November 1987 LIDS-P-1718

INFINITE DIMENSIONALITY RESULTS FOR TRAJECTORY MAP ESTIMATION BASED ON THE MALLIAVIN CALCULUS

by

Ofer Zeitouni

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

April 1987
Revised: March 1988

Abstract

The issue of infinite dimensionality for the nonlinear filtering problem has received a lot of attention in recent years, especially by using Lie algebraic methods. Recently, Ocone [5] has applied the Malliavin calculus in an abstract Hilbert space setting to prove results on this topic by showing "density" of the unnormalized density in function space. In this paper, an extension of Ocone's method to deal with first order stochastic Hamilton-Jacobi equations which arise in the maximum a posteriori estimation of trajectories of diffusions is made. The technique used can be easily generalized to a wider class of path-by-path stochastic control problems. The results of the analysis are that, in many cases, the MAP trajectory estimation problem does not possess a "universal" finite dimensional solution in the sense that the associated value function is not finite dimensionally computable, strengthening thus a conjecture made in [6].

This research was carried out while the author was with the Division of Applied Mathematics, Brown University and was supported in part by the Army Research Office under Contract No. DAAG29-84-K-0082, the Office of Naval Research under contract no. N00014-85-K-0607, and by the Weizmann Post-Doctoral Fellowship.

1. INTRODUCTION

The issue of finite dimensionality of solutions to the nonlinear filtering problem has been an active research field in the last few years. Following the ideas of Brockett [1], many authors have considered various Lie algebraic criterions for the existence of such finite dimensional solutions, cf. e.g. [2], [3] and the survey in [4].

Recently, Ocone [5] has proposed a new point of view on the same problem. Namely, Ocone suggested to use the Malliavin calculus of variation to show that under suitable Lie algebraic conditions, the filtering density is smooth in an appropriate infinite dimensional sense over some function spaces, therefore excluding the possibility of the existence of a finite dimensional filter.

The purpose of this paper is to extend Ocone's results to a class of nonlinear stochastic partial differential equations which arise in the study of maximum a posteriori estimation of diffusion processes [6]. This class of equations is basically a class of path-dependent Hamilton-Jacobi equations and is therefore related to an optimal control problem. This relation is exploited over and over throughout this paper, and enables us to convert Ocone's methods, which relied heavily on the linearity of Zakai's equation in nonlinear filtering, to cover the nonlinear class presented here.

The approach we take in this paper is easily generalized to a wider class of stochastic Hamilton-Jacobi equation arising from pathwise optimal control problems. In particular, for nondegenerate diffusions, some of Ocone's results concerning the non existence of finite-dimensional realizations for the conditional density may be re-derived using the relation between optimal filtering and stochastic control pointed out in [7]. We do not pursue this line of research further here.

A main difficulty in our approach is that, unlike in [5], the Hamilton-Jacobi equation dealt with does not have in general a unique solution for a.e. observation path, and moreover it does not even have in general a classical solution. We will work therefore with a subset of the observation space (with Wiener measure greater than zero), and for this subspace we will be able to ensure a local unique classical solution. Some of these problems may disappear in the case of second order Hamilton-Jacobi equations which appear in the filtering equation (or also in viscosity type approximations, cf. [8]). Those remarks are a subject for future investigation and are not dealt with here.

In order to make the presentation as clear as possible, we restrict ourselves to the one dimensional case. The extension to the multi-dimensional nondegenerate case with flat Riemannian structure associated with the diffusion being easy and not illuminating, we do not consider it here. We mention, however, that the general degenerate case or even the non-flat space case is not covered by our method, since the optimal control problem posed is ill-behaved and the trajectory MAP estimator does not always make sense in this case or in the case of non-flat metric, does not necessarily exist, c.f. the discussion in [12].

The organization of the paper is as follows: in the rest of the introduction, we describe the trajectory MAP estimator problem and state our basic assumptions. In section 2, we briefly state some definitions and theorems from [5] and [9] related to Malliavin's calculus. For a thorough introduction to this calculus, we refer the reader to [9], [10] and [11]. In Section 3, we present several auxiliary lemmas which allow us to represent Malliavin's covariance matrix in the spirit of Ocone's technique. As a by-product of our method, we obtain also new local uniqueness results for the MAP problem. We note that although the goal here and in [5] is the same, the techniques used differ basically. In Section 4, we prove Theorem 4.1, which is our basic "infinite dimensionality" criteria. An example and some concluding remarks follows.

We turn now to some definitions. Let w_t, θ_t be two independent <u>one-dimensional</u> Brownian motions adapted to some filtration F_t. Let x_t, $0 \le t \le T$ ("process") and y_t, $0 \le t \le T$ ("observation") satisfy the following pair of stochastic differential equations:

$$dx_t = f(x_t)dt + \sigma(x_t)dw_t \tag{1.1}$$

$$dy_t = h(x_t)dt + d\theta_t \tag{1.2}$$

where (1.1) is interpreted in its Stratonovich form and x_0 possesses the initial density $q_0(x_0)$. We make the following notations:

$C_b^k(\mathbb{R})$ denotes the space of functions $f:\mathbb{R} \to \mathbb{R}$ which are bounded together with their derivatives up to order k, with the k-th order derivatives being absolutely continuous.

$C^k[0,T]$ denotes the space of functions $f:[0,T] \to \mathbb{R}$ which are k times continuously differentiable.

$W^{1,2}$ denotes the Sobolev space of functions $f:[0,T] \to \mathbb{R}$ such that

$||f||_{1,2} < \infty$ where $||f||_{1,2}$ is the Sobolev norm defined by:

$$||f||_{1,2} = [\int_0^T (f^2(s) + (f'(s))^2)ds]^{1/2}$$

where $f'(t)$ denotes the distributional derivative of $f(t)$ w.r.t. t.
H denotes the space of function $f:[0,T] \to \mathbb{R}$ s.t. $f(0) = 0$ and $f \in W^{1,2}$.
$L^n(\mathbb{R}; e^{-x^2/2})$ denotes the space of functions $f:\mathbb{R} \to \mathbb{R}$ s.t.

$$||f||_n \triangleq [\int_{\mathbb{R}} e^{-x^2/2}(f(x))^n dx]^{1/n} < \infty.$$

$$L \triangleq \bigcap_{n \geq 1} L^n(\mathbb{R};\ e^{-x^2/2})$$

$$\mathbb{K} \triangleq L^2(\mathbb{R};\ e^{-x^2/2}).$$

$\mathbb{K}$ is a Hilbert space when equipped with the scalar product naturally associated with $||f||_2$. The Hermite polynomials $\{e_i\}$ form a C^∞, orthonormal, basis of $\mathbb{K}$, and $e_i \in L\ \forall i$.

L^2 denotes the usual Hilbert space with norm $[\int_{\mathbb{R}} f^2(x)dx]^{1/2}$

K denotes throughout a generic constant.

We make the following assumptions on the coefficients of Eqs. (1.1), (1.2):

$$f(x),\ \sigma(x)\ \varepsilon\ C_b^\infty(R),\ h(x)\ \varepsilon\ C^\infty(R),\ h'(x)\ \varepsilon\ C_b^\infty(R), \tag{A.1}$$

$$\exists K > 0 \text{ s.t. } \sigma(x) > K, \tag{A.2}$$

$$\exists K > 0 \text{ s.t. } \ln q_0(x) < K, \tag{A.3}$$

$$q_o(x)\ \varepsilon\ C^\infty_{(R)} \text{ and } \forall x,\ \ q_0(x) \neq 0, \tag{A.4}$$

