

Evaluation of Likelihood Functions†

TYRONE E. DUNCAN

*Information and Control Engineering, The University of Michigan,
Ann Arbor, Michigan 48104*

An expression is obtained for the likelihood function for the detection of a stochastic signal (diffusion process) in white noise. A stochastic differential equation is then obtained for the evolution of the likelihood function and the coefficients of this differential equation are related to a corresponding nonlinear filtering problem. Some extensions are noted to diffusion process signals in correlated noise and to more general stochastic signals.

1. INTRODUCTION

To solve many problems in statistical detection theory a likelihood function (Radon-Nikodym derivative) is calculated and evaluated against a threshold to determine a decision. Probably the most well known method for evaluating the likelihood function for the case of a Gaussian signal in Gaussian noise is to solve an integral equation for a function which is to be the kernel of a quadratic form in the observations. Difficulties with this method are that integral equations are usually difficult to solve and the solutions obtained often require storage of all the observations. Some work for continuous time detection problems has been done that removes these difficulties. Schweppe (1965) considered the case of a Gauss-Markov signal in white noise and by first solving the problem in discrete time and then formally passing to the limit to obtain the result for the continuous time problem he obtained a recursive method for evaluating the likelihood function which used the linear filtering results of Kalman and Bucy (1961). Sosulin

† This work was supported by the Applied Mathematics Division of the Air Force Office of Scientific Research under Contract AF 49(638)1517 at Stanford University and under Contract AF-814-66 at the University of Michigan and by JSEP at Stanford University. It is based on work done in partial fulfillment of the requirements for the PhD degree in the Department of Electrical Engineering, Stanford University, Stanford, California.

and Stratonovich (1965) considered the detection problem of a diffusion process in white noise and using the Stratonovich (1966) definition of stochastic integral formally obtained a differential equation for the likelihood function and related the terms of this differential equation to some nonlinear filtering results.

In this paper we provide a rigorous derivation of a stochastic differential equation for the evolution of the likelihood function for a diffusion process in white noise using the K. Itô (1951a) definition of stochastic integral. Some differences which are noted between our results and the results of Schweppe and the results of Sosulin and Stratonovich are related to the transformation calculus used and to the different definitions of stochastic integral. We also consider the case of a stochastic signal (diffusion process) in correlated noise (diffusion process) and obtain in some cases necessary and sufficient conditions for nonsingular detection. We relate the nonsingular problem to a white noise detection problem. We also indicate how the general diffusion process signal in diffusion process noise detection problem can be solved and the extension of our solution to more general stochastic signals.

2. PROBLEM STATEMENT

The detection problem will be described in terms of stochastic differential equations rather than white noise, because white noise does not exist as an ordinary random process, and in general when a random function is integrated with white noise, the integral can be defined in different ways (Stratonovich 1966). For a discussion of stochastic differential equations the reader is referred to K. Itô (1961) or Doob (1953).

It will be convenient to initially make some assumptions that will be continually used throughout this paper. Consider a stochastic differential equation

$$dx_t = a(t, x_t) dt + b(t, x_t) dB_t \quad (1)$$

where x_t , B_t and $a(t, x_t)$ are n vectors and $b(t, x_t)$ is an $n \times n$ matrix. We shall always assume the following hypothesis, H :

H : The process $\{B_t\} = \{(B_t^1, B_t^2, \dots, B_t^n)^T\}$ will be a vector of n independent standard Brownian motions. This will be called n dimensional Brownian motion. The components of the vector $a(t, x)$ and the matrix $b(t, x)$ are continuous in t and globally

Lipschitz continuous in x . The domain of the solution of the differential equation will be the bounded interval $[s, 1]$.

We note that given an initial condition independent of the future Brownian motion and assuming H the solution of (1) exists and is unique (K. Itô 1961).

The initial detection problem that we shall consider is described by the following stochastic differential equations

$$\begin{aligned} dy_t &= f(t)x_t dt + h(t) d\tilde{B}_t && \text{for signal present} \\ &= h(t) d\tilde{B}_t && \text{for signal not present} \end{aligned} \quad (2)$$

where

$$\begin{aligned} dx_t &= a(t, x_t) dt + b(t, x_t) dB_t \\ x(s) &= \alpha, y(s) = 0 \end{aligned} \quad (3)$$

We also assume, in addition to H , that $h^{-1}(t)$ exists and is continuous.

