ANNALES DE L'I. H. P., SECTION B

PATRICK CATTIAUX
CHRISTIAN LÉONARD

Minimization of the Kullback information of diffusion processes

Annales de l'I. H. P., section B, tome 30, n^o 1 (1994), p. 83-132.
<http://www.numdam.org/item?id=AIHPB_1994__30_1_83_0>

Ann. Inst. Henri Poincaré,

Vol. 30, n° 1, 1994, p. 83-132. *Probabilités et Statistiques*

Minimization of the Kullback information of diffusion processes

by

Patrick CATTIAUX and Christian LÉONARD

Université de Paris-Sud,
Département de Mathématiques, Bât. n° 425,
U.A.-C.N.R.S. n° 743, Équipe de Modélisation stochastique et statistique,
91405 Orsay Cedex, France

ABSTRACT. — In this paper we compute an explicit expression for the rate function of large deviations for the measure valued empirical process

$$\bar{X}^N(t)=\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i(t)}$$

where the X_i's are independent copies of a diffusion process in $\mathbb{R}^d$. This is done by minimizing the relative entropy (Kullback information) of a Probability measure Q with respect to the law P of X_i when all marginals of Q are fixed. The finiteness of the rate function is connected with the existence of conservative diffusions, with a general diffusion matrix. These diffusion processes are constructed in very general cases.

Key words : Large deviation, relative entropy, Kullback information, conservative diffusion process, Föllmer measure, Girsanov transformation.

RÉSUMÉ. — Dans cet article, nous donnons une formule explicite de la fonction de taux des grandes déviations du processus empirique à valeurs mesures

$$\bar{X}^N(t)=\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i(t)}$$

Classification A.M.S. : Primary 60 F 10, 60 J 60; secondary 60 G 44, 60 H 05, 60 J 57.

où les X_i sont des copies indépendantes d'un processus de diffusion dans $\mathbb{R}^d$. Ce programme est réalisé en minimisant l'entropie relative (information de Kullback) d'une probabilité Q par rapport à la loi P commune aux X_i, sous la contrainte des lois marginales temporelles de Q fixées. La finitude de la fonction de taux est mise en relation avec l'existence de diffusions conservatives pour des matrices de diffusion générales. Ces processus de diffusion sont construits dans une grande généralité.

0. INTRODUCTION

Let $\Omega = \mathbb{C}([0, T], \mathbb{R}^d)$ be the space of all continuous paths with values in $\mathbb{R}^d$ endowed with the uniform norm, and P be the law of a $\mathbb{R}^d$-valued diffusion process $X = (X(t))_{0 \leqq t \leqq T}$, considered as a random variable on Ω. We denote by $M(\Omega)$ and $M_1(\Omega)$ the sets of all measures and all Probability measures on Ω equipped with its Borel σ-algebra. Let us consider a sequence $(X_i)_{i \geqq 1}$ of independent copies of X, P^∞ the law of this sequence on the infinite tensor product Ω^∞ and the random empirical measures on $M_1(\Omega)$

$$\hat{X}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i}, \qquad N \geqq 1$$

where δ_z is the Dirac measure at point z. Sanov's theorem states that the laws of $\hat{X}^N (N \geqq 1)$ under P^∞ obey a large deviation principle in the space $M(\Omega)$ endowed with its usual weak topology, the rate function of which is the Kullback information $I(., P)$ defined for any $Q \in M(\Omega)$ by

$$I(Q, P) = \begin{cases} H(Q, P) = \int Z \log(Z)\, dP \text{ if } Q \in M_1(\Omega),\ Q \ll P, \\ \qquad \dfrac{dQ}{dP} = Z \text{ and } Z \log(Z) \in L^1(P), \\ +\infty \quad \text{otherwise.} \end{cases}$$

More precisely, if $\bar{A}$ and A^0 are the closure and the interior of the set A, for any Borel subset A of $M(\Omega)$

$$\begin{aligned} -\inf_{Q \in A^0} I(Q, P) &\leqq \liminf_{N \to \infty} \frac{1}{N} \log P^\infty(\hat{X}^N \in A) \\ &\leqq \limsup_{N \to \infty} \frac{1}{N} \log P^\infty(\hat{X}^N \in A) \leqq -\inf_{Q \in \bar{A}} I(Q, P). \end{aligned}$$

(*see* [DS] or [Léo]).

Now, let us consider the sequence of $M_1(\mathbb{R}^d)$-valued continuous processes

$$\bar{X}^N:\quad t\in[0, T]\mapsto \bar{X}^N(t)=\frac{1}{N}\sum_{i=1}^N \delta_{X_i(t)}\in M_1(\mathbb{R}^d),\qquad N\geqq 1.$$

Since the mapping

$$Q\in M_1(\Omega)\mapsto(t\mapsto Q_t)\in \mathbb{C}([0, T], M_1(\mathbb{R}^d))$$

where $Q_t=Q\circ X_t^{-1}$ is the Q-law of $X(t)$, is continuous when $\mathbb{C}([0, T], M_1(\mathbb{R}^d))$ is endowed with the topology of the weak convergence of all the t-marginal laws uniformly on $[0, T]$, the contraction principle provides us with the following large deviation principle. For any Borel subset B of $\mathbb{C}([0, T], M_1(\mathbb{R}^d))$, we have

$$\begin{aligned}-\inf_{\nu\in B^0} I(\nu)&\leqq\liminf_{N\to\infty}\frac{1}{N}\log P^\infty(\bar{X}^N\in B)\\ &\leqq\limsup_{N\to\infty}\frac{1}{N}\log P^\infty(\bar{X}^N\in B)\leqq-\inf_{\nu\in\bar{B}} I(\nu),\end{aligned}$$

where for any ν in $\mathbb{C}([0, T], M_1(\mathbb{R}^d))$

$$I(\nu)=\inf\left\{I(Q, P);\ Q\in M_1(\Omega),\ Q\circ X_t^{-1}=\nu_t,\ \text{for } 0\leqq t\leqq T\right\}.$$

In [DaG], D. A. Dawson and J. Gärtner have computed an explicit expression for $I(\nu)$ in the case where P is the law of a nonhomogeneous diffusion process on $\mathbb{R}^d$ with generator

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

assuming that $a=(a_{ij})$ and $b=(b_i)$ are locally Hölder continuous and that the symmetric matrix a is strictly positive definite. This expression is the following (at least formally):

$$\begin{aligned}I(\nu)=I(\nu_0, \mu_0)+\int_0^T \sup_{f\in\mathscr{D}}\,[\langle\dot{\nu}_t-A^*(t)\nu_t, f\rangle&\\ -\frac{1}{2}\int(\nabla f(x)/a(x, t)\nabla f(x))\,\nu_t(dx)]\,dt,&\end{aligned}$$

where $\mu_0=P\circ X_0^{-1}$ is the initial law of the process, $\mathscr{D}$ is the set of Schwartz test functions on $\mathbb{R}^d$, ∇ denotes the gradient operator, $A^*(t)$ is the formal adjoint of $A(t)$, $\dot{\nu}$ is the time derivative of ν in the sense of Schwartz distributions the first bracketting $\langle\,.\,,\,.\,\rangle$ being the duality $(\mathscr{D}, \mathscr{D}')$ and the second one $(./.)$ being the usual scalar product in $\mathbb{R}^d$.

In the special case where P is the Wiener measure starting from μ_0, H. Föllmer ([Föl]) proposed an alternate proof of the above result, under the assumption that the marginal flow $\nu=(\nu_t)_{0\leqq t\leqq T}$ is admissible in the sense that there exists $Q \in M_1(\Omega)$ such that $Q_t = \nu_t$ for $0 \leqq t \leqq T$ and $H(Q, P) < \infty$. In addition, he proved that for an admissible flow the infimum in the definition of $I(\nu)$ is attained for a Markov law Q^*, the drift (with respect to P) of which being characterized in terms of the flow ν. His constructive proof relies on an extension of a result of I. Csiszár [Cs] on the minimization of the Kullback information under linear constraints. By means of the same method, M. Brunaud [Bru] obtained similar results for a uniformly elliptic diffusion with smooth coefficients (also *see* [Mi]). The main reason for these assumptions is the employment of a time reversal argument which yields the approximation of Q^* by piecewise h-processes (or Schrödinger bridges). It should be remarked that the notion of admissible flow is strongly related to the "conservative diffusions" of E. Carlen [Ca], which appear in the stochastic interpretation of quantum mechanics.

Our aim is to recover the results of [DaG], [Föl], [Bru] and to extend them to a large class of Markov processes. In the present paper, we consider the case of diffusion processes in $\mathbb{R}^d$ with generator $A(t)$, where a and b are supposed to be locally bounded measurable, and we allow a to be degenerate. Since we want extend the methods and results to a larger class of processes, we decided not to use the usual time reversal of diffusion processes. We are here mainly interested in application to large deviation theory (*see* [Léo]), so, some of the natural developments in the direction of stochastic mechanics are not treated. They are more systematically studied in [CP] (also *see* [Pe]).

Let us now present briefly the organization and contents of the paper.

Section 1 is devoted to the introduction of the main notation.

In Section 2 we give some elementary properties of the relative entropy $H(Q, P)$ we will use throughout the paper. In particular we indicate in (2.6) a simple but efficient trick to build diffusion processes with given marginals under an entropy condition.

The main result of Section 3 is Theorem 3.1, which tells that one can associate to any Q, such that $H(Q, P)$ is finite, a Markov Probability measure $\bar{Q}$ with the same marginals and such that $H(\bar{Q}, P) \leqq H(Q, P)$. Therefore to solve the minimization problem, it is sufficient to restrict our attention to Markov Probability measures. Actually Theorem 3.1 is more precise. Indeed we prove that the "natural Markovian version" of Q (*see* § 3 (3.4) for the definition) is a Probability measure with the same marginals and smaller relative entropy. The proof of Theorem 3.1 is somewhat intricate because we have to use several approximation procedures. An alternate proof of this result can be obtained by the methods of Section 4, but we think that the one presented here (which is only based on stochastic calculus arguments) can be useful for other purpose.

In Section 4 we give some conditions for a given flow to be admissible. Let ν be a flow of Probability measures on $\mathbb{R}^d$. A Borel function B is said to be of ν-finite energy if

$$\int_0^T \int (B(x, t)/a(x, t) B(x, t)) \nu_t(dx)\, dt < \infty.$$

The results of the section are then the following:

Assume that $H(\nu_0, \mu_0) < \infty$ and that ν satisfies the weak forward equation for $A(t) + a B . \nabla$, for some B of ν-finite energy. If one of the following conditions is satisfied

* (H 1) $a = \sigma . \sigma^*$, σ and b are $C^{2, 1, \alpha}$ vector fields,
* (H 2) a is strictly positive definite, a and b are locally Hölder continuous,
* (H 3) $d\nu_t/d\mu_t$ is locally bounded,

then ν is admissible.

This result extends previous results of Carlen [Ca], Zheng [Zhe] and other people. As was said before, we refer to [CP] for a more complete discussion.

Finally in Section 5 we describe the set of all Markov Probability measures associated to an admissible flow (Theorem 5.3), and derive in Theorems 5.9 and 5.20 the explicit formula for $I(\nu)$, extending [DaG], [Föl] and [Bru].

We want to underline that a large part of the ideas we are using in this article, are already contained, at least in implicit form, in the papers of [DaG], [Föl], [Mi], [Zhe], [AN] or [DS1], in particular cases of sometimes with incomplete proofs (*see* the comments at the end of Section 5). We think that one originaglity of the present paper is that it clearly establishes a link between some of these papers, in particular it gives a purely large deviations interpretation of the Nelson's processes which is very satisfactory from the physical point of view (for a direct approach of the construction of Nelson processes by large deviations methods, *see* [CL2]). Another very interesting large deviations approach of these processes is given in [RZ].

As the reader will see, the proofs we have tried to give are often independent of the specific form of the generator. This allows to think that the contents of this paper are still available in a more general (Markov) context. This will be the aim of a future work ([CL1]).

ACKNOWLEDGEMENTS

We wish to thank F. Comets for pointing out to us Lemma 3.16 and for his kind interest in this work. We also thank G. Ben Arous, M. Brunaud and H. Föllmer for very useful discussions. We finally thank an anonymous referee for pointing out an error in the first version of this paper and for his very interesting comments.

1. PRELIMINARIES

We introduce here some definitions and notations that will hold throughout the paper.

Let $(\Omega, \mathscr{F})$ be a measurable space. We denote by $M(\Omega)$ (resp. $M_1(\Omega)$) the set of all positive measures (resp. Probability measures) on Ω.

For Q and P elements of $M_1(\Omega)$, $H(Q, P)$ denotes the relative entropy of Q with respect to P, defined by:

$$(1.1)\qquad H(Q, P)=\begin{cases} \displaystyle\int Z\log(Z)\,dP & \text{if } Q \ll P, \quad \dfrac{dQ}{dP}=Z \\ & \text{and} \quad Z\log(Z)\in L^1(P), \\ +\infty & \text{otherwise.}\end{cases}$$

Notice that $H(Q, P)\geqq 0$, the equality implying $Q=P$ (Jensen's inequality). The following result is well known:

(1.2) If Q and P belong to $M_1(\Omega)$ then

$$H(Q, P)=\sup\left\{\int \phi\,dQ-\log\left(\int \exp(\phi)\,dP\right),\ \phi\in\mathscr{B}_b(\Omega)\right\}$$

where $\mathscr{B}_b(\Omega)$ is the set of real valued bounded measurable functions, defined on Ω.

Furthermore, if Ω is a Polish space equipped with its Borel σ-field $\mathscr{F}$, one can replace, in (1.2), $\mathscr{B}_b(\Omega)$ by $C_b(\Omega)$, the space of bounded continuous functions.

In the paper, Ω will be $\mathbb{C}([0, T], \mathbb{R}^d)$, the space of all continuous paths with values in $\mathbb{R}^d$, T being a positive (but finite) real number; $(X_t)_{t\in[0,T]}$ will be the canonical process, $\mathscr{B}$ the Borel σ-field of $\mathbb{R}^d$.

Let $P\in M_1(\Omega)$ be given.

(1.3) DEFINITION. – *Let $(\nu_t)_{t\in[0,T]}$ be a flow of Probability measures on $\mathbb{R}^d$. We denote by A_ν the set $\{Q\in M_1(\Omega)$ s. t. $Q\circ X_t^{-1}=\nu_t$ for all $t\in[0, T]\}$, and by $A_{\nu,H}$ the set $\{Q\in A_\nu$ s. t. $H(Q, P)<+\infty\}$.*

The flow will be called admissible if $A_{\nu,H}$ is non void.

We adopt classical notation for the usual spaces of Functional Analysis. In particular the subscript $_0$ will denote functions with compact support, C_b^k is the space of k-times differentiable functions which are bounded with bounded derivatives up to an including order k and we omit the target space when it is $\mathbb{R}$.

Our aim is to study $A_{\nu,H}$ for a given flow ν, when P is the law of a diffusion process with initial law $\mu_0\,(\in M_1(\mathbb{R}^d))$ and generator $A(t)$ given as:

$$(1.4)\qquad A(t)=\frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x, t)\frac{\partial^2}{\partial x_i\,\partial x_j}+\sum_{i=1}^{d} b_i(x, t)\frac{\partial}{\partial x_i}.$$

Here we assume that

(1.5.i) the coefficients a_{ij} and b_i belong to $\mathscr{B}_{b,\,\mathrm{loc}}(\mathbb{R}^d \times \mathbb{R})$, the set of locally bounded Borel functions,

(1.5.ii) the matrix $a=(a_{ij})$ is non negative symmetric, *i.e.* for all $x \in \mathbb{R}^d$, $t \in \mathbb{R}$, and $\xi=(\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$, $\sum_{i,j=1}^{d} a_{ij}(x, t)\,\xi_i \xi_j \geqq 0$.

More precisely we make the following assumptions:

Let $\Omega_t = \mathbb{C}([0, T], \mathbb{R}^d \times \mathbb{R})$ be the time-space, with additionnal time coordinate denoted by u_t.

The canonical space $(\Omega_t, \mathscr{F}, X_., u_.)$ is equipped with a strong Markov family $(P_{x,u})_{(x,u) \in \mathbb{R}^d \times R}$ of Probability measures, such that:

$$
(1.6)\quad \begin{cases}
\text{(i) } P_{x,u}(X_0 = x,\ u_0 = u) = 1, \\
\text{(ii) } M_t^i = X_t^i - X_0^i - \displaystyle\int_0^t b_i(X_s, u_s)\,ds, \text{ is for all } i \text{ a continuous} \\
\qquad P_{x,u}\text{-local martingale, and } \langle M^i, M^j \rangle_t = \displaystyle\int_0^t a_{ij}(X_s, u_s)\,ds, \\
\text{(iii) } u_t = u + t,\ P_{x,u}\,\text{p.s.}, \\
\qquad \text{Furthermore, } P_{x,u} \text{ is an extremal solution to the martingale} \\
\qquad \text{problem} \\
\text{(i), (ii) and (iii) for all } (x, u) \in \mathbb{R}^d \times \mathbb{R}.
\end{cases}
$$

In particular no explosion occurs up to and including time T.

The (formal) generator of $(P_{x,u})$ is then $A' = \dfrac{\partial}{\partial u} + A(u)$.

If strong uniqueness occurs, hypotheses (1.6) are satisfied.

Also notice that when A does not depend upon the time coordinate, this coordinate plays no rôle nor in uniqueness or extremality.

According to [Jac, Thm 13.55], the last requirement in (1.6) is equivalent to:

(1.7) $P_{x,u}$ is an extremal solution to the martingale problem $M(A', C_b^{2,1}(\mathbb{R}^d \times \mathbb{R}), \delta_{x,u})$ (*see* notation in [Jac, ch. 13]).

Now for $\mu_0 \in M_1(\mathbb{R}^d)$ we define

(1.8) $P' = P_{\mu_0 \otimes \delta_0}$ (δ_0 is the Dirac measure at 0).

In addition we assume that P′ is an extremal solution of (1.6) with initial condition $\mu_0 \otimes \delta_0$. Notice that $u_s = s$ for all s, P′ a.s.

Before going further, we have to make some remarks about the time-space process.

Let π be the projection operator of $\mathbb{R}^d \times \mathbb{R}$ onto $\mathbb{R}^d$. To any random variable γ defined on Ω corresponds $\gamma' = \gamma \circ \pi$ defined on Ω_t. To P′ corresponds its projection on Ω denoted by P. We denote by P_x the projection of $P_{x,0}$.

To any $Q \ll P$ corresponds Q′ defined on Ω_t by $Q' = \left(\frac{dQ}{dP} \circ \pi\right) P'$.

Of course relative entropy and spatial marginals are preserved.

We shall now recall some basic facts related to the Girsanov transform theory and to the Föllmer measure.

If we put

$$s_n = \inf\{t \geqq 0, \|X_t\| \geqq n\} \wedge T \tag{1.9}$$

then s_n is a localizing sequence of stopping times for M_t.

Let $Q \in M_1(\Omega)$ be absolutely continuous with respect to P, and $Q \circ X_0^{-1} = \nu_0$. Thanks to the extremality assumption in (1.6), it follows from theorems 12.17, 12.34 and 12.48 of [Jac] that there exists a $\mathbb{R}^d$-valued previsible process β_s, such that if we put

$$T_n = \inf\left\{t \geqq 0, \int_0^t (\beta_s / a(X_s, s)\, \beta_s)\, ds \geqq n\right\} \wedge T, \qquad n \in \mathbb{N} \cup \{+\infty\}, \tag{1.10}$$

where $(./.)$ denotes the euclidian scalar product in $\mathbb{R}^d$, the density process $Z_.$ of Q with respect to P is given by the exponential formula:

$$\left\{\begin{aligned} & Z_t = \frac{d\nu_0}{d\mu_0} \exp\left(\int_0^t (\beta_s / dM_s) - \frac{1}{2}\int_0^t (\beta_s / a(X_s, s)\, \beta_s)\, ds\right) \\ & \qquad \text{on } \bigcup_n [0, T_n] \\ & Z_t = \liminf Z_{T_n} \quad \text{on } \bigcap_n\,]T_n, T]. \end{aligned}\right. \tag{1.11}$$

Furthermore one can choose a continuous version of $Z_.$ such that $Z_t = 0$ if $t > T_n$ for all n. We shall sometimes write $Z(\beta, \nu_0, P)$ if confusions are possible.

