NON-PARAMETRIC ESTIMATION
FOR DIFFUSION PROCESSES

by

G. Banon

Memorandum No. ERL-M599

25 October 1976

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ABSTRACT

It is proved that under specific condition (so called condition G_2) on the transition probability operator of a stationary Markov process, a recursive Kernel estimate of the initial density is convergent in quadratic mean.

Assumptions on the differential stochastic equations driven by Brownian motion are derived under which the stationary solution satisfy condition G_2.

The above results are applied to solve a class of nonlinear identification problems.

Research sponsored by the National Science Foundation Grant OIP75-04371.

TABLE OF CONTENTS

† q.m. for quadratic mean.

‡ p. for probability.

0) Introduction

An important problem in control engineering is the identification of dynamical systems. In this paper we focus our interest on systems represented by stochastic differential equations. Thus far, primary emphasis has been on solving the identification problem for linear systems, and many techniques have been proposed (see e.g., [Kalman and Bucy 1961], [Banon 1971], [Aguilar-Martin 1974], [Alengrin 1974], [Salut 1976]). In contrast, our interest is to develop an approach for the identification of a class of non-linear systems. More precisely, let $\{X_t, t\in[0,\infty)\}$ be a stochastic process defined on a probability space $(\Omega, a, \mathcal{P})$ and satisfying the following stochastic differential equation:

$$dX_t = m(X_t)\, dt + \sigma(X_t) dW_t, \tag{0-1a}$$

with initial condition

$$X_0 = X \tag{0-1b}$$

In (0-1a) W_t represents a Brownian motion defined on the same probability space $(\Omega, a, \mathcal{P})$.

Hence forth we shall make the following assumptions relative to the problem (0-1):

A_1: The initial random variable X is defined on $(\Omega, a, \mathcal{P})$ and is of second order: $EX^2 < \infty$.

A_2: $m(\cdot)$ and $\sigma(\cdot)$ are Borel measurable function on R satisfying, for $x, y \in R$, the uniform Lipschitz condition:

$$|m(x)-m(y)| \leq K|x-y|$$
$$|\sigma(x)-\sigma(y)| \leq K|x-y| \qquad (0\text{-}2)$$

and the linear growth condition:

$$|m(x)| \leq K\sqrt{1+x^2}$$
$$|\sigma(x)| \leq K\sqrt{1+x^2} \qquad (0\text{-}3)$$

where K is a positive constant.

Under A_1 and A_2 we know (see [Wong 1971] p. 150, prop. 4.1, P_5 and P_6) that the X_t process, solution of problem (0-1) is unique with probability one and is a Markov process.

Under some additional conditions the unique solution of problem (0-1) must have a stationary transition density, say $p_{X_t|X_0=a}(\cdot)$ satisfying the forward equation of Kolmogorov and $p_{X_t|X_0=a}(\cdot)$ must tend to a limiting density, say $p(\cdot)$ as t goes to infinity.

A class of such X_t process for which $m(\cdot)$ and $\sigma^2(\cdot)$ are polynomials has been constructed and some specific processes of this class are given in [Wong 1964] (see examples B and E). In this paper, we propose a procedure to estimate point by point the function $m(\cdot)$ when the function $\sigma(\cdot)$ is known or when $\sigma(\cdot)$ is unknown but takes a constant value.

Considering the properties of the transition density of the X_t process, its limiting density $p(\cdot)$ can be explicitly related to the pair $(m(\cdot),\sigma(\cdot))$.

So the techniques of non-parametric estimation used here to estimate point by point the density function $p(\cdot)$ appears to be a powerful tool in solving our initial problem of non-linear identification.

For the case of a stationary Markov process having a initial density $p(\cdot)$, in section 1, we give a local (or point by point) recursive estimate of $p(\cdot)$ and its derivative $p'(\cdot)$, for which we show the quadratic mean convergence under a specific assumption on the process itself. For the case of stochastic processes defined by problem (0-1) in section 2, we derive sufficient conditions on the pair $(m(\cdot),\sigma(\cdot))$ which imply the assumptions made in the previous section.

Independently of the estimation of the density we give in section 3 a quadratic mean convergent recursive estimate of σ^2 (assuming that the function $\sigma(\cdot)$ is constant). This is done using the quadratic variation properties of the stochastic processes defined by (0-1). Finally, in section 4 applying the previous results, we give the solution to a class of non-linear identification problems by suggesting a local estimate of the function $m(\cdot)$.

1) Estimation of $p(x_0)$ and $p'(x_0)$ for Markov processes

In the case of a sequence $X_1, X_2,\ldots,X_n$ of independent and identically distributed random variables whose distribution is absolutely continuous, many non-parametric estimates of the density function have been proposed in the past. We may recall the Rosenblatt-Parzen estimate (see [Rosenblatt 1956] and [Parzen 1962]), the Yamato estimate (see [Yamato 1972]) and the Deheuvels estimate (see [Deheuvels 1973]). As far as we know, in the case of a dependent sequence, relatively few results have been obtained. Under specific conditions on the nature of the sequence it has been shown that the Rosenblatt-Parzen estimate is still convergent (see [Roussas 1969] and [Rosenblatt 1970]).

In this paper, instead of a sequence of random variables, we have to deal with a stochastic process $\{X_t, t\in[0,\infty)\}$. In this section we assume that the X_t process is a stationary Markov process which has a initial density $p_X(\cdot)$ or simply $p(\cdot)$, and we introduce an estimate of $p(x_0)$, where x_0 is any point of R, which has the same structure as the Deheuvels estimate. Such structure is "recursive" and seems to be well adapted to our problem of estimating the function $m(\cdot)$ (see section 4). In order to prove the convergence to zero of the variance of our estimate (see theorem 1-1) we have to impose a specific condition on the transition density of the stationary Markov process.

For each $t\in[0,\infty)$, we define the transition probability operator H_t of a stationary process X_t by:

$$(H_t f)(a) = E(f(X_t)|X_0=a) \qquad a\in R \quad ,$$

where $f(\cdot)$ is any Borel measurable bounded function on R.

The transition probability operator is said to satisfy the condition G_2(see [Rosenblatt 1970] p. 202 for the case of a sequence of random variables): if there is some $s>0$ such that

$$|H_s|_2 = \sup_{\{f:Ef(X)=0\}} \frac{E^{1/2}(H_s f)^2(X)}{E^{1/2}f^2(X)} \leq \alpha < 1 \; .$$

The H_t operator is in fact a contraction (for any $t\in[0,\infty)$: $|H_t|_2 \leq 1$).

For stationary Markov processes, the transition probability operator verifies the semigroup property, i.e. for $s,t\geq 0$: $H_{s+t}=H_sH_t=H_tH_s$ (see [Wong 1971] p. 183). As a consequence of the semigroup and contraction properties, the condition G_2 implies, for $t\in[0,\infty)$:

$$|H_t|_2 < \frac{\beta^t}{\alpha} \text{ with } \beta=\alpha^{1/2} < 1. \tag{1-1}$$

Theorem 1-1 (quadratic mean convergence of $p_t(x_0)$)

Let $\{X_t, t\in[0,\infty)\}$ be a stationary Markov process having a continuous and bounded initial density $p(\cdot)$ on R and satisfying condition G_2.

Let $K(\cdot)$ be a probability density function (i.e., non negative, Borel measurable function such that $\int_R K(y)dy=1$) and be bounded on R, and let h. be a strictly positive function on R^+ such that, $t\in[0,\infty)$:

$$h_t \underset{t\to\infty}{\downarrow} 0, \tag{1-2a}$$

$$b_t = \int_0^t h_s \, ds < \infty, \tag{1-2b}$$

$$\text{and } b_t \underset{t\to\infty}{\longrightarrow} \infty \tag{1-3}$$

$$\text{For } t>0, \text{ let } p_t(x_0) = 1/b_t \int_0^t K\left(\frac{x_0-X_s}{h_s}\right) ds, \tag{1-4}$$

then $p_t(x_0) \xrightarrow[t\to\infty]{q.m.} p(x_0)$. #

Proof

To prove the convergence in quadratic mean it is necessary and sufficient to show that:

$$E\, p_t(x_0) \xrightarrow[t\to\infty]{} p(x_0) \tag{1-5}$$

$$\text{and} \quad \text{Var } p_t(x_0) \xrightarrow[t\to\infty]{} 0. \tag{1-6}$$

Because we may write:

$$E\, p_t(x_0)-p(x_0)=\frac{1}{b_t}\int_0^t h_s\left(E\,\frac{1}{h_s}\,K\left(\frac{x_0-X_0}{h_s}\right)-p(x_0)\right)ds,$$

and because $b_t\to\infty$ as $t\to\infty$, a sufficient condition to show the asymptotic unbiased (1-5) is that $E\,\frac{1}{h_s}\,K\left(\frac{x_0-X_0}{h_s}\right)-p(x_0)$ converges to zero as s goes to infinity.

