Stochastic Processes and their Applications 19 (1985) 327-339
North-Holland

REVERSE TIME DIFFUSIONS

Robert J. ELLIOTT,
Department of Pure Mathematics, University of Hull, Hull, HU6 7RX, England

Brian D.O. ANDERSON
Department of Systems Engineering, Australian National University, Canberra A.C.T. 2601, Australia

Received 24 October 1983
Revised 25 September 1984

The paper considers a diffusion evolving in $\mathbb{R}^n$. The stochastic differential equations giving the same process, but with the time parameter evolving in the negative direction, are obtained under a certain integrability hypothesis when the diffusion has a density function on a time varying submanifold of $\mathbb{R}^n$.

diffusion * stochastic differential equation * Brownian motion * time-reversedprocess * Markov process * Gaussian process * transition density

1. Introduction

This paper is a continuation of work commenced in [1] and considers a diffusion in $\mathbb{R}^n$ given as the solution of a family of stochastic differential equations. The problem discussed is to suppose the direction of the time parameter is reversed, that is, time evolves in a negative direction and the same diffusion process is observed, with the filtration generated by the reversed process.

There is an extensive literature on time reversal of Markov processes. See [3, 16]. The Markov property states that past and future are conditionally independent given the present state, so the reverse time process is again Markov. However, simple examples, (see [16]), show that, for example, the strong Markov property is not preserved by time reversal. Consequently, some basic properties are not preserved, and for diffusions it is not clear that the reverse time process is again a diffusion. This paper gives conditions under which this is so, and in that case derives the reverse time stochastic differential equations giving the reverse time diffusion.

Results of this kind have been obtained by, among others, Lindquist and Picci [10], for the stationary linear case, and by Anderson [1], under the assumptions that the transition density exists and the associated Kolmogorov equations have unique solutions. There is a rather formal derivation, [2], by Castanon, who also assumes the transition densities exist. However, there are gaps in Castanon's work, about

0304-4149/85/$3.30

which we say more later. The first work containing an approach similar to the present paper is the article by Itô [8], where a semimartinagle decomposition of a Brownian motion with respect to a reverse time filtration is discussed. This is briefly described at the beginning of the next section.

It is not too difficult, at least formally, to determine the form of the reverse time Kolmogorov equations but, as stated above, we consider the more difficult problem of obtaining the stochastic differential equation form for the reverse time diffusions. We consider the singular case and suppose the diffusion evolves on a submanifold of $\mathbb{R}^n$, on which it has a density. An integrability condition in terms of this density, Hypothesis 4.4, is also required. Following Castanon [2], we originally claimed that Theorem 4.2 established almost sure convergence. We are grateful to Mark Davis for pointing out that in fact it proves weak convergence in L^1. However, a weakly convergent set is bounded, so we were able to dispense with an awkward hypothesis used in the first draft of this paper and (using Hypothesis 4.4) obtain the reverse time decomposition of Theorem 4.6.

The first author wishes to thank the Department of Systems Engineering of the Australian National University for its warm hospitality and support as a Visiting Fellow through June to August, 1983, when this work was carried out.

2. The forward stochastic system

To motivate what follows we shall first describe the situation discussed by Itô [8]. Suppose $\{B_t\}, 0 \leq t \leq T$, is a real Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$ with $B_0 = 0$ a.s. Write $\{\mathcal{F}_t\}$ for the completion of the forward filtration $\sigma\{B_s: 0 \leq s \leq t\}$, and $\{G_t\}$ for the completion of the reverse time filtration $\sigma\{B_s: t \leq s \leq T\}$. The problem treated by Itô is whether $\{B_t\}$ is a semimartingale with respect to $\{G_t\}$ for reverse time t running from T to 0. For $s \leq t$, $E[B_s | G_t] = E[B_s | B_t] = (E[B_s B_t]/E[B_t^2])B_t = (s/t)B_t$ and Itô shows that $B_t = M_t + \int_t^T B_u/u \, du$, where M_t is a reverse time $\{G_t\}$ martingale. We shall discuss a similar problem for a more general diffusion.

Consider a process $x_t = (x_t^1, \ldots, x_t^n)$ in $\mathbb{R}^n$ which is the solution of a Stratronovich system of stochastic differential equations with $0 \leq t \leq T$ and with driving processes $W_t^1, \ldots, W_t^m$ which are independent Brownian motions on a probability space $(\Omega, \mathcal{F}, P)$. At $t = 0$ we suppose the initial condition x_0 is an $\mathbb{R}^n$-valued random variable which is independent of $W_t^k - W_s^k$ for $0 \leq s \leq t \leq T$, $1 \leq k \leq m$.

