PATRICK CATTIAUX
CHRISTIAN LÉONARD

Minimization of the Kullback information for some Markov processes

Séminaire de probabilités (Strasbourg), tome 30 (1996), p. 288-311.
<http://www.numdam.org/item?id=SPS_1996__30__288_0>

Minimization of the Kullback Information for some Markov Processes

P. Cattiaux[(1),(3)] and C. Léonard[(2),(3)]

(1) *Ecole Polytechnique, CMAP, 91128 Palaiseau Cedex, France.*
URA CNRS 756.

(2) *Equipe de Modélisation Stochastique et Statistique, Université de Paris-Sud, Bâtiment 425, 91405 Orsay Cedex, France.*
URA CNRS 743.

(3) *Equipe ModalX, Université de Paris X, Bâtiment G, 200 Av. de la République, 92001 Nanterre Cedex, France.*

E-mail: *cattiaux@paris.polytechnique.fr*
Christian.Leonard@math.u-psud.fr

Abstract. *We extend previous results of the authors ([CaL1] and [CaL2]) to general Markov processes which admit a "carré du champ" operator. This yields variational characterizations for the existence of Markov processes with a given flow of time marginal laws which is the stochastic quantization problem, extending previous results obtained by P.A. Meyer and W.A. Zheng or S. Albeverio and M. Röckner in the symmetric case to nonsymmetric processes.*

0. Introduction

In two previous papers ([CaL1], [CaL2]), we have studied the problem of minimizing the Kullback information (or relative entropy) with respect to the law P of a $\mathbb{R}^d$-valued diffusion process, when the flow of its time marginal laws is fixed. This problem is natural when one looks at the large deviations for the empirical process associated with independent copies of such diffusions (see [Föl] and the introduction of [CaL1]). At the same time, the finiteness of the rate function of this large deviation principle was connected in [CaL1] to the existence of diffusion processes with singular drifts of finite energy, encountered in Nelson's approach of Schrödinger's equation (see [Car]). The existence of such diffusions with a given flow of time marginal laws and a given drift is sometimes called "stochastic quantization". The problem of describing the minimizing element in the class of such singular diffusions is connected with the "critical" diffusions of Nelson.

After giving a "stochastic calculus" approach of this construction in [CaL1], we proved in [CaL2] that the finiteness of the rate function can be obtained using direct large deviations techniques.

We refer to the introductions of both papers [CaL1] and [CaL2], for a precise statement of what is written above and for the connection with Schrödinger's original ideas.

In these works, we announced that, in contrast with known methods (of [Car], [MeZ] and other references in [CaL1]), ours could be extended to more general frameworks. The present paper shows how to extend these results to a general strong Markov process with a Polish state space, provided that it admits a *carré du champ* operator. In particular, Section 3 follows closely the lines of [CaL1] (with a lot of simplifications) and Section 4 follows closely the lines of [CaL2]. Of course, the main difference is the use of an "intrisic" gradient operator which is connected to the stochastic structure of the reference process, and of the "Markov differential calculus" associated with this gradient instead of the usual euclidian structure on $\mathbb{R}^d$.

We want to underline that this generalization is not only a quest of abstraction. Actually, we show on some examples in Section 5, that this problem is also strongly connected with recent developments in the theory of symmetric and asymmetric Dirichlet forms on infinite dimensional state spaces. To keep this paper into a reasonable size, we shall not develop this point here, but somewhere else.

Let us present the organization of the paper.

Section 1 describes our framework and in Section 2 are recalled some elementary facts on the relative entropy.

In Section 3, assuming that the set A_ν^I (see (2.2)) is non empty, we describe its structure and characterize its minimal element (as in [CaL1], Sections 3 and 5).

In Section 4, we connect the weak Fokker-Planck equation with a large deviation principle, for which we give various expressions of the rate function. This leads to the natural non variational characterization of the existence of singular processes (in the spirit of singular diffusion processes) stated in the Corollary 4.7 of Theorem 4.6.

In Section 5, we study on some (generic) examples, how to fulfill the main hypothesis (HC) of Theorem 4.6. General statements are given for manifold-valued diffusion processes, reflected diffusion processes, symmetric processes (see Theorem 5.5) as well as particular infinite dimensional processes (in the nonsymmetric case).

1. The framework

$(E, \mathcal{E})$ is a Polish space equipped with its Borel σ-field $\mathcal{E}$. $\mathcal{E}^*$ is the universal completion of $\mathcal{E}$, $\Omega = C([0,T], E)$ is the set of E-valued continuous paths. $(\Omega, \mathcal{F}, (\mathcal{F}_t, X_t, \theta_t)_{t\in[0,T]}, (P_x)_{x\in E})$ is the canonical realization of a strongly Markov continuous E-valued process. As usual, the abbreviation a.e. (almost everywhere) stands for "except on a set of potential zero".

The sets $B(E), B^*(E), C(E), C_u(E)$ are respectively the sets of Borel, universally measurable, continuous and uniformly continuous real valued functions on E. The subscript b will mean bounded.

For $f \in B_b^*(E)$, we define $P_t f(x) = E^x[f(X_t)]$ (E^x stands for the expectation with respect to P_x) which is assumed to be $\mathcal{E}^*$-measurable for all $t \in [0,T]$. Then, $(P_t)_{t\in[0,T]}$ is a strongly continuous semigroup on the set $\mathcal{C} = \{f \in B_b^*(E)\,;\, \|P_t f - f\|_\infty \to 0 \text{ as } t \downarrow 0\}$. Let $(A, D(A))$ be the generator with its domain of the semigroup. The extended domain $D_e(A)$ is defined as

$$D_e(A) = \{f \in B_b^*(E)\,;\; \text{such that there exists } g \in B^*(E) \text{ with}$$

$$\int_0^T |g(X_s)|\,ds < \infty \ P_x\text{-a.s. for all } x, \text{ and}$$

$$C_t^f := f(X_t) - f(X_0) - \int_0^t g(X_s)\,ds \text{ is a } P_x\text{-local martingale}\}.$$

If $f \in D_e(A)$, we put $Af = g$, noticing that Af is defined up to a set of potential zero.

Definition. *Let Θ be a subset of $D_e(A)$. We shall say that Θ is a core if*

i) for all $x \in E$, P_x is an extremal solution of the martingale problem $\mathcal{M}(A, \Theta, \delta_x)$ (see [Jac] for the notation),

ii) Θ is a subalgebra of $C_{b,u}(E)$,

iii) there is no signed measure η, except 0, such that $\int f\,d\eta = 0$ for all $f \in \Theta$.

From now on, we shall assume that

(H) There exists a core.

Here are some well known consequences (see ([DeM], Chap. XV), ([Jac], pp. 421–431) or [MeZ]).

Properties of the process.

i) For any $\mu \in M_1(E)$ (probability measures on E) and any local P_μ-martingale M, M admits a continuous modification and its increasing process $\langle M\rangle_t$ is absolutely continuous.

ii) $D_e(A)$ is an algebra and we may define the *carré du champ* operator Γ on $D_e(A) \times D_e(A)$ as: $\Gamma(f,g) = A(fg) - fAg - gAf$.

iii) There exists a sequence $(\varphi_n)_{n\geq 1}$ of elements of $D_e(A)$ such that the local martingales $C_t^n = C_t^{\varphi_n}$, $n \geq 1$, for all P_μ, generate the space $\mathcal{M}^2_{\mathrm{loc}}(P_\mu)$ of square integrable local P_μ-martingales starting from 0, i.e. if $M \in \mathcal{M}^2_{\mathrm{loc}}(P_\mu)$, there exists a sequence $(m^n)_{n\geq 1}$ of previsible processes such that for all P_μ, and all localizing sequence $(T_k)_{k\geq 1}$ of stopping times

$$M_{t\wedge T_k} = \sum_{n\geq 1} \int_0^{t\wedge T_k} m_s^n\,dC_s^n \qquad (\text{in the sense of } \mathcal{M}^2_{\mathrm{loc}}(P_\mu)).$$

Furthermore, if M is a local matingale which is an additive functional, one can find functions in $B^*(E)$, still denoted m^n, such that $m_s^n = m^n(X_s)$ in the previous decomposition.

iv) Any $f \in D_e(A)$ is continuous along the paths.

These properties of the process allow us to define the natural gradient operators $(\nabla^n)_{n\geq 1}$. Indeed,

there exists a sequence $(\nabla^n)_{n\geq 1}$ of operators defined on $D_e(A)$ with values in $B^*(E)$ such that for all $f \in D_e(A)$ and all $t \geq 0$

$$C_t^f = \sum_{n\geq 1} \int_0^t \nabla^n f(X_s)\,dC_s^n$$

in the sense of $\mathcal{M}^2_{\mathrm{loc}}(P_\mu)$ for all P_μ.

Here again, $\nabla^n f$ is defined up to a set of potential zero. It follows that

$$\forall f,g \in D_e(A),\ \Gamma(f,g) = \sum_{n,k\geq 1} \nabla^n f\ \Gamma(\varphi_n, \varphi_k)\ \nabla^k f, \text{ a.e.}$$

For simplicity, we shall write $\Gamma_{n,k}$ instead of $\Gamma(\varphi_n, \varphi_k)$ and $\Gamma(f)$ instead of $\Gamma(f,f)$. The gradients ∇^n satisfy the usual rules of derivations and if Φ is a C_b^2 function on $\mathbb{R}$, $\nabla^n(\Phi \circ f) = \Phi' \circ f \dot{\nabla}^n f$ thanks to Itô's chain rule.

It is also easy to see that for any sequence $(\xi_n)_{n\geq 1}$ of real numbers, $\sum_{n,k\geq 1} \xi_n \Gamma_{n,k} \xi_k$ is nonnegative a.e.For such sequences $(\xi_n)_{n\geq 1}$ and $(\eta_n)_{n\geq 1}$, we write

$$\gamma(\xi,\eta) = \sum_{n,k\geq 1} \xi_n \, \Gamma_{n,k}(x) \, \eta_k. \tag{1.1}$$

Finally, for all $\mu \in M_1(E)$, P_μ is an extremal solution of the martingale problem associated with $(C^n)_{n\geq 1}$ and the initial law μ. Thus, the usual Girsanov theory is available. If $Q \ll P_\mu$, there exists a sequence $\beta = (\beta^n)_{n\geq 1}$ of real valued previsible processes (it will be called the drift of Q) such that, if we define

$$T_k = \inf\{t \geq 0, \int_0^t \gamma(\beta_s)\, ds \geq k\}, \;\; k \in \mathbb{N} \cup \{+\infty\}, \tag{1.2}$$

(where $\gamma(\beta_s) = \gamma(\beta_s, \beta_s)(X_s)$ as defined in (1.1)), the density process Z of Q is given by

$$Z_t = \frac{d(Q \circ (X_0)^{-1})}{d\mu} \exp\left(\sum_{n\geq 1} \int_0^{t\wedge T_\infty} \beta_s^n \, dC_s^n - \frac{1}{2}\int_0^{t\wedge T_\infty} \gamma(\beta_s)\, ds\right). \tag{1.3}$$

Furthermore, Z is a continuous P_μ-martingale, hence

$$T_\infty > T_k, P_\mu\text{-a.s.}, \forall k \geq 1 \quad \text{and} \quad T_\infty = \inf\{t \geq 0\,;\, Z_t = 0\}. \tag{1.4}$$

According to the usual Girsanov transform theory

> $T_\infty \wedge T = T$, Q-a.s., $N_t^n = C_t^n - \int_0^t \sum_{k\geq 1} \Gamma_{n,k}(X_s)\beta_s^k \, ds$ is a local Q-martingale with $(T_k)_{k\geq 1}$ as a localizing sequence of stopping times and
>
> $$\langle N^n, N^k\rangle_t = \int_0^t \Gamma_{n,k}(X_s)\, ds, \;\; Q\text{-a.s.}$$
>
> In addition, Q is an extremal solution of the martingale problem associated with $(N^n)_{n\geq 1}$ (see [Jac], 12.22).

