PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 4, Fasc. 1 (1984), p. 91-107

A STUDY OF A ONE-DIMENSIONAL BILINEAR DIFFERENTIAL MODEL FOR STOCHASTIC PROCESSES

BY

A. LE BRETON (GRENOBLE) AND M. MUSIELA (WROCŁAW)

Abstract. This paper* is concerned with a study of a one-dimensional bilinear differential model for stochastic processes in continuous time. We provide conditions for second-order and strict-sense stationarities of the state process. We obtain a linear representation of the state process, derive the optimal linear filter, and investigate its asymptotic behaviour. We consider the problem of parameter estimation for the autonomous version of the model. By the use of the quadratic variation of the process we compute the diffusion coefficient parameters. In the reduced model, under the additional assumption that the parameters of the diffusion coefficient are known, we use the maximum likelihood method and the method of moments in order to estimate the drift coefficient parameters. We prove consistency and asymptotic normality of the estimates.

1. Introduction. Bilinear deterministic models for dynamical systems in discrete time (see, e.g., [6]) and in continuous time (see, e.g., [4]) have been intensively studied. Analogous stochastic models have also been considered in time series analysis (see, e.g., [5]) and in theory of stochastic differential equations (see, e.g., [2]).

In this paper we try to develop a probabilistic and statistical study of a one-dimensional stochastic model which is given by a bilinear differential equation (in the Itô sense) of the form

$$(1.1)\quad dX_t = [A(t)X_t + a(t)]dt + [B(t)X_t + b(t)]dW_t, \qquad t \geqslant 0,\ X_0 = X(0),$$

where $W = (W_t;\ t \geqslant 0)$ is a standard brownian motion in $\boldsymbol{R}$ defined on some basic probability space $(\Omega, \mathscr{A}, (\mathscr{A}_t)_{t \geqslant 0}, P)$, the deterministic functions A, a,

* Work done during a stay of M. Musiela in Laboratoire I. M. A. G. Grenoble.

B, and b are measurable and satisfy for every $T > 0$ the conditions

$$\int_0^T |A(t)|dt < \infty, \quad \int_0^T |a(t)|dt < \infty, \quad \int_0^T |B(t)|^2 dt < \infty, \quad \int_0^T |b(t)|^2 dt < \infty,$$

and the initial state $X(0)$ is a random variable, defined on $(\Omega, \mathscr{A}, P)$, which is independent of W and admits an expectation $m(0)$ and a variance $K(0)$.

Section 2 is devoted to the analysis of the state process $(X_t; t \geqslant 0)$ of model (1.1). We compute its mean and covariance functions, we obtain a linear representation of $(X_t; t \geqslant 0)$ with respect to a wide-sense Wiener process (cf. [11]), and we give also conditions for second-order and strict-sense stationarities in model (1.1) (cf. [7]).

In Section 3, by the use of the linear representation and results in [11], we provide the equations for the optimal linear filter of $(X_t; t \geqslant 0)$; moreover, under the assumption that the state process is second-order stationary, we study the behaviour and the stability of the filter error.

In Section 4 we deal with the problem of parameter estimation in the autonomous version of the model when a strict-sense stationary solution exists, the state process being observed in continuous time. We use a two-step procedure (cf. [1], [3], and [8]): first the diffusion coefficient parameters are estimated by the use of the quadratic variation of the observed process and then the maximum likelihood method and the method of moments provide consistent and asymptotically normal estimates of the drift coefficient parameters (cf. [10]).

2. Properties of the state process. In this section we summarize the properties of the state process generated by model (1.1) which will be used in the next parts. For more details we refer to [9], § 2, 3A and 4A.

A. *Moments and linear representation.* The existence and uniqueness of the solution process $(X_t; t \geqslant 0)$ of equation (1.1) is ensured by general results on stochastic differential equations (see, e.g., [11], Ch. 4) and the conditions listed in Section 1. By the Itô formula (see, e.g., [2], Theorem 8.4.2, [9], Theorem 2.1) the process $(X_t; t \geqslant 0)$ is given by

$$X_t = \Phi_t \{X(0) + \int_0^t \Phi_s^{-1} [a(s) - B(s)b(s)] ds + \int_0^t \Phi_s^{-1} b(s) dW_s\}, \quad t \geqslant 0,$$

where

$$\Phi_t = \exp \{\int_0^t [A(s) - \tfrac{1}{2} B^2(s)] ds + \int_0^t B(s) dW_s\}, \quad t \geqslant 0.$$

Denote by $(\Psi_t; t \geqslant 0)$ the mean function of $(\Phi_t; t \geqslant 0)$, i.e.

$$\Psi_t = \exp \{\int_0^t A(s) ds\}, \quad t \geqslant 0.$$

The second-order structure of $(X_t;\, t \geqslant 0)$ is described in the following

LEMMA 2.1. *The mean function* $(m_t = \mathrm{E}X_t,\, t \geqslant 0)$ *is given by*

$$(2.1)\qquad m_t = \Psi_t\Big[m(0)+\int_0^t \Psi_s^{-1}a(s)ds\Big], \qquad t \geqslant 0.$$

The covariance function $\big(K(t, s) = \mathrm{E}[X_t - m_t)(X_s - m_s)],\, t \geqslant 0,\, s \geqslant 0\big)$ *is equal to*

$$(2.2)\qquad K(t, s) = \Psi_t \Psi_s^{-1} K_s, \qquad t \geqslant s,$$

where the variance function $\big(K_t = K(t, t),\, t \geqslant 0\big)$ *is expressed by*

$$(2.3)\qquad K_t = \exp\Big\{\int_0^t [2A(u)+B^2(u)]du\Big\}\Big\{K(0)+ \\ +\int_0^t \exp\Big\{-\int_0^s [2A(u)+B^2(u)]du\Big\}[B(s)m_s+b(s)]^2 ds\Big\}, \qquad t \geqslant 0.$$

Proof. Formula (2.1) follows immediately from (1.1). The derivation of equations (2.2) and (2.3) can be based on the representation

$$X_t - m_t = \Psi_t V_t, \qquad t \geqslant 0,$$

where

$$dV_t = \{B(t)V_t + \Psi_t^{-1}[B(t)m_t + b(t)]\}dW_t,\ t \geqslant 0, \qquad V_0 = X(0) - m(0).$$

