STOCHASTIC DERIVATIVES AND GENERALIZED h-TRANSFORMS OF MARKOV PROCESSES

CHRISTIAN LÉONARD

ABSTRACT. Let R be a continuous-time Markov process on the time interval $[0,1]$ with values in some state space $\mathcal{X}$. We transform this reference process R into $P :=$ $f_0(X_0) \exp\left(-\int_{[0,1]} V_t(X_t)\, dt\right) g_1(X_1)\, R$ where f_0, g_1 are nonnegative measurable functions on $\mathcal{X}$ and V is some measurable function on $[0,1] \times \mathcal{X}$. It is easily seen that P is also Markov. The aim of this paper is to identify the Markov generator of P in terms of the Markov generator of R and of the additional ingredients: f_0, g_1 and V in absence of regularity assumptions on f_0, g_1 and V.

As a first step, we show that the extended generator of a Markov process is essentially its stochastic derivative. Then, we compute the stochastic derivative of P to identify its generator, under a finite entropy condition. The abstract results are illustrated with continuous diffusion processes on $\mathbb{R}^d$ and Metropolis algorithms on a discrete space.

CONTENTS

1. INTRODUCTION

We consider continuous-time Markov processes with values in some Polish space $\mathcal{X}$ equipped with its Borel σ-field.

Notation. Let us fix some notation. The path space is the set

$$\Omega = D([0,1], \mathcal{X})$$

of all right continuous and left limited (càdlàg) $\mathcal{X}$-valued trajectories $\omega = (\omega_t)_{t \in [0,1]} \in \Omega$. It is equipped with the cylindrical σ-field: $\sigma(X_t; t \in [0,1])$ which is generated by the canonical process $X = (X_t)_{t \in [0,1]}$ defined for each $t \in [0,1]$ and $\omega \in \Omega$ by $X_t(\omega) = \omega_t \in \mathcal{X}$. We denote $\mathrm{P}(\Omega)$ the set of all probability measures on Ω. As usual, we call process any $P \in \mathrm{P}(\Omega)$ or any random element of Ω as well. For any $\mathcal{T} \subset [0,1]$, we denote $X_{\mathcal{T}} = (X_t)_{t \in \mathcal{T}}$ and the push-forward measure $P_{\mathcal{T}} = (X_{\mathcal{T}})_\# P$. In particular, for any $0 \leq r \leq s \leq 1$, $X_{[r,s]} = (X_t)_{r \leq t \leq s}$, $P_{[r,s]} = (X_{[r,s]})_\# P$ and $P_t = (X_t)_\# P \in \mathrm{P}(\mathcal{X})$ denotes the law of the

position X_t at time t where $P(\mathcal{X})$ the set of all probability measures on $\mathcal{X}$. The filtration is the canonical one: $\left(\sigma(X_{[0,t]}); t \in [0,1]\right)$.

Aim of the article. Let R be the law of some nicely behaved Markov process. We take this probability measure R as our *reference* law (this explains its unusual name R) and call *generalized h-transform* of R, any $P \in P(\Omega)$ which is absolutely continuous with respect to $R : P \ll R$, and with its Radon-Nikodym derivative of the special form:

$$P = f_0(X_0) \exp\left(-\int_{[0,1]} V_t(X_t)\, dt\right) g_1(X_1)\ R \tag{1}$$

where $f_0, g_1 : \mathcal{X} \to [0, \infty)$ are nonnegative Borel measurable functions on $\mathcal{X}$, the potential $V : [0,1] \times \mathcal{X} \to \mathbb{R}$ is also assumed to be Borel measurable on $[0,1] \times \mathcal{X}$ and all of them satisfy integrability conditions such that (1) defines a probability measure. We also say for short that P is an *h-process*.

It is easy to show (Proposition 4.2 below) that P inherits the Markov property from R. Consequently, it is tempting to know more about its infinitesimal generator. The aim of this article is to derive the generator of the Markov process P without assuming too many regularity conditions on R, f_0, g_1 and V.

Usual h-transform. Motivated by potential theory, the special case when $V \equiv 0$ but the terminal time $t = 1$ is replaced by some stopping time τ :

$$P = f_0(X_0)h(X_\tau)\ R^{(\tau)},$$

has been introduced in 1957 by J.L. Doob [Doo57, Doo00] with $R^{(\tau)}$ a Wiener process R killed at the exit time τ of a bounded domain D of $\mathbb{R}^d$. In this situation, for all $t \geq 0$ and x in D, the transition probability distributions of P are given by

$$P(X_t \in dz \mid X_0 = x) \propto h_t(z)\, R_t^{(\tau)}(dz \mid X_0 = x)$$

where $\propto$ means "proportional to" and $z \mapsto h_t(z) = E_R[\mathbf{1}_{\{\tau > t\}} h(X_\tau) \mid X_t = z]$ is a *space-time harmonic* function on D; this explains the letter h.

An example. In this paper, we shall only be concerned with the transform defined by (1), without stopping times. As an example, suppose that the reference process R is the unique solution of some stochastic differential equation

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dW_t$$

with locally Lipschitz coefficients b and σ, where W is a standard Wiener process on $\mathcal{X} = \mathbb{R}^d$. This implies that R is a solution of the martingale problem $\mathrm{MP}(b, a)$:

$$R \in \mathrm{MP}(b, a),$$

with b an adapted (drift) vector field and $a = \sigma\sigma^*$ an adapted (diffusion) matrix field. Since $P \ll R$, Girsanov's theory tells us that there exists some adapted vector field β such that P solves

$$P \in \mathrm{MP}(b + a\beta, a).$$

Now the problem is to express β in terms of the ingredients a, b, f_0, g_1 and V. Specifying the abstract results of this article to this continuous diffusion case leads to the next result (see Theorem 5.4 below): The additional drift term β can be written as

$$\beta(t, x) = \widetilde{\nabla}^P \psi(t, x), \quad dt P_t(dx)\text{-a.e.} \tag{2}$$

where

$$\psi(t,x) := \log E_R\Big[\exp\Big(-\int_{[t,1]} V_s(X_s)\, ds\Big) g_1(X_1) \mid X_t = x\Big], \quad dtP_t(dx)\text{-a.e.} \qquad (3)$$

is defined $dtP_t(dx)$-a.e. and $\widetilde{\nabla}^P$ is some linear operator which we call the P-*extended gradient*. This gradient coincides with the usual one on smooth functions: $\widetilde{\nabla}^P u = \nabla u$, for all $u \in \mathcal{C}_c^2(\mathbb{R}^d)$, when the diffusion matrix a has full rank. Of course, if R admits a regularizing and positivity improving transition probability density (for instance if R is the Wiener measure) and $V = 0$, then $\psi(t,x) := \log E_R(g_1(X_1) \mid X_t = x)$ is well-defined and smooth on $[0,1) \times \mathbb{R}^d$ and $\beta = \nabla\psi$. This situation is investigated in details by H. Föllmer [Föl88]. On the other hand, when V is a non-regular measurable function, even if R admits a regularizing semigroup, ψ may be a non-regular continuous function and (2) has an unusual meaning.

Non-regularity of V. The transition probability distributions in both directions of time of the generalized h-transform P are the Euclidean analogues [CZ91, CZ08] of the Feynman propagators [FH65] in the sense that for all $t \in [0,1]$

$$P(X_t \in dz \mid X_0 = x) \ \propto \ E_R\left[\exp\left(-\int_{[t,1]} V_s(X_s)\, ds\right) g_1(X_1) \mid X_t = z\right] R(X_t \in dz \mid X_0 = x)$$

$$P(X_t \in dz \mid X_1 = y) \ \propto \ E_R\left[f_0(X_0)\exp\left(-\int_{[0,t]} V_s(X_s)\, ds\right) \mid X_t = z\right] R(X_t \in dz \mid X_1 = y).$$

As non-regular potentials V are usual in physics, for instance discontinuous potentials with vertical asymptotic directions, we do not even assume that V is continuous.

From another view point, (1) is the generic form of the solution of the minimizer of the relative entropy

$$H(P|R) := \int_\Omega \log\left(\frac{dP}{dR}\right)\, dP \in [0,\infty]$$

which is seen as a function of P, subject to the constraints that its initial law P_0 is equal to some given $\mu_0 \in \mathrm{P}(\mathcal{X})$ and its flow of time-marginal laws $(P_t)_{t\in[0,1]}$ solves some prescribed Fokker-Planck evolution equation. In this convex optimization problem, f_0, g_1 and V act like Lagrange multipliers. See [Csi75, Föl88, CL94, CL95, CL96] for related entropy minimization problems and [Léo01] for a convex analytic derivation of this statement. For instance, when motivated by stochastic mechanics [Nel88], the above mentioned Fokker-Planck equation is related to the solution of some Schrödinger equation and its drift term explodes on the (nodal) set where the wave function vanishes. This enforces irregularities of V. See the introduction of [MZ85] for a brief explanation of this point and also Eq. (8) of [MZ85] where the potential $V_t(x) = \frac{A^R \Phi_t}{\Phi_t}(x)$ appears, with A^R the Markov generator of R and Φ the wave function.

Previous approaches to this problem. Let $P \in \mathrm{P}(\Omega)$ be a Markov process and $T_{s,t}^P u(x) := E_P[u(X_t) \mid X_s = x]$, $u \in U$, $0 \le s \le t$, be its semigroup on some Banach function space $(U, \|\cdot\|_U)$. For instance U may be the space of all bounded Borel measurable functions on $\mathcal{X}$ equipped with the topology of uniform convergence. Its infinitesimal generator is $A^P = (A_t^P)_{t\in[0,1]}$ with

$$A_t^P u(x) := \|\cdot\|_U\text{-}\lim_{h\downarrow 0}\frac{1}{h}E_P[u(X_{t+h}) - u(X_t) \mid X_t = x], \quad u \in \mathrm{dom}\, A^P \qquad (4)$$

where the domain dom A^P of A^P is precisely the set of all functions $u \in U$ such that the above strong limit exists for all $t \in [0,1)$ and $x \in \mathcal{X}$. We have seen with (2) and (3) that the function g defined by

$$g_t(x) := E_R \left[\exp \left(- \int_{[t,1]} V_s(X_s)\,ds \right) g_1(X_1) \mid X_t = x \right], \quad dt P_t(dx)\text{-a.e.} \qquad (5)$$

plays an important role in the description of the dynamics of P. One can prove rather easily (see [RY99, Ch. VIII,§3] for instance, in the diffusion case) that when V is zero and g is positive and regular enough, the generator A^P of the Markov semigroup associated with P is given for regular enough functions u on $\mathcal{X}$, by

$$A_t^P u(x) = A^R u(x) + \frac{\Gamma(g_t, u)}{g_t}(t, x), \quad (t, x) \in [0,1] \times \mathcal{X} \qquad (6)$$

where Γ is the carré du champ operator, defined for all functions u, v such that u, v and the product uv belong to the domain dom A^R of A^R, by

$$\Gamma(u, v) = A^R(uv) - uA^R v - vA^R u.$$

For Eq. (6) to be meaningful, it is necessary that for all $t \in [0,1]$, g_t and the product $g_t u$ belong to dom A^R. But we have already noticed that with a non-regular potential V, g might be non-regular as well. There is no reason why g_t and $g_t u$ are in dom A^R in general.

Clearly, one must drop the semigroup approach and work with semimartingales or Dirichlet forms. The Dirichlet form theory is natural for constructing irregular processes and has been employed in similar contexts, see [Alb03]. But it is made-to-measure for reversible processes and not very efficient when going beyond reversibility. Let us have a look at the semimartingale approach. Working with semimartingales means that instead of the infinitesimal semigroup generators A^R and A^P, we consider *extended generators* in the sense of the Strasbourg school [DM87], see Definition 2.2 below. This natural idea has already been implemented by P.-A. Meyer and W.A. Zheng [MZ84, MZ85] in the context of stochastic mechanics and also by P. Cattiaux and the author in [CL94, CL96] for solving related entropy minimization problems. But one still had to face the remaining problem of giving some sense to $\Gamma(g_t, u)$. Consequently, restrictive assumptions were imposed: reversibility in [MZ85] and, in [CL96], the standard hypothesis that the domain of the extended generator of R contains a "large" subalgebra. In practice this last requirement is not easy to verify, except for standard regular processes. In particular, it is difficult to find criteria for this property to be inherited by P when $P \ll R$.

In the present article, we overcome these limitations by choosing a different strategy which is based on stochastic derivatives and in some sense is more direct.

Further developments. Generalized h-processes are not only designed for Euclidean quantum mechanics [CZ08] or stochastic mechanics [Nel88].

  (i) They are a valuable tool for obtaining a new look at Hamilton-Jacobi-Bellman equations, by comparing the definition (1) with the usual Girsanov exponential Radon-Nikodym density.
 (ii) Because of the time symmetry of their definition when R is assumed to be reversible and V doesn't depend on t, they may bring interesting information about time reversal.
(iii) Even when V is zero, (1) provides an interesting process P which is sometimes called a Schrödinger bridge. It minimizes $H(P|R)$ subject to the marginal constraints $P_0 = \mu_0$ and $P_1 = \mu_1$. A connection with optimal transport is described in [**?**]. The flow $(P_t)_{t \in [0,1]}$ of this bridge is similar to the displacement interpolation introduced

by R. McCann [McC95] which is used for deriving functional inequalities or as a heuristic guideline in the so-called Otto calculus, see [Vil09]. This suggests that using $(P_t)_{t\in[0,1]}$ instead of the displacement interpolation could yield interesting results.

These potential developments will be investigated in future works.

Outline of the paper. The *stochastic derivative* L^P of P :

$$L_t^P u(x) := \lim_{h\downarrow 0} \frac{1}{h} E_P\left[u(X_{t+h}) - u(X_t) \mid X_t = x\right], \quad u \in \mathrm{dom}\, L^P$$

(compare (4)) was introduced by E. Nelson in [Nel67]. As usual, $\mathrm{dom}\, L^P$ is defined to be the set of all functions u such that the above limit exists, for the exact definition see Definition 2.6.

As a first step, we show that for a Markov process P, the stochastic derivative is equal to the extended generator $\mathcal{L}^P$ on a large class of functions u on $\mathcal{X}$:

$$\mathcal{L}_t^P u(x) = L_t^P u(x), \quad dt P_t(dx)\text{-a.e.}$$

This identity is the purpose of next Section 2 whose main results are Theorem 2.9 and Proposition 2.10. The key of Theorem 2.9's proof is the convolution Lemma 2.7.

With this general tool at hand, it remains to compute $L^P u$ for sufficiently many functions u to determine the martingale problem associated with P. And in view of (6), with g_t defined at (5), this essentially amounts to :

  (i) Prove that $g_t \in \mathrm{dom}\, L^R$ and compute $L^R g_t$;
  (ii) Prove that $g_t u \in \mathrm{dom}\, L^R$ for many "regular" functions u.

Problem (i) is solved at Section 3 by means of standard integration technics.

Problem (ii) is trickier. We solve it at Section 4 by assuming that the relative entropy of P with respect to R is finite:

$$H(P|R) < \infty.$$

The main technical step for solving this problem is Lemma 4.3 which allows us not to rely on Girsanov's theory in its usual form. In particular our abstract results are valid without assuming that R has the representation property (any R-martingale can be represented as some stochastic integral).

The main result of this paper is Theorem 4.12. It extends (6).

At Sections 5 and 6 we examplify our abstract results by means of continuous diffusion processes on $\mathbb{R}^d$ and time-continuous Markov chains. The main results of these sections are Theorem 5.4 which states (2) and Theorem 6.1 which describes the dynamics of the h-transforms of *Metropolis algorithms* on a discrete countable state space $\mathcal{X}$.

2. STOCHASTIC DERIVATIVES

We denote for any $t \in [0,1]$, $\overline{X}_t := (t, X_t) \in [0,1] \times \mathcal{X}$ and for any stopping time $Y_t^\tau := Y_{t\wedge\tau}$ and $\overline{X}_t^\tau := (t \wedge \tau, X_{t\wedge\tau})$.

Let P be a probability measure on Ω. Recall that a process M is called a *local P-martingale* if there exists a sequence $(\tau_k)_{k\geq 1}$ of $[0,1] \cup \{\infty\}$-valued stopping times such that $\lim_{k\to\infty} \tau_k = \infty$, P-a.s. and for each $k \geq 1$, the stopped process M^{τ_k} is a uniformly integrable P-martingale. A process Y is called a *special P-semimartingale* if $Y = B + M$, P-a.s. where B is a predictable bounded variation process and M is a local P-martingale.