Now

(1.12) $T_n \to T_\infty = T - Q$ a.s., $N_t = M_t - \int_0^t a(X_s, s)\, \beta_s\, ds$ is a continuous d-dimensional Q-local martingale admitting $s_n \wedge T_n$ as a localizing sequence of stopping times and $\langle N^i, N^j \rangle_t = \langle M^i, M^j \rangle_t$.

Conversely, if β_s is a previsible process, one can define T_n and $Z_t(\beta, \nu_0, P)$ as in (1.10) and (1.11). Then $Z_.$ is a nonnegative P-local martingale and we can choose a version which is a supermartingale. Once again we can choose a continuous version of Z and $Z_t = 0$ if $t > T_n$ for all n (for all these results see [Jac chap. 8 and chap. 12.3 §c]).

We can associate to this supermartingale its Föllmer measure ([Fö]). To this end, we must consider the space Ω_ξ of explosive trajectories with explosion time ξ which belongs to $[0, T] \cup \{+\infty\}$.

(1.13) NOTATION. – The Föllmer measure Q defined on Ω_ξ associated to the P-exponential supermartingale $Z_.(\beta, \nu_0, P)$ will be called the (β, ν_0, P)-FM. If $\beta_s = B(X_s, s)$ for a Borel function B, we write B instead of β.

More generally the Föllmer measure associated to a nonnegative P-supermartingale Z will be called the (Z, P)-FM.

Recall that the (β, ν_0, P)-FM Q satisfies the following equality

(1.14) For all stopping time τ, $\tau \leqq T$, and all $\mathscr{F}_\tau$ measurable F

$$E^Q[F\, 1_{\tau<\xi}] = E^P[Z_\tau(\beta, \nu_0, P)\, F].$$

Actually, one deduces from what precedes that

(1.15) $E^Q[F\, 1_{\tau<\xi}] = E^P[Z_\tau(\beta, \nu_0, P)\, F] = E^P[Z_{\tau\wedge T_\infty}(\beta, \nu_0, P)\, F].$

Finally, if $\beta_s = B(X_s, s)$ for a Borel function B, Z is a multiplicative functional. It follows that the (B, ν_0, P)-FM is a Markov Probability measure on Ω, as soon as $Q(\xi < +\infty) = 0$. By Theorem 24.36 in [Sha], it is in fact strong Markov.

2. SOME PROPERTIES OF THE RELATIVE ENTROPY

In this section we shall discuss some elementary properties related to the finite entropy condition.

(2.1) PROPOSITION. – *Assume that* Q *is a Probability measure with* $Q \ll P$. *Let* β_s *denotes its drift. Then*

$$H(Q, P) = H(\nu_0, \mu_0) + \frac{1}{2} E^Q\left[\int_0^T (\beta_s / a(X_s, s)\, \beta_s)\, ds\right].$$

Proof (cf. [Fö 2]). – We may assume that $H(\nu_0, \mu_0)$ is finite. Let Q_n be defined by $Q_n = Z_{T_n} P$. By Novikov criterion one knows that Q_n is a Probability measure. Furthermore $H(Q_n, P)$ is finite, $Q \ll Q_n$ and these two measures are the same in restriction to $\mathscr{F}_{T_n}$. Thus

$$\begin{aligned} H(Q_n, P) &= H(\nu_0, \mu_0) + E^{Q_n}\left[\int_0^{T_n} (\beta_s / dM_s) - \frac{1}{2}\int_0^{T_n} (\beta_s / a(X_s, s)\, \beta_s)\, ds\right] \\ &= H(\nu_0, \mu_0) + E^Q\left[\int_0^{T_n} (\beta_s / dN_s) - \frac{1}{2}\int_0^{T_n} (\beta_s / a(X_s, s)\, \beta_s)\, ds\right] \end{aligned}$$

But, according to (1.12), $\int_0^{t \wedge T_n} (\beta_s / dN_s)$ is a Q-martingale with zero expectation. It follows:

$$H(Q_n, P) = H(\nu_0, \mu_0) + \frac{1}{2} E^Q\left[\int_0^{T_n} (\beta_s / a(X_s, s)\, \beta_s)\, ds\right] = E^P[Z_{T_n} \log (Z_{T_n})].$$

Notice that the lim inf of Z_{T_n} is P almost surely equal to Z_T because $Z_s = 0$ if $s > T_n$ for all n, and recall that $T_n \to T$, Q a. s.

Since the function $x \log (x)$ is bounded from below, one can apply Fatou's lemma and obtain

$$H(Q, P) \leqq H(\nu_0, \mu_0) + \frac{1}{2} E^Q\left[\int_0^{T} (\beta_s / a(X_s, s)\, \beta_s)\, ds\right]$$

The conserve inequality is obtained by taking the supremum over n, in the right hand side of the following chain of relations:

$$(2.2) \quad H(Q, P) = H(Q, Q_n) + E^Q[\log (Z_{T_n})] \geqq E^Q[\log (Z_{T_n})] = H(Q_n, P).$$

Indeed we may assume that $H(Q, P)$ is finite. Since $\log (dQ_n/dP)$ belongs to $L^1(Q_n)$ it also belongs to $L^1(Q)$ and (2.2) makes sense. □

The next Proposition completes the preceding one.

(2.3) Proposition. – *Let* Q *be the* (β, ν_0, P)-FM (*see* (1.13)) *where* β *is a previsible process and* ν_0 *a Probability measure such that* $H(\nu_0, \mu_0) < +\infty$.

If $E^Q\left[\int_0^{T_\infty} (\beta_s / a(X_s, s)\, \beta_s)\, ds\right] < +\infty$, *then* $Q(\xi = \infty) = 1$. *So* Q *is a Probability measure on* Ω *with* $H(Q, P) < +\infty$, $T_\infty = T - Q$ a. s. *and* (2.1) *holds.*

Proof. – Applying (1.14) with $\tau = T_n$ and $F = 1$, we get $T_n < \xi - Q$ a. s. Thus $T_\infty \leqq \xi - Q$ a. s. Furthermore Q and $Q_n = Z_{T_n} P$ are the same in restriction to $\mathscr{F}_{T_n}$. Hence

$$\begin{aligned} H(Q_n, P) &= H(\nu_0, \mu_0) + \frac{1}{2} E^Q\left[\int_0^{T_n} (\beta_s / a(X_s, s)\, \beta_s)\, ds\right] \\ &= E^P[Z_{T_n} \log (Z_{T_n})] \\ &\leqq H(\nu_0, \mu_0) + \frac{1}{2} E^Q\left[\int_0^{T_\infty} (\beta_s / a(X_s, s)\, \beta_s)\, ds\right]. \end{aligned}$$

This implies that Z_{T_n} is a bounded sequence in the Orlicz space $L_{\tau^*}(P)$ with $\tau^*(x) = (x+1) \log (x+1) - x$. In particular Z_{T_n} is uniformly integrable. Since this sequence goes P a. s. to Z_{T_∞} we get, according to (1.15),

$$Q(\xi = \infty) = E^P[Z_{T_\infty}] = \lim E^P[Z_{T_n}] = 1. \quad \square$$

In view of what precedes it is natural to put the following definition:

(2.4) DEFINITION. – (i) *Let* Q *be a Probability measure on* Ω. *We denote by* $\mathscr{L}^2_Q$ *the following space:*

$$\mathscr{L}^2_Q = \left\{ \gamma_.;\ \mathbb{R}^d\text{-valued measurable process s. t.}\quad E^Q\left[\int_0^T (\gamma_s / a(X_s, s)\,\gamma_s)\,ds\right] < +\infty \right\}.$$

(ii) *Let* $(\nu_t)_{t\in[0,\,T]}$ *be a flow of Probability measures on* $\mathbb{R}^d$. *We denote by* $\mathscr{L}^2_\nu$ *the following space:*

$$\mathscr{L}^2_\nu = \left\{ \varphi;\ \mathscr{B}(\mathbb{R}^d\times[0, T])\ \text{measurable, with values in}\ \mathbb{R}^d\ \text{s. t.}\quad \int_0^T \int_{\mathbb{R}^d} (\varphi(x, s)/a(x, s)\,\varphi(x, s))\,\nu_s(dx)\,ds < +\infty \right\}$$

The quadratic forms introduced in (2.4) define hilbertian seminorms on the spaces $\mathscr{L}^2$.

(2.5) We shall denote by L^2_Q and L^2_ν the factor spaces ($\mathscr{L}^2$/Ker form) which are Hilbert spaces. If γ (resp. φ) belongs to L^2_Q (resp. L^2_ν) we shall say that γ (resp. φ) is of finite Q-energy (resp. ν-energy).

Putting together (2.1) and (2.3), we get the following trick, which we shall use several times in the sequel.

(2.6) *Trick.* – Let B be a Borel function defined on $\mathbb{R}^d\times\mathbb{R}$, and $\nu_t\,(0\leqq t\leqq T)$ be a flow of Probability measures on $\mathbb{R}^d$ such that $H(\nu_0, \mu_0) < +\infty$. Define $Z_.(B, \nu_0, P)$ as in (1.11) with $\beta_s = B(X_s, s)$, and denote by Q the (B, ν_0, P)-FM. Assume that

(i) *Energy condition.* – $\displaystyle\int_0^T \int_{\mathbb{R}^d} (B(x, s)/a(x, s)\,B(x, s))\,\nu_s(dx)\,ds < +\infty$,

i.e. B is of finite ν-energy,

(ii) *Domination.* – For all nonnegative f in $\mathscr{B}_b(\mathbb{R}^d\times\mathbb{R})$ and all t in $[0, T]$,

$$E^Q[f(X_t)\,1_{t\leqq T_\infty}] \leqq \int_{\mathbb{R}^d} f(x)\,\nu_t(dx).$$

Then by the monotone convergence theorem ii) extends to any nonnegative measurable f, so that

$$E^Q\left[\int_0^{T_\infty} (B(x, s)/a(x, s)\,B(x, s))\,ds\right] \leqq \int_0^T \int_{\mathbb{R}^d} (B(x, s)/a(x, s)\,B(x, s))\,\nu_s(dx)\,ds < +\infty.$$

Thanks to (2.3), Q is then a Probability measure on Ω, $H(Q, P) < +\infty$ and $T_\infty = T$ Q a.s. But now (ii) implies that $Q \circ X_t^{-1} = \nu_t$ for all t.

In other words, the finite energy condition together with the domination relation impose that Q has the prescribed marginals ν_t. This idea appears in a slightly different form in Zheng's paper [Zhe].

We complete this section with a useful remark on the behaviour of martingales in connection with an entropy condition.

(2.7) PROPOSITION. – *Let* Q *and* Q* *be two Probability measures on* Ω *such that* $H(Q, Q^*) < +\infty$. *Let* S_t *be a bounded* Q*-*martingale. Then* S_t *is a* Q-*semi martingale with decomposition* $S_t = K_t + V_t$, *where* K_t *is a* L^2 Q-*martingale with* $\langle K \rangle_t = \langle S \rangle_t$ *and* $E^Q[\langle K \rangle_T] < +\infty$.

Proof. – Everything is well known except (perhaps) that $E^Q[\langle S \rangle_T] < +\infty$ which implies that K_t is a (true) Q-martingale. Let C be a bound for $S_.$. Applying the Burkholder-Davis-Gundy inequalities we get

$$E^{Q^*}[\langle S \rangle_T^p] \leqq (4p)^p C^{2p}, \qquad \text{for} \quad 1 < p < +\infty.$$

It follows from Stirling formula that if $C < (4e)^{-1/2}$ then

$$E^{Q^*}[\exp \langle S \rangle_T] < +\infty.$$

Hence $\langle S \rangle_T$ belongs to the Orlicz space $L_\tau(Q^*)$ with $\tau(x) = e^x - x - 1$. But $\dfrac{dQ}{dQ^*}$ belongs to $L_{\tau^*}(Q^*)$ with $\tau^*(x) = (x+1)\log(x+1) - x$, since $H(Q, Q^*) < +\infty$. Hölder's inequality in Orlicz spaces leads to the desired result, since τ and τ^* are Legendre conjugate. In the general case it suffices to divide $S_.$ by a large enough constant. $\square$

(2.8) *Remarks.* – 1) The above proof shows that there exists a universal constant K such that for all Q, Q* and S as in (2.7)

$$E^Q[\langle S \rangle_T] \leqq K\,(H(Q, Q^*) + 1)\,C^2.$$

2) If $\dfrac{dQ}{dQ^*} = Z_T$ belongs to some $L^q(Q^*)$ with $q > 1$, then $\langle S \rangle_T$ belongs to all the $L^p(Q)$ and $E^Q[\langle S \rangle_T^p] \leqq \| Z_T \|_q (4pq^*)^p C^{2p}$ with $(1/q) + (1/q^*) = 1$.

3. ENTROPY AND THE MARKOV PROPERTY

Let us first discuss the following easy exercise.

Let $\lambda = dx \otimes dy$ be the Lebesgue measure on the unit square $S = [0, 1] \times [0, 1]$. Consider two Probability measures $\nu_0(dx)$ and $\nu_1(dy)$ on $[0, 1]$ such that $H(\nu_0, dx)$ and $H(\nu_1, dy)$ are finite. Then among all Probability measures ν on S such that $H(\nu, \lambda)$ is finite and $(\nu \circ x^{-1} = \nu_0$,

$\nu \circ y^{-1} = \nu_1$), the one which minimizes the relative entropy with respect to λ is $\nu_0 \otimes \nu_1$.

If we replace $(S, \lambda, (\nu_0, \nu_1))$ by $(\Omega, P, (\nu_t)_{t \in [0, T]})$ for a given admissible flow $\nu_.$, the natural analogous statement would be that the Probability measure Q in $A_{\nu, H}$ which minimizes the relative entropy with respect to P, has the Markov property.

This statement is actually true and follows from the following more precise result.

(3.1) THEOREM. – *Let* Q *be any Probability measure in* $A_{\nu, H}$. *Then there exists a (strong) Markov Probability measure* $\bar{Q}$ *such that* $\bar{Q} \in A_{\nu, H}$ *and* $H(\bar{Q}, P) \leqq H(Q, P)$. (*Here, strong stands for the time-space version of* $\bar{Q}$, see § 1).

Actually we shall prove that if P is a (non necessarily extremal) solution to the martingale problem (1.6) and if Q is given by the density $Z_T(\beta, \nu_0, P)$ of (1.11), a natural Markovian version $\bar{Q}$ [*see* (3.4) below] of Q satisfy (3.1). This result is more precise that the one in [Föl] or [AN] which only says that the minimizing Probability measure has the Markov property. Before proving (3.1), we shall first explain what we mean by a "natural Markovian version of Q".

Since H(Q, P) is finite, $\beta \in L^2_Q$. Thus, we can apply Riesz representation theorem to obtain:

(3.2) there exists a Borel function B such that for all $\varphi \in L^2_\nu$,

$$\int_0^T \int_\Omega (\varphi(X_s, s)/a(X_s, s)\,\beta_s)\,dQ\,ds = \int_0^T \int_{\mathbb{R}^d} (\varphi(x, s)/a(x, s)\,B(x, s))\,\nu_s(dx)\,ds.$$

Furthermore B is unique in L^2_ν.

Of course, this is nothing else than a particular form of multidimensional conditional expectation. Clearly we get

$$(3.3)\quad \int_0^T \int_{\mathbb{R}^d} (B(x, s)/a(x, s)\,B(x, s))\,\nu_s(dx)\,ds \leqq \int_0^T \int_\Omega (\beta_s/a(X_s, s)\,\beta_s)\,dQ\,ds < +\infty.$$

We choose once for all a version of B. Furthermore we extend B to the whole space $\mathbb{R}^d \times \mathbb{R}$ by defining

$$B(x, u) = B(x, T) \text{ fo } u \geqq T \qquad \text{and} \qquad B(x, u) = B(x, 0) \text{ for } u \leqq 0.$$

Define Q′ as in Section 1, a "natural Markovian version of Q′" is then $\bar{Q}'$ defined as

(3.4) $\bar{Q}'$ is defined on $\Omega_{t, \xi}$ as the (B, ν_0, P')-FM.

Notice that $\bar{Q}'$ may depend on the choice of B.

According to the final remark of Section 1, $\bar{Q}'$ is then strong Markov.

We shall now discuss some questions related to this $\bar{Q}'$.

Define T_n, S_n and τ_n as follows

(3.5.i) $T_n = \inf\left\{t \geqq 0, \int_0^t (\beta_s/a(X_s, u_s)\beta_s))\,ds \geqq n\right\} \wedge T$,

(3.5.ii) $S_n = \inf\left\{t \geqq 0, \int_0^t (B(X_s, u_s)/a(X_s, u_s)B(X_s, u_s))\,ds \geqq n\right\} \wedge T$,

(3.5.iii) $\tau_n = S_n \wedge T_n$.

Once again $\tau_n \to T$, Q' a.s. [by (3.3)]. Let us define the process G_t by:

$$
(3.6)\quad \begin{cases} G_t = \exp\left\{\int_0^t (B(X_s, u_s) - \beta_s/dN_s) \right.\\ \left. -\frac{1}{2}\int_0^t (B(X_s, u_s) - \beta_s/a(X_s, u_s)(B(X_s, u_s) - \beta_s))\,ds\right\} \quad \text{on } [0, \tau_n], \\ G_t = \liminf G_{\tau_n}, \qquad \text{if} \quad t > \tau_n \text{ for all } n, \end{cases}
$$

where N_t is defined in (1.12). Of course $u_s = s$ for all s, Q' a.s. $Q'(\tau_\infty = T) = 1$, thus G_t is Q' a.s. defined by the exponential formula and we can choose a continuous version of G_t which is a nonnegative Q-supermartingale.

(3.7) $\bar{Q}''$ is then defined on $\Omega_{t,\xi}$ as the (G, Q')-FM.

$\bar{Q}''$ does not depend of our choice of B. According to (3.3), $Q'(G_T = 0) = 0$. Therefore, $Q' \ll \bar{Q}''$. Now define $\bar{Z}_t$ as:

$$
(3.8)\quad \begin{cases} \bar{Z}_t = \dfrac{d(\nu_0 \otimes \delta_0)}{d(\mu_0 \otimes \delta_0)} \exp\left\{\int_0^t (B(X_s, u_s)/dM_s) \right.\\ \left. -\frac{1}{2}\int_0^t (B(X_s, u_s)/a(X_s, u_s)B(X_s, u_s))\,ds\right\} \quad \text{on } [0, \tau_n], \\ \bar{Z}_t = Z_t \liminf G_{\tau_n}, \qquad \text{if} \quad t > \tau_n \text{ for all } n. \end{cases}
$$

Since $E^{P'}[Z_t \liminf G_{\tau_n} 1_{t > \sup \tau_n}] = E^{Q'}[\liminf G_{\tau_n} 1_{t > \sup \tau_n}] = 0$, it holds

(3.9) $\qquad \bar{Z}_t = 0, \qquad \text{if} \quad t > \sup \tau_n, \quad P'$ a.s.

$\bar{Z}$ is a P'-supermartingale and $\bar{Q}''$ is actually the $(\bar{Z}, P')$-FM. Notice that $\bar{Z}$ is not necessarily continuous (there is a jump to 0 at time τ_∞), and non-multiplicative. In particular $\bar{Q}''(\xi = +\infty) \leqq E^{P'}[\bar{Z}_{\tau_\infty}]$, and not equal in general. $\bar{Z}_t$ is smaller or equal to the supermartingale $Z(B, \nu_0, P)$. Indeed they may differ on the set $\{\sup \tau_n < t < \sup S_n\}$ where $\bar{Z}_t = 0$ while $Z(B, \nu_0, P) \geqq 0$. As a consequence it is thus clear that $Q' \ll \bar{Q}'$ and that

(3.10) if $\tau_\infty = T - P'$ a.s. then $\bar{Q}' = \bar{Q}''$.