The coefficient functions a^i, g^{ik} are C^∞, and we suppose they satisfy growth conditions of the form $\sum_i |a^i(x)|^2 + \sum_{i,j} |g^{ij}(x)|^2 \leq K^2(1 + |x|^2)$, so that (2.1) has a unique solution for each initial condition x_0. Clearly conditions weaker than C^∞ would suffice here. Consider the vector fields on $\mathbb{R}^n$:

$$A(x) = \sum_{i=1}^n a^i(x) \frac{\partial}{\partial x_i}, \qquad X_k(x) = \sum_{i=1}^n g^{ik}(x) \frac{\partial}{\partial x_i}, \quad 1 \leq k \leq m.$$

Then, formally, the system can be written:

$$dx_t = A(x_t)\,dt + \sum_{k=1}^{m} X_k(x_t)\circ d W_t. \tag{2.1}$$

Recalling the $\circ\, dW$ represent Stratonovich integrals, for any C^2 function $F:\mathbb{R}^n \to \mathbb{R}$,

$$F(x_t) = F(x_0) + \int_0^t A(x_s)F(x_s)\,ds + \sum_{k=1}^{m} \int_0^t X_k(x_s)F(x_s)\circ d W_s^k. \tag{2.2}$$

Write $L = A + \frac{1}{2}\sum_{k=1}^{m} X_k^2$. Then the Itô integral form of (2.2) is

$$F(x_t) = F(x_0) + \int_0^t (LF)(x_s)\,ds + \sum_{k=1}^{m} \int_0^t (X_kF)(x_s)\,d W_s^k. \tag{2.3}$$

Notation 2.1. For $0 \leq t \leq T$, $\mathscr{F}_t$ will denote the completion of the forward time filtration

$$\sigma\{x_s: 0 \leq s \leq t\} = \sigma\{x_0, W_s: 0 \leq s \leq t\}.$$

G_t will denote the completion of the reverse time filtration

$$\sigma\{x_s: t \leq s \leq T\} = \sigma\{x_T, W_u - W_v: t \leq v \leq u \leq T\}.$$

(This equality may be justified using (4.2) below.)

3. The augmented Lie algebra

Consider now an augmented system with state $x_t^* = (x_t, t) \in \mathbb{R}^{n+1}$. The associated vector fields, now considered as vector fields on $\mathbb{R}^{n+1}$, are

$$A + \frac{\partial}{\partial t} \quad \text{and} \quad X_k, \quad 1 \leq k \leq m.$$

Write $H = LA(A + \partial/\partial t, X_1, \ldots, X_k)$ for the Lie algebra of vector fields (on $\mathbb{R}^{n+1}$), generated by $A + \partial/\partial t, X_1, \ldots, X_k$.

Hypothesis 3.1. Suppose H is nonsingular. That is, suppose the dimension of H is constant for all $x^* = (x, t) \in \mathbb{R}^{n+1}$. This dimension is the *rank* of H.

By the global version of Frobenius' complete integrability theorem, (cf. Hermann [7]), there is a maximal connected integral manifold M of H containing x_0^*. If rank $H = r+1$ then the dimension of M is $r+1$.

Let M_t denote the slice of M at constant t. Clark, [4], states that because $\{t = \text{constant}\}$ is not an integral manifold of H, M_t has dimension r. He proves, [4, Proposition 5.1]:

Theorem 3.2. *If rank $H = r+1$, then M_t is an r-dimensional manifold and the process x_t satisfies*

$$P(x_t \in M_t \text{ for } 0 \leq t \leq T) = 1.$$

That is, the process lives on M.

Remarks 3.3. We now wish to discuss integration over M_t. The treatment in Gelfand and Shilov [6, p. 239], will be followed.

Because the manifold M_t is a submanifold of $\mathbb{R}^n$ there is certainly an immersion of M_t into $\mathbb{R}^n$. Consequently, the Riemann structure of $\mathbb{R}^n$ pulls back to give a Riemann structure on M_t, and so a local volume element. (See Loomis and Sternberg [1].) This is presumably what D. Elliott [5], means when he talks of the "r-dimensional measure induced on M by the $(n+1)$ dimensional Lebesgue measure".