By simple computations we get

$$(2.4)\qquad \mathrm{E}(V_t V_s) = K(0) + \int_0^s \{B^2(u)\mathrm{E}(V_u^2) + \Psi_u^{-2}[B(u)m_u + b(u)]^2\}du, \\ t \geqslant s \geqslant 0.$$

Setting $t = s$, multiplying by Ψ_s^2, and noting $\Psi_u^2 \mathrm{E} V_u^2 = K_u$, we obtain

$$K_s = \Psi_s^2\Big\{K(0) + \int_0^s \Psi_u^{-2}[B^2(u)K_u + \{B(u)m_u + b(u)\}^2]du\Big\}, \qquad s \geqslant 0.$$

This leads to (2.3) and, moreover, together with (2.4), shows that $\Psi_s^{-2}\mathrm{E}(V_t V_s) = K_s,\ t \geqslant s$. Finally, (2.2) holds since $K(t, s) = \Psi_t \Psi_s \mathrm{E}(V_t V_s)$.

Following the ideas presented in [11], Ch. 15, we provide now a linear representation of the state process (for a detailed proof see [9], § 3A).

LEMMA 2.2. *There exists a wide-sense Wiener process* $W^* = (W_t^*;\, t \geqslant 0)$ *in* R *which is uncorrelated with* $X(0)$ *and such that the state process* $(X_t;\, t \geqslant 0)$ *admits the representation*

$$(2.5)\qquad dX_t = [A(t)X_t + a(t)]dt + \\ + [B^2(t)K_t + \{B(t)m_t + b(t)\}^2]^{1/2} dW_t^*,\ t \geqslant 0, \qquad X_0 = X(0).$$

Proof. It is easy to verify that Lemma 2.1 ensures that the assumptions of Theorem 15.2 in [11] are satisfied for the process $(X_t;\ t \geqslant 0)$; then representation (2.5) holds. Moreover, from the proof of Theorem 15.2 in [11] we can see that W^* is uncorrelated with $X(0)$.

B. *Second-order and strict-sense stationarities.* A characterization of second-order stationary of the state process is given in

LEMMA 2.3. *The process $(X_t;\ t \geqslant 0)$ generated by model* (1.1) *is second-order stationary if and only if either it is deterministic with $X(0) = m(0)$ a.s. and $A(t)m(0)+a(t) = B(t)m(0)+b(t) = 0$ a.e. or there exist two constants A and a such that*

$$Am(0)+a = 0 \tag{2.6}$$

and the following conditions hold for almost every $t \geqslant 0$:

$$A(t) = A, \quad a(t) = a, \tag{2.7}$$

$$[2A+B^2(t)]K(0)+[B(t)m(0)+b(t)]^2 = 0. \tag{2.8}$$

Moreover, if conditions (2.6)-(2.8) *are satisfied, then the covariance function of $(X_t;\ t \geqslant 0)$ is given by*

$$K(t, s) = e^{A|t-s|}K(0), \quad t \geqslant 0,\ s \geqslant 0. \tag{2.9}$$

Proof. If $K(0) = 0$ (i.e. $X(0) = m(0)$ a.s.), then $(X_t;\ t \geqslant 0)$ is second-order stationary if and only if $m_t = m(0)$ and $K_t = 0$, $t \geqslant 0$. In view of (2.1) and (2.3) this is true if and only if

$$\int_0^t [A(s)m(0)+a(s)]ds = \int_0^t \Psi_s^{-2}[B(s)m(0)+b(s)]^2 ds = 0, \quad t \geqslant 0,$$

or, equivalently,

$$A(t)m(0)+a(t) = B(t)m(0)+b(t) = 0 \text{ a.e.}$$

Let us now consider the case where $K(0) > 0$. It is clear from Lemma 2.1 that conditions (2.6)-(2.8) ensure second-order stationarity and that (2.9) holds. Conversely, if $(X_t;\ t \geqslant 0)$ is second-order stationary, then $K_t = K(0)$ and $K(t, s) = R(t-s)$, $t \geqslant 0$, $s \geqslant 0$. Then (2.2) implies that for $t \geqslant s$

$$\int_s^t A(u)du = \operatorname{Log} \frac{R(t-s)}{K(0)}$$

from which it follows that there exists a constant A such that $A(t) = A$ a.e. Consequently, $\Psi_t = e^{At}$, $t \geqslant 0$, and since $m_t = m(0)$, $t \geqslant 0$, formula (2.1) leads to $a(t) = -Am(0)$ a.e. Moreover, since $K_t = K(0)$, $t \geqslant 0$, formula (2.3) implies that (2.8) holds a.e.

Examples 2.1. (a) If $A(t) \equiv A < 0$, $a(t) \equiv a$, $B(t) \equiv (-2A)^{1/2}$, $b(t)$

$\equiv A^{-1}Ba$ $t \geqslant 0$, then for $m(0) = -a/A$ and every $K(0)$ the process is second-order stationary.

(b) If $A(t) \equiv A$, $a(t) \equiv a$, $B(t) \equiv B$, $b(t) \equiv b$, $t \geqslant 0$, with $2A+B^2 < 0$, then for

$$m(0) = -\frac{a}{A} \quad \text{and} \quad K(0) = \frac{(Ab-Ba)^2}{A^2|2A+B^2|}$$

the process is second-order stationary.

(c) If $A(t) \equiv A < 0$, $a(t) \equiv 0$, $B(t) \equiv (-2A)^{1/2} \sin t$, $b(t) = (-2A)^{1/2} \cos t$, $t \geqslant 0$, then for $m(0) = 0$ and $K(0) = 1$ the process is second-order stationary.