Remark that $\tau_\infty = T - P'$ a.s. if and only if Q′ and P′ are equivalent (assuming here that P′ is an extremal solution, *see* [Jac]). If $\bar{Q}'(\xi = +\infty) = 1$ then $\bar{Q}' \sim Q'$. But also, if $\bar{Q}''(\xi < +\infty) = 0$, *i.e.* $\bar{Q}''$ is a Probability measure on Ω_t, then $\bar{Z}_t$ is a martingale and so is $Z(B, \nu_0, P)$. This implies easily that $Z(B, \nu_0, P)$ and $\bar{Z}$ coïncide. Hence $\bar{Q}' = \bar{Q}''$ and $\bar{Q}' \sim Q'$.

In general one only has $Q' \ll \bar{Q}'$.

$\bar{Q}$ will denote the projection $\bar{Q}' \circ \pi$ (*see* § 1).

We shall show that Q and $\bar{Q}$ (thus Q′ and $\bar{Q}'$) have same marginals. $\bar{Q}'$ will thus be a Probability measure on Ω_t and the inequality $H(\bar{Q}, P) \leqq H(Q, P)$ will then follow from (3.3) and proposition 2.3.

Proof of (3.1). – *Step* 1. – Assume that

(3.11.*a*) a and b are bounded, so that

(3.11.*b*) M is a $L^2(P)$-(true) martingale, and assume furthermore that

(3.11.*c*) $|\beta_t| \leqq K$, for all t in $[0, T]$.

Under these assumptions the process $Z(\beta, \nu_0, P)$ is a P martingale which belongs to any $L^p(P)$, $1 \leqq p < +\infty$. Here are some other consequences:

(3.12.*a*) $N_.$ is a $L^2(Q)$ (true) martingale.

(3.12.*b*) One can choose for B a version of $E^{Q \times dt}[\beta_s \mid (X_s = x, s)]$, so that

$$|B(X_s, s)| \leqq K, \qquad Q \times dt \quad \text{a.s.}$$

(3.12.c) Since Ω is a Polish space we know that for almost all s,

$$B(X_s, s) = E^Q[\beta_s \mid X_s] \quad Q\,\text{a.s.}$$

It follows from (3.12.*b*) that G_t is a Q′-martingale which belongs to any $L^p(Q')$, $1 \leqq p < +\infty$, so that $\bar{Q}' = \bar{Q}''$ are Probability measures equivalent to P′. Define

$$\bar{N}_t = M_t - \int_0^t a(X_s, s)\, B(X_s, s)\, ds. \tag{3.13}$$

$\bar{N}$ is then a $L^2(\bar{Q})$ martingale. Furthermore $Q = \bar{G}_T \bar{Q}$, where $\bar{G}_t = E^{\bar{Q}}[\bar{G}_T \mid \mathscr{F}_t]$ is given by

$$\bar{G}_t = \exp\left\{ -\int_0^t (B(X_s, s) - \beta_s / d\bar{N}_s) - \frac{1}{2}\int_0^t (B(X_s, s) - \beta_s / a(X_s, s)(B(X_s, s) - \beta_s))\, ds \right\} \tag{3.14}$$

Hence for $f \in \mathscr{B}_b(\mathbb{R}^d)$ and $t \in]0, T]$ the following holds:

$$E^Q[f(X_t)] = E^{\bar{Q}}[f(X_t)] - E^{\bar{Q}}\left[f(X_t) \int_0^t \bar{G}_s (B(X_s, s) - \beta_s / d\bar{N}_s) \right]. \tag{3.15}$$

It remains to show that the second term of the sum vanishes. This result would be clear if one could exchange the stochastic integral and the conditional expectation $E^{\bar{Q}}[.\,|X_t]$ (stochastic conditional Fubini's theorem). Indeed, according to (3.12.*c*), for all $g\in\mathscr{B}(\mathbb{R}^d)$ and almost all s,

$$(3.16)\quad E^{\bar{Q}}[g(X_s)E^{\bar{Q}}[\bar{G}_s(\beta_s-B(X_s,s))\,|\,X_s]]$$
$$=E^{\bar{Q}}[g(X_s)\bar{G}_s(\beta_s-B(X_s,s))]=E^{Q}[g(X_s)(\beta_s-B(X_s,s))]=0.$$

So that

$$(3.17)\qquad E^{\bar{Q}}[\bar{G}_s(\beta_s-B(X_s,s))\,|\,X_s]=0\quad \bar{Q}\times ds\quad \text{a.s.}$$

To prove that the stochastic conditional Fubini's theorem holds, we shall use an approximation procedure of the stochastic integral (which appears p. 23 of Mac Kean's book "Stochastic integrals"). We thank Francis Comets for pointing out this result to us.

(3.18) LEMMA. – *Put* $t_i=it2^{-n}$, $i=0,\ldots,2^n$. *Let* h_s *be a previsible process in* $L^2(\bar{Q}\,x\,dt)$ (*with* $h_s=h_0$ *for* $s\leqq 0$). *Then*

$$\int_0^t (h_s/d\bar{N}_s)=\lim_n \frac{2^n}{t}\sum_{i=0}^{2^n-1}\left(\left(\int_{t_{i-1}}^{t_i} h_s\,ds\right)\Big/\bar{N}_{t_{i+1}}-\bar{N}_{t_i}\right)$$

where the limit is taken in $L^2(\bar{Q})$.

Proof of the Lemma. – The mapping

$$h_u\to h_u^n=\frac{2^n}{t}\sum_{i=0}^{2^n-1}\left(\int_{t_{i-1}}^{t_i} h_s\,ds\right)1_{u\in]t_i,t_{i+1}]}+h_0\,1_{u=0}$$

is continuous from $L^2(\bar{Q}\times dt)$ into itself, with norm equal to 1.

Furthermore if h is constant on each dyadic interval of length 2^{-k}, then for $n\geqq k$, h and h^n may differ on at most 2^{-k} intervals of length 2^{-n} (the left side of the dyadic intervals at level 2^{-k}). Hence for such an h

$$\|h-h^n\|_{L^2(\bar{Q}\times dt)}\leqq 2\,\|h\|_\infty\,t\,2^{k-n},\qquad \text{for}\quad n\geqq k.$$

Finally if h is a previsible process in $L^2(\bar{Q}\times dt)$, we can approximate h in $L^2(\bar{Q}\times dt)$ by a sequence of h_{dyad}. Since

$$\|h^n-h^n_{\text{dyad}}\|_{L^2(\bar{Q}\times dt)}\leqq\|h-h_{\text{dyad}}\|_{L^2(\bar{Q}\times dt)}\quad \text{for all } n$$

we immediatly see that h^n converges to h in $L^2(\bar{Q}\times dt)$. Remark that h^n is left continuous (will right limits) and hence previsible.

But a being bounded $E^{\bar{Q}}\left[\int_0^T (h_s/a(X_s,s)\,h_s)\,ds\right]\leqq C\;E^{\bar{Q}}\left[\int_0^T (h_s/h_s)\,ds\right]$, which proves that h^n converges to h in $L^2_{\bar{Q}}$ (*i.e.* $L^2(\bar{N},\bar{Q})$). The Lemma now follows from the isometry property of the L^2 stochastic integral. □

The previous approximation procedure is of course very useful because it does not require any kind of continuity property on the integrand. Recall that the approximation by the Riemann sums for instance, requires the left continuity of the paths.

We apply Lemma (3.18) with $h_s = \bar{G}_s(\beta_s - B(X_s, s))$ which satisfies the hypotheses of (3.18) since $\bar{G}$ is continuous, $\sup_s \bar{G}_s$ is square integrable with respect to $\bar{Q}$, and $\beta_s - B(X_s, s)$ is bounded and previsible. This yields

$$(3.19)\quad E^{\bar{Q}}[f(X_t)\int_0^t (h_s/d\bar{N}_s)]$$
$$= \lim_n \frac{2^n}{t} \sum_{i=0}^{2^n-1} E^{\bar{Q}}\left[f(X_t)\left(\left(\int_{t_{i-1}}^{t_i} h_s\,ds\right)\Big/\bar{N}_{t_{i+1}} - \bar{N}_{t_i}\right)\right]$$

But

$$E^{\bar{Q}}\left[f(X_t)\left(\left(\int_{t_{i-1}}^{t_i} h_s\,ds\right)\Big/\bar{N}_{t_{i+1}} - \bar{N}_{t_i}\right)\right] = \int_{t_{i-1}}^{t_i} E^{\bar{Q}}[f(X_t)(h_s/\bar{N}_{t_{i+1}} - \bar{N}_{t_i})]\,ds.$$

In the right hand term we can now take the conditional expectation with respect to $\sigma(X_v, v \geqq s)$. Of course $f(X_t)$ and $\bar{N}_{t_{i+1}} - \bar{N}_{t_i}$ are $\sigma(X_v, v \geqq s)$-measurable for $s \leqq t_i$. On the other hand, since $\bar{Q}$ is Markov and h_s is $\mathscr{F}_s$ measurable

$$E^{\bar{Q}}[h_s \mid \sigma(X_v, v \geqq s)] = E^{\bar{Q}}[h_s \mid X_s].$$

Finally

$$E^{\bar{Q}}\left[f(X_t)\left(\left(\int_{t_{i-1}}^{t_i} h_s\,ds\right)\Big/\bar{N}_{t_{i+1}} - \bar{N}_{t_i}\right)\right]$$
$$= \int_{t_{i-1}}^{t_i} E^{\bar{Q}}[f(X_t)(E^{\bar{Q}}[h_s/X_s]/\bar{N}_{t_{i+1}} - \bar{N}_{t_i})]\,ds.$$

Applying (3.17), we see that each term of the sum in (3.19) is equal to 0, which proves

$$E^{\bar{Q}}\left[f(X_t)\int_0^t \bar{G}_s(B(X_s, s) - \beta_s/d\bar{N}_s)\right] = 0.$$

(3.15) becomes $E^Q[f(X_t)] = E^{\bar{Q}}[f(X_t)]$. The proof is finished under the assumptions (3.11). □

Short comments on hypotheses (3.11). – The time reversal argument we used in the above proof (recall the conditionning) refrains us from using a localization procedure with stopping times. This explains (3.11.*a*) and (3.11.*b*). Lemma (3.18) requires to work in a L^2 framework. In order to ensure that the Girsanov densities are sufficiently integrable, it is natural to ask β and B to be bounded. Remark that in view of the definition of B, the boundedness of the action integral $\int_0^T (\beta_s/a(X_s, s)\,\beta_s)\,ds$ is not sufficient to get the boundedness of $\int_0^T (B(X_s, s)/a(X_s, s)\,B(X_s, s))\,ds$. This is the main reason for (3.11.*c*).

Step 2. — We keep (3.11.a), but replace (3.11.c) by

$$\int_0^{\mathrm{T}} (\beta_s / a(\mathrm{X}_s, s)\,\beta_s)\, ds \leqq \mathrm{C} \quad (\mathrm{P\ a.s.}) \tag{3.20}$$

Accordingly $\bar{\mathrm{Q}}' = \bar{\mathrm{Q}}''$ and Q' and $\bar{\mathrm{Q}}'$ are equivalent (*see* the discussion above). Define the sequence β_s^k of previsible processes as $\beta_s^k = \beta_s\, 1_{|\beta_s| < k}$. Then

$$\begin{cases} \text{(i)}\ |\beta_s^k| \leqq |\beta_s| \wedge k \qquad (|\ \ |\ \text{stands for the euclidean norm}), \\ \text{(ii)}\ \displaystyle\int_0^{\mathrm{T}} (\beta_s^k / a(\mathrm{X}_s, s)\,\beta_s^k)\, ds \leqq \int_0^{\mathrm{T}} (\beta_s / a(\mathrm{X}_s, s)\,\beta_s)\, ds \leqq \mathrm{C}, \end{cases} \tag{3.21}$$

and obviously β_s^k converges to β_s, P almost surely.

Let Q^k be the (β^k, ν_0, P)-FM. Q^k is a Probability measure thanks to (3.21.ii) and to Novikov criterion.

Moreover $\mathrm{H}(\mathrm{Q}^k, \mathrm{P}) \leqq \mathrm{C}$ [*see* proposition (2.1)].

We shall prove that a subsequence of $\mathrm{Q}^k \circ \mathrm{X}_t^{-1}$ converges to $\mathrm{Q} \circ \mathrm{X}_t^{-1}$.

Indeed thanks to (3.21) one can apply the bounded convergence theorem to get

$$\mathrm{E}^{\mathrm{P}}\left[\int_0^{\mathrm{T}} (\beta_s - \beta_s^k / a(\mathrm{X}_s, s)(\beta_s - \beta_s^k))\, ds\right] \underset{k}{\rightarrow} 0. \tag{3.22}$$

It follows that $\displaystyle\int_0^t (\beta_s^k / d\mathrm{M}_s) \underset{k}{\rightarrow} \int_0^t (\beta_s / d\mathrm{M}_s)$ in $\mathrm{L}^2(\mathrm{P})$, so that we can find a subsequence of $\mathrm{Z}_t^k \left(= \dfrac{d\mathrm{Q}^k}{d\mathrm{P}}\bigg|_{\mathscr{F}_t}\right)$ that converges P almost surely to $\mathrm{Z}_t \left(= \dfrac{d\mathrm{Q}}{d\mathrm{P}}\bigg|_{\mathscr{F}_t}\right)$. But thanks to (3.21.ii)

$$\begin{cases} \mathrm{E}^{\mathrm{P}}[(\mathrm{Z}_t^k)^p] \leqq \exp(p(p-1)\mathrm{C}/2) \\ \text{for all } k, \text{ and } p \in [1, +\infty[. \end{cases} \tag{3.23}$$

Thus for all $f \in \mathscr{B}_b(\mathbb{R}^d)$, the sequence $f(\mathrm{X}_t)\,\mathrm{Z}_t^k$ is bounded in all the $\mathrm{L}^p(\mathrm{P})$ and is thus uniformly integrable. It follows that for the above subsequence

$$\mathrm{E}^{\mathrm{Q}^k}[f(\mathrm{X}_t)] = \mathrm{E}^{\mathrm{P}}[f(\mathrm{X}_t)\,\mathrm{Z}_t^k] \underset{k}{\rightarrow} \mathrm{E}^{\mathrm{P}}[f(\mathrm{X}_t)\,\mathrm{Z}_t] = \mathrm{E}^{\mathrm{Q}}[f(\mathrm{X}_t)]. \tag{3.24}$$

One can prove that in fact, $\mathrm{H}(\mathrm{Q}, \mathrm{Q}^k)$ goes to 0. Hence this sequence converges to Q for the variation distance on $\mathrm{M}_1(\Omega)$. This provides a more direct proof of (3.24). But the idea of the above proof will be used in the sequel. Notice that in this case, we know that Z_t is a density of Probability. Since Z_t^k converges P a.s. to Z_t, the convergence of Q^k to Q is a consequence of Scheffé's theorem [Bi]. But the above argument is still true without assuming that Z (or Z^k) is a density of Probability, and will be used in this situation in the next sections.

To Q^k (resp. Q) corresponds $\bar{Q}^k$ (resp. $\bar{Q}$) builded as before. By step 1 we know that $\bar{Q}^k$ and Q^k have same marginals (β_s^k is bounded). But it is much more difficult to prove the convergence of $\bar{Q}^k \circ X_t^{-1}$ to $\bar{Q} \circ X_t^{-1}$. Indeed

$$(3.25) \qquad B^k(x, s) = E^{Q^k \times dt}[\beta_s^k \mid (X_s = x, s)]$$

and we do not have any uniform bound for

$$\int_0^T (B^k(X_s, s)/a(X_s, s)\, B^k(X_s, s))\, ds.$$

The key idea of the proof will be to use the Trick (2.6). Indeed the energy condition (2.6.i) is satisfied. So, in order to prove that $\bar{Q}$ is a Probability measure with marginals ν_t, it suffices to prove the domination relation

(3.26) for all nonnegative $f \in \mathscr{B}_b(\mathbb{R}^d)$, $E^{\bar{Q}}[f(X_t)\, 1_{t<S_\infty}] \leq E^{Q}[f(X_t)]$, where S_∞ is defined in (3.5.ii).

This will be proved by showing that for a subsequence $\bar{Q}^k$

$$(3.27) \qquad \begin{cases} E^{\bar{Q}}[f(X_t)\, 1_{t<S_\infty}] \leq \liminf_k E^{\bar{Q}^k}[f(X_t)] \\ (= \liminf_k E^{Q^k}[f(X_t)] = E^{Q}[f(X_t)]), \end{cases}$$

[Of course it is a subsequence of the subsequence for which (3.24) holds]. It is easy to guess that the inequality (3.27) will be obtained thanks to Fatou's lemma, if we prove the almost sure convergence of some densities. Hence, first notice that $Q \sim Q^k \sim \bar{Q}^k$, the density processes being given by:

$$(3.28) \qquad \begin{cases} \left.\dfrac{dQ^k}{dQ}\right|_{\mathscr{F}_t} = \Lambda_t^k = \exp\left\{ -\displaystyle\int_0^t (\beta_s - \beta_s^k / dN_s) \right. \\ \qquad \left. - \dfrac{1}{2} \displaystyle\int_0^t (\beta_s - \beta_s^k / a(X_s, s)(\beta_s - \beta_s^k))\, ds \right\} \\ \left.\dfrac{d\bar{Q}^k}{dQ}\right|_{\mathscr{F}_t} = \bar{\Lambda}_t^k = \exp\left\{ -\displaystyle\int_0^t (\beta_s - B^k(X_s, s) / dN_s) \right. \\ \left. - \dfrac{1}{2} \displaystyle\int_0^t (\beta_s - B^k(X_s, s)/a(X_s, s)(\beta_s - B^k(X_s, s)))\, ds \right\} \end{cases}$$

We shall prove that a subsequence of $\bar{\Lambda}_t^k$ converges Q almost surely to G_t [*see* (3.6)], since here $\left.\dfrac{d\bar{Q}}{dQ}\right|_{\mathscr{F}_t} = G_t$. Just as before it is enough to prove

$$(3.29) \qquad \lim_k E^Q\left[\int_0^T (B(X_s, s) - B^k(X_s, s) / a(X_s, s)(B(X_s, s) - B^k(X_s, s)))\, ds \right] = 0.$$

Recall the Kalman-Bucy formula

(3.30) $$B^k(X_s, s) = E^Q[\Lambda_s^k \beta_s^k \mid X_s]\, E^{Q^k}[(\Lambda_s^k)^{-1} \mid X_s] \quad Q\text{ a.s.}$$

noticing that $(\Lambda_s^k)^{-1}$ admits Q^k moments of any order. For simplicity we write

$$m_s^k = E^Q[\Lambda_s^k \beta_s^k \mid X_s]; \qquad n_s^k = E^{Q^k}[(\Lambda_s^k)^{-1} \mid X_s];$$
$$\gamma(h_s) = (h_s / a(X_s, s)\, h_s), \text{ for any previsible process } h.$$

For a better understanding of the computations, the reader can keep in mind that any quantity α whose vocation is to tend to 1 (for instance n_s^k), will be written as $1 + (\alpha - 1)$.