Conversely, let $\beta = (\beta^n)_{n\geq 1}$ be a sequence of previsible processes. Let $Z(\beta, \nu_o, P_{\mu_o})$ stand for the process defined by (1.3), with (ν_o, μ_o) in place of $(Q \circ (X_0)^{-1}, \mu)$. Then $Z(\beta, \nu_o, P_{\mu_o})$ is a nonnegative local P_{μ_o}-martingale, hence a P_{μ_o}-supermartingale, which is continuous P_{μ_o}-a.s. To this supermartingale corresponds its Föllmer measure.

(1.5) Notation. Let Ω_ξ be the space of explosive paths with explosion time ξ, the above Föllmer measure is called $(\beta, \nu_o, P_{\mu_o})$-FM (as in [CaL1]). If $\beta_s = B(X_s)$, we write (B, ν_o, P_{μ_o})-FM.

(1.6) If $P_{\mu_o}(T_\infty > T_k, \forall k \geq 1) = 1$, then $T_\infty = \inf\{t \geq 0\,;\, Z_t = 0\}$ and thanks to ([Sha], Theorem 24.36), the family Q_x of all the (B, δ_x, P_x)-FM, for a given B in $B^*(E)$, defines a strong Markov process on Ω_ξ.
Notice that P_{μ_o} and $(\beta, \nu_o, P_{\mu_o})$-FM are equivalent on $\{T_\infty = T_k\}$, hence $P_{\mu_o}(T_\infty > T_k, \forall k \geq 1) = 1$ is equivalent to the same condition replacing P_{μ_o} by $(\beta, \nu_o, P_{\mu_o})$-FM.

2. Relative entropy

We collect and adapt some results on the relative entropy which have been proved in [CaL1]. Let Q and P be in $M_1(\Omega)$, $I(Q \mid P)$ denotes the relative entropy of Q with respect to P defined by

$$I(Q \mid P) = \begin{cases} \int Z \log Z \, dP & \text{if } Q \ll P, Z = \frac{dQ}{dP}, Z \log Z \in L^1(P) \\ +\infty & \text{otherwise.} \end{cases}$$

It is well known that

$$I(Q \mid P) = \sup_{\Phi} \{ \int \Phi \, dQ - \log \int e^{\Phi} \, dP \}$$

where the supremum is taken either on $C_b(\Omega)$ or $B_b(\Omega)$.

Proposition 2.1. ([CaL1], 2.1). *Assume that $Q \ll P_{\mu_o}$. Set $\nu_o = Q \circ (X_0)^{-1}$. Then, if β is the drift of Q (see (1.2), (1.3)):*

$$I(Q \mid P_{\mu_o}) = I(\nu_o \mid \mu_o) + \frac{1}{2} E^Q [\int_0^T \gamma(\beta_s) \, ds]. \tag{2.1}$$

Proposition 2.2. ([CaL1] 2.3 and correction to [CaL1]). *Let $\beta = (\beta^n)_{n \geq 1}$ be a sequence of previsible processes, Q be the $(\beta, \nu_o, P_{\mu_o})$-FM (see (1.5)). Assume that*
i) $I(\nu_o \mid \mu_o) < +\infty$,
ii) $P_{\mu_o}(\cup_{k \geq 1} \{T_\infty = T_k\}) = 0$ *(or equivalently* $Q(\cup_{k \geq 1} \{T_\infty = T_k\}) = 0$),
iii) $E^Q [\int_0^{T_\infty \wedge T} \gamma(\beta_s) \, ds] < +\infty$.
Then $Q(\xi = +\infty) = 1$, $T_\infty \wedge T = T$, *Q-a.s.,* $I(Q \mid P_{\mu_o}) < +\infty$ *and (2.1) holds.*

We also introduce

i) For $Q \in M_1(E)$, $\mathcal{L}_Q^2 = \{(\beta^n)_{n \geq 1}$ previsible; $E^Q[\int_0^T \gamma(\beta_s) \, ds] < +\infty\}$

ii) Let $(\nu_t)_{t \in [0,T]}$ be a measurable flow of elements of $M_1(E)$,

$$\mathcal{L}_\nu^2 = \{B \in B^*(E \times [0,T])^{\mathbf{N}^*} \, ; \int_{E \times [0,T]} \gamma(B(t, \cdot))(x) \, \nu_t(dx) dt < +\infty\}$$

(B is allowed to be an infinite sequence).

The associated quotient Hilbert spaces are L_Q^2, L_ν^2 with the norms $\| \cdot \|_{L_Q^2}$ and $\| \cdot \|_{L_\nu^2}$. A β (resp. B) which belongs to L_Q^2 (resp. L_ν^2) is said to be of finite Q (resp. ν)-energy.

If $I(Q \mid P_{\mu_o}) < +\infty$, the drift β of Q is of finite Q-energy. Also notice that if β (or B) is a *finite sequence* of bounded previsible processes (functions in $B^*([0,T] \times E)$), then $\beta \in L_Q^2$ ($B \in L_\nu^2$).

We now recall a technical but useful result.

Proposition 2.3. ([CaL1], 2.7 and 2.8). *Let Q and Q^* be elements of $M_1(\Omega)$ such that $Q \ll Q^*$. Put $Z = \frac{dQ}{dQ^*}$. Let S be a bounded Q^*-martingale (with bound C). In the two cases described below, S is a Q-semimartingale with decomposition $S_t = K_t + V_t$ where K is a square integrable Q-martingale and $\langle K \rangle_t = \langle S \rangle_t$. Moreover,*

Case 1. If $I(Q \mid Q^*) < +\infty$, then $E^Q[\langle K \rangle_t] \leq c(1 + I(Q \mid Q^*)) C^2$ where c is a universal constant.

Case 2. If $Z_T \in L^q(Q^*)$ for some $q > 1$, then for all $p \in]1, +\infty[$, $E^Q[\langle K\rangle_t^p] \le \|Z_T\|_q (4pq^*)^p C^{2p}$ with $1/q + 1/q^* = 1$.

In the rest of the paper we are interested in the existence of a probability measure with given time marginals and finite relative entropy. Namely

(2.2) Definition. Let $\nu = (\nu_t)_{t\in[0,T]}$ be a measurable flow of probability measures on E. We define

$$A_\nu = \{Q \in M_1(\Omega), Q \circ (X_t)^{-1} = \nu_t, \forall t \in [0,T]\} \quad \text{and}$$
$$A_\nu^I(\mu_o) = \{Q \in A_\nu\,;\, I(Q \mid P_{\mu_o}) < +\infty\}.$$

If $A_\nu^I(\mu_o)$ is non empty, we shall say that the flow is μ_o-admissible.

Since the μ_o-admissibility implies $I(\nu_o \mid \mu_o) < +\infty$, we only need to look at this case. Conversely, if a flow is ν_o-admissible, it is also μ_o-admissible for all μ_o such that $I(\nu_o \mid \mu_o) < +\infty$. We shall only consider the case $\mu_o = \nu_o$ and denote A_ν^I the set $A_\nu^I = A_\nu^I(\nu_o)$.

3. Some properties of the admissible flows

In this section, we fix a given measurable flow $\nu = (\nu_t)_{t\in[0,T]}$ and we assume that $A_\nu^I \neq \emptyset$. Notice that this implies that $s \mapsto \nu_s$ is weakly continuous. We shall describe the set A_ν^I.

Proposition and Definition 3.1. *Let $Q \in A_\nu^I$ and β be the drift of Q. Then, there exists a unique $B \in L_\nu^2$ such that for all $\varphi \in L_\nu^2$:*

$$\int_{[0,T]\times\Omega} \gamma(\varphi(s,X_s),\beta_s)\, dQds = \int_{[0,T]\times E} \gamma(\varphi,B)(s,x)\, \nu_s(dx)ds.$$

B will be called the Markovian version of β, the (B,ν_o,P_{μ_o})-FM denoted by $\overline{Q}$ will be called the Markovian version of Q.

Proof. Apply Riesz's projection theorem. ∎

Remark. $I(Q \mid P_{\nu_o}) \ge \|B\|_{L_\nu^2}^2$. So that if $\overline{Q}$ is a probability measure on Ω and if $\overline{Q} \in A_\nu$, then $\overline{Q} \in A_\nu^I$ and $I(\overline{Q} \mid P_{\mu_o}) = \|B\|_{L_\nu^2}^2 \le I(Q \mid P_{\nu_o})$.

Since B depends on (t,x), we see that it is convenient to introduce the following time-space process.

Definition. The time-space process (u_t, X_t) with $u_t = u_0 + t$ is defined on the set of the time-space paths: Ω'. $P_{u,x}$ is the law of this process with initial point $(u,x) \in \mathbb{R} \times E$. The family $(P_{u,x})_{(u,x)\in\mathbb{R}\times E}$ is again a strong Markov process with generator $\frac{\partial}{\partial u} + A = A'$ and its domain: $D(A')$, contains $C^1(\mathbb{R}, D(A)) = \{f \in C^1(\mathbb{R}, B_b^*(E))\,;\, f(u,\cdot) \in D(A), \forall u \text{ and } Af \in C^0(\mathbb{R}, B_b^*(E))\}$.

Actually, as in [CaL1], we essentially need the strong Markov property of $P_{u,x}$ (not of P_x), i.e. we shall assume that (P_x) is a non-homogeneous Markov process. So, the main hypothesis we require is

(H') There exists a core Θ' for the time-space process.

All what has been done in Section 1 is still available since $\mathbb{R} \times E$ is a Polish space. Furthermore, for all $\mu \in M_1(E)$, $P_\mu = P_{\mu\otimes\delta_0} \circ X^{-1}$ so that the discussion of Section 2 is easy to transpose.

Definition. $D_{e,\nu}(A')$ is the set of functions $f \in D_e(A')$ such that

$$i) \quad \nabla f \in L^2_\nu \text{ and}$$
$$ii) \quad \int_{[0,T]\times E} |A'f(s,x)|\, \nu_s(dx)ds < +\infty.$$

Proposition 3.2. $D(A') \subset D_{e,\nu}(A')$.

