# FILTERING OF PROCESSES WHICH HAVE CONTINUOUS NOISE

by

Jan Van Schuppen and Pravin Varaiya

Memorandum No. ERL-M421

2 January 1974

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

FILTERING OF PROCESSES WHICH HAVE CONTINUOUS NOISE*

Jan Van Schuppen and Pravin Varaiya

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory,
University of California, Berkeley, California 94720

Abstract. This paper presents the structure of optimal nonlinear filters for the case where the unknown signal process and the observation process have additive noise which is a continuous martingale. Thus it generalizes recent work where the additive noise is Brownian motion. The results depend upon a representation theorem which states that all martingales of the observation process are stochastic integrals of the "innovations."

*Research supported in part by National Science Foundation under Grant GK-10656X3 and in part by U.S. Army Research Office-Durham, under Contract DAHC04-67-C-0046.

1. Introduction and Summary. This paper extends previous results [1,5-7,9-14] based on the theory of martingales for filtering and estimation where the "noise" process is an additive Brownian motion, to the case where it is an arbitrary additive martingale with continuous sample paths.

The first crucial result (Theorem 2.2) states that if $m_t$ is a n-dimentional, continuous martingale, then every martingale $z_t$, which is adapted to the σ-fields $\mathcal{F}_t^m$ generated by $m_t$, can be expressed as a stochastic integral

$$z_t = z_0 + \int_0^t \phi_s dm_s$$

This is a generalization of the corresponding celebrated result for the case where $m_t$ is Brownian motion [16], and is in fact an easy consequence of the latter result and an important theorem due to Knight [15].

Next, in Section 3, we show that if the probability $\mathcal{P}$ on the measurable space on which $m_t$ is defined is replaced by another measure $\mathcal{P}_1$, mutually absolutely continuous with respect to $\mathcal{P}$, then $m_t$ can be expressed as

$$(1.1) \qquad m_t = \int_0^t \phi_s d \langle m \rangle_s + n_t$$

where $n_t$ is a martingale under $\mathcal{P}_1$. Furthermore, the quadratic variations of $m_t$ and $n_t$ are equal, $\langle m \rangle_t \equiv \langle n \rangle_t$, and all processes, adapted to $\mathcal{F}_t^m$, which are martingales under $\mathcal{P}_1$ are stochastic integrals of $n_t$ (Theorem 3.1). More importantly, we show that this representation as stochastic integrals of $n_t$ continues to hold when $m_t$ is given by (1.1) whether or not $\mathcal{P}_1$ is absolutely continuous with respect to $\mathcal{P}$ (Theorem 3.2). Both these results use techniques first developed in [7] as refined later in [19,20].

At this point we are ready to use these concepts to formulate models for estimating an 'unknown' process given an 'observed' process of the form (1.1), and to use these results to obtain the stochastic differential equation satisfied by the optimum least squares estimates (Theorem 4.1). This portion of the exercise is worked out in Section 4.

There are various miscellaneous results of minor nature in the paper. Thus, for example, Section 3 contains a formula for the conditional likelihood ratio which extends the one derived in [5,10,11].

2. Martingale representation theorem. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and let $(\mathcal{F}_t)$, $t \in R_+$, be an increasing, right-continuous family of sub-$\sigma$-fields of $\mathcal{F}$. Let $\mathcal{F}_\infty = \bigvee_t \mathcal{F}_t$. Every family $(x_t)$, $t \in R_+$, of real-valued functions on $\Omega$ such that $x_t$ is $\mathcal{F}_t$-measurable, defines a <u>stochastic process</u> $(x_t, \mathcal{F}_t, \mathcal{P})$. The same family $(x_t)$ defines a <u>different</u> process if either the family $(\mathcal{F}_t)$ or the probability $\mathcal{P}$ changes. In particular, if $(x_t, \mathcal{F}_t, \mathcal{P})$ is a process, so is $(x_t, \mathcal{F}^x_t, \mathcal{P})$ where $\mathcal{F}^x_t = \sigma\{x_s | 0 \leq s \leq t\}$. When the context makes it clear we write $(x_t, \mathcal{F}_t, \mathcal{P})$ as $(x_t, \mathcal{F}_t)$ or $(x_t, \mathcal{P})$ or $(x_t)$. Finally it will always be assumed that $\sigma$-fields are complete with respect to the associated probability measures.

For a discussion of any unfamiliar terms or concepts used below please refer to [2] or [17].

Let $(\Omega, \mathcal{F}_t, \mathcal{P})$, $t \in R_+$ be a fixed family.

DEFINITION 2.1. $\mathcal{M}^1(\mathcal{F}_t,P)$ is the set of all processes $(m_t,\mathcal{F}_t,P)$ which are uniformly integrable martingales. $\mathcal{M}^2(\mathcal{F}_t,P) = \{(m_t) \in \mathcal{M}^1(\mathcal{F}_t,P) \mid \sup_t Em_t^2 < \infty\}$. For $i = 1,2$ $\mathcal{M}^i_{loc}(\mathcal{F}_t,P)$ is the set of all processes $(m_t)$ for which there exists a sequence of stopping times $S_k$ such that $S_k \to \infty$ a.s. and $(m_{t\wedge S_k} I_{\{S_k>0\}}) \in \mathcal{M}^i$ for all k.

Without loss of generality, we will assume that every $(m_t) \in \mathcal{M}^1_{loc}$ has sample paths which are right-continuous and have left-hand limits. It is evident that $\mathcal{M}^2 \subset \mathcal{M}^1$ and hence $\mathcal{M}^2_{loc} \subset \mathcal{M}^1_{loc}$. Also, if $(m_t) \in \mathcal{M}^1_{loc}$ has continuous sample paths, then $(m_t) \in \mathcal{M}^2_{loc}$.