Using the same procedure as in [Bochner 1955] (see Theorem 1.1.1. p. 2) or [Rosenblatt 1971] (see p. 1816) we have:

$$\left|E\,\frac{1}{h_s}\,K\left(\frac{x_0-X_0}{h_s}\right)-p(x_0)\right|=\left|\int_R \frac{1}{h_s}\,K\left(\frac{x_0-x}{h_s}\right)p(x)dx-p(x_0)\right|$$

$$\leq \int_R K(y)\left|p(x_0-h_s y)-p(x_0)\right|dy.$$

We now split the region of integration in two:

$$\leq 2\sup_{x\in R} p(x)\int_{h_s|y|>\varepsilon} K(y)dy$$

$$+\sup_{h_s|y|\leq\varepsilon}\left|p(x_0-h_s y)-p(x_0)\right|. \tag{1-7}$$

Because $p(\cdot)$ is bounded, the first term in (1-7) converges to zero as $s\to\infty$ (under condition (1-2a)), then by letting $\varepsilon\to 0$ the last term in (1-7) converges to zero since $p(\cdot)$ is continuous. Now we show the convergence to zero of the variance (statement (1-6)).

Denoting by:

$$f_s(x) = K\left(\frac{x_0-x}{h_s}\right) - EK\left(\frac{x_0-X_0}{h_s}\right)\quad \text{for } x\in R \tag{1-8}$$

and $C(s_1,s_2) = E\, f_{s_1}(X_{s_1}) f_{s_2}(X_{s_2})$ for $s_1, s_2 \in [0,\infty)$ (1-9)

we may write the variance of $p_t(x_0)$ as:

$$\mathrm{Var}\; p_t(x_0) = \frac{1}{b_t^2} \int_0^t \int_0^t C(s_1,s_2) d_{s_1}\; d_{s_2} \tag{1-10}$$

Using the stationarity property of the X_t process, (1-9) becomes:

$$C(s_1,s_2) = E\, f_{s_1}(X_0) f_{s_2}(X_{|s_1-s_2|}) \;. \tag{1-11}$$

In order to use later on condition G_2 we introduce in (1-11) the transition probability operator:

$$C(s_1,s_2) = E(f_{s_1}(X_0)(H_{|s_1-s_2|} f_{s_2})(X_0)).$$

By using Schwarz inequality we get:

$$C(s_1,s_2) \le E^{1/2} f_{s_1}^2(X_0)\; E^{1/2}(H_{|s_1-s_2|} f_{s_2})^2(X_0). \tag{1-12}$$

Because X_t is a Markov process and $E\, f_{s_2}(X_0) = 0$, and more specifically from (1-1), (1-12) becomes

$$C(s_1,s_2) < \frac{\beta}{\alpha}^{|s_1-s_2|} \quad E^{1/2} f_{s_1}^2(X_0) E^{1/2} f_{s_2}^2(X_0). \tag{1-13}$$

By construction of $f_s(\cdot)$, expression (1-8), we have for any $s \in [0,\infty)$:

$$\frac{1}{h_s} E\, f_s^2(X_0) \le \frac{1}{h_s} E\, K^2\left(\frac{x_0 - X_0}{h_s}\right)$$

$$= \int_R K^2(y) p(x_0 - h_s y) dy$$

$$\le \sup_{x \in R} p(x) \int_R K^2(y) dy,$$

which is bounded since $K(\cdot)$ and $\int_R K(y)dy$ are bounded.

Denoting $C = \sup_{x\in R} p(x) \int_R K^2(y)dy$, (1-13) becomes:

$$C(s_1,s_2) < \frac{C}{\alpha} \sqrt{h_{s_1} h_{s_2}}\ \beta^{|s_1-s_2|}$$

and (1-10) becomes:

$$\text{Var } p_t(x_0) < \frac{C}{\alpha b_t^2} \int_0^t \int_0^t \sqrt{h_{s_1} h_{s_2}}\ \beta^{|s_1-s_2|}\ d_{s_1} d_{s_2},$$

which may be written:

$$= \frac{2C}{\alpha b_t^2} \int_0^t \int_{s_2}^t \sqrt{h_{s_1} h_{s_2}}\ \beta^{(s_1-s_2)}\ d_{s_1} d_{s_2},$$

because h_t is a decreasing function:

$$\leq \frac{2C}{\alpha b_t^2} \int_0^t h_{s_2} \int_{s_2}^t \beta^{(s_1-s_2)}\ d_{s_1} d_{s_2},$$

by changing of variable

$$= \frac{2C}{\alpha b_t^2} \int_0^t h_{s_2} \int_0^{t-s_2} \beta^s\ d_s d_{s_2},$$

which can be bounded by:

$$\leq \frac{2C}{\alpha b_t} \int_0^t \beta^s\ ds$$

$$< \frac{2C}{\alpha\ \ln \frac{1}{\beta}\ b_t}.$$

Since $\beta<1$ (see (1-1)), we must have $\ln \frac{1}{\beta} > 0$ and therefore the variance of $p_t(x_0)$ must tend to zero as b_t goes to infinity, which completes the proof of the theorem. #

We now study the properties of $p_t'(x_0)$ as an estimate of $p'(x_0)$:

Theorem 1-2 (quadratic mean convergence of $p_t'(x_0)$)

Let $\{X_t, t\in[0,\infty)\}$ be a stationary Markov process having a continuous and bounded initial density $p(\cdot)$ on R and satisfying condition G_2. Let $K(\cdot)$ be a continuous probability density function of bounded variation on R and such that $K'(\cdot)$ is bounded on R. Let h. be a positive function on R^+ such that condition (1-2) is verified and so is:

$$h_t^2 \, b_t \xrightarrow[t\to\infty]{} \infty . \tag{1-14}$$

For $t>0$, let

$$p_t'(x_0) = \frac{1}{b_t} \int_0^t \frac{1}{h_s} K'\left(\frac{x_0-X_s}{h_s}\right) ds . \tag{1-15}$$

If $p'(\cdot)$ is continuous and bounded on R then:

$$p_t'(x_0) \xrightarrow[t\to\infty]{q.m.} p'(x_0). \qquad \#$$

Proof

As in the proof of Theorem 1-1 we first show the asymptotic unbiasedness of $p_t'(x_0)$, a sufficient condition is that:

$$E \frac{1}{h_s^2} K'\left(\frac{x_0-X_0}{h_s}\right) - p'(x_0)$$

converges to zero as s goes to infinity. To show that point we use similar argument as in [Bhattacharya, 1967] or [Shuster 1969]. Because $K(\cdot)$ is of bounded variation, $\lim\limits_{|y|\to\infty} K(y)$ exist (see [Natanson 1955] p. 239) and these limits must be zero since $\int_R K(y)dy=1$, therefore integrating by parts we get:

$$E \frac{1}{h_s^2} K'\left(\frac{x_0-X_0}{h_s}\right) = \frac{1}{h_s} \int_R K\left(\frac{x_0-x}{h_s}\right) p'(x)dx \tag{1-16}$$

From (1-16), the above needed convergence to zero follows by the same argument as in the proof of theorem 1-1 (see inequality 1-7).

The convergence to zero of $\operatorname{Var} p_t'(x_0)$ as $t \to \infty$ follows in the same way as we have proved the convergence of $\operatorname{Var} p_t(x_0)$ but now under the stronger condition (1-14)

$$h_t^2 \, b_t \longrightarrow \infty$$
$$t \longrightarrow \infty .$$

The only point which remains to be shown is that $\int_R K'^2(y)dy$ is bounded.