Definition 2.1 (Nice semimartingale). *A process Y is called a nice[1] P-semimartingale if $Y = B + M$ is a special P-semimartingale and the bounded variation process B has absolutely continuous sample paths P-a.s.*

Definition 2.2 (Extended generator of a Markov process). *Let P be a Markov process. A measurable function u on $[0,1] \times \mathcal{X}$ is said to be in the domain of the extended generator of P if there exists a measurable function v on $[0,1] \times \mathcal{X}$ such that $\int_{[0,1]} |v(t, X_t)| \, dt < \infty$, P-a.e. and the process*

$$M_t^u := u(t, X_t) - u(0, X_0) - \int_{[0,t]} v(s, X_s) \, ds, \quad 0 \le t \le 1,$$

is a local P-martingale. We denote

$$v(t, x) =: \mathcal{L}^P u(t, x)$$

and call $\mathcal{L}^P$ the extended generator of P. The domain of the extended generator of P is denoted by $\operatorname{dom} \mathcal{L}^P$.

Remarks 2.3.

(a) In other words, the measurable function u on $[0,1] \times \mathcal{X}$ is in $\operatorname{dom} \mathcal{L}^P$ if the process $u(t, X_t)$ is a nice P-semimartingale.

(b) The adapted process $t \mapsto \int_{[0,t]} v(s, X_s) \, ds$ is predictable since it is continuous.

(c) M^u admits a càdlàg P-version as a local P-martingale (we always choose this regular version).

(d) In many situations it is enough to consider continuous functions u. But it will be useful at some point to consider $\mathcal{L}^P g$ with g given by (5) and it is not clear a priori that g is continuous in the general case, see Theorem 4.12 and Lemma 5.3 below for instance. This is the reason why we do not restrict $\operatorname{dom} \mathcal{L}^P$ to continuous functions.

(e) The notation $v = \mathcal{L}u$ almost rightly suggests that v is a function of u. Indeed, when u is in $\operatorname{dom} \mathcal{L}^P$, the Doob-Meyer decomposition of the special semimartingale $u(t, X_t)$ into its predictable bounded variation part $\int v_s \, ds$ and its local martingale part is unique. But one can modify $v = \mathcal{L}^P u$ on a small (zero-potential) set without breaking the martingale property. As a consequence, $u \mapsto \mathcal{L}^P u$ is a multivalued operator and $u \mapsto \mathcal{L}^P u$ is an almost linear operation.

(f) Suppose that t_o is a fixed time of discontinuity of P, i.e. $P(X_{t_o} \ne X_{t_o^-}) > 0$. Then, in general a continuous function u cannot be in $\operatorname{dom} \mathcal{L}^P$. For this reason, one should think of the notion of extended generator for processes P that do not have any fixed time of discontinuity: $P(X_t \ne X_{t^-}) = 0$, for all $t \in [0,1]$.

The notion of generator is tightly connected with that of martingale problem.

Definition 2.4 (Martingale problem). *Let $\mathcal{C}$ be a class of measurable real functions u on $[0,1] \times \mathcal{X}$ and for each $u \in \mathcal{C}$, let $\mathcal{L}u : [0,1] \times \mathcal{X} \to \mathbb{R}$ be a measurable function such that $\int_{[0,1]} |\mathcal{L}u(t, \omega_t)| \, dt < \infty$ for all $\omega \in \Omega$. Take also a probability measure $\mu_0 \in \mathrm{P}(\mathcal{X})$. One says that $Q \in \mathrm{P}(\Omega)$ is a solution to the martingale problem $\mathrm{MP}(\mathcal{L}, \mathcal{C}; \mu_0)$ if $Q_0 = \mu_0 \in \mathrm{P}(\mathcal{X})$ and for all $u \in \mathcal{C}$, the process*

$$u(t, X_t) - u(0, X_0) - \int_{[0,t]} \mathcal{L}u(s, X_s) \, ds$$

is a local Q-martingale.

---

[1]This is a "local" definition in the sense that this notion probably appears somewhere else with another name.

As in Definition 2.2, this local martingale admits a càdlàg Q-version.

Playing with the definitions, it is clear that any Markov law $Q \in \mathrm{P}(\Omega)$ is a solution to $\mathrm{MP}(\mathcal{L}^Q, \mathcal{C}; Q_0)$ where $\mathcal{L}^Q$ is the extended generator of Q and $\mathcal{C}$ is any nonempty subset of dom $\mathcal{L}^Q$.

Our aim is to show that the extended generator can be computed by means of a stochastic derivative.

Definition 2.5 (Integration time). *Let u be a measurable real function on $[0,1] \times \mathcal{X}$ and τ be a stopping time. We say that τ is a P-integration time of u if the family of random variables $\{u(\overline{X}_t^\tau); t \in [0,1]\}$ is uniformly P-integrable.*

Definition 2.6 (Stochastic derivative of a Markov process). *Let P be a Markov process and u be a measurable real function on $[0,1] \times \mathcal{X}$. We say that u admits a stochastic derivative under P at time $t \in [0,1]$ if for P_t-almost all $x \in \mathcal{X}$ there exists a P-integration time σ^x of u such that $\sigma^x \geq t$, P-a.e. and for any P-integration time τ of u satisfying $\tau > \sigma^x$, P-a.e. the following limit*

$$L^P u(t,x) := \lim_{h \downarrow 0} E_P \left(\frac{1}{h} [u(\overline{X}_{t+h}^\tau) - u(t,x)] \mid X_t = x \right)$$

exists and does not depend on τ.

If u admits a stochastic derivative for $dt P_t(dx)$-almost all (t,x), we say that u belongs to the domain dom L^P of the stochastic derivative L^P of the Markov process P.

If the function u does not depend on the time variable t, we denote

$$L_t^P u(x) = L^P u(t,x).$$

This extension of Nelson's definition by means of integration times seems to be new. It is consistent since the supremum of two integration times is still an integration time. Indeed, the supremum of two stopping times is a stopping time and for all t, $|u(\overline{X}_t^{\tau \vee \tau'})| \leq |u(\overline{X}_t^\tau)| + |u(\overline{X}_t^{\tau'})|$.

As in Definition 2.2, we do not restrict the domain of the stochastic derivative to continuous functions, see Remark 2.3-(d).

Since P is a Markov process, we have also

$$L^P u(t,x) = \lim_{h \downarrow 0} E_P \left(\frac{1}{h} [u(\overline{X}_{t+h}^\tau) - u(t,x)] \mid \tau > t, X_t = x \right).$$

We denote $\overline{P}$ the product of the Lebesgue measure on $[0,1]$ by the process P: $\overline{P}(dt d\omega) = dt P(d\omega)$. In the sequel, we shall be concerned with the function space $L^p([0,1] \times \Omega, \overline{P})$.

Lemma 2.7. *For all $h > 0$, let $k^h \geq 0$ be a measurable convolution kernel such that $\operatorname{supp} k^h \subset [-h, h]$ and $\int_{\mathbb{R}} k^h(s) \, ds = 1$.*
Let P be a bounded positive measure on Ω (which may not be a probability measure) and $v(t, \omega)$ be a function in $L^p([0,1] \times \Omega, \overline{P})$ with $1 \leq p < \infty$. Define for all $h > 0$ and $t \in [0,1]$, $k^h * v(t) = \int_{\mathbb{R}} k^h(t-s) v_s \, ds$ where v is extended by putting $v_s = 0$ for all $s \notin [0,1]$.
Then, $k^h * v$ is in $L^p([0,1] \times \Omega, \overline{P})$ and $\lim_{h \downarrow 0} k^h * v = v$ in $L^p([0,1] \times \Omega, \overline{P})$.

We see that $k^h(s) \, ds$ is a probability measure on $\mathbb{R}$ which converges narrowly to the Dirac measure δ_0 as h tends down to zero.

Proof. In this lemma, we endow as usual Ω with the Skorokhod topology which turns it into a Polish space and has the interesting property that its Borel σ-field matches with the cylindrical σ-field.

We denote $L^p([0,1] \times \Omega, \overline{P}) = L^p(\overline{P})$ and start the proof by showing that $k^h * v \in L^p(\overline{P})$. For P-almost all ω, $v(\cdot, \omega) \in L^p([0,1])$ so that $k^h * v(\cdot, \omega)$ is also in $L^p([0,1])$ with $\|k^h * v(\cdot, \omega)\|_{L^p([0,1])} \leq \|v(\cdot, \omega)\|_{L^p([0,1])}$. It remains to integrate with respect to $P(d\omega)$ to obtain

$$\|k^h * v\|_{L^p(\overline{P})} \leq \|v\|_{L^p(\overline{P})} < \infty. \tag{7}$$

Now, we prove the convergence. As p is finite, the space $C_c([0,1] \times \Omega)$ of all continuous functions with a compact support in $[0,1] \times \Omega$ is dense in $L^p(\overline{P})$. We approximate v in $L^p(\overline{P})$ by a sequence $(v_n)_{n \geq 1}$ in $C_c([0,1] \times \Omega)$. For all h and n

$$\begin{aligned} \|k^h * v - v\|_{L^p(\overline{P})} &\leq \|k^h * (v - v_n)\|_{L^p(\overline{P})} + \|k^h * v_n - v_n\|_{L^p(\overline{P})} + \|v_n - v\|_{L^p(\overline{P})} \\ &\leq \|k^h * v_n - v_n\|_{L^p(\overline{P})} + 2\|v - v_n\|_{L^p(\overline{P})} \end{aligned}$$

where we used (7).
Take an arbitrary small $\eta > 0$ and choose n large enough for $\|v - v_n\|_{L^p(\overline{P})} \leq \eta$ to hold. Then,

$$\|k^h * v - v\|_{L^p(\overline{P})} \leq \|k^h * v_n - v_n\|_{L^p(\overline{P})} + 2\eta. \tag{8}$$

Fix this n. Since v_n is in $C_c([0,1] \times \Omega)$, it is a uniformly continuous function. Therefore, for all $\eta > 0$, there exists $h(\eta) > 0$ such that for any t, t', ω, ω' satisfying $|t - t'| + d_\Omega(\omega, \omega') \leq h(\eta)$, we have $|v_n(t', \omega') - v_n(t, \omega)| \leq \eta$, where d_Ω is the Skorokhod metric on Ω. In particular, with $\omega = \omega'$, we see that

$$|t' - t| \leq h(\eta) \Rightarrow \sup_{\omega \in \Omega} |v_n(t', \omega) - v_n(t, \omega)| \leq \eta.$$

Because of the property: $\operatorname{supp} k^h \subset [-h, h]$, we deduce from this that for any $\omega \in \Omega$, $|k^h * v_n(t) - v_n(t)| \leq \int_{\mathbb{R}} |v_n(t-s) - v_n(t)| k^h(s) \, ds \leq \eta$ as soon as $h \leq h(\eta)/2$. Consequently $\|k^h * v_n - v_n\|_{L^p(\overline{P})} \leq P(\Omega)\eta$. Finally, with (8) this leads us to $\|k^h * v - v\|_{L^p(\overline{P})} \leq (2 + P(\Omega))\eta$. Since η is arbitrary, this shows that $\lim_{h \to 0} \|k^h * v - v\|_{L^p(\overline{P})} = 0$, which is the desired result $\qquad \square$

Proposition 2.8. *Let P be a Markov process and u be a function in the domain $\operatorname{dom} \mathcal{L}^P$ of the extended generator $\mathcal{L}^P$ of P. We suppose in addition that there exists $1 \leq p < \infty$ such that $E_P \int_{[0,1]} |\mathcal{L}^P u(t, X_t)|^p \, dt < \infty$. Then,*

$$\lim_{h \downarrow 0} E_P \int_{[0,1-h]} \left| \frac{1}{h} E_P[u(t+h, X_{t+h}) - u(t, X_t) \mid X_t] - \mathcal{L}^P u(t, X_t) \right|^p dt = 0. \tag{9}$$

Proof. We denote $v_t = \mathcal{L}^P u(t, X_t)$. Choosing the specific convolution kernel $k^h = \frac{1}{h} \mathbf{1}_{[-h,0]}$, and relying on the very definition of the extended generator, we obtain

$$\begin{aligned} \frac{1}{h} E_P[u(t+h, X_{t+h}) - u(t, X_t) \mid X_t] &= \frac{1}{h} E_P[u(t+h, X_{t+h}) - u(t, X_t) \mid X_{[0,t]}] \\ &= E_P[k^h * v(t) \mid X_{[0,t]}] = E_P[k^h * v(t) \mid X_t]. \end{aligned}$$

On the other hand, by Jensen's inequality and Fubini's theorem

$$\begin{aligned} E_P \int_{[0,1]} \left| E_P[k^h * v(t) \mid X_t] - v_t \right|^p dt &= E_P \int_{[0,1]} \left| E_P[k^h * v(t) - v_t \mid X_t] \right|^p dt \\ &\leq E_P \int_{[0,1]} |k^h * v(t) - v_t|^p \, dt. \end{aligned}$$

Our hypothesis $v \in L^p([0,1] \times \Omega, \overline{P})$ is precisely the assumption of previous Lemma 2.7 which insures that $\lim_{h \downarrow 0} E_P \int_{[0,1]} |k^h * v(t) - v_t|^p \, dt = 0$. Gathering these considerations, we obtain (9). $\qquad \square$

A variant of this proposition already appears in [Föl86]. But it seems to the author that its proof is incomplete and that it is difficult to avoid a convolution argument such as Lemma 2.7.

Theorem 2.9. *Let P be a Markov process and u be a function in the domain $\operatorname{dom} \mathcal{L}^P$ of the extended generator $\mathcal{L}^P$ of P. Then, u belongs to $\operatorname{dom} L^P$ and*

$$\mathcal{L}^P u = L^P u, \quad dt P_t(dx)\text{-a.e.}$$

Proof. By the definition of the extended generator, there exists a localizing sequence $(\tau_k)_{k \geq 1}$ of stopping times, i.e. such that $\lim_{k \to \infty} \tau_k = \infty$, P-a.e. and for all $k \geq 1$, the stopped process M^{τ_k} where

$$M_t = u(t, X_t) - \int_{[0,t]} \mathcal{L}^P u(s, X_s) \, ds,$$

is a uniformly integrable martingale. By considering the sequence of stopping times $\inf\{t \in [0,1]; \int_{[0,t]} |\mathcal{L}^P u(s, X_s)| \, ds \geq k\} \in [0,1] \cup \{\infty\}$ indexed by $k \geq 1$, it is easy to show that $(\tau_k)_{k \geq 1}$ can also be chosen such that for each k, τ_k is also an integration time of u.

Let us consider a fixed integration time τ of u such that M^τ is a uniformly integrable martingale. Denoting $v^\tau(t) = \mathbf{1}_{\{t \leq \tau\}} \mathcal{L}^P u(t, X_t)$ and choosing $k^h = \frac{1}{h} \mathbf{1}_{[-h,0]}$ as in the proof of Proposition 2.8, we see that $h \mapsto [u(\overline{X}^\tau_{t+h}) - u(\overline{X}^\tau_t)] - h k^h * v^\tau(t)$ is a martingale. It follows that

$$\frac{1}{h} E_P[u(\overline{X}^\tau_{t+h}) - u(t, X^\tau_t) \mid X_{[0,t]}] = E_P[k^h * v^\tau(t) \mid X_{[0,t]}] = \mathbf{1}_{\{t \leq \tau\}} E_P[k^h * v(t) \mid X_t].$$

Remark for future use that this implies that

$$\frac{1}{h} E_P[u(\overline{X}^\tau_{t+h}) - u(t, X^\tau_t) \mid X_{[0,t]}] = \mathbf{1}_{\{t \leq \tau\}} \frac{1}{h} E_P[u(\overline{X}^\tau_{t+h}) - u(t, X^\tau_t) \mid X^\tau_t]. \qquad (10)$$

Then, as for (9) with $p = 1$, we obtain

$$\lim_{h \downarrow 0} E_P \int_{[0, \tau \wedge (1-h)]} \left| \frac{1}{h} E_P[u(\overline{X}^\tau_{t+h}) - u(t, X^\tau_t) \mid X_{[0,t]}] - \mathcal{L}^P u(t, X_t) \right| dt = 0$$

and with Fatou's lemma

$$E_P \int_{[0,1]} \liminf_{h \downarrow 0} \mathbf{1}_{\{t \leq \tau\}} \left| \frac{1}{h} E_P[u(\overline{X}^\tau_{t+h}) - u(t, X^\tau_t) \mid X_{[0,t]}] - \mathcal{L}^P u(t, X_t) \right| dt = 0.$$

But, since u is in $\operatorname{dom} \mathcal{L}^P$, $\lim_{h \downarrow 0} \frac{1}{h} E_P[u(\overline{X}^\tau_{t+h}) - u(t, X^\tau_t) \mid X_{[0,t]}]$ appears as the computation of the derivative of an absolutely continuous function. Therefore, this limit exists for Lebesgue-almost all t and the $\liminf_{h \downarrow 0}$ arising from the application of Fatou's lemma is a genuine limit[2]. Hence,

$$E_P \int_{[0,1]} \mathbf{1}_{\{t \leq \tau\}} \lim_{h \downarrow 0} \left| \frac{1}{h} E_P[u(\overline{X}^\tau_{t+h}) - u(t, X_t) \mid X_{[0,t]}] - \mathcal{L}^P u(t, X_t) \right| dt = 0.$$

---

[2]The absolute continuity plays a crucial role. Note that it is also of primary importance in the definition of the extended generator.

