probability μ_V is admissible.

Define

$$\frac{1}{T}H_T(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) = -\frac{1}{T}\int_{\mathscr{D}}\log\left(\frac{d\tilde{\mathbb{P}}_{\tilde{\mu}}}{d\mathbb{P}_{\tilde{\mu}}}\Big|_{\mathscr{F}_T}\right)(\omega)\,d\tilde{\mathbb{P}}_{\tilde{\mu}}(\omega), \tag{38}$$

for $\tilde{\mathbb{P}}_{\tilde{\mu}}$ admissible and $\mathbb{P}_{\tilde{\mu}}$ the probability on $\mathscr{D}$ induced by the continuous time Markov chain with infinitesimal generator L, defined in (4), and initial probability $\tilde{\mu}$.

It is possible to compute

$$\log\left(\frac{d\tilde{\mathbb{P}}_{\tilde{\mu}}}{d\mathbb{P}_{\tilde{\mu}}}\Big|_{\mathscr{F}_T}\right)(\omega) = \int_0^T[1-\tilde{\gamma}(\omega_s)]\,ds + \big[\log\big(\phi(\omega_T)\big)-\log\big(\phi(\omega_0)\big)\big].$$

Then

$$H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) = \int[\tilde{\gamma}(x)-1]\,d\tilde{\mu}(x). \tag{39}$$

Note that

$$H(\mathbb{P}^V_{\mu_V}|\mathbb{P}_{\mu_V}) = \lambda_V - \int V(x)\,d\mu_V(x).$$

We denote the Pressure (or, Free Energy) of V as the value

$$\mathbf{P}(V) := \sup_{\substack{\tilde{\mathbb{P}}_{\tilde{\mu}}\\ \text{admissible}}}\left\{H(\tilde{\mathbb{P}}_{\tilde{\mu}}|\mathbb{P}_{\tilde{\mu}}) + \int V(x)\,d\tilde{\mu}(x)\right\}. \tag{40}$$

Using (39), the pressure of V, $\mathbf{P}(V)$, is equal to

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

Recalling the definition of $\tilde{\gamma}$, in (36), and $\tilde{\mu}$, in (37), we have

$$\sup_{\phi>0}\int(L+V)\left(\tfrac{\phi}{\|\phi\|_2}\right)(x)\tfrac{\tilde{r}_\phi}{\|\tilde{r}_\phi\|_2}(x)dx = \lambda. \tag{41}$$

Indeed, first we can assume $\|\tilde{r}_\phi\|_2 = 1$.

We can also assume ϕ is such that $\int\frac{\phi(x)}{|\phi|^2}\tilde{r}_\phi(x)dx = 1$.

5 Time-Reversal Process and entropy production

Related results can be find in [37], [14], [34], [26], [28] and [27].

Before we begin the study of duality on the Skorokhod space we will state results for the detailed balance condition when the continuous time Markov Chain takes values on $\{1,2..,k\}$. Denote by $\sigma=(\sigma_1,...,\sigma_k)$ the initial invariant probability for the line sum zero matrix $W=(W_{i,j})_{i,j=1,..,k}$.

The detailed balance condition for W is: for all $i,j=1,...,k$

$$\sigma_i W_{i,j}=\sigma_j W_{j,i}.$$

Consider the inner product

$$< x,y >_\sigma = \sum_{j=1}^{k} \sigma_j x_j y_j .$$

It is easy to see that W satisfies the detailed balance condition, if and only if, W is self-adjoint for the inner product $< .,. >_\sigma$.

We assume in this section that L is of the form (4).

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. As mentioned before, we have that μ is somehow invariant with respect to L^*:

$$\int Lf(x)\mu(dx)=\iint [f(y)-f(x)]\,P(y,x)dy\ \theta(x)dx$$

$$=\int f(y)\int P(y,x)\theta(x)dx\ dy-\int f(x)\int P(y,x)dy\ \theta(x)dx$$

$$=\int f(y)\theta(y)dy-\int f(x)\theta(x)dx=0.$$

More precisely, $L^*(\theta)=0$, where L^* acts on $\mathscr{L}^2(dx)$. We will consider the dual process associated with μ on what follows. The substantial change 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

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

The dual operator for L (using $< f,g >_\mu$), will be denoted by $\mathfrak{L}^*:\mathscr{L}^2(\mu)\to\mathscr{L}^2(\mu)$. One can show that

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

To verify this, notice that

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

$$=\iint f(y)P(y,x)g(x)dy\ \theta(x)dx-\int f(x)g(x)\ \theta(x)dx$$

$$=\iint g(x)P(y,x)\theta(x)dx\ f(y)dy-\int f(x)g(x)\ \theta(x)dx$$

$$= \iint g(y)P(x,y)\theta(y)dy\ f(x)dx - \int f(x)g(x)\ \theta(x)dx$$
$$= \iint g(y)P(x,y)\frac{\theta(y)}{\theta(x)}dy\ f(x)\theta(x)dx - \int g(x)\ f(x)\theta(x)dx.$$
$$= \int \left[\int g(y)P(x,y)\frac{\theta(y)}{\theta(x)}dy - g(x) \right]\ f(x)\theta(x)dx$$

From (3) we have $\int P(x,y)\theta(y)dy = \theta(x)$. Thus, $\int P(x,y)\frac{\theta(y)}{\theta(x)}dy = 1$ and

$$< Lf, g >_\mu = \iint [g(y) - g(x)]P(x,y)\frac{\theta(y)}{\theta(x)}dy\ f(x)\theta(x)dx$$
$$= \iint [g(y) - g(x)]P^*(y,x)dy\ f(x)\theta(x)dx,$$

where $P^*(y,x) = P(x,y)\frac{\theta(y)}{\theta(x)}$. To fix ideas, we write again the expression for $\mathfrak{L}^*$:

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

Notice that we write L^* for the dual over $\mathscr{L}^2(dx)$ and $\mathfrak{L}^*$ for the one over $\mathscr{L}^2(\mu)$.

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

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

It has transition family $\hat{P}_t$ satisfying

$$\int g(x)(\hat{P}_t f(x))\ d\mu(x) := \mathbb{E}_\mu[g(X_T)f(X_{T-t})],\ \forall f, g \in \mathscr{L}^2(\mu).$$

This object is not at all new. In fact, notice that

$$\mathbb{E}_\mu[g(X_T)f(X_{T-t})] = \mathbb{E}_\mu[f(X_{T-t})\,\mathbb{E}_\mu[g(X_T)|\mathscr{F}_{T-t}]\,]$$
$$= \mathbb{E}_\mu[f(X_{T-t})\,\mathbb{E}_{X_{T-t}}[g(X_t)]\,] = \mathbb{E}_\mu[f(X_0)\,\mathbb{E}_{X_0}[g(X_t)]\,]$$
$$= \int f(x)\,\mathbb{E}_x[g(X_t)]\,d\mu(x)$$
$$= \int f(x)\,P_t(g(x))\,d\mu(x).$$

Since the last is true for all $f, g \in \mathscr{L}^2(\mu)$, we already get that $\hat{P}_t = P_t^*$, the transition family of $\mathfrak{L}^* : \mathscr{L}^2(\mu) \to \mathscr{L}^2(\mu)$. This also means that $\hat{L}f(x) = \mathfrak{L}^* f(x)$, $dx - a.s.$, where $\hat{L}$ is the infinitesimal generator of the semigroup $\hat{P}_t$.