This property follows from the fact that $K'(\cdot)$ is bounded and $K(\cdot)$ is of bounded variation which imply (see [Natanson 1955] p. 259):

$$\int_R |K'(y)|dy < \infty . \qquad \#$$

Remark 1-1

Both estimates $p_t(x_0)$ and $p_t'(x_0)$ defined in theorem 1-1 and 1-2 respectively are recursive, i.e. solutions to $(t>0)$:

$$\frac{d\, p_t(x_0)}{dt} = -\frac{h_t}{b_t}\, p_t(x_0) + \frac{1}{b_t} K\left(\frac{x_0 - X_t}{h_t}\right)$$

and

$$\frac{d\, p_t'(x_0)}{dt} = -\frac{h_t}{b_t}\, p_t'(x_0) + \frac{1}{b_t h_t} K'\left(\frac{x_0 - X_t}{h_t}\right) .$$

The initial conditions of these two differential equations can be arbitrary (when no apriori information is available), they do not affect the final value of both estimates.

2) Estimation of $p(x_0)$ and $p'(x_0)$ for Diffusion Processes

In the previous section we have shown the quadratic mean convergence of $p_t(x_0)$ and $p_t'(x_0)$ under the assumptions that the X_t Markov process is stationary and satisfies condition G_2.

Indeed, the stationarity assumption is not essential here, since we are dealing with asymptotic properties. A sufficient condition should be the existence of the limit of $p_{X_t}(\cdot)$ as t goes to infinity.

We now assume that the X_t process to be estimated is defined by problem (0-1) under A_1-A_2 and has a transition density $p_{X_t|X_0=a}(\cdot)$ which converges, for all $a\in R$, to a limiting density $p(\cdot)$ as t goes to infinity.

As we have seen in the introduction, such process is a Markov process. More, the limit of $p_{X_t}(\cdot)$ must be equal to $p(\cdot)$ since we have for all $x\in R$:

$$
\begin{aligned}
\lim_{t\to\infty} p_{X_t}(x) &= \lim_{t\to\infty} \int_R p_{X_t|X_0=a}(x)\, p_{X_0}(a)da \\
&= \int_R \lim_{t\to\infty} p_{X_t|X_0=a}(x)\, p_{X_0}(a)da \\
&= \int_R p(x)\, p_{X_0}(a)da = p(x).
\end{aligned}
$$

Hence, to have the stationarity of the X_t process defined by (0-1) under A_1-A_2 and the condition that the limiting density $p(\cdot)$ exists, it is necessary and sufficient to choose the initial density $p_{X_0}(\cdot)$ equal to $p(\cdot)$.

For the sake of simplicity, from now on we assume that the above X_t process is stationary.

In this section we derive sufficient conditions on the pair $(m(\cdot),\sigma(\cdot))$ to have the transition density convergence and condition G_2 satisfied.

Let us denote $P(x,t|a,s) = \mathcal{P}(X_t<x|X_s=a)$ the transition function of the unique X_t process solution of problem (0-1) where $a,x\in R$ and $t>s$.

Under A_1,A_2 (see section 0) and the additional condition:

$A_3: x\in R \quad \sigma(x)\geq\sigma_0>0,$

we know (see [Wong 1971] p. 173 prop. 7.1,(a) and (e)) that $P(x,t|\cdot,\cdot)$ is the unique solution of the backward equation of Kolmogorov:

$$\frac{1}{2}\,\sigma^2(a)\,\frac{\partial^2 P(x,t|a,s)}{\partial a^2} + m(a)\,\frac{\partial P(x,t|a,s)}{\partial a} = -\,\frac{\partial P(x,t|a,s)}{\partial s} \tag{2-1a}$$

with the terminal condition: $$\lim_{s\uparrow t} P(x,t|a,s) = \begin{cases} 1 & x>a \\ 0 & x<a \end{cases} \tag{2-1b}$$

and is absolutely continuous, that is $P(x,t|a,s)$ can be written as:

$$P(x,t|a,s) = \int_{-\infty}^{x} p(y,t|a,s)dy \quad a,x\in R \quad t>s.$$

Because the functions $m(\cdot)$ and $\sigma(\cdot)$ do not depend on time, we see from (2-1a) that $p(x,t|a,s)$ depends only on t-s and not on t and s separately. Denote by $p_a(\cdot,\cdot)$ $a\in R$, the transition density on $R\times R^+$ of the X_t process satisfying (0-1) (we will use the shorter notation $p_a(x,t)$ instead of $p_{X_{t+s}|X_s=a}(x)$; we can drop the s because of the stationarity of the transition density).

If in addition to A_1, A_2 and A_3, we assume that:

A_4:$m'(\cdot)$, $\sigma'(\cdot)$ and $\sigma''(\cdot)$ satisfy the conditions of type (0-2) and (0-3), then (see [Wong 1971] p. 173, prop. 7.1, (d)) $p_a(\cdot,\cdot)$ is the unique fundamental solution of the forward equation of Kolmogorov (Fokker-Planck equation):

$$\frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2(x)p_a(x,t)) - \frac{\partial}{\partial x}(m(x)p_a(x,t)) = \frac{\partial}{\partial t}p_a(x,t) \tag{2-2a}$$

with the initial condition $\lim\limits_{t\downarrow 0} p_a(x,t) = \delta(x-a)$. (2-2b)

We now want to show that under sufficient conditions on the functions $m(\cdot)$ and $\sigma(\cdot)$, $p_a(\cdot,\cdot)$, the unique solution of problem (2-2), converges to a limiting density as t goes to infinity. In other words we want to find stability conditions of problem (2-2). To find out these conditions we need the following lemma:

<u>Lemma 2-1</u> - (expansion formula)

Let $p_a(\cdot,\cdot)$ be the unique solution of problem (2-2), then $p_a(\cdot,\cdot)$ can be written in the form for $a,x\in R$, $t\in[0,\infty)$

$$p_a(x,t) = \pi(x)\int_R e^{-\lambda t}\sum_{j,k=1}^{2}\phi_j(x,\lambda)\,\phi_k(a,\lambda)\,d\,\rho_{jk}(\lambda) \tag{2-3}$$

Where $\pi(\cdot)$ is any positive solution of the equation:

$$\frac{1}{2}\frac{d}{dx}(\sigma^2(x)\pi(x)) = m(x)\pi(x) \qquad x\in R\ , \tag{2-4}$$

$\phi_1(\cdot,\lambda)$ and $\phi_2(\cdot,\lambda)$ are solutions of the Sturm-Liouville equation:

$$\frac{1}{2}\frac{d}{dx}\left(\sigma^2(x)\pi(x)\frac{du}{dx}(x)\right) + \lambda\pi(x)u(x)=0 \quad \lambda\in R,\ x\in R, \tag{2-5a}$$

and satisfying the conditions:

$$\phi_1(0,\lambda)=1 \qquad \phi_1'(0,\lambda)=0\ ,$$

$$\phi_2(0,\lambda)=0 \qquad \phi_2'(0,\lambda)=1\ ,$$

$(\rho_{jk}(\lambda))$ is the limiting matrix of the spectral matrix $(\rho_{\gamma,jk}(\lambda))$ associated with equation (2-5a) together with the boundary conditions (corresponding to the reflecting barriers in the regular case of problem (2-2)):

$$u'(-\gamma) = u'(\gamma)=0 \quad \gamma\in R \tag{2-5b}$$

once $\phi_1(\cdot,\lambda)$ and $\phi_2(\cdot,\lambda)$ are chosen as basis for the solutions of (2-5a), as γ goes to ∞. #

Proof

We shall give a proof by constructing a solution (see also solution of problem 9, Chapter 4, p. 178 in [Wong 1971]). We can verify that any function $f_a(\cdot,t)$ of the form

$$\pi(\cdot)\int_R e^{-\lambda t}\sum_{j,k=1}^{2}\phi_j(\cdot,\lambda)\psi_k(a,\lambda)\, d\,\rho_{jk}(\lambda)$$ is a solution of equation (2-5a) where $\pi(\cdot)$, $\phi_j(\cdot,\lambda)$ $j=1,2$ and $(\rho_{jk}(\lambda))$ are defined in the same way as in lemma 2-1, and $\psi_j(\cdot,\lambda)$ $j=1,2$ are any function of the same class as the ϕ's.