With (10) this shows us that for $\overline{P}$-almost all (t, ω) we have

$$\mathbf{1}_{\{\tau(\omega) \geq t\}} \lim_{h \downarrow 0} \frac{1}{h} E_P \left[u(\overline{X}_{t+h}^\tau) - u(t, X_t(\omega)) \mid X_t^\tau \right](\omega) = \mathbf{1}_{\{\tau(\omega) \geq t\}} \mathcal{L}^P u(t, X_t(\omega)).$$

As the left-hand side vanishes when $\tau(\omega) = t$, we obtain

$$\mathbf{1}_{\{\tau > t\}} \lim_{h \downarrow 0} \frac{1}{h} E_P \left[u(\overline{X}_{t+h}^\tau) - u(t, X_t) \mid \tau > t, X_t \right] = \mathbf{1}_{\{\tau > t\}} \mathcal{L}^P u(t, X_t).$$

This results holds true for any integration time τ of u such that M^τ is a uniformly integrable martingale.

By assumption, for $\overline{P}$-almost all (t, ω) there exists $k(t, \omega)$ large enough for the localizing time $\tau_{k(t,\omega)}$ to satisfy $\tau_{k(t,\omega)}(\omega) \geq t$. Choosing $\sigma^{X_t(\omega)} = \tau_{k(t,\omega)}$, we obtain for $dt P_t(dx)$-almost all (t, x) an integration time $\sigma^x \geq t$ such that any integration time $\tau > \sigma^x$ satisfies

$$\lim_{h \downarrow 0} \frac{1}{h} E_P \left[u(\overline{X}_{t+h}^\tau) - u(t, x) \mid X_t = x \right](\omega) = \mathcal{L}^P u(t, x).$$

This completes the proof of the theorem.   □

Let us investigate a partial converse of Theorem 2.9.

Proposition 2.10. *Let P be a Markov process, u and v be measurable real functions on $[0, 1] \times \mathcal{X}$ which satisfy the following requirements. The function v verifies $\int_{[0,1]} |v(t, X_t)| \, dt < \infty$, P-a.s. and there exists a sequence $(\tau^k)_{k \geq 1}$ of integration times of u such that $\lim_{k \to \infty} \tau_k = \infty$, P-a.s. and for each $k \geq 1$,*

$$\lim_{h \downarrow 0} E_P \int_{[0, 1-h]} \left| \frac{1}{h} E_P[u(\overline{X}_{t+h}^{\tau_k}) - u(\overline{X}_t^{\tau_k}) \mid X_t] - \mathbf{1}_{\{t \leq \tau_k\}} v(t, X_t) \right| dt = 0 \qquad (11)$$

Then, u belongs to $\operatorname{dom} \mathcal{L}^P$ and $\operatorname{dom} L^P$ and

$$\mathcal{L}^P u = L^P u = v, \quad dt P_t(dx)\text{-a.e.}$$

Note that if P admits a fixed time of discontinuity, there might be many continuous functions u which do not verify (11).

Proof. The proof relies on the subsequent easy analytic result.

Claim. *Let a, b be two measurable functions on $[0, 1]$ such that a is right continuous, b is Lebesgue-integrable and $\lim_{h \downarrow 0} \int_{[0, 1-h]} \left| \frac{1}{h} \{ a(t+h) - a(t) \} - b(t) \right| dt = 0$. Then, a is absolutely continuous and its distributional derivative is $\dot{a} = b$.*

To see this, remark first that $t \mapsto \mathbf{1}_{\{0 \leq t \leq 1-h\}} \frac{1}{h} \{ a(t+h) - a(t) \}$ is integrable for any $0 < h \leq 1$. Take any $0 \leq r \leq s < 1$. On one hand, we have $\lim_{h \downarrow 0} \int_{[r,s]} \frac{1}{h} \{ a(t+h) - a(t) \} \, dt = \int_{[r,s]} b(t) \, dt$ and on the other one: $\int_{[r,s]} \frac{1}{h} \{ a(t+h) - a(t) \} \, dt = \frac{1}{h} \int_{[r,r+h]} a(t) \, dt - \frac{1}{h} \int_{[s,s+h]} a(t) \, dt$, so that with the assumed right continuity of a we have $\lim_{h \downarrow 0} \int_{[r,s]} \frac{1}{h} \{ a(t+h) - a(t) \} \, dt = a(s) - a(r)$. Therefore $a(s) - a(r) = \int_{[r,s]} b(t) \, dt$ which is the claimed property.

Let us fix τ^k as in the assumption of the proposition. We write $E = E_P$, $u_t = u(\overline{X}_t^{\tau_k})$ and $v_t = \mathbf{1}_{\{t \leq \tau_k\}} v(\overline{X}_t^{\tau_k})$ to simplify the notation. Define the family of stopping times $\sigma_k := \inf\{s \in [0, 1]; \int_{[0, s]} |v(t, X_t)| \, dt \geq k\}$ where k describes the integers. By considering the stopping times $\sigma_k \wedge \tau_k$, we can assume without loss of generality that

$$v \in L^1(\overline{P}).$$

Fix $0 \le r < 1$. We have

$$\left| E \left[\int_{[r,1-h]} \left(\frac{1}{h}\{u_{t+h} - u_t\} - v_t \right) dt \mid X_r \right] \right|$$

$$\le E \left[\int_{[r,1-h]} E \left(\left| \frac{1}{h}\{u_{t+h} - u_t\} - v_t \right| \mid X_t \right) dt \mid X_r \right]$$

With (11) and Fatou's lemma, we obtain

$$E \left(\liminf_{h \downarrow 0} \left| E \left[\int_{[r,1-h]} \left(\frac{1}{h}\{u_{t+h} - u_t\} - v_t \right) dt \mid X_r \right] \right| \right)$$

$$\le \lim_{h \downarrow 0} E \int_{[r,1-h]} E \left(\left| \frac{1}{h}\{u_{t+h} - u_t\} - v_t \right| \mid X_t \right) dt = 0.$$

Hence, there exists a sequence $(h_n)_{n \ge 1}$ of positive numbers such that $\lim_{n \to \infty} h_n = 0$ and

$$\lim_{n \to \infty} \int_{[r,1-h_n]} E \left[\left(\frac{1}{h_n}\{u_{t+h_n} - u_t\} - v_t \right) \mid X_r \right] dt = 0, \quad P\text{-a.e.}$$

It remains to apply the result of the above claim to $a(t) = E\left[u_t \mid X_r \right]$ and $b(t) = E\left[v_t \mid X_r \right]$ to see that for all $0 \le r \le s < 1$, $E\left[u_s - u_r - \int_{[r,s]} v_t \, dt \mid X_{[0,r]} \right] = 0$. This proves that M^{τ_k} is a P-martingale where

$$M_s := u(s, X_s) - u(0, X_0) - \int_{[0,s]} v(t, X_t) \, dt.$$

With the assumptions that $\lim_{k \to \infty} \tau_k = \infty$, P-a.s., $\mathbf{1}_{[0,\tau_k]} v \in L^1(\overline{P})$ and the fact that $\{u(\overline{X}_t^{\tau_k}); t \in [0,1]\}$ is uniformly P-integrable by the very definition of the integration time τ_k, we conclude that M is a local P-martingale. Therefore, u belongs to $\operatorname{dom} \mathcal{L}^P$ and $\mathcal{L}^P u = v$. And we also have $u \in \operatorname{dom} L^P$ and $L^P u = \mathcal{L}^P u$ by Theorem 2.9. $\qquad\square$

3. Feynman-Kac processes

Let R be a probability measure on Ω which is a stationary Markov process with the invariant *probability* measure

$$m := R_t \in \mathrm{P}(\mathcal{X}), \quad \forall t \in [0,1]. \tag{12}$$

We also consider a lower bounded potential V, i.e. a measurable function $V : [0,1] \times \mathcal{X} \to \mathbb{R}$ such that

$$\inf_{[0,1] \times \mathcal{X}} V \ge -\lambda_o \tag{13}$$

with $0 \le \lambda_o < \infty$. Let g_1 be a nonnegative m-integrable function on $\mathcal{X}$. In this section we look at the real valued process

$$G_t := E_R \left[\exp \left(-\int_{[t,1]} V_s(X_s) \, ds \right) g_1(X_1) \mid X_{[0,t]} \right] =: g_t(X_t), \quad t \in [0,1], \quad g_1 \ge 0 \tag{14}$$

which we call a Feynman-Kac process. Last equality, where $g_t : \mathcal{X} \to [0, \infty)$ is a measurable function, is a consequence of the Markov property of R.

Orlicz spaces. The mere integrability of g_1 is sufficient for defining G, but it will not be enough in general for our purpose. We are going to assume that g_1 is in some Orlicz space

$$L^\gamma(m) := \left\{ u : \mathcal{X} \to \mathbb{R}; \text{ measurable}, \int_\mathcal{X} \gamma(a_o|u|)\, dm < \infty, \text{ for some } a_o > 0 \right\}$$

associated with the Young function γ. Recall that $\gamma : \mathbb{R} \to [0, \infty]$ is a Young function if it is convex, even, lower semicontinuous and $\gamma(0) = 0$. Important instances are

- $\gamma(a) = \gamma_p(a) := |a|^p/p$, with $1 \le p < \infty$, then $L^{\gamma_p}(m) = L^p(m)$;
- $\gamma(a) = \gamma_\infty(a) := \begin{cases} 0 & \text{if } |a| \le 1 \\ \infty & \text{otherwise} \end{cases}$, then $L^{\gamma_\infty} = L^\infty(m)$.

Let us introduce the functions

$$\begin{aligned} \theta(a) &:= e^a - a - 1, \quad a \in \mathbb{R}, \\ \theta^*(a) &:= (a+1)\log(a+1) - a, \quad a \in [-1, \infty) \end{aligned} \tag{15}$$

with the convention $0 \log 0 = 0$. They are convex conjugate to each other and $\theta(a) = \log \mathbb{E} e^{a(N-1)}$ where N is a Poisson(1) random variable. Moreover, $\theta(|a|)$ and $\theta^*(|b|)$ are Young functions which are also convex conjugate to each other.

Two other important Orlicz spaces are

- $\gamma(a) = \theta(|a|)$ corresponds to the following L^γ :

$$L^{\exp}(m) := \left\{ u : \mathcal{X} \to \mathbb{R}; \text{ measurable}, \int_\mathcal{X} e^{a_o|u|}\, dm < \infty, \text{ for some } a_o > 0 \right\},$$

- $\gamma(a) = \theta^*(|a|)$ corresponds to the following L^γ :

$$L \log L(m) := \left\{ u : \mathcal{X} \to \mathbb{R}; \text{ measurable}, \int_\mathcal{X} |u| \log_+ |u|\, dm < \infty \right\},$$

where we use the assumed boundedness of the positive measure m in the above expressions. The Luxemburg norm of $L^\gamma(m)$ is defined by $\|u\|_{L^\gamma(m)} := \inf\{\alpha > 0; \int_\mathcal{X} \gamma(|u|/\alpha)\, dm \le 1\}$. Let $\gamma^*(b) := \sup_{a \ge 0}\{ab - \gamma(a)\} \in [0, \infty]$, $b \ge 0$, be the convex conjugate of γ. It follows immediately from Fenchel's inequality $ab \le \gamma(a) + \gamma^*(b)$, that the Hölder inequality

$$\|uv\|_{L^1(m)} \le 2\|u\|_{L^\gamma(m)}\|v\|_{L^{\gamma^*}(m)}, \quad u \in L^\gamma(m), v \in L^{\gamma^*}(m)$$

holds true. In particular, since $\theta(|\cdot|)$ and $\theta^*(|\cdot|)$ are convex conjugate to each other, we have $\|uv\|_{L^1(m)} \le 2\|u\|_{L \log L(m)}\|v\|_{L^{\exp}(m)}$, for all $u \in L \log L(m), v \in L^{\exp}(m)$.

The Young function γ is said to satisfy the *condition* Δ_2 if there exist constants $C, A > 0$ such that $\gamma(2a) \le C\gamma(a)$, for all $a \ge A$. The spaces $L \log L(m)$ and $L^p(m)$ with $1 \le p < \infty$ satisfy Δ_2. But $L^\infty(m)$ and $L^{\exp}(m)$ do not.

Preliminary results. We assume that the next finite entropy condition is satisfied

$$g_1 \ge 0, \quad \int_\mathcal{X} g_1 \log_+ g_1\, dm < \infty$$

and we pick a Young function γ such that

$$\int_\mathcal{X} \gamma(g_1)\, dm < \infty \quad \text{and} \quad L \log L(m) \subset L^\gamma(m) \subset L^p(m) \text{ for some } 1 < p < \infty. \tag{16}$$

In particular, we have $\gamma \in \Delta_2$.

Because of (12), (13) and (16), with G_t given by (14), we have for all $t \in [0,1]$ and $\alpha > 0$,

$$\int_{\mathcal{X}} \gamma(g_t/\alpha)\, dm = E_R \gamma(G_t/\alpha) \leq E_R \gamma(e^{\lambda_o} G_1/\alpha) \leq C_{\gamma,\lambda_o} E_R(G_1/\alpha)$$

where $C_{\gamma,\lambda_o} > 0$ is some finite constant which can be derived by means of the condition Δ_2. Optimizing in α leads us to

$$\|g_t\|_{L^\gamma(m)} \leq C_{\gamma,\lambda_o} \|g_1\|_{L^\gamma(m)}, \quad \forall t \in [0,1].$$

Recall that a real valued process G is said to admit a càdlàg version if there exists a modification G' of G, i.e. $R(G_t \neq G'_t) = 0$ for all $t \in [0,1]$, with its sample paths in $D_{\mathbb{R}} := D([0,1], \mathbb{R})$.

Lemma 3.1. *Let us assume that in addition to* (12), (13) *and* (16), *we have*

$$\int_{[0,1]} \|V_t\|_{L^1(m)}\, dt < \infty. \tag{17}$$

Then, the process G admits a càdlàg version. In the sequel G will always be assumed to be this $D_{\mathbb{R}}$-valued version.

It is a nonnegative semimartingale which satisfies the so-called Feynman-Kac semigroup property:

$$E_R \left[\exp \left(-\int_{[s,t]} V_r(X_r)\, dr \right) G_t \mid X_{[0,s]} \right] = G_s, \quad 0 \leq s \leq t \leq 1. \tag{18}$$

Moreover, denoting $G_* := \sup_{t \in [0,1]} G_t$, we have

$$\|G_*\|_{L^\gamma(R)} \leq C_{\gamma,\lambda_o} \|g_1\|_{L^\gamma(m)}$$

for some finite positive constant C_{γ,λ_o}.

This implies that $\{\gamma(G_t); t \in [0,1]\}$ is uniformly integrable in $L^1(R)$.

Proof. Let us prove (18). For all $0 \leq s \leq t \leq 1$,

$$E_R \left[\exp \left(-\int_{[s,t]} V_r(X_r)\, dr \right) G_t \mid X_{[0,s]} \right]$$

$$= E_R \left[\exp \left(-\int_{[s,t]} V_r(X_r)\, dr \right) E_R \left\{ \exp \left(-\int_{[t,1]} V_r(X_r)\, dr \right) G_1 \mid X_{[0,t]} \right\} \mid X_{[0,s]} \right]$$

$$= E_R \left[E_R \left\{ \exp \left(-\int_{[s,1]} V_r(X_r)\, dr \right) G_1 \mid X_{[0,t]} \right\} \mid X_{[0,s]} \right]$$

$$= E_R \left[\exp \left(-\int_{[s,1]} V_r(X_r)\, dr \right) G_1 \mid X_{[0,s]} \right]$$

$$= G_s$$

which is (18).