Remark 2.1. If the state process generated by (1.1) is second-order stationary with $K(0) > 0$, then from Lemmas 2.2 and 2.3 we infer that there exist a wide-sense Wiener process W^* and a constant $A \leqslant 0$ such that

$$dX_t = A[X_t - m(0)]dt + \left(-2AK(0)\right)^{1/2} dW_t^*, \quad t \geqslant 0, \ W_0 = X(0).$$

Now we consider the autonomous version of model (1.1):

$$dX_t = (AX_t + a)dt + (BX_t + b)dW_t, \quad t \geqslant 0, \ X_0 = X(0). \tag{2.10}$$

We study the problem of existence of a strict-sense stationary solution with first two moments. First, let us eliminate the deterministic case. In view of Lemma 2.3 we know that the solution process of (2.10) is strictly stationary with $K(t) \equiv 0$ if and only if $Am(0)+a = Bm(0)+b = 0$ and $X(0) = m(0)$ a.s. Now let us work with $K(0) > 0$ and, taking into account Lemma 2.3, assume that $2A+B^2 \leqslant 0$. In fact, when $B = 0$, in order to avoid the trivial case

$$A = a = B = b = 0, \quad \text{i.e.} \quad X_t \equiv X(0), \quad t \geqslant 0, \tag{C0}$$

we shall assume

$$B = 0, \quad A < 0, \quad b \neq 0. \tag{C1}$$

Moreover, when $B \neq 0$, since the deterministic case is eliminated, we have

$$B \neq 0, \quad 2A+B^2 < 0, \quad Ab \neq aB. \tag{C2}$$

We are able to prove the following

LEMMA 2.4. *In case* (C1) *or* (C2) *there exists a unique invariant probability distribution* $\mu_{a,A}^{b,B}$ *for the Markov process generated by* (2.10). *In case* (C1) $\mu_{a,A}^{b,B}$ *is the gaussian distribution* $N(-a/A, b^2/2|A|)$, *and in case* (C2) *it is the distribution of a random variable of the form*

$$\{\text{sign}\, U^{-1}(b/B - a/A) - b/B\},$$

where U *has the gamma-distribution*

$$\Gamma\left(1 - \frac{2A}{B^2}, \frac{B^2}{2|A|\,|b/B - a/A|}\right).$$

Moreover, if $X(0)$ is distributed along $\mu_{a,A}^{b,B}$, then the state process $(X_t;\ t \geqslant 0)$ is a second-order (stationary) process with

$$(2.11)\qquad m_t \equiv -\frac{a}{A}, \quad K(t, s) = \frac{(Ab-aB)^2}{A^2|2A+B^2|} e^{A|t-s|}, \quad s \geqslant 0,\ t \geqslant 0.$$

Proof. In case (C1) we can write

$$d(X_t + a/A) = A(X_t + a/A)dt + bdW_t, \quad t \geqslant 0,\ X_0 = X(0).$$

Since, as is well known, an Ornstein-Uhlenbeck process admits a unique invariant probability measure which is also gaussian, the result holds.

In case (C2) we can write

$$X_t = Z_t - b/B, \quad t \geqslant 0,$$

where

$$dZ_t = (AZ_t + \tilde{a})dt + BZ_t dW_t, \quad t \geqslant 0,\ Z_0 = X(0) + b/B,$$

and

$$\tilde{a} = a - bA/B \neq 0.$$

Let us assume that $\tilde{a}$ is positive. From the formula

$$(2.12)\qquad Z_t = \Phi_t\{Z_0 + \tilde{a}\int_0^t \Phi_s^{-1} ds\}, \quad t \geqslant 0,$$

it is clear that if a stationary probability distribution for $(Z_t;\ t \geqslant 0)$ exists, then it is necessarily concentrated on $]0, \infty[$. It is easy to show (using, for instance, the results of [7], § 13) that $]0, \infty[$ is a non-singular interval of positive recurrent type for the Markov process generated by (2.12) and, therefore, that a unique stationary distribution confined on $]0, \infty[$ for this process exists. Moreover (see, e.g., [13], p. 274, or [16], Example F), this distribution admits a density f which satisfies the Pearson equation

$$\frac{\dot{f}(z)}{f(z)} = \frac{2}{B^2}\frac{(A-B^2)z+\tilde{a}}{z^2},$$

and then is of the form

$$Kz^{2(AB^{-2}-1)} \exp\left\{-\frac{2\tilde{a}}{B^2}\frac{1}{z}\right\}, \quad z > 0.$$

It follows that the stationary distribution for $(Z_t;\ t \geqslant 0)$ is that of a random variable U^{-1}, where U has the $\Gamma(1-2A/B^2, B^2/2\tilde{a})$-distribution. Moreover, simple computations show that, since $2A+B^2 < 0$,

$$\mathrm{E}U^{-1} = \frac{b}{B} - \frac{a}{A}, \quad \mathrm{Var}\ U^{-1} = \frac{(Ab-aB)^2}{A^2|2A+B^2|}.$$

Finally, coming back to $(X_t;\ t \geqslant 0)$, we obtain easily the announced result. Similar arguments lead to the result for $\tilde{a} < 0$.

3. Optimal linear filtering of the state process. In this section we consider the linear filtering problem (in the sense of [11], Ch. 15) for the state X_t of the process generated by (1.1) by the use of observations on $[0, t]$ of the process $(Y_t;\ t \geqslant 0)$ which is given by

$$dY_t = [A_1(t)X_t + a_1(t)]dt + B_1(t)d\tilde{W}_t, \qquad t \geqslant 0,\ Y_0 = Y(0).$$

Here $\tilde{W} = (\tilde{W}_t;\ t \geqslant 0)$ is a standard brownian motion in R independent of W, the functions A_1, a_1, B_1 satisfy analogous conditions as A, a, B, respectively, and the initial state $Y(0)$ is a square-integrable random variable such that $(X(0), Y(0))$ is independent of $(W, \tilde{W})$. We assume also that $B_1(t) \neq 0$ for all $t > 0$.