Under (A.1) - (A.4), it was demonstrated in [6] that the following holds

$$\lim_{\varepsilon \to 0} \frac{P(\|\phi. - x.\| < \varepsilon | y_0^T)}{P(\|w\| < \varepsilon)} = \exp J(\phi,T) \tag{1.3a}$$

where

$$J(\phi,T) \triangleq \ln(q_0(\phi_0)\sigma(\phi_0)) - \frac{1}{2}\int_0^T \frac{(\dot{\phi}_s - f(\phi_s))^2}{\sigma^2(\phi_s)}\, ds \tag{1.3b}$$

$$- \frac{1}{2}\int_0^T \sigma(\phi_s)\left(\frac{f(\phi_s)}{\sigma(\phi_s)}\right)' ds$$

$$- \frac{1}{2}\int_0^T h^2(\phi_s) ds + h(\phi_T)y_T - \int_0^T y_s h'(\phi_s)\dot{\phi}_s ds$$

for each $\phi.\ \varepsilon\ C^1[0,T]$ which satisfies an equation of the type

$$\dot{\phi}_t - \dot{\phi}_0 = \int_0^t k_1(\phi_s, \dot{\phi}_s, s)ds + N(\phi_s)y_s - \int_0^t N'(\phi_s)\dot{\phi}_s y_s ds \qquad (1.3c)$$

for appropriate k_1, N. We refer to [6] and to the appendix of [12] for details. We interpret therefore $J(\phi,T)$ as the "conditional posterior density" of pathes. Now, under the additional hypotheses:

$$|\ln q_0(x)| \text{ is at most of quadratic growth as } |x| \to \infty, \qquad (A.5)$$

$$\exists K > 0 \text{ s.t. } \lim_{|x| \to \infty} \frac{h^2(x)}{x^2} > K. \qquad (A.6)$$

it was demonstrated in [12] that a maximum a posteriori estimator, defined as

$$\overset{*}{\phi}_0 = \underset{\phi \varepsilon W^{1,2}}{\operatorname{argmax}} J(\phi,T) \qquad (1.4)$$

exists and satisfies (1.3c).

We note that the results in [12] apply to our case under consideration here (even with $\sigma(\cdot) \neq \sigma$) since in the one dimensional case, the scalar curvature associated with $\sigma^2(\cdot)$ is identically zero.

We will impose later additional restriction in order to show our infinite dimensionality results, c.f. cases A,B,C below. Only under those additional restrictions we will be able to push through our analysis.

Following Bellman's optimality principle, let us define

$$S(t,x) = \inf_{\substack{\phi \varepsilon W^{1,2} \\ \phi_t = x}} (-J(\phi,t)) \tag{1.5}$$

Note that by definition, $J(\hat{\phi}_0(\cdot),\ t) = -\inf_x S(t,x)$.

Then, as one might suspect from optimal control theory, c.f., e.g., [14],

$$v(t,x) \triangleq S(t,x) + \int^x \frac{f(\theta)}{\sigma^2(\theta)}\, d\theta$$

satisfies (whenever it is differentiable) the following stochastic P.D.E.:

$$dv(t,x) = [-\frac{\sigma^2(x)}{2}\,(\frac{\partial v(t,x)}{\partial x})^2 + \frac{1}{2}\,l(x)]dt - h(x)dy_t;$$

$$v(0,x) = -\ln(q_0(x)\sigma(x)) + \int^x \frac{f(\theta)}{\sigma^2(\theta)}\, d\theta \tag{1.6}$$

where

$$l(x) \triangleq \frac{f^2(x)}{\sigma^2(x)} + \sigma(x)(\frac{f(x)}{\sigma(x)})' + h^2(x).$$

Note that as pointed out in [12], (1.6) is related to the logarithmic transformation of Zakai's equation (which possesses, however, additional smoothness due to an additional Laplacian term in the l.h.s. of (1.6)). Many times throughout, we use the robust form of equation (1.6), i.e. by defining $u(t,x) = v(t,x) + h(x)y_t$ we have:

$$\frac{\partial u(t,x)}{\partial t} = -\frac{\sigma^2(x)}{2}\left(\frac{\partial u(t,x)}{\partial x}\right)^2$$

$$+ \sigma^2(x)h'(x)y_t \frac{\partial u(t,x)}{\partial x}$$

$$-\frac{1}{2}\sigma^2(x)h'^2(x)y_t^2 + \frac{1}{2}\,l(x); \tag{1.7}$$

$$u(0,x) = -\ln[q\ (x)\sigma(x)] + \int^x \frac{f(\theta)}{\sigma^2(\theta)}\,d\theta$$

In the sequel, we concentrate on analyzing the existence, moment bounds smoothness (in the Malliavin sense) and existence of "densities" (over function space, c.f. below) of solutions to (1.6). As pointed out in [5], the existence of a "density" over function space question is related to the question of the existence of finite- dimensional realizations of solutions to (1.6). To demonstrate that such realizations are possible, let $\sigma(x) = 1$, $h(x) = x$, $f(x) = -x$ (which implies that $l(x)$ is quadratic), and $\ln q_0(x) = -x^2$. Substituting $v(x,t) = a_t x^2 + b_t x + c_t$, one easily convinces itself that $v(x,t)$ is indeed a solution of (1.6) and that a_t, b_t, c_t are represented as an output of simple recursive filters driven by the observation y. Note that the non-existence of finite dimensional realizations of $u(t,x)$ does not imply that such realizations for $\hat{\phi}_0$ do not exist. What we show is similar to the "universal" nonexistence of finite dimensional filters for densities in the filtering problem, which does not eliminate in general the possibility of "specific" realizations. Indeed, in our case $v(x,t)$, like $\rho(x,t)$ in the filtering case, cannot be finite-dimensionally computed, whereas $\hat{\phi}_0$ or $\hat{\phi}_T$ might be. For a discussion of this point, c.f., e.g., [3].

In our analysis, the following special cases (which satisfy assumptions (A.1) - (A.5)) turn out to be important; it is only under those more restrictive conditions that we will be able to push through the analysis:

Case A: $h(x) = h.x$, $\sigma(x) = 1$, $\forall x \left|\frac{\partial^2 \ln q_0(x)}{\partial x^2}\right| < K$,

Case B: $\sigma(x) = 1$, $\forall x \left|\frac{\partial \ln q_0(x)}{\partial x}\right| < K$, $\left|\frac{\partial^2 \ln q_0(x)}{\partial x^2}\right| < K$, $|(h^2(x))''| < K$

Case C: $\sigma(x) > K_1$, $\forall x \left|\frac{\partial \ln q_0(x)}{\partial x}\right| < K$, $\left|\frac{\partial^2 \ln q_0(x)}{\partial x^2}\right| < K$, $|(h^2(x))''| < K$

(Note that Case C is equivalent to Case B if $h'(x) > 0$ by making the coordinate transformation $z = \int^x \frac{1}{\sigma(\theta)} d\theta$).

2. STOCHASTIC CALCULUS OF VARIATIONS IN HILBERT SPACE

We present here some of the definitions and results from [5] and [6] which we need. Additional references on the same topic are [9], [10], and [11].

Let Θ denote the space of continuous, $\mathbb{R}$ valued functions on $[0,T]$ starting at zero. Let $\mathcal{B}$ be the Borel σ-algebra of Θ with the sup norm topology, let μ be the Wiener measure on Θ and $\mathcal{B}_t = \sigma\{\theta(s) \mid 0 \leq s \leq t\}$. For an arbitrary Hilbert space X, $L^q(\mu,X)$ denotes the space of X valued random functions F with $E||F||_X^q < \infty$.

Definition 2.1: Let X be a separable Hilbert space, and let $F: \Theta \to X$ be measurable. Then:

a) $F \in w^1(X)$ and has an H differential $DF: \Theta \to HS(H;X)$ (where $HS(H;X)$ denotes the space of Hilbert-Schmidt operators from H to X) if

$$\forall \varepsilon > 0, \forall \gamma \in H, \lim_{t \downarrow 0} \mu\{\theta \mid ||(F(\theta+t\gamma) - F(\theta))/t - DF(\theta)(\gamma)||_X > \varepsilon\} = 0 \tag{2.1}$$

provided $t \to F(\theta + t\gamma)$ is absolutely continuous $\forall \theta,\gamma$. For $\gamma \in H$, denote $DF(\theta)(\gamma)$ by $D\gamma F(\theta)$.

b) For $F \in w^1(X)$, define the Malliavin covariance derivative $\nabla F^* \nabla F: \Theta \to X \otimes X$ as follows:

$$\nabla F^* \nabla F(\theta) \triangleq \sum_{i=1}^{\infty} DF(\theta)(\gamma_i) \otimes DF(\theta)(\gamma_i). \tag{2.2}$$

where $\{\gamma_i\}$ is an orthonormal basis for H.

Various other interpretations of (2.1), (2.2) in terms of the dual of D and in terms of the Ito-Wiener chaos decomposition are given in [5], [9]. We do not use those here.