The n dimensional process $\{x_t\}$ corresponds to the (stochastic) signal which is obtained by (3) and $\{\tilde{B}_t\}$ is an m dimensional Brownian motion independent of the n dimensional Brownian motion $\{B_t\}$. The process $\{y_t\}$ is an m dimensional process of our observations.

3. SOME PRELIMINARY RESULTS

Before establishing our main result we must note a few preliminaries. The first theorem gives some sufficient conditions for the absolute continuity of the measures corresponding to solutions of stochastic differential equations. The result stated is due to Girsanov (1960) while similar, less general results have been given by Prokhorov (1956) and Skorokhod (1960).

THEOREM 1. *Suppose that*

$$dx_t = a(t, x_t) dt + b(t, x_t) dB_t \quad (4)$$

$$dy_t = (a(t, y_t) + b(t, y_t)h(t, y_t)) dt + b(t, y_t) dB_t \quad (5)$$

where

- i) $t \in [s, 1] x(s) = y(s)$
- ii) a and h are n vectors and b is an $n \times n$ matrix
- iii) $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ and $h(\cdot, \cdot)$ are measurable in both variables, in particular a and b are continuous in their first variable and globally Lipschitz continuous in their second variable

$$\text{iv) } \int_s^1 |h(t, x_t)|^2 dt < \infty \quad \text{a.s.}$$

$$\text{v) } |h(t, x_t)| < h_0(|x_t|)$$

where h_0 is a nondecreasing function of a real variable.

Then the measures μ_X and μ_Y induced on $C_n[s, 1]$ (the space of all continuous functions with values in \mathbb{R}^n) by $\{x_t\}$ and $\{y_t\}$ respectively are mutually absolutely continuous.

The Radon-Nikodym derivative, $d\mu_Y/d\mu_X$, is given by

$$\frac{d\mu_Y}{d\mu_X} = \exp \left[\int_s^1 h^T(u, x_u) dB_u - \frac{1}{2} \int_s^1 |h(u, x_u)|^2 du \right] \quad (6)$$

If b^{-1} exists, this can be rewritten entirely in terms of $\{x_t\}$ as

$$\begin{aligned} \frac{d\mu_Y}{d\mu_X} = \exp \left[\int_s^1 h^T(u, x_u) b^{-1}(u, x_u) dx_u \right. \\ \left. - \int_s^1 h^T(u, x_u) b^{-1}(u, x_u) a(u, x_u) du - \frac{1}{2} \int_s^1 |h(u, x_u)|^2 du \right] \end{aligned} \quad (7)$$

We next describe a result from nonlinear filtering theory. Given the following nonlinear filtering problem

$$dx_t = a(t, x_t) dt + b(t, x_t) dB_t \quad (8)$$

$$dy_t = f(t)x_t dt + h(t) d\tilde{B}_t \quad (9)$$

where (8) describes our state and (9) describes our observations (This is our detection problem with signal present with the assumptions as made there).

THEOREM 2. *The conditional mean for the above filtering problem is given by the following expression*

$$E[x_t | y_u ; s \leq u \leq t] = \hat{x}_t = \frac{E_{\mu_X} [\Psi_t x_t]}{E_{\mu_X} [\Psi_t]} \quad (10)$$

where E_{μ_X} corresponds to integration with respect to the function space measure μ_X generated by the solution of (8) and

$$\begin{aligned} \Psi_t = \Psi_t(x_u, y_u ; s \leq u \leq t) \\ = \exp \left[\int_s^t x_u^T f_u^T g_u^{-1} dy_u - \frac{1}{2} \int_s^t x_u^T f_u^T g_u^{-1} f_u x_u du \right] \end{aligned} \quad (11)$$

and $g = h^T h$.

The proof of this result will not be included here but it can be easily proved from our absolute continuity results and the fact that the process $\{x_t\}$ has a finite second moment. A stochastic differential equation can also be obtained for its evolution in time. These results are discussed by Kushner (1967) and by Duncan (1968).