Proof. Let $P = P_{\nu_o \otimes \delta_0}$, $Q \in A^I_\nu$. Then, the law of $(\cdot, X.)$ under Q, still denoted: Q, satisfies $I(Q \mid P) < +\infty$. If $f \in D(A')$, C^f_t is a bounded martingale and $A'f$ is bounded. According to Proposition 2.3: $\nabla f \in L^2_\nu$. ■

Definition. Let $B \in L^2_\nu$ and $\Lambda \subset D_{e,\nu}(A')$. We say that ν satisfies the (B,Λ)-weak Fokker-Planck equation: (B,Λ)-wFP, if for all $0 \le s \le t \le T$ and all $f \in \Lambda$

$$\int_E f(t,x)\, \nu_t(dx) - \int_E f(s,x)\, \nu_s(dx) = \int_s^t \int_E (A'f + \gamma(B, \nabla f))(u,x)\, \nu_u(dx)du.$$

Proposition 3.3.

1. *Let $B \in L^2_\nu$ and $\Lambda \subset D_{e,\nu}(A') \cap C_b(\mathbb{R} \times E)$. Then, ν satisfies the (B,Λ)-wFP if and only if for all $f \in \Lambda$*

$$\int_E f(T,x)\, \nu_T(dx) - \int_E f(0,x)\, \nu_0(dx) = \int_{[0,T]\times E} (A'f + \gamma(B, \nabla f))(s,x)\, \nu_s(dx)ds.$$

2. *Let $Q \in A^I_\nu$, β its drift, B the Markovian version of β (see Proposition 3.1). Then, ν satisfies the $(B, D_{e,\nu}(A'))$-wFP equation.*

Proof. 1) Choose a sequence $(\psi_n)_{n\ge1}$ in $C^\infty([0,T])$ such that $(\psi_n)_{n\ge1}$ is pointwise convergent with limit $\mathbb{1}_{[0,t]}$, $0 \le \psi_n \le 1$ and $(-\psi'_n)_{n\ge1}$ considered as a sequence of measures is weakly convergent with limit δ_t. For $f \in \Lambda$, $\psi_n f \in \Lambda$ and

$$\begin{aligned}&\int_E \psi_n(T) f(T,x)\nu_T(dx) - \int_E \psi_n(0) f(0,x)\nu_0(dx)\\ &= \int_0^T \psi'_n(s)\Big(\int_E f(s,x)\nu_s(dx)\Big)\, ds + \int_{[0,T]\times E} \psi_n(s)(A'f + \gamma(B,\nabla f))(s,x)\, \nu_s(dx)ds.\end{aligned}$$

Since $f \in C_b$, $s \mapsto \int_E f(s,x)\, \nu_s(dx)$ is continuous ($s \mapsto \nu_s$ is weakly continuous) and we are allowed to take the limit in n.

2) Let $f \in D_{e,\nu}(A')$. Applying Girsanov transform theory and Itô's formula, we obtain that

$$f(t,X_t) - f(0,X_0) - \int_0^t A'f(s,X_s)\, ds - \int_0^t \gamma(\beta_s, \nabla f)(s,X_s)\, ds$$

is a square integrable Q-martingale, since $\nabla f \in L^2_\nu$. One completes the proof, taking the expectation with respect to Q and then using Proposition 3.1. ■

Notation. $P = P_{\nu_o \otimes \delta_0}$.

The rest of this section is devoted to the extension the results of ([CaL1], Sections 3 and 5) (except those concerned with the minimization problem) to our general setting. Though the proofs are very similar, we prefer giving almost all the details, not only for the reader's conveniency, but also because some points are presented in a simpler way.

The following construction will be used several times.

(3.1) Auxiliary construction. This construction works under the assumptions (3.1.iii) and (3.1.iv) stated below. Let $B^* = (B^{*1}, \dots, B^{*n})$ be a finite sequence of elements of $B_b^*(\mathbb{R}\times E)$. We also assume that $\gamma(B^*)$ is bounded. Then, for all $(u,x) \in \mathbb{R}\times E$, $(B^*, \delta_x \otimes \delta_u, P_{u,x})$-FM is a probability measure $Q^*_{u,x}$ on Ω' and $Q^*_{u,x} \sim P_{u,x}$ with $I(Q^*_{u,x} \mid P_{u,x}) < +\infty$ and $I(P_{u,x} \mid Q^*_{u,x}) < +\infty$. We thus have a homogeneous strong Markov family $(Q^*_{u,x})$ associated with a strongly continuous Markov semigroup (Q^*_t) on $\mathcal{C}(B^*)$:

$$Q^*_t f(u,x) = E^{Q^*_{u,x}}[f(u_t, X_t)],\ t \in [0,T],\ f \in \mathcal{C}(B^*) \text{ with}$$
$$\mathcal{C}(B^*) = \{f \in B_b^*(\mathbb{R}\times E)\,;\ \|Q^*_t f - f\|_\infty \to 0 \text{ as } t \downarrow 0\}.$$

We denote $(A(B^*), D(B^*))$ the generator of this semigroup with its domain. According to (1.4), one can define the gradient operator ∇^* and then easily prove that $D_e(A') = D_e(A(B^*))$, $\nabla^* = \nabla$ and the extended generators A' and $A(B^*)$ satisfy $A(B^*) = A' + \gamma(B^*, \nabla)$, where $\gamma(B^*, \nabla)(f) = \gamma(B^*, \nabla f)$ (all of this holds a.e. for both families $(Q^*_{u,x})$ and $(P_{u,x})$).

Let us take $t \in [0,T]$ and define the following:

(3.1.i)

For $f \in C_{b,u}([0,T]\times E)$, we define $F(s;(u,x)) = E^{Q^*_{u,x}}[f(u_{t-s}, X_{t-s})]$, $s \in [0,t]$.

(3.1.ii)

We set $G(s,x) = F(s;(s,x))$ on $[0,t]\times E$, $G((s,x)) = G(t,x)$ if $s \ge t$.

Let $0 \le a \le b, u \ge 0$. First assume that $u + b \le t$ and pick $\theta \in B_b^*(\mathbb{R}\times E)$. Then

$$\begin{aligned} E^{Q^*_{u,x}}[G(u_b,X_b)\theta(u_a,X_a)] &= E^{Q^*_{u,x}}\left[\theta(u_a,X_a)E^{Q^*_{u_a,X_a}}[G(u_{b-a},X_{b-a})]\right] \\ &= E^{Q^*_{u,x}}\left[\theta(u_a,X_a)E^{Q^*_{u_a,X_a}}\left(E^{Q^*_{u_{b-a},X_{b-a}}}\left[f(t,X_{t-u-b})\right]\right)\right] \\ &= E^{Q^*_{u,x}}\left[\theta(u_a,X_a)E^{Q^*_{u_a,X_a}}[f(t,X_{t-u-a})]\right]. \end{aligned}$$

So $G(u_s, X_s)$ is a bounded $Q^*_{u,x}$-martingale up to time $t-u$.

If $u + b > t$, $G(u_b, X_b) = f(t, X_b)$ $Q^*_{u,x}$-a.s. It follows that $G \in D_e(A(B^*))$ and that $A(B^*)G(s,x) = 0$ if $s \in [0,t]$ for all x. Thus $G \in D_e(A')$ and $A'G(s,x) = -\gamma(B^*, \nabla G)(s,x)$ for $s \in [0,t]$. Furthermore $G(t,x) = f(t,x)$ and $G(0,x) = E^{Q^*_x}[f(t,X_t)]$.

Now, if

(3.1.iii) $$\nabla G \in L^2_\nu,$$

$G \in D_{e,\nu}(A')$ since $\gamma(B^*)$ is bounded (apply Cauchy-Schwarz inequality). If furthermore

(3.1.iv) ν satisfies the $(B, D_{e,\nu}(A'))$-wFP equation for some B in L^2_ν,

it follows that

$$\int_E f(t,x)\,\nu_t(dx) - \int_E E^{Q^*_{x,0}}[f(u_t,X_t)]\,\nu_0(dx) = \int_{[0,t]\times E} \gamma(B - B^*, \nabla G)(s,x)\,\nu_s(dx)ds.$$

Applying Cauchy-Schwarz inequality, we obtain for a nonnegative f

$$E^{Q^*}[f(t,X_t)] \le \int_E f(t,x)\,\nu_t(dx) + \|B-B^*\|_{L^2_\nu}\|\nabla G\|_{L^2_\nu}$$

where $Q^* = Q^*_{\nu_o\otimes\delta_0}$.

The key point now is that we can get a bound for $\|\nabla G\|_{L^2_\nu}$ which only involves f, B and B^*. Indeed, $G^2 \in D_{e,\nu}(A')$ with $\nabla G^2 = 2G\nabla G$ and $A'G^2 = 2GA'G + \Gamma(G)$. It follows that

$$\begin{aligned}\int_{[0,t]\times E} \Gamma(G)(s,x)\,\nu_s(dx)ds &= \int_E f^2(t,x)\,\nu_t(dx) - \int_E G^2(0,x)\,\nu_0(dx)\\ &\quad - 2\int_{[0,t]\times E} G(s,x)\gamma(B-B^*,\nabla G)(s,x)\,\nu_s(dx)ds,\end{aligned}$$

which yields (see ([CaL1], 4.21–4.23) for the details)

$$\|\nabla G\|_{L^2_\nu} \le 2\|f\|_\infty(1+\|B-B^*\|_{L^2_\nu})^{1/2}.$$

Finally, for any nonnegative $f \in C_{b,u}([0,T]\times E)$

$$\text{(3.1.v)} \qquad E^{Q^*}[f(t,X_t)] \le \int_E f(t,x)\,\nu_t(dx) + 2\|f\|_\infty\,\|B-B^*\|_{L^2_\nu}(1+\|B-B^*\|_{L^2_\nu})^{1/2}.$$

But (3.1.v) extends to any nonnegative $f \in B_b^*(\mathbb{R}\times E)$ since $\mathbb{R}\times E$ is a Polish space and $Q^*\circ(X_t)^{-1} + \nu_t$ is regular. ■

We shall immediately use this construction to prove the following extension of Theorem 3.1 in [CaL1].

Theorem 3.4. *Let $Q \in A^I_\nu$ and $\overline{Q}$ its Markovian version. Then, $\overline{Q} \in A^I_\nu$ and $I(\overline{Q} \mid P_{\nu_o}) \le I(Q \mid P_{\nu_o})$.*

Proof. It should be possible to adapt the (intricate) proof of ([CaL1], Theorem 3.1). We prefer following the scheme of proof suggested in the Section 5 of [CaL1] after the remark 5.5 and using the construction (3.1).