DEFINITION 2.2. $\mathcal{A}(\mathcal{F}_t,P)$ is the set of all right-continuous processes $(a_t,\mathcal{F}_t,P)$ with $a_0 = 0$, and such that they have integrable variation i.e., $E\int_0^\infty |da_t| < \infty$. $\mathcal{A}_{loc}(\mathcal{F}_t,P)$ is the set of all processes $(a_t)$ for which there exists a sequence of stopping times $S_k$ that $S_k \to \infty$ a.s. and $(a_{t\wedge S_k}) \in \mathcal{A}(\mathcal{F}_t,P)$.

Now if $(m_t) \in \mathcal{M}^2_{loc}(\mathcal{F}_t,P)$, then there exists a unique $\mathcal{F}_t$-predictable process $(\langle m\rangle_t) \in \mathcal{A}_{loc}(\mathcal{F}_t,P)$ such that $(m_t^2 - \langle m\rangle_t) \in \mathcal{M}^1_{loc}(\mathcal{F}_t,P)$. $(\langle m\rangle_t)$ is called the (predictable) <u>quadratic variation</u> of $(m_t)$. It is important to note that unless $(m_t)$ is continuous, $(\langle m\rangle_t)$ depends crucially upon the family $(\mathcal{F}_t,P)$. These statements are proved in [16]. For $(m_t)$ and $(n_t)$ in $\mathcal{M}^2_{loc}(\mathcal{F}_t,P)$, let $\langle m,n\rangle_t = \frac{1}{2}(\langle m+n\rangle_t - \langle m\rangle_t - \langle n\rangle_t)$. If $\langle m,n\rangle_t \equiv 0$ a.s. we say that $(m_t)$, $(n_t)$ are <u>orthogonal</u>; it then follows that $(m_t n_t) \in \mathcal{M}^1_{loc}$.

Our representation theorem is a consequence of the following result.

THEOREM 2.1. (Knight [15]) Let $(m_t^k) \in \mathcal{M}^2(\mathcal{F}_t,P)$, $k = 1,\ldots,n$, be a collection of continuous, pairwise orthogonal martingales. Let $(B_t,\mathcal{F}_t,P)$ be a n-dimensional Brownian motion which is independent of

$(m_t^1), \ldots, (m_t^n)$. Let

$$T_t^k(\omega) = \inf \{s \mid \langle m^k \rangle_s(\omega) > t\}, \text{ if this is finite}$$
$$= \infty, \text{ otherwise}$$

Set

$$x_t^k(\omega) = m^k_{T_t^k(\omega)}(\omega), \text{ if } T_t^k(\omega) < \infty$$

(2.1) $$= m_\infty^k(\omega) + B^k_{t - \langle m^k \rangle_\infty(\omega)}(\omega), \text{ otherwise,}$$

and let $\mathcal{F}_t^x$ be the sub-σ-field of $\mathcal{F}$ generated by $x_s^k$, $0 \leq s \leq t$, $1 \leq k \leq n$.

Then $(x_t, \mathcal{F}_t^x, P)$ is an n-dimensional Brownian motion.

Remark 2.1. The sample paths of $(x_s^k)$ are obtained from those of $(m_t^k)$ by a random time change up to the time $s = \langle m^k \rangle_\infty$, and after this time an independent Brownian motion $B^k$ is attached to $x^k$. Note that $\langle m^k \rangle_{T_t^k(\omega)}(\omega) = t$ if $\langle m^k \rangle_\infty(\omega) > t$. Also, if $t_1 < t_2$, then on the set $\{\omega \mid \langle m^k \rangle_{t_1}(\omega) = \langle m^k \rangle_{t_2}(\omega)\}$ we have $m_t^k(\omega) = m_{t_1}^k(\omega)$ a.s. for $t_1 \leq t \leq t_2$. Different components $x^k$ of x are obtained by <u>different</u> time changes, and the importance of the theorem lies precisely in permitting this, since the result has been long known for the case when $\langle m^1 \rangle_t \equiv \cdots \equiv \langle m^n \rangle_t$. It should be clear that there is no general inclusion relation between $\mathcal{F}_t^m$, the σ-field generated by $m_s^k$, $0 \leq s \leq t$, $1 \leq k \leq n$, and $\mathcal{F}_t^x$. Of course $\mathcal{F}_\infty^m \subset \mathcal{F}_\infty^x$. Meyer [2] has given a simple proof of this theorem when $\langle m^k \rangle_\infty = \infty$ a.s. for all k.

DEFINITION 2.3. For $a_t \in \mathcal{A}_{loc}(\mathcal{F}_t, \mathcal{P})$, let $L^2(a_t, \mathcal{F}_t, \mathcal{P})$ ($L^2_{loc}(a_t, \mathcal{F}_t, \mathcal{P})$) denote the set of all $\mathcal{F}_t$-predictable processes $(\phi_t, \mathcal{F}_t, \mathcal{P})$ such that $E\int_0^\infty \phi_s^2 |da_s| < \infty$ ($\int_0^t \phi_s^2 |da_s| < \infty$ a.s. for all $t \in R_+$).

COROLLARY 2.1. Under the hypothesis of Theorem 2.1 every random variable $z \in L^2(\Omega, \mathcal{F}^m_\infty, \mathcal{P})$ has a representation

$$z = E(z|\mathcal{F}^m_0) + \sum_{k=1}^{n}\int_0^\infty \phi_t^k dm_t^k \tag{2.2}$$

for some processes $(\phi_t^k) \in L^2(\langle m^k \rangle_t, \mathcal{F}^m_t, \mathcal{P})$. (The integral in (2.2) is a stochastic integral, see [4]).