With the help of the inequality $\gamma(f+g) \leqq 2(\gamma(f) + \gamma(g))$ and of the conditional Jensen inequality, it holds:

$$\begin{aligned} E^Q[\gamma(B^k(X_s, s) - B(X_s, s))] &\leqq 2\, E^Q[\gamma(B(X_s, s) - m_s^k)] + 2\, E^Q[(1-n_s^k)^2 \gamma(m_s^k)], \\ &\leqq 2\, E^Q[\gamma(\beta_s - \Lambda_s^k \beta_s^k)] + 2\, E^Q[(1-n_s^k)^2 (\Lambda_s^k)^2 \gamma(\beta_s^k)]. \end{aligned}$$

Thus

$$\begin{aligned} &\int_0^T E^Q[\gamma(B(X_s, s) - B^k(X_s, s))]\, ds \\ &\quad \leqq 4 \int_0^T E^Q[\gamma(\beta_s - \beta_s^k)]\, ds \\ &\qquad + 4 \int_0^T E^Q[(\sup_s (1-\Lambda_s^k))^2 \gamma(\beta_s^k)]\, ds \\ &\qquad + 2 \int_0^T E^Q[(\sup_s \Lambda_s^k (1-n_s^k))^2 \gamma(\beta_s^k)]\, ds. \end{aligned}$$

Using (3.21.ii), we see that the above quantity is less than

(3.31) $$4 \int_0^T E^Q[\gamma(\beta_s - \beta_s^k)]\, ds + 4\, C\, E^Q[(\sup_s (1-\Lambda_s^k))^2] + 2\, C\, E^Q[(\sup_s \Lambda_s^k (1-n_s^k))^2]$$

Likewise (3.22) the first term in (3.31) goes to 0. Since $(1-\Lambda_s^k)$ is a Q-martingale, one can use Doob's inequality to obtain

$$E^Q[(\sup_s (1-\Lambda_s^k))^2] \leqq 4\, E^Q[(1-\Lambda_T^k))^2].$$

It is easy to see that (a subsequence of) $(1-\Lambda_T^k)$ goes Q almost surely to 0 and is uniformly bounded (in k) in all the $L^p(Q)$ for $1 \leqq p < +\infty$ [*cf.* (3.23)]. Thus the second term of (3.31) also goes to 0.

We come now to the third and last term. Since $\sup_s \Lambda_s^k$ is also uniformly bounded in all the $L^p(Q)$, it suffices to control $E^Q[(\sup_s (1-n_s^k))^4]$ and to apply Cauchy-Schwarz inequality. Once again we use the conditional Jensen inequality, Cauchy Schwarz and Doob inequalities to get:

$$\begin{aligned}
&E^Q[(\sup_s (1-n_s^k))^4] \\
&\leqq E^{Q^k}[(\Lambda_T^k)^{-1} \sup_s E^{Q^k}[(1-(\Lambda_s^k)^{-1})^4 \mid X_s]], \\
&\leqq (E^{Q^k}[(\Lambda_T^k)^{-2}])^{1/2} (E^{Q^k}[\sup_s (1-(\Lambda_s^k)^{-1})^8])^{1/2}, \\
&\leqq c\,(E^{Q^k}[(\Lambda_T^k)^{-2}])^{1/2} (E^{Q^k}[(1-(\Lambda_T^k)^{-1})^8])^{1/2}, \\
&\leqq c\,(E^Q[(\Lambda_T^k)^{-1}])^{1/2} (E^Q[\Lambda_T^k(1-(\Lambda_T^k)^{-1})^8])^{1/2}.
\end{aligned}$$

Since (a subsequence of) $(\Lambda_T^k)^{-1}$ goes Q almost surely to 1 and since all quantities appearing in the previous bound are uniformly bounded (in k) in all the $L^p(Q)$, we deduce that the last term of (3.31) tends to 0.

We have thus proved that a subsequence of $\bar{\Lambda}_t^k$ converges Q almost surely to G_t. Now (3.27) is a consequence of Fatou's lemma. □

Short comments. — The computations of step 2 are merely tedious but unfortunately we cannot avoid this technical step. Indeed (3.20) is the natural hypothesis in view of a localization procedure and as we already said, it is unsufficient for the first step techniques to be available.

Step 3. — We shall now remove (3.11.a) and (3.20). Let $\beta \in L_Q^2$ and put

(3.32) $$\beta_s^n = \beta_s 1_{s \leqq \varepsilon_n} \quad \text{where} \quad \varepsilon_n = T_n \wedge s_n;$$

so that

(3.33) $$\int_0^T (\beta_s^n / a(X_s, s)\, \beta_s^n)\, ds \leqq n.$$

Denote by Q^n the (β_s^n, ν_0, P)-FM which is a Probability measure on Ω, with density Z^n. B^n is defined by

(3.34) $$B^n(x, s) = E^{Q^n \times dt}[\beta_s^n \mid (X_s = x, s)],$$

where the conditional expectation is taken in the sense of (3.2). Notice that $B^n(x, s)$ is $\nu_n \times ds$ almost surely equal to 0 on $\{|x| > n\}$, where ν_n denotes the flow of the marginals of Q^n. So we can choose B^n such that B^n is identically equal to 0 on $\{|x| > n\}$.

(3.35) Let $\bar{Q}^n$ denote the (B^n, ν_0, P)-FM and

$$\theta_{n,m} = \inf\left\{t \geqq 0, \int_0^t (B^n(X_s, u_s)/a(X_s, u_s)\, B^n(X_s, u_s))\, ds \geqq m\right\} \wedge s_m.$$

(3.36) Lemma. — *$\bar{Q}^n$ and Q^n have same marginals.*

Proof. — We introduce a test function $\varphi^k \in C_0^\infty(\mathbb{R}^d)$ such that $\varphi^k(x) = x$ if $|x| \leqq k-1$ and $\varphi^k(x) = 0$ if $|x| \geqq k$. Denote by P^k the P law of $\varphi^k(X)$.

Ito's formula shows that P^k is a (not necessarily extremal) solution to the martingale problem (1.6.ii) with bounded coefficients a^k and b^k.

Next define $Q^{n,k}$ as the (β^n, ν_0, P^k)-FM, $Z_T^{n,k}$ denoting the corresponding density. For $k>n$, one can remove the k in $Z_T^{n,k}$, and $Q^{n,k}$ coincides with Q^n on $\mathscr{F}_{\varepsilon_n}$, because P^k coincides with P on $\mathscr{F}_{s_n}$. Furthermore for $k \geqq n$ and $f \in \mathscr{B}_b(\mathbb{R}^d)$, it holds

$$(3.37) \qquad E^{Q^{n,k}}[f(X_t)] = E^{P^k}[Z_t^n f(X_t)] = E^P[Z_t^n f(\varphi^k(X_t))].$$

Indeed $Z_t^n = Z_{t \wedge \varepsilon_n}^n$, P coincides with P^k on $\mathscr{F}_{\varepsilon_n}$ and $\varphi^k(X_t) = X_t$ on $\{t \leqq \varepsilon_n\}$, so that

$$E^{P^k}[Z_t^n f(X_t) 1_{t \leqq \varepsilon_n}] = E^P[Z_t^n f(\varphi^k(X_t)) 1_{t \leqq \varepsilon_n}],$$

and

$$\begin{aligned} E^{P^k}[Z_t^n f(X_t) 1_{t > \varepsilon_n}] &= E^{P^k}[Z_{t \wedge \varepsilon_n}^n f(X_t) 1_{t > \varepsilon_n}] \\ &= E^{P^k}[Z_{t \wedge \varepsilon_n}^n 1_{t > \varepsilon_n} f(X_{t-\varepsilon_n} \circ \theta_{\varepsilon_n})] \\ &= E^P[Z_{t \wedge \varepsilon_n}^n 1_{t > \varepsilon_n} f(\varphi^k(X_{t-\varepsilon_n} \circ \theta_{\varepsilon_n})] \\ &= E^P[Z_t^n f(\varphi^k(X_t)) 1_{t > \varepsilon_n}], \end{aligned}$$

thanks to the strong Markov property of P and P^k.

We can apply the bounded convergence theorem in (3.37) and get

$$(3.38) \qquad \lim_k E^{Q^{n,k}}[f(X_t)] = \lim_k E^P[Z_t^n f(\varphi^k(X_t))] = E^{Q^n}[f(X_t)].$$

In the same way, one can define $\bar{Q}^{n,k}$, which coincides with $\bar{Q}^n$ on $\mathscr{F}_{\theta_{n,m}}$ for $k>m$, thanks to our choice of B^n. We know that $\bar{Q}^{n,k}$ and $Q^{n,k}$ have same marginals, according to step 2. Now

$$E^{P^k}[\bar{Z}_t^n f(X_t) 1_{t \leqq \theta_{n,m}}] = E^P[\bar{Z}_t^n f(\varphi^k(X_t)) 1_{t \leqq \theta_{n,m}}] \qquad \text{for} \quad k>m.$$

By taking the limit in k, we first obtain that

$$\lim_k E^{\bar{Q}^{n,k}}[f(X_t) 1_{t \leqq \theta_{n,m}}] = \lim_k E^P[\bar{Z}_t^n f(\varphi^k(X_t)) 1_{t \leqq \theta_{n,m}}] = E^{\bar{Q}^n}[f(X_t) 1_{t \leqq \theta_{n,m}}].$$

Together with (3.38) this yields

$$E^{\bar{Q}^n}[f(X_t) 1_{t \leqq \theta_{n,m}}] \leqq E^{Q^n}[f(X_t)],$$

for any nonnegative f and any m.

But $\theta_{n,m}$ goes P a.s. to

$$\inf\left\{ t \geqq 0, \int_0^t (B^n(X_s, u_s)/a(X_s, u_s) B^n(X_s, u_s))\, ds = +\infty \right\}$$

when m goes to infinity, and *we can thus apply the trick* (2.6) *to obtain that* $\bar{Q}^n$ *and* Q^n *have same marginals.* □

Now let us take the limit in n.

The norm of $Z_{t \wedge \varepsilon_n}$ is bounded by $H(Q, P)$ in $L_{\tau^*}(P)$ and $Z_{t \wedge \varepsilon_n} \to Z_t - P$ a.s. Hence $Q^n \circ X_t^{-1}$ converges to $Q \circ X_t^{-1}$ when n goes to infinity.

Once again it is much more difficult to study the behaviour of $\bar{Q}^n$. Introduce

(3.39) $$B_k = B\, 1_{|B| < k},\ \bar{Q}_k \text{ the } (B_k, \nu_0, P)\text{-FM}.$$

Let m be a fixed integer, S_m is defined by (3.5.ii). Since $B_k \to B$, P a.s. on the set $\{|B| < +\infty\}$, one has

(3.40) $$\lim_{k \to 0} E^P \left[\int_0^T 1_{s \leqq S_m} (B - B_k / a (B - B_k)) (X_s, s)\, ds \right] = 0,$$

thanks to the bounded convergence theorem. Notice that

(3.41) $$\int_0^T 1_{s \leqq S_m} (B_k / a B_k) (X_s, s)\, ds \leqq m, \quad \text{for all } k.$$

We then deduce that (a subsequence of) $Z_{t \wedge S_m}(B_k, \nu_0, P)$ converges P almost surely to $Z_{t \wedge S_m}(B, \nu_0, P)$. Furthermore $Z_{t \wedge S_m}(B_k, \nu_0, P)$ is a bounded sequence of $L^2(P)$, thanks to (3.41). For the above subsequence and any bounded f, we get

(3.42) $$E^{\bar{Q}}[f(X_t)\, 1_{t < S_m}] = \lim_{k \to \infty} E^{\bar{Q}_k} [f(X_t)\, 1_{t < S_m}].$$

We shall now compare the right hand side with $E^{\bar{Q}^n}$.

Since $\bar{Q}_k$ and $\bar{Q}^n$ are equivalent, we denote by $G^{n, k}$ the density of $\bar{Q}^n$ with respect to $\bar{Q}_k$. For any measurable bounded F defined on Ω, it holds

(3.43) $$E^{\bar{Q}^n}[F] = E^{\bar{Q}_k}[F] + E^{\bar{Q}_k}[F\, (G^{n, k} - 1)].$$

We shall now calculate an upper bound for the last term. According to [Föl] p. 133,

(3.44) $$|E^{\bar{Q}_k}[F (G^{n, k} - 1)]| \leqq 2 \| F \|_\infty [H(\bar{Q}^n, \bar{Q}_k)]^{1/2}.$$

Assume for the moment that the following lemma holds.

(3.45) LEMMA. – $\limsup_{n \to \infty} H(\bar{Q}^n, \bar{Q}_k) \leqq (2 H(Q, P))^{1/2} \| B - B_k \|_{L^2_V}.$

Since $\bar{Q}^n$ and Q^n have same marginals, and Q^n converges to Q, it holds

(3.46) $$E^Q[f(X_t)\, 1_{t < S_m}] \leqq E^{\bar{Q}_k}[f(X_t)\, 1_{t < S_m}] + 2 \| f \|_\infty [(2 H(Q, P))^{1/2} \| B - B_k \|_{L^2_V}]^{1/2}.$$

Now apply (3.42), letting k go to infinity, in order to get

(3.47) $$E^Q[f(X_t)\, 1_{t < S_m}] \leqq E^{\bar{Q}}[f(X_t)\, 1_{t < S_m}].$$

This proves first that $T = S_\infty \leqq \xi - \bar{Q}$ a.s. (make $f = 1$ and $m \to \infty$), next that Q and $\bar{Q}$ have same marginals for $t \in [0, T[$. One can thus use the trick (2.6) to prove that $\xi = \infty - \bar{Q}$ a.s.

To finish the proof of (3.1), it remains to prove Lemma (3.45). As the patient reader will see, the proof is desperately technical.

Proof of (3.45). – By Proposition 2.1

$$2\,\mathrm{H}(\bar{\mathrm{Q}}^n, \bar{\mathrm{Q}}_k)$$

$$= \mathrm{E}^{\bar{\mathrm{Q}}^n}\left[\int_0^{\mathrm{T}} (\mathrm{B}_k(\mathrm{X}_s, s) - \mathrm{B}^n(\mathrm{X}_s, s)/a(\mathrm{X}_s, s)(\mathrm{B}_k(\mathrm{X}_s, s) - \mathrm{B}^n(\mathrm{X}_s, s)))\,ds\right]$$

$$= \mathrm{E}^{\mathrm{Q}^n}\left[\int_0^{\mathrm{T}} (\mathrm{B}_k(\mathrm{X}_s, s) - \mathrm{B}^n(\mathrm{X}_s, s)/a(\mathrm{X}_s, s)(\mathrm{B}_k(\mathrm{X}_s, s) - \mathrm{B}^n(\mathrm{X}_s, s)))\,ds\right],$$

since $\bar{\mathrm{Q}}^n$ and Q^n have same marginals. But

$$\lim_{n\to\infty} \mathrm{E}^{\mathrm{Q}^n}\left[\int_0^{\mathrm{T}} (\mathrm{B}_k/a\,\mathrm{B}_k)(\mathrm{X}_s, s)\,ds\right] = \mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} (\mathrm{B}_k/a\,\mathrm{B}_k)(\mathrm{X}_s, s)\,ds\right]. \tag{3.49}$$

$$\begin{aligned} &\mathrm{E}^{\mathrm{Q}^n}\left[\int_0^{\mathrm{T}} (\mathrm{B}^n/a\,\mathrm{B}_k)(\mathrm{X}_s, s)\,ds\right] \\ &\quad = \mathrm{E}^{\mathrm{Q}^n}\left[\int_0^{\mathrm{T}} 1_{s\leqq\varepsilon_n}(\beta_s/a(\mathrm{X}_s, s)\,\mathrm{B}_k(\mathrm{X}_s, s))\,ds\right] \\ &\quad = \mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} 1_{s\leqq\varepsilon_n}(\beta_s/a(\mathrm{X}_s, s)\,\mathrm{B}_k(\mathrm{X}_s, s))\,ds\right], \end{aligned} \tag{3.50}$$

so that by the monotone convergence theorem and (3.2)

$$\begin{aligned} &\lim_{n\to\infty} \mathrm{E}^{\mathrm{Q}^n}\left[\int_0^{\mathrm{T}} (\mathrm{B}^n/a\,\mathrm{B}_k)(\mathrm{X}_s, s)\,ds\right] \\ &\quad = \mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} (\beta_s/a(\mathrm{X}_s, s)\,\mathrm{B}_k(\mathrm{X}_s, s))\,ds\right] \\ &\quad = \mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} (\mathrm{B}/a\,\mathrm{B}_k)(\mathrm{X}_s, s))\,ds\right] \\ &\quad = \mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} (\mathrm{B}_k/a\,\mathrm{B}_k)(\mathrm{X}_s, s)\,ds\right]. \end{aligned} \tag{3.51}$$

Finally, define $h^n(x, s) = \mathrm{E}^{\mathrm{Q}}[1_{s\leqq\varepsilon_n}/(\mathrm{X}_s = x, s)]$ (usual conditional expectation), by choosing a version which is always less than 1. Then $h^n\mathrm{B}^n \in \mathrm{L}^2_\nu$ [defined in (2.5)]. Indeed if we define $\theta_k = 1_{|\mathrm{B}^n|<k}$, one has

$$\begin{aligned} &\mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} (h^n\,\mathrm{B}^n\,\theta_k/ah^n\,\mathrm{B}^n\,\theta_k)(\mathrm{X}_s, s)\,ds\right] \\ &\quad \leqq \mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} (h^n\,\theta_k(\mathrm{B}^n/a\,\mathrm{B}^n))(\mathrm{X}_s, s)\,ds\right] \\ &\quad \leqq \mathrm{E}^{\mathrm{Q}}\left[\int_0^{\mathrm{T}} 1_{s\leqq\varepsilon_n}(\theta_k(\mathrm{B}^n/a\,\mathrm{B}^n))(\mathrm{X}_s, s)\,ds\right] \end{aligned}$$

$$\leq E^{Q^n}\left[\int_0^T 1_{s \leq \varepsilon_n} (\theta_k (B^n/a B^n)) (X_s, s)\, ds\right]$$
$$\leq 2 H (Q^n, P) \leq 2 H (Q, P);$$

so that

$$\text{(3.52)} \qquad E^Q\left[\int_0^T (h^n B^n/ah^n B^n) (X_s, s)\, ds\right] \leq 2 H (Q, P),$$

by the monotone convergence theorem. Now

$$\begin{aligned}
\text{(3.53)} \quad & E^{Q^n}\left[\int_0^T (h^n (B^n/a B^n)) (X_s, s)\, ds\right] \\
&= E^{Q^n}\left[\int_0^T 1_{s \leq \varepsilon_n} h^n (X_s, s) (\beta_s/(a B^n) (X_s, s))\, ds\right] \\
&= E^Q\left[\int_0^T 1_{s \leq \varepsilon_n} h^n (X_s, s) (\beta_s/a (X_s, s) B^n (X_s, s))\, ds\right] \\
&= E^Q\left[\int_0^T 1_{s \leq \varepsilon_n} (h^n (B/a B^n)) (X_s, s))\, ds\right] \\
&\quad + E^Q\left[\int_0^T 1_{s \leq \varepsilon_n} h^n (X_s, s) (\beta_s - B (X_s, s))/a (X_s, s) B^n (X_s, s))\, ds\right] \\
&= E^Q\left[\int_0^T 1_{s \leq \varepsilon_n} (h^n (B_k/a B^n)) (X_s, s))\, ds\right] \\
&\quad + E^Q\left[\int_0^T 1_{s \leq \varepsilon_n} (h^n ((B - B_k)/a B^n)) (X_s, s))\, ds\right] \\
&\quad + E^Q\left[\int_0^T 1_{s > \varepsilon_n} h^n (X_s, s) (\beta_s - B (X_s, s))/a (X_s, s) B^n (X_s, s))\, ds\right].
\end{aligned}$$

The last equality is obtained by noticing that

$$E^Q\left[\int_0^T h^n (X_s, s) ((\beta_s - B (X_s, s))/a (X_s, s) B^n (X_s, s))\, ds\right] = 0.$$

The introduction of B_k is necessary for the sequel because we do not know if $B \in L^2_{\nu_n}$. We now evaluate the last two terms in (3.53).

$$\begin{aligned}
\text{(3.54)} \quad & E^Q\left[\int_0^T 1_{s \leq \varepsilon_n} (h^n (B - B_k)/a B^n)) (X_s, s))\, ds\right] \\
&\qquad\qquad \leq (2 H (Q, P))^{1/2} \| B - B_k \|_{L^2_\nu}
\end{aligned}$$

thanks to (3.52) and Cauchy Schwarz. Similarly

$$(3.55)\quad E^Q\left[\int_0^T 1_{s>\varepsilon_n} h^n(X_s, s)(\beta_n - B(X_s, s))/a(X_s, s)\, B^n(X_s, s))\, ds\right]$$
$$\leqq (2\,H(Q, P))^{1/2}\left| E^Q\left[\int_0^T 1_{s>\varepsilon_n}((\beta_s - B(X_s, s))/ a(X_s, s)(\beta_s - B(X_s, s))\, ds\right]\right|^{1/2},$$

which goes to 0 when n goes to infinity since $\varepsilon_\infty = T - Q$ a.s.