Let us define $\widetilde{V} := V + \lambda_o \geq 0$ and for all $t \in [0,1]$

$$\widetilde{G}_t := e^{-\lambda_o(1-t)} G_t = E_R \left[\exp \left(-\int_{[t,1]} \widetilde{V}_s(X_s)\, ds \right) \widetilde{G}_1 \mid X_{[0,t]} \right]$$

where $\widetilde{G}_1 = G_1 = g_1(X_1)$. Because $\widetilde{V} \geq 0$, we see that for all $0 \leq s \leq t \leq 1$,

$$
\begin{aligned}
E_R\left(\widetilde{G}_t \mid X_{[0,s]}\right) &= E_R\left[\exp\left(-\int_{[t,1]} \widetilde{V}_r(X_r)\,dr\right)\widetilde{G}_1 \mid X_{[0,s]}\right] \\
&\geq E_R\left[\exp\left(-\int_{[s,1]} \widetilde{V}_r(X_r)\,dr\right)\widetilde{G}_1 \mid X_{[0,s]}\right] \\
&= \widetilde{G}_s.
\end{aligned}
$$

In other words, $\widetilde{G}$ is a nonnegative submartingale.

It follows from the fact that the forward filtration satisfies the standard assumptions and from a well-known result of the general theory of stochastic processes that $\widetilde{G}$ admits a càdlàg modification (still denoted by $\widetilde{G}$) if $t \in [0,1] \mapsto E_R\widetilde{G}_t \in [0,\infty)$ is a right continuous real function. But this latter property is a direct consequence of Lebesgue's dominated convergence theorem and the pathwise right continuity of

$$
t \in [0,1] \mapsto \exp\left(-\int_{[t,1]} \widetilde{V}_r(X_r)\,dr\right) \in (0,1]
$$

which is satisfied under the assumption (17): $E_R \int_{[0,1]} |V_t(X_t)|\,dt < \infty$, which implies that $\int_{[0,1]} |\widetilde{V}_t(X_t)|\,dt < \infty$, R-a.s.[3]

Furthermore, we have $E_R\gamma(\widetilde{G}_t) \leq E_R\gamma(\widetilde{G}_1) < \infty$ by Jensen's inequality and the submartingale property. Doob's maximal inequality, which holds for any nonnegative submartingale and any Young function γ which verifies (16)[4], tells us that there exists a positive finite constant $c_\gamma < \infty$ such that

$$
\|\sup_{t \in [0,1]} \gamma(\widetilde{G}_t)\|_{L^1(R)} \leq c_\gamma \sup_{t \in [0,1]} \|\gamma(\widetilde{G}_t)\|_{L^1(R)} = c_\gamma \|\gamma(\widetilde{G}_1)\|_{L^1(R)} = c_\gamma \|\gamma(g_1)\|_{L^1(m)} < \infty.
$$

Hence $\{\gamma(\widetilde{G}_t); t \in [0,1]\}$ is uniformly integrable in $L^1(R)$. Since the product of two semimartingales is still a semimartingale, we deduce that $G_t = e^{\lambda_o(1-t)}\widetilde{G}_t$ is a càdlàg semimartingale such that $\{\gamma(G_t); t \in [0,1]\}$ is uniformly integrable in $L^1(R)$. This completes the proof of the lemma. $\qquad\square$

Recall that since R is a bounded nonnegative measure, a family $\{H_t; t \in [0,1]\}$ of real valued measurable functions is uniformly integrable in $L^1(R)$ if and only if there exists an increasing convex function $\xi : [0,\infty) \to [0,\infty)$ such that $\lim_{a\to\infty} \xi(a)/a = +\infty$ and $\sup_{t \in [0,1]} E_R\xi(|H_t|) < \infty$.

Claim.   *Let A_t, B_t, $t \in [0,1]$ be two random variables such that both $\{\gamma(A_t); t \in [0,1]\}$ and $\{\gamma^*(B_t); t \in [0,1]\}$ are uniformly integrable in $L^1(R)$. Then, the family of products $\{A_tB_t; t \in [0,1]\}$ is uniformly integrable in $L^1(R)$.*

Let us prove this claim. By hypothesis there exist two functions ξ_1 and ξ_2 as above such that $\sup_t E\xi_1(\gamma(A_t)) < \infty$ and $\sup_t E\xi_2(\gamma^*(B_t)) < \infty$ where we wrote $\sup_t = \sup_{t \in [0,1]}$ and $E = E_R$ for short. Let ξ be the convex envelope of $x \mapsto \xi_1(x/2) \wedge \xi_2(x/2)$. It is convex as a definition and still increasing and satisfies $\lim_{x\to\infty} \xi(x)/x = \infty$. We also obtain with Fenchel's inequality $\xi(|A_tB_t|) \leq \xi(\gamma(A_t) + \gamma^*(B_t)) \leq \xi(2\gamma(A_t))/2 + \xi(2\gamma^*(B_t))/2 \leq$

---

[3]Remark that without the assumption that $\int_{[0,1]} |\widetilde{V}_t(X_t)|\,dt < \infty$, R-a.s. and with the convention $e^{-\infty} = 0$, $t \in [0,1] \mapsto \exp\left(-\int_{[t,1]} \widetilde{V}_r(X_r)\,dr\right) \in [0,1]$ is well-defined R-a.s., but it might fail to be right continuous.

[4]For Doob's inequality in the class $L\log L$, see [RY99, p. 54] for instance.

$\xi_1(\gamma(A_t)) + \xi_2(\gamma^*(B_t))$ for each $t \in [0, 1]$. Consequently, $\sup_t E\xi(|A_t B_t|) < \infty$. This shows that $\{A_t B_t; t \in [0, 1]\}$ is uniformly integrable and completes the proof of the claim.

The assumption (17) will not be strong enough for our purpose. We strengthen it in the next lemma.

Lemma 3.2. *Let us assume in addition to* (12), (13) *and* (16) *that the family* $\{\gamma^*(V_t); t \in [0, 1]\}$ *is uniformly integrable in* $L^1(m)$. *Then,*

  (1) $\{\frac{1}{h}\int_{[t,t+h]} |V_s|\, ds\, |G_{t+h} - G_t|; t \in [0, 1], h > 0\}$ *is uniformly integrable in* $L^1(R)$;
  (2) $\int_{[0,1]} G_t V_t(X_t)\, dt$ *is in* $L^1(R)$;
  (3) $\{V_t G_t(X_t); t \in [0, 1]\}$ *is uniformly integrable in* $L^1(R)$.

Proof. We write $V_t = V_t(X_t)$, $\sup_t = \sup_{t \in [0,1]}$ and $E = E_R$ for short.

  ● Proof of (1). There exists a function ξ as above such that $\sup_t E\xi(\gamma^*(V_t)) < \infty$. But $E\xi\left(\frac{1}{h}\int_{[t,t+h]} \gamma^*(V_s)\, ds\right) \leq \frac{1}{h}\int_{[t,t+h]} E\xi(\gamma^*(V_s))\, ds \leq \sup_t E\xi(\gamma^*(V_t)) < \infty$. This shows that $\{\frac{1}{h}\int_{[t,t+h]} \gamma^*(V_s)\, ds; t \in [0, 1], h > 0\}$ is uniformly integrable. On the other hand, we already know by Lemma 3.1 that $\{\gamma(|G_{t+h} - G_t|); t \in [0, 1], h > 0\}$ is also uniformly integrable. The above claim permits us to conclude.

  ● Proof of (2). We see that

$$E \int_{[0,1]} G_t |V_t|\, dt \leq E(\gamma(G_*)) + E \int_{[0,1]} \gamma^*(V_t)\, dt \leq E(\gamma(G_*)) + \sup_t E\gamma^*(V_t) < \infty$$

which is finite by Lemma 3.1 and the assumption that $\{\gamma^*(V_t); t \in [0, 1]\}$ is uniformly integrable.

  ● Proof of (3). The result directly follows from the above Claim, Lemma 3.1 and our assumptions on V. $\qquad\square$

The extended Feynman-Kac generator. The main result of this section is the next theorem.

Theorem 3.3. *Let us take the following ingredients.*

  (i) $R \in \mathrm{P}(\Omega)$ *is a stationary Markov process with invariant law* $m = R_t \in \mathrm{P}(\mathcal{X})$ *for all* $t \in [0, 1]$;
  (ii) γ *is a Young function which satisfies* (16) *and* γ^* *is its convex conjugate;*
  (iii) V *is a measurable function on* $[0, 1] \times \mathcal{X}$ *which is bounded below and is such that* $\{\gamma^*(V_t); t \in [0, 1]\}$ *is uniformly integrable in* $L^1(m)$;
  (iv) g_1 *is a nonnegative function on* $\mathcal{X}$ *in* $L^\gamma(m)$.

Then, the function $g : (t, x) \in [0, 1] \times \mathcal{X} \mapsto g_t(x) \in [0, \infty)$ *which is defined for all* $t \in [0, 1]$, m-*almost everywhere by* (14):

$$g_t(X_t) := E_R \left[\exp\left(- \int_{[t,1]} V_s(X_s)\, ds \right) g_1(X_1) \mid X_{[0,t]} \right], \quad R\text{-a.s.},$$

is a nonnegative function in $L^\gamma([0, 1] \times \mathcal{X}, dt m(dx))$ *which is in* dom L^R *and in* dom $\mathcal{L}^R$. *Moreover, it satisfies*

$$L^R g(t, x) = \mathcal{L}^R g(t, x) = V_t(x) g_t(x), \quad dt m(dx)\text{-a.e.}$$

and $\int_{[0,1] \times \mathcal{X}} |V_t(x)| g_t(x)\, dt m(dx) < \infty$.

Proof. The proof is based on an application of Proposition 2.10 with $u(t, x) = g_t(x)$ and $v(t, x) = V_t(x) g_t(x)$. We write $V_t = V_t(X_t)$ and $E = E_R$ for short. We know by Lemma 3.2 that $\int_{[0,1] \times \mathcal{X}} |V_t(x)| g_t(x)\, dt m(dx) = E \int_{[0,1]} |V_t G_t|\, dt < \infty$. This

implies that $\int_{[0,1]} |V_t G_t|\, dt < \infty$, R-a.s. We have also seen at Lemma 3.1 that G_t is a right continuous uniformly integrable process. It follows that we can choose $\tau_k = \infty$ R-a.s. for all $k \geq 1$ in formula (11) and that for all $0 \leq s \leq t \leq 1$, $t \in [s,1] \mapsto E(G_t \mid X_s)$ is a right continuous real function. Therefore, to obtain the announced results, it is sufficient to show that

$$\lim_{h \downarrow 0} E \int_{[0,1-h]} \left| \frac{1}{h} E_P[G_{t+h} - G_t \mid X_t] - V_t G_t \right| dt = 0.$$

We decompose

$$-\frac{1}{h} E[G_{t+h} - G_t \mid X_t] + V_t G_t = E_P[A_h^t + B_h^t + C_h^t \mid X_t]$$

where

$$A_h^t := \frac{1}{h} \theta\left(-\int_{[t,t+h]} V_s\, ds\right) G_t$$

$$B_h^t := \frac{1}{h} \left(e^{-\int_{[t,t+h]} V_s\, ds} - 1\right)(G_{t+h} - G_t)$$

$$C_h^t := G_t \frac{1}{h} \int_{[t,t+h]} (V_t - V_s)\, ds$$

with $\theta(a) := e^a - a - 1$, $a \in \mathbb{R}$ which we already met at (15). It remains to prove that

$$\lim_{h \downarrow 0} E \int_{[0,1-h]} |A_h^t|\, dt = \lim_{h \downarrow 0} E \int_{[0,1-h]} |B_h^t|\, dt = \lim_{h \downarrow 0} E \int_{[0,1-h]} |C_h^t|\, dt = 0.$$

• Proof of $\lim_{h \downarrow 0} E \int_{[0,1-h]} |A_h^t|\, dt = 0$. We have

$$0 \leq \frac{1}{h} \theta\left(-\int_{[t,t+h]} V_s\, ds\right)$$
$$\leq \left|\left(\frac{1}{h} \int_{[t,t+h]} V_s\, ds\right)\left(e^{-\int_{[t,t+h]} V_s\, ds} - 1\right)\right| \leq \lambda_o(e^{\lambda_o h} - 1) + \frac{1}{h} \int_{[t,t+h]} |V_s|\, ds.$$

But $E\gamma^*\left(\frac{1}{h} \int_{[t,t+h]} |V_s|\, ds\right) \leq \frac{1}{h} \int_{[t,t+h]} E\gamma^*(|V_s|)\, ds \leq \sup_t E\gamma^*(|V_t|) < \infty$ by assumption. Since $\lim_{b \to \infty} \gamma^*(b)/b = \infty$ because γ doesn't grow too fast, $\{A_h^t; t \in [0,1], h > 0\}$ is uniformly integrable. This leads us to the desired convergence result since $\lim_{h \downarrow 0} A_h^t(\omega) = 0$ for $dt R(d\omega)$-almost all $(t,\omega) \in [0,1] \times \Omega$.

• Proof of $\lim_{h \downarrow 0} E \int_{[0,1-h]} |B_h^t|\, dt = 0$. Since $t \mapsto \int_{[0,t]} V_s\, ds$ is absolutely continuous and $t \mapsto G_t$ is right continuous R-a.s., we see that

$$|B_h^t| \leq e^{\lambda_o h} \frac{1}{h} \int_{[t,t+h]} |V_s|\, ds\, |G_{t+h} - G_t| \underset{h \downarrow 0}{\to} 0, \quad R\text{-a.s.}$$

On the other hand we have shown at Lemma 3.2 that $\{B_h^t; t \in [0,1], h > 0\}$ is uniformly integrable.

• Proof of $\lim_{h \downarrow 0} E \int_{[0,1-h]} |C_h^t|\, dt = 0$. We have

$$E \int_{[0,1-h]} G_t \left|\frac{1}{h} \int_{[t,t+h]} (V_t - V_s)\, ds\right| dt \leq E\left[G_* \int_{[0,1-h]} \left|V_t - \frac{1}{h} \int_{[t,t+h]} V_s\, ds\right| dt\right]$$

where we put $V_t = 0$ for all $t > 1$. By Lemma 3.1, $G_* \in L^\gamma(R)$. Therefore the measure $G_* R$ is a bounded measure and we can apply Lemma 2.7 with $v(t,\omega) = V_t(\omega)$ in $L^1([0,1] \times \Omega, G_* R)$ and $k^h = \frac{1}{h} \mathbf{1}_{[-h,0]}$. This completes the proof of the theorem.   □

4. GENERALIZED h-TRANSFORMS OF A MARKOV PROCESS

Let $R \in P(\Omega)$ be a stationary Markov process with the invariant probability measure $m \in P(\mathcal{X})$ as in Section 3. In the present section we consider the process

$$P := f_0(X_0) \exp\left(-\int_{[0,1]} V_t(X_t) \, dt\right) g_1(X_1) \, R \in P(\Omega) \tag{19}$$

where $V : [0,1] \times \mathcal{X} \to \mathbb{R}$ is a lower bounded measurable potential and

$$f_0 \in L^{\gamma^*}(m), g_1 \in L^\gamma(m), \quad f_0, g_1 \geq 0. \tag{20}$$

It is assumed once for all that

$$R(f_0(X_0)g_1(X_1) > 0) > 0$$

to discard the uninteresting trivial situation where $P = 0$. We normalize f_0 and g_1 to obtain $P(\Omega) = 1$.

Remark that $\exp\left(-\int_{[0,1]} V_t(X_t) \, dt\right)$ is bounded. It follows with the assumption (20) that $f_0(X_0) \in L^{\gamma^*}(R), g_1(X_1) \in L^\gamma(R)$ and that $f_0(X_0) \exp\left(-\int_{[0,1]} V_t(X_t) \, dt\right) g_1(X_1)$ is a nonnegative R-integrable function. Hence, it can be normalized such that P is a probability measure.

Definition 4.1 (Generalized h-transform of R). *Let $R \in P(\Omega)$ be a stationary Markov process which admits an invariant probability measure.*
A process $P \in P(\Omega)$ which is specified by formula (19) is called a generalized h-transform of R, or a generalized h-process for short.

It is not essential that R is assumed to be a stationary Markov process in this definition. Our aim is to identify P as the solution of a martingale problem. To do it, we are going to derive the extended generator $\mathcal{L}^P$ of the generalized h-process P on a class of functions $\mathcal{C}$ which is large enough to characterize P. With Theorem 2.9, we see that we are on the way to compute its stochastic derivative L^P on $\mathcal{C}$.