We use the following lemmas of [5], [16]

Lemma 2.1

(a) Let $F: \Theta \to \mathbb{R}^n$ belong to $w^1(\mathbb{R}^n)$. If $||DF(\theta)||_{HS}^2 < \infty$, $||D||DF(\theta)||_{HS}^2||_{HS}^2 < \infty$ and $\nabla F^* \nabla F > 0$ μ a.s., then $d(\mu \circ F^{-1})/dx$ exists, i.e. the $\mathbb{R}^n$ valued r.v. $F(\theta)$ possesses a density w.r.t. Lebesgue measure.

(b) Let F be as in (a), but without satisfying the extra condition $||D||DF(\theta)||_{HS}^2||_{HS}^2 < \infty$. Then the conclusion still holds (c.f. [16]).

(c) Let F be as in (b), with however the conditions $\nabla F^* \nabla F > 0$ and $||DF(\phi)||_{HS}^2 < \infty$ holding only on an open ret Λ with $\mu(\Lambda) > 0$. Let $\hat{\mu}$ denote the restriction of μ to Λ. Then $d(\hat{\mu} \circ F^{-1})/dx$ exists, i.e. the $\mathbb{R}^n$ valued r.v. $F(\theta)$, restricted to Λ, possesses a density w.r.t. Lebesgue measure ([16, theorems 3.9]).

We remark that one could prove easily a local version of (a) (in the same way as (c) is a local version of (b)). Indeed, let $A \varepsilon \mathbb{R}^n$ be an open set with $m(A) = 0$ and $0 \notin A$, where m is Lebesgue measure on $\mathbb{R}^n$. Using [18, pg. 23, Theorem 4.4], and the fact that on Λ, $\nabla F^* \nabla F > 0$ together with the appropriate technical conditions,

$\mu(\theta | F(\theta).1_{\theta\varepsilon\Lambda} \; \varepsilon A) = 0$

However, if $F(\theta)1_{\theta\varepsilon\Lambda} \; \varepsilon$ A, clearly one has that $\theta\varepsilon\Lambda$. Therefore, one has that $\mu(\theta | F(\theta)\varepsilon A \cap \theta\varepsilon\Lambda) = 0$, which implies existence of density everywhere but near {0}. To show the existence of density in 0, repeat the above argument with $(F(\theta)+1)$.

Finally, we define what we mean by densities in arbitrary Hilbert space, again following [5].

<u>Definition 2.2</u>: Let p(t) ε X. Let s denote any closed finite dimensional subspace of X, and let $\mathrm{Proj}|_s$ denote the projection onto s. The cylinder sets based on s are the elements of the σ-algebra:

$B_s = (\mathrm{Proj}^{-1}|_s(u) | u$ is a Borel set of s)

and define a Lebesgue measure m_s on (X,B_s) by

$\forall u \subset s$, Borel, $m_s(\mathrm{Proj}^{-1}|_s(u)) = m(u)$

where m is Lebesgue measure on s. Then $\mu \circ p(t)^{-1}$ is said to admit a density w.r.t. cylinder sets if $\mu \circ p(t)^{-1}|_B \ll m_s \; \forall s$ as above.

As discussed in [5], the fact that $\mu \circ p(t)^{-1}$ admits a density w.r.t. cylinder sets implies the non-existence of a finite dimensional representation of p(t). For if such a finite-dimensional representation existed, clearly the law of p(t) would be based on the finite dimensional subspace $\tilde{s}$ defined by it; therefore, for any subspace s' strictly bigger than $\tilde{s}$, $\mu \circ p(t)^{-1}$ could not satisfy $\mu o p(t)^{-1}|_{B'} \ll m_{s'}$, in contradiction with the existence of densities. Our goal is the sequel will therefore be towards proving statements like definition (2.2) for v(x,t).

3. SOME PROPERTIES OF v(t,x), AND A STOCHASTIC GRADIENT REPRESENTATION

In this section, we prove some properties of v(t,x) (and specifically, smoothness results w.r.t. x) which will turn out to be useful in the computation of the stochastic gradient of v(t,x). As a corollary we derive a result on the local uniqueness of the MAP trajectories estimator. This result is not used in the sequel.

Throughout, we need the following property:

(P) There exists a set $\Lambda \in C_0[0,T]$ and a constant $\tau > 0$ s.t.:

$$\mu(\Lambda) > 0, \tag{3.1a}$$

$$\forall y. \in \Lambda, \; u(t,x) \text{ is } C^1 \text{ w.r.t. } t \text{ and } C^\infty \text{ w.r.t. } x \text{ for } (t,x) \in [0,\tau] \times \mathbb{R} \tag{3.1b}$$

Our approach in proving P will be to use the classical method of characteristics, a short account of which is brought in the appendix. We note that by the general results of the method of characteristics, there exists a neighborhood of the line (0,x) where u(t,x) is C^1 w.r.t. t and C^∞ w.r.t. x (c.f., e.g., [17, pg. 24, thm. 8.1]). In (3.1b) we require however this neighborhood to be uniform in x, and to achieve that we need the restrictions imposed by cases A,B,C.

We have the following lemma:

Lemma 3.1. For cases A, B, C, property (P) holds.

Proof: We treat separately Cases A, B, and Case C. Our method of proof will be to show that the characteristic curves (c.f. lemma 1 in the appendix) define, for $y. \varepsilon \Lambda$, a diffeomorphism in $[0,\tau] \times \mathbb{R}$. Therefore, by the appendix, P holds for $y. \varepsilon \Lambda$.

<u>Case A,B.</u> Let $\sigma(x) = 1$. Then, the Hamiltonian for the robust equation (1.7) is:

$$H(t,x,p) = \frac{1}{2} p^2 - y_t h'(x)p - \frac{1}{2} l(x) + \frac{1}{2} y_t^2 h'^2(x). \tag{3.2}$$

The characteristic equations are therefore:

$$\dot{X} = P - y_t h'(X), \quad X(0) = x, \tag{3.3a}$$

$$\dot{P} = y_t h''(X)P + \frac{1}{2} l'(X) - y_t^2 h'(X)h''(X); \; P(0) = -\frac{\partial}{\partial x} \ln q_0(x) + f(x) \tag{3.3b}$$

Define $\xi_t \triangleq \frac{\partial X(t)}{\partial t}$; $\pi_t \triangleq \frac{\partial P(t)}{\partial x}$. Then, one has

$$\dot{\xi}_t = \pi_t - y_t h''(X)\xi_t; \quad \xi(0) = 1$$

$$\dot{\pi}_t = y_t h^{(3)}(X)P\xi_t + y_t h''(X)\pi_t + \frac{1}{2} l''(X)\xi_t - \frac{y^2{}_t((h'(X))^2)''}{2} \xi_t;$$

$$\pi(0) = -\frac{\partial^2}{\partial x^2} \ln q_0(x) + \frac{\partial}{\partial x} (f(x))$$

Take $\Lambda \triangleq \{y. \,|\sup_{0 \leq s \leq T} |y_s| < 1\}$. Clearly, $\mu(\Lambda) > 0$. Let K denote a bound on $l''(x)$, $f(x)$, $h'(x)$, $|\partial^2 \ln q_0(x)/\partial x^2| \; q_0(x)|$ and, in case B, also on $|\partial \ln q_0(x)/\partial x| \; q_0(x)|$ and $h''(x)$.

In Case A, note that $\dot{\xi}_t = \pi_t$, $|\dot{\pi}_t| \leq K\, \xi_t$. Let

$$m \triangleq \sup\{\xi_s | \pi_s = 0, \quad 0 \leq s \leq 1, \ \xi_\tau > 0 \text{ in } \tau \in [0,s]\}$$

$$\theta = \inf\{s | \xi_s = m\}$$

with the convention that $m = \sup_{s \in [0,1]} \xi_s$ and $\theta=1$ if $\pi_s > 0$ in [0,1] and m=1, $\theta=0$ if $\pi_s < 0$ as long as $\xi_s > 0$. Clearly, one has

$$|\pi_{t+\theta}| \leq mK + mKt$$

$$|\xi_{\theta+t}| \geq m - \frac{mKt^2}{2} - mKt$$

Choosing $\tau = 2(\sqrt{1+2/K} - 1)$, one has $\xi_s > 0$ for $s \in [0,\tau]$, and therefore $x \rightarrow X(t,x)$ is a diffeomorphism (w.r.t. x), and the assertion follows.