It remains to evaluate the first term. Define

$$g^n(x, s) = E^{Q^n}[1_{s\leqq\varepsilon_n}/(X_s = x, s)].$$

$$(3.56)\quad E^Q\left[\int_0^T 1_{s\leqq\varepsilon_n}(h^n(B_k/a\,B^n))(X_s, s)\, ds\right]$$
$$= E^{Q^n}\left[\int_0^T 1_{s\leqq\varepsilon_n}(h^n(B_k/a\,B^n))(X_s, s))\, ds\right]$$
$$= E^{Q^n}\left[\int_0^T (g^n h^n(B_k/a\,B^n))(X_s, s))\, ds\right]$$
$$= E^{Q^n}\left[\int_0^T 1_{s\leqq\varepsilon_n}(g^n h^n(B_k/a(X_s, s)\,\beta_s)\, ds\right]$$
$$= E^{Q}\left[\int_0^T 1_{s\leqq\varepsilon_n}(g^n h^n(B_k/a(X_s, s)\,\beta_s)\, ds\right]$$
$$= E^{Q}\left[\int_0^T 1_{s\leqq\varepsilon_n}(g^n h^n(B_k/a\,B))(X_s, s))\, ds\right]$$
$$+ E^{Q}\left[\int_0^T 1_{s>\varepsilon_n}(g^n h^n)(X_s, s)((\beta_s - B(X_s, s))/a(X_s, s)\, B_k(X_s, s))\, ds\right]$$
$$\leqq E^Q\left[\int_0^1 (B_k/a\,B_k)(X_s, s))\, ds\right] + y(n, k),$$

where $y(n, k)$ goes to 0 when n goes to infinity, uniformly in k [same proof as (3.55)]. The inequality follows from $0\leqq g^n h^n \leqq 1$.

We have thus obtained

$$(3.57)\quad \limsup_{n\to\infty} E^{Q^n}\left[\int_0^T (h^n(B^n/a\,B^n))(X_s, s)\, ds\right]$$
$$\leqq E^Q\left[\int_0^T (B_k/a\,B_k)(X_s, s))\, ds\right] + (2\,H(Q, P))^{1/2}\,\|B - B_k\|_{L^2_V}.$$

What about $E^{Q^n}\left[\int_0^T ((1-h^n)(B^n/a\,B^n))(X_s, s)\,ds\right]$? Well replacing h^n by $(1-h^n)$ in the above computations, one get

$$(3.58)\quad \limsup_{n\to\infty} E^{Q^n}\left[\int_0^T (1-h^n)(B^n/a\,B^n))(X_s, s)\,ds\right] \leqq (2\,H(Q, P))^{1/2}\,\|B-B_k\|_{L^2_v} + \limsup_{n\to\infty} E^Q\left[\int_0^T 1_{s\leqq\varepsilon_s}(g^n(1-h^n)(B_k/a\,B_k)(X_s, s))\,ds\right] \leqq (2\,H(Q, P))^{1/2}\,\|B-B_k\|_{L^2_v},\quad \text{since } h^n\to 1\ Q\text{ a.s.}$$

Putting (3.58), (3.57), (3.51) and (3.49) together, we finally obtain

$$(3.59)\quad \limsup_{n\to\infty} 2\,H(\bar{Q}^n, \bar{Q}_k)\leqq 2\,(2\,H(Q, P))^{1/2}\,\|B-B_k\|_{L^2_v}.\quad \square$$

Remarks. – If we assume that a and b are bounded, Lemma (3.36) is unnecessary. Another proof of (3.1) will be given in Section 5, by using some results on Markov semi groups. Though the last step of the above proof is unreadable, we think that at least the first two steps have some interest from the stochastic calculus point of view.

The next Theorem, which relies Markov Probability measures with Markovian drifts, is very intuitive but requires a proof.

(3.60) THEOREM. – *Let* $Q\in M_1(\Omega)$ *be a Markov Probbility measure with* $H(Q, P)<+\infty$, *and* $v_.$ *the flow of its marginals. Then the drift* β_s *of* Q *can be written* $\beta_s=B(X_s, s)$, *where* B *belongs to* L^2_v.

Proof. – Since $H(Q, P)$ is finite, $\beta\in L^2_Q$. Define B as in (3.2). Let $f\in C_0^\infty(\mathbb{R}^d\times[0, T])$, and denote by C_t^f the bounded P martingale

$$C_t^f=f(X_t, t)-f(X_0, 0)-\int_0^t\left(\frac{\partial f}{\partial s}+A(s)\,f\right)(X_s, s)\,ds.$$

By Ito's formula, $C_t^f-\int_0^t(\nabla_x f(X_s, s)/a(X_s, s)\,\beta_s)\,ds$ is a local Q-martingale with localizing sequence $s_n\wedge T_n$, where ∇_x denotes the gradient operator relative to the spatial coordinates. Hence, for all $h\in\mathscr{B}_b(\mathbb{R}^d\times[0, T])$, and $t>u$,

$$(3.61)\quad E^Q\left[\left\{C^f_{t\wedge T_n\wedge s_n}-C^f_{u\wedge T_n\wedge s_n}-\int_{u\wedge T_n\wedge s_n}^{t\wedge T_n\wedge s_n}(\nabla_x f(X_s, s)/a(X_s, s)\,\beta_s)\,ds\right\} h(X_{u\wedge T_n\wedge s_n}, u\wedge T_n\wedge s_n)\right]=0.$$

But $\nabla_x f$ is bounded, with compact support. An application of Cauchy-Schwarz inequality yields:

$$\left|\int_{u \wedge T_n \wedge s_n}^{t \wedge T_n \wedge s_n} (\nabla_x f(X_s, s)/a(X_s, s)\,\beta_s)\,ds\right| \leq C \left|\int_0^T (\beta_s/a(X_s, s)\,\beta_s)\,ds\right|^{1/2}.$$

The right hand term belongs to $L^2(Q)$ ($\subseteq L^1(Q)$) because of the finite entropy assumption. Since all the other terms of (3.61) are bounded, one can use Lebesgue bounded convergence theorem. In the limit we get:

$$(3.62) \quad E^Q\left[\left\{C_t^f - C_u^f - \int_u^t (\nabla_x f(X_s, s)/a(X_s, s)\,\beta_s)\,ds\right\} h(X_u, u)\right] = 0.$$

Now applying (3.2) with $\varphi(x, s) = \nabla_x f(x, s)\, h(x, u)\, \mathbf{1}_{[u, t]}(s)$ we get:

$$(3.63) \quad E^Q\left[\left\{C_t^f - C_u^f - \int_u^t (\nabla_x f(X_s, s) \Big/ a(X_s, s)\,B(X_s, s))\,ds\right\} h(X_u, u)\right] = 0.$$

But since Q is Markov, (3.63) means that

$$C_t^f - \int_0^t (\nabla_x f(X_s, s)/a(X_s, s)\,B(X_s, s))\,ds$$

is a Q-martingale. Just as in part *a*) of the proof of Theorem 13.55 in [Jac], this shows that the process

$$X_t - X_0 - \int_0^t (b(X_s, s) + a(X_s, s)\,B(X_s, s)\,ds$$

is a local Q-martingale. Now Theorem 12.48 of [Jac] applies, which proves (3.60). □

Remark. – Theorem 12.48 in [Jac] also shows that for each $B \in L_v^2$, there exists at most one Q in $A_{v,H}$ with drift B.

Theorems (3.1) and (3.60) are interseting for two reasons. First the study of the non vacuity of $A_{v,H}$ reduces to the construction of diffusions with a Markovian drift. Second, the minimization problem also reduces to the case of Markov diffusions in $A_{v,H}$. We shall study both problems in the next two sections.

4. ADMISSIBLE FLOWS AND ASSOCIATED DIFFUSIONS

In this section we give some sufficient conditions for $A_{v,H}$ to be non empty. Connection with Nelson's diffusions (or Carlen's conservative diffusions [Ca], [Zhe], [MZ1], [Mi]...) will be briefly discussed. A systematic study of $A_{v,H}$ in the spirit of stochastic mechanics is done in [CP]. Recall that A′ is (formally) defined as

$$(4.1) \qquad A' = A(s) + \frac{\partial}{\partial s}.$$

(4.2) DEFINITION. — *We say that* $f \in \mathscr{B}_b(\mathbb{R}^d \times \mathbb{R})$ *belongs to the extended domain* D_e *of* A', *if there exists* $g \in \mathscr{B}(\mathbb{R}^d \times \mathbb{R})$ *with*

$$\int_0^T |g(X_s, u_s)|\, ds < +\infty - P_{x,u} \text{ a.s.} \qquad \textit{for all} \quad (x, u),$$

and

$$C_t^f = f(X_t, u_t) - f(X_0, u_0) - \int_0^t g(X_s, u_s)\, ds$$

is for all (x, u) *a local* $P_{x,u}$*-martingale. If* $f \in D_e$ *we put* $A' f = g$.

Of course if $f \in D_e$, C_t^f is a P′ local martingale.

Notice that Definition (4.2) is not exactly the one in [DM4, p. 244], but an extension of it (*see* the "commentaire 21", p. 244 in [DM4]). Also notice that $A' f$ is defined up to a set of potential zero.

If $f \in D_e$, C_t^f is a local martingale, additive functional, and $P_{x,u}$ is, for all (x, u), an extremal solution to the martingale problem (1.6). It follows that one can find a function denoted by ∇f, with $\nabla f \in \mathscr{B}(\mathbb{R}^d \times \mathbb{R}, \mathbb{R}^d)$, such that

$$(4.3) \quad \text{for all } (x, u),\ C_t^f = \int_0^t (\nabla f(X_s, u_s) / dM_s) \text{ for all } t,\ P_{x,u} \text{ a.s.}$$

Of course $C_0^\infty \subseteqq D_e$, and for a smooth f, $\nabla f = \nabla_x f$, where ∇_x is the usual gradient operator in the space direction. Here again ∇f is defined up to a set of potential zero.

(4.4) DEFINITION. — *Let* $(\nu_t)_{t \in [0, T]}$ *be a flow of Probability measures on* $\mathbb{R}^d$. *We call* f *a* ν*-good function if* $f \in D_e$ *and if there exist choices of* ∇f *and* $A' f$ *such that*

(i) $\int_0^T \int_{\mathbb{R}^d} |A' f(x, s)|\, d\nu_s\, ds < +\infty$,

(ii) $\nabla f \in L_\nu^2$.

The set of ν*-good functions is denoted by* $D_{e, \nu}$.

Any smooth function with compact support is a ν-good funtion.

If $\nu_.$ is admissible, (i) and (ii) hold for any choices of ∇f and $A' f$.

(4.5) We fix once for all a given flow $\nu_.$, such that $H(\nu_0, \mu_0) < +\infty$.

The next proposition gives a necessary condition for $\nu_.$ to be admissible.

(4.6) PROPOSITION. — *Assume that* $A_{\nu, H}$ *is non void. Let* Q *be any Markov Probability measure in* $A_{\nu, H}$ [*such a* Q *exists thanks to Theorem* (3.1)], *and* B *be its associated drift. Then for any* $f \in D_{e,\nu}$

$$\bar{C}_t^f = f(X_t, t) - f(X_0, 0) - \int_0^t [A' f(X_s, s) + (\nabla f(X_s, s)/a(X_s, s)B(X_s, s))]\, ds$$

is a L^2 Q-*martingale.*

(4.7) Corollary. — *If* $A_{\nu, H}$ *is non void, there exists* $B \in \mathscr{B}(\mathbb{R}^d \times \mathbb{R}, \mathbb{R}^d)$ *such that*

(i) $B \in L^2_\nu$,

(ii) *for any* $f \in D_{e,\nu}$ *and any* $0 \leqq s \leqq t \leqq T$,

$$\int_{\mathbb{R}^d} f(x, t)\, d\nu_t - \int_{\mathbb{R}^d} f(x, s)\, d\nu_s = \int_s^t \int_{\mathbb{R}^d} [A' f + (\nabla f/a\, B)](x, u)\, d\nu_u\, du.$$

(4.8) *Notation.* — Let Λ be a set of Borel functions such that $C_0^\infty \subseteqq \Lambda \subseteqq D_{e,\nu}$. If the relation (4.7 ii) holds for any $f \in \Lambda$, we say that $\nu_.$ satisfies the (B, Λ)-weak forward equation.

Proof. — To prove (4.6) it suffices to mimic the proof of (3.60). Indeed let ε_n be a localizing sequence of stopping times for the local P-martingale C_t^f. Then (3.61) remains true with $T_n \wedge \varepsilon_n$, B and ∇f instead of $T_n \wedge s_n$, β and $\nabla_x f$. Since f is a ν-good function, $\nabla f \in L^2_\nu$ and $\int_0^T \int_{\mathbb{R}^d} |A' f(x, s)|\, d\nu_s\, ds < +\infty$, so that one can again apply the bounded convergence theorem to get (3.63), which proves that $\bar{C}_t^f$ is a Q-martingale. The rest of (4.6) and (4.7) is then immediate. □

The converse statement of (4.7) would be the following:

(4.9) *Statement.* — Let $B \in \mathscr{B}(\mathbb{R}^d \times \mathbb{R}, \mathbb{R}^d)$, Λ be a set of Borel functions such that $C_0^\infty \subseteqq \Lambda \subseteqq D_{e,\nu}$. Assume that

(i) $H(\nu_0, \mu_0) < +\infty$;

(ii) $B \in L^2_\nu$;

(iii) ν satisfies the (B, Λ)-weak forward equation.

Then the (B, ν_0, P)-FM [*see* (1'.13)] belongs to $A_{\nu, H}$.

This statement is an extension of Nelson's problem solved by Carlen [Ca] in the special case where $A(t) = \frac{1}{2}\Delta$ (with additional assumptions such as the backward energy condition for instance). Various other references are relevant. Among them let us mention Meyer and Zheng [MZ1] where the problem is solved for a general symmetric Markov process, Zheng [Zhe] for a probabilistic approach of Carlen's case and Mikami [Mi] for an improvement of Carlen's method.

The results we shall derive are more general (for ordinary diffusions) since we are dealing with a general second order generator and we do not impose backward conditions. Moreover, as in [Zhe], we build a solution which is absolutely continuous with respect to the initial P, in the spirit of the Girsanov transform theory (recall the definition of the Föllmer measure). This answers to a question formulated by Nelson [Ne]. Actually our method consists essentially in piecing together small parts of Zheng's and Mikami's papers, with a bit of original touch. We also want to say that Mikami's proof could be extended to our general setting (with some slight restrictions), but would not provide all the consequences one can deduce from ours. For a complete discussion and bibliography on the subject see [CP].

Although we did not succeed in proving (4.9) in full generality, the method provides various results in different contexts. Hence, before stating a first result we shall give the spirit of the proof. We use the word "good" when a technical difficulty appears.

(4.9 *bis*) *Outline of the "proof" of* (4.9). — First, we follow [Mi].

Let B_k be a sequence of bounded Borel functions with compact support such that B_k converges to B in L^2_v. We define $Q^k_{x,u}$ as the $(B_k, \delta_x \otimes \delta_u, P_{x,u})$-FM. Thanks to the compactness of the support of B_k, $(Q^k_{x,u})$ is a family of Markov Probability laws on Ω, such that $Q^k_{x,u} \sim P_{x,u}$.

Let $Q^k = Q^k_{\mathrm{v}_0}$.

To any $f \in \mathscr{B}_b(\mathbb{R}^d)$ we associate $f_k(x, s) = E^{Q^k}_{x,s}[f(X_{t-s})]$ defined for $s \in [0, t]$

If B_k and f are "good", we expect that

(4.10.i) $f_k \in \Lambda$,

(4.10.ii) ∇f_k is a bounded sequence of L^2_v,

(4.10.iii) $A' f_k + (B_k/a \nabla f_k) = 0$ on $\mathbb{R}^d \times [0, t]$.

Applying the (B, Λ)-weak forward equation (4.9.iii), (4.10.i) and (4.10.iii) we thus get

$$(4.11)\quad \int_{\mathbb{R}^d} f_k(x, t)\, d\mathrm{v}_t - \int_{\mathbb{R}^d} f_k(x, 0)\, d\mathrm{v}_0 = \int_0^t \int_{\mathbb{R}^d} [\nabla f_k/a(B - B_k)](x, u)\, d\mathrm{v}_u\, du.$$

But $f_k(x, t) = f(x)$ and $\int_{\mathbb{R}^d} f_k(x, 0)\, d\mathrm{v}_0 = E^{Q^k}[f(X_t)]$, therefore

$$(4.12)\quad E^{Q^k}[f(X_t)] = \int_{\mathbb{R}^d} f(x)\, d\mathrm{v}_t + \int_0^t \int_{\mathbb{R}^d} [\nabla f_k/a(B_k - B)](x, u)\, d\mathrm{v}_u\, du,$$

which yields for a nonnegative "good" f, considering (4.10.ii)

$$(4.13)\quad E^{Q^k}[f(X_t)] \leqq \int_{\mathbb{R}^d} f(x)\, d\mathrm{v}_t + C \| B_k - B \|_{L^2_\mathrm{v}}.$$

If we could prove that Q^k converges weakly to Q, the proof would be finished. Zheng [Zhe] proves in his case that Q^k is tight, using the tightness criterion of [MZ2]. Here we cannot use directly this criterion, but we shall use the Trick (2.6) which is inspired by Zheng's paper, as we already said. Assume that B^k is such that

$$(4.14) \qquad \begin{cases} E^P\left[\int_0^{T_n} (B - B_k/a(B-B_k))(X_s, s)\,ds\right] \\ \quad \text{goes to 0 for each fixed } n. \end{cases}$$

Then we know that for all t,

(4.15) $Z^k_{t \wedge T_n}$ goes P a.s. to $Z_{t \wedge T_n}$ (at least for a subsequence, *see* § 3).

Here of course Z^k_t is the P (super)martingale associated to B_k.

Now we may apply Fatou's lemma to get

$$(4.16) \qquad \begin{aligned} &E^Q[f(X_t)\,1_{t \leqq T_n}] \\ &= E^P[Z_{t \wedge T_n} f(X_t)\,1_{t \leqq T_n}] \leqq \liminf_k E^{Q^k}[f(X_t)\,1_{t \leqq T_n}] \\ &\leqq \liminf_k E^{Q^k}[f(X_t)] \leqq \int_{\mathbb{R}^d} f(x)\,d\nu_t. \end{aligned}$$

By the monotone convergence theorem we obtain for all nonnegative "good" f

$$(4.17) \qquad E^Q[f(X_t)\,1_{t \leqq T_\infty}] \leqq \int_{\mathbb{R}^d} f(x)\,d\nu_t.$$

If the set of "good" functions is large enough, (4.17) extends to any nonnegative measurable f, and we may apply (2.6). □

The point now is to overcome the technical difficulties in order to make rigorous the above sketch of proof. The first result we give, uses regularity results for stochastic flows.

(4.18) THEOREM. – *Assume that* $a = \sigma.\sigma^*$, *and that* σ *and* b *are* $C_b^{2,1,\alpha}$, i.e. *are twice continuously differentiable in* x, *once in* t, *bounded with bounded derivatives, and that furthermore the derivatives satisfy a global Hölder condition of order* α *for some positive* α.

Let $\nu_.$ *be a flow of Probability measures such that* $H(\nu_0, \mu_0) < +\infty$, *and* B *be of finite* ν-*energy. If* ν *satisfies the* (B, C_0^∞)-*weak forward equation, then the* (B, ν_0, P)-Föllmer Measure [*see* (1.13)] *belongs to* $A_{\nu, H}$.