Proof. First of all $L^2(\Omega, \mathcal{F}^m_\infty, \mathcal{P}) \subset L^2(\Omega, \mathcal{F}^x_\infty, \mathcal{P})$. Next, since $(x_t, \mathcal{F}^x_t, \mathcal{P})$ is Brownian motion, therefore, by the well-known result for the representation of functionals of Brownian motion [16],

$$z = E(z|\mathcal{F}^x_0) + \sum_{k=1}^{n}\int_0^\infty \psi_t^k dx_t^k \tag{2.3}$$

for some processes $(\psi_t^k) \in L^2(\langle x^k \rangle_t, \mathcal{F}^x_t, \mathcal{P})$. Now, using (2.1), rewrite (2.3) as

$$z = E(z|\mathcal{F}^m_0) + \sum_{k=1}^{n}\int_0^{\langle m^k \rangle_\infty} \psi_t^k dx_t^k + \sum_{k=1}^{n}\int_{\langle m^k \rangle_\infty}^\infty \psi_t^k dB^k_{t-\langle m^k \rangle_\infty}.$$

Since $(B_t)$ is independent of $\mathcal{F}^m_\infty$, therefore the second integral above vanishes a.s.,

$$(2.4) \qquad z = E(z|\mathcal{F}_0^m) + \sum_{k=1}^{n} \int_0^{\langle m^k \rangle_\infty} \psi_t^k dx_t^k$$

Finally, set

$$\phi_t^k(\omega) = \psi^k_{\langle m^k \rangle_t(\omega)}(\omega), \; k = 1,\dots,n$$

It is easily checked (see [18]) that $\phi_t^k$ is $\mathcal{F}_t^m$-predictable, and

$$\int_0^\infty \phi_t^k dm_t^k = \int_0^{\langle m^k \rangle_\infty} \psi_t^k dx_t^k,$$

$$E\int_0^\infty (\phi_t^k)^2 d\langle m^k \rangle_t = E\int_0^{\langle m^k \rangle_\infty} (\psi_t^k)^2 d\langle x^k \rangle_t.$$

□

Notation. We use the notation $(\phi o m)_t$ to denote stochastic integral

$$(\phi o m)_t = \int_0^t \phi_s dm_s$$

Our representation theorem is an easy consequence of the Corollary above.

THEOREM 2.2. Let $(m_t^k) \in \mathcal{M}^2(\mathcal{F}_t^m, P)$, k=1,...,n be pairwise orthogonal martingales with continuous sample paths. Then

(i) $(z_t) \in \mathcal{M}^2(\mathcal{F}_t^m, P) \Leftrightarrow$ there exist $(\phi_t^k) \in L^2(\langle m^k \rangle_t, \mathcal{F}_t^m, P)$ such that

$$(2.5) \qquad z_t = z_0 + \sum_{k=1}^{n} (\phi^k o m^k)_t$$

(ii) $\mathcal{M}^2_{loc}(\mathcal{F}^m_t,\mathcal{P}) = \mathcal{M}^1_{loc}(\mathcal{F}^m_t,\mathcal{P})$ and every $(z_t) \in \mathcal{M}^1_{loc}(\mathcal{F}^m_t,\mathcal{P})$ has continuous sample paths.

(ii) $(z_t) \in \mathcal{M}^1_{loc}(\mathcal{F}^m_t,\mathcal{P}) \Leftrightarrow$ there exist $(\phi^k_t) \in L^2_{loc}(\langle m^k\rangle_t,\mathcal{F}^m_t,\mathcal{P})$ such that

$$z_t = z_0 + \sum_{k=1}^{n} (\phi^k o m^k)_t \tag{2.6}$$

Remark 2.2. (i) This result is certainly false if the sample path continuity assumption is dropped. For suppose that $N^1_t$ and $N^2_t$ are two independent Poisson processes. Then $m_t \equiv N^1_t - N^2_t$ is a discontinuous martingale and $x^i_t \equiv N^i_t - t$, i=1,2 both belong to $\mathcal{M}^2(\mathcal{F}^m_t)$. But neither of these can be represented as stochastic integrals of $(m_t)$. However the result continues to hold without the orthogonality restriction. To see this, suppose that the $(m^k_t)$ are not orthogonal. Then we can construct orthogonal martingales $(\tilde{m}^k_t) \in \mathcal{M}^2(\mathcal{F}^m_t,\mathcal{P})$ such that

$$\tilde{m}^1_t \equiv m^1_t,$$

$$\tilde{m}^k_t \equiv m^k_t - \sum_{i=1}^{k-1} (\phi^{k,i} o m^i)_t, \quad k > 1, \tag{2.7}$$

and such that the $(\tilde{m}^k_t)$ and the $(m^k_t)$ generate the same family of martingales (see [16, p. 223]). Using Theorem 2.2 we obtain the representation in terms of stochastic integrals of the $(\tilde{m}^k_t)$ and then substitute the relations (2.6) to obtain a representation in terms of the $(m^k_t)$.

(ii) It is easy to see that the processes $(\phi^k_t)$ in (2.5) or (2.6) satisfy

$$\langle z, m^k \rangle_t = \int_0^t \phi_s^k d\langle m \rangle_s,$$

hence

$$\left(\frac{d\langle z, m\rangle}{d\langle m\rangle}\right) = (\phi_t^k)$$

3. Translation of martingales by an absolutely continuous change of measure.

We need two results of a somewhat general nature, the first of these is an immediate consequence of [3, Theorem 1].