By setting t=0 in the above expression and using the expansion theorem (see [Coddington and Levinson 1955] p. 251 Theorem 5.2) the ψ's may be regarded as the transform of $f_a(x,0)$ by the means of the ϕ's i.e.:

$$\psi_j(a,\lambda) = \int_R f_a(x,0)\ \phi_j(x,\lambda)dx \quad j=1,2 \quad a\in R$$

If we now assume that $f_a(x,0)$ must be taken as $\delta(x-a)$ (i.e. $f_a(\cdot,\cdot)$ satisfy the initial condition of problem (2-2)) then we get $\psi_j(\cdot,\lambda) \equiv \phi_j(\cdot,\lambda)$, j=1,2 which completes the proof of the lemma.

#

Remark 2-1

We know (see [Coddington and Levinson 1955] p. 251 theorem 5.1) that the limiting matrix $(\rho_{jk}(\lambda))$ defined in lemma 2-1 always exists but could be non-unique (in the so called limit-circle case), actually we are not going to introduce more conditions to have the uniqueness, because we only need here the existence property.

#

Let Λ be the set of non-constancy points of $(\rho_{jk}(\lambda))$.

The set Λ is called the spectrum of the problem (2-5) with $\gamma\to\infty$. Despite the fact that the spectrum could not be completely defined (since ρ could be non-unique) we can say something about the nature of the spectrum.

We know that there exists an increasing sequence of eigenvalues $\{\lambda_{\gamma,n}\}$ and a complete orthonormal set of corresponding eigenfunctions $\{\theta_{\gamma,n}(\cdot)\}$ associated with the Sturm-Liouville problem (2-5) n=0,1,2,... . Using the boundary conditions (2-5b) we have:

$$\lambda_{\gamma,n} = \int_{-\gamma}^{\gamma} \frac{\sigma^2(z)}{2}\ \pi(z) \left(\frac{d\ \theta_{\gamma,n}(x)}{dx}\right)_z^2 dz,$$

which shows, since the integrant is non-negative, that

$$\lambda_{\gamma,n} \geq 0 \quad n=0,1,2\ldots\ .$$

Letting $\gamma\to\infty$, we see that the spectrum Λ cannot lie on the negative part of the real line.

Further, we can verify directly that for every $\gamma \in R$ $\lambda_{\gamma,0}=0$ and $\theta_{\gamma,0}(\cdot)$ is a constant, say θ_γ, such that:

$$\int_{-\gamma}^{\gamma} \theta_\gamma^2 \, \pi(z)dz=1.$$

To say something about $\lambda_{\gamma,0}$ as γ goes to infinity, we must consider the integrability of $\pi(\cdot)$. If $\pi(\cdot)$ is not integrable on R then zero cannot remain an eigenvalue as $\gamma \to \infty$ since the square integrability with respect to $\pi(\cdot)$ of the corresponding solution of problem (2-5) cannot be maintained any more.

We can now state the following assumption under which we shall show the convergence of the transition density to a limiting density:

A_6: The pair $(m(\cdot),\sigma(\cdot))$ is such that $\pi(\cdot)$, the solution of equation (2-4), is integrable on R.

For example, $m(x) = Ax$ together with $\sigma(x)=B$, where A and B are two constants such that $A<0$ and $B\neq 0$, satisfy A_5.

<u>Lemma 2-2</u> (convergence to a limiting density)

Let $\{X_t, t\in[0,\infty)\}$ be the process defined by problem (0-1) under A_1-A_4, then under the additional assumption A_5, the transition density $p_a(\cdot,t)$ of the X_t process converges for all $a\in R$ to a continuous limiting density $p(\cdot)$ on R solution of equation (2-4) as t goes to infinity.

#

<u>Proof</u>

From the above results concerning the spectrum and using lemma 2-1, we may conclude that $p_a(\cdot,\cdot)$, the unique solution of problem (2-2) converges identically to zero if $\pi(\cdot)$ is not integrable and to a density $p(\cdot)$ if A_5 is satisfied. This can be seen more explicitly in letting $t\to\infty$ in expression (2-3).

If zero is not an eigenvalue then there are no jumps in the matrix $(\rho_{j,k}(0))$ and the right hand side of (2-3) must vanish as $t\to\infty$. If zero is an eigenvalue ($\lambda_{\infty,0}=0$), then we have for all a, $x\in R$.

$$\lim_{t\to\infty} p_a(x,t) = p(x)\left(\sum_{j,k=1}^{2} \phi_j(x,0)\ \phi_k(a,0)r_j r_k\right)$$

where $r_j r_k$ is the jump of $\rho_{jk}(0)$. By construction the corresponding eigenfunction $\theta_{\infty,0}(\cdot)$ can be written for $x\in R$:

$$\theta_{\infty,0}(x) = r_1\phi_1(x,0) + r_2\phi_2(x,0).$$

Because of the conditions on the ϕ's at x=0 and the fact that $\theta_{\infty,0}(x)=1$, we have $r_1=1$ and $r_2=0$ which proves that $\lim_{t\to\infty} p_a(x,t) = p(x)$ under A_5.

Finally, since the limiting density $p(\cdot)$ is solution of equation (2-4), $p(\cdot)$ must be continuous on R under A_2.

#

Indeed, under specific assumptions on $m(\cdot)$ and $\sigma(\cdot)$ we can say more about the nature of the spectrum.

Lemma 2-3 (nature of the spectrum)

Let, for $x\in R$:

$$\mu(x) = \frac{m^2(x)}{2\sigma^2(x)} + \frac{m'(x)}{2} - \frac{m(x)\sigma'(x)}{\sigma(x)} + \frac{\sigma'^2(x)}{8} - \frac{\sigma(x)\sigma''(x)}{4} \qquad (2\text{-}6)$$

If $\min(\lim_{x\to -\infty} \mu(x), \lim_{x\to\infty} \mu(x)) = \mu$ for some $\mu\in(-\infty,\infty]$, then Λ can only be discrete in the interval $(-\infty,\mu)$. #

Proof

We shall give an outline of the proof. By using the standard transformation (see [Birkhoff and Rota 1969] p. 296 or [Titchmarch 1946] p. 22) the equation (2-5a) can be rewritten in the form of the Schroedinger equation:

$$\frac{d^2v(y)}{dy^2} + (\lambda - q(y))\, v(y) = 0 \tag{2-7}$$

with

$$y(x) = \int_0^x \frac{\sqrt{2}}{\sigma(z)}\, dz \quad x\in R\ , \tag{2-8}$$

$$v(y(x)) = \left(\frac{\sigma(x)\pi(x)}{\sqrt{2}}\right)^{1/2} u(x),$$

and

$$q(y(x)) = \frac{1}{\sqrt{\sigma(x)\pi(x)}} \frac{d^2\sqrt{\sigma(x)\pi(x)}}{dy^2}$$

The spectrum being unchanged in the transformation, we may study the nature of the spectrum from equation (2-7). We know (see [Coddington and Levinson 1955] problem 2 p. 255, [Titchmarch 1946] p. 113 or [Schiff 1955] Ch. II, sec. 8) that if the potential function $q(y)$ is bounded from below say by μ, as y tends to either end points of its domain then the spectrum is discrete in the interval bounded above by μ.

More precisely, if $\min(\lim_{x\to -\infty} q(y(x)), \lim_{x\to\infty} q(y(x))) = \mu$ for some $\mu\in[-\infty,\infty)$, then Λ can only be discrete in the interval

$(-\infty,\mu)$. Using (2-4), (2-7) and (2-8) we obtain $q(y(x)) = \mu(x)$ $x \in R$, where $\mu(x)$ is given by (2-6) which completes the proof of the lemma. #

We now state the last assumption on $m(\cdot)$ and $\sigma(\cdot)$.

A_6: the pair$(m(\cdot),\sigma(\cdot))$ is such that μ of the lemma 2-3 is strictly positive.