Proof. – Since all coefficients of A′ are bounded, $C_b^{2,1} \subseteqq D_{e,\nu}$. An easy density argument shows that ν satisfies the (B, $C_b^{2,1}$)-weak forward equation.

Let $B^* \in \mathscr{B}_b(\mathbb{R}^d \times \mathbb{R})$ with compact support, and denote by μ_t the flow $\mu_t = P \circ X_t^{-1}$. One can choose a sequence B_k of $C_0^\infty(\mathbb{R}^d \times \mathbb{R})$ such that B_k

converges to B^* in the space $L^2((d\nu + d\mu) \otimes dt)$, *i. e.*

$$\lim_k \int_0^T \int_{\mathbb{R}^d} |B_k - B^*|^2 (x, t)(\nu_t + \mu_t)(dx)\,dt = 0.$$

In particular, since a is bounded, B_k goes to B^* both in L^2_ν and L^2_μ. Define $Q^k_{x,u}$ (resp. $Q^*_{x,u}$) as the Markov family of Probability laws on Ω_t associated to B_k (resp. B^*) as in (4.9 *bis*), denoting by Z^k (resp. Z^*) the associated exponential martingale, and finally define Q^k (resp. Q^*) as the projection of $Q^k_{\nu_0 \otimes \delta_0}$ (resp. $Q^*_{\nu_0 \otimes \delta_0}$) on Ω.

Since $E^P\left[\int_0^T (B^* - B_k / a(B^* - B_k))(X_s, s)\,ds\right]$ goes to 0 by construction, one can find, for each fixed t, a subsequence of Z^k_t which goes P a.s. to Z^*_t.

For $f \in C_0^\infty(\mathbb{R}^d)$, we define

$$f_k(x, s) = E^{Q^k_{x,s}}[f(X_{t-s})] \qquad \text{for} \quad s \in [0, t]. \tag{4.19}$$

With our assumptions on a and b, we know, by means of Theorem 3.4 in [Ku] and the method in [IW, p. 255-259] that $f_k \in C_b^{2,1}(\mathbb{R}^d \times [0, t])$. The boundedness of the derivatives of f_k follows from the support property of f. Indeed, the expectation of the squared norm of the derivatives of the associated stochastic flow are of polynomial growth while $Q^k_{x,u}(\tau \leq T)$ behaves like $\exp(-x^2/cT)$ (uniformly in u) if τ is the first hitting time of the support of f.

Furthermore (*see* [IW, p. 255-259])

(4.20.i) $A' f_k + (B_k / a\, \nabla_x f_x) = 0$ on $\mathbb{R}^d \times [0, t]$,

(4.20.ii) $f_k(x, t) = f(x)$ and $\int_{\mathbb{R}^d} f_k(x, 0)\,d\nu_0 = E^{Q^k}[f(X_t)]$.

In view of (4.10) the main thing we have to prove now, is that the sequence $\nabla_x f_k$ is bounded in L^2_ν, restricting the time interval to $[0, t]$.

Following Mikami's idea, this can be done by applying the $(B, C_b^{2,1})$-weak forward equation to $(f_k)^2$ which again belongs to $C_b^{2,1}$. If we do that, we get

$$\begin{aligned} \int_0^t \int_{\mathbb{R}^d} [(\nabla_x f_k / a\, \nabla_x f_k)](x, u)\,d\nu_u\,du &= \int_{\mathbb{R}^d} (f_k)^2 (x, t)\,d\nu_t - \int_{\mathbb{R}^d} (f_k)^2 (x, 0)\,d\nu_0 \\ &\quad - 2 \int_0^t \int_{\mathbb{R}^d} f_k(x, u)\,[(\nabla_x f_k / a(B - B_k))(x, u)\,d\nu_u\,du, \end{aligned} \tag{4.21}$$

thanks to (4.19). Therefore

$$(\| \nabla_x f_k \|_{L^2_\nu})^2 \leq \| f \|^2_\infty + 2 \| f \|_\infty \| B - B_k \|_{L^2_\nu} \| \nabla_x f_k \|_{L^2_\nu}, \tag{4.22}$$

which yields (by using $x^2 \leqq c^2 + 2\,cdx \Rightarrow |x| \leqq 2|c|(1+d^2)^{1/2}$)

$$(4.23)\quad \|\nabla_x f_k\|_{L^2_\nu} \leqq 2\|f\|_\infty [1 + (\|B - B_k\|_{L^2_\nu})^2]^{1/2} \leqq 2\|f\|_\infty [2 + (\|B - B^*\|_{L^2_\nu})^2]^{1/2}$$

for k large enough. We then deduce as in (4.11)-(4.13) that for any nonnegative smooth f with compact support, and k large enough

$$(4.24)\quad E^{Q^k}[f(X_t)] \leqq \int_{\mathbb{R}^d} f(x)\,d\nu_t + 2\|f\|_\infty [2 + (\|B - B^*\|_{L^2_\nu})^2]^{1/2} (\|B - B^*\|_{L^2_\nu} + \|B^* - B_k\|_{L^2_\nu}).$$

Since Z_t^k goes P a.s. to Z_t^*, Fatou's lemma yields

$$(4.25)\quad E^{Q^*}[f(X_t)] \leqq \int_{\mathbb{R}^d} f(x)\,d\nu_t + 2\|f\|_\infty [2 + (\|B - B^*\|_{L^2_\nu})^2]^{1/2} \|B - B^*\|_{L^2_\nu}.$$

Let Q be the (B, ν_0, P)-FM. We choose a sequence θ_n of $C_0^\infty(\mathbb{R}^d)$ such that $|\theta_n(x)| \leqq 1$ for all x, $\theta_n(x) = 1$ if $|x| < n$ and 0 if $|x| > n+1$. Applying (4.25) to $B^n = B\,1_{|B| \leqq n}\,\theta_n(x)$ which is bounded with compact support, we obtain

$$(4.26)\quad E^{Q^n}[f(X_t)] \leqq \int_{\mathbb{R}^d} f(x)\,d\nu_t + 2\|f\|_\infty [2 + (\|B - B^n\|_{L^2_\nu})^2]^{1/2} \|B - B^n\|_{L^2_\nu}.$$

But $\|B - B^n\|_{L^2_\nu}$ goes to 0, as n tends to infinity, and for each fixed m

$$(4.27)\quad \begin{cases} E^P\left[\displaystyle\int_0^{T_m} (B - B^n / a(B - B^n))(X_s, s)\,ds\right] \\ \text{goes to 0, as } n \text{ tends to infinity.} \end{cases}$$

Thus we can follow (4.14)-(4.17) and obtain

$$(4.28)\quad \begin{cases} E^Q[f(X_t)\,1_{t \leqq T_\infty}] \leqq \displaystyle\int_{\mathbb{R}^d} f(x)\,d\nu_t, \\ \text{for } f \in C_0^\infty(\mathbb{R}^d) \text{ and nonnegative.} \end{cases}$$

Let f be a bounded nonnegative Borel function. One can approximate f both in $L^1(\nu_t)$ and in $L^1(Q \circ X_t^{-1})$ by nonnegative step functions. Thus it suffices to prove that (4.28) extends to $f = 1_A$ for Borel subsets A. Since $\nu_t + Q \circ X_t^{-1}$ is regular, one can restrict our attention to closed A, then to continuous nonnegative f (*see* [Bi]). A truncation argument shows that we may consider only functions f with compact support, which can be uniformly approximated by nonnegative smooth functions. Hence (4.28) extends to any nonnegative Borel function and we may apply the Trick (2.6) to conclude. □

Remark. — The key points in the above derivation are first that f_k is smooth enough, second that $\nabla_x f_k \in L^2_\nu$ (which is immediate in the above proof because a is bounded). Indeed as soon as $\nabla_x f_k \in L^2_\nu$ the derivation (4.21)-(4.23) is available, and so the sequence $\nabla_x f_k$ is bounded in L^2_ν.

If a and b are not anymore bounded or if f_k does not have bounded derivatives, we have to localize in the above proof.

(4.29) THEOREM. – *Assume that* $a=\sigma.\sigma^*$, σ *and* b *are* $C^{2,1,\alpha}$, *with a local Hölder condition. Let* $\nu_{\bullet}$ *be a flow of Probability measures such that* $H(\nu_0, \mu_0) < +\infty$, *and* B *be of finite* ν*-energy. If* ν *satisfies the* (B, C_0^∞)*-weak forward equation, then the* (B, ν_0, P)-Föllmer Measure [*see* (1.13)] *belongs to* $A_{\nu, H}$.

Proof. – We adapt the proof of (4.18).

For $\lambda > 0$ we define

$$h(\lambda) = \sup\left\{ |x| \vee \max_i \left|\frac{\partial \varphi}{\partial x_i}(x)\right| \vee \max_{ij} \left|\frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x)\right|;\ |x| \leqq 1/\lambda \right\},$$

Let χ be a C_0^∞ function such that $0 \leqq \chi(x) \leqq 1$, $\chi(x)=1$ if $|x| \leqq 1/2$ and $\chi(x)=0$ if $|x| \geqq 1$. C stands for an upper bound for the first and the second derivatives of χ. For $\lambda>0$, let us define

$$\psi_\lambda(x) = \chi(\lambda x) \exp - [(\lambda^2/h(\lambda))\,|x|^2].$$

Then ψ_λ is a family of smooth functions with pointwise limit 1 when λ goes to 0. Furthermore

$$(4.30)\qquad \left|\frac{\partial}{\partial x_i}\psi_\lambda(x)\right| \leqq \left(\lambda \left|\frac{\partial}{\partial x_i}\chi(\lambda x)\right| + 2(\lambda^2/h(\lambda))\,|x_i|\,\chi(\lambda x) \exp - [(\lambda^2/h(\lambda))\,|x|^2]\right),$$
$$\leqq [\lambda\, C \exp - (1/4\, h(\lambda)) + 2\lambda/h(\lambda)]\, 1_{|x| \leqq 1/\lambda},$$

since $\dfrac{\partial}{\partial x_i}\chi(\lambda x) = 0$ if $1/2\lambda \geqq |x|$ or $|x| \geqq 1/\lambda$ and $\chi(\lambda x)=0$ if $|x| \geqq 1/\lambda$.

In the same way we get

$$(4.31)\qquad \left|\frac{\partial^2}{\partial x_i \partial x_j}\psi_\lambda(x)\right| \leqq [(C\lambda^2 + 4C(\lambda^2/h(\lambda)))\exp - (1/4\, h(\lambda)) + 4\lambda^2/h^2(\lambda) + 2\lambda^2/h(\lambda)]\, 1_{|x| \leqq 1/\lambda}.$$

Since $h(\lambda)$ goes to infinity when λ goes to 0, it follows that one can find a constant K such that

(4.32.i) $|a(x,t)|\,|\nabla_x \psi_\lambda(x)| \leqq K\lambda$ for all (x, t),
(4.32.ii) $|A' \psi_\lambda(x)| \leqq K\lambda$ for all (x, t),
(4.32.iii) $\|\nabla_x \psi_\lambda\|_{L^2_\nu} \leqq K\lambda$.

Since we assumed that for all (x, u) no explosion occurs up to and including time T, for the initial martingale problem (1.6), the (time-space) diffusion process is strictly conservative [Ku]. If B_k is smooth with compact support, $(Q^k_{x,u})$ remains strictly conservative. Thus, $(Q^k_{x,u})$ is the image law of a $C^{2,1,\beta}$ stochastic flow for $\beta<\alpha$ ([Ku], Thm. 5.4). But we do not know if f_k defined by (4.19) is a $C^{2,1}$ function, because we do not know if we can differentiate under the expectation sign. Hence we must slightly modify f_k. Let σ^λ and b^λ be $C^{2,1}_b$ functions which coincide with σ and b on $\{|x|\leqq 1/\lambda\}$. Define $Q^{k,\lambda}_{x,u}$ as the law of the (time-space) diffusion with diffusion matrix $a^\lambda=\sigma^\lambda(\sigma^\lambda)^*$ and drift $b^\lambda+a^\lambda B_k$ starting from (x, u). For a compactly supported smooth f we define

$$f_{k,\lambda}(x, s)=E^{Q^{k,\lambda}_{x,s}}[f(X_{t-s})] \qquad \text{for} \quad s\in[0, t]. \tag{4.33}$$

Then $f_{k,\lambda}(x, s)$ is a $C^{2,1}_b$ function and satisfies

(4.34.i) $A' f_{k,\lambda}+(B_k/a\nabla_x f_{k,\lambda})=0$ on $\{|x|\leqq 1/\lambda\}\times[0, t]$,

(4.34.ii) $f_{k,\lambda}(x, t)=f(x)$ and

$$\int_{\mathbb{R}^d} f_{k,\lambda}(x, 0)\, d\nu_0=E^{Q^{k,\lambda}}[f(X_t)],$$

since A'^λ and a^λ coincide with A' and a on $\{|x|\leqq 1/\lambda\}\times[0, T]$.

(4.35) For $\lambda=1/k$ we define $f_k=f_{k,\lambda}=f_{k,\,1/k}$.

An easy density argument shows that ν satisfies the $(B, C^{2,1}_0)$-weak forward equation, so that we may apply the weak forward equation to $\psi^2_\lambda f_k$. In virtue of (4.34.i) this yields

$$\begin{aligned}
&\int_{\mathbb{R}^d}\psi^2_\lambda f_k(x, t)\, d\nu_t-\int_{\mathbb{R}^d}\psi^2_\lambda f_k(x, 0)\, d\nu_0\\
&\quad=\int_0^t\int_{\mathbb{R}^d}[(\nabla_x(\psi^2_\lambda f_k)/a(B-B_k))](x, u)\, d\nu_u\, du\\
&\quad\quad+2\int_0^t\int_{\mathbb{R}^d} f_k(x, u)\\
&\quad\quad\times[(\nabla_x\psi^2_\lambda/a B_k)+A'\psi^2_\lambda-2(\nabla_x\psi^2_\lambda/a\nabla_x\psi^2_\lambda)](x, u)\, d\nu_u\, du\\
&\quad\quad+2\int\int_d(\nabla_x\psi_\lambda/a\nabla_x(\psi_\lambda f_k))(x, u)\, d\nu_u\, du,
\end{aligned} \tag{4.36}$$

since the support of ψ^2_λ is included in $\{|x|\leqq 1/\lambda\}$.

Applying the weak forward equation to $\psi^2_\lambda f^2_k$ (resp. $\psi^4_\lambda f^2_k$) we get as in (4.22) and (4.23), considering (4.34.i), and restricting the time interval to $[0, t]$

$$\begin{aligned}
\|\nabla_x(\psi_\lambda f_k)\|_{L^2_\nu}\leqq\|f\|_\infty[1+(\|B-B_k\|_{L^2_\nu})^2+2t\|A'\psi_\lambda\|_\infty\\
+2\|B_k\|_{L^2_\nu}\|\nabla_x\psi_\lambda\|_{L^2_\nu}+4(\|\nabla_x\psi_\lambda\|_{L^2_\nu})^2]^{1/2}
\end{aligned} \tag{4.37}$$

For k large enough, $\|B^*-B_k\|_{L^2_\nu}\leqq 1$ and $\|B_k\|_{L^2_\nu}\leqq 1+\|B^*\|_{L^2_\nu}$. Recall that later on, B^* will be $B\,1_{|B|\leqq n}\theta_n(x)$, so that we may assume that

$\| B^* \|_{L^2_\nu} \leqq \| B \|_{L^2_\nu}$. Finally, using (4.32), we obtain the existence of a constant K' such that for k large enough and $\lambda \leqq 1$,

$$(4.38)\qquad \| \nabla_x (\psi_\lambda f_k) \|_{L^2_\nu} \leqq K' \| f \|_\infty [1 + (\| B - B^* \|_{L^2_\nu})^2 + \| B \|_{L^2_\nu}]^{1/2}.$$

The same upper bound is valid for $\nabla_x (\psi_\lambda^2 f_k)$. Hence $\nabla_x (\psi_\lambda f_k)$ [resp. $\nabla_x (\psi_\lambda^2 f_k)$] is bounded in L^2_ν independently of k and λ. Applying all the above estimates in (4.36) we get for any nonnegative f

$$(4.39)\qquad \int_{\mathbb{R}^d} \psi_\lambda^2 f_k (x, 0)\, d\nu_0 \leqq \int_{\mathbb{R}^d} \psi_\lambda^2 f(x)\, d\nu_t + C_2 (\lambda + \lambda^2) \| f \|_\infty + C_3 \lambda \\ + C_1 [1 + (\| B - B^* \|_{L^2_\nu})^2 + \| B \|_{L^2_\nu}]^{1/2} (\lambda + \| B - B^* \|_{L^2_\nu} + \| B_k - B^* \|_{L^2_\nu}),$$

where C_1, C_2, C_3 only depend on $\| f \|_\infty$ and $\| B \|_{L^2_\nu}$.

Letting λ tend to 0 (hence k tends to ∞), by Lebesgue's bounded convergence theorem and Fatou's lemma, we obtain

$$(4.40)\qquad E^{Q^*} [f(X_t)] \leqq \int_{\mathbb{R}^d} f(x)\, d\nu_t \\ + C [1 + (\| B - B^* \|_{L^2_\nu})^2 + \| B \|_{L^2_\nu}]^{1/2} \| B - B^* \|_{L^2_\nu},$$

where C only depends on $\| f \|_\infty$ and $\| B \|_{L^2_\nu}$.

Indeed, let s_k be the first exit time of the ball of radius k. $Q^{k, 1/k}$ and Q^k are the same in restriction to $\mathscr{F}_{s_k}$. Hence for $k > n$,

$$(4.41)\qquad E^{Q^k} [f(X_t)\, 1_{t \leqq s_n}] = E^{Q^{k, 1/k}} [f(X_t)\, 1_{t \leqq s_n}] \leqq \int_{\mathbb{R}^d} \psi_\lambda^2 f_k (x, 0)\, d\nu_0.$$

(4.40) follows since Z_t^k goes P a.s. to Z_t^* and since s_n goes Q^* a.s. to T. The rest of the proof is identical to the proof of (4.18). □

The next result lies on p.d.e. result as in [DaG].

(4.42) THEOREM. – *Assume that a and b are locally Hölder continuous, and that a is uniformly elliptic. Let* (ν, B) *be as in* (4.29)*; then the* (B, ν_0, P)-*Föllmer Measure* [see (1.13)] *belongs to* $A_{\nu, H}$.

Proof. – Here again the difficulty is that we do not know if f_k in the proof of (4.18) satisfies (4.20). Actually we do not know if there exists a solution of $A' f + (B_k / a \nabla_x f) = 0$ on the whole unbounded space $\mathbb{R}^d \times [0, t]$.

Furthermore, although we think that the modified f_k in the proof of (4.29) again satisfies (4.32), we did not find in the litterature any argument to prove it. Notice that here we do not have anymore the stochastic flow argument. Consequently, once again we have to modify f_k.

Let f be a nonnegative smooth function with compact support. For k such that supp.$f \subset \{ |x| < k \}$, let s_k be the first exit time of the ball of radius k. Define

$$(4.43)\qquad f_k (x, s) = E^{Q^k_{x, s}} [f(X_{(t-s)}\, 1_{(t-s) < s_k}] \qquad \text{for} \quad s \in [0, t].$$

By means of a standard p.d.e. result, it is proved in [DaG] that f_k belongs to $C^{2,1}(\{|x|\leq k\}\times[0,t])$ and satisfies

(4.44.i) $A' f_k + (B_k/a\nabla_x f_k) = 0$ on $\{|x|\leq k\}\times[0,t]$,

(4.44.ii) $f_k(x,t)=f(x)$ on $\{|x|\leq k\}$,

(4.44.iii) $f_k(x,s)=0$ if $|x|=k$.