However, following [6], consider a partition of unity of $\mathbb{R}^n$, $\{a \in C_0^\infty(\mathbb{R}^n)\}$, such that in the support $U_a \subset \mathbb{R}^n$ of a we can suppose (by Frobenius' Theorem), that there is a local coordinate system $y = (y_1, \ldots, y_n)$ such that the set $M_t \cap U_a$ is the set on which $y_{r+1}, \ldots, y_n$ are zero. Consequently, using the notation of Gelfand and Shilov [6], for $\varphi \in C^\infty(U_a)$ the distribution $\delta(M_t)$ on M_t is

$$\langle \varphi, \delta(M_t) \rangle = \int_{\mathbb{R}^n} \varphi(x)\delta(M_t) = \int_{\mathbb{R}^r} \Phi(y_1, \ldots, y_r, 0, \ldots, 0)\, \mathrm{d}y_1 \cdots \mathrm{d}y_r.$$

Here $\Phi(y_1, \ldots, y_n) = \varphi(y_1, \ldots, y_n) D(\substack{x\\y})$ where $D(\substack{x\\y}) = \det(\partial x_i / \partial y_j)$ is the Jacobian determinant of the transformation expressing the x_i in the terms of the y_j. The r-form $D(\substack{x\\y})\, \mathrm{d}y_1, \ldots, \mathrm{d}y_r$ is an expression for the volume element on M_t in terms of this chart; clearly it transforms using the Jacobian, as in the treatment of Loomis and Sternberg [11].

Hypothesis 3.4. Suppose that the probability law of the process x_t has a smooth density on M_t, which is absolutely continuous with respect to the above measure induced on M_t by the Riemann structure. That is, we suppose there is a C^∞ function $p(t, x)$ defined on M_t such that the density of the probability law of x_t is

$$p(t, x)\delta(M_t).$$

Indeed, in the terminology of Gelfand and Shilov, [6], p. 239, p is called a density on M.

Remark 3.5. The existence of smooth transition densities on M is discussed in D. Elliott's paper [5], where the hypoellipticity of the operator $-(\partial/\partial t) + L$ is related to the nonsingularity of H.

Lemma 3.6. *If Z is a smooth vector field on M_t then $Z\delta(M_t) = 0$.*

Proof. Consider any $\phi \in C_0^\infty(\mathbb{R}^n)$ and a partition of unity $\{a\}$ as above. Then

$$\langle \phi, Z\delta(M_t)\rangle = \sum_\alpha \int_{\mathbb{R}^n} (\alpha\phi)(x) Z\delta(M_t).$$

With the above choice of coordinate chart $(y_1, \ldots, y_n)$ on U_α, Z has the form

$$Z = \sum_{j=1}^r b_j(y_1, \ldots, y_r)\frac{\partial}{\partial y_j}.$$

Therefore,

$$\int_{\mathbb{R}^n} (\alpha\phi)(x) Z\delta(M_t)$$

$$= \int_{\mathbb{R}^n} (\alpha\phi)(y) D\begin{pmatrix} x \\ y \end{pmatrix}\left\{\sum_{j=1}^r b_j(y_1, \ldots, y_r)\frac{\partial}{\partial y_j}\delta(y_{r+1}, \ldots, y_n)\right\} \mathrm{d}y_1 \cdots \mathrm{d}y_r.$$

We immediately see $(\partial/\partial y_j)\delta(y_{r+1}, \ldots, y_n) = 0$, or, working one step further it equals

$$-\sum_{j=1}^r \int_{\mathbb{R}^r} \frac{\partial}{\partial y_j}\{b_j(y_1, \ldots, y_r)\Phi(y_1, \ldots, y_r, 0, \ldots, 0)\} \,\mathrm{d}y_1 \cdots dy_r,$$

which is zero because the functions have compact support. Here, $\Phi = (\alpha\phi)D$.

Remarks 3.7. In the terminology of [6], $Z\delta(M_t)$ is a multiplet layer on M_t.

4. Quasimartingale decompositions

Recall that if $f(u)$, $0 \leq u \leq T$, is a $\{G_t\}$-predictable process (in reverse time) such that $\int_0^T E[f(u)^2]\,\mathrm{d}y < \infty$ then the backward Itô integral is defined by Kunita [9], if f is continuous in probability, as

$$\int_s^t f(u)\,\hat{\mathrm{d}} W_u^k := \lim_{|\Delta|\to 0} \sum_{j=0}^{n-1} f(t_{j+1})(W_{t_{j+1}}^k - W_{t_j}^k).$$

Here $\Delta = \{s = t_0 < t_1 < \cdots < t_n = t\}$ and

$$|\Delta| = \max_k |t_{k+1} - t_k|.$$

The backward Stratonovich integral is defined as

$$\int_s^t f(u) \circ \hat{\mathrm{d}} W_u^k := \lim_{|\Delta|\to 0} \sum_{j=0}^{n-1} \tfrac{1}{2}(f(t_{j+1}) + f(t_j))(W_{t_{j+1}}^k - W_{t_j}^k).$$

The two integrals are related by the formula

$$\int_s^t f(u) \circ \hat{\mathrm{d}} W_u^k = \int_s^t f(u)\,\hat{\mathrm{d}} W_u^k + \tfrac{1}{2}(\langle f, W^k\rangle_t - \langle f, W^k\rangle_s).$$

(The definition of $\langle\ \rangle_t$ is similar to (4.5) below.)