For a fixed $T > 0$, we are interested in the quantity $\frac{1}{T}H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu)$ where $H_T(.|.)$ is the relative entropy between two probability measures over the space of trajectories $\mathscr{D}(S,[0,T])$. Recall the definition

$$H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) := \int_{\mathscr{D}} \log \frac{d\mathbb{P}_\mu}{d\hat{\mathbb{P}}_\mu}\bigg|_{\mathscr{F}_T} d\mathbb{P}_\mu.$$

The general formula of the above Radon-Nykodin derivative for a càdlàg process with general state space S is given by

$$\left.\frac{d\mathbb{P}_\mu}{d\hat{\mathbb{P}}_\mu}\right|_{\mathscr{F}_T} = \exp\left\{\int_0^T [\hat{\lambda}(X_s)-\lambda(X_s)]ds + \sum_{s\leq T}\log\left(\frac{\lambda(X_{s^-})}{\hat{\lambda}(X_{s^-})}\frac{dP}{d\hat{P}}(X_{s^-},X_s)\right)\right\} \tag{42}$$

Above, for each fixed $x \in S$, $\frac{dP}{d\hat{P}}(x,y)$ is the Radon-Nykodin derivative of $P(x,dy)$ with respect to $\hat{P}(x,dy)$. The summation over $s \leq T$ stands for all jumps until time T. Here we are omitting technicalities that guarantee the existence of this derivative

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

$$\hat{P}(x,dy) = P^*(x,dy) = P^*(y,x)dy = P(x,y)\frac{\theta(y)}{\theta(x)}dy$$

$$\Rightarrow \frac{dP}{d\hat{P}}(X_{s^-},X_s) = \frac{P(X_s,X_{s^-})\theta(X_{s^-})}{P(X_{s^-},X_s)\theta(X_s)}.$$

After these simplifications, we get

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

$$= \mathbb{E}_\mu\left[\sum_{s\leq T}\log\left(\frac{P(X_s,X_{s^-})}{P(X_{s^-},X_s)}\right) + \log(\theta(X_{s^-})) - \log(\theta(X_s))\right].$$

Notice the telescopic summation

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

since μ is invariant. Therefore, it is natural to consider the expression

$$\mathbb{E}_\mu\left[\sum_{s\leq T}\log\left(\frac{P(X_s,X_{s^-})}{P(X_{s^-},X_s)}\right)\right].$$

For this one, we use the underlying structure of the Markov chain given by P. Invoking the jump times T_n and the skeleton ξ_n of the discrete time Markov chain, we write

$$\sum_{n=0}^{\infty}\mathbb{E}_\mu\left[\sum_{s\leq T}\log\left(\frac{P(X_s,X_{s^-})}{P(X_{s^-},X_s)}\right)\mathbf{1}_{[T_n\leq T\leq T_{n+1}]}\right]$$

If until T there are no jumps, then we will simply get 0 from the expression above. This means that the summation could start at the first jump, $n = 1$.

$$= \sum_{n=1}^{\infty}\mathbb{E}_\mu\left[\sum_{k=0}^{n-1}\log\left(\frac{P(\xi_{k+1},\xi_k)}{P(\xi_k,\xi_{k+1})}\right)\mathbf{1}_{[T_n\leq T\leq T_{n+1}]}\right]$$

To simplify the calculations below, we denote for every $n \geq 1$, $\varphi(x_0,...,x_n) := \sum_{k=0}^{n-1}\log\frac{P(x_{k+1},x_k)}{P(x_k,x_{k+1})}$. Then

$$H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) = \sum_{n=1}^{\infty} \int d\mu(x_0) \int P(x_1,x_0)dx_1 \int \dots \int P(x_n,x_{n-1})\, dx_n\ \varphi(x_0,\dots,x_n)$$

$$\times \int_0^\infty ds_0\, e^{-s_0} \dots \int_0^\infty ds_n\, e^{-s_n}\ \mathbf{1}_{[0\le T-\sum_{i=0}^{n-1} s_i \le s_n]}. \tag{43}$$

Above, the first line covers spatial integrals that involve P. The second line integrals are independent of P and are easier to compute now. Notice that

$$\int_0^\infty ds_n e^{-s_n}\ \mathbf{1}_{[0\ \le T-\sum_{i=0}^{n-1} s_i\ \le\ s_n]} = \int_{T-\sum_{i=0}^{n-1} s_i}^\infty ds_n e^{-s_n}\ \mathbf{1}_{[0\ \le T-\sum_{i=0}^{n-1} s_i]}$$
$$= e^{-(T-\sum_{i=0}^{n-1} s_i)}\, \mathbf{1}_{[0\le T-\sum_{i=0}^{n-1} s_i]} = e^{-T}\ e^{\sum_{i=0}^{n-1} s_i}\, \mathbf{1}_{[0\le T-\sum_{i=0}^{n-1} s_i]}.$$

The second line in (43) becomes

$$e^{-T} \int_0^\infty \dots \int_0^\infty ds_0 \dots ds_{n-1}\ \mathbf{1}_{[\sum_{i=0}^{n-1} s_i \le T]},$$

where 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. This means those integrals sum up to $\frac{1}{2^n}\frac{2^n}{n!}T^n$. Thus, the entire second line of (43) is equal to $e^{-T}\ \frac{T^n}{n!}$.