We can now follow the proof of (4.29) with this new f_k but without modifying Q^k (*i.e.* we do not need $Q^{k,\lambda}$). $\square$

Remark. – 1) We cannot use the f_k of the above proof in the proof of (4.29). Indeed without ellipticity (or more generally hypoellipticity) assumption, the uniqueness of the classical solution of the underlying Cauchy-Dirichlet problem fails and the argument of [DaG] does not hold anymore. Of course one can use Stroock-Varadhan's result [SV] to prove that f_k is a weak solution in the degenerate case, but in that case it is not known wether f_k is regular enough (*i.e.* $C^{2,1}$) or not, except in the hypoelliptic case [Cx].

2) One can remove in (4.42) the local Hölder continuity of b, assuming only that b is continuous. This can be done by approximating b and by using the Girsanov formula. We refer to [DaG] for the details.

In all previous theorems we used the usual differential calculus, and so we were obliged to make some regularity assumptions on the coefficients. In return, the only assumption on ν is the finite entropy condition at time 0. As a consequence, we obtain that if the conditions of Theorems (4.29) or (4.42) are fulfilled, any solution ν of the B-weak forward equation with B of finite ν-energy, satisfies $H(\nu_t,\mu_t)<+\infty$ for almost all t.

The last theorem of this section deals with the other point of view, assuming some conditions on ν but no more restrictions on a and b (except those ones required from the beginning). First we have to introduce some definitons.

(4.45) DEFINITION. – *Let $\nu_.$ be a flow of Probability measures and $B\in L^2_\nu$. We say that ν satisfies the extended* (B, Λ)*-weak forward equation, for $C_0^\infty \subseteqq \Lambda \subseteqq D_{e,\nu}$, if for any $0\leqq s\leqq t\leqq T$ and any* $F\in C^1([s,t],\Lambda)$

$$\int_{\mathbb{R}^d\times\mathbb{R}} F(t,x,u)\,d\nu_t(x)\,\delta_t(u) - \int_{\mathbb{R}^d\times\mathbb{R}} F(s,x,u)\,d\nu_s(x)\,\delta_s(u)$$
$$= \int_s^t \int_{\mathbb{R}^d\times\mathbb{R}} \left[\frac{\partial}{\partial v} + A'(x,u) + (B(x,u)/a(x,u)\nabla)\right] \times F(v,x,u)\,d\nu_v(x)\,\delta_v(u)\,dv.$$

In other words the flow $\nu_t\otimes\delta_t$ satisfies the weak forward equation.

Notice that if $\Lambda=C_0^{2,1}$, $f(x,t)=F(t,x,t)$ belongs to $C_0^{2,1}$ and the above definition is nothing else than the (non extended) weak forward equation.

In fact the problem is that we do not know, in general, if for $F \in C^1([0, T], D_{e, \nu})$ the function $f(x, t) = F(t, x, t)$ belongs to $D_{e, \nu}$.

Of course if ν satisfies the extended weak forward equation, it also satisfies the non extended one. Furthermore, if $A_{\nu, H}$ is non empty, ν satisfies the extended $(B, D_{e, \nu})$-weak forward equation for some B of finite ν-energy, as it can be seen by applying Ito's formula.

The next space will provide a good core.

(4.46) *Notation.* – We denote by Λ_0 the following space

$$\Lambda_0 = \{ f \in D_e, \text{s.t. } f \text{ has compact support, } A' f \in L^1(\mu_t \otimes dt) \\ \text{and } C_t^f \text{ [see (4.2)] is a } L^2\, P_{x, u}\text{-martingale for all } (x, u) \}.$$

(4.47) LEMMA. – *Let* $\nu_{\cdot}$ *be a flow of Probability measures such that* $\nu_t \ll \mu_t$ *for all* $t \in [0, T]$. *Denote by* $\rho(x, t)$ *the density* $\dfrac{\partial \nu_t}{\partial \mu_t}$. *Assume that* $\rho \in \mathscr{B}_{b, \text{loc}}(\mathbb{R}^d \times \mathbb{R})$. *Then* $\Lambda_0 \subseteqq D_{e, \nu}$.

Proof. – It is easy to see that for f in Λ_0, one can find versions of $A' f$ and of ∇f which are compactly supported. The proof is now obvious since $\nabla f \in L^2_\mu$, $A' f \in L^1(\mu_t \otimes dt)$ and have compact supports and $\rho \in \mathscr{B}_{b, \text{loc}}(\mathbb{R}^d \times \mathbb{R})$. □

This leads us to the last result of the section.

(4.48) THEOREM. – *Let* $\nu_{\cdot}$ *be a flow of Probability measures such that* $\nu_t \ll \mu_t$ *for all* $t \in [0, T]$. *Denote by* $\rho(x, t)$ *the density* $\dfrac{\partial \nu_t}{\partial \mu_t}$. *Let* B *be of finite* ν-*energy*.

Assume that $\rho \in \mathscr{B}_{b, \text{loc}}(\mathbb{R}^d \times \mathbb{R})$, *that* $H(\nu_0, \mu_0) < +\infty$ *and that* ν *satisfies the extended* (B, Λ_0)-*weak forward equation.*

Then the (B, ν_0, P)-*Föllmer Measure* [see (1.13)] *belongs to* $A_{\nu, H}$.

Proof. – Let B^* be a bounded Borel function with compact support. The associated $Q^*_{x, u}$ is then a family of Probability measures. Let $Q^* = Q^*_{\nu_0 \otimes \delta_0}$. The associated semigroup $Q^*_t f(x, u) = E^{Q^*_{x, u}}[f(X_t, u_t)]$, defined on $\mathscr{B}_p(\mathbb{R}^d \times \mathbb{R})$ is then strongly continuous on the closed subspace

$$\mathscr{C}(B^*) = \{ f \in \mathscr{B}_b(\mathbb{R}^d \times \mathbb{R}),\ \| Q^*_t f - f \|_\infty \to 0 \text{ when } t \to 0 \}$$

with generator $A(B^*)$ and domain $D(B^*)$. $C_0^\infty(\mathbb{R}^d \times \mathbb{R})$ is included in $\mathscr{C}(B^*)$, and $A(B^*)$ coincides with $A' + (B^*/a \nabla_x)$ on $C_0^\infty(\mathbb{R}^d \times \mathbb{R}) \cap D(B^*)$.

Let $f \in D(B^*)$. Then the function F

$$F(s, x, u) = E^{Q^*_{x, u}}[f(X_{t-s}, u_{t-s})]$$

belongs to $C^1([0, t], D(B^*))$ and satisfies

$$\text{(4.49.i)} \quad \left(\frac{\partial}{\partial s} + A(B^*) \right) F = 0 \text{ on } [0, t] \times \mathbb{R}^d \times \mathbb{R},$$

$$\text{(4.49.ii)} \quad F(t, x, u) = f(x, u);\ F(0, x, u) = E^{Q^*_{x, u}}[f(X_t, u_t)].$$

All these facts are well known in semigroup theory (*see* [Ta]). We want to apply the extended weak forward equation. To this end, we need the following lemma.

(4.50) LEMMA. – *For all* $\psi \in C_0^\infty(\mathbb{R}^d \times \mathbb{R})$ *and* $g \in D(B^*)$, $\psi . g$ *belongs to* Λ_0.

The proof of the lemma is postponed to the end of the proof theorem (4.43). Let $G \in C^1([0, t], \Lambda_0)$. Then $G^2 \in C^1([0, t], \Lambda_0)$. Indeed using the change of variables formula (Ito's formula), $G^2(s)$ belongs to D_e and $\nabla(G^2)(s) = 2G.\nabla G(s)$. Hence we can apply the extended (B, Λ_0)-weak forward equation to $(\psi_\lambda F)^2$ (where ψ_λ is as in the proof of (4.29)) and get as in (4.37), that for any nonnegative f

$$(4.51)\quad \left[\int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} [(\nabla(\psi_\lambda F)/a\nabla(\psi_\lambda F))](s, x, u)\, d\nu_s(x)\, \delta_s(u)\, ds\right]^{1/2} \\ K' \| f \|_\infty [1 + (\| B - B^* \|_{L^2_\nu})^2 + 2t \| A' \psi_\lambda \|_\infty \\ + 2 \| B^* \|_{L^2_\nu} \| \nabla_x \psi_\lambda \|_{L^2_\nu} + 4(\| \nabla_x \psi_\lambda \|_{L^2_\nu})^2]^{1/2}$$

for some constant K'. This yields as in (4.40) ($\lambda \to 0$)

$$(4.52)\quad E^{Q^*}[f(X_t, t)] \leqq \int_{\mathbb{R}^d} f(x, t)\, d\nu_t \\ + C[1 + (\| B - B^* \|_{L^2_\nu})^2 + \| B \|_{L^2_\nu}]^{1/2} \| B - B^* \|_{L^2_\nu}$$

for any nonnegative f in $D(B^*)$.

Recall that if $f \in \mathscr{C}(B^*)$, $f_\varepsilon = \varepsilon^{-1} \int_0^\varepsilon Q_s^* f\, ds$ belongs to $D(B^*)$ and converges to f in the uniform topology. In particular, if f is nonnegative, f_ε is also nonnegative. So (4.52) extends to $\mathscr{C}(B^*)$, and in particular to C_0^∞. The rest of the proof is now similar to the proof of (4.18).

It remains to prove Lemma (4.50).

Let $g \in D(B^*)$. Then

$$C_t^*(g) = g(X_t, u_t) - g(X_0, u_0) - \int_0^t A(B^*) g(X_s, u_s)\, ds$$

is a bounded $Q^*_{x, u}$-martingale. As in (4.3), one can find a Borel function $\nabla^* g$ such that $C_t^*(g) = \int_0^t (\nabla^* g(X_s, u_s)/dN_s^*)$ for all t, $Q^*_{x, u}$ a.s. for all (x, u), where $N_t^* = M_t - \int_0^t a B^*(X_s, u_s)\, ds$. Since $H(P_{x, u}, Q^*_{x, u}) < +\infty$ (B^* is bounded with compact support), we can apply the Girsanov transform and Proposition (2.7) in order to get

(4.53) $C_t^*(g)-\int_0^t (\nabla^* g/a\,B^*)(X_s, u_s)\,ds$ is a $L^2 P_{x,u}$-martingale for all (x, u), with brackets $\int_0^t (\nabla^* g/a\,\nabla^* g)(X_s, u_s)\,ds$.

Thus g belongs to D_e, $\nabla g=\nabla^* g$ and $A' g=A(B^*)g-(B^*/a\nabla g)$.

Furthermore, according to Remark (2.8.1),

$$\|\nabla g\|_{L^2_\mu}\leqq C\|g\|_\infty (1+H(P, Q^*_{\mu_0\otimes\delta_0}))^{1/2}.$$

So ψg has compact support, $\|\nabla(\psi g)\|_{L^2_\mu}$ is finite and $A'(\psi g)$ belongs to $L^1(d\mu_t\otimes dt)$. □

Remark that we only used the local boundedness of ρ to prove that Λ_0 is included in $D_{e,\nu}$ via Hölder's inequality for the pair (L^1, L^∞) in (4.47). In the next section we shall see another existence result, assuming that $A_{\nu,H}$ is non empty, as well as applications to the large deviation problem.

5. ENTROPY MINIMIZATION AND RELATED TOPICS

In the preceding section we gave some sufficient conditions for $A_{\nu,H}$ not to be empty. In this one we shall first describe $A_{\nu,H}$, then study the minimization problem and finally apply all the results to the initial large deviation problem (*see* § 0).

(5.1) DEFINITION. – *Let* ν *be a flow of Probability measures and* $C_0^\infty\subseteqq\Lambda\subseteqq D_{e,\nu}$. *We define* $H^{-1}(\Lambda,\nu)$ *as the* L^2_ν *closure of the set* $\{\nabla f, f\in\Lambda\}$.

If $\Lambda=D_{e,\nu}$ (*resp.* C_0^∞) *we put* $H^{-1}(\Lambda,\nu)=H^{-1}(\nu)$ (resp. $H_0^{-1}(\nu)$).

Of course $H_0^{-1}(\nu)\subseteqq H^{-1}(\nu)\subseteqq L^2_\nu$. In general $H_0^{-1}(\nu)$ is not equal to the whole L^2_ν. Take for instance

$$d=2,\qquad A=\frac{1}{2}\Delta,\qquad \nu_t(dx)=(2\pi t)^{-1}\exp-(|x|^2/2t)\,dx.$$

It is then easy to check that $B(x, t)=(x_2, -x_1)$ belongs to $(H_0^{-1}(\nu))^\perp$, where $\perp$ denotes the orthogonal set for the L^2_ν norm.

A particular case is the one dimensional case. Indeed we can state

(5.2) PROPOSITION. – *For* $d=1$, $H_0^{-1}(\nu)=L^2_\nu$, *provided that a is bounded.*

Proof. – If $B\in(H_0^{-1}(\nu))^\perp$, then for Lebesgue almost all t in $[0, T]$ the derivative (in the sense of Schwartz distributions) of $a(x)B(x, t)\nu_t$ is equal to 0. Thus this distribution is actually a constant function λ_t. But a being bounded, $1_{[u,t]}\in L^2_\nu$ for all $0\leqq u\leqq t\leqq T$, which implies that $\lambda_t=0$ for

Lebesgue almost all t, *i.e.* $\int_0^T \int_R f(x, t)\, a(x)\, B(x, t)\, d\nu_t\, dt = 0$ *for any test function* f. By approximation this extends first to any bounded continuous f with compact support, and then to any bounded continuous f by Lebesgue's theorem. This proves that the bounded signed measure $a(x)\, B(x, t)\, d\nu_t$ is equal to 0 and finally that $B=0$ in L^2_ν. $\square$

The next result gives a precise description of the set of Markov Probability measures in $A_{\nu, H}$.

(5.3) THEOREM. – *Assume that* $A_{\nu, H}$ *is non void. Then, there exists* $B^* \in H^{-1}(\nu)$ *such that*

(i) *for any Markov Probability measure* Q *in* $A_{\nu, H}$, *the associated drift* B *verifies* $(B-B^*) \in (H^{-1}(\nu))^\perp$,

(ii) *conversely for any* $B = B^* + B^\perp$ *with* $B^\perp \in (H^{-1}(\nu))^\perp$, *the* (B, ν_0, P)-FM Q *belongs to* $A_{\nu, H}$.

So there is a biunivoque correspondence between the set of Markov elements of $A_{\nu, H}$ *and* $(H^{-1}(\nu))^\perp$.

We denote by Q^* *the* (B^*, ν_0, P)-FM, *which belongs to* $A_{\nu, H}$ *thanks to* (ii).

Proof. – (i) Consider two Markovian Probabilities Q_1 and Q_2 of $A_{\nu, H}$, with their drifts B_1 and B_2 (3.34). ν satisfies the $(B_i, D_{e, \nu})$-weak forward equation [*see* (4.6) and (4.7)], so that for all f in $D_{e, \nu}$

$$\int_0^T \int_{R^d} (\nabla f(x, s)/a(x, s) \{ B_1(x, s) - B_2(x, s) \})\, \nu_s(dx)\, ds = 0.$$

Thus $B_1 - B_2$ belongs to $(H^{-1}(\nu))^\perp$, in other words B_1 and B_2 have the same projection on $H^{-1}(\nu)$. We denote this projection by B^*.

(ii) Let Q be a Markov element of $A_{\nu, H}$ [*see* (3.1)], B its associated drift, $B^\perp$ any element in $(H^{-1}(\nu))^\perp$. Then ν satisfies the $(B+B^\perp, D_{e, \nu})$-extended weak forward equation.

In particular if the hypotheses of one of the Theorems (4.18), (4.29) or (4.48) are satisfied, the proof of (ii) is finished.

We shall see that (ii) is true in full generality.

Indeed consider

$$B_k = (B\, 1_{|B| \leq k} + B^\perp\, 1_{|B^\perp| \leq k})\, \theta_k$$

where θ_k is as in the proof of (4.18), and define $Q^k_{x, u}$ as the Markov family of Probability laws associated with B_k. Q^k starts again from $\nu_0 \otimes \delta_0$.

We adopt the notation of the proof of (4.48), in particular, to any $f \in D(B_k)$ one associates $F(s, x, u)$. We should follow the proof of (4.48) replacing Λ_0 by $D_{e, \nu}$, provided $F \in C^1([0, T], D_{e, \nu})$. But, we cannot prove the last assertion. Indeed, $F \in C^1([0, T], D_e)$ but we do not know wether, for each fixed s, $\nabla F(s, ., .)$ belongs to L^2_ν or not. Fortunately, as we shall see later, $\nabla F(s, x, s) \in L^2_\nu$. Looking at the weak forward equation, it is easily seen that the latest weaker property is exactly what is needed for the proof.

Let Z be the density process of Q with respect to P. Define

$$T_n = \inf\left\{ t \geqq 0, \int_0^t (B(x, s)/a(x, s)\, B(x, s))\, ds \geqq n \right\} \wedge T,$$

and

$$Q_n = Z_{T_n} P.$$

Then $Q^k \sim Q_n$ and

$$\begin{aligned} 2H(Q_n, Q^k) &= E^{Q_n}\left[\int_0^T (B_k - B\, 1_{s \leqq T_n}/a(B_k - B\, 1_{s \leqq T_n}))(X_s, s)\, ds\right], \\ &\leqq 2E^{Q_n}\left[\int_0^{T_n} (B/aB)(X_s, s)\, ds\right] + 4k\, E^{Q_n}\left[\int_0^T (\theta_k/a\, \theta_k)(X_s, s)\, ds\right], \\ &\leqq 2E^{Q}\left[\int_0^{T_n} (B/aB)(X_s, s)\, ds\right] + 4kT \sup_{|x| \leqq k+1} |a(x, s)|, \\ &\leqq 2(\|B\|_{L^2_\nu})^2 + 4kT \sup_{|x| \leqq k+1} |a(x, s)|. \end{aligned}$$

Hence

$$E^{Q^n}\left[\int_0^t (\nabla F/a \nabla F)(s, X_s, s)\, ds\right] \leqq C(f)(1 + (\|B\|_{L^2_\nu})^2 + 2kT \sup_{|x| \leqq k+1} |a(x, s)|),$$

thanks to (2.8). But

$$\begin{aligned} \int_0^t \int_{\mathbb{R}^d} (\nabla F/a \nabla F)(s, x, s)\, d\nu_s\, ds &= E^{Q}\left[\int_0^t (\nabla F/a \nabla F)(s, X_s, s)\, ds\right] \\ &\leqq \liminf E^{Q^n}\left[\int_0^t (\nabla F/a \nabla F)(s, X_s, s)\, ds\right] \leqq C(f, k, B) < +\infty. \end{aligned}$$

Hence $\nabla F(s, x, s) \in L^2_\nu$. It remains to follow the proof of (4.48) to conclude. Since $B - B^* \in (H^{-1}(\nu))^\perp$, the proof of (ii) is complete. □

(5.4) COROLLARY. – *Let* $A_{\nu, H}$ *be non void.*

(i) *Assume that either* $a = \sigma . \sigma^*$ *with* σ *and* b *in* $C^{2, 1, \alpha}$ *for some* $\alpha > 0$ *or* a *and* b *are locally Hölder continuous and* a *is uniformly elliptic.*

Then $H_0^{-1}(\nu) = H^{-1}(\nu)$.

(ii) *Assume that* $\dfrac{\partial \nu_t}{\partial \mu_t} = \rho(t, .)$ *and that* $\rho(., .) \in \mathscr{B}_{b, \text{loc}}$.

Then $H^{-1}(\Lambda_0, \nu) = H^{-1}(\nu)$.

Proof. – Define B'^* the projection of B^* onto $H_0^{-1}(\nu)$ [resp. $H^{-1}(\Lambda_0, \nu)$]. Since $\nu_.$ satisfies the (extended) (B'^*, C_0^∞) [resp. (B'^*, Λ_0)]-weak forward equation, we may apply Theorem (4.29) [resp. (4.48)] and get that Q'^* the (B'^*, ν_0, P)-FM belongs to $A_{\nu, H}$. Thus $B^* = B'^*$.