4. MAIN RESULT

We have now established sufficient preliminaries so we shall return to the detection problem. We first derive an expression for the likelihood function (Radon-Nikodym derivative).

LEMMA 1. *The likelihood function, Λ_t , for the detection problem (2) is given by*

$$\Lambda_t = E_{\mu_X}[\Psi_t] \quad (12)$$

where Ψ_t and E_{μ_X} are defined in Theorem 2.

Proof. The likelihood function is the Radon-Nikodym derivative of the measures, say ρ_1 and ρ_0 , corresponding to the two hypotheses, signal present and signal not present. Fix t and let $\Gamma \in \mathfrak{B}(y_u; s \leq u \leq t)$ (the augmented Borel field generated by $\{y_u; s \leq u \leq t\}$). Since the likelihood function is a martingale of Brownian motion we have

$$\begin{aligned} \rho_1(\Gamma) &= \iint_{\Gamma} \Psi_1 d(\mu_X \times \rho_0) \\ &= \iint_{\Gamma} E[\Psi_1 | \mathfrak{B}(x_u, y_u; s \leq u \leq t)] d(\mu_X \times \rho_0) \quad (13) \\ &= \iint_{\Gamma} \Psi_t d(\mu_X \times \rho_0) \end{aligned}$$

Since $\{x_u; s \leq u \leq t\}$ is a functional only of the Brownian motion $\{B_u; s \leq u \leq t\}$ (cf. the construction of $\{x_t\}$ K. Itô 1961) the measure generated by the Borel field $\mathfrak{B}(x_u - x_s, \bar{B}_u; s \leq u \leq t)$ is the product of the measures μ_X and ρ_0 . Therefore

$$\begin{aligned} \rho_1(\Gamma) &= \iint_{\Gamma} \Psi_t d\mu_X d\rho_0 \\ &= \int E_{\mu_X}[\Psi_t] d\rho_0 \end{aligned} \quad (14)$$

By definition of the Radon-Nikodym derivative we have then

$$\Lambda_t = E_{\mu_X}[\Psi_t] \quad \text{a.s. } \rho_0 \quad (15)$$

This was done for fixed t but it follows immediately for a countable dense set $\{t_i\}$ in $[s, 1]$. Since $x_t, y_t,$ and Ψ_t are continuous in t (a.s.) and the sequence $\langle \Psi_{t_i} \rangle$ is uniformly integrable we have for all t

$$\Lambda_t = E_{\mu_X}[\Psi_t] \quad \text{a.s.} \quad \rho_0 \tag{16}$$

To obtain a stochastic differential equation for the likelihood function we shall use the following result of K. Itô (1951b) for obtaining differentials of smooth functions of solutions of stochastic differential equations.

LEMMA 2. *If $G(t, x)$ has a continuous first partial derivative with respect to t and a continuous second partial derivative with respect to $x, -\infty < x < \infty, s \leq t \leq 1$ and if the functions $f(t, \omega)$ and $g(t, \omega)$ are independent of the future Brownian motion and $\int |f|^2 dt < \infty, \int |g| dt < \infty$ a.s. and*

$$dz(t, \omega) = g(t, \omega) dt + f(t, \omega) dB(t, \omega) \tag{17}$$

then a.s.

$$\begin{aligned} G(1, z_1) - G(s, z_s) &= \int_s^1 \frac{\partial G}{\partial z}(t, z_t) f(t, \omega) dB_t \\ &+ \int_s^1 \left\{ \frac{\partial G(t, z_t)}{\partial z} g(t, \omega) + \frac{\partial G(t, z_t)}{\partial t} + \frac{1}{2} f^2(t, \omega) \frac{\partial^2 G}{\partial z^2}(t, z_t) \right\} dt \end{aligned} \tag{18}$$

We now establish a simple result for interchanging an expectation and a stochastic integral.