A. *Equations for the optimal linear filter.* We have the following

THEOREM 3.1. *The optimal linear filter $\hat{X}_t$ of the state X_t from the observation $(Y_s;\ 0 \leqslant s \leqslant t)$ is given by*

$$d\hat{X}_t = [A(t)\hat{X}_t + a(t)]dt + \gamma_t \frac{A_1(t)}{B_1^2(t)} \{dY_t - [A_1(t)\hat{X}_t + a_1(t)]dt\}, \qquad t > 0,$$

$$\dot{\gamma}_t = 2A(t)\gamma_t + \Sigma_t^2 - \left\{\frac{A_1(t)}{B_1(t)}\right\}^2 \gamma_t^2, \qquad t > 0, \tag{3.1}$$

with

$$\hat{X}_0 = m(0) + \mathrm{Cov}\,(X(0), Y(0))\,\mathrm{Cov}^+(Y(0), Y(0))[Y(0) - \mathrm{E}\,Y(0)],$$

$$\gamma_0 = K(0) - \mathrm{Cov}^2(X(0), Y(0))\,\mathrm{Cov}^+(Y(0), Y(0)),$$

$$\Sigma_t^2 = B^2(t)K_t + [B(t)m_t + b(t)]^2, \qquad t \geqslant 0,$$

while

$$\gamma_t = \mathrm{E}(X_t - \hat{X}_t)^2, \qquad t \geqslant 0.$$

Proof. Taking into account Lemma 2.2, we obtain the result directly by the use of Theorem 15.3 in [11].

B. *Behaviour of the filter error.* Let $(\gamma_t^x;\ t \geqslant 0)$ denote the solution of (3.1) when $\gamma_0 = x$, i.e. $\gamma_t^x = \mathrm{E}(X_t - \hat{X}_t)^2$ when

$$K(0) - \mathrm{Cov}^2(X(0), Y(0))\,\mathrm{Cov}^+(Y(0), Y(0)) = x \geqslant 0.$$

Then the following holds:

THEOREM 3.2. *Assume that the state process $(X_t;\ t \geqslant 0)$ is non-deterministic second-order stationary and that the equation for the observation is autonomous with $A_1 \neq 0$. Then*

(i) *under* (C0), *for* $t \geqslant 0$,

$$\gamma_t^x = \begin{cases} [1/x+(A_1/B_1)^2 t]^{-1} & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases}$$

and, for every $x \geqslant 0$,

$$\lim_{t\uparrow\infty} \downarrow \gamma_t^x = 0.$$

(ii) *under* (C1) *or* (C2)

$$\gamma_t^x = \frac{\gamma^*(x-\gamma_*)-\gamma_*(x-\gamma^*)\exp\{-2\lambda_* t\}}{(x-\gamma_*)-(x-\gamma^*)\exp\{-2\lambda_* t\}}, \quad t \geqslant 0,\ x \geqslant 0,$$

where

$$\gamma^* = \frac{B_1^2}{A_1^2}(A+\lambda_*), \quad \gamma_* = \frac{B_1^2}{A_1^2}(A-\lambda_*), \quad \lambda_* = \left[A^2+2|A|K(0)\frac{A_1^2}{B_1^2}\right]^{1/2}.$$

Moreover, for every $x \leqslant \gamma^*$,

$$\lim_{t\uparrow\infty} \uparrow \gamma_t^x = \gamma^*$$

and, for every $x \geqslant \gamma^*$,

$$\lim_{t\uparrow\infty} \downarrow \gamma_t^x = \gamma^*.$$

Proof. Under (C0) the results are obvious. Under (C1) or (C2), since

$$\Sigma_t^2 \equiv -2AK(0) \quad \text{and} \quad 2A\gamma^*+2|A|K(0)-\frac{A_1^2}{B_1^2}(\gamma^*)^2 = 0,$$

either $x = \gamma^*$ and then $\gamma_t^x = \gamma^*$, $t \geqslant 0$, or $x \neq \gamma^*$ and then the function $(\gamma_t^x;\ t \geqslant 0)$ satisfies

$$\frac{\dot{\gamma}_t^x}{-A_1^2 B_1^{-2}(\gamma_t^x)^2+2A\gamma_t^x-2AK(0)} = 1, \quad t \geqslant 0,\ \gamma_0^x = x.$$

In that case, if $F(y)$ is an integral of the expression

$$\{-A_1^2 B_1^{-2} y^2+2Ay-2AK(0)\}^{-1},$$

then we have $F(\gamma_t) = t+\text{const}$. We can easily compute $F(y)$ and then invert the last equation in order to obtain $(\gamma_t^x;\ t \geqslant 0)$ in the form stated in the theorem. The last assertion is a simple consequence of that form.

Remarks 3.1. (a) Theorem 3.2 precises Theorem 16.2 from [11] in the one-dimensional case.

(b) The stability property $\lim_{t\uparrow\infty} \gamma_t^x = \gamma^*$ still holds (cf. [9], Theorem 3.7) if model (1.1) is autonomous and such that a second-order stationary solution exists even if the state process $(X_t;\ t \geqslant 0)$ itself is not assumed to be second-order stationary.

4. Parametric estimation in the autonomous model. In this part of the paper we want to demonstrate how the two-step procedure (cf. Section 1) works in model (2.10) and also to compare the method of moments with that of maximum likelihood for estimating the drift coefficient parameters. We assume that the state process of (2.10) is observed in continuous time and we are interested in estimation of all the parameters a, A, b, and B or some functions of these parameters.

Let $C(\boldsymbol{R}_+, \boldsymbol{R})$ be the canonical space endowed with the σ-algebra $\mathscr{B} = \sigma\{\pi_s;\ s \geqslant 0\}$ generated by the coordinate functions π_s, $s \geqslant 0$. Let $P_{\nu,a,A}^{b,B}$ be the distribution of the state process of (2.10) when the distribution of $X(0)$ is ν and let ${}_TP_{\nu,a,A}^{b,B}$ be the restriction of $P_{\nu,a,A}^{b,B}$ to the σ-algebra $\mathscr{B}_T = \sigma\{\pi_s;\ 0 \leqslant s \leqslant T\}$. Then the statistical space under study when we observe the process from 0 to T is

$$\big(C(\boldsymbol{R}_+, \boldsymbol{R}),\ \mathscr{B}_T,\ \{{}_TP_{\nu,a,A}^{b,B};\ (\nu, a, A, b, B) \in \Lambda\}\big),$$

where Λ is some subset of the set $\mathscr{P} \times \boldsymbol{R}^4$ while $\mathscr{P}$ is the set of probability measures on $\boldsymbol{R}$.