Let

$$\begin{cases} S_k = \inf\{t\ge 0;\ \int_0^t \gamma(B)(s,X_s)\,ds \ge k\} & \text{and put}\quad S_k = +\infty \quad \text{if } \int_0^T \gamma(B)(s,X_s)\,ds < k.\\ T_k = \inf\{t\ge 0;\ \int_0^t \gamma(\beta_s)\,ds \ge k\} & \text{and put}\quad T_k = +\infty \quad \text{if } \int_0^T \gamma(\beta_s)\,ds < k.\end{cases}$$

We define $Q^k = Z_{T\wedge T_k}(\beta,\nu_o\otimes\delta_0,P)\cdot P$ (see the notation in Section 1), which is a probability measure, thanks to Novikov's criterion.

We then define the sequence $(B^i_{n,p})_{i\ge 1}$ as follows

$$\begin{cases} B^i_{n,p} = B^i(u,x)\mathbb{1}_{\{|B^i(u,x)|\le n\}}\mathbb{1}_{\{\max_{1\le j\le n}\Gamma_{ij}(u,x)\le p\}} & \text{if } i\le n\\ B^i_{n,p} = 0 & \text{if } i>n.\end{cases}$$

For each (n,p), $B_{n,p}$ is a finite sequence of bounded measurable functions and $\gamma(B_{n,p})$ is bounded. $Q_{n,p}$ is the $(B_{n,p},\nu_o\otimes\delta_0,P)$-FM. According to the second part of Proposition 3.3, to apply the construction (3.1), it suffices to prove that for G as in (3.1.ii) (with $B^* = B_{n,p}$), ∇G belongs to L^2_ν (Cf. (3.1.iii)). But

$$\|\nabla G\|^2_{L^2_\nu} = E^Q[\int_0^t \Gamma(G)(s,X_s)\,ds] \leq \liminf_{k\to\infty} E^{Q^k}[\int_0^{t\wedge T_k} \Gamma(G)(s,X_s)\,ds].$$

Easy computations show that $I(Q^k \mid Q_{n,p}) \leq 2I(Q \mid P_{\nu_o})$, so we may apply the Case 1 of Proposition 2.3 which yields

$$E^{Q^k}[\int_0^{t\wedge T_k} \Gamma(G)(s,X_s)\,ds] \leq c(1+2I(Q \mid P_{\nu_o}))\|G\|_\infty \leq c(1+2I(Q \mid P_{\nu_o})).$$

So $\nabla G \in L^2_\nu$ and thus for any nonnegative $f \in B^*_b(\mathbb{R} \times E)$

$$E^{Q_{n,p}}[f(t,X_t)] \leq \int_E f(t,x)\,d\nu_t + 2\|f\|_\infty \|B - B_{n,p}\|_{L^2_\nu}(1+\|B-B_{n,p}\|_{L^2_\nu})^{1/2}. \tag{3.2}$$

Applying twice the bounded convergence theorem, on one hand one has

$$\lim_{n\to\infty}\lim_{p\to\infty} \|B - B_{n,p}\|_{L^2_\nu} = 0.$$

On the other hand, the same argument yields for each k

$$\lim_{n\to\infty}\lim_{p\to\infty} E^{P_{\nu_o}}[\int_0^{t\wedge S_k} \gamma(B-B_{n,p})(s,X_s)\,ds] = 0.$$

It follows that a subsequence of $Z_{t\wedge S_k}(B_{n,p},\nu_o\otimes\delta_0,P)$ tends P-a.s. as p tends to infinity to $Z_{t\wedge S_k}(B_{n,\infty},\nu_o\otimes\delta_0,P)$ and that a subsequence of the latest Girsanov density tends P-a.s. as p tends to infinity to $Z_{t\wedge S_k}(B,\nu_o\otimes\delta_0,P)$. Applying twice Fatou's lemma , we obtain that for any nonnegative $f \in B^*_b(\mathbb{R}\times E)$:

$$\begin{aligned} E^{\overline{Q}}[f(t,X_t)\mathbb{1}_{\{t<S_k\}}] &\leq \liminf_{n\to\infty}\liminf_{p\to\infty} E^{Q_{n,p}}[f(t,X_t)\mathbb{1}_{\{t<S_k\}}] \\ &\leq \liminf_{n\to\infty}\liminf_{p\to\infty} E^{Q_{n,p}}[f(t,X_t)] \\ &\leq \int_E f(t,x)\,\nu_t(dx). \end{aligned}$$

By monotone convergence, we can replace S_k by S_∞ and then take $f(u,x) = \gamma(B)(u,x)$, which yields

$$E^{\overline{Q}}[\int_0^{S_\infty\wedge T} \gamma(B)(s,X_s)\,ds] \leq \|B\|^2_{L^2_\nu}.$$

Furthermore, starting with (3.2), one can prove exactly as in the Correction of [CaL1] that $P(\cup_{k\in\mathbb{N}}\{T_\infty = T_k\}) = 0$.

According to the Proposition 2.2, $\overline{Q}$ is a probability measure on Ω and $S_\infty = +\infty, \overline{Q}$-a.s. Hence: $I(\overline{Q} \mid P_{\nu_o}) < +\infty$. This implies

$$E^{\overline{Q}}[f(t,X_t)] \leq \int_E f(t,x)\,\nu_t(dx) \tag{3.3}$$

for f as before. But, since $\overline{Q}\circ(X_t)^{-1}$ and ν_t are probability measures on E, it follows that $\overline{Q}\circ(X_t)^{-1}=\nu_t$. So, $\overline{Q}\in A^I_\nu$. Finally, thanks to the Proposition 2.1, $I(\overline{Q}\mid P_{\nu_o})=\|B\|^2_{L^2_\nu}\le I(Q\mid P_{\nu_o})$. ■

Thanks to (1.6), (the extension of) $\overline{Q}$ (to Ω') is a strong Markov probability measure. Conversely, we have the following result.

Proposition 3.5. (see [CaL1], Theorem 3.60). *The drift β of any Markov probability measure $Q\in A^I_\nu$ is Markovian, i.e.: $\beta_s=B(s,X_s)$.*

Finally, we can give a full description of A^I_ν. To this end, we first state the following

Definition. $H^{-1}(\nu)$ is defined as the L^2_ν-closure of the set $\{\nabla f\,;\, f\in D_{e,\nu}(A')\}$.

Theorem 3.6. *We denote by $\perp$ the orthogonality in L^2_ν.*

a) *There exists (a unique) $B^\nu\in H^{-1}(\nu)$ such that for any Markov probability measure Q in A^I_ν, the drift B of Q satisfies $B-B^\nu\in[H^{-1}(\nu)]^\perp$ and conversely, for any $B^\perp\in[H^{-1}(\nu)]^\perp$, the $(B^\nu+B^\perp,\nu_o,P_{\nu_o})$-FM belongs to A^I_ν.*

b) *All the Markov elements of A^I_ν are equivalent. All the elements of A^I_ν are absolutely continuous with respect to Q^ν : the (B^ν,ν_o,P_{ν_o})-FM.*

Proof. a) The second part of Proposition 3.3 shows that all the Markovian drifts B of the Markov elements of A^I_ν have the same projection B^ν onto $H^{-1}(\nu)$. Let $Q\in A^I_\nu$ be Markov, B be its drift and $B^\perp\in[H^{-1}(\nu)]^\perp$. First, for all $t\in[0,T]$ and $f\in D_{e,\nu}(A')$

$$\int_{[0,t]\times E}\gamma(B^\perp,\nabla f)(s,x)\,\nu_s(dx)ds=0.$$

Indeed, we may apply the orthogonality property (which holds on the whole time interval $[0,T]$) to $\psi_n\nabla f=\nabla(\psi_n f)$ for $\psi_n\in C^\infty([0,T])$ which converges pointwise to $\mathbb{1}_{[0,t]}$ and satisfies $0\le\psi_n\le 1$.

So ν satisfies the $(B+B^\perp,D_{e,\nu}(A'))$-wFP equation. The proof is then the same as in Theorem 3.4, if we replace $B_{n,p}$ by $B_{n,p}+B^\perp_{n,p}$, since $I(Q^k\mid Q_{n,p})$ is less than $\|B\|^2_{L^2_\nu}+\|B^\perp\|^2_{L^2_\nu}$.

b) See the Proposition 5.6 in [CaL1]. ■

In [CaL1], we showed that in some cases, one can replace $D_{e,\nu}(A')$ in the definition of $H^{-1}(\nu)$ by a smaller set. Looking at the above proof, we see that what is really needed is that $B^\perp$ is orthogonal to any ∇G, obtained in the auxiliary construction (3.1). Actually, in this construction $\frac{dQ^*_{u,x}}{dP_{x,u}}$ belongs to all the L^p-spaces and we may apply the Case 2 of Proposition 2.3 to show that $C^G_\cdot$ is a (true) $P_{x,u}$-martingale which belongs to all the $L^p(P_{x,u})$. This yields

$$H^{-1}(\nu)=\{\nabla f\,;\, f\in D_{e,\nu}(A')\text{ such that for all }(u,x),\ C^f\text{ is a }P_{x,u}\text{-martingale}$$
$$\text{which belongs to all the }L^p\}.$$

Another interesting point would be to know if this is possible to replace $D_{e,\nu}(A')$ by $D(A')$. In general, we do not know if it is possible. We shall study some specific examples later.

Remark. Assume that for all $t\in[0,T]$, supp $\nu_t=E$ (or more generally $=F$, for a fixed closed subset of E). Let Q be a Markov element of A^I_ν, $Q_x=Q(\cdot\mid X_0=x)$. Then $Q_x(\xi=+\infty)=1$,

ν_o-a.e. An application of the Markov property shows that actually $Q_x(\xi = +\infty) = 1$ for all x (resp. all $x \in F$), except for some "exceptional" x's. For such nice realization of a (F-valued) Markov process with generator $A + \Gamma(B, \nabla)$ see e.g. [CaF].

If we define a generalized "nodal set": $N = \{B = +\infty\}$, we do not know whether the process hits N or not. For such a study see [MeZ] (and also [CaF]).

4. Large deviations and applications

We follow the lines of [CaL2] by proving the equality of various functionals with the help of large deviations results. We will then study the nonvacuity of A^I_ν.

In this section, we fix once for all a weakly continuous flow $\nu = (\nu_t)_{t\in[0,T]}$ and write P in place of P_{ν_o}. The origin of the study of A^I_ν is the large deviation principle for the $M_1(E)$-valued empirical processes

$$\overline{X}^N : t \in [0,T] \mapsto \overline{X}^N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)},\ N \geq 1$$

where $(X_i)_{i\geq 1}$ is a sequence of independent E-valued processes. Here, instead of identically distributed $(X_i)_{i\geq 1}$, we consider particles with laws $(P_{u_i})_{i\geq 1}$ and assume that

$$(4.1)\qquad \begin{cases} i) \quad \text{either} \quad \lim\limits_{N\to\infty} \dfrac{1}{N}\sum\limits_{i=1}^N \delta_{u_i} = \nu_o \text{ for the topology } \sigma(M_1(E), B^*_b(E)) \\ ii) \quad \text{or} \quad \lim\limits_{N\to\infty} \dfrac{1}{N}\sum\limits_{i=1}^N \delta_{u_i} = \nu_o \text{ for the topology } \sigma(M_1(E), C_b(E)) \\ \qquad\qquad \text{and } (P_x)_{x\in E} \text{ is Feller continuous.} \end{cases}$$

By Feller continuous, it is understood that the semigroup maps C_b into C_b.