As an example of functions $m(\cdot)$ and $\sigma(\cdot)$ satisfying assumption A_6 we can mention the class of functions such that $m(x) \sim Ax^{\alpha}$ and $\sigma(x) \sim Bx^{\beta}$ as $|x| \to \infty$ with $\alpha,\beta = 0,1$, $A<0$ and $B \neq 0$. For this class of functions we may simplify the study of $\mu(x)$ at the infinity by noting from (2-6) that:

if $\alpha=0,1$; $\beta=0$

then

$$\mu(x) \sim \frac{A^2}{2B^2} \; x^{2\alpha}$$

if $\alpha=\beta=1$

then

$$\mu(x) \sim \frac{1}{2B^2} \left(A - \frac{B^2}{2}\right)^2$$

if $\alpha=0$, $\beta=1$

then

$$\mu(x) \sim \frac{B^2}{8}$$

We notice that the exponent in the first expression is even and the coefficients in the three cases are always strictly positive so is μ of lemma 2-3.

Assumption A_6, as it can be seen from lemma 2-3, implies that the spectrum Λ can only be discrete at the beginning of the interval $[0,\infty)$. Such result will allow us to prove the following lemma:

Lemma 2-4 (the condition G_2)

Let $\{X_t, t\in[0,\infty)\}$ be the process defined by problem (0-1) under A_1-A_5, then under the additional assumption A_6, the X_t process satisfies condition G_2 (see section 1). #

Proof

Using expression (2-3) of lemma 2-1 we may write, for any function $f(\cdot)$ on R which is Borel measurable, bounded and such that $Ef(X)=0$, and any $s>0$:

$$(H_s f)(a) = \int_R \sum_{j,k=1}^{2} \phi_k(a,\lambda)e^{-\lambda s} \int_R f(x)\,\phi_j(x,\lambda)\,p(x)dx\,d\rho_{jk}(\lambda),$$

where H_s is the transition probability operator defined in section 1.

Using the Parseval equality we get:

$$E(H_s f)^2(X) = \int_R e^{-2\lambda s} \sum_{j,k=1}^{2} g_j(\lambda)\,g_k(\lambda)\,d\rho_{jk}(\lambda)$$

with

$$g_j(\lambda) = \int_R f(x)\,\phi_j(x,\lambda)\,p(x)dx \qquad j=1,2. \tag{2-9}$$

Since, under A_5, zero is an eigenvalue and the corresponding jumps of $\rho_{jk}(\lambda)$ are 1 for j=k=1 and zero otherwise, we have:

$$E(H_s f)^2(X) = \left(\int_R f(x)p(x)dx\right)^2 + \int_{R-\{0\}} e^{-2\lambda s} \sum_{j,k=1}^{2} g_j(\lambda)$$

$$g_k(\lambda)\,d\rho_{jk}(\lambda).$$

Since $Ef(X)=0$, we have:

$$E(H_s f)^2(X) = \int_{R-\{0\}} e^{-2\lambda s} \sum_{j,k=1}^{2} g_j(\lambda)\ g_k(\lambda)\ d\rho_{jk}(\lambda).$$

Let λ_0 be the lower bound of $\Lambda-\{0\}$, then

$$E(H_s f)^2(X) \leq e^{-2\lambda_0 s} \int_R \sum_{j,k=1}^{2} g_j(\lambda)\ g_k(\lambda)\ d\rho_{jk}(\lambda).$$

Recalling expression (2-9) and using once more the Parseval equality we get the following bound:

$$E(H_s f)^2(X) \leq e^{-2\lambda_0 s} \int_R f^2(x)\ p(x)dx = e^{-2\lambda_0 s}\ Ef^2(X). \quad (2\text{-}10)$$

Inequality (2-10) implies:

$$|H_s|_2 \leq e^{-\lambda_0 s} \quad (2\text{-}11)$$

Under A_6, we have seen that the spectrum Λ can only be discrete at the beginning of $[0,\infty)$, therefore the lower bound λ_0 of $\Lambda-\{0\}$ must be strictly positive, so that from (2-11) condition G_2 is certainly satisfied. #

Finally, using lemma 2-4, we may rewrite theorem 1-1 and 1-2 of section 1, for the special case of the X_t Markov process defined by problem (0-1).

<u>Theorem 2-1</u> (quadratic mean convergence of $p_t(x_0)$)

Let $\{X_t, t\in[0,\infty)\}$ be the process defined by problem (0-1) under A_1-A_6.

Let $K(\cdot)$ and $h.$ be the functions defined in theorem 1-1.

Let $p_t(x_0)$ be the estimate defined by (1-4). If $p(\cdot)$ is bounded on R then:

$$p_t(x_0) \xrightarrow[t\to\infty]{q.m.} p(x_0). \qquad \#$$

Theorem 2-2 (quadratic mean convergence of $p_t'(x_0)$)

Let $\{X_t, t\in[0,\infty)\}$ be the process defined by problem (0-1) under A_1-A_6.

Let $K(\cdot)$ and h. be the function defined in theorem 1-2.

Let $p_t'(x_0)$ be the estimate defined by (1-15).

If $p(\cdot)$ and $p'(\cdot)$ are bounded on R then:

$$p_t'(x_0) \xrightarrow[t\to\infty]{q.m.} p'(x_0). \qquad \#$$

In the statement of theorem 2-2, we may forget the condition that $p'(\cdot)$ is continuous because we are now under assumptions A_2 and A_5.

Remark 2-2

Assumptions A_2-A_6 are only sufficient conditions. The case $m(x)= -\text{sgn}(x)$, and $\sigma(x)=1$ for example which does not satisfy the A_2 condition still works on (see solution of problem 12, Chapter 4, p. 179 in [Wong 1971]). #

From the above results (more specifically from lemma 2-4) we may draw figure 2-1.

3) Estimation of σ^2

We now assume that the function $\sigma(\cdot)$ in (0-1a) takes a constant value σ. By using a property of the quadratic variation of the X_t process defined by problem (0-1) we suggest and prove the convergence of a recursive estimate of σ^2.

<u>Theorem 3-1</u> (quadratic mean convergence of σ_n^2)

Let $\{X_t, t\in[0,\infty)\}$ be the solution to the stochastic differential equation (0-1a) where $m(\cdot)$ satisfy (0-2) and (0-3), and $\sigma(x)=\sigma$ $\forall x\in R$. Let the initial condition satisfy $EX^4<\infty$.

Let $\{\tau_i\}_{i=1}^{\infty}$ be a bounded sequence of positive numbers such that:

$$\tau_i \underset{i\to\infty}{\to} 0 \tag{3-1}$$

and $\{t_i\}_{i=1}^{\infty}$ a sequence such that $0\leq t_1$ and t_i, and $t_i+\tau_i \leq t_{i+1}$ $i=1,2\ldots$.

Let, $n=1,2,\ldots$

$$\sigma_n^2 = \frac{1}{n} \sum_{i=1}^{m} \frac{1}{\tau_i}(x_{t_i+\tau_i} - X_{t_i})^2 \tag{3-2}$$

then: $\sigma_n^2 \xrightarrow[n\to\infty]{q.m.} \sigma^2$. #

<u>Proof</u> (here, we follow closely the same arguments as those given in [Wong and Zakai 1965]).

Let Q_n be equal to $\sigma_n^2-\sigma^2$, we have to prove that $EQ_n^{\ 2}$ tends to zero as n goes to infinity.