A. *First step: parameter estimation in the diffusion coefficient*. By the use of the quadratic variation of the observed process we can identify the diffusion coefficient of the model before we estimate the parameters in the drift coefficient. We know that if $(\langle\pi\rangle(t):\ t \geqslant 0)$ is the quadratic variation of the canonical process, then the following equalities hold for every $t \in]0, T]$, $(\nu, a, A, b, B) \in \Lambda$:

$$\langle\pi\rangle(t) = \lim_{N\to+\infty} \sum_{i=1}^{2^N} (\pi_{it/2^N} - \pi_{(i-1)t/2^N})^2 = \int_0^t (b + B\pi_s)^2\, ds \qquad {}_TP_{\mu,a,A}^{b,B}\text{-a.s.}$$

So we are able to compute, by the use of the observation of the state process on a finite time, the functions of parameters b and B on which the distribution of the process is dependent. First, since $B = 0$ if and only if $(t^{-1}\langle\pi\rangle(t);\ t > 0)$ is a constant function, we can identify whether $B = 0$ or not. Now, when $B = 0$, we can compute b^2 by the formula $b^2 = t^{-1}\langle\pi\rangle(t)$. Similarly, when $B \neq 0$, since the formula

$$\int_0^t (b + B\pi_s)^2\, ds = \langle\pi\rangle(t), \qquad t \in]0, T],$$

holds ${}_TP^{b,B}_{\nu,a,A}$-a.s. for every ν, a, and A, we can, for instance, first compute b/B by solving the equation

$$[t_2\langle\pi\rangle(t_1)-t_1\langle\pi\rangle(t_2)]\frac{b^2}{B^2}+2\Big[\langle\pi\rangle(t_1)\int_0^{t_2}\pi_s ds-\langle\pi\rangle(t_2)\int_0^{t_1}\pi_s ds\Big]\frac{b}{B}+$$
$$+\Big[\langle\pi\rangle(t_1)\int_0^{t_2}\pi_s^2 ds-\langle\pi\rangle(t_2)\int_0^{t_1}\pi_s^2 ds\Big]=0$$

and then compute B^2 by the formula

$$B^2=\frac{\langle\pi\rangle(t_1)}{b^2B^{-2}t_1+\int_0^{t_1}\pi_s^2 ds+2bB^{-1}\int_0^{t_1}\pi_s ds}.$$

Consequently, we are allowed to assume that the diffusion coefficient is known because it would eventually have been computed previously (with probability one) on some finite time interval.

B. *Second step: parameter estimation in the drift coefficient.* In the following we are concerned with the reduced statistical model

$$(C(\boldsymbol{R}_+,\boldsymbol{R}),\ \mathscr{B}_T,\ \{{}_TP_{\nu,a,A};\ (\nu,a,A)\in\Theta\}),$$

where Θ is some subset of $\mathscr{P}\times\boldsymbol{R}^2$, and b and B have been cancelled in ${}_TP^{b,B}_{\nu,a,A}$ (and will be in $P^{b,B}_{\nu,a,A}$) since these parameters are now fixed known. Since we deal with the case where the model admits a stationary second-order distribution, we consider a parametrization of (2.10) given in terms of the natural parameters of the corresponding process. Taking into account the discussion in Section 2B, when $2A+B^2<0$, it is better to set

$$\beta=-A,\qquad m=-\frac{a}{A},\qquad \sigma^2=\frac{(b+Bm)^2}{2\beta-B^2},$$

and then to write model (2.10) in the form

$$dX_t=\beta(m-X_t)dt+(b+BX_t)dW_t,\qquad t\geqslant 0,\ X_0=X(0), \tag{4.1}$$

with $\beta>B^2/2$. A second-order stationary solution of (4.1) has mean and covariance functions given by

$$m_t\equiv m,\qquad K(t,s)=\sigma^2e^{-\beta|t-s|},\qquad t\geqslant 0,\ s\geqslant 0.$$

A strict-sense stationary solution of (4.1) corresponds to $\nu=\nu_{\beta,m}$, where $\nu_{\beta,m}=\mu^{b,B}_{\beta m,-\beta}$ is given in Lemma 2.4. When we consider the case $B\neq 0$, we shall assume that $m+b/B>0$ (since the case $m+b/B<0$ is quite similar) and, moreover, that $\mathscr{N}$ is a subset of the set $\mathscr{P}(b/B)$ of probability measures on $\boldsymbol{R}$ which are concentrated on $]-b/B,\infty[$. Finally, rather to look for estimates of a and A we shall try to estimate β, m, and σ^2. Then we are concerned with the statistical model

$$(C(\mathbb{R}_+, \mathbb{R}), \mathscr{B}_T, \{{}_TP_{\nu,\beta,m};\ (\nu, \beta, m) \in \mathscr{C}\}), \tag{4.2}$$

where $\mathscr{C}$ is some subset of $\mathscr{P} \times \mathbb{R}_+^* \times \mathbb{R}$ (in fact, of $\mathscr{P}(b/B) \times]B^2/2, \infty[\times]-b/B, \infty[$ when $B \neq 0$), ${}_TP_{\nu,\beta,m}$ (resp. $P_{\nu,\beta,m}$) standing for ${}_TP_{\nu,\beta m,-\beta}$ (resp. $P_{\nu,\beta m,-\beta}$).

An immediate consequence of Girsanov's Theorem (see, e.g., [11]) is that if ν is absolutely continuous with respect to some fixed probability measure ν_0, then the statistical space (4.2) is dominated by ${}_TP_{\nu_0,0,0}$ with the likelihood function

$$\frac{d\,{}_TP_{\nu,\beta,m}}{d\,{}_TP_{\nu_0,0,0}} = H_{\nu_0,\nu}(\pi_0) \exp\{L_T(\beta, m)\},$$

where

$$H_{\nu_0,\nu}(x) = \frac{d\nu}{d\nu_0}(x), \qquad x \in \mathbb{R},$$

and

$$L_T(\beta, m) = \int_0^T \beta(m-\pi_t)(b+B\pi_t)^{-2}\,d\pi_t - \tfrac{1}{2}\int_0^T \beta^2(m-\pi_t)^2(b+B\pi_t)^{-2}\,dt,$$

the stochastic integral being defined with respect to ${}_TP_{\nu_0,0,0}$.