Note that if f is Markov $\int_s^t f(u)\circ\hat{\mathrm{d}}W_u^k=\int_s^t f(u)\circ\mathrm{d}W_u^k$.

Recall from (2.1) that

$$x_t=x_0+\int_0^t A(x_s)\,\mathrm{d}s+\sum_{k=1}^m\int_0^t X_k(x_s)\circ\mathrm{d}W_s^k. \tag{4.1}$$

Our object is to obtain a reverse time Markov representation for x_t. As a first step we can write, following Castanon [2],

$$x_t=x_T-\int_t^T A(x_s)\,\mathrm{d}s-\sum_{k=1}^m\int_t^T X_k(x_s)\circ\mathrm{d}W_s^k$$

and because the integrand is Markov, this is

$$=x_T+\int_T^t(-A(x_s))\,\mathrm{d}(-s)+\sum_{k=1}^m\int_T^t X_k(x_s)\circ\hat{\mathrm{d}}W_s^k.$$

The stochastic integrals here are Stratonovich integrals; the Itô form is

$$x_t=x_T+\int_T^t\Big(A(x_s)-\tfrac{1}{2}\sum_{k=1}^m X_k^2(x_s)\Big)\,\mathrm{d}s+\sum_{k=1}^m\int_T^t X_k(x_s)\,\hat{\mathrm{d}}W_s^k. \tag{4.2}$$

However, this is not of the form required because each integral $\int_T^t X_k(x_s)\,\hat{\mathrm{d}}W_s^k$ is not a G_t-martingale.

Remarks 4.1. It is at this point that Castanon [2], makes some incompletely justified assumptions. He claims that the representation (4.2) describes x_t as a G_t-semimartingale, that is, as the sum of a (local) G_t-martingale and a process of bounded variation, and he then proceeds to determine what he calls the Doob–Meyer decomposition of

$$\sum_{k=1}^m\int_T^t X_k(x_s)\,\hat{\mathrm{d}}W_s^k$$

(in reverse time). However, the Doob–Meyer decomposition relates to supermartingales. The quantity investigated is certainly not a supermartingale. Indeed, it is not clear that each term $\int_T^t X_k(x_s)\,\hat{\mathrm{d}}W_s^k$, or even each component W_t^k of the Brownian motion, is a semimartingale with respect to the reverse time filtration G_t. Under an integrability assumption, Hypothesis 4.4 below, we are able to show below that each component W_t^k is a quasimartingale with respect to the reverse time filtration. However, first we investigate a weak limit by adapting a technique of Castanon [2].

Theorem 4.2. *Under the above hypotheses,*

$$\lim_{h\to0+}h^{-1}E[W_t^k-W_{t-h}^k\,|\,x_t]=-\Big(\operatorname{div}X_k+I_{p(t,\cdot)\neq0}\frac{X_kp(t,\cdot)}{p(t,\cdot)}\Big)(x_t)$$

weakly in $L^1(\mathbb{R}^n)$ where the density of the probabilty law of x_t on M_t is $p(t,x_t)\delta(M_t)$.

Note the right-hand side is independent of any chart because div X_k *is invariant.* (*See* [11]).

Proof. Consider the reverse time filtration G_t and suppose $0 \leq t-h \leq t \leq T$. By the Markov property,

$$E[W_t^k - W_{t-h}^k \mid G_t] = E[W_t^k - W_{t-h}^k \mid x_t] = \Lambda^h(x_t), \quad \text{say}.$$

Consider any $\phi \in C_0^\infty(\mathbb{R}^n)$.

$$E[\phi(x_t)\Lambda^h(x_t)] = E[\phi(x_t)(W_t^k - W_{t-h}^k)] = E[E[\phi(x_t)(W_t^k - W_{t-h}^k) \mid x_{t-h}]].$$

Itô's formula (see McKean [12]) then gives:

$$\begin{aligned}\phi(x_t)(W_t^k - W_{t-h}^k) &= \int_{t-h}^t \phi(x_s)\,\mathrm{d}W_s^k + \int_{t-h}^t (W_s^k - W_{t-h}^k)(L\phi)(x_s)\,\mathrm{d}s \\ &\quad + \sum_{j=1}^m \int_{t-h}^t (W_s^k - W_{t-h}^k)(X_j\phi)(x_s)\,\mathrm{d}W_s^j \\ &\quad + \int_{t-h}^t (X_k\phi)(x_s)\,\mathrm{d}s.\end{aligned}$$