Then, by ([DaG], Theorem 3.5 and Lemma 4.6) (see also the Theorem 2.1 of [CaL2]) and by the contraction principle, we have:

Theorem 4.1. *Let $(u_i)_{i\geq 1}$ be a sequence in E, $\mathbb{P} = \otimes_{i\geq 1} P_{u_i}$ and assume that (4.1) is fulfilled. Then, for any Borel subset A of $C([0,T], M_1(E))$ endowed with the topology of the weak convergence of time marginal laws uniformly on $[0,T]$, we have*

$$-\inf_{\nu'\in A^o} J_2(\nu') \leq \liminf_{N\to\infty} \frac{1}{N}\log \mathbb{P}(\overline{X}^N \in A) \leq \limsup_{N\to\infty} \frac{1}{N}\log \mathbb{P}(\overline{X}^N \in A) \leq -\inf_{\nu'\in \bar{A}} J_2(\nu')$$

where A^o and $\bar{A}$ are respectively the interior and the closure of A, J_2 being defined by

$$J_2(\nu') = \begin{cases} \min\limits_{Q\in A_{\nu'}} I(Q \mid P) & \text{if } \nu'_o = \nu_o \\ +\infty & \text{if } \nu'_o \neq \nu_o. \end{cases}$$

The rate function J_2 has compact level sets.

Of course, if $\nu'_o = \nu_o$ and $A^I_{\nu'} = \emptyset$, then $J_2(\nu') = +\infty$. The notation J_2 is taken from [CaL2], we use it for an easier comparison.

In order to obtain an alternate expression for J_2, we shall use a Cramér type theorem. Indeed, we may consider $\overline{X}^N$ as a random linear functional on $C_b([0,T] \times E)$, given by

$$\overline{X}^N(f) = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{T} \int_0^T f(t, X_i(t))\, dt \right), \; f \in C_b([0,T] \times E).$$

(This relaxation procedure is well known in Control theory). Now, using general results of D.A. Dawson and J. Gärtner ([DaG]), as explained in the Section 2.b of [CaL2], one gets the large deviation principle stated in Theorem 4.2 below (see [CaL2], Lemma 2.2).

Consider the relaxed flow

$$\nu'(f) = \frac{1}{T} \int_{[0,T]\times E} f(t,x)\, \nu'_t(dx)dt, \; f \in C_b([0,T] \times E).$$

The space of relaxed flows, denoted $\mathcal{RF}$, is endowed with the relative topology $\sigma(C_b^\sharp([0,T] \times E), C_b([0,T] \times E))$, where $C_b^\sharp$ stands for the algebraic dual space of C_b.

Theorem 4.2. *Suppose that (4.1) holds. Then, for any Borel subset A of $\mathcal{RF}$, we have*

$$- \inf_{\nu' \in A^o} J_1(\nu') \le \liminf_{N\to\infty} \frac{1}{N} \log I\!P(\overline{X}^N \in A) \le \limsup_{N\to\infty} \frac{1}{N} \log I\!P(\overline{X}^N \in A) \le - \inf_{\nu' \in \bar{A}} J_1(\nu')$$

where

$$J_1(\nu') = \begin{cases} \sup\limits_{c \in C_b([0,T]\times E)} \Big\{ \dfrac{1}{T} \displaystyle\int_{[0,T]\times E} c(t,x)\, \nu'_t(dx)dt \\ \qquad\qquad - \displaystyle\int_E \log E^{P_u}[\exp \frac{1}{T} \int_0^T c(t, X_t)\, dt]\, \nu_o(du) \Big\} & \text{if } \nu'_o = \nu_o \\ +\infty & \text{if } \nu'_o \neq \nu_o. \end{cases}$$

Notice that in Theorem 4.1, we can replace the topology by Theorem 4.2's one, since the transformation arising in the contraction is still continuous. Both J_1 and J_2 are then lower semicontinuous (J_2 has compact level sets). Therefore, by the uniqueness of a lower semicontinuous rate function on a regular space (see [DeZ]), one obtains the following

Corollary 4.3. *Under the assumption (4.1), $J_1 = J_2 (= J)$.*

Remark. In J_1, we can replace the supremum over all $c \in C_b([0,T] \times E)$ by the supremum over all $c \in \mathcal{C}$, provided that

i) $\mathcal{C}$ is a subalgebra of $C_b([0,T] \times E)$,

ii) $1\!\!1$ belongs to $\mathcal{C}$,

iii) $\mathcal{C}$ generates the Borel σ-field of $[0,T] \times E$.

This fact is easily seen, building a sequence $(c_n)_{n\ge 1}$ of $\mathcal{C}$ which converges pointwise to c and such that $\|c_n\|_\infty \le 1 + \|c\|_\infty$ (thanks to the properties of $\mathcal{C}$) and then applying the bounded convergence theorem.

As in [CaL1], [CaL2] and [DaG] (see also [Föl] and other references in [CaL1]), we want to give other variational descriptions of $J(\nu)$. If $J_2(\nu) < +\infty$, then $A_\nu^I \neq \emptyset$ and the minimizing Q^ν was described at Theorem 3.6. As in [DaG] or ([CaL1], Theorem 5.9, 1) and 2)), one can immediately state:

Theorem 4.4. *Let us define*

$$J_3(\nu) = \sup_{f \in D_{e,\nu}(A')} \{\int_E f(T,x)\,d\nu_T - \int_E f(0,x)\,d\nu_0 - \int_{[0,T]\times E} A'f(s,x)\,\nu_s(dx)ds - \frac{1}{2}\|\nabla f\|^2_{L^2_\nu}\}.$$

We have

1. $J_2(\nu) \geq J_3(\nu)$.
2. *If* $J_2(\nu) < +\infty$, *then* $J_2(\nu) = J_3(\nu) = I(Q^\nu \mid P_{\nu_o})$.
3. *For all* $Q \in A^I_\nu$, $I(Q \mid P_{\nu_o}) = I(Q \mid Q^\nu) + I(Q^\nu \mid P_{\nu_o})$.

Notice that for this result, the hypothesis (4.1) is unnecessary.

It is thus natural to ask wether J_2 and J_3 match everywhere or not, which is equivalent to the fact that $\{\nu\,;\,J_2(\nu) = +\infty\} = \{\nu\,;\,J_3(\nu) = +\infty\}$. The next proposition states that the finiteness of J_3 is equivalent to the existence of a solution to a weak Fokker-Planck equation.

Proposition 4.5. *The following statements are equivalent.*

1. *There exists* $B \in L^2_\nu$ *such that* ν *satisfies the* $(B, D_{e,\nu}(A'))$*-wFP equation.*
2. *There exists* $B \in L^2_\nu$ *such that for all* $f \in D_{e,\nu}(A')$,

$$(*) \qquad \int_{[0,T]\times E} (A'f + \gamma(B,\nabla f))(s,x)\,\nu_s(dx)ds = \int_E f(T,x)\,d\nu_T - \int_E f(0,x)\,d\nu_0.$$

3. *There exists* $B^\nu \in H^{-1}(\nu)$ *such that* $(*)$ *is satisfied.*
4. $J_3(\nu) < +\infty$.

Proof. 1) $\Rightarrow$ 2) $\Rightarrow$ 3) (projection onto $H^{-1}(\nu)$) $\Rightarrow$ 1) (see the first part of Proposition 3.3).
3) $\Rightarrow$ 4) since $J_3(\nu) = \frac{1}{2}\|B^\nu\|^2_{L^2_\nu}$.
4) $\Rightarrow$ 3) thanks to the following argument. Assume that $J_3(\nu) < +\infty$. Then, if $f \in D_{e,\nu}(A')$ and $\nabla f = 0$, for all $\lambda \in \mathbb{R}$ we have

$$\lambda(\int_E f(T,x)\,d\nu_T - \int_E f(0,x)\,d\nu_0 - \int_{[0,T]\times E} A'f(s,x)\,\nu_s(dx)ds) \leq J_3(\nu)$$

and so the left hand side vanishes identically. This shows that the map

$$\begin{cases} \{\nabla f\,;\,f \in D_{e,\nu}(A')\} & \to \quad \mathbb{R} \\ \qquad \nabla f & \mapsto \quad \mathcal{L}(f) = \int_E f(T,x)\,d\nu_T - \int_E f(0,x)\,d\nu_0 \\ & \qquad\qquad - \int_{[0,T]\times E} A'f(s,x)\,\nu_s(dx)ds \end{cases}$$

is well defined, linear and continuous if $\{\nabla f\,;\,f \in D_{e,\nu}(A')\}$ is equipped with the hilbertian seminorm $\|\nabla f\|_{L^2_\nu}$. By Riesz' representation theorem, there exists $B^\nu \in H^{-1}(\nu)$ such that $\mathcal{L}(f) = \gamma(B^\nu, \nabla f)$. ■

Looking at J_3, we recognize a Hamilton-Jacobi operator whose inverse can be easily computed. Indeed, for $c \in C_b([0,T]\times E)$ (define $c(s,x) = c(T,x)$ if $s \geq T$) define

$$(4.2) \qquad \begin{cases} g_c(t,x) &= E^{P_{t,x}}[\exp \int_0^{T-t} c((u_s, X_s))\,ds] & \text{if } t < T \\ g_c(t,x) &= 1 & \text{if } t \geq T \\ f_c(t,x) &= \log g_c(t,x) \end{cases}$$

(notice that g_c is bounded from below by a positive constant). Applying the Markov property, we get

$$E^{P_{u,x}}[g_c(u_t, X_t)] = E^{P_{u,x}}[\exp \int_{u_t}^{T} c(s, X_{s-u})\, ds].$$

But, since c and X are continuous, for $t \in [0, T-u]$, $t \mapsto \int_{u_t}^{T} c(s, X_{s-u})\, ds$ is of class C^1 and $\frac{d}{dt} \exp(\int_{u_t}^{T} c(s, X_{s-u})\, ds) = -c((u_t, X_t)) \exp(\int_{u_t}^{T} c(s, X_{s-u})\, ds)$. Therefore, $E^{P_{u,x}}[g_c(u_t, X_t) - g_c(u_0, X_0) + \int_0^t \mathbb{I}_{s<T} c(u_s, X_s) g_c(u_s, X_s)\, ds] = 0$, which yields, thanks again to the Markov property,

$$g_c \in D(A') \quad \text{and} \quad A' g_c(t,x) = -c(t,x) g_c(t,x) \mathbb{I}_{t<T}, \tag{4.3}$$

so that

$$f_c \in D(A'), f_c(T,x) = 0, \nabla f_c = \frac{\nabla g_c}{g_c} \quad \text{and} \quad A' f_c + \frac{1}{2}\Gamma(f_c) + c = 0 \text{ on } [0,T[\times E. \tag{4.4}$$

It follows that

$$\int_{[0,T]\times E} c(s,x)\, \nu_s(dx) ds - \int_E \log E^{P_x}[\exp \int_0^T c(s, X_s)\, ds]\, \nu_o(dx)$$
$$= \int_E f_c(T,x)\, d\nu_T - \int_E f_c(0,x)\, d\nu_0 - \int_{[0,T]\times E} (A' f_c + \frac{1}{2}\Gamma(f_c))(s,x)\, \nu_s(dx) ds.$$

Hence, provided that $f_c \in D_{e,\nu}(A')$, which is actually equivalent to $\nabla f_c \in L^2_\nu$ (or $\nabla g_c \in L^2_\nu$), one gets $J_1(\nu) \le J_3(\nu)$ (the normalizing constant $\frac{1}{T}$ is irrelevant). Let us summarize our results.