LEMMA 3. *Consider the likelihood function Λ_t . The following equality is valid a.s. ρ_0*

$$\begin{aligned} \Lambda_t &= E_{\mu_X}[\Psi_t] = E_{\mu_X} \left[1 + \int_s^t x_u^T f_u^T g_u^{-1} \Psi_u dy_u \right] \\ &= 1 + \int_s^t E_{\mu_X} [x_u^T f_u^T g_u^{-1} \Psi_u] dy_u \end{aligned} \tag{19}$$

Proof. We obtain the first line by applying the stochastic differential rule (Lemma 2) to

$$\Psi = e^{z_t}$$

where

$$dz_t = x_t^T f_t^T g_t^{-1} dy_t - \frac{1}{2} x_t^T f_t^T g_t^{-1} f_t x_t dt$$

and

$$g = h^T h.$$

We can define the stochastic integral

$$\int_s^t x_u^T f_u^T g_u^{-1} \Psi_u dy_u$$

as a limit in $L^1(dP)$ of finite sum approximations from the martingale property of this stochastic integral, i.e.,

$$\lim_n \sum_{i=1}^n x_{t_i}^T f_{t_i}^T g_{t_i}^{-1} \Psi_{t_i} (y_{t_{i+1}} - y_{t_i}) = \int_s^t x_u^T f_u^T g_u^{-1} \Psi_u dy_u$$

as the partitions become dense in $[s, 1]$. Since $\{y_u; s \leq u \leq t\}$ is the variable of integration for ρ_0 it is independent of $\{x_u; s \leq u \leq t\}$. Therefore for the finite sum approximations we can interchange the expectation, E_{μ_X} , and the finite sum. Therefore we have

$$E_{\mu_X} \int_s^t x_u^T f_u^T g_u^{-1} dy_u = \int_s^t E_{\mu_X} x_u^T f_u^T g_u^{-1} dy_u \quad \text{a.s. } \rho_0 \quad (20)$$

To establish this result for all $t \in [s, 1]$ use the continuity of the stochastic integral. ■

We are now prepared to establish our main result.

THEOREM 3. *Consider the detection problem (2). The process $\{\Gamma_t\}$ defined as*

$$\Gamma_t = \ln \Lambda_t = \ln E_{\mu_X}[\Psi_t] \quad (21)$$

satisfies the following stochastic differential equation

$$d\Gamma_t = \hat{x}_t^T f_t^T g_t^{-1} dy_t - \frac{1}{2} \hat{x}_t^T f_t^T g_t^{-1} f_t \hat{x}_t dt \quad (22)$$

where

$$\hat{x}_t = \frac{E_{\mu_X}[\Psi_t x_t]}{E_{\mu_X}[\Psi_t]} \quad (23)$$

Proof. (Note that Λ_t is strictly positive and finite a.s. so $\ln \Lambda_t$ is well defined and finite a.s.) Applying the stochastic differential rule to $\ln \Lambda_t$ we obtain

$$d\Gamma_t = \frac{E_{\mu_X} x_t^T f_t^T g_t^{-1} \Psi_t dy_t}{E_{\mu_X}[\Psi_t]} - \frac{1}{2} \frac{E_{\mu_X} [x_t^T f_t^T \Psi_t] g_t^{-1} E_{\mu_X} [f_t x_t \Psi_t] dt}{\{E_{\mu_X}[\Psi_t]\}^2}$$

Thus

$$d\Gamma_t = \hat{x}_t^T f_t^T g_t^{-1} dy_t - \frac{1}{2} \hat{x}_t^T f_t^T g_t^{-1} f_t \hat{x}_t dt. \quad \blacksquare$$

Remark. We shall now briefly mention some comparisons with results that were mentioned in the introduction. Our results differ in form from both the results of Schweppe and of Sosulin and Stratonovich. Sosulin and Stratonovich (1965) indicate that their stochastic integral is to be interpreted in the Stratonovich sense and it is known that in general this integral is not identical to the K. Itô stochastic integral. Schweppe (1965) obtained his result from a formal passage to the limit from discrete time using ordinary calculus. If we do some simple manipulations and use the correction term (Stratonovich 1966) that relates Stratonovich stochastic integrals to K. Itô stochastic integrals we can reconcile our result with Schweppe's result and verify that his result must also be interpreted in the Stratonovich sense.