Now $\int_{t-h}^t \phi(x_s)\,\mathrm{d}W_s^k$ and each of $\int_{t-h}^t (W_s^k - W_{t-h}^k)(X_j\phi)\,\mathrm{d}W_s^j$ are martingales, and so have expectation 0. Therefore,

$$\begin{aligned}&E[E[\phi(x_t)(W_t^k - W_{t-h}^k) \mid x_{t-h}]] \\ &\quad = E\left[\int_{t-h}^t (W_s^k - W_{t-h}^k)(L\phi)(x_s)\,\mathrm{d}s + \int_{t-h}^t (X_k\phi)(x_s)\,\mathrm{d}s\right].\end{aligned} \tag{4.3}$$

Now

$$\begin{aligned}&\left|E\left[\int_{t-h}^t (W_s^k - W_{t-h}^k)(L\phi)(x_s)\,\mathrm{d}s\right]\right| \\ &\quad \leq K \int_{t-h}^t E|W_s^k - W_{t-h}^k|\,\mathrm{d}s \leq K \int_{t-h}^t (E|W_s^k - W_{t-h}^k|^2)^{1/2}\,\mathrm{d}s \leq Kh^{3/2}\end{aligned}$$

where K is a uniform bound for $|(L\phi)(x)|$.

Dividing (4.3) by h and letting $h \to 0+$ we see that for all $\phi \in C_0^\infty(\mathbb{R}^n)$

$$\lim_{h\to 0^+} E[\phi(x_t)h^{-1}\Lambda^h(x_t)] = E[(X_k\phi)(x_t)]. \tag{4.4}$$

We wish to integrate the right-hand side of (4.4) by parts. From hypothesis 3.4 the probability law of x_t has density $p(t,x)\delta(M_t)$, so

$$E[(X_k\phi)(x)] = \langle X_k\phi, p\delta(M_t)\rangle.$$

As X_k has the expression

$$X_k(x)=\sum_{i=1}^n g^{ik}(x)\frac{\partial}{\partial x_i}$$

we see that X_k is a vector field on M_t and

$$\langle X_k\phi, p\delta(M_t)\rangle=\int_{\mathbb{R}^n}\left(\sum_{i=1}^n g^{ik}(x)\frac{\partial\phi}{\partial x_i}\right)p(t,x)\delta(M_t).$$

Integrating by parts, this is

$$=-\sum_{i=1}^n\int_{\mathbb{R}^n}\phi(x)\frac{\partial}{\partial x_i}(g^{ik}(x)p(t,x)\delta(M_t))$$

$$=-\int_{\mathbb{R}^n}\phi(x)((\operatorname{div}X_k)p(t,x)\delta(M)$$

$$+(X_kp)(x)\delta(M_t)+p(t,x)X_k\delta(M_t)).$$

However, from Lemma 3.6 the distribution $X_t\delta(M_t)$ is zero, so

$$\lim_{h\to 0}E[\phi(x_t)h^{-1}\Lambda^h(x_t)]$$

$$=-\int_{\mathbb{R}^n}\phi(x)((\operatorname{div}X_k)p(t,x)\delta(M_t)$$

$$+\frac{(X_kp)(x)}{p(t,x)}I_{p(t,\cdot)\neq 0}\,p(t,x)\delta(M_t))$$

$$=-E\left[\phi(x_t)(\operatorname{div}(X_k)(x_t)+I_{p(t,x_t)\neq 0}\frac{(X_kp)(t,\cdot)(x_t)}{p(t,x_t)}\right)\Bigg].$$

As this holds for all $\phi\in C_0^\infty(\mathbb{R}^n)$,

$$\lim_{h\to 0+}h^{-1}\Lambda^h(x)=-\left(\operatorname{div}X_k+I_{p(t,x)\neq 0}\frac{(X_kp)(t,\cdot)(x)}{p(t,x)}\right),\quad\text{weakly in } L^1(\mathbb{R}^n).$$

Note that this formulation is invariant.

Remarks 4.3. Write

$$\chi_t^k(x)=-\left(\operatorname{div}X_k(x)+I_{p\neq 0}\frac{(X_kp)(x)}{p(t,x)}\right).$$

Then what we should like to conclude from the above calculation is that there is a decomposition

$$W_t^k=\bar{W}_t^k-\int_T^t\chi_u^k(x_u)(x_u)\,\mathrm{d}u,$$

where $\bar{W}_t^k$ is a G_t reverse-time martingale, that is, for $h>0$, $E[\bar{W}_{t-h}^k\,|\,G_t]=\bar{W}_t^k$.