Theorem 4.6. *Assume that*

$$\begin{cases} i) & \text{there exists a sequence } (x_i)_{i\ge 1} \text{ such that } \nu_o = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \delta_{x_i} \\ & \text{for the topology } \sigma(M_1(E), B_b(E)) \\ \text{or} & \\ ii) & (P_x)_{x\in E} \text{ is Feller continuous.} \end{cases}$$

(the other condition in (4.1), ii) is always satisfied in a Polish space).
Moreover, assume that
(HC) there exists a subalgebra $\mathcal{C}$ of $C_b([0,T]\times E)$ with $\mathbb{I} \in \mathcal{C}$ which generates the Borel σ-field of $[0,T]\times E$ and such that for $c \in \mathcal{C}$, the function ∇f_c defined by (4.2) belongs to L^2_ν.

Then,

$$J_1(\nu) = J_2(\nu) = J_3(\nu).$$

(4.5) Remark. If in addition $(P_x)_{x\in E}$ is Feller continuous, then g_c and f_c are continuous.

Corollary 4.7. *Under the hypotheses of Theorem 4.6, if ν satisfies the $(B, D_{e,\nu}(A'))$-wFP equation for some $B \in L^2_\nu$, then the (B, ν_o, P_{ν_o})-FM belongs to A^I_ν, i.e. there exists a Markovian probability measure Q such that $I(Q \mid P_{\nu_o}) < +\infty$, $Q \circ (X_t)^{-1} = \nu_t$ for all t and Q is a solution to the martingale problem $\mathcal{M}(A + \gamma(B,\nabla), D_{e,\nu}(A'), \nu_o)$. Using the terminology of Section 2: ν is admissible.*

Proof of Corollary 4.7. Apply Proposition 4.5, Theorem 4.6 and Theorem 3.6. ■

Corollary 4.7 is a general setting of E. Carlen's existence result ([Car]). Notice that, in contrast with [Car], we do not assume any "dual energy condition" on the backward drift. Moreover, we obtained in Section 3, a complete description of all possible Markovian Schrödinger (or Nelson) processes associated with a given flow ν. But, let us go on for a while discussing large deviations properties.

In order to compare J_1 and J_3, we used (4.4). But, it is also possible to directly use (4.3) and write

$$\begin{aligned}(4.6)\qquad &\int_{[0,T]\times E} c(s,x)\,\nu_s(dx)ds - \int_E \log E^{P_x}[\exp\int_0^T c(s,X_s)\,ds]\,\nu_o(dx)\\ &= -\int_{[0,T]\times E} \frac{A'g_c}{g_c}(s,x)\,\nu_s(dx)ds - \int_E \log g_c(0,x)\,\nu_o(dx).\end{aligned}$$

Define

$$\begin{aligned}C_{\exp} = \{g \in D_{e,\nu}(A')\,;\, g \geq 1, g(t,x) = 1,\ \forall x \in E, t \geq T,\ C^g \text{ is a bounded}\\ P_{u,x}\text{-martingale for all } (u,x),\ g \text{ and } \frac{A'g}{g} \in C_b([0,T]\times E)\}.\end{aligned}$$

Theorem 4.8. *Assume that $(P_x)_{x\in E}$ is Feller continuous. Then,*

$$C_{\exp} = \{g_c\,;\, c \geq 0, c \in C_b\}$$

and

$$J_1(\nu) = J_4(\nu) := \sup_{g\in C_{\exp}} \left\{ -\int_{[0,T]\times E} \frac{A'g}{g}(s,x)\,\nu_s(dx)ds - \int_E \log g(0,x)\,\nu_o(dx). \right\}$$

Proof. If $c \in C_b$ and $c \geq 0$, then g_c satisfies all the properties of $C_{\exp}$ except perhaps the continuity assumption. This last property is ensured by the continuity of $x \mapsto P_x$ (Feller property). Conversely, let $g \in C_{\exp}$ and $c = -\dfrac{A'g}{g}$. We can define g_c as in (4.2). We are going to prove that $g = g_c$.

Define $\tau = \inf\{t \geq 0\,;\, g_c(u_t,X_t) = g(u_t,X_t)\}$. τ is less than $T-u$, so it is a bounded stopping time and for all (u,x), $g_c(u_\tau,X_\tau) = g(u_\tau,X_\tau)$ thanks to the continuity assumptions. From the optional sampling theorem, it comes out that for all (u,x)

$$(g_c - g)(u,x) + E^{P_{u,x}}[\int_0^\tau c(g_c-g)(u_s,X_s)\,ds] = 0.$$

Since $c \geq 0$ and by continuity: $c(g_c-g)(u_s,X_s)$ and $c(g_c-g)(u,x)$ have the same sign up to time τ, $P_{u,x}$-a.s., so both terms in the above sum are equal to 0. In particular $g_c(u,x) = g(u,x)$ for all (u,x).

Finally, $J_1 = J_4$ thanks to (4.6), since the supremum in J_1 can be taken over all nonnegative c. ■

Remarks. i) This theorem (as well as nonentropic cases) can be derived using another large deviations approach: the MEM's method introduced by [DcG] and developed by F. Gamboa and E. Gassiat (see e.g. [GaG]). For a finite flow (i.e. discrete time) see [CaG]. But a relaxation method similar to the present paper's one, should allow to consider the general continuous flow of marginals with the methods of [CaG].

ii) At least at a formal level, Theorem 4.8 is similar to the results of Lemma 4.2.35 and Theorem 4.2.23 of [DeS].

5. Examples of admissible flows

Here again, ν is a weakly continuous flow of marginal laws and $P = P_{\nu_o}$. We shall assume throughout this section that

(5.1) The hypothesis of Theorem 4.6 is satisfied.

The goal of this section is to give sufficient conditions for ν to be admissible, i.e. for A^I_ν to be nonempty, i.e. for $J_2(\nu)$ to be finite. According to Theorem 4.6, when (HC) holds, it is enough to check the finiteness of $J_3(\nu)$, which is equivalent, thanks to Proposition 4.5, to the following:

(5.2) There exists $B \in L^2_\nu$ such that ν satisfies the $(B, D_{e,\nu}(A'))$-wFP equation.

Assuming (5.2), we thus have two possibilities:

α. to find sufficient conditions on ν for (HC) to hold with $\mathcal{C} = C_b([0,T] \times E)$, or

β. to find sufficient conditions on P for (HC) to hold for a well chosen $\mathcal{C}$ and any ν satisfying (5.1) and (5.2).

Another possibility would be to use the "approximation procedure" of Section 3 (see (3.1)) as in [CaL1] in order to give a direct construction. But, here again, the main point is to prove that ∇G belongs to L^2_ν for some suitable G (see (3.1.iii)), and this is of course of the same nature as proving that (HC) holds.

In the Section 4 of [CaL1] and in the Section 3 of [CaL2], we have studied these situations in the case where P is the law of a $\mathbb{R}^d$-valued diffusion process. Here, we shall only give some examples for which answers to the questions α or β are not too hard to get. As was expected, these examples cover a large part of the "usual processes".

In a general setting, the most natural approach is the one in α, and we will start our study with this problem.

A. When does (HC) hold with $\mathcal{C} = C_b([0,T] \times E)$?

Since relative entropy does not increase under measurable transforms, for any admissible ν, we have

for all $t \in [0,T]$, $I(\nu_t \mid \mu_t) < +\infty$ (in particular $\nu_t \ll \mu_t$) where $\mu_t = P \circ (X_t)^{-1}$.

Conversely, assume that for all $t \in [0,T]$, $\nu_t \ll \mu_t$ and define

$$\rho(\cdot, t) = \frac{d\nu_t}{d\mu_t}.$$

We want to find sufficient conditions on ρ for (HC) to hold with $\mathcal{C} = C_b([0,T] \times E)$, i.e. for $\int_{[0,T]\times E} \Gamma(g_c)\, \nu_s(dx)ds$ to be finite for all $c \in C_b([0,T] \times E)$. But

(5.3)

$$\int_{[0,T]\times E} \Gamma(g_c)\, \nu_s(dx)ds = \int_{[0,T]\times E} \Gamma(s,x)\rho(s,x)\, \mu_s(dx)ds = E^P[\int_0^T \Gamma(g_c)(s, X_s)\rho(s, X_s)\, ds].$$

The main estimate is given in the following lemma.

Lemma 5.1. *There exists $\lambda_o > 0$ such that for all $\lambda < \lambda_o$, $\sup_x E^{P_x}[\exp \lambda \langle C^{g_c}\rangle_T] < +\infty$.*

Proof. Since $A'g_c = -cg_c$ is bounded, C^{g_c} is a bounded P_x-martingale for all $x \in E$, with a uniform bound K (which does not depend on x). Applying BDG inequalities, we obtain for $1 \le q < +\infty$

$$\sup_{x\in E} E^{P_x}[\langle C^{g_c}\rangle_T^q] \le (4q)^q K^{2q}.$$

Thus, applying Stirling's formula (see [CaL1], 2.7 for the argument) and taking the quadratic growth of $\lambda \mapsto \langle C^{\lambda g_c}\rangle_T = \lambda^2 \langle C^{g_c}\rangle_T$ into account, we obtain for λ small enough: $\sup_{x\in E} E^{P_x}[\exp \lambda \langle C^{g_c}\rangle_T] < +\infty$. ■

One can then obtain:

Proposition 5.2. *Assume that (5.1), (5.2) and (5.3) hold. Assume in addition that $\operatorname{ess\,sup}_{t\in[0,T]} \rho(t, X_t) \in L^{\tau^*}(P)$, where L^{τ^*} is the Orlicz space associated with: $\tau^*(u) = (u + 1)\log(u + 1) - u, u \ge 0$. Then, ν is admissible.*

Corollary 5.3. *If $\rho \in B_b([0, T] \times E)$, then ν is admissible.*

See ([CaL1], 4.48) for the same result in the case of $\mathbb{R}^d$-valued diffusions.