Turning our attention to Case B, note that P is bounded under our assumptions. Let K denote also the bound on $h^{(3)}(\cdot)P$. One has then

$$|\dot{\pi}_t| \leq K|\pi_t| + K|\xi_t|$$

$$\dot{\xi}_t \geq \pi_t - K|\xi_t|$$

Let m, θ be defined as above. One gets

$$\pi_t \geq -Km(1+\frac{1}{K}) \exp(xt)$$

$$\xi_{\theta+t} \geq m \exp(-Kt) - 2(1+\frac{1}{K}) \sinh(Kt) m$$

Choosing $\tau = \ln\frac{K+1}{2K+1}/2K$, one has $\xi_s > 0$ for $s \in [0,\tau]$, with the same conclusion as above.

Case C. Let $\sigma(x) \neq 1$. In this case, the characteristic curve equations are:

$$\dot{X} = \sigma^2(X)P - y_t h'(X)\sigma^2(X), \qquad X(0) = x \tag{3.4a}$$

$$\dot{P} = -\sigma(X)\sigma'(X)P^2(X) + y_t h''(X)P\sigma^2(X) + \frac{1}{2} 1'(X) - \frac{1}{2} y_t((h'(X))^2)'\sigma^2(X) \tag{3.4b}$$

$$-2\sigma(X)\sigma'(X)(-y_t h'(X)P + \frac{1}{2} y_t^2 h'^2(X)); \; P(0) = -\frac{\partial}{\partial x}(\ln q_0(x)\sigma(x)) + \frac{f(x)}{\sigma(x)}$$

Therefore, $\sigma(X)P$ satisfies:

$$(\sigma(\dot{X})P) = +\sigma^2(X)\sigma'(X)y_t h'(X) + \sigma^2(X)y_t h^{(3)}(\sigma(X)P)$$

$$+ \frac{1}{2} 1'(X)\sigma(X) - \frac{1}{2} y_t(h'(X)^2)'\sigma^3(X) - \sigma^2(X)\sigma'(X)y_t^2(h'(X))^2 \tag{3.5}$$

which guarantees that $|P| < K$ for some K independent of x.

Therefore, one has, by defining $\xi = \partial X/\partial x$, $\pi = \partial\sigma(X)P/\partial x$:

$$\dot{\xi}_t = \sigma\pi_t + (\sigma P)\sigma'\xi_t - y_t h''\sigma^2\xi_t - 2y_t h'\sigma\sigma'\xi_t; \; \xi(0) = 1,$$

$$\dot{\pi}_t = F_1(X,y_t)\xi_t + F_2(X,y_t)\pi_t \quad ;$$

$$\pi(0) = \frac{\partial f(x)}{\partial x} - \frac{\partial}{\partial x}(\sigma(x)\frac{\partial}{\partial x} \ln q_0(x)\sigma(x))$$

where $F_1(X,y_t)$, $F_2(X,y_t)$ are some bounded functions. The conclusion follows under our assumptions exactly as in Cases A,B.

We define now the MAP trajectory with endpoint x, $\hat{\phi}^x$, as:

$$\hat{\phi}^x = \arg \max_{\substack{\phi \varepsilon w^{1,2} \\ \phi_t = x}} J(\phi,t)$$

The definition is justified by the:

Corollary 3.2. The MAP trajectory estimator with endpoint x exists and is unique in Cases A, B, C for all t,x in some strip $[0,\tau] x \mathbb{R}$ and all y. ε Λ.

Proof: The fact that $\hat{\phi}^x$ exists is an easy adaptation of the proof of [12]: indeed, the proof in [12] consisted of the following two steps:

a) demonstrating that $J(\phi)$ is bounded in $w^{1,2}$ which trivially still holds in the restriction of $w^{1,2}$ to those paths with $\phi_t = x$.

b) demonstrating that $J(\phi)$ is l.s.c. in the weak topology in $w^{1,2}$- which still holds on the restriction due to the fact that ϕ_t is a continuous functional in $w^{1,2}$.

The uniqueness follows by the method of proof of theorem 1 in [14]. cf. remark 2 in the appendix.

Lemma 3.3: A generalized solution v(x,t) which is the value function of the optimal control problem (1.4) (modulo the deterministic C^∞ drift shift $\int^x \frac{f(\tau)}{\sigma^2(\tau)} d\tau$) exists in L. Moreover,

$$(\forall n > 1)\ E\int_{\mathbf{R}}(v(t,x))^n e^{-x^2/2}dx < \infty,\quad E\int_{\mathbf{R}} e^{-x^2/2}\left[\frac{\partial v(t,x)}{\partial x}\right]^n dx < \infty \tag{3.7}$$

and the same bounds apply also for the time integrals (in [0,T]) of these expressions.

Moreover, in Cases A,B,C, the solution $v(t,x)$ is a classical solution with the smoothness properties of Lemma 3.1 and (3.7) holds in $[0,\tau]$ when the expectations are conditioned on the cylinder set defined by Λ.

Finally, in Λ, $\partial^2 v(t,x)/\partial x^2 \leq K(1+|x|^2)$ and $|\partial v(t,x)/\partial x| < K(1+|x|)$ for some K.

<u>Proof</u>: Note that by our assumptions,

$$-K(1 + (y_s^*)^2 + |x|^2 \leq J(\phi_x) \leq K[1 + \int_0^T y_s^2 ds + |x|]$$

where y_s^* denotes the maximum of $|y_\cdot|$ on [0,T]. Therefore, by the boundedness of $f(x)$, one has

$$|v(t,x)| \leq K(1 + (y_s^*)^2 + |x|^2)$$

and

$$(\forall n)\quad \int_{\mathbf{R}}(v(t,x))^n e^{-x^2/2}dx < \infty$$

with the same bound on the expectation of the last quantity. Therefore, the first half of the lemma is proved.

Rewriting now $J(\phi,T)$ as

$$J(\phi,T) = \ln(q_0(\phi_0)\sigma(\phi_0)) + h(\phi_T)y_T \tag{3.8}$$

$$- \frac{1}{2}\int_0^T \sigma(\phi_s)\left[\frac{f(\phi_s)}{\sigma(\phi_s)}\right]' ds$$

$$- \frac{1}{2}\int_0^T \frac{(\dot{\phi}_s - f(\phi_s) + y_s h'(\phi_s)\sigma^2(\phi_s))^2}{\sigma^2(\phi_s)}$$

$$+ \frac{1}{2}\int_0^T y_s^2 h'^2(\phi_s)\sigma^2(\phi_s)ds$$

$$- \int_0^T f(\phi_s)y_s h'(\phi_s)ds - \frac{1}{2}\int_0^T h^2(\phi_s)ds,$$

one has $\underset{||\phi||_{1,2}\to\infty}{J(\phi^x,T) \to -\infty}$. Therefore, $||\overset{*}{\phi}{}^x||_{1,2} \leq K_x(y_0^T)$, where for some $K_1 > 0$

$$K_x(y_0^T) \leq K_1(1 + y_s^{*2} + |x|)$$

by (A.1) - (A.4) as in [12], and therefore

$$|(\overset{*}{\phi}{}^x)_t| \leq K_2(1 + y_s^{*2} + |x|)$$

due to the fact that $w_{1,2}$ is a reproducing kernel Hilbert space (RKHS). Following [14], note that, by direct computation,

$$|J(\overset{*}{\phi}^x + x' - x,T) - J(\overset{*}{\phi}^x, T)| \leq K_3|x - x'|(||\overset{*}{\phi}^x||_{1,2} + 1)^2 \tag{3.9}$$

$$+ |\ln q_0(\overset{*}{\phi}^x + x - x') - \ln q_0(\overset{*}{\phi})|$$

$$\leq K_4|x - x'|(1 + (y_s^*)^2 + |x|)^2$$

where the last inequality is due to (A.5). Note now that $S(t,x) \leq -J(\overset{*}{\phi}^x + x' - x)$, whereas $S(t,x) = -J(\overset{*}{\phi}^x)$. Therefore,

$$S(t,x') - S(t,x) \leq K_4|x - x'|(1 + (y_s^*)^2 + |x|)^2. \tag{3.10}$$

Interchanging the role of x, x', we get the opposite inequality, and hence

$$|S(t,x') - S(t,x)| \leq K_4|x - x'|(1 + (y_s^*)^2 + |x|)^2.$$

which implies the same relation on $v(t,x)$ due to the boundedness of $f(x)$ and $1/\sigma(x)$. Therefore,

$$\int_{\mathbf{R}} e^{-x^2/2}\left[\frac{\partial v(t,x)}{\partial x}\right]^n dx \leq K_5(1 + (y_s^*)^n).$$

Since our boundedness assumptions ensure that $E(y_s^*)^m < \infty$ for all m, one has the second half of the lemma, and the proof is completed for the general case.