We make the following assumption:

Hypothesis 4.4. For any $\delta > 0$, $1 \leq k \leq m$,

$$\chi_t^k(x) \in L^1([\delta, T] \times \Omega),$$

that is

$$p(t, x) \operatorname{div} X_k(x) + (X_k p(t, x)) \in L^1([\delta, T] \times \mathbb{R}^n).$$

If we adapt the characterization obtained by Stricker in [14, Theorem 2], we have in the present situation:

Theorem 4.5. *A process $\{z_t\}$, $t \in [\delta, T]$, is a reverse time G_t-quasimartingale if, for $h > 0$,*

$$\int_\delta^T E|E[z_t - z_{t-h} | G_t]| \, dt = O(h).$$

Theorem 4.6. *Under hypothesis 4.4 the process $\{W_t^k\}$ is a reverse time G_t-quasimartingale with a decomposition*

$$W_t^k = \bar{W}_t^k + \int_t^T \chi_u^k(x_u) \, du. \tag{4.5}$$

Here $\{\bar{W}_t^k\}$ is a reverse time G_t-Brownian motion, that is $\{\bar{W}_t^k\}$ is a reverse time G_t-martingale with $\bar{W}_T^k = W_T^k$, and $\langle \bar{W}^k \rangle_t - \langle \bar{W}^k \rangle_2 = t - s$.

Proof. Adding integration to the steps between (4.2) and (4.3) in the proof of Theorem 4.3 we see that for every $\phi \in C_0^\infty(\mathbb{R}^n)$

$$\lim_{h \to 0^+} \int_\delta^T E[\phi(x_t) h^{-1} \Lambda^h(x_t)] \, dt$$
$$= \int_\delta^T E[(X_k \phi)(x_t)] \, dt = \int_\delta^T E[\phi(x_t) \chi_t^k(x_t)] \, dt.$$

Consequently, if the limit function $\chi_t^k \in L^1([\delta, T] \times \Omega)$, we have

$$\lim_{h \to 0+} h^{-1} \Lambda^h(x_t) = \chi_t^k(x_t)$$

weakly in $L^1([\delta, T] \times \Omega)$. However, a weakly convergent set is bounded, so $\{h^{-1} \Lambda^h(x_t)\}$ is bounded in $L^1([\delta, T] \times \Omega)$.

That is, $\int_\delta^T E|E[W_t^k - W_{t-h}^k | G_t]| dt$ is $O(h)$.
By Theorem 4.5 $\{W_t^k\}$ is a G_t quasimartingale with a decomposition

$$W_t^k = \bar{W}_t^k + A_t.$$

Here $\{\bar{W}_t^k\}$ is a reverse time G_t martingale and $\{A_t\}$ is a G_t-predictable process of bounded variation. A quasimartingale is the difference of two supermartingales and the variables $\{W_t^k\}$, $\delta \leq t \leq T$, give a process in class D. $\{A_t\}$ is the difference of the predictable increasing processes in their Doob–Meyer decompositions and so, as in Meyer [13, VII T. 29], for each $t \in [\delta, T]$, A_t is the weak limit in $L^1(\Omega)$ as $h \to 0+$ of $\int_t^T h^{-1}\Lambda^h(x_u)\,du$. Consequently, for any $\theta \in L^\infty(\Omega)$ and $t \in [\delta, T]$, with $\langle\ \rangle$ denoting the inner product between $L^\infty(\Omega)$ and $L^1(\Omega)$,

$$\left\langle \theta, \int_t^T \chi_u^k\,du - A_t \right\rangle$$

$$= \left\langle \theta, \int_t^T (\chi_u^k - h^{-1}\Lambda^h(x_u))\,du \right\rangle + \left\langle \theta, \int_t^T h^{-1}\Lambda^h(x_u)\,du - A_t \right\rangle.$$

The limit is zero, because the first $\langle\ \rangle$ on the right has limit zero as $h \to 0+$, using weak convergence in $L^1([\delta, T] \times \Omega)$. Similarly, the second $\langle\ \rangle$ on the right has limit zero as $h \to 0+$, using weak convergence in $L^1(\Omega)$. Consequently we have a decomposition

$$W_t^k = \bar{W}_t^k + \int_t^T \chi_u^k(x_u)\,du.$$

For $\delta \leq s \leq t \leq T$ the quadratic variation process

$$\langle W^k \rangle_t - \langle W^k \rangle_s = \lim \sum_{i=1}^N (W_{t_{i+1}}^k - W_{t_i}^k)^2 = \lim \sum_{i=1}^N (\bar{W}_{t_{i+1}}^k - \bar{W}_{t_i}^k)^2 \tag{4.6}$$

is independent of the filtration, where the limit (in probability) is taken over partitions $s = t_0 < t_1 < \cdots < t_N = t[s, t]$. Consequently, we have that

$$\langle W^k \rangle_t - \langle W^k \rangle_s = t - s = \langle \bar{W}^k \rangle_t - \langle \bar{W}^k \rangle_s,$$

Therefore, $\{\bar{W}_t^k\}$ is a reverse time G_t-Brownian motion and the Theorem is proved.