Playing with the Markov property. *Recall that $P \in P(\Omega)$ is a Markov process if and only if for all $t \in [0,1]$, $X_{[0,t]}$ and $X_{[t,1]}$ are independent with respect to the conditional law $P(\cdot \mid X_t)$. In other words, if and only if the past and future are independent conditionally on the present.* This property is invariant with respect to time reversal. In particular the time reversed process of R is still Markov. As with the definition of g at (14), one can define a measurable function $f_t(x)$ on $[0,1] \times \mathcal{X}$ by the formula

$$E_R\left[f_0(X_0) \exp\left(-\int_{[0,t]} V_s(X_s) \, ds\right) \mid X_{[t,1]}\right] =: f_t(X_t), \quad t \in [0,1] \quad R\text{-a.s.} \tag{21}$$

since $E_R(a \mid X_{[t,1]}) = E_R(a \mid X_t)$ for any $X_{[0,t]}$-measurable and integrable function a. As $\exp\left(-\int_{[0,t]} V_s(X_s) \, ds\right)$ is bounded and $f_0(X_0) \in L^{\gamma^*}(R)$, we see that $f_t \in L^{\gamma^*}(m)$ for all $t \in [0,1]$.

Proposition 4.2.
  (1) *The generalized h-process P is Markov.*
  (2) *For every $t \in [0,1]$, $P_t \ll m$ and*

$$\frac{dP_t}{dm} = f_t g_t \tag{22}$$

where f_t and g_t are defined respectively by (21) *and* (14) *and stand respectively in* $L^{\gamma^*}(m)$ *and* $L^\gamma(m)$.

(3) For every $0 \le s \le t \le 1$,

$$\frac{dP_{[s,t]}}{dR_{[s,t]}} = \frac{dP_s}{dm}(X_s)g_s(X_s)^{-1}\exp\left(-\int_{[s,t]} V_r(X_r)\,dr\right)g_t(X_t) \tag{23}$$

$$= f_s(X_s)\exp\left(-\int_{[s,t]} V_r(X_r)\,dr\right)f_t(X_t)^{-1}\frac{dP_t}{dm}(X_t) \tag{24}$$

$$= f_s(X_s)\exp\left(-\int_{[s,t]} V_r(X_r)\,dr\right)g_t(X_t)$$

where no division by zero occurs in the sense that $g_s > 0$, P_s-a.s. and $f_t > 0$, P_t-a.s.

Proof. • Proof of (1). Fix $0 < t < 1$ and take two bounded nonnegative functions a and b such that a is $X_{[0,t]}$-measurable and b is $X_{[t,1]}$-measurable. Let us write $\alpha = f_0(X_0)\exp\left(-\int_{[0,t]} V_s(X_s)\,ds\right) \in \sigma(X_{[0,t]})$ and $\beta = \exp\left(-\int_{[t,1]} V_s(X_s)\,ds\right)g_1(X_1) \in \sigma(X_{[t,1]})$ so that $P = \alpha\beta R$ and

$$E_P(ab \mid X_t) = \frac{E_R(ab\alpha\beta \mid X_t)}{E_R(\alpha\beta \mid X_t)} \overset{\checkmark}{=} \frac{E_R(a\alpha \mid X_t)E_R(b\beta \mid X_t)}{E_R(\alpha \mid X_t)E_R(\beta \mid X_t)} = E_P(a \mid X_t)E_P(b \mid X_t)$$

where we used the Markov property of R at the marked equality. This proves that P is Markov.

• Proof of (2) and (3). As a general result of integration theory, if $P = ZR$ with $Z \in L^1(R)$, then the push-forward $P_\phi := \phi_\# P$ of the measure P by the measurable application ϕ is absolutely continuous with respect to $R_\phi := \phi_\# R$ and $P_\phi = E_R(Z \mid \phi)\,R_\phi$ where $E_R(Z \mid \phi) := E_R(Z \mid \sigma(\phi))$ is the conditional expectation of Z with respect to the σ-field $\sigma(\phi)$ generated by ϕ. In particular, with $\phi = X_{[s,t]}$ we obtain

$$P_{[s,t]} = E_R(dP/dR \mid X_{[s,t]})\,R_{[s,t]}.$$

We have

$$E_R(dP/dR \mid X_{[s,t]})$$
$$= E_R\left[f_0(X_0)\exp\left(-\int_{[0,1]} V_r(X_r)\,dr\right)g_1(X_1) \mid X_{[s,t]}\right]$$
$$= E_R\left[f_0(X_0)\exp\left(-\left\{\int_{[0,s]}+\int_{[s,t]}+\int_{[t,1]}\right\}V_r(X_r)\,dr\right)g_1(X_1) \mid X_{[s,t]}\right]$$
$$= f_s(X_s)\exp\left(-\int_{[s,t]} V_r(X_r)\,dr\right)g_t(X_t)$$

where the Markov property of R is used at last equality. In particular, when $s = t$ this gives us (22). But with (22), we see that for all t, $f_t > 0$ and $g_t > 0$, P_t-a.s., $f_s(X_s) = g_s(X_s)^{-1}\frac{dP_s}{dm}(X_s)$ and $g_t(X_t) = f_t(X_t)^{-1}\frac{dP_t}{dm}(X_t)$. $\qquad\square$

A preliminary result under a finite entropy condition. A seemingly innocent result is proved at Proposition 4.7 below. But in fact it is a masndatory technical key to our approach. It states that, provided that the canonical process is a *nice R-semimartingale*

(see Definition 2.1), under the assumption that the *relative entropy*

$$H(P|R) := \int \log \frac{dP}{dR} \, dP < \infty$$

is finite, if a large class of regular functions stands in $\operatorname{dom} \mathcal{L}^R$, then it is also in $\operatorname{dom} \mathcal{L}^P$.

Let $\mathbf{r}$ be a probability on $D_{\mathbb{R}}$ such that the canonical process x on $D_{\mathbb{R}}$ is a nice semimartingale

$$\mathrm{x} = \mathrm{x}_0 + B + M^{\mathbf{r}}, \quad \mathbf{r}\text{-a.s.} \tag{25}$$

where B is an absolutely continuous process and $M^{\mathbf{r}}$ is a local $\mathbf{r}$-martingale. Suppose also that the quadratic variation and the jump compensator are absolutely continuous. More precisely, there exists a nonnegative adapted process a such that $\int_{[0,1]} a_t \, dt < \infty$, $\mathbf{r}$-a.s. and

$$d[\mathrm{x}, \mathrm{x}]_t^c = a_t \, dt, \quad \mathbf{r}\text{-a.s.}$$

and the dual predictable projection $\overline{\ell}$ of the jump measure $\sum_{0 \le s \le t} \delta_{(s, \Delta \mathrm{x}_s)}$ has the following form

$$\overline{\ell}_t(dtdq) = dt\ell_t(dq), \quad \mathbf{r}\text{-a.s.}$$

This means that $\ell_t = \ell(t, \mathrm{x}_{[0,t)}; \cdot)$ is a predictable nonnegative measure on $\mathbb{R}_* := \mathbb{R} \setminus \{0\}$ such that

$$\sum_{0 \le s \le t} f(s, \mathrm{x}_{[0,s)}; \Delta \mathrm{x}_s) = \int_{[0,t] \times \mathbb{R}_*} f(s, \mathrm{x}_{[0,s)}; q) \, ds\ell_s(dq) + M_t^f$$

where M^f is a local $\mathbf{r}$-martingale and this decomposition is valid for any measurable function f such that $\int_{[0,1] \times \mathbb{R}_*} |f(t, \mathrm{x}_{[0,t)}; q)| \, dt\ell_t(dq) < \infty$, $\mathbf{r}$-a.s.

It is also assumed that

$$\int_{[0,1] \times \mathbb{R}_*} \theta(\alpha|q|) \, dt\ell_t(dq) < \infty, \ \forall \alpha \ge 0 \quad \mathbf{r}\text{-a.s.} \tag{26}$$

where $\theta(a) := e^a - a - 1$, $a \in \mathbb{R}$ already appeared at (15).

Lemma 4.3. *Let $\mathbf{r}$ be as above and $\mathbf{p}$ be a probability on $D_{\mathbb{R}}$ such that $H(\mathbf{p}|\mathbf{r}) < \infty$. Then, x is also a nice $\mathbf{p}$-semimartingale.*

Remarks 4.4.
  (1) Girsanov's theorem tells us that if x is an $\mathbf{r}$-semimartingale and $\mathbf{p} \ll \mathbf{r}$, then x is also a $\mathbf{p}$-semimartingale. This lemma tells us that the property of being a *nice* semimartingale is also hereditary under the stronger condition that $H(\mathbf{p}|\mathbf{r}) < \infty$.
  (2) In case when no jump occurs and the $\mathbf{r}$-semimartingale is built on a Brownian filtration, it is well-known that Lemma 4.3 is still valid with the weaker assumption that $\mathbf{p} \ll \mathbf{r}$ instead of $H(\mathbf{p}|\mathbf{r}) < \infty$. This follows from Girsanov's theorem and a martingale representation theorem.
  (3) The assumption $H(\mathbf{p}|\mathbf{r}) < \infty$ is not very restrictive. Indeed, $\mathbf{p} \ll \mathbf{r}$ means that $d\mathbf{p}/d\mathbf{r} \in L^1(\mathbf{r})$, while $H(\mathbf{p}|\mathbf{r}) < \infty$ means that $(d\mathbf{p}/d\mathbf{r}) \log_+ (d\mathbf{p}/d\mathbf{r}) \in L^1(\mathbf{r})$.
  (4) For more details about extensions of this result, see [Léo12].

Proof. The proof is based on the variational representation

$$H(\mathbf{p}|\mathbf{r}) = \sup\{E_{\mathbf{p}}u - \log E_{\mathbf{r}}e^u; u \text{ measurable} : E_{\mathbf{r}}e^u < \infty\} \tag{27}$$

of the relative entropy which holds true for any probability measure $\mathbf{p}$ such that $H(\mathbf{p}|\mathbf{r})$ is finite, see for instance [Léo12, Lemma 3.1] for a proof.

Let h belong to the space $\mathcal{S}$ of all simple predictable processes:

$$h_t = h_0 \mathbf{1}_{\{0\}}(t) + \sum_{i=1}^{k} h_i \mathbf{1}_{(T_i, T_{i+1}]}(t)$$

with k a finite integer, $h_i \in \sigma(\mathrm{x}_{[0,T_i)})$, $|h_i| < \infty$ and $0 \leq T_1 \leq \cdots \leq T_{k+1} = 1$ an increasing sequence of stopping times. Its stochastic integral with respect to $M^{\mathbf{r}}$ is $h \cdot M_t^{\mathbf{r}} = \sum_{i=1}^{k} h_i(M_{T_{i+1} \wedge t}^{\mathbf{r}} - M_{T_i \wedge t}^{\mathbf{r}})$, $t \in [0,1]$ and the stochastic exponential of $h \cdot M^{\mathbf{r}}$ is

$$\mathcal{E}(h \cdot M^{\mathbf{r}})_t = \exp\left(h \cdot M_t^{\mathbf{r}} - \int_{[0,t]} \frac{h_s^2}{2} a_s \, ds - \int_{[0,t] \times \mathbb{R}_*} \theta(h_s q) \, ds \ell_s(dq) \right)$$

Under the assumption (26), the integrals in the exponential are finite $\mathbf{r}$-a.s. so that the sequence of stopping times $\inf\{t \in [0,1]; \int_{[0,t]} \frac{h_s^2}{2} a_s \, ds + \int_{[0,t] \times \mathbb{R}_*} \theta(h_s q) \, ds \ell_s(dq) \geq k\}$ tends to infinity $\mathbf{r}$-a.s. as k tends to infinity. It follows that $\mathcal{E}(h \cdot M^{\mathbf{r}})$ is a positive supermartingale and in particular that: $E_{\mathbf{r}} \mathcal{E}(h \cdot M^{\mathbf{r}})_1 \leq 1$. Therefore, for any h in $\mathcal{S}$, $\log E_{\mathbf{r}} \mathcal{E}(h \cdot M_1^{\mathbf{r}}) \leq 0$ and with (27) we obtain

$$\begin{aligned} E_{\mathbf{p}}(h \cdot M_1^{\mathbf{r}}) &\leq H(\mathbf{p}|\mathbf{r}) + E_{\mathbf{p}}\left(\int_{[0,1]} \frac{h_t^2}{2} a_t \, dt + \int_{[0,1] \times \mathbb{R}_*} \theta(h_t q) \, dt \ell_t(dq) \right) \\ &\leq H(\mathbf{p}|\mathbf{r}) + \int_{[0,1] \times D_{\mathbb{R}}} \Phi(t, \eta; h_t(\eta)) \, \overline{\mathbf{p}}(dt d\eta) \end{aligned}$$

where

$$\overline{\mathbf{p}}(dt d\eta) = dt \mathbf{p}(d\eta)$$

and for all $t \in [0,1], \eta \in D_{\mathbb{R}}, x \in \mathbb{R}$,

$$\Phi(t, \eta; x) := a_t(\eta) x^2 / 2 + \int_{\mathbb{R}_*} \theta(|qx|) \ell(t, \eta; dq).$$

A standard convexity argument (note that $\theta(|x|)$ is a convex nonnegative even function) proves that the gauge functional

$$|h|_{\mathbf{p}} := \inf\left\{ \alpha > 0; \int_{[0,1] \times D_{\mathbb{R}}} \Phi(t, \eta; h_t(\eta)/\alpha) \, \overline{\mathbf{p}}(dt d\eta) \leq 1 \right\} \in [0, \infty)$$

is a seminorm on $\mathcal{S}$. Considering $h/|h|_{\mathbf{p}}$ and $-h/|h|_{\mathbf{p}}$ in the above inequality, it is easy to deduce that

$$|E_{\mathbf{p}}(h \cdot M_1^{\mathbf{r}})| \leq (H(\mathbf{p}|\mathbf{r}) + 1)|h|_{\mathbf{p}}, \quad \forall h \in \mathcal{S}.$$

This means that if $H(\mathbf{p}|\mathbf{r}) < \infty$, $h \mapsto E_{\mathbf{p}}(h \cdot M_1^{\mathbf{r}})$ is a $|\cdot|_{\mathbf{p}}$-continuous linear form on $\mathcal{S}$. But, $|\cdot|_{\mathbf{p}}$ is the seminorm of an Orlicz space and by assumption (26), $\int \Phi(ah) \, d\overline{\mathbf{p}} < \infty$ for all $a \geq 0$ and $h \in \mathcal{S}$. This implies that $\mathcal{S}$ is a subspace of the "small" Orlicz space $S^{\Phi}(\overline{\mathbf{p}}) := \{f : [0,1] \times D_{\mathbb{R}} \to \mathbb{R}, \text{ measurable}, \int \Phi(t, \eta; af_t(\eta)) \, \overline{\mathbf{p}}(dt d\eta) < \infty, \forall a \geq 0\}$ whose dual representation is well-known, see [RR91]: There exists a measurable function k on $[0,1] \times D_{\mathbb{R}}$ which stands in the "large" Orlicz space $\{k : [0,1] \times D_{\mathbb{R}} \to \mathbb{R}, \text{ measurable}, \int \Phi^*(t, \eta; a_o k_t(\eta)) \, \overline{\mathbf{p}}(dt d\eta) < \infty, \text{ for some } a_o > 0\} =: L^{\Phi^*}(\overline{\mathbf{p}}) \subset L^1(\overline{\mathbf{p}})$ associated with the convex conjugates $\Phi^*(t, \eta; \cdot)$ of $\Phi(t, \eta; \cdot)$, such that

$$E_{\mathbf{p}}(h \cdot M_1^{\mathbf{r}}) = \int_{[0,1] \times D_{\mathbb{R}}} k_t(\eta) h_t(\eta) \, \overline{\mathbf{p}}(dt d\eta), \quad \forall h \in \mathcal{S}. \tag{28}$$

Since h is predictable, we also have $\int kh\,d\overline{\mathbf{p}} = \int_{[0,1]} E_{\mathbf{p}}(k_t h_t)\,dt = E_{\mathbf{p}}\int_{[0,1]} E_{\mathbf{p}}(k_t \mid \mathrm{x}_{[0,t)})h_t\,dt$ and taking $\tilde{b}_t = E_{\mathbf{p}}(k_t \mid \mathrm{x}_{[0,t)})$ we see with (28) that

$$E_{\mathbf{p}}\left(\int_{[0,1]} h_t\,dM_t^{\mathbf{r}} - \int_{[0,1]} h_t\tilde{b}_t\,dt\right) = 0, \quad \forall h \in \mathcal{S}.$$

It follows that $M_t^{\mathbf{P}} := M_t^{\mathbf{r}} - \widetilde{B}_t$ with $\widetilde{B}_t := \int_{[0,t]} \tilde{b}_s\,ds$ is a local $\mathbf{p}$-martingale and with (25) we finally obtain that

$$\mathrm{x} = \mathrm{x}_0 + B + \widetilde{B} + M^{\mathbf{P}}, \quad \mathbf{p}\text{-a.s.}$$

where $B + \widetilde{B}$ has absolutely continuous sample paths $\mathbf{p}$-a.s. □

Let us go back to R and P given at (19).