In Cases A,B,C, the conclusion of the lemma follows from the combination of the above considerations and Lemma 3.1. We note that in those cases, we have shown (in lemma 3.1) that either P is bounded (cases B,C) or $|P| \leq K(1+x)$. Since, by (A.7) in the appendix, $P = \partial v/\partial x|_{\hat{\phi}}$, one has a tighter bound; namely, $|\partial v/\partial x| \leq K(1+|x|)$ for y. ε Λ.

The bound on the second derivative of v(t,x) is proven in a similar way. Indeed, for any function f, let

$$\Delta_\alpha f(x) \triangleq \frac{f(x+\alpha) + f(x-\alpha) - \partial f(x)}{\alpha^2} ;$$

repeating the arguments above, one has:

$$\alpha^2\Delta_\alpha S(x,t) \le -J(\overset{*}{\phi}{}^x+\alpha,T)-J(\overset{*}{\phi}{}^x-\alpha,T) + 2J(\overset{*}{\phi}{}^x,T)$$

$$\le \left|J(\overset{*}{\phi}{}^x+\alpha,T)-J(\overset{*}{\phi}{}^x,T) + J(\overset{*}{\phi}{}^x-\alpha,T)-J(\overset{*}{\phi}{}^x,T)\right|$$

$$\le \alpha^2\left|\Delta_\alpha \ln(q_0(x)\sigma(x))\right| + \alpha^2\left|\Delta_\alpha h(\overset{*}{\phi}{}^x_T)y_T\right|$$

$$+ \frac{1}{2}\int_0^T \alpha^2\left|\Delta_\alpha(\sigma(\cdot)(\tfrac{f}{\sigma})'(\cdot))\big|_{\overset{*}{\phi}{}^x_s}\right|ds$$

$$+ \frac{1}{2}\int_0^T y_s^2\ \alpha^2\left|\Delta\alpha(h'^2(\cdot)\sigma^2(\cdot))\big|_{\overset{*}{\phi}{}^x_s}\right|ds$$

$$+ \int_0^T |y_s|\alpha^2\left|\Delta_\alpha(f(\cdot)h'(\cdot))\big|_{\overset{*}{\phi}{}^x_s}\right|ds$$

$$+ \frac{1}{2}\int_0^T \alpha^2(\Delta_\alpha h^2(\cdot))\big|_{\overset{*}{\phi}{}^x_s}\ ds$$

$$+ \int_0^T \dot{\phi}_s^2\alpha^2(\Delta_\alpha \frac{1}{\sigma^2}(\cdot))\big|_{\overset{*}{\phi}{}^x_s}\ ds$$

$$+ 2\int_0^T \alpha^2\left|\Delta_\alpha \frac{f(\cdot)}{\sigma^2(\cdot)}\right|\big|_{\overset{*}{\phi}{}^x_s}\ ds$$

$$+ 2\int_0^T y_s^2\alpha^2\left|\Delta_\alpha(h'(\cdot)\sigma^2(\cdot))\right|\big|_{\overset{*}{\phi}{}^x_s}\ ds$$

and we deduce that $\frac{\partial^2 S(x,t)}{\partial x^2} \le K(1+|x|^2)$ on Λ. Again, the same bounds hold for $v(t,x)$ by the boundedness of derivatives of $f(\cdot)/\sigma^2(\cdot)$.

[ ]

Finally, we prove three additional lemmas, which are needed for the stochastic gradient representation.

Lemma 3.4. In Cases A, B, C, a classical solution to the following equation exists in the strip $[0,\tau] \times \mathbb{R}$ for each $\nu \in H$, $||u||_{1,2} = 1$ and $y. \in \Lambda$:

$$\frac{\partial \mu(t,x)}{\partial t} = -\sigma^2(x) \frac{\partial v(t,x)}{\partial x} \frac{\partial \mu(t,x)}{\partial x} - h(x)\nu'(t); \quad \mu_0(x) = 0. \tag{3.13}$$

This solution is unique as a classical solution. Moreover,

$$|\mu(t,x)| < K_\nu |x|^\beta, \quad \left|\frac{\partial \mu(t,x)}{\partial x}\right| < K_\nu |x|^\beta, \quad \forall y \in \Lambda \tag{3.14}$$

for some K_ν and $\beta < \infty$, independent of $y. \in \Lambda$. Finally,

$$E_\Lambda \sum_{i=1}^{\infty} ||\mu_{\nu_i}(t,x)||_K < \infty$$

where $\mu_\nu(t,x)$ denotes the solution of (3.13) with ν_i substituted instead of ν and ν_i is a complete orthonormal basis of H.

Proof. We use again the method of characteristics: The characteristic curve $X^x(t)$ for (3.13) satisfies

$$\frac{dX(t)}{dt} = -\sigma^2(x(t))\,\frac{\partial v(t,X(t))}{\partial x}\,,\quad X(0) = x \tag{3.15}$$

since $\frac{\partial v}{\partial x}$ is locally Lipshitz, and $\left|\frac{\partial v}{\partial x}(t,x)\right| \lesssim K(1+|x|)$, a solution to (3.15) exists in $[0,\tau] \times \mathbb{R}$, and $|X(t)| \lesssim K(1+|x|)$. Moreover, defining $\eta = \partial X(t)/\partial x$, one has

$$\frac{d\eta}{dt} = \sigma^2(x(t))\,\frac{-\partial^2 v}{\partial x^2}(t,X(t))\eta - \frac{\partial v}{\partial x}(t,X(t))\frac{\partial}{\partial x}\sigma^2(X(t))\xi;\ \eta(0) = 1$$

and therefore, due to our bound on $\frac{\partial^2 v}{\partial x^2}$, $\eta(t) > 0$ for all $t \in [0,\tau]$ which implies that $X(t)$ is a diffeomorphism on $[0,\tau] \times \mathbb{R} \to \mathbb{R}$ and therefore $X(t)$ is a characteristic curve and a classical solution to (3.13) exists. Moreover, the bounds (3.14) follow easily from the method of characteristics exactly as in lemma 3.3 above.

Finally, invoking the characteristic representation of the appendix, one finds that, for the system (3.13)

$$\dot{P} = - Pv_{xx}(X,t) - h'(X)\nu_i; \quad P(0) = 0 \tag{3.17a}$$

$$\dot{U}_{\nu_i} = Pv_x(X) + q_t = Pv_x(X) - H(X,t,P) + H(X,0,0) \tag{3.17b}$$

$$= PV_x(X) - PV_x(X) - h(X)\nu_t$$

$$= -h(x)\nu \,, \quad U_{\nu i}(0) = 0$$

Therefore,

$$U_{\nu_i}(t,x) = -\int_0^t h(X(x,s),s)\nu_i(s)ds \tag{3.18}$$

where $|h(X(x,s)| \lesssim K|X(x,s)|$; since, for each x, $\int_0^t h^2(X(x,s))ds < \infty$, (3.18) states that U_ν (t,x) is the projection of $-h(X(x,s),s)$ in the direction $\nu_i(s)$. By our bounds on X we therefore get

$$\sum_{i=1}^{\infty} ||U_{\nu_i}(t,X^{-1}(x,t))||_K^2 \lesssim \int_0^t ||h^2(X(X^{-1}(x,t),s)||_K^2 ds \lesssim K < \infty$$

with the same K for all y. ε Λ. The lemma is proved.

<u>Lemma 3.5</u>. Let v_ν^ε denote the solution of Eq. (1.7) with $y_t + \varepsilon\nu_t$ substituted instead of ν_t. Then, for y. ε Λ and $||\nu|| < 1$, and for a small enough ε,

$$|v - v^{\varepsilon}_{\nu}| < \varepsilon K_{\nu}(1 + |x|) \tag{3.21}$$

$$\left|\frac{\partial}{\partial x}(v - v^{\varepsilon}_{\nu})\right| < \varepsilon K_{\nu}(1 + |x|). \tag{3.22}$$

where K_v is independent of y. ε Λ.