LEMMA 3.1 Let $(n_t) \in \mathcal{M}^2_{loc}(\mathcal{F}_t, P)$ have continuous sample paths. Then

(3.1) $$\Lambda_t = \exp(n_t - \langle n \rangle_t), t \in R_+$$

is the unique process which satisfies the integral equation

(3.2) $$\Lambda_t = 1 + \int_0^t \Lambda_{s-} n_s$$

Furthermore $\Lambda_t \geq 0$ a.s., $(\Lambda_t, \mathcal{F}_t, P)$ is a supermartingale, and

(3.3) $$E(\Lambda_t) \leq 1 \text{ for all } t$$

The next result is a special case of [20, Theorem 3.2]. For completeness, we present a proof in the Appendix.

LEMMA 3.2 Let $(n_t)$, $(z_t)$ be continuous local martingales in $\mathcal{M}^2_{loc}(\mathcal{F}_t, P)$ and let $\Lambda_t$ be given by (3.1). Suppose that

(3.4) $$E(\Lambda_\infty) = 1$$

and define the probability measure $\mathcal{P}_1$ on $(\Omega, \mathcal{F})$ by

$$\mathcal{P}_1(A) = \int_A \Lambda_\infty(\omega)\mathcal{P}(d\omega), \quad A \in \mathcal{F}$$

Then the process $(x_t) \in \mathcal{M}^2_{loc}(\mathcal{F}_t, \mathcal{P}_1)$, where $x_t$ is defined by

(3.5) $$x_t = z_t - \langle n,z \rangle_t$$

and, furthermore,

(3.6) $$\langle x \rangle_t = \langle z \rangle_t \text{ a.s.}$$

Remark 3.1 Lemma 3.2 is a generalization of [8, Theorem 1].

As a corollary of Lemmas 3.1, 3.2 and Theorem 2.2 we obtain our first interesting result. Let $(m^k_t) \in \mathcal{M}^2(\mathcal{F}_t, \mathcal{P})$, $k=1,\ldots,n$ be continuous, pairwise orthogonal martingales. Let $\mathcal{P}_1$ be another measure on $(\Omega, \mathcal{F})$ such that $\mathcal{P}_1 \sim \mathcal{P}$. Define the conditional likelihood ratio

(3.7) $$L_t = E\left(\frac{d\mathcal{P}_1}{d\mathcal{P}} \Big| \mathcal{F}^m_t\right), \quad t \in R_+$$

THEOREM 3.1 There exist processes $(\phi^k_t) \in L^2_{loc}(\langle m^k \rangle_t, \mathcal{F}^m_t, \mathcal{P})$ such that

(3.8) $$L_t = L_0 \exp \sum_{k=1}^n \left( \int_0^t \phi^k_s dm^k_s - \frac{1}{2}\int_0^t (\phi^k_s)^2 d\langle m^k \rangle_s \right)$$

Let

(3.9) $$x^k_t = m^k_t - m^k_0 - \int_0^t \phi^k_s d\langle m^k \rangle_s$$

Then

(3.10) $$(x_t^k) \in \mathcal{M}^2(\mathcal{F}_t^m, \mathcal{P}_1), \quad \langle x^k \rangle_t \equiv \langle m^k \rangle_t, \quad \langle x^k, x^j \rangle_t \equiv 0 \text{ for } k \neq j,$$

and we have the representations

$(y_t) \in \mathcal{M}^2(\mathcal{F}_t^m, \mathcal{P}_1) \Leftrightarrow$ there exist $(\psi_t^k) \in L^2(\langle x^k \rangle_t, \mathcal{F}_t^m, \mathcal{P}_1)$ such that

(3.11) $$y_t = y_0 + \sum_{k=1}^{n} (\psi^k o x^k)_t,$$

$(y_t) \in \mathcal{M}^2_{loc}(\mathcal{F}_t^m, \mathcal{P}_1) \Leftrightarrow$ there exist $(\psi_t^k) \in L^2_{loc}(\langle x^k \rangle_t, \mathcal{F}_t^m, \mathcal{P}_1)$ such that

(3.12) $$y_t = y_0 + \sum_{k=1}^{n} (\psi^k o x^k)_t$$

Proof. From (3.7) it is immediate that $(L_t) \in \mathcal{M}^1(\mathcal{F}_t^m, \mathcal{P})$. Also since $\mathcal{P}_1 \sim \mathcal{P}$, $L_t > 0$ a.s., and so

$$(\Lambda_t) = (\frac{L_t}{L_0}) \in \mathcal{M}^1(\mathcal{F}_t^m, \mathcal{P})$$

Hence, if we define $(n_t)$ by

$$n_t = \int_0^t \frac{1}{\Lambda_s} d\Lambda_s,$$

then $(n_t) \in \mathcal{M}^1_{loc}(\mathcal{F}_t^m, \mathcal{P})$ and $\Lambda_t$ satisfies (3.1). By Theorem 2.2 there exist $(\phi_t^k) \in L^2_{loc}(\langle m^k \rangle_t, \mathcal{F}_t^m, \mathcal{P})$ such that

(3.13) $$n_t = \sum_{k=1}^{n} (\phi^k o m^k)_t$$

Also $(n_t)$ has continuous sample paths. Hence by Lemma 3.1

$$\Lambda_t = L_t L_0^{-1} = \exp(n_t - \frac{1}{2} \langle n \rangle_t)$$

which upon substitution of (3.13) yields (3.8).