Definition 4.5 (The class $\mathcal{U}_R$). *Let the reference Markov process R be given. We say that the measurable function $u : [0,1] \times \mathcal{X} \to \mathbb{R}$ is in the class $\mathcal{U}_R$ (with respect to R): $u \in \mathcal{U}_R$, if*

(a) $u \in \operatorname{dom}\mathcal{L}^R$;
(b) $d[u_t(X_t), u_t(X_t)]^c \ll dt$, R-a.s.;
(c) the predictable dual projection $\overline{\ell^u}$ of $\sum_{t\in[0,1]} \delta_{(t,\Delta u_t(X_t))}$ satisfies $\overline{\ell^u}(dtdq) = dt\ell_t^u(dq)$ and $\int_{[0,1]\times\mathbb{R}_*} \theta(\alpha|q|)\,dt\ell_t^u(dq) < \infty$ for all $\alpha \geq 0$, R-a.s.

In other words, $u \in \mathcal{U}_R$ if the process $u(t, X_t)$ is a R-semimartingale and its law $\mathbf{r} \in \mathrm{P}(D_{\mathbb{R}})$ meets the assumptions of Lemma 4.3.

Remark 4.6. For the class $\mathcal{U}_R$, we have in mind $\mathcal{C}_c^{1,2}([0,1] \times \mathbb{R}^d)$ when R is such that the canonical process X is a *nice R-semimartingale* with its values in $\mathcal{X} = \mathbb{R}^d$. Indeed, at least in the continuous case when no exponential moments of ℓ^u are required, Itô's formula immediately implies that $\mathcal{C}_c^{1,2}([0,1] \times \mathbb{R}^d) \subset \mathcal{U}_R$.
Otherwise, if X is not a nice R-semimartingale, then it might happen that $\mathcal{U}_R$ reduces to the constant functions.

A useful result is the following

Proposition 4.7. *Let us assume that $H(P|R) < \infty$. Then any $u \in \mathcal{U}_R$ is also in $\operatorname{dom}\mathcal{L}^P$.*

Proof. Let $\Psi : \Omega \to D_{\mathbb{R}}$ be the application $\Psi = (u_t(X_t))_{t\in[0,1]}$. The measure $\mathbf{r} = \Psi_{\#}R \in \mathrm{P}(D_{\mathbb{R}})$ is the law of the process $(u_t(X_t))_{t\in[0,1]}$ when the canonical process is governed by $R \in \mathrm{P}(\Omega)$. By the definition of the class $\mathcal{U}_R$, $\mathbf{r}$ satisfies the assumptions of Lemma 4.3. Let $\mathbf{p} = \Psi_{\#}P$ be the law of $(u_t(X_t))_{t\in[0,1]}$ under $P \in \mathrm{P}(\Omega)$. By contraction of the relative entropy (an easy consequence of (27)), we have $H(\mathbf{p}|\mathbf{r}) = H(\Psi_{\#}P|\Psi_{\#}R) \leq H(P|R) < \infty$. This is the second assumption of Lemma 4.3, and this lemma tells us that $(u_t(X_t))_{t\in[0,1]}$ is a nice P-semimartingale, i.e. $u \in \operatorname{dom}\mathcal{L}^P$. □

Lemma 4.8. *Let $P \in \mathrm{P}(\Omega)$ be specified by (19) with $\inf V > -\infty$ and $f_0, g_1 \geq 0$. Then, for $H(P|R) < \infty$, it is sufficient that $\int_{\mathcal{X}} f_0^2 \log_+^p(f_0)\,dm < \infty$ and $\int_{\mathcal{X}} g_1^2 \log_+^p(g_1)\,dm < \infty$ for some $p > 1$.*

Proof. Since m and R are bounded positive measures, only the large values of the functions are important as regards integrability issues. As $\exp(-\int_{[0,1]} V_t\,dt)$ is bounded, all we have to show is that if two nonnegative functions $F = f_0(X_0)$ and $G = g_1(X_1)$ satisfy $\int F^2 \log_+^p(F)\,dR < \infty$ and $\int G^2 \log_+^p(G)\,dR < \infty$ with $p > 1$, then $\int FG \log_+(FG)\,dR < \infty$.

For all $x, y \geq 0$, we have $xy \leq x^2 \log_+ x$ when $y \leq x \log_+ x$ and in the alternate case when

$y \geq x \log_+ x$, we see that for any $0 < q < 1$ and $y \geq y_q$ large enough, $x \leq y(\log_+ y)^{-q}$. Hence,

$$xy \leq x^2 \log_+ x + y^2 (\log_+ y)^{-q}, \quad \forall x \geq 0, y \geq y_q.$$

Now, for F, G large enough we have

$$
\begin{aligned}
FG \log_+(FG) &\leq (F \log_+ F)G + FG \log_+ G \\
&\leq F^2 \log_+ F + G^2 \log_+ G + \frac{(F \log_+ F)^2}{\log_+^q(F \log_+ F)} + \frac{(G \log_+ G)^2}{\log_+^q(G \log_+ G)} \\
&\leq 2F^2 \log_+^{2-q} F + 2G^2 \log_+^{2-q} G
\end{aligned}
$$

which completes the proof of the lemma. $\square$

Let $\chi(a)$ be a Young function, then

$$\gamma_f(a) := \chi(|a| \log_+ |a|) \quad \text{and} \quad \gamma_g(b) := \chi^*(|b| \log_+ |b|) \tag{29}$$

are also Young functions. Clearly, $ab \log_+(ab) \leq (a \log_+ a)b + a(b \log_+ b) \leq 2[\gamma_f(a) + \gamma_g(b)]$ for any large enough positive numbers a, b. Therefore, if $f \in L^{\gamma_f}$ and $g \in L^{\gamma_g}$, then $fg \in L \log L$.

Gathering our last results leads us to the following statement.

Theorem 4.9. *Let P be the generalized h-process given at (19) with $\inf V > -\infty$ and the functions f_0 and g_1 such that one of the following conditions is satisfied:*

(i) *$f_0 \in L^{\gamma_f}(m)$ and $g_1 \in L^{\gamma_g}(m)$ where γ_f and γ_g satisfy (29);*
(ii) *$\int_\mathcal{X} f_0^2 \log_+^p(f_0) \, dm < \infty$ and $\int_\mathcal{X} g_1^2 \log_+^p(g_1) \, dm < \infty$ for some $p > 1$.*

Then, $H(P|R) < \infty$ and any function $u \in \operatorname{dom} \mathcal{L}^R$ which is in the class $\mathcal{U}_R$ is also in the extended domain $\operatorname{dom} \mathcal{L}^P$ associated with P.

As particular cases of condition (i) above, we have $f_0 \in L^\infty(m), g_1 \in L \log L(m)$ and $f_0 \in L \log L(m), g_1 \in L^\infty(m)$. Condition (ii) is a slight improvement of condition (i) with $\chi(x) = x^2$.

The stochastic derivative of P. Let us start saying some words about the carré du champ operator Γ^R of a Markov process P. It is a general result of the theory of stochastic processes that the product of two real semimartingales is still a semimartingale. More precisely, if Y and Z are semimartingales, then

$$YZ = YZ_0 + \int Y_- dZ + \int Z_- dY + [Y, Z]$$

where $[Y, Z]_t$ is the limit along refining finite partitions of the time interval by means of stopping times: $0 \leq T_1 \leq \cdots \leq T_k = 1$, of the cross variation $\sum_i (Y_{T_{i+1} \wedge t} - Y_{T_i \wedge t})(Z_{T_{i+1} \wedge t} - Z_{T_i \wedge t})$. It is a remarkable result that $[Y, Z]$ is again a semimartingale. Its compensator is denoted by $\langle Y, Z \rangle$, this means that

$$[Y, Z] = \langle Y, Z \rangle + M^{Y,Z}$$

where $\langle Y, Z \rangle$ is a predictable bounded variation process and $M^{Y,Z}$ is a local martingale. Nevertheless, the product of two *nice* semimartingales might not be nice anymore. Let Y and Z be nice. Clearly, the stochastic integrals $\int Y_- dZ$ and $\int Z_- dY$ are nice so that YZ is nice if and only if $\langle Y, Z \rangle$ is absolutely continuous.

Definition 4.10 (Carré du champ operator). *Let u and v be two measurable real functions on $[0,1] \times \mathcal{X}$. Going back to the canonical process X on Ω, suppose that the processes $u(X) = (u_t(X_t))_{t \in [0,1]}$ and $v(X) = (v_t(X_t))_{t \in [0,1]}$ are P-semimartingales such that $\langle u(X), v(X) \rangle$ is absolutely continuous P-a.s. Then, we say that the couple of functions (u,v) is in the domain $\mathrm{dom}\,\Gamma^P$ of the carré du champ operator Γ^P which is defined by*

$$d\langle u(X), v(X) \rangle_t =: \Gamma^R(u,v)(t, X_{t^-})\, dt, \quad P\text{-a.s.}$$

This identity determines the function $(t,x) \in [0,1] \times \mathcal{X} \mapsto \Gamma^P(u,v)(t,x) \in \mathbb{R}$, $dtP_t(dx)$-almost everywhere.

As a direct consequence of this definition, we obtain the following result which is often used as a definition of Γ^P.

Proposition 4.11. *Let u and v be two continuous functions on $[0,1] \times \mathcal{X}$ such that u,v and their product uv belong to $\mathrm{dom}\,\mathcal{L}^P$. Then, $(u,v) \in \mathrm{dom}\,\Gamma^P$ and*

$$\Gamma^P(u,v) = \mathcal{L}^P(uv) - u\mathcal{L}^P v - v\mathcal{L}^P u.$$

Proof. We denote $U_t = u(t, X_t)$ and $V_t = v(t, X_t)$. By hypothesis, we have $dU_t = \mathcal{L}^P u(t, X_t)\, dt + dM_t^u$, $dV_t = \mathcal{L}^P v(t, X_t)\, dt + dM_t^v$ and $d(UV)_t = \mathcal{L}^P(uv)(t, X_t)\, dt + dM_t^{uv}$ where M stands for any local R-martingale. Therefore,

$$
\begin{aligned}
d[U,V]_t &= d(UV)_t - U_t dV_t - V_t dU_t \\
&= [\mathcal{L}^P(uv) - u\mathcal{L}^P(v) - v\mathcal{L}^P(u)](t, X_{t^-})\, dt + dM_t
\end{aligned}
$$

with $dM_t = dM_t^{uv} - U_t dM_t^v - V_t dM_t^u$. Hence, $d\langle U,V \rangle_t = [\mathcal{L}^P(uv) - u\mathcal{L}^P v - v\mathcal{L}^P u](t, X_{t^-})\, dt$, which is the announced result. $\square$

There are no tractable general conditions on P which imply that $\langle u(X), v(X) \rangle$ is absolutely continuous P-a.s. whenever $u,v \in \mathrm{dom}\,\mathcal{L}^P$. Counterexamples are known, see [Mok89]; $u,v \in \mathrm{dom}\,\mathcal{L}^P$ doesn't imply in general that $(u,v) \in \mathrm{dom}\,\Gamma^P$. Some additional assumptions are needed.

Theorem 4.12. *Let the h-process P and the function $g_t(x)$ be defined by (19) and (14). Let the hypotheses of Theorem 3.3 and Proposition 4.7 be satisfied:*

(i) $R \in \mathrm{P}(\Omega)$ is a stationary Markov process with invariant law $m = R_t \in \mathrm{P}(\mathcal{X})$ for all $t \in [0,1]$;

(ii) γ is a Young function which satisfies (16) and γ^* is its convex conjugate;

(iii) V is a measurable function on $[0,1] \times \mathcal{X}$ which is bounded below and is such that $\{\gamma^*(V_t); t \in [0,1]\}$ is uniformly integrable in $L^1(m)$;

(iv) g_1 is a nonnegative function on $\mathcal{X}$ in $L^\gamma(m)$.

We also assume that f_0 and g_1 satisfy the hypotheses of Theorem 4.9 to insure that $H(P|R) < \infty$.

Then, $\mathcal{U}_R \subset \mathrm{dom}\,\mathcal{L}^P \subset \mathrm{dom}\,L^P$ and for all $u \in \mathcal{U}_R$ which satisfies for almost all $t \in [0,1)$ and m-almost all x,

$$\sup_{s \in [t, t+h_o]} E_R(|\mathcal{L}^R u_s|^p \mid X_t = x) < \infty, \quad \text{for some } h_o > 0 \text{ and } p > 1, \tag{30}$$

we have

$$(g, u) \in \mathrm{dom}\,\Gamma^R$$

and

$$\mathcal{L}^P u(t,x) = L^P u(t,x) = L^R u(t,x) + \frac{\Gamma^R(g,u)(t,x)}{g_t(x)}, \quad dtP_t(dx)\text{-a.e.}$$

where no division by zero occurs since $g_t > 0$, P_t-a.s.

Proof. Let u be in $\mathcal{U}_R$, then we know by Theorems 2.9 and 4.9 that

$$u \in \mathrm{dom}\, \mathcal{L}^P \subset \mathrm{dom}\, L^P. \tag{31}$$

With (23) we see that for all $0 \le t < t+h \le 1$ and P_t-almost all x,

$$E_P(u_{t+h}(X_{t+h}) - u_t(x) \mid X_t = x)$$
$$= E_R\left(g_t(x)^{-1}[u_{t+h}(X_{t+h}) - u_t(x)] \exp\left(-\int_{[t,t+h]} V_r(X_r)\, dr\right) g_{t+h}(X_{t+h}) \mid X_t = x\right)$$

with $g_t(x) > 0$, $P_t(dx)$-a.s. We write for simplicity $u_s(X_s) = U_s$, $g_s(X_s) = G_s$, $V_s = V_s(X_s)$, $D^h U_t = u_{t+h}(X_{t+h}) - u_t(x)$, $D^h G_t = g_{t+h}(X_{t+h}) - g_t(x)$ and $D^h F_t = \int_{[t,t+h]} V_r(X_r)\, dr$. The inner term in the right-hand side expectation is

$$g_t(x)^{-1} D^h U_t e^{-D^h F_t} G_{t+h}$$
$$= D^h U_t(1 + [e^{-D^h F_t} - 1])(1 + D^h G_t / g_t(x))$$
$$= D^h U_t + D^h U_t D^h G_t / g_t(x) + [e^{-D^h F_t} - 1][D^h U_t + D^h U_t D^h G_t / g_t(x)]. \tag{32}$$

As it is assumed that $u \in \mathrm{dom}\, \mathcal{L}^R$, $(U_r)_{r \in [0,1]}$ is a R-semimartingale. Since its sample paths are in $D_{\mathbb{R}}$, they are bounded R-a.s. and the sequence of stopping times $\inf\{r \in [0,1]; |U_r| + G_r \ge k\}$ converges R-a.s. to infinity. Therefore, we can assume without loss of generality that U and G are bounded without introducing integration times.

The contribution of the first term $D^h U_t$ of (32) is well understood. Since $u \in \mathrm{dom}\, L^R$, we have

$$\lim_{h \downarrow 0} \frac{1}{h} E_R^x D^h u_t = L^R u(t, x) \tag{33}$$

where we denote $E_R^x = E_R(\cdot \mid X_t = x)$ for simplicity.