Proof: $n^{\varepsilon} = v - v^{\varepsilon}_{\nu}$ satisfies the following equation:

$$\frac{\partial n^{\varepsilon}(t,x)}{\partial t} = \frac{-\sigma^2(x)}{2}\left[\frac{\partial v(t,x)}{\partial x} + \frac{\partial v^{\varepsilon}_{\nu}(t,x)}{\partial x}\right]\frac{\partial n^{\varepsilon}(t,x)}{\partial x} + \varepsilon h'(x)\nu_t;\ n^{\varepsilon}(0,x) = 0.$$

The characteristic method and corollary A1 of the appendix may be applied here to yield the conclusion. We omit the details.

Define now

$$m^{\varepsilon}_{\nu}(t,x) = \frac{v(t,x) - v^{\varepsilon}_{\nu}(t,x)}{\varepsilon} - \mu(t,x).$$

We have then

Lemma 3.6. $||m^{\varepsilon}_{\nu}(t,x)||_K \underset{\varepsilon \to 0}{\to} 0$. Moreover, $E_{\Lambda}(||m^{\varepsilon}_{\nu}(t,x)||_K)^n < \infty\ \forall n$.

Proof: $m^{\varepsilon}_{u}(t,x)$ is the solution to the following equation:

$$\frac{\partial m_u^\varepsilon(t,x)}{\partial t} = -\frac{\sigma^2(x)}{2}\left[\frac{\partial v(t,x)}{\partial x} + \frac{\partial v_\nu^\varepsilon(t,x)}{\partial x}\right]\frac{\partial m_\nu^\varepsilon(t,x)}{\partial x}$$

$$+\frac{\sigma^2(x)}{2}\left[\frac{\partial v(t,x)}{\partial x} - \frac{\partial v_\nu^\varepsilon(t,x)}{\partial x}\right]\frac{\partial \mu}{\partial x}\ ;\ m_\nu^\varepsilon(0,x) = 0.$$

The assertion follows now easily from Lemma 3.4, 3.5, again by the method of characteristics and lemma A1 of the appendix.

We conclude this section by combining the results of the above lemmas to:

<u>Theorem 3.1</u>: $D_\nu v(t,x)$ exists and satisfies Eq. (3.13). Moreover $Dv(t,x)$ is a HS operator $H \to K$. Finally, $\|D\|Dv(t,x)\|_{HS}^2\|^2 < \infty$.

<u>Proof</u>: The existence of $D_\nu v(t,x)$ follows directly from Lemma 3.6. The existence of $Dv(t,x)$ then follows from the bound on its HS norm presented in Lemma 3.4. We have therefore only to prove the last statement of the theorem, i.e. we have to show that

$$\sum_{i,j} \|D_{\nu_j} D_{\nu_i} V(x,t)\|_K^2 < \infty$$

for a complete orthonormal basis $\{\nu_i\}$. However, let $\eta_j^i \triangleq D_{\nu_j} D_{\nu_i} v(x,t)$, then η_j^i satisfies

$$\frac{d\eta_j^i}{dt} = -\,[\frac{\partial v}{\partial x}][\frac{\partial \eta_j^i}{\partial x}] - \frac{\partial}{\partial x}(\mu_{\nu_j}(t,x))[\frac{\partial}{\partial x}\mu_{\nu_i}(t,x)]. \tag{3.19}$$

Note that we have

$$\frac{\partial}{\partial x}\mu_{\nu_j}(t,x) = \int_0^t k_y(x,s)\nu_j'(s)ds \triangleq \beta_j(t,x) \tag{3.20}$$

for some uniformly (w.r.t. t) bounded kernel $k_y(x,s)$ (and, therefore, $\sum_j ||\beta_j(t,x)||_K^2 < \infty$, uniformly in $[0,\tau]$). Therefore, exactly as in the derivation of (3.18) one has

$$\eta_j^i = \int_0^t \tilde{k}_y(x,s)\beta_i(s,x)\beta_j(s,x)ds$$

with, again, $\tilde{k}_y(x,s)$ being a bounded kernel. Therefore,

$$\sum_{i,j} ||\eta_j^i||^2 \le K \int_0^t ||\beta_i(s,x)||_K^2 ||\beta_j(s,x)||_K^2 ds \le Kt||\beta_i(t,x)||_K^{*2}||\beta_j(t,x)||_K^{*2}$$

where $||\beta_i(t,x)||_K^*$ denotes the supremum (over t) of $||\beta_i(t,x)||_K$, which exists due to the uniform boundedness of $\sum_i ||\beta_i(t,x)||_K^2$. One concludes therefore that

$$\sum_{i,j} ||\eta_j^i||^2 < \infty$$

and the theorem is proved.

Remarks: (1) Actually, Theorem 3.1 is nothing but the extension of the usual rules of Frechet derivation to the case at hand: indeed, substituting the Frechet derivative of $(\partial/\partial X)^2$ in equation (1.7) would yield the correct expression for $D_y((\partial v/\partial X)^2)$.

(2) The last part of the theorem is not needed in the sequel, due to the results of [16]. However, since it yields somewhat stronger information, it is given here for the sake of completeness.

4. A SMOOTHNESS CONDITION

In this section, we derive a smoothness criterion for those solutions of (1.6) which are the value function of the optimal control problem (1.4). By Lemma 2.1(c), a sufficient condition for a density of (1.6) to exist w.r.t. cyclinder sets on a set Λ with $\mu(\Lambda) > 0$ is that

$$\forall a \in \mathbf{K}, \sum_{i=1}^{\infty} (D_{\alpha_i} u, a)^2 > 0 \tag{4.1}$$

where α_i is an orthonormal basis for H and the scalar product in (4.1) is taken in $\mathbf{K}$. See also [5] for a discussion of this point.

Throughout, we assume Cases A, B, or C, so that v(t,x) is C^∞ w.r.t. x and possesses an Itô representation in some strip $[0,\tau]\mathrm{x}\mathbf{R}$ for $y_\cdot \in \Lambda$ with $\mu(\Lambda) > 0$. We need the following definitions:

<u>Definition 4.1</u>. Σ will denote the space of functions f(t,x) which are C^∞ w.r.t. x and possess an Itô representation, i.e., for $f(t,x) \in \Sigma$

$$f(t,x) = \int_0^t (d_R f(s,x))ds + \int_0^t (d_I f(s,x))dy_s \tag{4.2}$$

<u>Definition 4.2</u>. Let the operator L_v^t be defined as

$$L_v^t \triangleq \sigma^2(x) \frac{\partial v(t,x)}{\partial x} \frac{\partial}{\partial x} .$$

Define the set $\Phi_V(t)$ of functions f(t,x) as follows:

(a) $L_v^t h(x) \in \Phi_v(t)$,

(b) if $f(t,x) \in \Phi_v(t)$, so does $L_v^t f(t,x) + d_R f(t,x)$,

(c) if $f(t,x) \in \Phi_v(t)$, so does $d_I f(t,x)$.

Definition 4.2 is a reminiscent of the usual definition of Lie brackets of time invariant operators. Indeed, had not v(t,x) been time varying, one could have considered $L_v^t f(x)$ as

$$[(\sigma^2(x)\frac{\partial^2}{\partial x^2})\partial,\ f(x)\partial]$$

where ∂ denotes a Frechet derivation. The inclusion of time varying operators leads to (b) and (c).

We are ready now to state our main result:

<u>Theorem 4.1</u>: Assume that $\Phi_v(t)|_{t=0}$ is dense in $\mathbb{K}$. Then, v(x,t) possesses a density w.r.t. to cylinder sets on Λ (for $t \in [0,\tau]$), i.e. no finite dimensional solution exists for problem (1.4).