Notation 4.7. The equality $\stackrel{r}{=}$ below will indicate the processes are considered in reverse time and the stochastic integrals are the corresponding backward stochastic integrals.

Using equations (4.2) and (4.6) we can now write down the reverse time stochastic differential equations describing the process.

Theorem 4.8. *For* $t \in]0, T]$,

$$x_t \stackrel{r}{=} x_T + \int_T^t \left(A(x_s) - \tfrac{1}{2}\sum_{k=1}^m X_k^2(x_s) - \sum_{k=1}^m X_k(x_s)\chi_s^k(x_s)\right) ds$$

$$+ \sum_{k=1}^m \int_T^t X_k(x_s)\,\hat{\mathrm{d}}\bar{W}_s^k.$$

Here the $\bar{W}_t^k = W_t^k + \int_T^t \chi_u^k(x_u)\,\mathrm{d}u$, $1 \leq k \leq m$, *are the above Brownian motion processes for* $t \in]0, T]$, *with respect to the filtration* G_t.

Proof. From (4.2),

$$x_t = x_T + \int_T^t \Big(A(x_s) - \tfrac{1}{2}\sum_{k=1}^m X_k^2(x_s)\Big)\,\mathrm{d}s + \sum_{k=1}^m \int_T^t X_k(x_s)\,\mathrm{d}W_s^k.$$

Therefore,

$$x_t \overset{r}{=} x_T + \int_T^t \Big(A(x_s) - \tfrac{1}{2}\sum_{k=1}^m X_k^2(x_s) - \sum_{k=1}^m X_k(x_s)\chi_s^k(x_s)\Big)\,\mathrm{d}s$$

$$+ \sum_{k=1}^n \int_T^t X_k(x_s)\,\hat{\mathrm{d}}\bar{W}_s^k.$$

Remark 4.9. It is trivial to verify that the main result of [1] follows by identifying M_t with $\mathbb{R}^n$. Because we have constructed the reverse time diffusion by showing the forward-time trajectories can be described by reverse-time equations, the following result, proved using a technique similar to that of Clarke [4], is almost a triviality.

Lemma 4.10. *The process* x_t *remains on* M *for* t *evolving in reverse time.*

Proof. Consider any point $x_s^* = (x_s, s)$ of the (augmented) forward diffusion. Then M_t with $\mathbb{R}^n$. Because we have constructed the reverse time diffusion by showing the $(y_1, \ldots, y_{n+1})$ in $\mathbb{R}^{n+1}$, with domain U containing x_s^*, such that the sets on which $y_{r+1}, \ldots, y_{n+1}$ are constant are the integral manifolds of H in U. Therefore,

$$A(y_{r+j}(x_t^*)) = 0, \quad X_k(y_{r+j}(x_t^*)) = 0, \qquad 1 \leq j \leq n-r+1,$$

$$\text{for } t \leq s \text{ and } x_t^* \text{ in } U.$$

Consequently, on U, for $r+1 \leq i \leq n+1$,

$$\mathrm{d}y_i(x_t^*) \overset{r}{=} A(y_i(x_t^*))\,\mathrm{d}t - \tfrac{1}{2}\sum_{k=1}^m X_k^2(y_i(x_t^*))\,\mathrm{d}t - \sum_{k=1}^m \chi^k(x_s)X_k(y_i(x_t^*))\,\mathrm{d}t$$

$$+ \sum_{k=1}^m X_k(y_i(x_t^*))\,\mathrm{d}\bar{W}_t^k = 0.$$

That is, on U, $y_{r+1}, \ldots, y_{n+1}$ stay constant along x_t^* for the reverse time flow, and so the process remains on M.

Remark 4.11. It might be hoped that the local formulation used in the above Lemma would indicate how the methods of [1] could be adapted to the present situation. For example, on the open set $U \subset \mathbb{R}^{n+1}$ the coordinates $y_{r+1}, \ldots, y_{n+1}$ of the local coordinate system are constant along the trajectories x_t^*, and, in the forward-time direction, the coordinates y_i, $1 \leq i \leq r$, satisfy

$$\mathrm{d}y_i(x_t^*) \overset{f}{=} L(y_i(x_t^*))\,\mathrm{d}t + \sum_{k=1}^{m} X_k(y_i(x_t^*))\,\mathrm{d}W_t^k,$$

where the stochastic integrals are Itô integrals. Write $\hat{y}(x)$ for the components $(y_1(x), \ldots, y_r(x))$. The coefficients L and X_k above are, in general, only defined for $\hat{y}$ in the image $V = \hat{y}(U)$ of U.