5. SOME GENERALIZATIONS

We shall now discuss some detection problems of stochastic signals (diffusion processes) in correlated noise (diffusion processes). Consider a detection problem described by the following stochastic equations

$$\begin{aligned} y_t &= H_t x_t + z_t && \text{for signal present} \\ &= z_t && \text{for signal not present} \end{aligned} \quad (24)$$

where

$$dx_t = a(t, x_t) dt + b(t, x_t) dB_t \quad (25)$$

$$dz_t = g(t, z_t) dt + h(t) d\tilde{B}_t \quad (26)$$

$$x(s) = \alpha z(s) = 0$$

We assume H and that the inverse of the diffusion matrix $h(t)$, $h^{-1}(t)$, exists and is continuous and that the derivative of the time-varying matrix H_t , H_t' , exists and is continuous in t .

For this detection problem we shall give necessary and sufficient conditions for nonsingular detection and relate the nonsingular problem to a white noise problem.

We first note a result for the quadratic (second order) variation of the solution of a stochastic differential equation. This result is obtained by K. Itô (1951b), Wang (1964) and Wong and Zakai (1965). Some related results have been discussed by Baxter (1956), Gladyshev (1961), Bühlman (1963), Cogburn and Tucker (1961) and Pierre (1967).

LEMMA 4. *Let $\{x_t\}$ be the process which satisfies*

$$dx_t = \phi(t, x_t) dt + \Gamma(t, x_t) dB_t \quad (27)$$

where $t \in [s, 1]$, $x(s) = \alpha$, $\{x_t\} = (x_t^1, x_t^2, \dots, x_t^n)^T$ is an n dimensional diffusion process and $\{B_t\}$ is an n dimensional Brownian motion. The n -vector $\phi(t, x)$ and the $n \times n$ matrix $\Gamma(t, x)$ have components which are continuous in t and globally Lipschitz continuous in x . Then

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{k(n)-1} (x^j(t_{i+1}^{(n)}) - x^j(t_i^{(n)}))^2 = \sum_{r=1}^n \int_s^t \gamma_{jr}^2(t, x_t) dt \quad (28)$$

in the mean square and almost surely where the partition $\{t_i^{(n)}\}$ is a refinement of $\{t_i^{(n-1)}\}$ for all n and these partitions become dense in $[s, 1]$, ($\Gamma(t, x_t) = \{\gamma_{ij}(t, x_t)\}$).

LEMMA 5. Consider the detection problem (24). For this detection problem to be nonsingular it is necessary and sufficient that for all t

$$H(t)b(t, x_t) \equiv 0 \quad \text{a.s.} \quad (29)$$

Proof (sufficiency). It will be convenient to change the form of the above detection problem by describing the hypotheses by stochastic differential equations (i.e., apply the stochastic differential rule to the two hypotheses).

$$\begin{aligned} dy_t &= H'(t)x_t dt + H(t)a(t, x_t) dt + H(t)b(t, x_t) dB_t + g(t, y_t) \\ &\quad - H(t)x_t dt + h(t) d\tilde{B}_t \quad \text{for signal present} \\ &= g(t, y_t) dt + h(t) d\tilde{B}_t \quad \text{for signal not present} \end{aligned} \quad (30)$$

Now let $H(t)b(t, x_t) \equiv 0$ a.s. Then the two hypotheses are a.s.

$$\begin{aligned} dy_t &= H'(t)x_t dt + H(t)a(t, x_t) dt + g(t, y_t) - H(t)x_t dt \\ &\quad + h(t) d\tilde{B}_t \quad \text{for signal present} \\ &= g(t, y_t) dt + h(t) d\tilde{B}_t \quad \text{for signal not present} \end{aligned} \quad (31)$$

Since the process $\{x_t\}$ is generated by the n dimensional Brownian motion $\{B_t\}$ it is independent of the process $\{y_t\}$ satisfying

$$\begin{aligned} dy_t &= g(t, y_t) dt + h(t) d\tilde{B}_t \\ y(s) &= 0 \end{aligned} \quad (32)$$

With this independence and the fact that h^{-1} exists for all t we can apply Girsanov's theorem as we did in the white noise case (Lemma 1) and obtain for the likelihood function

$$\Lambda_t = E_{\mu_X}[\tilde{\Psi}_t] \quad (33)$$

where $\tilde{\Psi}_t$ is obtained from Girsanov's theorem (Theorem 1).