Let us control, the last term of (32). As G and U can be assumed to be bounded, DU_t and $DG_t DU_t$ are also bounded. Hence, $DU_t + DG_t DU_t / g_t(x)$ is bounded and by right continuity of the sample paths, it tends to zero R-a.s.. By dominated convergence, we obtain

$$\lim_{h \downarrow 0} E_R^x \gamma(DU_t + DG_t DU_t / g_t(x)) = 0.$$

On the other hand, $|e^{-D^h F_t} - 1| = |[e^{-D^h F_t} - 1]/(-D^h F_t)|\, |D^h F_t| \le e^{\lambda_o h}|\int_{[t,t+h]} V_r\, dr|$ and

$$E_R^x \gamma^*\left(\frac{1}{h}[e^{-D^h F_t} - 1]\right) \le c_{\gamma^*,\lambda_o} E_R^x \gamma^*\left(\frac{1}{h}\int_{[t,t+h]} V_r\, dr\right) \le c_{\gamma^*,\lambda_o} E_R^x \frac{1}{h}\int_{[t,t+h]} \gamma^*(V_r)\, dr$$
$$= c_{\gamma^*,\lambda_o} \frac{1}{h}\int_{[t,t+h]} E_R^x \gamma^*(V_r)\, dr \le c_{\gamma^*,\lambda_o} \sup_{r \in [0,1]} E_R^x \gamma^*(V_r) < \infty.$$

It follows with Hölder's inequality that

$$\lim_{h \downarrow 0} \frac{1}{h} E_R^x \left|[e^{-D^h F_t} - 1][D^h U_t + D^h U_t D^h G_t / g_t(x)]\right|$$
$$\le 2 \lim_{h \downarrow 0} \left\|\frac{1}{h}[e^{-D^h F_t} - 1]\right\|_{L^{\gamma^*}(R^x)} \left\|DU_t + DG_t DU_t / g_t(x)\right\|_{L^\gamma(R^x)} = 0. \tag{34}$$

Let us look at $D^h U_t D^h G_t$ coming from the second term of (32). By means of basic stochastic calculus we arrive at

$$DG_t DU_t = \int_{[t,t+h]} (G_r - G_t)\, dU_r + \int_{[t,t+h]} (U_r - U_t)\, dG_r + [G, U]_{t+h} - [G, U]_t.$$

With $dU_r = \mathcal{L}^R u_r \, dr + dM_r^u$ and $dG_r = \mathcal{L}^R g_r \, dr + dM_r^g = V_r G_r \, dr + dM_r^g$ where we relied on Theorem 3.3 in last equality, taking the expectation leads us to

$$E_R^x(DG_t DU_t)$$
$$= \underbrace{E_R^x \int_{[t,t+h]} (G_r - G_t) \, \mathcal{L}^R u_r \, dr}_{A_h} + \underbrace{E_R^x \int_{[t,t+h]} (U_r - U_t) \, V_r G_r \, dr}_{B_h} + \underbrace{E_R^x([G,U]_{t+h} - [G,U]_t)}_{C_h}.$$

Let us control A_h, B_h and C_h. By Hölder's inequality with $1/p + 1/q$ and $q \geq 1$,

$$A_h \leq \left(E_R^x \int_{[t,t+h]} |G_r - G_t|^q \, dr \right)^{1/q} \left(E_R^x \int_{[t,t+h]} |\mathcal{L}^R u_r|^p \, dr \right)^{1/p}.$$

But $E_R^x \int_{[t,t+h]} |G_r - G_t|^q \, dr = o(h)$ since $\{G_r; r \in [0,1]\}$ is bounded and G is right continuous. We also obtain, $E_R^x \int_{[t,t+h]} |\mathcal{L}^R u_r|^p \, dr = \int_{[t,t+h]} E_R^x |\mathcal{L}^R u_r|^p \, dr = O(h)$, under the condition that (30) holds. It follows that $A_h = o(h)^{1/q} O(h)^{1/p} = o(h)$.

Let us control B_h. We can take U bounded and we already know by Lemma 3.2 that $\{V_t G_t; t \in [0,1]\}$ is uniformly integrable. Since U is right continuous, it follows that $B_h = o(h)$.

We know by (31) that the limit

$$\lim_{h \downarrow 0} \frac{1}{h} E_R^x \big\{ D^h U_t + D^h U_t D^h G_t / g_t(x)$$
$$+ [e^{-D^h F_t} - 1][D^h U_t + D^h U_t D^h G_t / g_t(x)] \big\} =: L^P u(t,x)$$

exists. We have also shown (33) and (34) which imply that, $dt P_t(dx)$-a.e. :

$$L^P u(t,x) = L^R u(t,x) + g_t(x)^{-1} \lim_{h \downarrow 0} \frac{1}{h} C_h$$
$$= L^R u(t,x) + g_t(x)^{-1} \lim_{h \downarrow 0} \frac{1}{h} E_R^x([G,U]_{t+h} - [G,U]_t)$$
$$= L^R u(t,x) + g_t(x)^{-1} \lim_{h \downarrow 0} \frac{1}{h} E_R^x(\langle G,U \rangle_{t+h} - \langle G,U \rangle_t)$$

and in particular that the limit $\lim_{h \downarrow 0} \frac{1}{h} E_R^x(\langle G,U \rangle_{t+h} - \langle G,U \rangle_t)$ exists. Since this is true for all t and x, this shows that (g,u) belongs to the domain of Γ^R. We conclude noticing that by definition $\lim_{h \downarrow 0} \frac{1}{h} E_R^x(\langle G,U \rangle_{t+h} - \langle G,U \rangle_t) = \Gamma^R(g,u)(t,x)$. $\square$

We note for future use the following result.

Corollary 4.13. *Under the assumptions of Theorem 4.12, we have*

$$\Gamma^R(g,u)(t,x) = \lim_{h \to \infty} \frac{1}{h} E_R \Big([g_{t+h}(X_{t+h}) - g_t(x)][u_{t+h}(X_{t+h}) - u_t(x)] | X_t = x \Big), \; dtm(dx)\text{-a.e.}$$
(35)

The product ug is in $\mathrm{dom}\,\mathcal{L}^R$ and $\Gamma^R(g,u) = \mathcal{L}^R(gu) - g\mathcal{L}^R u - u\mathcal{L}^R g$.

Proof. The identity (35) has been proved during the previous proof of Theorem 4.12. Next assertion follows from $D(GU) = DGDU + UDG + GDU$ and the convergences which are implied by $u,g \in \mathrm{dom}\,\mathcal{L}^R$ and $(g,u) \in \mathrm{dom}\,\Gamma^R$. $\square$

Remark 4.14. Let us also remark that applying Lemma 4.3 to the quadratic variation $[u(X)]$, under the assumption $H(P|R) < \infty$ we see that $d\langle u(X)\rangle_t^R \ll dt$, R-a.s. implies that $d\langle u(X)\rangle_t^P \ll dt$, P-a.s. It follows that

$$\operatorname{dom}\Gamma^P \subset \operatorname{dom}\Gamma^R$$

in the sense that we consider $dtP_t(dx)$-a.e.-defined functions instead of $dtm(dx)$-a.e.-defined functions.

In the special case when X is continuous R-a.s., we also have $\langle u(X)\rangle^P = \langle u(X)\rangle^R$, P-a.s., which implies that $\Gamma^P(u,v)(t,x) = \Gamma^R(u,v)(t,x)$, $dtP_t(dx)$-a.e., $(u,v) \in \operatorname{dom}\Gamma^P$.

5. CONTINUOUS DIFFUSION PROCESSES ON $\mathbb{R}^d$

In this section we examplify the previous abstract results with simple continuous diffusion processes on $\mathbb{R}^d$.

The reference process R. The reference process R is the law of a Markov continuous diffusion process on the state space $\mathcal{X} = \mathbb{R}^d$ which admits an invariant probability measure m. To fix the ideas, we assume in the whole section that it is the solution of the stochastic differential equation (SDE)

$$X_t = X_0 + \int_{[0,t]} b(X_s)\,ds + \int_{[0,t]} \sigma(X_s)\,dW_s, \quad t \in [0,1]$$

where W is a $\mathbb{R}^d$-valued Wiener process, $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to M_{d\times d}$ are locally Lipschitz functions which are respectively vector-valued and matrix-valued. We also assume that R-a.s., X doesn't explode on the time interval $[0,1]$.

Result 5.1. *Under these hypotheses on R, it is known that R is the unique solution of the martingale problem $\mathrm{MP}(\mathcal{L},\mathcal{C};\mu_o)$ in the sense of Definition 2.4 with the initial measure $\mu_o = m$ and the generator $\mathcal{L}^R$ given for all $u \in \mathcal{C} = \mathcal{C}_c^{1,2}([0,1]\times\mathbb{R}^d)$ by*

$$\mathcal{L}^R u(t,x) = \partial_t u(t,x) + \sum_{i=1}^d b_i(x)\partial_{x_i}u(t,x) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial_{x_i}\partial_{x_j}u(t,x)$$

where $(a_{ij})_{1\le i,j\le d} = a := \sigma\sigma^* \in M_{d\times d}$.
We denote this martingale problem $\mathrm{MP}(b,a;m)$.

Extended gradients. We introduce the notion of extended gradient. Let P be a solution to the martingale problem $\mathrm{MP}(b^P, a; P_0)$, for some drift vector field $b^P : [0,1]\times\mathbb{R}^d \to \mathbb{R}^d$. A simple computation based on Proposition 4.11 gives us

$$\Gamma^P(\varphi,v)(t,x) = \nabla\varphi_t(x)\cdot a(x)\nabla v(x), \ dtP_t(dx)\text{-a.e.}, \ \varphi \in \mathcal{C}_c^{1,2}([0,1]\times\mathbb{R}^d), v \in \mathcal{C}_c^2(\mathbb{R}^d). \quad (36)$$

One proves the Cauchy-Schwarz type inequality

$$\left(\int_{[s,t]} d\langle A,B\rangle_r\right)^2 \le \int_{[s,t]} d\langle A,A\rangle_r \int_{[s,t]} d\langle B,B\rangle_r, \quad 0 \le s \le t \le 1$$

with the usual discriminent argument. Let us take u,v in $\mathcal{C}_c^2(\mathbb{R}^d)$ and ψ a measurable function on $[0,1]\times\mathcal{X}$ such that (ψ,u) and (ψ,v) are in $\operatorname{dom}\Gamma^P$ and such that $\Gamma^P(v-u,v-u) = 0$. Then, the above Cauchy-Schwarz inequality implies that $\Gamma^P(\psi_t,u)(t,x) = \Gamma^P(\psi_t,v)(t,x)$, $dtP_t(dx)$-a.e. Consequently, the linear operator $u \mapsto \Gamma^P(\psi,u)$ only depends on the equivalence class defined by $u \sim v \overset{\mathrm{def}}{\Leftrightarrow} \Gamma^P(v-u,v-u) = 0, dtP_t(dx)$-a.e. $\Leftrightarrow$

$a \cdot \nabla(v - u) = 0, dt P_t(dx)$-a.e., and it follows that there exists some vector field β on $[0,1] \times \mathbb{R}^d$ such that $\Gamma^P(\psi, \cdot)$ is represented by

$$\Gamma^P(\psi, v)(t, x) = \beta_t(x) \cdot a(x) \nabla v(x), \quad dt P_t(dx)\text{-a.e.,} \quad v \in \mathcal{C}^2(\mathbb{R}^d). \tag{37}$$

Moreover, up to $dt P_t(dx)$-a.e. equality, there is a unique such β with its values in the range of a.

Comparing (36) and (37), it is natural to introduce the following definition.

Definition 5.2 (Extended gradient). *Let ψ be a measurable function on $[0,1] \times \mathbb{R}^d$ such that for all $u \in \mathcal{C}_c^2(\mathbb{R}^d)$, (ψ, u) is in $\operatorname{dom} \Gamma^P$. The unique vector field β which satisfies (37) and $\beta_t(x) \in \operatorname{Range} a(x)$ up to $dt P_t(dx)$-a.e. equality is denoted by $\beta = \widetilde{\nabla}^P \psi$ and it is called the P-extended gradient of ψ.*
When no confusion can occur, we simply drop P and write $\widetilde{\nabla}^P \psi = \widetilde{\nabla} \psi$.

It is clear with our previous discussion that for any $u \in \mathcal{C}_c^2(\mathbb{R}^d)$, $\widetilde{\nabla} u$ is the orthogonal projection of ∇u on the range of the diffusion matrix a. In particular, $\widetilde{\nabla} u = \nabla u$, $dt P_t(dx)$-a.e., when $a(x)$ is invertible for all $x \in \mathbb{R}^d$.

The martingale problem which is solved by P. Now we consider the generalized h-process P. We are going to see that P solves a martingale problem $\mathrm{MP}(b + a\beta, a)$ and that the additional drift β has the special form

$$\beta = \widetilde{\nabla}^P \psi, \quad P\text{-a.s.}$$

with $\psi = \log g$, i.e.

$$\psi(t, x) := \log E_R \left[\exp \left(- \int_{[t,1]} V_s(X_s) \, ds \right) g_1(X_1) \mid X_t = x \right], \quad dt P_t(dx)\text{-a.e.} \tag{38}$$

which is well-defined $dt P_t(dx)$-a.e. since $g(t, x) > 0$, $dt P_t(dx)$-a.e., but might not be defined $dt m(dx)$-a.e. in general.

Lemma 5.3. *Assume that R satisfies the hypotheses of Result 5.1 and P defined by (19) satisfies the hypotheses of Theorem 4.12. Then, for all $u \in \mathcal{U}_R$ which verifies (30), (ψ, u) is in $\operatorname{dom} \Gamma^P$ and*

$$\frac{\Gamma^R(g, u)}{g}(t, x) = \Gamma^P(\psi, u)(t, x), \quad dt P_t(dx)\text{-a.e.}$$

Proof. Let us denote $Z_t = dP_{[0,t]}/dR_{[0,t]}$. As Z admits a continuous version and $Z_t = f_0(X_0) \exp \left(- \int_{[0,t]} V_s(X_s) \, ds \right) G_t$ with $G_t := g_t(X_t)$, G also admits a continuous version. Applying Itô's formula to the continuous process $\psi_t(X_t) = \log G_t$, we obtain

$$d\psi_t(X_t) = \frac{dG_t}{G_t} - \frac{1}{2} \frac{d\langle G \rangle_t}{G_t^2} \quad P\text{-a.s.} \tag{39}$$

We deduce from this with Theorem 4.12 that for any $u \in \mathcal{U}_R$ which verifies (30), $d\langle \psi(X), u(X) \rangle_t = d\langle G, u(X) \rangle_t / G_t = [\Gamma^R(g, u)/g](t, X_t) \, dt$, P-a.s. This completes the proof of the lemma. $\qquad \square$

Theorem 5.4. *Assume that R satisfies the hypotheses of Result 5.1 and let P be the generalized h-process which is defined by (19). Assume also that f_0, g_1 and V satisty the hypotheses of Theorem 4.12.*
Then P is the unique solution in $\{ Q \in \mathrm{P}(\Omega); H(Q|R) < \infty \}$ of

$$P \in \mathrm{MP}(b + a\widetilde{\nabla}^P \psi, a; P_0)$$

with $P_0 = f_0 g_0 \, m$ and where the function

$$\psi(t, x) := \log g(t, x) = \log E_R \left[\exp \left(- \int_{[t,1]} V_s(X_s) \, ds \right) g_1(X_1) \mid X_t = x \right], \quad dt P_t(dx)\text{-a.e.}$$

is defined by (14) *and* (38).

Proof. Choosing $\mathcal{U}_R = \mathcal{C}_c^{1,2}([0, 1] \times \mathbb{R}^d)$ in Theorem 4.12, the assumption (30) holds true for all $u \in \mathcal{C}_c^{1,2}([0, 1] \times \mathbb{R}^d)$. The result now follows from Theorem 4.12 and Lemma 5.3. The assertion $P_0 = f_0 g_0 \, m$ is (22).

The uniqueness is implied by a general result of Girsanov's theory since R is the unique solution to its own martingale problem and $H(P|R) < \infty$. For an entropic point of view under the present requirement that $H(P|R) < \infty$, see [Léo12]. Otherwise, when $P \ll R$ is only assumed this is a standard result of Girsanov's theory, see [JS87]. □

Kolmogorov diffusion process. We illustrate this theorem by means of a diffusion process which plays an important role in the area of functional equalities connected with the concentration of measure phenomenon [Bak94, Roy99, Led01, Vil09].