Proof. According to (5.3)

$$\int_{[0,T]\times E} \Gamma(g_c)(s, x)\, \nu_s(dx)ds = E^P[\int_0^T \Gamma(g_c)(s, X_s)\rho(s, X_s)\, ds]$$

$$\le E^P[\operatorname*{ess\,sup}_{t\in[0,T]} \rho(t, X_t) \int_0^T \Gamma(g_c)(s, X_s)\, ds].$$

Thanks to Lemma 5.1, $\int_0^T \Gamma(g_c)(s, X_s)\, ds = \langle C^{g_c}\rangle_T$ belongs to $L^\tau(P)$ with $\tau(u) = e^u - u - 1, u \ge 0$. It remains to apply Hölder's inequality in Orlicz spaces to conclude. ■

Of course, except in the bounded case of Corollary 5.3, Proposition 5.2 is not really tractable. If E is compact and ρ is continuous, one may apply Corollary 5.3. As in [CaL1], one can expect to relax the boundedness assumption into a local boundedness one, in the case of a σ-compact space E. A natural method to improve Corollary 5.3 would be to show that

(5.4) if $(\nu_t)_{t\in[0,T]}$ $(= \rho(\cdot, t)\mu_t)$ satisfies (5.2), one can find a sequence $(\rho_n)_{n\ge1}$ of bounded densities satisfying (5.2) and such that $J_3(\nu_n) \underset{n\to\infty}{\longrightarrow} J_3(\nu)$.

In the general case, we do not even know whether (5.4) is true or not. But, if ρ and P_x are smooth enough, one can show that (5.4) holds. Here, we will restrict ourselves to the simpler symmetric case, as in [MeZ].

Theorem 5.5. *Assume that $(P_x)_{x\in E}$ is a μ-symmetric Feller process. Let ν be a probability measure on $(E, \mathcal{E})$ and the associated stationary flow $\nu = \nu_t$ for $t \in [0, T]$. Assume that $I(\nu \mid \mu) < +\infty$, $\frac{d\nu}{d\mu} = \rho$ and $\rho^{1/2}$ belongs to the domain of the Dirichlet form associated with (P_x, μ). Then, ν is admissible, i.e. there exists Q such that $I(Q \mid P_\nu) < +\infty$ and $Q \circ (X_t)^{-1} = \nu$ for all $t \in [0, T]$.*

Proof. Denote $(\mathcal{D}, D(\mathcal{D}))$ the above Dirichlet form. Recall that for $f \in D(\mathcal{D})$, there exists $\widetilde{\nabla} f = (\widetilde{\nabla}^n f)_{n\geq 1}$ such that $\mathcal{D}(f) = \frac{1}{2}\sum_{n\geq 1}\int |\widetilde{\nabla}^n f|^2\, d\mu$. Furthermore, if $f \in D_e(A)$ and $Af \in L^2(\mu)$, then $f \in D(\mathcal{D})$ and $\widetilde{\nabla} f = \nabla f$ (in $L^2(\mu)$) (see [BoH]). Actually, this fact extends to the functions $f \in D_{e,\mu}(A)$ which are continuous. Indeed, for such an f

$$\begin{aligned}\frac{1}{t}\int f(f - T_t f)\, d\mu &= \frac{1}{t}\int E^{P_x}[f(X_0)(f(X_0) - f(X_t))]\, d\mu \\ &= -\frac{1}{t}\int E^{P_x}[f(X_0)\int_0^t Af(X_s)\, ds]\, d\mu \\ &= -\frac{1}{t}\int Af(x)E^{P_x}[\int_0^t f(X_s)\, ds]\, d\mu.\end{aligned}$$

Thanks to Lebesgue bounded convergence theorem and since f is continuous and bounded, it follows that

$$\lim_{t\downarrow 0}\frac{1}{t}\int f(f - T_t f)\, d\mu = -\int f(x)Af(x)\, d\mu.$$

According to a well known result for Dirichlet forms (see e.g. [Fuk], Lemma 1.3.4), this means that $f \in D(\mathcal{D})$.

Now, since $\rho^{1/2} \in D(\mathcal{D})$, $\rho'_k := [(\rho \vee 1/k) \wedge k]^{1/2}$ and $\rho_k = [(\rho \vee 1/k) \wedge k]$ also belong to $D(\mathcal{D})$. Let $d\nu_k = c_k \rho_k d\mu$ (where c_k is a normalizing constant, which clearly tends to 1 as $k \to \infty$.) Then, $D_{e,\nu_k}(A) = D_{e,\mu}(A)$. In particular, if $f \in D_{e,\nu_k}(A)$ is continuous, $f \in D(\mathcal{D})$ and

$$\int Af\, \rho_k\, d\mu = -\frac{1}{2}\int \nabla f \cdot \widetilde{\nabla}\rho_k\, d\mu = -\int \nabla f \cdot \widetilde{\nabla}\rho'_k\, \rho'_k\, d\mu = -\int \nabla f \cdot \frac{\widetilde{\nabla}\rho'_k}{c_k \rho'_k}\, d\nu_k.$$

It follows that ν_k satisfies the $(\frac{\widetilde{\nabla}\rho'_k}{\rho'_k}, D_{e,\nu_k}(A) \cap C^0)$-wFP equation, i.e. $J_3(\nu_k) < +\infty$ (if one restricts the supremum to the continuous functions f). But ρ_k is bounded, $\frac{\widetilde{\nabla}\rho'_k}{\rho'_k} \in L^2_{\nu_k}$ (since $\widetilde{\nabla}\rho'_k \in L^2(\mu)$) and (P_x) is Feller continuous. According to Corollary 5.3 and the Remark (4.5), it follows that ν_k is admissible, i.e. $J_2(\nu_k) < +\infty$ (with respect to the measure P_{ν_k}). Furthermore

$$J_2(\nu_k) = \inf\{I(Q \mid P_{\nu_k})\,;\, Q \circ (X_t)^{-1} = \nu_k\} \leq \frac{1}{2}\mathcal{D}(\rho'_k).$$

We cannot directly take the limit in k because the reference measure P_{ν_k} depends on k. But, if $Q \circ (X_0)^{-1} = \eta$,

$$I(Q \mid P_\mu) = I(Q \mid P_\eta) + I(\eta \mid \mu).$$

Since $I(\eta \mid \mu) < +\infty$, $I(\nu_k \mid \mu) = \int(\log c_k + \log \rho_k)c_k\rho_k\, d\mu < +\infty$, and $\lim_{k\to\infty} I(\nu_k \mid \mu) = I(\nu \mid \mu)$ by Lebesgue's theorem again. Denote

$$J'_2(\nu') = \inf\{I(Q \mid P_\mu)\,;\, Q \circ (X_t)^{-1} = \nu'_t\}.$$

Then,

$$J'_2(\nu_k) = I(\nu_k \mid \mu) + J_2(\nu_k) \leq I(\nu_k \mid \mu) + \frac{1}{2}\mathcal{D}(\rho'_k).$$

But $I(\nu_k \mid \mu)$ and $\mathcal{D}(\rho_k')$ converge respectively to $I(\nu \mid \mu)$ and $\mathcal{D}(\rho^{1/2})$. It follows that $J_2'(\nu_k)$ is bounded. Since J_2' is lower semicontinuous for the weak topology and ν_k converges weakly to ν, then $J_2'(\nu)$ is finite. ■

Remarks. i) One can prove that $J_2'(\nu) = I(\nu \mid \mu) + \frac{1}{2}\mathcal{D}(\rho^{1/2})$ (see [CaF] for the Brownian case).

ii) It can be proved in many cases (for instance, the finite dimensional case as in [CaF]) that the assumption $\rho^{1/2} \in D(\mathcal{D})$ is also necessary for $\rho d\mu$ to be admissible. This situation is quite satisfactory in the symmetric case.

iii) This result is related to recent works on Dirichlet forms on non locally compact spaces (see [MaR], [AlR], [Son]), especially to the extension of the Girsanov formula in this context (see in particular [ARZ]). Notice that an hypothesis $I(\nu \mid \mu) < +\infty$ also appears in these works, since $\log \rho$ is assumed to be in $L^2(\rho d\mu)$.

iv) Our approach can be extended to nonsymmetric cases with additional material.

We shall see how to deal with question β.

B. How to use a differential structure

To choose a $\mathcal{C}$ in such a way that (HC) holds, seems to be hard to do unless one can use a "universal" differential structure on E which is connected to the stochastic structure, i.e. to $(\nabla^n)_{n\geq 1}$. This leads us to require that E is equipped with a linear structure (i.e. a tangent space at each point). Here again, we shall only consider a few examples without giving all the details.

B1. Finite dimensional manifolds

Assume that E is a d-dimensional connected C^∞ manifold (without boundary, but possibly $E = \mathbb{R}^d$ since we do not assume any compactness). The natural candidate for $\mathcal{C}$ would be $C_o^\infty \vee \mathbb{1}$: the algebra generated by the constant function $\mathbb{1}$ and the space of compactly supported C^∞ functions defined on E. But, it is known (see [Jac] 13.53.3) that if C_o^∞ is included in the (true) domain of the generator A of $(P_x)_{x\in E}$ and if the semigroup is Feller continuous, then A has (in local coordinates) the form

$$A = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i}$$

with the coefficients $a_{i,j}$ and b_i bounded and continuous. Furthermore, in this case

$$\Gamma(f) = \sum_{i,j=1}^{d} a_{i,j}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}$$

for $f \in C_o^\infty(E)$ with its support in a local chart. (One can choose $\Theta = C_o^\infty(E) \vee \mathbb{1}$ or relax the boundedness assumption in the definition of $D_e(A)$ and take for $(\varphi_n)_{n\geq 1}$ a countable family of coordinate changes.)

This has already been studied in [CaL1] and [CaL2], at least for $E = \mathbb{R}^d$. Notice that in the uniformly elliptic case, one can use known regularity results on Hamilton-Jacobi equations, see e.g. [Lio], in order to recover Theorem 4.42 of [CaL1] by means of the method which is developped in Section 4.

The manifold-valued case is completely similar. Indeed, imbed E into $\mathbb{R}^m$ $(m \geq 2d+1)$ appealing to Whitney's theorem and assume that A is the restriction to E of the operator

$$\frac{1}{2}\sum_{i,j=1}^{d} \sigma_{i,j}\frac{\partial}{\partial x_i}\left(\sum_{k=1}^{d}\sigma_{i,j}\frac{\partial}{\partial x_k}\right)+\sum_{i=1}^{d} b_i\frac{\partial}{\partial x_i} \tag{5.5}$$

where $\sigma_{i,j}$ and b_i belong respectively to C_b^2 and C_b^1 (C_b^k is the space of C^k functions with bounded derivatives of order 0 to k). Then, applying the differentiability result with respect to the initial data (see e.g. [Kun] or [IkW], pp. 254–255), it is easily shown that g_c belongs to $C_b^1(E)$ for any $c \in C_o^\infty(E) \vee \mathbb{1}$ and $t \in [0,T]$. Accordingly, (HC) holds with $\mathcal{C} = C_o^\infty(E) \vee \mathbb{1}$. In the time-dependent case, one can relax the differentiability assumption in the time direction. In the elliptic case again, known results on Hamilton-Jacobi equations could be used. Notice that W. A. Zheng ([Zhe]) obtained a similar result in the case of a compact manifold, compactness being a key point in his approach.