The method of [1] assumes the Kolmogorov equations have unique solutions, and at time t the density function $p(t, x)$ is concentrated on the slice M_t. Unfortunately, because M_t is not, in general, time invariant it is not clear what form the Kolmogorov equations, or even their local forms on V, should take. Consequently this approach does not appear too promising in the present situation.

Linear systems 4.12. Consider, for $0 \leq t \leq T$, the linear system

$$\mathrm{d}x = Ax\,\mathrm{d}t + B\,\mathrm{d}W.$$

Here again $x \in \mathbb{R}^n$, $W_t = (W_t^1, \ldots, W_t^m)$ is an m-dimensional Brownian motion and A and B are constant matrices of appropriate dimensions. The initial condition is a gaussian random variable x_0 independent of W_t, $t > 0$, and of zero mean, so

$$x_t = \mathrm{e}^{At}x_0 + \int_0^t \mathrm{e}^{A(t-s)}B\,\mathrm{d}W_s. \tag{4.7}$$

Suppose rank $(B, AB, \ldots, A^{n-1}B) = r$; then we see from (4.7) that x_t lies on a moving r-dimensional hyperplane M_t which passes through $\mathrm{e}^{At}x_0$. The covariance of x_t is given by

$$\operatorname{Cov} x_t = P(t) = P_0 + \int_0^t \mathrm{e}_{A_s} BB^*\,\mathrm{e}^{A_s^*}\,\mathrm{d}s.$$

Write $P(t)^{\#}$ for the Moore-Penrose pseudo inverse and $\hat{P}(t)$ for the determinant of the restriction of $P(t)$ to M_t. Then the density function of the probability law of x_t on the hyperplane M_t is given by

$$p(t, x) = \frac{1}{(2\pi)^{r/2}\hat{P}(t)} \exp(-\tfrac{1}{2}x^* P(t)^{\#} x).$$

The Hypothesis 4.4 is satisfied and the reverse time Brownian motion is then $\bar{W}_t$, where

$$\mathrm{d}\bar{W}_t = \mathrm{d}W_t - B^* P(t)^{\#} x_t\,\mathrm{d}t.$$

References

[1] B.D.O. Anderson, Reverse time diffusion equation models, Stochastic Process. Appl. 12 (1982) 313–326.

[2] D.A. Castanon, Reverse time diffusion processes. IEEE Trans. Inf. Theory, 28 (1982), 953–956.
[3] K.L. Chung and J.B. Walsh, To reverse a Markov process. Acta. Math. 123 (1969), 225–251.
[4] J.M.C. Clarke, An introduction to stochastic differential equations on manifolds, in: D.Q. Mayne and R.W. Brockett, eds., Geometric Methods in System Theory (Reidel, Dordrecht, Holland, 1973).
[5] D. Elliott, Diffusions on manifolds arising from controllable systems, in: D.Q. Mayne and R.W. Brockett, eds., Geometric Methods in System Theory (Reidel, Dordrecht, Holland, 1973).
[6] I.M. Gelfand and G.E. Shilov, Generalized Functions. Vol. 1 (Academic Press, New York, 1964).
[7] R. Hermann, Differential geometry and the calculus of variations (Academic Press, New York, 1968).
[8] K. Itô, Extension of Stochastic Integrals (Proc. International Symposium Kyoto, 1976, Kinokuniya, Tokyo, Wiley, New York, 1978), 95–109.
[9] H. Kunita, Ón backward stochastic differential equations, Stochastics 6 (1982) 293–313.
[10] A. Lindquist and G. Picci, On the stochastic realization problem, SIAM J. Control Opt. 17 (1979) 365–389.
[11] L.H. Loomis and S. Sternberg, Advanced Calculus (Addison-Wesley, Reading, MA, 1968).
[12] H.P. McKean, Stochastic Integrals (Academic Press, New York, 1969).
[13] P.A. Meyer, Probabilités et Potential (Hermann Cie, Paris 1966).
[14] C. Stricker, Une characterization des quasimartingales, Séminaire de Probabilités IX, Lecture Notes in Math. 465 (Springer-Verlag, Berlin, 1975).
[15] G. Stricker, Quasimartingales, martingales locales et filtration naturelle, Z. Warsch. Verw. Geb. 63 (1977) 55–63.
[16] D. Williams, Diffusions, Markov Processes and Martingales (J. Wiley and Sons, Chichester, 1979).