Proof (necessity). If for all t , $H(t)b(t, x_t) \neq 0$ a.s. then for some t , say t^* ,

$$H(t^*)b(t^*, x_{t^*}) \neq 0 \quad \text{for } x_{t^*} \in \Lambda, \quad P(\Lambda) > 0$$

From the continuity conditions there exists an interval $[t_0(\omega), t_1(\omega)]$ such that for $t \in [t_0(\omega), t_1(\omega)]$

$$H(t)b(t, x_t) \neq 0 \quad \text{for } x_t \in \Lambda$$

Therefore applying our quadratic test statistic (Lemma 4) we are able on the set Λ to distinguish with zero error between the two hypotheses. If signal is present we have

$$\sum_j \int_s^1 \bar{b}_{ij}^2(s, x_s) ds + \int_s^1 h_{ij}^2(s) ds \quad i = 1, 2, \dots, n$$

where $\bar{b}(t, x_t) = \{\bar{b}_{ij}(t, x_t)\} = H(t)b(t, x_t)$ and for signal not present we have

$$\sum_j \int_s^1 h_{ij}^2(s) ds \quad i = 1, 2, \dots, n \quad \blacksquare$$

We note that in the nonsingular case we have the two hypotheses

$$dy_t = H'(t)x_t dt + H(t)a(t, x_t) dt + g(t, y_t - H(t)x_t) dt + h(t) d\tilde{B}_t$$

for signal present

$$= g(t, y_t) dt + h(t) d\tilde{B}_t \quad \text{for signal not present}$$

For the hypothesis for signal not present we can use the absolute continuity results (Theorem 1.) to show that the measure corresponding to the solution of (33) is absolutely continuous with respect to the measure corresponding to the solution of

$$\begin{aligned} dy_t &= h(t) d\tilde{B}_t \\ y_s &= 0 \end{aligned} \quad (34)$$

So we can consider this correlated noise problem as a detection problem with respect to white noise.

Remark 1. For the necessary and sufficient conditions for nonsingular detection we have to determine if $H(t)b(t, x_t) \equiv 0$ a.s. Without additional assumptions on the coefficients of the stochastic differential equations this is not equivalent to $H(t)b(t, x) = 0 \quad x \in \mathcal{R}^n \quad t \in T \equiv [s, 1]$, though this latter condition is clearly sufficient. The detection problem may still be nonsingular even if $H(t)b(t, x) \neq 0$ for some $x \in \mathcal{R}^n \quad t \in T$.

This can occur if the process $\{x_t\}$ does not take values (a.s.) in the region of $\mathcal{R}^n \otimes T$ where $H(t)b(t, x) \neq 0$. Trivial examples can be constructed, for example, when $\{x_t\}$ is constant (a.s.). Some sufficient conditions do exist to prove that $\{x_t\}$ "fills" the state space \mathcal{R}^n but these conditions will not be described here.

Remark 2. If we consider a more general correlated noise problem

$$\begin{aligned} y_t &= H(t, x_t) + G(t, z_t) && \text{for signal present} \\ &= G(t, z_t) && \text{for signal not present} \end{aligned} \quad (5)$$

where x_t and z_t are given by (25) and (26) respectively. Then it should be "clear" how to proceed. If sufficient smoothness is assumed on ζ and H then we proceed by taking derivatives or differentials (whichever is appropriate) using Lemma 2 in the vector case (K. Itô 1951b) until we obtain some Brownian motion. Depending upon whether it is $\{\tilde{y}_t\}$ or $\{B_t\}$ will determine whether our problem is nonsingular or singular. For the Gaussian (linear) detection problem other techniques are available to determine singular or nonsingular detection (cf. e.g., Root 1963).

Remark 3. The expression for the likelihood function (e^{r_t}) also arises in the nonlinear filtering problem and can be used to simplify especially the nonlinear smoothing problem.

Remark 4. The analogy with the detection of a sure signal should be noted, i.e., the expression for the likelihood function (e^{r_t}) uses, instead of the known signal for the sure signal case, the best estimate (in a mean square sense) of the signal for the stochastic signal case.