<u>Proof</u>: Let $\phi(t,s)\colon \mathbb{K} \cap C^1(\mathbb{R}) \to \mathbb{K} \cap C^1(\mathbb{R})$ satisfy:

$$\frac{d}{dt}(\phi(t,s) \circ a) = -\frac{\partial v}{\partial x}(t,x)\,\frac{\partial}{\partial x}(\phi(t,s) \circ a);\ \phi(s,s)(a) = a \tag{4.3}$$

for all $a \in \mathbb{K}$. One should consider (4.3) as a definition of $\phi(t,s)$. Note that a solution to (4.3) exists, again by the method of characteristics, for each $a \in \mathbb{K} \cap C^1(\mathbb{R})$. It is straightforward to check that the solution of Eq. (3.13) is, for each $\alpha \in H$,

$$D_\alpha v(t,x) = \int_0^t [\phi(t,s) \circ (h(x)\alpha'(s))]ds, \qquad (4.4)$$

and v(t,x) will possess a density w.r.t. cylinder sets on Λ (for $t \in [0,\tau]$) if $\sum_i (D_{\alpha_i} v(t,x),a)^2 > 0$, $\forall a \in \mathbf{K}$. Assume the contrary, i.e. assume that

$$\exists a \in \mathbf{K} \text{ s.t. } \forall \alpha \in H \ (D_\alpha v(t,x),a) = 0 \qquad (4.5)$$

However, by (4.4) one has:

$$0 = (D_\alpha u,a) = \int_0^t (\phi(t,s) \circ h(x)\alpha'_s,a)ds$$

$$= \int_0^t (\phi(t,s) \circ h(x),a)\alpha'_s ds \qquad (t < \tau).$$

Therefore, $(\phi(t,s) \circ h(x),a) = 0$ in $[0,\tau]$. Moreover, for $s \lesssim t$, one has

$$(\phi(t,s) \circ h(x),a) = -(h(x), \phi^*(t,s) \circ a) \qquad (4.6)$$

where $\phi^*(t,s)$ is the adjoint of $\phi(t,s)$, and satisfies the equation:

$$\frac{d(\phi^*(t,s)oa)}{ds} = \frac{\partial(\phi^*(t,s)oa)}{\partial x}\frac{\partial v(s,x)}{\partial x}; \ (\phi^*(t,t) \circ a = a). \qquad (4.7)$$

Expanding the r.h.s. of (4.7), one has

$$-(h(x),\phi^*(t,s) \circ a) = -(h(x),\phi^*(t,0) \circ a) - \int_0^t [\frac{\partial v}{\partial x}\frac{\partial h}{\partial x}, \phi^*(t,r) \circ a]dr \tag{4.8}$$

and therefore, one has again $\forall r \in [0,t]$, $0 = [\frac{\partial v}{\partial x}\frac{\partial h}{\partial x}, \phi^*(t,r) \circ a]$. Using the same procedure over and over and the fact that $v(x) \in C^\infty(\mathbb{R})$ and that all the coefficients involved are $C^\infty(\mathbb{R})$ by assumption, one concludes that for each $\delta \in \Phi_V(r)$,

$$(\delta(r,x), \phi^*(t,r) \circ a) = 0. \tag{4.9}$$

Taking $r = 0$ and noting that $\phi^*(t,0) \circ a \neq 0$ on Λ where 0 denotes the fucntion whose value is zero everywhere (by the characteristics method and the fact that $\phi^*(t,t) \circ a = a$), one has a contradiction to the fact that $\Phi_V(t)|_{t=0}$ is dense in K, and the proof is complete.

Remark. As pointed out by Ocone, the explicit computation of the condition on the span of $\Phi_V|_{t=0}$ is usually difficult to check out. (The cases where the infinite dimensionality is easy to check, e.g. $h(x) = x^3$ or $f(x)$ a high order polynomial, do not satisfy the technical assumptions (A.1) - (A.6) or the growth conditions imposed by Cases A, B, C.) The following corollary of Theorem 4.1 is an example of the kind of results one may expect.

Corollary 4.1. Assume that there exists an $n_0 > 1$ s.t., for all $n \geq n_0$,

$$h^{(n)}(x) \in L^2. \tag{4.10a}$$

The support of $H_{n_o}(\omega)$ is (up to a set of measure zero) all of $\mathbb{R}$. (4.10b)

$$\exists\delta > 0 \text{ s.t. } H_{n_o}(\omega)e^{+\delta\omega^2/4} \text{ is bounded} \qquad (4.10c)$$

where $H_n(\omega) \triangleq F(h^{(n)}(x))$ is the Fourier transform of $h^{(n)}(x)$.

Then $\Phi_V(t)|_{t=0}$ is dense in $\mathbb{K}$ and no finite dimensional solution exists to the MAP problem.

Proof. We will prove that $\{h^{(n)}(x)\}$, $n \geq n_0$, which is easily checked to belong to $\Phi_V(t)|_{t=0}$, spans $\mathbb{K}$. Indeed, assume the contrary. Then

$$(\exists g \in \mathbb{K}) \text{ s.t. } \forall n \geq n_0, \quad \int_{\mathbb{R}} h^{(n)}(x)g(x)e^{-x^2/2}dx = 0. \qquad (4.11)$$

Since $h^{(n)}(x) \in L^2$ and $g(x)e^{-x^2} \in L^2$, (4.11) is a statement in L^2, which means that $\{h^{(n)}(x)\}_{n\geq n_o}$ does not span L^2, either. Therefore, working on Fournier space, one has

$$\exists\tilde{G} \neq 0 \in L^2 \text{ s.t. } \forall n \geq n_o, \int_{\mathbb{R}} H_n(\omega)\tilde{G}(\omega)d\omega = 0. \qquad (4.12)$$

However, $H_n(\omega) = (j\omega)^{n-n_0}H_{n_0}(\omega)$. Rewriting (4.12), one has

$$\exists G \neq 0 \in \mathbf{K} \text{ s.t. } (\forall n \geq n_0) \int_{\mathbf{R}} [\omega^n e^{(\omega^2/4)(1-\delta)}][H_{n_0}(\omega)e^{(\omega^2/4)\delta}G(\omega)]e^{-\omega^2/2}d\omega = 0 \tag{4.13}$$

where $G(\omega) = \tilde{G}(\omega)e^{\omega^2/4}$. By (4.10a), $[H_{n_0}(\omega)e^{(\omega^2/4)\delta}G(\omega)] \in \mathbf{K}$. Since, for $1>\delta\geq 0$, one may easily check that $\omega^n e^{(\omega^2/4)(1-\delta)}$ spans $\mathbf{K}$, (4.13) implies that $(H_{n_0}(\omega)e^{(\omega^2/4)\delta}G(\omega)) = 0$ a.e., which is clearly a contradiction to (4.10b). Hence, the corollary is proved.

Acknowledgements. I would like to thank D. Ocone for providing me with the reprint of [5], which initiated this work, and for some helpful hints regarding it, in particular for bringing [16] to my attention. Also, I would like to thank O. Hijab, E. Pardoux and A. Dembo for several discussions and suggestions concerning this work. Finally, the remarks of the reviewer provided a strong stimulus for making the presentation clearer and more correct.

References

[1] Brockett, R.W., Nonlinear systems and nonlinear estimation theory, in: Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Hazewinkel & Willems, eds., Reidel-Dordrecht, 1981.

[2] Ocone, D., Finite dimensional estimation in nonlinear filtering, Stochastic Systems, loc. cit.

[3] Chaleyat-Maurel, M. and Michel, D., Des resultats de non existence de filtre de dimension finie, Stochastics, 13, 1984, pp. 83-102.

[4] Marcus, S.I., Algebraic and geometric methods in nonlinear filtering, SIAM J. Control and Optimization, 22, 1984, pp. 817-844.

[5] Ocone, D., Stochastic calculus of variations for stochastic partial differential equations, Preprint, 1986.

[6] Zeitouni, O. and Dembo, A., A maximum a posteriori estimator for trajectories of diffusion processes, Stochastic, 20, 1987, pp. 211-246.

[7] Fleming, W.H. and Mitter, S.K., Optimal control and nonlinear filtering for nondegenerate diffusion processes, Stochastics, 8, 1982, pp. 63-77.

[8] Lions, P.L.L., Generalized solutions to Hamilton-Jacobi equations, Pitman, 1982.

[9] Watanabe, S., Stochastic Differential Equations and Malliavin Calculus, Tata Inst. of Fundamental Research, Springer, 1984.

[10] Strook, D.W., The Malliavin calculus - a functional analytic approach, J. Funct. Anal., 44, 1981, pp. 212-257.

[11] Zakai, M., The Malliavin calculus, Acta Appl. Math., 3, 1985, pp. 175-207.

[12] Zeitouni, O. and Dembo, A., An existence theorem and some properties of maximum a posteriori estimators of trajectories of diffusions, Stochastics, Vol. 23, pp. 197-218, 1988.