We can write

$$L_T(\beta, m) = \beta m S_0(T) - \beta S_1(T) - \frac{\beta^2}{2} m^2 I_0(T) + \beta^2 m I_1(T) - \frac{\beta^2}{2} I_2(T),$$

where

$$I_j(T) = \int_0^T \pi_t^j (b+B\pi_t)^{-2}\,dt, \qquad j = 0, 1, 2,$$

and the stochastic integrals

$$S_j(T) = \int_0^T \pi_t^j (b+B\pi_t)^{-2}\,d\pi_t, \qquad j = 0, 1,$$

are explicitly given by

$$S_0(T) = \begin{cases} b^{-2}[\pi_T - \pi_0] & \text{if } B = 0, \\ B^{-1}[(b+B\pi_0)^{-1} - (b+B\pi_T)^{-1}] + B[bI_0(T) + BI_1(T)] & \text{if } B \neq 0, \end{cases}$$

and

$$S_1(T) = \begin{cases} 2^{-1} b^{-2}[\pi_T^2 - \pi_0^2 - b^2 T] & \text{if } B = 0, \\ T/2 + B^{-2}[\operatorname{Log}(b/B + \pi_T) - \operatorname{Log}(b/B + \pi_0)] - & \\ \quad - B^{-2} b[(b+B\pi_0)^{-1} - (b+B\pi_T)^{-1}] - & \\ \quad - b[bI_0(T) + BI_1(T)] & \text{if } B \neq 0. \end{cases}$$

Then, if v is known, it is easy to see that the maximum likelihood estimate of (β, m) is given by

$$\hat{m}_T = \frac{S_0(T)I_2(T)-S_1(T)I_1(T)}{S_0(T)I_1(T)-S_1(T)I_2(T)},$$

(4.3)

$$\hat{\beta}_T = \frac{S_0(T)I_1(T)-S_1(T)I_0(T)}{I_0(T)I_2(T)-I_1^2(T)}.$$

The asymptotic properties of these estimates are described in the following

THEOREM 4.1. *The estimate $(\hat{\beta}_T, \hat{m}_T)$ given by (4.3) is strongly consistent and is asymptotically normal, i.e. with respect to $P_{v,\beta,m}$ we have*

$$\lim_{T\to+\infty} (\hat{\beta}_T, \hat{m}_T) \overset{\text{a.s.}}{=} (\beta, m),$$

$$\lim_{T\to+\infty} T^{1/2}\left[\begin{pmatrix}\hat{\beta}_T\\ \hat{m}_T\end{pmatrix}-\begin{pmatrix}\beta\\ m\end{pmatrix}\right] \overset{\mathscr{D}}{=} N\left(0, \begin{bmatrix}-1 & 0\\ m & \beta\end{bmatrix}^{-1} \Lambda_{\beta,m}^{-1}\begin{bmatrix}-1 & m\\ 0 & \beta\end{bmatrix}^{-1}\right),$$

where

$$\Lambda_{\beta,m} = \int_{\mathbb{R}} \begin{bmatrix}x(b+Bx)^{-1}\\ (b+Bx)^{-1}\end{bmatrix}\begin{bmatrix}x(b+Bx)^{-1}\\ (b+Bx)^{-1}\end{bmatrix}' dv_{\beta,m}(x).$$

Proof. If $L_T^{(1)}(\beta, m)$ stands for the gradient of $L_T(\beta, m)$, we can write

$$L_T^{(1)}(\beta, m) = \begin{bmatrix}-1 & m\\ 0 & \beta\end{bmatrix}\left\{\begin{bmatrix}S_1(T)\\ S_0(T)\end{bmatrix}-\beta m\begin{bmatrix}I_1(T)\\ I_0(T)\end{bmatrix}+\beta\begin{bmatrix}I_2(T)\\ I_1(T)\end{bmatrix}\right\}.$$

Similarly, if $L_T^{(2)}(\beta, m)$ denotes the matrix of second-order partial derivatives of $L_T(\beta, m)$, we have

$$L_T^{(2)}(\beta, m) = -\begin{bmatrix}-1 & m\\ 0 & \beta\end{bmatrix}\begin{bmatrix}I_2(T) & I_1(T)\\ I_1(T) & I_0(T)\end{bmatrix}\begin{bmatrix}-1 & 0\\ m & \beta\end{bmatrix}+$$
$$+\begin{bmatrix}0 & S_0(T)-\beta m I_0(T)+\beta I_1(T)\\ S_0(T)-\beta m I_0(T)+\beta I_1(T) & 0\end{bmatrix}.$$

Usual arguments in the maximum likelihood method will provide the announced results if we prove that, with respect to $P_{v,\beta,m}$,

(4.4) $$\lim_{T\to+\infty} T^{-1} L_T^{(1)}(\beta, m) \overset{\text{a.s.}}{=} 0,$$

(4.5) $$\lim_{T\to+\infty} T^{-1} L_T^{(2)}(\beta, m) \overset{\text{a.s.}}{=} -\begin{bmatrix}-1 & m\\ 0 & \beta\end{bmatrix}\Lambda_{\beta,m}\begin{bmatrix}-1 & 0\\ m & \beta\end{bmatrix} = -\Gamma_{\beta,m},$$

(4.6) $$\lim_{T\to+\infty} T^{-1/2} L_T^{(1)}(\beta, m) \overset{\mathscr{D}}{=} N(0, \Gamma_{\beta,m}).$$

Let us first note that

(4.7) $$\lim_{T\to+\infty} T^{-1} I_j(T) = \int x^j (b+Bx)^{-2} d\nu_{\beta,m}(x) < +\infty, \quad j = 0, 1, 2,$$

because of Lemma 2.4 and of the ergodic properties of the process under study that it ensures (cf. Theorem 4.1 in [12]). Now let us notice that we can write

(4.8) $$d\pi_t = \beta(m-\pi_t)\,dt + (b+B\pi_t)\,d\tilde{W}_t, \quad t \geqslant 0,$$

where $(\tilde{W}_t;\ t \geqslant 0)$ is some brownian motion with respect to $P_{\nu,\beta,m}$. It follows that

(4.9) $$S_j(T) = \beta m I_j(T) - \beta I_{j+1}(T) + \int_0^T \pi_t^j (b+B\pi_t)^{-1} d\tilde{W}_t, \quad j = 0, 1.$$

Then, using Lemma 17.4 of [11] and taking into account (4.7) for $j = 0, 2$, we obtain

$$\lim_{T\to+\infty} T^{-1}[S_j(T) - \beta m I_j(T) - \beta I_{j+1}(T)] \overset{\text{a.s.}}{=} 0, \quad j = 0, 1,$$

which implies (4.4) and (4.5).