B2. Finite dimensional manifolds with boundary

Let E be a d-dimensional connected C^∞ manifold with a smooth boundary ∂E which is locally on one side. For simplicity, one can assume that $E = \overline{D} = D \cup \partial D$ where $D = \{x\,;\,\psi(x) > 0\}$ and $\partial D = \{x\,;\,\psi(x) = 0\}$ for a given $\psi \in C_b^\infty(\mathbb{R}^d)$, but the results still hold in more general contexts. For more details on what follows, we refer to [Cat] and the references contained therein.

We consider $(P_x)_{x\in E}$: the law of a reflected diffusion, i.e. whose generator coincides with the one defined in (5.5) for all $f \in C_b(E)$ satisfying an oblique derivative condition on ∂E, i.e.

$$\beta\cdot\frac{\partial}{\partial n}f = 0 \ \text{ on } \partial E$$

for a given vector field β defined on ∂E, $\frac{\partial}{\partial n}$ being the inward pointed normal derivative on ∂E (for instance, if $|\nabla\psi| \equiv 1$ on ∂E, one can identify $\frac{\partial}{\partial n}$ with $\nabla\psi$, this will done in the sequel).

For simplicity, we assume that

(5.6) $\sigma_{i,j}, b_i$ and β are C_b^∞ functions

(more precisely: are the restrictions to E of smooth functions defined on the whole space, but after imbedding; this is not a restriction, thanks to Whitney's theorem). We refer to [Cat] for the minimal differentiability assumptions required for the following to hold.

In addition, it is assumed that

(5.7)
$$\begin{cases} i) & |\nabla\psi| \equiv 1 \text{ on } \partial E, & \\ ii) & \beta\cdot\nabla\psi \geq c_o > 0 \text{ on } \partial E, & \text{(strong transversality assumption)} \\ iii) & \sum_i\left(\sum_j \sigma_{i,j}\frac{\partial\psi}{\partial x_j}\right)^2 \geq c_1 > 0 \text{ on } \partial E, & \text{(i.e. } \partial E \text{ is uniformly noncharacteristic).} \end{cases}$$

Under all these assumptions, one knows that $(P_x)_{x\in E}$ exists and can be built via the resolution of a stochastic differential system with reflection (see [IkW]). Moreover, the solution is (weakly) unique and Feller continuous. Let

$$\Theta = \{f \in C_b^\infty(E)\,;\,\beta\cdot\frac{\partial f}{\partial n} = 0 \text{ on } \partial E\}.$$

Then, Θ is a core for $(P_x)_{x\in E}$ and we can use all the material of this work, with $\Gamma(f) = \sum_{i,j=1}^{d} a_{i,j} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}$ $(a = \sigma^*\sigma)$ as before. Here again, we may take $\mathcal{C} = C_0^\infty(E) \vee \mathbb{1}$, but we cannot anymore apply the arguments of the previous part to prove that $g_c \in C_b^1(E)$, since there is no regular flow associated with the reflected diffusion. Hence, we have to make additional assumptions.

Theorem 5.5. *In addition to (5.6) and (5.7), assume that A satisfies a uniform Hörmander's condition. Then,* $g_c \in C_b^\infty(E)$ *for all* $t \in [0,T]$ *and* $c \in C_o^\infty(E) \vee \mathbb{1}$, *and (HC) holds for* $\mathcal{C} = C_o^\infty(E) \vee \mathbb{1}$.

The above result follows from Theorem 4.4 of [Cat]. We also refer to this paper for the precise meaning of a uniform Hörmander's condition (called (HG.unif), there) as well as for known analytical results in the uniformly elliptic case. Notice that one cannot treat the case of a Ventcel like boundary condition, since in this case, the corresponding Θ is not an algebra and the "carré du champ" Γ is not anymore absolutely continuous with respect to ds.

B3. Infinite dimensional linear spaces

The method of B1 can be extended to any linear space provided that one can represent P_x by a stochastic process $(X_t(x))_{t\in[0,T]}$ which depends smoothly on x. This can be done, for instance, for the solutions of stochastic differential equations in Hilbert spaces with smooth coefficients.

The same method also applies in the case of an abstract Wiener space (μ, H, E) with P_x the law of the standard Brownian motion (or the Ornstein-Uhlenbeck process) starting from x. In this case, the "usual" gradient is the Gâteaux derivative in the directions of H (the Cameron-Martin space) and $\mathcal{C}$ can be chosen as

$$\mathcal{C} = \{c = \varphi(\langle l_1,\cdot\rangle,\ldots,\langle l_n,\cdot\rangle), n \geq 1, \varphi \in C_b^\infty(\mathbb{R}^n), l_1,\ldots,l_n \in E^*\},$$

where E^* stands for the dual space of E. This result can be extended to the more general situation of a symmetric process associated with an "admissible" Dirichlet form (see [BoH] or [MaR]) and a non necessarily stationary flow ν (in contrast with the situation of A1 where ν was stationary). But a precise discussion would need to introduce additional material and we shall not enter into the details here.

Another interesting situation would be the case when $E = (\mathbb{R}^d)^{\mathbb{Z}^k}$, i.e. particle systems as in [LeR], [ShS], [MNS] or [CRZ]. But, even it is trivial to extend B1 to an infinite collection of independent Brownian motions, the existence result we obtain via the Theorem 4.6 has no real interest, because the "global" finite entropy conditon is too strong. Indeed, all interesting systems will satisfy a "local" finite entropy condition (see e.g. [FöW]) but not a "global" one, or involve the "specific" entropy rather than the relative one (see e. g. [Föl]).

6. References

[AlR] S. Albeverio and M. Röckner. *Classical Dirichlet forms on topological vector spaces. Closability and a Cameron-Martin formula.* J. Funct. Anal. 88, (1990), 395–436.

[ARZ] S. Albeverio, M. Röckner and T.S. Zhang. *Girsanov transform for symmetric diffusions with infinite dimensional state space.*Ann. Probab. 21, (1993), 961–978.

[BoH] N. Bouleau and F. Hirsch. *Dirichlet forms and analysis on Wiener space.* Ed. De Gruyter, (1991).

[Car] E. Carlen. *Conservative diffusions.* Comm. Math. Phys. 94. (1984), 293–315.

[Cat] P. Cattiaux. *Stochastic calculus and degenerate boundary value problems.* Ann. Inst. Fourier 42. (1992), 541–624.

[CaF] P. Cattiaux et M. Fradon. *Entropy, reversible diffusion processes and Markov uniqueness.* Preprint Orsay & Polytechnique, (1995).

[CaG] P. Cattiaux et F. Gamboa. *Large deviations and variational theorems for marginal problems.* Preprint Orsay & Polytechnique, (1995).

[CaL1] P. Cattiaux and C. Léonard. *Minimization of the Kullback information of diffusion processes.* Ann. Inst. Henri Poincaré, Vol. 30, (1994), 83–132.
Correction. To appear in Ann. Inst. Henri Poincaré, (1995).

[CaL2] P. Cattiaux and C. Léonard. *Large deviations and Nelson processes.* Forum Math., Vol. 7, (1995), 95–115.

[CRZ] P. Cattiaux , S. Roelly and H. Zessin. *Une approche Gibbsienne des diffusions Browniennes infinidimensionnelles.*to appear in Prob. Th. Rel. Fields.

[DaG] D.A. Dawson and J. Gärtner. *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions.* Stochastics and Stoch. Rep., Vol. 20, (1987), p 247–308.

[DcG] D. Dacunha-Castelle et F. Gamboa. *Maximum d'entropie et problèmes de moments.* Ann. Inst. Henri Poincaré 26, (1990), 567–596.

[DeM] C. Dellacherie et P.A. Meyer. *Probabilités et Potentiel, Ch. XII–XVI.* (1987), Hermann, Paris.

[DeS] J.D. Deuschel and D.W. Stroock. *Large Deviations.* Pure and Applied Mathematics, 137, (1989), Academic Press, Boston.

[DeZ] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* (1993), Jones and Bartlett Publishers.

[Föl] H. Föllmer. *Random Fields and Diffusion Processes.* Cours à l'Ecole d'été de Probabilités de Saint-Flour. (1988), Lecture Notes in Mathematics 1362, Springer Verlag.

[FöW] H. Föllmer and A. Wakolbinger. *Time reversal of infinite dimensional diffusion processes.* Stoch. Proc. and Appl. 22, (1986), 59–78.

[Fuk] M. Fukushima. *Dirichlet forms and Markov processes.* Second Ed., North-Holland, (1988).

[GaG] F. Gamboa and E. Gassiat. *Bayesian methods for ill posed problems.* Preprint Orsay, (1994).

[IkW] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* (1989), North Holland.

[Jac] J. Jacod. *Calcul stochastique et problèmes de martingales.* Lecture Notes in Mathematics 714, (1979), Springer Verlag.

[Kun] H. Kunita. *Stochastic differential equations and stochastic flow of diffeomorphisms.* Ecole d'été de probabilités de Saint-Flour. Lecture Notes in Mathematics 1907, (1984), Springer-Verlag.

[LeR] G. Leha and G. Ritter. *On solutions to stochastic differential equations with discontinuous drift in Hilbert spaces.* Math. Ann., Vol. 270, (1985), 109–123.

[Lio] P.L. Lions. *Generalized solutions of Hamilton-Jacobi equations.* Research Notes in Mathematics, 69. Pitman, Boston, (1982).

[MaR] Z. Ma and M. Röckner. *An introduction to the theory of (non symmetric) Dirichlet forms.* Springer, (1992).

[MeZ] P.A. Meyer et W.A. Zheng. *Construction de processus de Nelson réversibles.* Séminaire de Probabilités XIV, Lecture Notes in Mathematics 1123, (1985), 12–26.

[MNS] A. Millet, D. Nualart and M. Sanz. *Time reversal for infinite dimensional diffusions.* Prob. Th. Rel. Fields 82, (1989), 315–347.

[Sha] M.J. Sharpe. *general theory of Markov processes.* Academic Press, (1988).

[ShS] T. Shiga and A. Shimizu. *Infinite dimensional stochasticdifferential equations and their applications.* J. Math. Kyoto Univ. 20, (1980), 395–416.

[Son] S.Q. Song. *Habilitation.* Univ. Paris VI, (1993).

[Zhe] W.A. Zheng. *Tightness results for laws of diffusion processes. Application to stochastic mechanics.* Ann. Inst. Henri Poincaré, Vol. 21, No. 2, (1985), 103–124.