Next, in (3.5), if we identify $z_t$ with $m_t^k - m_0$, then (3.5) and (3.6) yield the first and second assertions of (3.10). To prove the orthogonality, note that, by (3.6)

$$\langle x^k + x^j \rangle_t = \langle m^k + m^j \rangle_t,$$

hence

$$\langle x^k, x^j \rangle_t = \frac{1}{2} ( \langle x^k + x^j \rangle_t - \langle x^k \rangle_t - \langle x^j \rangle_t) = \frac{1}{2} ( \langle m^k + m^j \rangle_t - \langle m^k \rangle_t - \langle m^j \rangle_t)$$

$$= \langle m^k, m^j \rangle_t = 0 \text{ a.s. for } k \neq j \text{ by hypothesis.}$$

Finally to obtain the representations we begin by the well-known fact that

$$(y_t) \in \mathcal{M}^1_{loc}(\mathcal{F}^m_t, \mathcal{P}_1) \Leftrightarrow (y_t L_t) \in \mathcal{M}^1_{loc}(\mathcal{F}^m_t, \mathcal{P})$$

By Theorem 2.2 there exist $(\eta_t^k) \in L^2_{loc}( \langle m^k \rangle_t, \mathcal{F}^m_t, \mathcal{P})$ such that

$$y_t L_t = y_0 L_0 + \sum_{k=1}^{n} (\eta^k \circ m^k)_t$$

or, dividing throughout by $L_0$,

$$y_t \Lambda_t = y_0 + \sum_{k=1}^{n} (L_0^{-1} \eta^k \circ m^k)_t \qquad (3.14)$$

From (3.2), (3.9), (3.13), and the differentiation formula [4], it is easy to show that

$$(3.15) \qquad d\Lambda_t^{-1} = - \sum_{k=1}^{n} \Lambda_t^{-1} \phi_t^k \, dx_t^k$$

which together write (3.14) and the differentiation formula gives

$$dy_t = (y_t\Lambda_t)d(\Lambda_t^{-1}) + \Lambda_t^{-1}d(y_t\Lambda_t) - \sum_{k=1}^{n} \Lambda_t^{-1}\phi_t^k\eta_t^k L_0^{-1} d\langle m^k\rangle_t$$

$$= \sum_{k=1}^{n} (\Lambda_t^{-1}\eta_t^k L_0^{-1} - y_t\phi_t^k)dx_t^k$$

$$= \sum_{k=1}^{n} (L_t^{-1}\eta_t^k - y_t\phi_t^k)dx_t^k$$

Hence

$$y_t = y_0 + \sum_{k=1}^{n} (\psi^k o x^k)_t$$

where $\psi_t^k = L_t^{-1}\eta_t^k - y_t\phi_t^k$. This proves (3.12). To prove (3.11) we merely need to observe that, because the $(x_t^k)$ are orthogonal, therefore

$$E(\sum_{k=1}^{n} (\psi^k o x^k)_t)^2 = \sum_{k=1}^{n} E \int_0^t (\psi_s^k)^2 d\langle x^k\rangle_s .$$

□

Remark 3.2. The significance of this result is twofold. In the first place it shows that when $\mathcal{P}$ is replaced by $\mathcal{P}_1$, a continuous martingale $(m_t) \in \mathcal{M}^2(\mathcal{F}_t^m, \mathcal{P})$ takes the form of a semimartingale

$$(3.16) \qquad m_t = m_0 + \int_0^t \phi_s d\langle m\rangle_s + x_t ,$$

where $(\phi_t) = \left(\frac{d\langle m,n\rangle}{d\langle m\rangle}\right)$, and where $(x_t) \in \mathcal{M}^2(\mathcal{F}^m_t, \mathcal{P}_1)$. Furthermore, the linear map $(m_t, \mathcal{F}^m_t, \mathcal{P}) \mapsto (x_t, \mathcal{F}^m_t, \mathcal{P}_1)$ is an isometry i.e., $(\langle x\rangle_t) = (\langle m\rangle_t)$. Secondly, all processes $(y_t)$ which are martingales on the family $(\Omega, \mathcal{F}^m_t, \mathcal{P}_1)$ are stochastic integrals of the new "basis" martingales $(x^k_t)$. It turns out that this result holds whenever the semimartingales $(m^k_t)$ have the form (3.9), as we see below.

Once again let $(m^k_t) \in \mathcal{M}^2(\mathcal{F}_t, \mathcal{P})$, k=1,...,n be continuous, pairwise orthogonal martingales. Let $(h^k_t) \in L^2(\langle m^k\rangle_t, \mathcal{F}_t, \mathcal{P})$ and define the semimartingales $(x^k_t)$ by

(3.17) $$x^k_t = x^k_0 + \int_0^t h^k_s d\langle m^k\rangle_s + m^k_t, \quad t \in R_+$$

Let $\mathcal{F}^x_t = \sigma\{x^k_s | 0 \leq s \leq t,\ 1 \leq k \leq n\}$.

Notation. For any process $(f_t, \mathcal{F}_t, \mathcal{P})$ such that $E|f_t| < \infty$ for all t, let

$$\hat{f}_t = E(f_t | \mathcal{F}^x_t)$$

THEOREM 3.2. There exist continuous martingales $(\mu^k_t) \in \mathcal{M}^2(\mathcal{F}^x_t, \mathcal{P})$ with

(3.18) $$\langle \mu^k\rangle_t \equiv \langle m^k\rangle_t, \quad \langle \mu^k, \mu^j\rangle_t \equiv 0 \quad k \neq j$$

such that

(3.19) $$x^k_t = x^k_0 + \int_0^t \hat{h}^k_s d\langle \mu^k\rangle_s + \mu^k_t$$

Furthermore, we have the representations

$(y_t) \in \mathcal{M}^2(\mathcal{F}^x_t, P) \Leftrightarrow$ there exist $(\phi^k_t) \in L^2(\langle \mu^k \rangle_t, \mathcal{F}^x_t, P)$ such that

$$y_t = y_0 + \sum_{k=1}^{n} (\phi^k \circ \mu^k)_t, \tag{3.20}$$

$(y_t) \in \mathcal{M}^2_{loc}(\mathcal{F}^x_t, P) \Leftrightarrow$ there exist $(\phi^k_t) \in L^2_{loc}(\langle \mu^k \rangle_t, \mathcal{F}^x_t, P)$ such that

$$y_t = y_0 + \sum_{k=1}^{n} (\phi^k \circ \mu^k)_t \tag{3.21}$$

Note. In analogy with the case where in (3.17) $(m^k_t)$ is Brownian motion [10], we call the $(\mu^k_t)$ the <u>innovations</u> of the $(x^k_t)$.