The first line in (43) is

$$\sum_{n=1}^{\infty}\sum_{k=0}^{n-1} \int d\mu(x_0) \int P(x_1,x_0)dx_1 \int \dots \int P(x_n,x_{n-1})\, dx_n\ \log\frac{P(x_{k+1},x_k)}{P(x_k,x_{k+1})}. \tag{44}$$

To illustrate what happens in general, we will analyze the term of $k=0$ of the above sum, for general n:

$$\int d\mu(x_0) \int P(x_1,x_0) \log\frac{P(x_1,x_0)}{P(x_0,x_1)}dx_1 \left[\int P(x_2,x_1)\, dx_2 \dots \int P(x_n,x_{n-1})\, dx_n\right]$$

The integrals inside [ ] are all equal to 1. This happens in general for the integrals where the variable of integration is x_i for $k+1 < i \le n$. We can rewrite the general term of (44) as

$$\int d\mu(x_0) \int P(x_1,x_0)dx_1 \int \dots \int P(x_k,x_{k-1})\, dx_k \int \log\frac{P(x_{k+1},x_k)}{P(x_k,x_{k+1})} P(x_{k+1},x_k)\, dx_{k+1}.$$

Now we will handle those integrals of variables x_i with $i<k$. This is allowed since there are no functions that depend on $x_0,\dots,x_{k-1}$ on the right. Notice that

$$\int d\mu(x_0) \int P(x_1,x_0)dx_1 \int P(x_2,x_1)dx_2 \int \dots \int P(x_k,x_{k-1})\, dx_k$$
$$= \iint \theta(x_0)P(x_1,x_0)dx_0\ dx_1 \int P(x_2,x_1)dx_2 \int \dots \int P(x_k,x_{k-1})\, dx_k$$
$$= \int \theta(x_1)dx_1 \int P(x_2,x_1)dx_2 \dots \int P(x_k,x_{k-1})\, dx_k$$

$$= \int d\mu(x_1) \int P(x_2,x_1) dx_2 \int \ldots \int P(x_k,x_{k-1})\, dx_k$$

$$\ldots$$

$$= \int d\mu(x_{k-1}) \int P(x_k,x_{k-1})\, dx_k$$

$$= \iint \theta(x_{k-1}) P(x_k,x_{k-1})\, dx_{k-1}\, dx_k$$

$$= \int \theta(x_k)\, dx_k.$$

After all of this, (44) becomes

$$\int \theta(x_k) dx_k \int \log \frac{P(x_{k+1},x_k)}{P(x_k,x_{k+1})} P(x_{k+1},x_k)\, dx_{k+1}$$

$$= \int \theta(x) dx \int \log \frac{P(y,x)}{P(x,y)} P(y,x)\, dy$$

$$= \iint \log \frac{P(y,x)}{P(x,y)}\; P(y,x)\; dy\; d\mu(x).$$

Therefore,

$$H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) = \sum_{n=1}^{\infty} e^{-T}\; \frac{T^n}{n!}\; \sum_{k=0}^{n-1} \iint \log \frac{P(y,x)}{P(x,y)}\; P(y,x)\; dy\; d\mu(x)$$

$$= \sum_{n=1}^{\infty} e^{-T}\; \frac{T^n}{(n-1)!}\; \iint \log \frac{P(y,x)}{P(x,y)}\; P(y,x)\; dy\; d\mu(x)$$

$$= T\; \iint \log \frac{P(y,x)}{P(x,y)}\; P(y,x)\; dy\; d\mu(x)$$

and finally, the entropy production rate is

$$ep = \lim_{T\to\infty} \frac{1}{T}\; H_T(\mathbb{P}_\mu|\hat{\mathbb{P}}_\mu) = \iint \log \frac{P(y,x)}{P(x,y)}\; P(y,x)\; dy\; d\mu(x). \tag{45}$$

Notice that $P(x,y) = P(y,x) \Rightarrow ep = 0$.

Proposition 5.1. *In the above conditions, $ep \geq 0$, for all transition functions $P(x,y) > 0$.*

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

$$\iint [\log(\theta(x)) - \log(\theta(y))]\, P(y,x)\, dy\; d\mu(x) = 0. \tag{46}$$

Therefore, we can include this term into the entropy production rate of (45) as

$$ep = \iint \log \left[\frac{P(y,x)\theta(x)}{P(x,y)\theta(y)} \right]\; P(y,x)\; dy\; d\mu(x)$$

$$= \iint \left[\frac{P(y,x)\theta(x)}{P(x,y)\theta(y)} \right] \log \left[\frac{P(y,x)\theta(x)}{P(x,y)\theta(y)} \right]\; P(x,y) \frac{\theta(y)}{\theta(x)}\; dy\; d\mu(x)$$

Since $\iint P(x,y)\frac{\theta(y)}{\theta(x)}\;dy\;d\mu(x)=\int 1\,d\mu(x)=1$, we can use this as a probability measure (in fact it is $P^*(y,x)\;dy\;d\mu(x)$) in order to apply the Jensen inequality for the convex function $\psi(x)=\;x\log x$ on $\mathbb{R}_+$. In this way,

$$\psi\left(\int\left[\frac{P(y,x)\theta(x)}{P(x,y)\theta(y)}\right]P(x,y)\frac{\theta(y)}{\theta(x)}\;dy\;d\mu(x)\right)$$
$$\leq\int\psi\left[\frac{P(y,x)\theta(x)}{P(x,y)\theta(y)}\right]P(x,y)\frac{\theta(y)}{\theta(x)}\;dy\;d\mu(x)=ep.$$

Finally,

$$ep\geq\psi\left(\int P(y,x)\;dy\;d\mu(x)\right)=\psi(1)=0.$$

The idea here was similar to the one in Lemma 3.3 in [31]. □

Proposition 5.2. $ep^*=\lim_{T\to\infty}\frac{1}{T}H_T(\hat{\mathbb{P}}_\mu|\mathbb{P}_\mu)=ep.$

Proof. Recall that $P^*(y,x)=P(x,y)\theta(y)/\theta(x)$. The calculation in (46) allows us to add a term into ep:

$$ep=\iint\log\frac{P(y,x)}{P(x,y)}\;P(y,x)\;dy\;d\mu(x)+\iint[\log(\theta(x)-\log(\theta(y))]\,P(y,x)\,dy\;d\mu(x)$$
$$=\iint\log\frac{P(y,x)\theta(x)}{P(x,y)\theta(y)}\,P(y,x)\,dy\;d\mu(x)$$
$$=\iint\log\frac{P(y,x)}{P^*(y,x)}\,P(y,x)\,dy\;d\mu(x).$$

In the formula above, we can recognize the transition functions, $P(y,x)$ and $P^*(y,x)$, associated to the processes $\mathbb{P}_\mu$ and $\hat{\mathbb{P}}_\mu$, respectively. Now, to proceed the change to $H_T(\hat{\mathbb{P}}_\mu,\mathbb{P}_\mu)$, we change the role of them:

$$ep^*=\iint\log\frac{P^*(y,x)}{P(y,x)}\,P^*(y,x)\,dy\;d\mu(x)=\iint\log\frac{P(x,y)\theta(y)}{P(y,x)\theta(x)}\,P^*(y,x)\,dy\;d\mu(x)$$
$$=\iint\log\frac{P(x,y)}{P(y,x)}\,P^*(y,x)\,dy\;d\mu(x)+\iint[\log\theta(y)-\log\theta(x)]\,P^*(y,x)\,dy\;d\mu(x)$$
$$=\iint\log\frac{P(x,y)}{P(y,x)}\,P(x,y)\frac{\theta(y)}{\theta(x)}\,dy\;d\mu(x)$$
$$=\iint\log\frac{P(x,y)}{P(y,x)}\,P(x,y)\;dx\;d\mu(y)=ep.$$