In order to prove (4.6), because of (4.9) we have only to show that

$$\lim_{T\to+\infty} T^{-1/2} \int_0^T \begin{bmatrix} \pi_t \\ 1 \end{bmatrix} (b+B\pi_t)^{-1}\, dW_t \overset{\mathcal{D}}{=} N(0, \Lambda_{\beta,m}).$$

This fact follows from (4.7) and the results of [14] or [15].

Remarks 4.1. (a) These results can be applied, for instance, in the case where $\nu = \delta_{x(0)}$ for given $x(0)$ (with $x(0) > -b/B$ in the case $B \neq 0$). When ν is not known, it is clear that we can still use the estimate $(\hat{\beta}_T, \hat{m}_T)$ defined by (4.3); in that case it can be considered as an approximate maximum likelihood estimate (see [1] and [9] for the classical stationary Gauss-Markov case).

(b) We can use the consistent estimate

$$\hat{\sigma}_T^2 = \frac{(b+\hat{m}_T B)^2}{2\hat{\beta}_T - B^2}$$

for the parameter σ^2. It is possible to obtain the asymptotic distribution for $(\hat{\beta}_T, \hat{m}_T, \hat{\sigma}_T^2)$ (see [9] for explicit formulas of the asymptotic covariance in cases $B = 0$ and $B \neq 0$).

Now we use the method of moments for parameter estimation. Let $\tilde{m}_T$

and $\tilde{\sigma}_T^2$ be defined on (4.2) by

$$
\tilde{m}_T = T^{-1} J_1(T) = T^{-1} \int_0^T \pi_t \, dt,
$$

(4.10)

$$
\tilde{\sigma}_T^2 = T^{-1} J_2(T) - \tilde{m}_T^2 = T^{-1} \int_0^T (\pi_t - \tilde{m}_T)^2 \, dt.
$$

We have the following

THEOREM 4.2. *The estimate $(\tilde{m}_T, \tilde{\sigma}_T^2)$ given by (4.10) is strongly consistent and the estimate $\tilde{m}_T$ is asymptotically normal with*

$$
\lim_{T\to+\infty} T^{1/2}(\tilde{m}_T - m) \overset{\mathscr{D}}{=} N(0, 2\sigma^2/\beta).
$$

Moreover, if $B = 0$ or if $B \neq 0$ and $\beta > 3B^2/2$, we also have

$$
\lim_{T\to+\infty} T^{1/2}\left[\begin{pmatrix} \tilde{m}_T \\ \tilde{\sigma}_T^2 \end{pmatrix} - \begin{pmatrix} m \\ \sigma^2 \end{pmatrix}\right] \overset{\mathscr{D}}{=} N(0, \Psi_{\beta,m} \Delta_{\beta,m} \Psi'_{\beta,m})
$$

where

$$
\Psi_{\beta,m} = \begin{bmatrix} 1/\beta & 0 \\ \dfrac{2[B(mB+b) - \beta m]}{\beta(2\beta - B^2)} & \dfrac{2}{2\beta - B^2} \end{bmatrix}
$$

and

$$
\Delta_{\beta,m} = \int_{\mathbb{R}} \begin{bmatrix} b+Bx \\ x(b+Bx) \end{bmatrix} \begin{bmatrix} b+Bx \\ x(b+Bx) \end{bmatrix}' d\nu_{\beta,m}(x).
$$

Proof. The ergodic properties of the process under study provide the consistency of the estimate since

(4.11)
$$
\lim_{T\to+\infty} T^{-1} J_i(T) \overset{\text{a.s.}}{=} \int x^i \, d\nu_{\beta,m}(x) < +\infty, \quad i = 1, 2.
$$

Now, let us look at $\tilde{m}_T - m$ and $\tilde{\sigma}_T^2 - \sigma^2$: by the use of representation (4.8) and the fact that

$$
\sigma^2 + m^2 = \frac{b^2 + 2m(\beta m + bB)}{2\beta - B^2}
$$

it is easy to prove that

(4.12)
$$
T^{1/2}\begin{pmatrix} \tilde{m}_T - m \\ \tilde{\sigma}_T^2 - \sigma^2 \end{pmatrix}
= \begin{bmatrix} \beta^{-1} T^{-1/2}(\pi_0 - \pi_T) \\ \dfrac{1}{2\beta - B^2} \dfrac{1}{T^{1/2}} \left[\pi_0^2 - \pi_T^2 + \dfrac{2(\beta m + bB)}{\beta}(\pi_0 - \pi_T)\right] + \dfrac{\tilde{m}_T + m}{\beta} \dfrac{1}{T^{1/2}}(\pi_T - \pi_0) \end{bmatrix} +
$$

$$+\begin{bmatrix} 1/\beta & 0 \\ \dfrac{2(\beta m+bB)}{\beta(2\beta-B^2)}-\dfrac{\tilde{m}_T+m}{\beta} & \dfrac{2}{2\beta-B^2} \end{bmatrix} \frac{1}{T^{1/2}} \int_0^T \begin{bmatrix} b+B\pi_t \\ \pi_t(b+B\pi_t) \end{bmatrix} d\tilde{W}_t.$$