Proof. First of all, it is clear from the proof of (3.6) that $\langle m^k \rangle_t$ is $\mathcal{F}^x_t$-measurable. Next define the continuous processes $(\mu^k_t, \mathcal{F}^x_t, P)$ by

$$\mu^k_t = x^k_t - x^k_0 - \int_0^t \hat{h}^k_s d\langle m^k \rangle_s. \tag{3.22}$$

It is easy to check, using this relation and (3.17), that $(\mu^k_t) \in \mathcal{M}^2(\mathcal{F}^x_t, P)$. From (3.17) and (3.22)

$$\mu^k_t = \int_0^t (h^k_s - \hat{h}^k_s) ds + m^k_t = \int_0^t \tilde{h}_s ds + m^k_t, \text{ say,}$$

which, using the differentiation rule, yields

$$\mu^k_t \mu^j_t - \mu^k_s \mu^j_s = \int_s^t (\mu^k_\tau \tilde{h}^j_\tau + \mu^j_\tau \tilde{h}^k_\tau) d\tau + \int_s^t (\mu^k_\tau dm^j_\tau + \mu^j_\tau dm^k_\tau) + \\ + \langle m^k, m^j \rangle_t - \langle m^k, m^j \rangle_s,$$

so that

$$E(\mu_t^k\mu_t^j - \mu_s^k\mu_s^j|\mathcal{F}_s^x) = \langle m^k,m^j\rangle_t - \langle m^k,m^j\rangle_s$$

It follows that

$$\langle \mu^k,\mu^j\rangle_t = \langle m^k,m^j\rangle_t \text{ a.s.}$$

and this proves (3.18).

We prove the representations (3.20), (3.21) in a sequence of stages. Let $(y_t) \in \mathcal{M}^1_{loc}(\mathcal{F}_t^x,\mathcal{P})$. Firstly, define

$$\hat{z}_t^k = (\hat{h}^k o \mu^k)_t$$

Evidently $(\hat{z}_t^k) \in \mathcal{M}^2(\mathcal{F}_t^x,\mathcal{P})$ since $(\hat{h}_t^k) \in L^2(\langle\mu^k\rangle_t,\mathcal{F}_t^x,\mathcal{P})$. For each integer $\ell$, let $T_\ell$ be the stopping time

$$T_\ell = \inf\{t|\ \langle\hat{z}^k\rangle_t = \int_0^t (\hat{h}_s^k)^2 d\langle\mu^k\rangle_s > \ell \text{ for some } k\}, \text{ if this is finite}$$
$$= \infty \text{ otherwise}$$

Evidently $T_\ell \uparrow \infty$ a.s. Define

$$y_{\ell,t} = y_{t\wedge T_\ell},\ \hat{z}^k_{\ell,t} = \hat{z}^k_{t\wedge T_\ell},\ \hat{z}_{\ell,t} = \sum_{k=1}^n \hat{z}^k_{\ell,t},\ \mu^k_{\ell,t} = \mu^k_{t\wedge T_\ell},\ x^k_{\ell,t} = x^k_{t\wedge T_\ell}$$

Secondly, let $(\Lambda_{\ell,t})$ be the unique solution of the integral equation

$$\Lambda_{\ell,t} = 1 + \int_0^t \Lambda_{\ell,s} d\hat{z}_{\ell,t}$$

Since $\langle\hat{z}_\ell\rangle_t \le n\ell$ a.s. for all t, it follows from an adaptation [19, 3.3.6] of a technique due to Clark that

$$E(\Lambda_{\ell,\infty}) = 1$$

Hence by Lemma 3.2

$$(3.23) \qquad (\mu^k_{\ell,t} - \langle \hat{z}_\ell, \mu^k_\ell \rangle_t) = (\mu^k_{\ell,t} + \int_0^{t\wedge T_\ell} \hat{h}^k_s d\langle \mu^k \rangle_s) = (x^k_{t\wedge T_\ell} - x_0) \in \mathcal{M}^2(\mathcal{F}^x_{t\wedge T_\ell}, \mathcal{P}_\ell)$$

where the probability measure $\mathcal{P}_\ell$ is given by $d\mathcal{P}_\ell = \Lambda_{\ell,\infty} d\mathcal{P}$; furthermore the $(x^k_{\ell,t})$ are orthogonal martingales since

$$\langle x^k_\ell, x^j_\ell \rangle_t \equiv \langle \mu^k_\ell, \mu^j_\ell \rangle_t \equiv \langle \mu^k, \mu^j \rangle_{t\wedge T_\ell}$$

Thirdly, by Lemma 3.2 again

$$(y_{\ell,t} - \langle \hat{z}_\ell, y_\ell \rangle_t) \in \mathcal{M}^2(\mathcal{F}^x_{t\wedge T_\ell}, \mathcal{P}_\ell)$$