□

Remmark 5.3. *In case one wants to symmetrize the process generated by L by taking the average* $\frac{L+\mathfrak{L}^*}{2}$*, the following would apply. Consider the transition function for this operator*

$$Q(y,x)=\frac{P(y,x)}{2}+\frac{P(x,y)\theta(y)}{2\theta(x)}=\frac{P(y,x)\theta(x)+P(x,y)\theta(y)}{2\theta(x)}.$$

Notice that

$$\int Q(y,x)dy = \frac{1}{2}\int P(y,x)dy + \frac{1}{2\theta(x)}\int P(x,y)\theta(y)dy = 1$$

and

$$Q(x,y) = Q(y,x)\frac{\theta(x)}{\theta(y)}.$$

We arrive at an equation that can be understood as a balance condition:

$$Q(x,y)\theta(y) = Q(y,x)\theta(x) \tag{47}$$

We already knew that $\frac{L+\mathfrak{L}^*}{2}$ is symmetric by construction, but the above is sufficient to conclude the symmetry for any operator.

Proposition 5.4. *Every operator $\mathscr{A}$ which has a transition function Q that satisfies the balance condition (47) is symmetric in $\mathscr{L}^2(\mu)$.*

Proof. Indeed,

$$\begin{aligned}
&\int (\mathscr{A} f)(x)\, g(x)\, d\mu(x) \\
&= \iint [f(y) - f(x)]Q(y,x)\, dy\; g(x)\theta(x)dx \\
&= \iint [f(y) - f(x)]\, g(x)\, Q(y,x)\theta(x)dy\, dx \\
&= \iint [f(y) - f(x)]\, g(x)\, Q(x,y)\theta(y)dy\, dx \\
&= \iint f(y)g(x)\; Q(x,y)\theta(y)\, dy\, dx - \int f(x)g(x)\theta(x) \int Q(y,x)\, dy\, dx \\
&= \iint f(y)g(x)\; Q(x,y)\theta(y)\, dy\, dx - \int f(x)g(x)\theta(x)\, dx \\
&= \iint f(y)g(x)\; Q(x,y)\theta(y)\, dy\, dx - \int f(y)g(y)\theta(y)\, dy \\
&= \iint f(y)[g(x) - g(y)]\; Q(x,y)\theta(y)\, dy\, dx \\
&= \int f(y) \int [g(x) - g(y)]\; Q(x,y)dx\; \theta(y)\, dy \\
&\int f(y)\, (\mathscr{A} g)(y)\, d\mu(y).
\end{aligned}$$

□

6 Expansiveness of the semi-flow Θ_t, $t \geq 0$, on $\mathscr{D}$

From now on we assume in this section that L is of the form (4).

In this section, we consider the Skorokhod space $\hat{\mathscr{D}}$ of paths $w : (-\infty, \infty) \to \mathbb{R}$ continuous at right and with limit at left (also called *càdlàg*).

$\hat{\Theta}_t$, $t \in \mathbb{R}$, denotes the bidirectional flow on $\hat{\mathscr{D}}$, acting on w by translation to the left on time t. That is, for fixed t, then $\tilde{w} = \hat{\Theta}_t(w)$ is such that $\tilde{w}(s) = w(s+t)$.

Let $\mathscr{D}$, the set of path $w_2 : [0, \infty) \to \mathbb{R}$ continuous at right and with limit at left.

Θ_t, $t \geq 0$, denotes the shift on $\mathscr{D}$, acting on w_2 by translation to the left on time t. That is, for fixed t, then $\tilde{w}_2 = \Theta_t(w_2)$ is such that $\tilde{w}_2(s) = w_2(s+t)$.

Let $\mathscr{D}^*$, the set of path $w_1 : (0, \infty) \to \mathbb{R}$ continuous at left and with limit at right.

It is necessary to make more clear our notation: by $w =< w_1 | w_2 >= (w_1, w_2)$ we mean a path $w : \mathbb{R} \to \mathscr{D}$ such that $w(t) = w_2(t)$ for $t \geq 0$ and $w(t) = w_1(-t)$ for $t < 0$. In our notation $w_1 : [0, \infty) \to \mathbb{R}$ is continuous at left and with limit at right, and $w_2 : (0, \infty) \to \mathbb{R}$ is continuous at right and with limit at left.

Using the above notation we can write $\hat{\mathscr{D}}$, the Skorokhod space of *càdlàg* paths $w : (-\infty, \infty) \to \mathbb{R}$, as $\hat{\mathscr{D}} = \mathscr{D}^* \times \mathscr{D}$. A typical path in $\hat{\mathscr{D}}$ will be written in the form $w =< w_1 | w_2 >= (w_1, w_2) \in \mathscr{D}^* \times \mathscr{D}$. By convention w_1 will be at left of $t = 0$ and w_2 at right of $t = 0$.

Denote, for $s \geq 0$,

$$(w_1|_t w_2)(s) = \begin{cases} w_1(t-s), & s < t \\ w_2(s-t), & s \geq t \end{cases} .$$

We denote by $\Pi_1 : \mathscr{D}^* \times \mathscr{D} \to \mathscr{D}^*$ the projection $\Pi_1(w) = \Pi_1(< w_1 | w_2 >) = w_1$. If we denote by $\Pi_2 : \mathscr{D}^* \times \mathscr{D} \to \mathscr{D}$ the projection $\Pi_2(w) = \Pi_2(< w_1 | w_2 >) = w_2$, then $(w_1|_t w_2) = \Pi_2(\hat{\Theta}_{-t}(w_1, w_2))$.

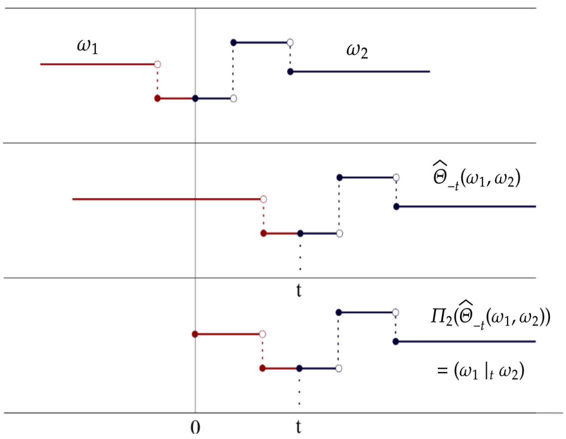


Figure 2: The bilateral shift and the projection Π_2

We will show that the semi-flow Θ_t, $t \geq 0$, is *expanding* (see (49)).