The Kolmogorov diffusion process is the unique solution of the SDE

$$dX_t = -\nabla U(X_t) \, dt + dW_t$$

where U is a $\mathcal{C}^2$-differentiable function on $\mathbb{R}^d$ such that $Z_U := \int_{\mathbb{R}^d} e^{-2U(x)} \, dx < \infty$. This SDE admits the Boltzmann-Gibbs probability measure

$$m^U(dx) := Z_U^{-1} e^{-2U(x)} \, dx$$

as a reversing measure. We take this reversible Kolmorov diffusion as the reference process R. Hence, the initial law is $R_0 = m^U$ and

$$R \in \mathrm{MP}(-\nabla U, \mathrm{Id}).$$

The generalized h-process to be considered here is P specified by (19) with the assumptions of Theorem 5.4. This theorem tells us that

$$P \in \mathrm{MP}(-\nabla U + \widetilde{\nabla}^P \psi, \mathrm{Id}).$$

In the special case when the potential V is zero, we have for all $0 \leq t < 1$,

$$g_t(x) = E_R(g_1(X_1) \mid X_t = x) = [2\pi(1 - t)]^{-d/2} \int_{\mathbb{R}^d} g_1(y) \exp \left(\frac{|y - x|^2}{2(1 - t)} \right) \, dy.$$

Therefore, $g \in \mathcal{C}^\infty([0, 1) \times \mathbb{R}^d)$ and g_t is positive for all $0 \leq t < 1$. It follows with

$$\psi_t(x) = \log E_R(g_1(X_1) \mid X_t = x), \quad t \in [0, 1), x \in \mathbb{R}^d,$$

that $\widetilde{\nabla}^P \psi_t = \nabla \psi_t$ and

$$P \in \mathrm{MP}\big(-\nabla[U - \psi], \mathrm{Id}\big)$$

and with Theorem 3.3 we see that ψ is a classical solution of the Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{cases} \mathcal{L}^R \psi(t, x) + \frac{1}{2} |\nabla \psi_t(x)|^2 = 0, & t \in [0, 1), x \in \mathbb{R}^d \\ \lim_{t \uparrow 1} \psi_t(x) := \psi_1(x) = \log g_1(x), & t = 1, x \in \{g_1 > 0\} \end{cases}$$

where

$$\mathcal{L}^R u(t, x) = \left(\partial_t - \nabla U(x) \cdot \nabla + \frac{1}{2} \Delta \right) u(t, x).$$

Let us go back to the general case when V is not constant. The positivity improving property of the heat kernel implies that ψ_t is well-defined for all $t \in [0, 1)$. But it might not be smooth enough to be a classical solution of the HJB equation:

$$\begin{cases} \mathcal{L}^R\psi(t, x) + \frac{1}{2}|\nabla\psi_t(x)|^2 - V(t, x) = 0, & t \in [0, 1), x \in \mathbb{R}^d \\ \lim_{t\uparrow 1} \psi_t(x) := \psi_1(x) = \log g_1(x), & t = 1, x \in \{g_1 > 0\} \end{cases}$$

Because of its semigroup representation (38), ψ is a continuous viscosity solution of this equation, see [FS93, Thm II.5.1] for instance.

6. CONTINUOUS-TIME MARKOV CHAINS

In this section we examplify our results with simple Markov jump processes on a countable discrete space $\mathcal{X}$ which are analogous to the Kolmogorov diffusion processes. The set of paths is $\Omega = D([0, 1], \mathcal{X})$.

The reference process R. Since $\mathcal{X}$ is a countable discrete space, every function is measurable and continuous. Let $\mathrm{B}(\mathcal{X})$ denote the space of all real bounded functions on $\mathcal{X}$. The first ingredient is a Markov generator

$$\int_{\mathcal{X}} [u(y) - u(x)] \, J^0(x; dy), \quad u \in \mathrm{B}(\mathcal{X}) \tag{40}$$

where J^0 is a kernel of positive measures on $\mathcal{X}$ such that $J^0(x; \{x\}) = 0$ for all $x \in \mathcal{X}$ and

(i) $J^0(x; \mathcal{X}) < \infty$, for all $x \in \mathcal{X}$;
(ii) J^0 induces an irreducible process in the sense that $J^0(x; \mathcal{X}) > 0$ for all $x \in \mathcal{X}$ and for any couple of distinct states (x, y), there exists a finite chain $x = z_1, z_2, \ldots, z_n = y$ such that $J(z_i; \{z_{i+1}\}) > 0$ for all i;
(iii) J^0 satisfies the detailed balance condition

$$m^0(dx)J^0(x; dy) = m^0(dy)J^0(y; dx) \tag{41}$$

for some nonnegative measure m^0 on $\mathcal{X}$ (possibly with an infinite mass).

We say that $Q \in \mathrm{P}(\Omega)$ solves the *martingale problem* $\mathrm{MP}(K)$ associated with the predictable jump kernel $K = K(t, X_{[0,t)}; dy)$, if

$$u(t, X_t) - u(0, X_0) - \int_{[0,t]} ds \int_{\mathcal{X}} [u(s, y) - u(s, X_{s^-})] \, K(s, X_{[0,s)}; dy), \quad t \in [0, 1]$$

is a local Q-martingale for a large class of functions u.

Under the assumption (i), there is a unique law $R^0 \in \mathrm{P}(\Omega)$ which solves the martingale problem with a prescribed initial law and the Markov generator (40): $R^0 \in \mathrm{MP}(J^0)$. Under the assumption (iii), the measure m^0 is its invariant measure which is unique (up to scalar multiplication) under the irreducibility assumption (ii).

The second ingredient is a potential U on $\mathcal{X}$ such that $Z_U := \int_{\mathcal{X}} e^{-2U} \, dm < +\infty$. The reference process R is the law of the Markov jump process with generator

$$\mathcal{L}^R u(x) \quad := \quad \int_{\mathcal{X}} [u(y) - u(x)] \, J(x; dy), \quad u \in \mathrm{B}(\mathcal{X}) \quad \text{where}$$

$$J(x; dy) \quad := \quad \exp(-[U(y) - U(x)]) \, J^0(x; dy)$$

which is well defined for all $u \in \mathrm{B}(\mathcal{X})$ provided that

$$\int_{\mathcal{X}} e^{-U(y)} \, J^0(x; dy) < +\infty, \quad \forall x \in \mathcal{X},$$

as this last integrability assumption implies that

$$J(x; \mathcal{X}) < \infty, \quad \forall x \in \mathcal{X}. \tag{42}$$

It is easily seen that the Boltzmann-Gibbs probability measure

$$m^U(dx) := Z_U^{-1} e^{-2U(x)} m^0(dx)$$

is the reversing measure of the jump process R since the detailed balance conditions are satisfied. Indeed,

$$m^U(dx)J(x; dy) = e^{-2U(x)} e^{-[U(y)-U(x)]} m^0(dx)J^0(x; dy)$$
$$= e^{-[U(x)+U(y)]} m^0(dx)J^0(x; dy) = e^{-[U(x)+U(y)]} m^0(dy)J^0(y; dx) = m^U(dy)J(y; dx)$$

where (41) has been used at the last but one equality. Therefore,

$$R \in \mathrm{MP}(J; m^U).$$

Moreover, it is the unique solution of this martingale problem. Indeed, thanks to (42) it is possible to build a unique strong solution on some auxiliary probability space: a combination of a discrete-time Markov chain with transition probabilities $J(x; dy)/J(x; \mathcal{X})$ and independent exponential clocks with frequencies $J(x; \mathcal{X})$, $x \in \mathcal{X}$.

This reference law is sometimes called a *Metropolis dynamics* on the set $\mathcal{X}$. It is useful for estimating m^U when the very high cardinality of $\mathcal{X}$ prevents us from computing the normalizing constant Z_U.

The martingale problem which is solved by P. Now we consider the h-process P. Applying Theorem 4.12, we need to compute $\Gamma^R(g, u)/g$ for a large class of functions $u \in \mathcal{U}_R$. We choose this class to be $\mathrm{B}(\mathcal{X})$ for the following reasons. On one hand, we can see that $\mathrm{B}(\mathcal{X}) \subset \mathcal{U}_R$ because with (42) it is clear that $\mathrm{B}(\mathcal{X}) \subset \mathrm{dom}\,\mathcal{L}^R$ and for all $u \in \mathrm{B}(\mathcal{X})$ and $\alpha \geq 0$, $\int_{[0,1] \times \mathcal{X}} \exp(\alpha[u(y) - u(X_{t-})]) \, dt J(X_{t-}; dy) < \infty$. On the other hand, we also see immediately with (42) that (30) holds for any bounded function u.

Theorem 6.1. *Let $R \in \mathrm{MP}(J; m^U)$ be as above, P be the h-process specified at (19) and assume also that f_0, g_1 and V satisty the hypotheses of Theorem 4.12.*

Then P is the unique solution in $\{Q \in \mathrm{P}(\Omega); H(Q|R) < \infty\}$ of $\mathrm{MP}(J^P; P_0)$ with $P_0 = f_0 g_0\, m$ and

$$J^P(t, x; dy) = \exp\left(\psi_t(y) - \psi_t(x)\right) J(x; dy) = \frac{g_t(y)}{g_t(x)} J(x; dy), \quad dt P_t(dx)\text{-}a.e.$$

where the function

$$\psi(t, x) := \log g(t, x) = \log E_R \left[\exp\left(-\int_{[t,1]} V_s(X_s) \, ds\right) g_1(X_1) \mid X_t = x\right], \quad dt P_t(dx)\text{-}a.e.$$

is still defined by (14) and (38).

Proof. Corollary 4.13 tells us that for all $u \in \mathrm{B}(\mathcal{X})$, $\Gamma^R(g, u) = \mathcal{L}^R(ug) - u\mathcal{L}^R g - g\mathcal{L}^R u$. Hence, $\Gamma^R(g, u)(t, x) = \int_{\mathcal{X}} [u(y) - u(x)][g_t(y) - g_t(x)] J(x; dy)$, $dt m(dx)$ and

$$\frac{\Gamma^R(g, u)}{g}(t, x) = \int_{\mathcal{X}} [u(y) - u(x)]\left(\frac{g_t(y)}{g_t(x)} - 1\right) J(x; dy)$$
$$= \int_{\mathcal{X}} [u(y) - u(x)]\left(e^{\psi_t(y) - \psi_t(x)} - 1\right) J(x; dy), \quad dt P_t(x)\text{-}a.e.$$

We conclude with Theorem 4.12 that P solves the announced martingale problem. The uniqueness statement follows from the general Girsanov theory: because $P \ll R$, it is inherited from the fact that R is the unique solution of its martingale problem. $\square$

As with the continuous diffusion processes, we see that some gradient of ψ is involved in the shift from the dynamics of R to the dynamics of the h-process P. Indeed, denoting

$$Du(x; y) := u(y) - u(x)$$

the discrete gradient of u at x, we have

$$J^P(x; dy) = \exp\left(D\psi_t(x; y) \right) J(x; dy).$$

With Theorem 3.3 we know that $\mathcal{L}^R g = V g$. If g is time-differentiable and positive on $[0, 1) \times \mathcal{X}$, we deduce that ψ is a classical solution of the following integro-differential HJB equation

$$\mathcal{L}^R \psi(t, x) + \int_{\mathcal{X}} \theta(D\psi_t(x; y)) \, J(x; dy) - V(t, x) = 0$$

where $\theta(a) := e^a - a - 1$ and $\mathcal{L}^R$ is the generator whose value on any t-differentiable bounded function u is

$$\mathcal{L}^R u(t, x) = \partial_t u(t, x) + \int_{\mathcal{X}} Du_t(x; y) \, J(x; dy).$$

In the general case when g might not be time-differentiable and positive on $[0, 1) \times \mathcal{X}$, the semigroup representation of ψ implies that ψ is the unique continuous viscosity solution of the HJB equation

$$\begin{cases} \mathcal{L}^R \psi(t, x) + \int_{\mathcal{X}} \theta(D\psi_t(x; y)) \, J(x; dy) - V(t, x) = 0, & t \in [0, 1), x \in \mathbb{R}^d \\ \lim_{t \uparrow 1} \psi_t(x) := \psi_1(x) = \log g_1(x), & t = 1, x \in \{g_1 > 0\} \,. \end{cases}$$

References

[Alb03] S. Albeverio. *Theory of Dirichlet forms and applications, in Ecole d'Eté de Probabilités de Saint-Flour XXX-2000*, volume 1816 of *Lecture Notes in Mathematics*. Springer, Berlin, 2003.

[Bak94] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Ecole d'Eté de Probabilités de Saint-Flour (1992)*, number 1581 in Lecture Notes in Mathematics. Springer-Verlag, 1994.

[CL94] P. Cattiaux and C. Léonard. Minimization of the Kullback information of diffusion processes. *Ann. Inst. H. Poincaré. Probab. Statist.*, 30:83–132, 1994.

[CL95] P. Cattiaux and C. Léonard. Large deviations and Nelson's processes. *Forum Math.*, 7:95–115, 1995.

[CL96] P. Cattiaux and C. Léonard. Minimization of the Kullback information for some Markov processes. In *Seminar on Probability, tome 30 (Univ. Strasbourg, Strasbourg, 1996)*, volume 1626 of *Lecture Notes in Math.*, pages 288–311. Springer, Berlin, 1996.

[Csi75] I. Csiszár. *I*-divergence geometry of probability distributions and minimization problems. *Annals of Probability*, 3:146–158, 1975.

[CZ91] A.B. Cruzeiro and J.-C. Zambrini. Malliavin calculus and Euclidean quantum mechanics, I. *J. Funct. Anal.*, 96(1):62–95, 1991.

[CZ08] K.L. Chung and J.-C. Zambrini. *Introduction to Random Time and Quantum Randomness.* World Scientific, 2008.

[DM87] C. Dellacherie and P.A. Meyer. *Probabilités et Potentiel. Ch. XII à XVI. Théorie du potentiel associée à une résolvante, théorie des processus de Markov.* Hermann. Paris, 1987.

[Doo57] J.L. Doob. Conditional Brownian motion and the boundary limits of harmonic functions. *Bull. Soc. Math. France*, 85:431–458, 1957.

[Doo00] J.L. Doob. *Classical Potential Theory and Its Probabilistic Counterpart.* Classics in Mathematics. Springer, 2nd edition, 2000. (reprint of the 1984 first edition).

[FH65] R. Feynman and A. Hibbs. *Quantum Mechanics and Path Integrals.* McGraw-Hill, 1965.

[Föl86]   H. Föllmer. Time reversal on Wiener space. In *Stochastic Processes - Mathematic and Physics*, volume 1158 of *Lecture Notes in Math.*, pages 119–129. Springer, Berlin, 1986.

[Föl88]   H. Föllmer. *Random fields and diffusion processes, in École d'été de Probabilités de Saint-Flour XV-XVII-1985-87*, volume 1362 of *Lecture Notes in Mathematics*. Springer, Berlin, 1988.

[FS93]   W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Applications of Mathematics*. Springer, 1993.

[JS87]   J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der mathematischen Wissenshaften*. Springer, 1987.

[Led01]   M. Ledoux. *The Concentration of Measure Phenomenon*. Mathematical Surveys and Monographs 89. American Mathematical Society, Providence RI, 2001.

[Léo01]   C. Léonard. Minimizers of energy functionals. *Acta Math. Hungar.*, 93(4):281–325, 2001.

[Léo12]   C. Léonard. Girsanov theory under a finite entropy condition. In *Séminaire de probabilités de Strasbourg, vol. XLIV.*, pages 429–465. Lecture Notes in Mathematics 2046. Springer, 2012.

[McC95]   R.J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80:309–323, 1995.

[Mok89]   G. Mokobodzki. L'opérateur carré du champ: un contre-exemple. In *Séminaire de probabilités. Tome 23*, volume 1372 of *Lecture Notes in Mathematics*, pages 324–325. Springer, 1989.

[MZ84]   P.-A. Meyer and W. A Zheng. Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré. Probab. Statist.*, 20(4):353–372, 1984.

[MZ85]   P.-A. Meyer and W. A Zheng. Construction de processus de Nelson réversibles. In *Séminaire de probabilités. Tome 19*, volume 1123 of *Lecture Notes in Mathematics*, pages 12–26. Springer, 1985.

[Nel67]   E. Nelson. *Dynamical theories of Brownian motion*. Princeton University Press, 1967. Second edition (2001) at: www.math.princeton.edu/~nelson/books.html.

[Nel88]   E. Nelson. *Stochastic mechanics and random fields, in Ecole d'Eté de Probabilités de Saint-Flour XV-XVII-1985-87*, volume 1362 of *Lecture Notes in Mathematics*. Springer, Berlin, 1988.

[Roy99]   G. Royer. *Une initiation aux inégalités de Sobolev logarithmiques*, volume 5 of *Cours spécialisés*. Société Mathématique de France, 1999.

[RR91]   M.M. Rao and Z.D. Ren. *Theory of Orlicz spaces*, volume 146 of *Pure and Applied Mathematics*. Marcel Dekker, Inc., 1991.

[RY99]   D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 3rd edition, 1999.

[Vil09]   C. Villani. *Optimal Transport. Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.

Modal-X. Université Paris Ouest. Bât. G, 200 av. de la République. 92001 Nanterre, France

E-mail address: christian.leonard@u-paris10.fr