Hence, by Theorem 2.2, there exist processes $(\phi^k_{\ell,t}) \in L^2(\langle x^k_\ell \rangle_t, \mathcal{F}^x_{t\wedge T_\ell}, \mathcal{P}_\ell)$ such that

$$y_{\ell,t} - \langle \hat{z}_\ell, y_\ell \rangle_t = y_0 + \sum_{k=1}^{n} (\phi^k_\ell o x^k_\ell)_t$$

$$= y_0 + \sum_{k=1}^{n} \{(\phi^k_\ell o \mu^k_\ell)_t - \int_0^t \phi^k_{\ell,s} d\langle \hat{z}_\ell, \mu^k_\ell \rangle_s\}, \text{ by } (3.23),$$

which we can rewrite as

$$y_{\ell,t} - y_0 - \sum_{k=1}^{n} (\phi^k_\ell o \mu^k_\ell)_t = \langle \hat{z}_\ell, y_\ell \rangle_t - \sum_{k=1}^{n} \int_0^t \phi^k_{\ell,s} d\langle \hat{z}_\ell, \mu^k_\ell \rangle_s = w_t, \text{ say}$$

The left-hand side is a member of $\mathcal{M}^2(\mathcal{F}^x_{t\wedge T_\ell}, \mathcal{P}_\ell)$, whereas the right-hand side is a member of $\mathcal{A}(\mathcal{F}^x_{t\wedge T_\ell}, \mathcal{P}_\ell)$. Hence, by [16, p. 213], $w_t \equiv 0$. Thus we have shown that

$$y_{\ell,t} = y_0 + \sum_{k=1}^{n} (\phi^k_\ell o \mu^k_\ell)_t$$

Finally, define the processes $(\phi^k_t)$ by

$$\phi^k_t(\omega) = \phi^k_{\ell,t}(\omega) \text{ for } T_{\ell-1} < t \leq T_\ell .$$

It is easy to see that

$$y_t = y_0 + \sum_{k=1}^{n} (\phi^k o \mu^k)_t . \tag{3.24}$$

The integrability conditions on $(\phi^k_t)$ follows from the fact that if $y_t$ satisfies (3.24) then

$$\langle y \rangle_t = \sum_{k=1}^{n} \int_0^t (\phi^k_s)^2 d \langle \mu^k \rangle_s \text{ a.s.}$$

□

This is our most important result of this section. We use it to extend the likelihood ratio formula which has been obtained previously for the Brownian motion case [5,11].

Let $(\Omega, \mathcal{F}_t, \mathcal{P}_i)$, $t \in R_+$, $i=1,2$ be two families with the same measurable spaces $(\Omega, \mathcal{F}_t)$. Let $x^k_t$, $m^k_t$, $k=1,\ldots,n$ be families of real-valued functions on $\Omega$ such that

(i) the $(m^k_t)$ are continuous, pairwise orthogonal martingales in $\mathcal{M}^2(\mathcal{F}_t, \mathcal{P}_i)$ for both $i=1,2$,

(ii) there exist processes $(h^{k,i}_t) \in L^2(\langle m^k \rangle_t, \mathcal{F}_t, \mathcal{P}_i)$ $i=1,2$ such that under $\mathcal{P}_i$ $(x_t^k)$, $(m_t^k)$ are related according to

(3.25) $$x_t^k = x_0^k + \int_0^t h_s^{k,i} d\langle m^k\rangle_s + m_t^k, \quad t \in R_+$$

(iii) $\mathcal{P}_1 \sim \mathcal{P}_2$

THEOREM 3.3. The conditional likelihood ratio $L_t = E_1(\frac{d\mathcal{P}_2}{d\mathcal{P}_1} | \mathcal{F}_t^x)$ is given by the formula

(3.26) $$L_t = L_0 \exp \sum_{k=1}^n \{\int_0^t (\hat{h}_s^{k,2} - \hat{h}_s^{k,1}) d\mu_s^{k,1} - \frac{1}{2}\int_0^t (\hat{h}_s^{k,2} - \hat{h}_s^{k,1})^2 d\langle\mu^{k,1}\rangle_s\}$$

where $\hat{h}_t^{k,i} = E_i(h_t^{k,i} | \mathcal{F}_t^x)$ and $(\mu_t^{k,1})$ is the innovations of $(x_t^k, \mathcal{F}_t^x, \mathcal{P}_1)$ given by (3.19).

Proof. By Theorem 3.2 and the argument which led to (3.13) we can conclude that

$$\Lambda_t = L_t L_0^{-1} = 1 + \int_0^t \Lambda_s dn_s$$

where $n_t = \sum_{k=1}^n (\phi^k o \mu^{k,1})_t$ for some processes $(\phi_t^k) \in L^2(\langle\mu^{k,1}\rangle_t, \mathcal{F}_t^x, \mathcal{P}_1)$. So by Lemma 3.1

(3.27) $$\Lambda_t = \exp \sum_{k=1}^n (\int_0^t \phi_s^k \mu_s^{k,1} - \frac{1}{2}\int_0^t (\phi_s^k)^2 d\langle\mu^{k,1}\rangle_s)$$

By Lemma 3.2

$$(\mu_t^{k,1} - \int_0^t \phi_s^k d\langle\mu^{k,1}\rangle_s) \in \mathcal{M}_{loc}^2(\mathcal{F}_t^x, \mathcal{P}_2)$$

so that substitution using (3.19), (3.20) leads to

$$(x_t^k - x_0^k - \int_0^t \hat{h}_s^{k,1} d\langle m^k\rangle_s - \int_0^t \phi^{k,s} d\langle m^k\rangle_s) \in \mathcal{M}^2_{loc}(\mathcal{F}_t^x, \mathcal{P}_2)$$