Proposition 6.1. *The continuous time shift Θ_t, $t \geq 0$, acting on the Skorokhod space $\mathscr{D}$ equipped with the Skorokhod metric is expanding: given paths w_1, w_2 and t*

$$d\left((w_1|_t w_2), (w_1|_t w_2')\right) \leq \int_t^\infty e^{-u} du = e^{-t}. \tag{48}$$

Proof. Notice that $|A(w_1|_t w_2) - A(w_1|_t w_2')| \leq C_A d\left((w_1|_t w_2),(w_1|_t w_2')\right)$, where $d(x,y)$ denotes the Skorokhod distance on $\mathscr{D}$. This is the distance between two paths that coincides until time t. Recall the definition of the Skorokhod distance

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

Above, the Λ set is for the continuous functions such that the below function γ is finite.

So, if we choose λ as the identity function, we get

$$\gamma(\lambda) := \sup_{t\geq 0} \text{ ess } |\log \lambda'(t)| = 0.$$

Then

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

where $q = r \wedge 1$ with r denoting the metric on the state space, i.e., Lebesgue in $[0,1]$. For $u < t$, the distance q above is $q(w_1(s\wedge u), w_1(s\wedge u)) = 0$. Furthermore, the distance q is upper bounded by 1. Then,

$$d\left((w_1|_t w_2),(w_1|_t w_2')\right) \leq \int_t^{\infty} e^{-u}du = e^{-t}. \tag{49}$$

□

Proposition 6.2. *For a fix Lipschitz function $A : \mathscr{D} \to \mathbb{R}$ and a path $w_2' \in \mathscr{D}$ denote $W_{t_1}^0 = W_{A,t_1,w_2'}^0 : \mathscr{D}^* \times \mathscr{D} \to \mathbb{R}$ the function given by*

$$W^0(w_1,w_2)_{A,t_1,w_2'} = W_{t_1}^0(w_1,w_2) := \sum_{n=1}^{\infty} A(w_1|_{nt_1} w_2) - A(w_1|_{nt_1} w_2'). \tag{50}$$

Then, W^0 is well defined.

Proof. As consequence of (49), we have

$$|W_{t_1}^0(w_1,w_2)| \leq \sum_{n=1}^{\infty} |A(w_1|_{nt_1} w_2) - A(w_1|_{nt_1} w_2')| \leq C_A \sum_{n=1}^{\infty} e^{-nt_1} < \infty,$$

for $t_1 > 0$. We conclude that $W_{t_1}^0$ given by the expression (50) is well defined.

□

7 Appendix 1 - Existence of $K_t(x,y)$

In this section, we will show explicitly the existence of a function $K_t(x,y)$ which has the following relation with the e^{tL} operator: for every function $f : S \to \mathbb{R}$, it satisfies

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

Remember that $L = \mathscr{L} - I$, and $\mathscr{L}$ is acting on functions as $\mathscr{L}(f)(y) = \int f(x)P(x,y)dx$. 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 $\mathscr{L}$ and $-I$ commute:

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

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

$$P^k(x,y) := \int \cdots \int P(x, z_{k-1}) \cdots P(z_2, z_1) P(z_1, y) \, dz_1 dz_2 ... dz_{k-1}.$$

For example, $P^2(x,y) = \int P(x,z)P(z,y)dz$ and $P^{k+1}(x,y) = \int P(x,z)P^k(z,y)dz$. Of course, $P^1(x,y) = P(x,y)$. Now, we state that

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

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

$$\mathscr{L}^{k+1}(f)(y) = \mathscr{L}^k(\mathscr{L} f)(y) = \int \mathscr{L} f(x) P^k(x,y) dx$$

$$= \int \int f(z) P(z,x) dz P^k(x,y) dx = \int f(z) \int P(z,x) P^k(x,y) dx \, dz$$

$$= \int f(z) P^{k+1}(z,y) dz.$$

Above, to changing 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. We already have a kernel for L^k:

$$L^k(f)(y) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \mathscr{L}^j(f)(y)$$

$$= (-1)^k f(y) + \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \int f(x) P^j(x,y) dx$$

$$= (-1)^k f(y) + \int f(x) \left[\sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} P^j(x,y) \right] dx$$

$$= (-1)^k f(y) + \int f(x) \; Q_k(x,y) dx,$$

where $Q_k(x,y) = \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} P^j(x,y)$, for $k \geq 1$. Notice that

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

is our desired function, because

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

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

7.1 Properties of $K_t(x,y)$

In this section, we will deduce some properties of $K_t(x,y)$. Remember that $L(f)(y) = \int (f(x) - f(y))P(x,y)dx = \int f(x)P(x,y)dx - f(y)$. Since $P_t = e^{tL}$ is the homogeneous semigroup generated by the infinitesimal generator L we have that

$$\frac{d}{dt} P_t(f) = L(P_t f) = P_t(Lf).$$

So, we calculate

$$\frac{d}{dt} P_t f(y) = \frac{d}{dt} \left(e^{-t} f(y) + \int f(x) K_t(x,y) dx \right) = -e^{-t} f(y) + \int f(x) \frac{d}{dt} K_t(x,y) dx$$

and

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

So,

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

This last equation is valid for every $f : S \to \mathbb{R}$, which means that

$$\frac{d}{dt}K_t(x,y) = -K_t(x,y) + e^{-t}P(x,y) + \int K_t(x,z)P(z,y)dz. \tag{51}$$

The above is equal to

$$\frac{d}{dt}K_t(x,y) = L(K_t(x,\cdot))(y) + e^{-t}P(x,y) \tag{52}$$

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

$$\frac{d}{dt}K_t(x,y) = L^*(K_t(\cdot,y))(x) + e^{-t}P(x,y). \tag{53}$$

Just to exemplify, if $P(x,y) = q(x)$ were independent of y, we would have that $K_t(x,y) = K_t(x)$ also do not depend on y and $\frac{d}{dt}K_t(x) = e^{-t}q(x)$, which means $K_t(x) = (1-e^{-t})q(x)$ and $e^{tL}f(y) = e^{-t}f(y) + (1-e^{-t})\int f(x)q(x)dx$.

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

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

while

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

This means

$$K_{s+t}(x,y) = e^{-t}K_s(x,y) + e^{-s}K_t(x,y) + \int K_s(x,z)K_t(z,y)dz \tag{54}$$

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

8 Appendix 2 - Existence of K_t^V

In this section, we will show explicitly the existence of a function $K_t^V : S \times S \to \mathbb{R}^+$ which has the following relation with the $e^{t(L+V)}$ operator: for every function $f : S \to \mathbb{R}$, it satisfies

$$e^{t(L+V)} f(y) = \int f(x) K_t^V(x,y) dx + e^{t(V(y)-\lambda(y))} f(y).$$

Here we are considering with more generality an operator L acting on functions as

$$L(f)(x) = \lambda(x) \int (f(y) - f(x)) P(y, dx).$$

Using the Feynman-Kac formula (see Corollary 11 and further remark in A-G notes) for a bounded function $V : S \to \mathbb{R}$ we get

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

We will analyze this expression in terms of the graphic construction of the jump process (X_t).