By construction of σ_n^2:

$$Q_n = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\tau_i} (X_{t_i+\tau_i} - X_{t_i})^2 - \sigma^2,$$

writing equation (0-1a) in its integral form, we get:

$$Q_n = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\tau_i} \left(\int_{t_i}^{t_i+\tau_i} m(x_s)ds + \sigma \int_{t_i}^{t_i+\tau_i} dWs\right)^2 - \sigma^2 .$$

Denoting by

$$A_n = \sigma^2 \sum_{i=1}^{n} \frac{1}{\tau_i} \left(\int_{t_i}^{t_i+\tau_i} dW_s \right)^2 - n\sigma^2 \tag{3-3}$$

and

$$B_n = \sum_{i=1}^{n} \frac{1}{\tau_i} \left(\int_{t_i}^{t_i+\tau_i} m(X_s)ds \right)^2 , \tag{3-4}$$

we may rewrite Q_n as:

$$Q_n = \frac{1}{n} \left(A_n + B_n + 2 \sum_{i=1}^{n} \frac{\sigma}{\tau_i} \int_{t_i}^{t_i+\tau_i} m(X_s)ds \int_{t_i}^{t_i+\tau_i} dW_s \right)$$

By using Schwarz inequality in the expression of Q_n^2 we get:

$$Q_n^2 \leq \frac{3}{n^2} \left(A_n^2 + B_n^2 + 4\sigma^2 B_n \sum_{i=1}^{n} \frac{1}{\tau_i} \left(\int_{t_i}^{t_i+\tau_i} dWs \right)^2 \right),$$

introducing A_n in the last sum:

$$= \frac{3}{n^2} (A_n^2 + B_n^2 + 4n\sigma^2 B_n + 4A_n B_n) .$$

Taking the expectation of both sides in the last expression and using the Schwart inequality we get:

$$EQ_n{}^2 \leq \frac{3}{n^2} (EA_n^2 + EB_n^2 + 4n\sigma^2 E^{1/2} B_n^2 + 4E^{1/2} A_n^2 E^{1/2} B_n^2)$$

To prove the convergence to zero of EQ_n^2 we only have to show that $\frac{1}{n^2} EA_n^2$ and $\frac{1}{n^2} EB_n^2$ converge to zero.

From (3-3) we may rewrite A_n as:

$$A_n = \sigma^2 \sum_{i=1}^{n} \frac{1}{\tau_i} \left(\left(\int_{t_i}^{t_i+\tau_i} dWs \right)^2 - \int_{t_i}^{t_i+\tau_i} ds \right) .$$

Using a property of the stochastic integral (see [Ito 1951], or [Wong and Zakai 1965] expression 4, p. 104) A_n becomes:

$$= 2\,\sigma^2 \sum_{i=1}^{n} \frac{1}{\tau_i} \int_{t_i}^{t_i+\tau_i} \int_{t_i}^{s} dW_u dW_s .$$

Taking the expectation of A_n^2 and using properties of the Brownian motion:

$$EA_n^2 = 4\sigma^4 \sum_{i=1}^{n} \frac{1}{\tau_i^2} \int_{t_i}^{t_i+\tau_i} E\left(\int_{t_i}^{s} dW_u\right)^2 ds$$

$$= 4\sigma^4 \sum_{i=1}^{n} \frac{1}{\tau_i^2} \int_{t_i}^{t_i+\tau_i} (s-t_i)ds$$

Therefore

$$\frac{1}{n^2} EA_n^2 = \frac{2\sigma^4}{n}$$

and the right hand side converges to zero as n goes to ∞.

From (3-4) and using Schwarz inequality:

$$B_n^2 \leq n \sum_{i=1}^{n} \frac{1}{\tau_i^2} \left(\int_{t_i}^{t_i+\tau_i} m(X_s)ds\right)^4$$

$$\leq n \sum_{i=1}^{n} \tau_i \int_{t_i}^{t_i+\tau_i} m^4(X_s)ds$$

Taking the expectation of B_n^2

$$EB_n^2 \leq nEm^4(X_0) \sum_{i=1}^{n} \tau_i^2$$

Since EX_0^4 is bounded and $m(\cdot)$ satisfy (0-3), $Em^4(X_0)$ is bounded, say by C. Therefore

$$\frac{1}{n^2} EB_n^2 \leq \frac{C}{n} \sum_{i=1}^{n} \tau_i^2$$

and the right hand side converges to zero and n goes to ∞ under the condition (3-1) $\tau_i \underset{i \to \infty}{\to} 0$ because of a property of the sequence of arithmetic means.

#

Remark 3-1

The estimate σ_n^2 defined in theorem 3-1 is recursive, i.e. solution to:

$$\sigma_n^2 = \frac{n-1}{n} \sigma_{n-1}^2 + \frac{1}{n\tau_n} (X_{t_n+\tau_n} - X_{t_n})^2$$

with n=1,2...

The choice of the sequence $\{t_i\}$ can be determined by practical consideration (e.g. $t_{i+1}-t_i$ equals to computational time of the above difference equation).

4) Estimation of $m(x_0)$

In section 2 we have seen that the limiting density $p(\cdot)$, when it exists, must verify equation (2-4). So, since $p(\cdot)$ is strictly positive on R, we are able to express the function $m(\cdot)$ in term of $p(\cdot)$. At any point x_0 of R we have:

$$m(x_0) = \frac{1}{2}(\sigma^{2'}(x_0)+\sigma^2(x_0)q(x_0)), \tag{4-1a}$$

where

$$q(x_0) = \frac{p'(x_0)}{p(x_0)} \tag{4-1b}$$

Using the results of section 2, we may find a procedure to estimate $q(x_0)$.

Theorem 4-1 (convergence in probability of $q_t(x_0)$)

Let $\{X_t, t\in[0,\infty)\}$ be the process defined by problem (0-1) under A_1-A_6.

Let $\varepsilon>0$.

Let $K_1(\cdot)$ and $K_2(\cdot)$ be the functions $K(\cdot)$ defined in Theorem 1-1 and 1-2 respectively. Let h. be a positive function on R^+ satisfying conditions (1-2) and (1-14).

For $t>0$, let

$$q_t(x_0) = \frac{\int_0^t \frac{1}{h_s} K_2'\left(\frac{x_0-X_s}{h_s}\right)ds}{\int_0^t K_1\left(\frac{x_0-X_s}{h_s}\right)ds+\varepsilon} \quad . \tag{4-2}$$

If $p(\cdot)$ and $p'(\cdot)$ are bounded on R, then

$$q_t(x_0) \xrightarrow[t\to\infty]{p.} q(x_0) .$$ #

Proof

Multiplying numerator and denominator of (4-2) by $\frac{1}{b_t}$ $t>0$ and using the estimates given by (1-4) and (1-15) we get:

$$q_t(x_0) = \frac{p_t'(x_0)}{p_t(x_0) + \frac{\varepsilon}{b_t}}$$

Because the mean square convergence implies the convergence in probability, we may apply the property of the latter relative to continuous functions (here to the quotient), see [Lukacs 1975] p. 43.

Therefore using the results of theorems 2-1 and 2-2 we have:

$$q_t(x_0) \xrightarrow[t\to\infty]{p.} \frac{p'(x_0)}{p(x_0)}$$

which completes the proof by considering (4-1b). #

Finally we can state our last result which is relative to a local estimate of the function $m(\cdot)$:

Corollary 4-1 (convergence in probability of $m_t(x_0)$)

Let $\sigma(\cdot)$ be the function defined in section 0 (which we assume known).

Let, for $t>0$

$$m_t(x_0) = \frac{1}{2}(\sigma^{2'}(x_0)+\sigma^2(x_0)q_t(x_0)) \tag{4-3}$$

where $q_t(x_0)$ is given by (4-2), then under the conditions of theorem 4-1:

$$m_t(x_0) \xrightarrow[t\to\infty]{p.} m(x_0).$$ #

Proof

The convergence in probability of $m_t(x_0)$ follows from Theorem 4-1 and considering expressions (4-1a) and (4-3). #

In the case when the function $\sigma(\cdot)$ is unknown but takes a constant value σ, we may use the result of the previous section.

Corollary 4-2 (convergence in probability of $m_{t,n}(x_0)$)

Let $\sigma(x) = \sigma\ \forall x \in R$, (but σ unknown). Let $EX_4 < \infty$, and let σ_n^2 be the estimate defined by (3.2).

Let, for $t>0$ and $n=1,2\ldots$

$$m_{t,n}(x_0) = \frac{1}{2}\sigma_n^2 q_t(x_0) \tag{4-4}$$

where $q_t(x_0)$ is given by (4-2), then under the conditions of theorem 4-1:

$$m_{t,n}(x_0) \xrightarrow[t,n\to\infty]{P.} m(x_0). \qquad \#$$

Proof

The convergence in probability of $m_{t,n}(x_0)$ follows by using the theorems 3-1 and 4-1 and considering expressions (4-1a) and (4-4). #

Remark 4-1

The previous results are true under the specific condition that $p(\cdot)$ and $p'(\cdot)$ are bounded (see theorem 4-1). Actually we should introduce a last assumption, say A_7, on the pair $(m(\cdot),\sigma(\cdot))$ such that the above condition is satisfied.