[13] Rund, H., The Hamilton-Jacobi Theory in the Calculus of Variations, Van Nostrand, Princeton, N.J. (1966).

[14] Fleming, W.H., The Cauchy problem for a nonlinear first order partial differential equation, J. Diff. Eqns., 5, 1969, pp. 515-530.

[15] Fleming, W.H. and Rishel, R., Deterministic and Stochastic Optimal Control, Springer, Berlin, 1975.

[16] Bouleau, N. and Hirsch, F., Proprietes d'absolue continuite dans les espaces de Dirichlet et application aux equations differentielles stochastiques. In seminaire de probabilites xx, pp. 131-161, Lecture notes in Mathematics #1204, Springer, 1986.

[17] Benton, S.H., The Hamilton Jacobi Equation, Academic Press, New York, 1977.

[18] Ikeda, N. and Watanabe, S., An Introduction to Malliavin Calculus, Proc. Toniguchi Symposium, K. Ito Ed., North Holland, 1984.

Appendix

In this appendix, we bring a short exposition of the classical method of characteristics as needed in section 3. The exposition follows the concise exposition in [8, section 1.2]. Another more complete reference is [13].

Let $H(x,t,p)$ be C^1 in x, C^1 in p, and further assume that for fixed x,p, $dH(x,t,p)$ exists in the Ito sense. Also, let $p(x)$ be a C^1 function. Define

$$X'(t) = \frac{\partial H}{\partial p}(X,t,P), \qquad X(0) = x \tag{A.1a}$$

$$P'(t) = -\frac{\partial H}{\partial x}(X,t,P), \quad P(0) = D\phi(x) \tag{A.1b}$$

$$U'(t) = P\frac{\partial H}{\partial p}(X,t,P) + q_t, \quad U(0) = \phi(x) \tag{A.1c}$$

$$dq_t = -d_tH(X,t,P), \; q_0 = -H(x,0,D\phi(x)) \tag{A.1d}$$

where $d_tH(X,t,P)$ denotes the Ito differential of H with X,P fixed. We have then:

Lemma 1: Assume that for some strip $[0,\tau]$, (A.1a) and (A.1b) define a C^1 diffeomorphism $x \to X(t,x)$. Then

$$u(x,t) = U(X^{-1}(x,t) \tag{A.2}$$

is a solution of

$$u_t + H(x,t,u_x) = 0; \quad u(0,x) = \phi(x) \tag{A.3}$$

for all $t \varepsilon [0,\tau)$.

<u>Proof</u>. First, note that by our construction,

$$E = q_t + H(X,t,P) = 0 \quad \forall t \ \varepsilon \ [0,\tau], \quad x\varepsilon\mathbb{R} \tag{A.4}$$

since $dE = dq_t + d_tH + H_x \frac{\partial x}{\partial t} dt + H_p \frac{\partial P}{\partial t} dt$

$$= (H_xH_p - H_pH_x)dt = 0, \quad E(x,0) = H(x,0,D\phi(x))-H(x,0,h,D\phi(x)) = 0$$

Therefore, one also has $\frac{\partial E}{\partial x} = \frac{\partial E}{\partial t} = 0$. Next, we compute $\frac{\partial U}{\partial x}$: By (A.1c), we have

$$U(X^{-1}(x,t),t) = U(X^{-1}(x,t),0) + \int_0^t P(s,X^{-1}(t,x))\frac{\partial}{\partial S}(X(s,X^{-1}(t,x))ds + \int_0^t q_s ds$$

$$= U(X^{-1}(x,t),0) + P(t,X^{-1}(t,x))x - P(0,X^{-1}(t,x))X^{-1}(t,x)$$

$$+ \int_0^t q_s ds + \int_0^t X(s,X^{-1}(t,x))H_x(X(s,X^{-1}(t,x)), \ s, \ P(s,X^{-1}(t,x))ds \tag{A.5}$$

One has therefore:

$$\frac{\partial U}{\partial x} = P(t,X^{-1}(t,x)) + x\,\frac{\partial P(t,X^{-1}(t,x))}{\partial x} + \frac{\partial U(X^{-1}(x,t),0)}{\partial x} \tag{A.6}$$

$$- \frac{\partial}{\partial x}\,(P(t,X^{-1}(t,x)) + \int_0^t \frac{\partial}{\partial x}\,q_s ds + \int_0^t \frac{\partial X}{\partial x}\,H_x ds + \int_0^t X\,\frac{\partial H_x}{\partial x}\,ds$$

where we have omitted the arguments of several functions.

Note that $\frac{\partial X}{\partial x}\,Hx = \frac{\partial H}{\partial x} - Hp\,\frac{\partial P}{\partial x}$; also, $\frac{\partial H_x}{\partial x} = -\frac{\partial \dot{P}}{\partial x} = -\frac{\partial}{\partial s}\frac{\partial P}{\partial x}$.

Substituting in (A.6), one has

$$\frac{\partial U}{\partial x} = P(t,X^{-1}(t,x)) + x\,\frac{\partial P}{\partial x} + \frac{\partial U(X^{-1}(x,t),0)}{\partial x} - \frac{\partial}{\partial x}\,(P_0\cdot X^{-1}(t,x))$$

$$+ \int_0^t \frac{\partial E_s}{\partial x}\,ds - \int_0^t (\frac{\partial P}{\partial x}\,\dot{X} + (\frac{\partial \dot{P}}{\partial x})X)ds$$

$$= P(t,X^{-1}(t,x)) + x\,\frac{\partial P}{\partial x} + \frac{\partial U(X^{-1}(x,t),0)}{\partial x} - \frac{\partial}{\partial x}\,(P_0X^{-1}(t,x))$$

$$- x\,\frac{\partial P}{\partial x} + X^{-1}(x,t)\,\frac{\partial P_0}{\partial x} = P(t,X^{-1}(t,x)) \tag{A.7}$$

Where the last equality follows from the boundary conditions of (A.1b) and (A.1c).

In order to prove the lemma, it is therefore enough to check that

$$\frac{\partial U}{\partial t} + H(X^{-1}(t,x),\ t,\ P(t,x^{-1}(t,x)) = 0 \qquad (A.8)$$

The computation is similar to the above and therefore ommitted. [ ]

Remarks 1): From the above characterization of u(x,t), it is easy to see that if the boundary data and the Hamiltonian in (A.3) are C^∞ w.r.t. the space variable x, and if the map $x \rightarrow X(t,x)$ is C^∞ in $x \varepsilon \mathbb{R}$, so is the solution u(t,x).

2): C^2 solutions (in the space variable) of the Hamilton-Jacobi equation (A.3) are unique, as long as they exist everywhere, i.e. as long as $x \rightarrow X(t,x)$ is a diffeomorphism. For a general proof, c.f. [13, pg. 59, theorem]. An alternative proof, at least for convex (w.r.t. p) Hamiltonians would proceed as follows: To the Hamilton-Jacobi equation there corresponds a control problem (c.f. e.g., [8, pg. 27] or [14]). Whenever, the solution of the H-J equation is C^1, it is the value function of the control problem and therefore unique. For details, c.f. [14], [15].

3): We can prove the following corollary, whch turns out to be useful in lemmas (3.4), (3.5), (3.6):

Corollary A1: Assume one has $H(X,t,P) = Pf(t,x) + g(t,x)$ where f(t,x) and g(t,x) satisfy

$$|f(t,x)| \leq K_1(|x| + 1), \quad |g(t,x)| \leq K_2(|x|^{\alpha}+1, \quad f_x(t,x) \geq -K_3(1+|x|^2 \qquad \text{(A.9)}$$

Moreover, assume that $\phi(x) \equiv 0$. Then

$$|u(t,x)| \leq KK_2(|x|^{\alpha} + 1) \qquad \text{(A.10)}$$

for some K independent of K_2, in some strip $[\theta,\tau]x\mathbb{R}$.

<u>Pf</u>. Applying Lemma 1, one has

$$|U^x(t)| \leq |U^x(0) + \int_0^t \dot{U}(s)ds| \leq K|g(t,x)| \leq KK_2(1+|x|^{\alpha} \qquad \text{(A.11)}$$

By our bound on f_x, there exists a strip where X(x,t) is a diffeomorphism. In this strip, $\dot{X} = f(t,x)$ yields that $|X(x,t)| \leq K(|x|+1)$ substituting in (A.1) one get (A.10).