This means that we will use that the trajectories are piece-wise constant.

$$\mathbb{E}_x \left[e^{\int_0^t V(X_r) dr} f(X_t) \right] = \mathbb{E}_x \left[e^{\int_0^t V(X_r) dr} f(X_t) \, 1_{[T_0 \le t < T_1]} \right]$$

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

$$= e^{tV(x)} f(x) \, \mathbb{P}_x[\tau_0 > t] + \sum_{n=1}^{\infty} \mathbb{E}_x \left[e^{\int_0^t V(X_r) dr} f(X_t) 1_{[T_n \le t < T_{n+1}]} \right].$$

The first term is $e^{tV(x)} f(x) e^{-t\lambda(x)}$, since for $\mathbb{P}_x$, τ_0 is exponential with parameter $\lambda(x)$. In order to rewrite each term of the summation, we notice that in the case of $T_n \le t < T_{n+1}$ (the one in the above figure) we have

$$\int_0^t V(X_r) dr = \tau_0 V(x) + \tau_1 V(x_1) + ... + \tau_{n-1} V(x_{n-1}) + (t - T_n) V(x_n)$$

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

since $T_n = \tau_0 + ... + \tau_{n-1}$ and with the convention $x_0 = x$. Now, define

$$\varphi_t^{n,V}(x, x_1, x_2, ..., x_n) = \mathbb{E}_x \left[exp \left[\sum_{i=0}^{n-1} \tau_i V(x_i) + \left(t - \sum_{i=0}^{n-1} \tau_i \right) V(x_n) \right] \cdot 1_{[T_n \le t < T_{n+1}]} \right].$$

Notice that for a fixed t, $\varphi_t^{n,V} \ge 0$ in general and for only one $n(t)$ is true that $\varphi_t^{n(t),V} > 0$. In this way, the general term of summation becomes

$$\mathbb{E}_x \left[e^{\int_0^t V(X_r) dr} f(X_t) 1_{[T_n \le t < T_{n+1}]} \right]$$

$$= \int \cdots \int \varphi_t^{n,V}(x, x_1, ..., x_n) f(x_n) \, P(x_{n-1}, dx_n) ... P(x_1, dx_2) P(x, dx_1).$$

Going back to the beginning, we have

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

$$+\sum_{n=1}^{\infty}\int\cdots\int \varphi_t^{n,V}(x,x_1,...,x_n) f(x_n)\; P(x_{n-1},dx_n)...P(x_1,dx_2)P(x,dx_1).$$

To conclude, we are going to explicitly write the above expression in the case of $P(x,dy)=P(y,x)\,dy$, where P is a transition probability for a Markov chain in S with $P(y,x)>0$, for every $x,y\in S$.

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

$$+\sum_{n=1}^{\infty}\int\cdots\int \varphi_t^{n,V}(x,x_1,...,x_n) f(x_n)\; P(x_n,x_{n-1})dx_n\ldots P(x_2,x_1)dx_2\, P(x_1,x)dx_1.$$

Defining $Q_t^{n,V}(x_n,x)=\int\cdots\int \varphi_t^{n,V}\, P(x_n,x_{n-1})\ldots P(x_1,x)\; dx_{n-1}\ldots dx_1$, we get

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

where $K_t^V(y,x)=\sum_{n=1}^{\infty} Q_t^{n,V}(y,x)$. The hypothesis of $P(x,y)>0$ implies that $Q_t^{n,V}(y,x)\geq 0$ and for the special one $n(t)$, $Q_t^{n(t),V}(y,x)>0$ holds. Then $K_t^V(y,x)>0$, for every $y,x\in S$. This ends the discussion.

8.1 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 $\lambda\equiv 1$. Since the semigroup $P_t^V=e^{t(L+V)}$ is homogeneous in time, we have $(L+V)(P_t^V)(f)=\frac{d}{dt}P_t^V(f)=P_t^V((L+V)f)$. The middle term opens as

$$\begin{aligned}
\frac{d}{dt}P_t^V(f)(y) &= \frac{d}{dt}\left[\int f(x) K_t^V(x,y) dx + e^{-t+tV(y)} f(y)\right] \\
&= \int f(x)\frac{d}{dt}K_t^V(x,y)dx + (V(y)-1)e^{-t+tV(y)} f(y)
\end{aligned}$$

while

$$P_t^V((L+V)f)(y) = \int (L+V)(f)(z)\; K_t^V(z,y)dz + e^{-t+tV(y)}(L+V)(f)(y)$$

$$=\int\int f(x)P(x,z)dx\; K_t^V(z,y)dz - \int f(z)\; K_t^V(z,y)dz + \int V(z) f(z)\; K_t^V(z,y)dz$$

$$+e^{-t+tV(y)}V(y)f(y) + e^{-t+tV(y)}\int f(x)P(x,y)dx - e^{-t+tV(y)} f(y)$$

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

$$+ e^{-t+tV(y)} f(y)(V(y)-1)$$

Now, for every f it holds

$$\begin{aligned} 0 &= -\frac{d}{dt} P_t^V(f) + P_t^V((L+V)f) \\ &= \int f(x) \Big[-\frac{d}{dt} K_t^V(x,y) + \int P(x,z) K_t^V(z,y) dz - K_t^V(x,y) \\ &\quad + V(x) K_t^V(x,y) + e^{-t+tV(y)} P(x,y) \Big] dx. \end{aligned}$$

Which means that

$$\frac{d}{dt} K_t^V(x,y) = \int P(x,z) K_t^V(z,y) dz - K_t^V(x,y) + V(x) K_t^V(x,y) + e^{-t+tV(y)} P(x,y)$$

$$= (L^* + V)(K_t^V(\cdot,y))(x) + e^{t(V(y)-1)} P(x,y). \tag{55}$$

The other one opens as

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

Which gives us

$$\frac{d}{dt} K_t^V(z,y) = \int K_t^V(z,x) P(x,y) dx + e^{-t+tV(z)} P(z,y) + K_t^V(z,y)(V(y)-1)$$

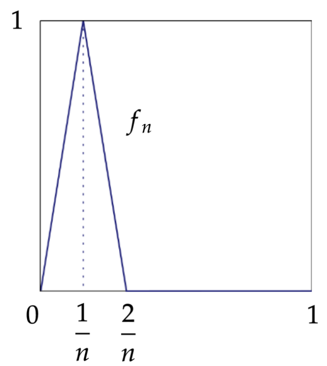
$$= (L+V)(K_t^V(z,\cdot))(y) + e^{t(V(z)-1)} P(z,y). \tag{56}$$

9 Appendix 3 - The need for Lipschitz or Hölder class on Theorem 1

We present an example of the operator $L + V + zI$ with domain the whole space of continuous functions. Without the restriction on the variation of the functions, we cannot assure that the operator is compact and therefore, we cannot use Krein-Rutman theorem.