When $\sigma(\cdot)$ takes a constant value on R we can verify that under the following condition:

$$\min\left(\lim_{x\to -\infty} m(x),\ -\lim_{x\to\infty} m(x)\right)>0$$

assumptions A_5 and A_6 are satisfied and $p(\cdot)$ and $p'(\cdot)$ are bounded on R.

This can be seen in writing explicitly the solutions of equation (2-4):

$$\pi(x) = C\, e^{\int_0^x \frac{2m(z)}{\sigma^2} dz} \qquad \#$$

Many functions $K(\cdot)$ satisfying conditions of Theorem 1-1 and 1-2 have been proposed in the past (see [Parzen 1962] and [Rosenblatt 1971]). Perhaps the simplest choices for $K_1(\cdot)$ and $K_2(\cdot)$ in (4-1) would be for $y\in R$:

$$K_1(y) = \frac{1}{2} I_{(-1,1]}(y)$$

and

$$K_2(y) = (1-|y|) I_{(-1,1]}(y)$$

i.e. $K_2'(y) = \mathrm{sgn}(y) I_{(-1,1]}(y) = 2\mathrm{sgn}(y) K_1(y)$

where

$$I_A(y) = 1 \text{ if } y \in A$$

$$= 0 \text{ otherwise}$$

and $\mathrm{sgn}(y) = 1$ if $y \geq 0$

$$= -1 \text{ otherwise.}$$

5) Conclusion

In Table 5-1 we give a summary of the main results obtained in this paper. Actually, many questions remain unanswered. Among them, how the identification procedure proposed here could be extended to multidimensional stochastic processes?

Could assumption A_5 and A_6 be related to one another?

Under which additional assumptions can we prove the convergence to zero of the integrated mean square error or the convergence with probability one of our estimates? Can we say something about the rate of convergence to zero of the mean square error?

Acknowledgement

I would like to express my thanks to Professor E. Wong for his very helpful guidance and many suggestions offered during my stay at the U. C. Berkeley.

I am especially grateful for instructive and illuminating conversations on diffusion processes.

The idea suggested by Professor Wong of using expansion formula for transition density representation in the density estimate convergence proof has been of great importance in this work.

I would like to thank J. Aguilar-Martin and H. T. Nguyen for their useful criticisms.

Special thanks are due to Martha for preparing the copy of the manuscript.

Finally, I would like to acknowledge the financial support provided by the National Science Foundation Grant No. 01P75-04371 and the "Centre National de la Recherche Scientifique."

References

1. Aguilar-Martin, J. (1974)*: "Estimation et filtrage dans la modélisation stochastique en automatique," Thèse, Université Paul Sabatier de Toulouse.

2. Alengrin, G. (1974)*: "La théorie du filtrage non linéaire et ses applications au traitement du signal et à l'identification en automatique," Thèse, Université Paul Sabatier de Toulouse.

3. Banon, G. (1971)*: "Etude d'algorithmes d'estimation des paramètres pour l'identification adaptive en temps réel des processus linéaires perturbés par un bruit corrélé," Thèse, Université Paul Sabatier de Toulouse.

4. Bhattacharya P. K. (1967): "Estimation of a probability density function and its derivatives," Sankhyā, Ser. A 29, p. 373-382.

5. Birkhoff, G. and Rota, G. C. (1969): "Ordinary differential equations," Blaisdell.

6. Bochner, S. (1955): "Harmonic analysis and the theory of probability," University of California.

7. Coddington, E. A. and Levinson, N. (1955): "Theory of ordinary differential equations," McGraw-Hill.

8. Deheuvels, P. (1973, 16 Avril): "Sur l'estimation séquentielle de la densité," C. R. Acad. Sc. Paris, t. 276, série A, p. 1119-1121.

9. Itô, K. (1951): "On a formula concerning stochastic differentials," Nagoya math. J. 3, p. 55-65.

10. Kalman, R. E. and Bury, R. S. (1961): "New results in linear filtering and prediction theory," Transactions of A.S.M.E., J. of Basic Engineering, 83D, p. 95-108.

11. Lukacs, E. (1975): "Stochastic convergence," Academic press, second edition.

12. Natanson, I. P. (1955): "Theory of Functions of a real variable," Ungar, N.Y.

13. Parzen, E. (1962): "On estimation of a probability density function and mode," Ann. Math. Stat. 33, p. 1065-1076.

14. Rosenblatt, M. (1956): "Remarks on some non-parametric estimates of a density function," Ann. Math. Stat. 27, p. 832-837.

15. Rosenblatt, M. (1970): "Density estimates and Markov sequences," Nonparametric Techniques in Statistical Inference, M. Puri, p. 199-210.

16. Rosenblatt, M. (1971): "Curve estimates," Ann. Math. Stat. 42, p. 1815-1842.

17. Roussas, G. (1969): "Nonparametric estimation in Markov processes," Ann. Inst. Statist. Math. 21, p. 73-87.

18. Salut, G. (1976)*: "Identification optimale des systèmes linéaires stochastiques," Thèse, Université Paul Sabatier de Toulouse.

19. Schiff, L. I. (1955): "Quantum mechanics," McGraw-Hill.

20. Schuster, E. F. (1969): "Estimation of a probability density function and its derivatives," Ann. Math. Stat. 40, p. 1187-1195.

21. Titchmarsh, E. C. (1946): "Eigenfunction expansions associated with second-order differential equations," Oxford.

22. Wong, E. (1964): "The construction of a class of stationary Markoff processes," Proc. Symp. in Appl. Math., Am. Math. Soc., 16, p. 264-276.

23. Wong, E. (1971): "Stochastic processes in information and dynamical systems," McGraw-Hill.

24. Wong, E. and Zakai, M. (1965): "The oscillation of stochastic integrals," Z. Wahrscheinlichheitstheorie Verw. Geb. 4, p. 103-112.

25. Yamato, H. (1972): "Sequential estimation of a continuous probability density function and mode," Bull. Math. State. Jap. 14, p. 1-12.

---

* may be found at

ESRIN, Space documentation service, Via Galileo Galilei
Casella Postale 64, I 00044 FRASCATI (ROMA), ITALY.

Figure 2-1 (consequence of lemma 2-4)