Remark 5. After submission of this paper, extensions of the likelihood function form (Theorem 3) for more general stochastic signals have been obtained independently by T. Kailath and by the author. In a paper to appear, Kailath uses a different technique to approach the detection problem by reformulating the stochastic differential equation for signal present. Absolute continuity results for these more general signals can be obtained by using some entropy results (cf. A. Perez, Notions generalisees d'incertitude, d'entropie et d'information du point de vue de la theorie de martingales, in Transactions of the First Prague Conference on Information Theory, Statistical Decision Functions, Random Processes," pp 183-208. Publishing House, Czech. Acad. Sci., Prague (1957). With uniqueness of the stochastic differential equation with signal present and the appropriate absolute continuity, the result in Theorem 3 easily generalizes. Kailath in his forthcoming paper also has some discussion on the stochastic integral of K. Itô as contrasted with some other proposed definitions.

ACKNOWLEDGEMENT

The author wishes to thank Professor T. Kailath who suggested this problem and who also made some helpful suggestions.

RECEIVED: April 9, 1968

REFERENCES

- BAXTER, G. (1956), A strong limit theorem for Gaussian processes. *Proc. Amer. Math. Soc.* **7**, 522-527.
- BÜHLMAN, H. (1963), L^2 -Martingales and orthogonal decomposition. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **1**, 394-414.
- COGBURN, R., AND TUCKER, H. A. (1961), A limit theorem for a function of the increments of a decomposable process. *Trans. Amer. Math. Soc.* **99**, 278-284.
- DOOB, J. L. (1953), "Stochastic Processes." John Wiley and Sons, Inc., New York.
- DUNCAN, T. E., On the nonlinear filtering problem. To appear.
- GLADYSHEV, E. A. (1961), A new limit theorem for random processes with gaussian increments. *Theor. Probability Appl.* **6**, 52-61.
- GIBSANOV, I. V. (1960), On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theor. Probability Appl.* **V**, 285-301.
- ITÔ, K. (1951a), On stochastic differential equations. *Mem. Amer. Math. Soc.* **4**.
- ITÔ, K. (1951b), On a formula concerning stochastic differentials. *Nagoya Math. J.* **3**, 55-65.
- ITÔ, K. (1961), Lectures on stochastic processes, Tata Institute for Fundamental Research, Bombay, India.
- KALMAN, R. E. AND BUCY, R. S. (1961), "New results in linear filtering theory," *J. Basic Engr. (ASME Trans)* **83D**, pp. 95-108.
- KUSHNER, H. J. (1967), Dynamical equations for optimum nonlinear filtering. *J. Differential Equations* **3**, 171-190.
- PIERRE, P. A. (1967), Representation and analysis of signals part XXV. properties of non-gaussian, continuous parameter, random processes as used in detection theory. Department of Electrical Engineering, The Johns Hopkins University.
- PROHOROV, YU. V. (1956), Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1**, 157-214, especially 207-212.
- ROOT, W. L. (1963). Singular gaussian measures in detection theory in "Proc. Sym. on Time Series Analysis," 293-315. John Wiley and Sons, Inc., New York.
- SCHWEPPE, FRED C. (1965), Evaluation of likelihood functions for gaussian signals. *IEEE Trans. Information Theory* **IT-11**, 61-70.
- SKOROKHOD, A. V. (1960), On the differentiability of measures corresponding to random processes. II. Markov Processes. *Theor. Probability Appl.* **V**, 40-49.
- SOSULIN, YU. AND STRATONOVICH, R. L. (1965), Optimum detection of a diffusion process in white noise. *Radiotekhn, i Èlekttron.* **10**, 827-838 [English transl.: *Radio Engrg. Electron. Phys.* **10**, 704-714].

- STRATONOVICH, R. L. (1966), A new representation for stochastic integrals. Vestnik Moskov Univ., Ser. I, *Mat. Meh.*, **1**, pp. 3-12 [English transl.: *SIAM J. Control*, **4**, 362-371].
- WANG, J. (1964), Properties and limit theorems of sample functions of the Itô stochastic integral. *Chinese Math.* **5**, 556-570.
- WONG, E., AND ZAKAI, M. (1965), The oscillation of stochastic integrals. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **4**, 103-112.