Let $f_n : [0,1] \to \mathbb{R}$ be defined as

$$f_n(x) = \begin{cases} nx & 0 \le x < \frac{1}{n} \\ 2 - nx & \frac{1}{n} \le x < \frac{2}{n} \\ 0 & \frac{2}{n} \le x \le 1 \end{cases}$$



The key idea here is that the set $\{f_n\}$ is contained in the unitary ball of $C_0([0,1])$, but it does not admit a uniform bound M on the variation, i.e., $|f_n(y) - f_n(x)| \le M|y-x|$, for all n, x, y.

Take $P(x,y) \equiv 1$. Then

$$Lf_n(y) = \int f_n(x)dx - f_n(y) = \frac{1}{n} - f_n(y).$$

For $V(x) = (x-1/2)^2$ and $z = 1 + \|V\|_\infty$, we get

$$g_n(y) := (L+V+zI)f_n(y) = \frac{1}{n} + f_n(y)[V(y) + z - 1]$$

$$= \frac{1}{n} + f_n(y)(V(y) + \|V\|_\infty) = \frac{1}{n} + f_n(y)(y^2 - y + 1/2).$$

If $L+V+zI$ is compact, $\{g_n\}$ must admit a convergent subsequence, since $\{f_n\}$ is a bounded sequence. Without loss of generality we assume that such subsequence g_{n_k} is rewritten as g_n.

It is a Cauchy sequence: take $\varepsilon = \frac{1}{4}$. There must be a N s.t. $m, n \ge N \Rightarrow \|g_n - g_m\|_\infty < \frac{1}{4}$. Notice that $f_n(\frac{1}{n}) = 1$ and $f_n(x) = 0$, for $x \ge \frac{2}{n}$. For $m \ge 2n$, we have $f_m(\frac{1}{n}) = 0$, since $\frac{1}{n} = \frac{2}{2n} \ge \frac{2}{m}$. So,

$$\left| g_n\left(\frac{1}{n}\right) - g_m\left(\frac{1}{n}\right) \right| = \left| 1\left(\frac{1}{n^2} - \frac{1}{n} + \frac{1}{2}\right) + \frac{1}{n} - 0 - \frac{1}{m} \right|$$

$$= \left| \frac{1}{2} + \frac{1}{n^2} - \frac{1}{m} \right| > \frac{1}{2} - \frac{1}{m}.$$

The latter contradicts $\|g_n - g_m\|_\infty < \frac{1}{4}$ for $n \ge 2$.

10 Appendix 4 - Proofs of claims of Example 1

In this appendix we will show the proofs of the claims mentioned in Example 1.

First we will show that

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

Note that $P(x,y) = \cos[(x-y)2\pi]/2+1$ is symmetric and continuous on $[0,1]$. Also note that

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

$$\Rightarrow P^2(x,y) := \int P(x,z)P(z,y)\,dz = \frac{\cos(2\pi(x-y))}{8} + 1$$

We state: $P^n(x,y) = \frac{\cos(2\pi(x-y))}{2^{2n-1}} + 1$. We proceed by indution.

$$P^n(x,y) = \int P^{n-1}(x,z)P(z,y)dz$$

$$= \int \left(\frac{\cos(2\pi(x-z))}{2^{2n-3}} + 1\right)\left(\frac{\cos(2\pi(z-y))}{2} + 1\right) dz$$

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

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

We continue to find $K_t(x,y)$. We refer to the appendix 1 to the general case.

$$Q_k(x,y) := \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} P^j(x,y)$$

$$= \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \left(\frac{\cos(2\pi(x-y))}{2^{2j-1}} + 1\right)$$

$$= 2\cos(2\pi(x-y)) \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \frac{1}{2^{2j}} + \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j}$$

$$= 2\cos(2\pi(x-y)) \left[(-1)^{k+1} + \left(-\frac{3}{4}\right)^k\right] + (-1)^{k+1}$$

$$= (2\cos(2\pi(x-y)) + 1)(-1)^{k+1} + 2\cos(2\pi(x-y)) \left(-\frac{3}{4}\right)^k$$

Finally,

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

$$= (2\cos(2\pi(x-y)) + 1) \sum_{k=1}^{\infty} \frac{t^k}{k!} (-1)^{k+1} + 2\cos(2\pi(x-y)) \sum_{k=1}^{\infty} \frac{t^k}{k!} \left(-\frac{3}{4}\right)^k$$

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

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

This way,

$$e^{tL}f(y) = e^{-t}f(y) + \int f(x)K_t(x,y)dx$$

$$= e^{-t}f(y) + (1-e^{-t})\int f(x)dx + 2(e^{-3t/4} - e^{-t})\int f(x)\cos(2\pi(x-y))dx.$$

Since $P(x,y) = P(y,x)$, we have $L = L^*$ and the normalized eigenfunctions are precisely the eigendensities. By construction, $L^*(1) = L(1) = 0$, so 1 is an eigenfunction, for which we get dx, the Lebesgue probability is invariant. Given the simplicity of P we can go further and find that they are unique.

Continuous functions $f : \mathbb{S}^1 \to \mathbb{R}$ can be seen as periodic functions $f : \mathbb{R} \to \mathbb{R}$ with period 1, so that we can employ Fourier Series. Write

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

where $a_0 = 2\int f(x)dx$, $a_n = 2\int f(x)\cos(2\pi nx)dx$ and $b_n = 2\int f(x)\sin(2\pi nx)dx$. Notice $\cos(2\pi(x-y)) = \cos(2\pi x)\cos(2\pi y) + \sin(2\pi x)\sin(2\pi y)$. Then

$$Lf(y) = \int f(x)dx + \frac{1}{2}\int f(x)\cos(2\pi(x-y))dx - f(y)$$

$$= \frac{a_0}{2} + \frac{1}{2}\cos(2\pi y)\frac{a_1}{2} + \frac{1}{2}\sin(2\pi y)\frac{b_1}{2} - f(y).$$

Therefore, $Lf = 0$ if and only if

$$f(y) = \frac{a_0}{2} + \cos(2\pi y)\frac{a_1}{4} + \sin(2\pi y)\frac{b_1}{4}. \tag{57}$$

and from this follows $a_1 = a_1/4$, $b_1 = b_1/4$ and $a_n = b_n = 0$, $\forall n \geq 2$. Conclusion: $Lf = 0 \Rightarrow f \equiv \frac{a_0}{2}$, constant. This means that the only eigendensity of the operator e^{tL} is that of Lebesgue measure dx.