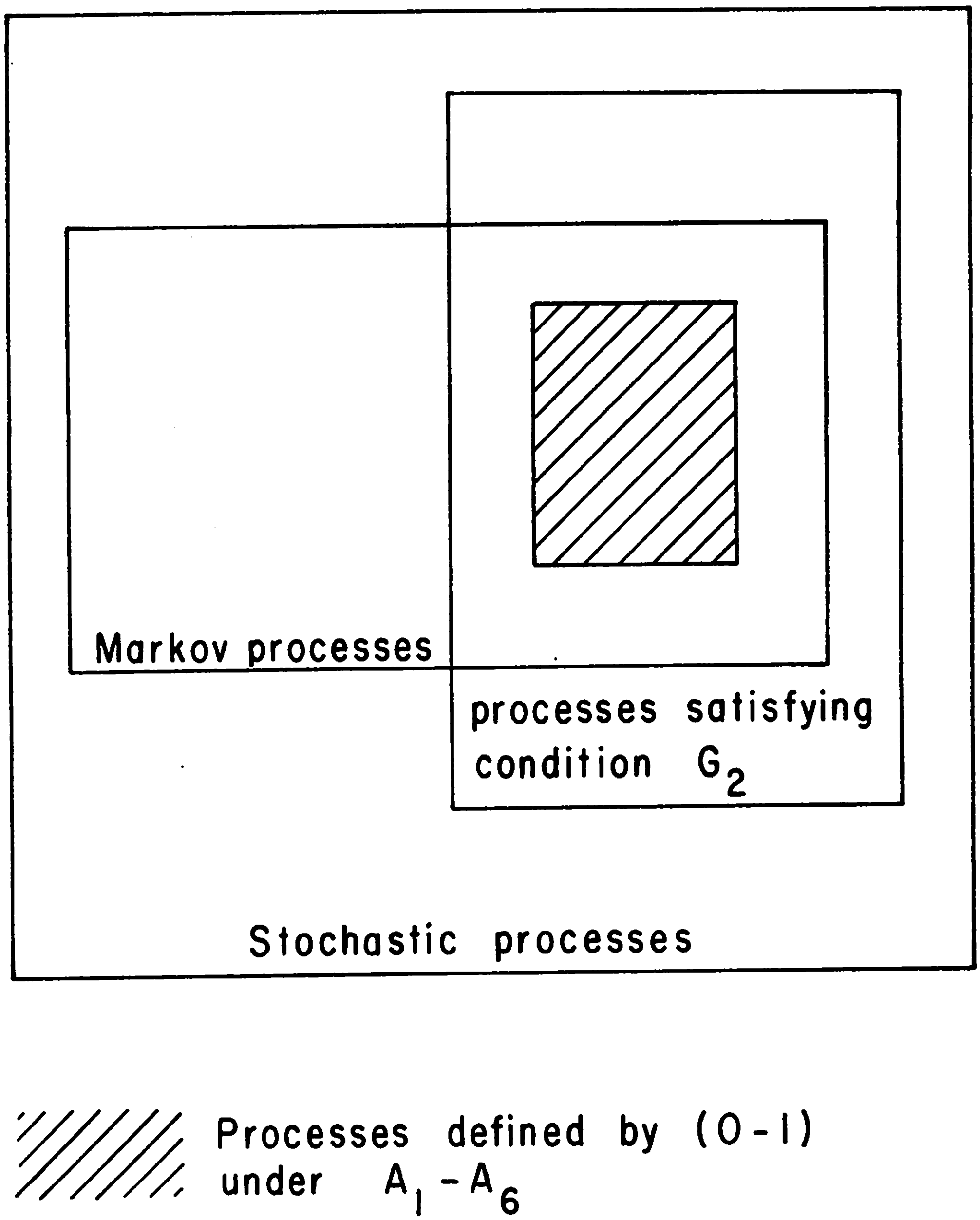

Table 5-1 (Summary of the main results)

Section	
0	$\left.\begin{array}{l} X_t \text{ defined by } (0\text{-}1) \\ A_1\text{-}A_2 \end{array}\right\} \Rightarrow \left\{ X_t \text{ is unique and Markov} \right.$
1	$\left.\begin{array}{l} X_t \text{ is stationary Markov} \\ X_t \text{ has an initial density} \\ p(\cdot) \text{ and satisfies } G_2 \end{array}\right\} \overset{\text{Th. 1-1, 1-2}}{\Rightarrow} \left\{\begin{array}{l} p_t(x_0)\to p(x_0) \\ p'_t(x_0)\to p'(x_0) \end{array}\right.$
2	$\left.\begin{array}{l} \left.\begin{array}{l} X_t \text{ defined by} \\ (0\text{-}1)\ A_1\text{-}A_3 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} X_t \text{ has a transition} \\ \text{density } p_a(\cdot,\cdot) \end{array}\right\} \\ A_4 \end{array}\right\} \Rightarrow \left.\left\{\begin{array}{l} p_a(\cdot,\cdot) \text{ is the} \\ \text{unique solution} \\ \text{of } (2\text{-}2) \end{array}\right\} \right. \overset{\text{lem. 2-1, 2-2}}{\Rightarrow} p_a(\cdot,t)\to p(\cdot)$ A_5 $\left.\begin{array}{l} p_a(\cdot,\cdot) \text{ is the} \\ \text{unique solution of} \\ (2\text{-}2) \\ A_5\text{-}A_6 \end{array}\right\} \overset{\text{lem 2-1, 2-3, 2-4}}{\Rightarrow} \{G_2\} \overset{\text{th. 2-1, 2-2}}{\Rightarrow} \left\{\begin{array}{l} p_t(x_0)\to p(x_0) \\ p'_t(x_0)\to p'(x_0) \end{array}\right.$
3	$\left.\begin{array}{l} X_t \text{ defined by } (0\text{-}1) \\ \sigma(x)=\sigma \\ EX^4<\infty \end{array}\right\} \overset{\text{Th. 3-1}}{\Rightarrow} \left\{ \sigma_n^2\to\sigma^2 \right.$
4	$\left.\begin{array}{l} X_t \text{ defined by } (0\text{-}1) \\ A_1\text{-}A_6 \end{array}\right\} \overset{\text{Th. 4-1}}{\Rightarrow}$ $\left.\begin{array}{l} q_t(x_0)\to q\ (x_0) \\ \sigma(\cdot) \text{ known} \end{array}\right\} \overset{\text{Cor. 4-1}}{\Rightarrow} \left\{ m_t(x_0)\to m(x_0) \right.$ $\left.\begin{array}{l} q_t(x_0)\to q\ (x_0) \\ \sigma_n\to\sigma \end{array}\right\} \overset{\text{Cor. 4-2}}{\Rightarrow} \left\{ m_{t,n}(x_0)\to m(x_0) \right.$

Table 5-1 (Summary of the main results)

Section	
0	$\left.\begin{array}{l} X_t \text{ defined by (0-1)} \\ A_1\text{-}A_2 \end{array}\right\} \Rightarrow \left\{ X_t \text{ is unique and Markov} \right.$
1	$\left.\begin{array}{l} X_t \text{ is stationary Markov} \\ X_t \text{ has an initial density} \\ p(\cdot) \text{ and satisfies } G_2 \end{array}\right\} \overset{\text{Th. 1-1, 1-2}}{\Rightarrow} \left\{\begin{array}{l} p_t(x_0) \to p(x_0) \\ p_t'(x_0) \to p'(x_0) \end{array}\right.$
2	$\left.\begin{array}{l} X_t \text{ defined by} \\ \text{(0-1) } A_1\text{-}A_3 \end{array}\right\} \Rightarrow \left.\begin{array}{l} \left\{\begin{array}{l} X_t \text{ has a transition} \\ \text{density } p_a(\cdot,\cdot) \end{array}\right. \\ A_4 \end{array}\right\} \Rightarrow \left.\begin{array}{l} \left\{\begin{array}{l} p_a(\cdot,\cdot) \text{ is the} \\ \text{unique solution} \\ \text{of (2-2)} \end{array}\right. \\ A_5 \end{array}\right\} \overset{\text{lem. 2-1, 2-2}}{\Rightarrow} p_a(\cdot,t) \to p(\cdot)$
	$\left.\begin{array}{l} p_a(\cdot,\cdot) \text{ is the} \\ \text{unique solution of} \\ \text{(2-2)} \\ A_5\text{-}A_6 \end{array}\right\} \overset{\text{lem 2-1, 2-3, 2-4}}{\Rightarrow} \left\{ G_2 \right\} \overset{\text{th. 2-1, 2-2}}{\Rightarrow} \left\{\begin{array}{l} p_t(x_0) \to p(x_0) \\ p_t'(x_0) \to p'(x_0) \end{array}\right.$
3	$\left.\begin{array}{l} X_t \text{ defined by (0-1)} \\ \sigma(x)=\sigma \\ EX^4<\infty \end{array}\right\} \overset{\text{Th. 3-1}}{\Rightarrow} \left\{ \sigma_n^2 \to \sigma^2 \right.$
4	$\left.\begin{array}{l} X_t \text{ defined by (0-1)} \\ A_1\text{-}A_6 \end{array}\right\} \overset{\text{Th. 4-1}}{\Rightarrow}$
	$\left.\begin{array}{l} q_t(x_0) \to q\ (x_0) \\ \sigma(\cdot) \text{ known} \end{array}\right\} \overset{\text{Cor. 4-1}}{\Rightarrow} \left\{ m_t(x_0) \to m(x_0) \right.$
	$\left.\begin{array}{l} q_t(x_0) \to q\ (x_0) \\ \sigma_n \to \sigma \end{array}\right\} \overset{\text{Cor. 4-2}}{\Rightarrow} \left\{ m_{t,n}(x_0) \to m(x_0) \right.$