Furthermore if $B^{\perp} \in (H_0^{-1}(\nu))^{\perp}$ [resp. $(H^{-1}(\Lambda_0, \nu))^{\perp}$] the same argument shows that $B^* + B^{\perp}$ is an admissible drift, which proves in virtue of (5.3), that $B^{\perp} \in (H^{-1}(\nu))^{\perp}$. □

(5.5) *Remark.* – *If* $A_{\nu, H}$ *is not empty*, $d=1$ *and a is bounded, there exists a unique Markovian Probability measure in* $A_{\nu, H}$. [Recall (5.2)].

Remark. – The proof of Theorem 5.3 can be adapted in order to give another proof of Theorem 3.1. Indeed let $Q \in A_{\nu, H}$ with drift β and Markovian version B. Then ν satisfies the $(B, D_{e, \nu})$-weak forward equation, and B is of finite ν-energy. Take $B_k = B\, 1_{|B| \leqq k}\, \theta_k$, Q^k as in (5.3) and Q_n the stopped Q at time T_n with T_n as in (3.5.i). Then $H(Q_n, Q^k) \leqq 2 H(Q, P)$. The rest of the proof of (5.3) is unchanged and the (B, ν_0, P)-FM belongs to $A_{\nu, H}$. With the material of Section 4, this proof is of course much more simple that the one given at Section 3, but it requires some analytical tools.

(5.6) PROPOSITION. – *All the Markovian elements in* $A_{\nu, H}$ *are equivalent. All the elements in* $A_{\nu, H}$ *are absolutely continuous with respect to* Q^*.

Proof. – To any Q in $A_{\nu, H}$ we may associate a Markov version $\bar{Q}$ builded in (3.1), and we know that $Q \ll \bar{Q}$ (*see* § 3). So it is enough to prove that all the Markov elements of $A_{\nu, H}$ are equivalent. Let Q be such an element associated to B. Define Q' as the $(B^* - B, \nu_0, Q)$-FM. We claim that $Q' = Q^*$.

Indeed, as in the discussion preceding the proof of (3.1) {precisely the discussion below (3.9)} we know that for all nonnegative bounded f

$$E^{Q'}[f(X_t)\, 1_{t<\xi}] \leqq E^{Q^*}[f(X_t)]$$

so that

$$E^{Q'}\left[\int_0^T (B^* - B/a(B^* - B))(X_s, s)\, ds\right] \leqq E^{Q^*}\left[\int_0^T (B^* - B/a(B^* - B))(X_s, s)\, ds\right] < +\infty,$$

noticing that $T = T_\infty$ Q' a.s. This implies that Q' is a Probability measure on Ω and so is equal to Q^*. Furthermore Q' (hence Q^*) and Q are equivalent. □

Remark (due to the referee). – In general all the measures of $A_{\nu, H}$ are not equivalent. Here is a nice counterexample due to the referee.

Let P be the Wiener measure, $a > 0$ and $A = \{\omega / \|\omega\|_\infty > a\}$. The measure $Q = P[\,.\,/A]$ satisfies $H(Q, P) = -\log P(A) < +\infty$. Furthermore Q minimizes $H(., P)$ among all Probability measures R with $R(A) = 1$, and the minimizing measure is unique. It can be shown that Q is not Markov (Exercise). According to 3.1, there exists a Markov measure $\bar{Q}$ with the same marginals as Q and $H(\bar{Q}, P) \leqq H(Q, P)$. $\bar{Q}(A)$ cannot be equal to 1, otherwise $H(\bar{Q}, P) > H(Q, P)$. So $\bar{Q}$ and Q are not equivalent. We know that $Q \ll \bar{Q}$. This also provides us with an example where $\bar{Q}''$ defined in Section 3 is not a Probability measure on Ω.

We conclude this section by solving the minimization problem. Define

$$(5.7) \qquad I(\nu) = \inf \{ H(Q, P),\ Q \in A_{\nu, H} \},$$

and for $C_0^\infty \subseteqq \Lambda \subseteqq D_{e, \nu}$,

$$(5.8) \quad J(\nu, \Lambda) = \sup_{f \in \Lambda} \left\{ \int_{\mathbb{R}^d} f(x, T)\, d\nu_T - \int_{\mathbb{R}^d} f(x, 0)\, d\nu_0 - \int_0^T \int_{\mathbb{R}^d} A' f(x, t)\, d\nu_t\, dt - \frac{1}{2} (\| \nabla f \|_{L^2_\nu})^2 \right\}$$

(5.9) THEOREM. – *Let* $(\nu_t)_{t \in [0, T]}$ *be a flow of Probability measures on* $\mathbb{R}^d$, *such that* $t \to \nu_t$ *is continuous from* $[0, T]$ *into* $\mathscr{D}'(\mathbb{R}^d)$. *Then*

1) $I(\nu) \geqq H(\nu_0, \mu_0) + J(\nu, D_{e, \nu})$.

2) *If* $I(\nu) < +\infty$ *then* $I(\nu) = H(\nu_0, \mu_0) + J(\nu, D_{e, \nu})$. *Furthermore the infimum in* (5.7) *is attained at the only point* Q^*, *and for all* Q *in* $A_{\nu, H}$ *it holds*

$$H(Q, P) = H(Q^*, P) + H(Q, Q^*).$$

3) *Conversely if one of the following assumptions is fulfilled*

(i) $a = \sigma . \sigma^*$ *with* σ *and* b *in* $C^{2, 1, \alpha}$ *for some* $\alpha > 0$,

(ii) *a and b are locally Hölder continuous and a is uniformly elliptic, then* $H(\nu_0, \mu_0) + J(\nu, C_0^\infty) = I(\nu) = H(\nu_0, \mu_0) + J(\nu, D_{e, \nu})$.

Proof. – 1) and 2). If $I(\nu) = +\infty$ there is nothing to prove, thus assume that $I(\nu)$ is finite or equivalently that $A_{\nu, H}$ is non void. To any Q in $A_{\nu, H}$ we associate the Markov version B of its drift, and the associated Markov $\bar{Q}$. Then for any $f \in D_{e, \nu}$

$$(5.10) \quad \frac{1}{2} (\| B \|_{L^2_\nu})^2 = \frac{1}{2} (\| B - \nabla f \|_{L^2_\nu})^2 - \frac{1}{2} (\| \nabla f \|_{L^2_\nu})^2 + \langle B, \nabla f \rangle_{L^2_\nu}$$

so that

$$(5.11) \qquad \frac{1}{2} (\| B \|_{L^2_\nu})^2 \geqq -\frac{1}{2} (\| \nabla f \|_{L^2_\nu})^2 + \langle B, \nabla f \rangle_{L^2_\nu}.$$

But since ν satisfies the $(B, D_{e, \nu})$-weak forward equation, it holds

$$\int_{\mathbb{R}^d} f(x, T)\, d\nu_T - \int_{\mathbb{R}^d} f(x, 0)\, d\nu_0 - \int_0^T \int_{\mathbb{R}^d} A' f(x, t)\, d\nu_t\, dt = \langle B, \nabla f \rangle_{L^2_\nu}.$$

It follows

$$H(Q, P) \geqq H(\nu_0, \mu_0) + \int_{\mathbb{R}^d} f(x, T)\, d\nu_T - \int_{\mathbb{R}^d} f(x, 0)\, d\nu_0 - \int_0^T \int_{\mathbb{R}^d} A' f(x, t)\, d\nu_t\, dt - \frac{1}{2}(\|\nabla f\|_{L^2_\nu})^2,$$

so that taking the infimum in Q on the left hand side and the supremum in f on the right hand side, we have proved that $I(\nu) \geqq H(\nu_0, \mu_0) + J(\nu, D_{e,\nu})$. Furthermore, since $B^* \in H^{-1}(\nu)$, (5.10) leads

$$(5.12)\quad \frac{1}{2}(\|B^*\|_{L^2_\nu})^2 \leqq \sup\left\{-\frac{1}{2}(\|\nabla f\|_{L^2_\nu})^2 + \langle B^*, \nabla f\rangle_{L^2_\nu}, f \in D_{e,\nu}\right\},$$

and finally

$$(5.13)\qquad I(\nu) \leqq H(Q^*, P) \leqq H(\nu_0, \mu_0) + J(\nu, D_{e,\nu}).$$

According to (5.6) $Q \ll Q^*$ so that

$$H(Q, P) = H(Q, Q^*) + E^Q\left[\log\left(\frac{dQ^*}{dP}\right)\right]$$

If at least one term in the right hand side makes sense. Actually both terms are well defined since $B^* \in L^2_\nu$. It is easy to see that

$$\begin{aligned} E^Q\left[\log\left(\frac{dQ^*}{dP}\right)\right] &= -\frac{1}{2}E^Q\left[\int_0^T (B^*/a B^*)\, ds\right] \\ &\quad + E^Q\left[\int_0^T (B^*/a\, \beta_s)\, ds\right] + H(\nu_0, \mu_0), \\ &= -\frac{1}{2}E^Q\left[\int_0^T (B^*/a B^*)\, ds\right] \\ &\quad + E^Q\left[\int_0^T (B^*/a B)\, ds\right] + H(\nu_0, \mu_0), \\ &= \frac{1}{2}\int_0^T \int (B^*/a B^*)(x, s)\, d\nu_s\, ds] + H(\nu_0, \mu_0) = H(Q^*, P), \end{aligned}$$

since $B - B^*$ and B^* are orthogonal in L^2_ν.

Conversely, assume that $J(\nu, C_0^\infty)$ is finite. Then there exists a constant C such that for any $f \in C_0^\infty$,

$$(5.14)\quad \mathscr{L}(f) = \int_{\mathbb{R}^d} f(x, T)\, d\nu_T - \int_{\mathbb{R}^d} f(x, 0)\, d\nu_0 - \int_0^T \int_{\mathbb{R}^d} A' f(x, t)\, d\nu_t\, dt \leqq C + \frac{1}{2}(\|\nabla f\|_{L^2_\nu})^2.$$

Since $\mathscr{L}$ is linear, (5.14) implies that $\mathscr{L}$ is a continuous linear form on C_0^∞ equipped with the seminorm $N(f) = \|\nabla f\|_{L^2_\nu}$. The completion of

$(C_0^\infty/N^{-1}(0))$ is isomorphic to $H_0^{-1}(\nu)$, so that the quotient map of $\mathscr{L}$ [which exists since $\mathscr{L}(f)$ is equal to 0 if $N(f)=0$] can be extended as a continuous linear form on the Hilbert space $H_0^{-1}(\nu)$. Hence, by Riesz representation theorem, there exists a B in $H_0^{-1}(\nu)$ such that for all $f\in C_0^\infty$

$$(5.15)\quad \int_{\mathbb{R}^d} f(x, T)\,d\nu_T - \int_{\mathbb{R}^d} f(x, 0)\,d\nu_0 - \int_0^T \int_{\mathbb{R}^d} A' f(x, t)\,d\nu_t\,dt = \int_0^T \int_{\mathbb{R}^d} (\nabla f/a\,B)(x, t)\,d\nu_t\,dt$$

We shall prove that (5.15) implies that ν satisfies the (B, C_0^∞)-weak forward equation. Indeed, for a fixed $f\in C_0^\infty$ consider the function

$$(5.16)\qquad \gamma(s, f) = \int_{\mathbb{R}^d} f(x, s)\,d\nu_s.$$

From the continuity hypothesis on the flow ν, $\gamma(., f)$ is continuous.

Now choose a sequence ψ_n of $C^\infty([0, T])$ such that $0\leqq\psi_n\leqq 1$, ψ_n is pointwise convergent with limit $1_{[0,\, t]}$ and $(-\psi_n')$, considered as a measure, is weakly convergent to δ_t, the Dirac measure at point t. Such a sequence exists. We may apply (5.15) to $\psi_n f$ which yields

$$(5.17)\quad \int_0^T \psi_n'(s)\,\gamma(s, f)\,ds = \int_{\mathbb{R}^d} \psi_n(T) f(x, T)\,d\nu_T - \int_{\mathbb{R}^d} \psi_n(0) f(x, 0)\,d\nu_0 - \int_0^T \int_{\mathbb{R}^d} \psi_n(s)\,[A' f + (\nabla f/a\,B)](x, s)\,d\nu_s\,ds.$$

We can pass to the limit in n, thanks to Lebesgue theorem in the right hand side and to the continuity property of γ in the left hand side. This proves that ν satisfies the weak forward equation.

Because of our hypotheses, we know that the (B, ν_0, P)-FMQ belongs to $A_{\nu,\,H}$. But

$$J(\nu, C_0^\infty) = \sup\left\{\langle B, \nabla f\rangle_{L^2_\nu} - \frac{1}{2}(\|\nabla f\|_{L^2_\nu})^2, f\in C_0^\infty\right\} \geqq \frac{1}{2}(\|B\|_{L^2_\nu})^2$$

since $B\in H_0^{-1}(\nu)$, and

$$H(Q, P) = \frac{1}{2}(\|B\|_{L^2_\nu})^2 \geqq I(\nu) - H(\nu_0, \mu_0) \geqq J(\nu, D_{e,\nu}) \geqq J(\nu, C_0^\infty). \quad \square$$

We also may state a slightly different but similar result

(5.18) THEOREM. — *Define* $J'(\nu, \Lambda)$ *as*

$$J'(\nu,\Lambda)=\sup\left\{\int_{\mathbb{R}^d\times\mathbb{R}} F(T,x,u)\,d\nu_T(x)\,\delta_T(u) - \int_{\mathbb{R}^d\times\mathbb{R}} F(0,x,u)\,d\nu_0(x)\,\delta_0(u) - \int_0^T\int_{\mathbb{R}^d\times\mathbb{R}}\left[\frac{\partial}{\partial t}+A'\right]F(t,x,u)\,d\nu_t(x)\,\delta_t(u)\,dt - \frac{1}{2}(\|\nabla F\|_{L^2_\nu})^2,\ F\in C^1([0,T],\Lambda)\right\}.$$

Assume that $\dfrac{\partial\nu}{\partial\mu_t}=\rho(t,\,.)$ *and that* $\rho(.,,.)\in\mathscr{B}_{b,\,\mathrm{loc}}$.

Then $I(\nu)=H(\nu_0,\mu_0)+J'(\nu,\Lambda_0)$.

The proof of (5.18) is the same as the previous one except for some changes in the notations. It is left to the reader.

Comments. — The first half of the proof of (5.9) [proofs of 1) and 2)] is similar to the exposition of [Fö1] in the case of Brownian motion. The Pythagoras equality is already known in great generality [Cs].

The ideas of the second half of the proof are partly contained in [DaG], and essentially contained in the Appendix of [DS1]. But, at our level of understanding, the derivation of the result in [DS1] is not completely rigorous. We shall discuss briefly some points and to this end we use the notation of the Appendix of [DS1].

The first point is the following. In order to build the diffusion with additional drift γ_k^φ (which satisfies the good required condition we stated before), the authors call upon Carlen's theorem. As we already said, this result is proved only for the Brownian motion in [Ca]. Also notice that we cannot use Carlen's result to build a Brownian motion with drift and then take its image by the flow of the diffusion, because the energy condition and the weak forward equation are satisfied for the flow of marginals of the perturbed diffusion and not of the perturbed Brownian motion. Thus, it seems that the results of Section 4 in the present paper are necessary to get [DS1]'s result. Furthermore uniqueness of the solution of the weak forward equation (with the given γ_k^φ) does not hold in general. This argument is thus unsufficient to prove that $\mathscr{W}^\varphi$ has the good marginals.

REFERENCES

[AN] R. AEBI and M. NAGASAWA, Large Deviations and the Propagation of Chaos for Schrödinger Processes, *Prob. Th. and Rel. Fields*, Vol. **94**, 1992, pp. 53-68.

[Bi] P. BILLINGSLEY, *Convergence of Probability Measures*, Wiley, New York, 1968.

[Bru] M. BRUNAUD, *Thèse*, Université Paris-XI, 1989.

[Ca] E. CARLEN, Conservative Diffusions, *Comm. Math. Phys.*, Vol. **94**, 1984, pp. 293-316.

[Cx] P. CATTIAUX, Calcul stochastique et opérateurs dégénérés du second ordre, part II, Problème de Dirichlet, *Bull. Sc. Math.*, Vol. **115**, 1991, pp. 81-122.

[CL1] P. CATTIAUX and C. LÉONARD, Minimization of the Kullback Information for Some Markov Processes (in preparation).

[CL2] P. CATTIAUX, C. LÉONARD, *Large Deviations and Nelson Processes*, Preprint, 1992.

[CP] P. CATTIAUX and F. PETIT, *Nelson Processes Revisited* (in preparation).

[Cs] I. CSISZAR, I-Divergence Geometry of Probability Distributions and Minimization Problems, *Ann. of Prob.*, Vol. **3**, 1984, pp. 146-158.

[DaG] D. DAWSON and J. GÄRTNER, Large Deviations from the MacKean Vlasov Limit for Weakly Interacting Diffusions, *Stochastics*, Vol. **20**, 1987, pp. 247-308.

[DM4] C. DELLACHERIE and P. A. MEYER, *Probabilités et potentiel*, chp. XII-XVI, Hermann, 1987.

[DS] J. D. DEUSCHEL and D. W. STROOCK, *Large deviations*, Academic Press, 1989.

[DS1] J. D. DEUSCHEL and D. W. STROOCK, A Function Space Large Deviation Principle for Certain Stochastic Integrals, *Probab. Th. Rel. Fields*, Vol. **83**, 1989, pp. 279-307.

[Fö] H. FÖLLMER, The Exit Measure of a Supermartingale, *Z.W.*, Vol. **21**, 1972, pp. 154-166.

[Fö1] H. FÖLLMER, Random Fields and Diffusion Processes, École d'été de Probabilités de St Flour, *L.N. Math.*, Vol. **1362,** 1988, pp. 101-204.

[Fö2] H. FÖLLMER, An Entropy Approach to the Time Reversal of Diffusion Processes, *Proc.* 4th *IFIP Workshop on Stoch. Diff. Eq.*, M. MÉTIVIER and E. PARDOUX Ed., L.N. in Control and Information, 1987.

[IW] N. IKEDA and S. WATANABE, *Stochastic Differential Equations and Diffusion Processes*, North Holland, 1981.

[Jac] J. JACOD, Problèmes de martingales, *L.N. Math.*, 714, 1979, Springer.

[Ku] H. KUNITA, Stochastic Differential Equations and Stochastic Flows..., École d'été de Probabilités de St Flour, *L.N. Math.*, Vol. **907**, 1984, pp. 144-305.

[Léo] C. LÉONARD, *Thèse d'État*, Université Paris-XI, 1991.

[MZ1] P. A. MEYER and W. A. ZHENG, Construction de processus de Nelson réversibles, Sem. Probab. XIX, *L.N. Math.*, 1123, 1985, pp. 12-26.

[MZ2] P. A. MEYER and W. A. ZHENG, Tightness Criteria for Laws of Semimartingales, *Ann. I.H.P.*, Vol. **20**, 1984, pp. 353-372.

[Mi] T. MIKAMI, *P.H.D. Thesis*, Brown University.

[Na] M. NAGASAWA, *Schrödinger Equations and Diffusion Theory* (to appear).

[Ne] E. NELSON, Stochastic Mechanics and Random Fields, École d'été de Probabilités de St Flour, *L.N. Math.*, 1362, 1988, pp. 429-450.

[Pe] F. PETIT, *Thèse*, Université Paris-VII, 1991.

[RZ] S. ROELLY and H. ZESSIN, *Sur la mécanique statistique d'une particule Browniennne sur le tore*, Preprint, 1991.

[Sha] M. J. SHARPE, *General Theory of Markov Processes*, Academic Press, 1988.

[SV] D. W. STROOCK and S. R. S. VARADHAN, On Degenerate Elliptic Parabolic Operators..., *Comm. Pure. Appl. Math.*, Vol. **25**, 1972, pp. 651-713.

[Ta] H. TANABE, *Equations of Evolution*, Pitman, 1979.
[Zhe] W. A. ZHENG, Tightness Results for Law of diffusion Processes, Application to Stochastic Mechanics, *Ann. I.H.P.*, Vol. **21**, 1985, pp. 103-124.

(*Manuscript received May* 12, 1992;
revised March 1, 1993.)