11 Appendix 5 - Another look of Feynman-Kac formula for symmetrical L

Consider X_t a continuous time process with state space S and infinitesimal generator L. Let f and V be two functions on S 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 $\hat{L}$ it's generator. For this process $\hat{X}$, we have that, by Feynman-Kac, 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 to 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

$$\int_{w(0)=x} e^{\int_0^t V(w(s))ds} f(w(t)) d\mathbb{P}(w) = \int_{w(T)=x} e^{\int_{T-t}^T V(w(s))ds} f(w(T-t)) d\mathbb{P}(w). \quad (58)$$

References

[1] P. Adamo, R. Belousov, and L. Rondoni, Fluctuation-Dissipation and Fluctuation Relations: From Equilibrium to Nonequilibrium and Back, *Large Deviations in Physics*, Lect. Notes in Physics, 92-133 Vol. 885, Springer Verlag (2014)

[2] 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).

[3] 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, *Rev. in Math Phys*, Vol. 23, N. 10, pp 1063-1113 (2011).

[4] D. Bakry, I. Gentil and M. Ledoux, Analysis and Geometry of Markov Diffusion Operators, Springer

[5] A. Bobrowski, Functional Analysis for Probability and Stochastic Processes. Cambridge Press (2005)

[6] 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 (2022)

[7] M. D. Donsker and S. R. S. Varadhan, On a Variational Formula for the Principal Eigenvalue for Operators with Maximum Principle. Proc. Nat. Acad, Scien, March 1, 1975 72 (3) 780–783

[8] K. Deimling. Nonlinear Functional Analysis, Springer Verlag, 1985

[9] S. N. Ethier and T. G. Kurtz, Markov processes - Characterization and Convergence, Wiley (2005)

[10] J. Feng and T. Kurtz, Large Deviations for Stochastic Processes, AMS

[11] G. Gallavotti and E. G. D. Cohen. Dynamical Ensembles in Nonequilibrium Statistical Mechanics, Phys. Rev. Lett. 74, 2694 (1995).

[12] G. Gallavotti, Fluctuation patterns and conditional reversibility in nonequilibrium systems Annales de l'I.H.P. Physique theorique (1999) Volume: 70, Issue: 4, page 429–443

[13] D. A. Gomes, A stochastic analogue of Aubry-Mather theory. Nonlinearity 15, 581–603, 2002

[14] Da-Quan Jiang Min Qian Min-Ping Qian, Mathematical Theory of Nonequilibrium Steady States, Lect. Notes in Math 1833 (2004)

[15] Da-quan Jiang, Min Qian and Min-ping Qian. Entropy Production and Information Gain in Axiom-A Systems. Commun. Math. Phys., 214, 389–409 (2000).

[16] M. Kac, Integration in some function spaces and some of its application, Acad Naz dei Lincei Scuola Superiore Normale Superiore, Piza, Italy (1980)

[17] Y. Kifer, Large Deviations in Dynamical Systems and Stochastic processes, TAMS, Vol 321, N.2, 505–524 (1990)

[18] Y. Kifer, Principal eigenvalues, topological pressure, and stochastic stabilty of equilibrium states, Israel Journal of Mathematics, Vol. 70, No. I, pp 1–47 (1990)

[19] T. Liggett, Continuous Time Markov Processes, AMS, (2010).

[20] A. O. Lopes, A. Neumann and Ph. Thieullen, A thermodynamic formalism for continuous time Markov chains with values on the Bernoulli Space: entropy, pressure and large deviations, Journ. of Statist. Phys. Volume 152, Issue 5, Page 894–933 (2013)

[21] A. O. Lopes and Ph. Thieullen, Transport and large deviations for Schrodinger operators and Mather measures, Modeling, Dynamics, Optimization and Bioeconomics III, Springer Verlag, 247-255 (2018)

[22] A. O. Lopes, J. Mohr, R. Souza and Ph. Thieullen, Negative Entropy, Zero temperature and stationary Markov chains on the interval, Bull. Soc. Bras. Math. Vol 40 n 1, (2009), 1-52

[23] A. O. Lopes, G. Muller and A. Neumann, Diffusion Processes: entropy, Gibbs states and the continuous time Ruelle operator, preprint

[24] A. O. Lopes and A. Neumann, Large Deviations for stationary probabilities of a family of continuous time Markov chains via Aubry-Mather theory, Journ. of Statistical Physics. Vol. 159 – Issue 4 pp 797-822 (2015)

[25] 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, *Erg. Theory and Dyn Systems*, 35 (6), 1925–1961 (2015)

[26] A. O. Lopes and J. K. Mengue, On information gain, Kullback-Leibler divergence, entropy production and the involution kernel, Disc. and Cont. Dyn. Syst. Series A. Vol. 42, No. 7, 3593–3627 (2022)

[27] C. Maes, K. Netocny, and B. Shergelashvili, A selection of nonequilibrium issues, Methods of Contemporary Mathematical Statistical Physics, 247-306 (2009)

[28] C. Maes and K. Netocny, Time-Reversal and Entropy, Journal of Statistical Physics, Vol. 110, Nos. 1/2, (2003)

[29] C. A. Morales, Bowen-Walters expansiveness for semigroups of linear operators, online - to appear in Ergodic Theory and Dynamical Systems.

[30] G. Muller and A. Neumann; Some general results for continuous-time Markov chain, preprint (2022).

[31] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* Vol 187-188 (1990)

[32] A. Politi, Stochastic Fluctuations in Deterministic Systems, *Large Deviations in Physics*, Lect. Notes in Physics, 243-261 Vol. 885, Springer Verlag (2014)

[33] W. Strock, An introduction the the theory of large deviations , Springer

[34] Da-Quan Jiang, Min Qian, Min-Ping Qian, Mathematical Theory of Nonequilibrium, Steady States, Lect notes in Math 1833

[35] D. Ruelle; Positivity of entropy production in nonequilibrium statistical mechanics, Journal of Statistical Physics, Vol. 85, 1996, 1996.3, pp. 1–25

[36] D. Ruelle; A generalized detailed balance relation. J. Stat. Phys. 164, no. 3, 463–471 (2016)

[37] Y. Wang and H. Qian, Mathematical Representation of Clausius' and Kelvin's Statements of the Second Law and Irreversibility, Jour. of Stat. Phys. (2020) 179:808–837

[38] G. Muller and A. Neumann; Some general results for continuous-time Markov chain, preprint (2022).