Moreover, $\mathrm{E}_{\nu,\beta,m}$ standing for the expectation with respect to $P_{\nu,\beta,m}$, setting $m_{i,t} = \mathrm{E}_{\nu,\beta,m}(\pi_t^i)$, $i = 1, 2$, we have

$$m_{1,t}-m = e^{-\beta t}(m_{1,0}-m)$$

and

$$m_{2,t}-(m^2+\sigma^2) = \exp\{(-2\beta+B^2)t\}\,[K(0)-(m^2+\sigma^2)+ \\ +2(\beta m+bB)(m_{1,0}-m)\int_0^t \exp\{(\beta-B^2)s\}\,ds].$$

Then

$$\lim_{T\to+\infty} m_{2,T} = m^2+\sigma^2,$$

which implies

$$\lim_{T\to+\infty} \mathrm{E}_{\nu,\beta,m}|T^{-1/2}(\pi_0^i-\pi_T^i)| = 0, \quad i = 1, 2. \tag{4.13}$$

Formula (4.12), together with (4.11), (4.13), and the results of [14], ensures that

$$\lim_{T\to+\infty} T^{1/2}(\tilde{m}_T-m) \overset{\mathscr{D}}{=} N\big(0, \beta^{-2}\int(b+Bx)^2\,d\nu_{\beta,m}(x)\big),$$

where

$$\int_{\mathbb{R}} (b+Bx)^2\,d\nu_{\beta,m}(x) = b^2+B^2(m^2+\sigma^2)+2bBm = 2\sigma^2\beta.$$

Moreover, if $B = 0$ or $B \neq 0$ and $\beta > 3B^2/2$, then by the inequality

$$\int x^4\,d\nu_{\beta,m}(x) < +\infty$$

the ergodic property

$$\lim_{T\to+\infty} T^{-1}\int_0^T \pi_t^j\,dt \overset{\text{a.s.}}{=} \int x^j\,d\nu_{\beta,m}(x) < +\infty, \quad j = 1, 2, 3, 4,$$

holds. Then, by the same arguments, the last assertion in the theorem holds because

$$\lim_{T\to+\infty}\left\{\frac{2(\beta m+bB)}{\beta(2\beta-B^2)}-\frac{\tilde{m}_T+m}{\beta}\right\} \overset{\text{a.s.}}{=} \frac{2[B(mB+b)-\beta m]}{\beta(2\beta-B^2)}.$$

Remarks 4.2. (a) It is easy to see that in the case $B = 0$ the asymptotic variance of $\tilde{m}_T$ is the same as that of $\hat{m}_T$ and in the case $B \neq 0$ the first one is greater than the second one.

(b) We can use the consistent estimate

$$\tilde{\beta}_T = \frac{(b+B\tilde{m}_T)^2}{2\tilde{\sigma}^2} + \frac{B^2}{2}$$

for the parameter β. It is still possible to obtain the asymptotic distribution for $(\tilde{\beta}_T, \tilde{m}_T, \tilde{\sigma}_T^2)$ (see [9] for explicit formulas of the asymptotic covariances).

REFERENCES

[1] M. Arato, *On the statistical examination of continuous state Markov processes. I*, Selected Transl. in Math. Statist. and Probability 14 (1978), p. 203-225.

[2] L. Arnold, *Stochastic differential equations (Theory and applications)*, J. Wiley, 1974.

[3] B. M. Brown and J. I. Hewitt, *Asymptotic likelihood theory for diffusion processes*, J. Appl. Probability 12 (1975), p. 228-238.

[4] C. Bruni, G. Di Pillo and G. Koch, *Bilinear systems: an appealing class of "nearly linear" systems in theory and applications*, IEEE Trans. Automatic Control Vol. AC-19, 4 (1974), p. 334-348.

[5] C. W. Granger and A. Andersen, *Non-linear time series modelling*, in: *Applied time series analysis* (D. F. Findley ed.), Academic Press, New York 1978.

[6] A. Isidori, *Direct construction of minimal bilinear realizations from non-linear input output maps*, IEEE Trans. Automatic Control Vol. AC-18, 6 (1973), p. 626-631.

[7] K. Ito and M. Nisio, *On stationary solutions of a stochastic differential equation*, J. Math. Kyoto Univ. 4 (1) (1964), p. 1-75.

[8] A. Le Breton, *Parameter estimation in a linear stochastic differential equation*, p. 353-366 in: Trans. 7th Prague Conf. and 1974 E.M.S., Acad. Publ. House of the Czechoslovak Acad. of Sci., Vol. A (1977).

[9] — and Musiela, *A study of a one-dimensional bilinear differential model for stochastic processes*, R.R. no. 221 Lab. I.M.A.G., Univ. Grenoble 1, France, 1980.

[10] T. S. Lee and F. Kozin, *Almost sure asymptotic likelihood theory for diffusion processes*, J. Appl. Probability 14 (1977), p. 527-537.

[11] R. Š. Lipcer and A. N. Širjajev, *Statistics of random processes. I, II*, Springer-Verlag, New York — Heidelberg 1977, 1978.

[12] G. Maruyama and H. Tanaka, *Some properties of one-dimensional diffusion processes*, Mem. Fac. Sci. Kyushu Univ., Ser. A., 11 (2) (1957), p. 117-141.

[13] Yu. V. Prohorov and Yu. A. Rozanov, *Probability theory*, Springer-Verlag, Berlin 1969.

[14] A. F. Taraskin, *On the asymptotic normality of vector stochastic integrals and estimates of the drift parameters of a many-dimensional diffusion process*, Teor. Verojatnost. i Mat. Statist. 2 (1970), p. 205-220.

[15] — *Some limit theorems for stochastic integrals*, in: *Theory of stochastic processes*, Vol. 1, J. Wiley, 1974.

[16] E. Wong, *The construction of a class of stationary Markoff processes*, p. 264 - 276 in: Proc. Symp. in Appl. Math., 16, Amer. Math. Soc., 1964.

A. Le Breton
Laboratorie I.M.A.G.
B.P. 53 X
38041 Grenoble - Cédex, France

M. Musiela
Institute of Mathematics
Polish Academy of Sciences
ul. Kopernika 18
51 - 617 Wrocław, Poland

Received on 1. 8. 1981