On the other hand, again by Theorem 3.2,

$$(x_t^k - x_0^k - \int_0^t \hat{h}_s^{k,2} d\langle m^k\rangle_s) \in \mathcal{M}^2_{loc}(\mathcal{F}_t^x, \mathcal{P}_2)$$

so that

$$(\int_0^t (\phi_s^k + \hat{h}_s^{k,1} - \hat{h}_s^{k,2}) d\langle m^k\rangle_s) \in \mathcal{M}^2_{loc}(\mathcal{F}_t^x, \mathcal{P}_2).$$

By [16,p.213] this can happen only if

$\int_0^t (\phi_s^k + \hat{h}_s^{k,1} - \hat{h}_s^{k,2}) d\langle m^k\rangle_s = 0$ a.s. for all t, which in turn implies by well-known properties of stochastic integrals that

$$(\phi^k o \mu^k)_t = ((\hat{h}^{k,1} - \hat{h}^{k,2}) o \mu^k)_t \text{ a.s.}$$

Substituting this into (3.27) leads to (3.26). ¤

4. Nonlinear filtering of processes. We use Theorem 3.2 to obtain the structure of optimal least-squares estimates of an "unknown" process $(z_t)$ when the observed processes $(x_t^k)$ have the form of a semimartingale (3.17). The unknown process is also modelled as a semimartingale but of a more general form. We begin with a more abstract result.

Throughout this section the unknown process $(z_t)$ is a semimartingale,

(4.1) $$z_t = z_0 + a_t + n_t, \; t \in R_+$$

with $E|z_0| < \infty$, $(a_t) \in \mathcal{A}(\mathcal{F}_t, \mathcal{P})$, $(n_t) \in \mathcal{M}^1(\mathcal{F}_t, \mathcal{P})$ with $n_0 = 0$ a.s.

The observed process is vector-valued with components $(x_t^k, \mathcal{F}_t, \mathcal{P})$, $k=1,\ldots,n$. $\mathcal{F}_t^x = \sigma\{x_s^k | k=1,\ldots,n,\ 0 \le s \le t\}$. For any t, s in $R_+$, and any process $(f_t, \mathcal{F}_t, \mathcal{P})$ denote

$$\hat{f}_{t|s} = E(f_t | \mathcal{F}_s^x),\ \hat{f}_t = \hat{f}_{t|t}.$$

LEMMA 4.1. There exist a unique predictable process $(\bar{a}_t) \in \mathcal{A}(\mathcal{F}_t^x, \mathcal{P})$ and a unique martingale $(\bar{n}_t) \in \mathcal{M}^1(\mathcal{F}_t^x, \mathcal{P})$ with $\bar{n}_0 = 0$ a.s. such that

$$\hat{z}_t = \hat{z}_0 + \bar{a}_t + \bar{n}_t,\ t \in R_+ \tag{4.2}$$

Proof. Write $(a_t)$ as $a_t = a_t^1 - a_t^2$ where $(a_t^1)$, $(a_t^2)$ are increasing processes in $\mathcal{A}(\mathcal{F}_t, \mathcal{P})$, and check that $(\hat{a}_t^1, \mathcal{F}_t^x, \mathcal{P})$ and $(\hat{a}_t^2, \mathcal{F}_t^x, \mathcal{P})$ are both submartingales. By [17,VII T31] there exist unique increasing, predictable processes $(\bar{a}_t^i) \in \mathcal{A}(\mathcal{F}_t^x, \mathcal{P})$ and unique martingales $(\bar{m}_t^i) \in \mathcal{M}^1(\mathcal{F}_t^x, \mathcal{P})$ with $\bar{m}_0^i = 0$ a.s. such that

$$\hat{a}_t^i = \bar{a}_t^i + \bar{m}_t^i \text{ a.s. } t \in R_+. \tag{4.3}$$

Then, $\hat{z} = \hat{z} + \bar{a}_t + \bar{n}_t$, where

$$\bar{a}_t = \bar{a}_t^1 - \bar{a}_t^2, \tag{4.4}$$

$$\bar{n}_t = \hat{n}_t + \bar{m}_t^1 - \bar{m}_t^2 + \hat{z}_{0|t} - \hat{z}_0,$$

satisfies the assertion. ◻

Remark 4.1. (i) If $(a_t)$ in (4.1) is given by $a_t = \int_0^t f_s ds$ for some process $(f_s)$, then it is easy to check that $(\bar{a}_t)$ in (4.2) is given by

$$\bar{a}_t = \int_0^t \hat{f}_s ds \text{ a.s.}$$

(ii) For $t > s$, the predictor $\hat{z}_{t|s}$ is given by

$$\hat{z}_{t|s} = \hat{z}_{t|s} + E(\bar{a}_t | \mathcal{F}_s^x) - \bar{a}_s$$

For the remainder of this section we make the following additional assumptions:

(A1) $(n_t) \in \mathcal{M}^2(\mathcal{F}_t, \mathcal{P})$,

(A2) the processes $(x_t^k)$ satisfy

(4.5) $$x_t^k = x_0^k + \int_0^t h_s^k d\langle m_s^k \rangle + m_t^k, \quad t \in R_+,$$

where the $(m_t^k)$ are continuous, pairwise orthogonal martingales in $\mathcal{M}^2(\mathcal{F}_t, \mathcal{P})$ with $m_0^k = 0$ a.s., and $(h_t^k) \in L^2(\langle m^k \rangle_t, \mathcal{F}_t, \mathcal{P})$.

By [16,p.223], there exist predictable processes $(\phi_t^k, \mathcal{F}_t, \mathcal{P})$ such that

(4.6) $$E\int_0^\infty |\phi_t^k| d\langle m^k \rangle_t < \infty \text{ and } \langle m^k, n \rangle_t = \int_0^t \phi_s^k d\langle m^k \rangle_s \text{ a.s.}$$

THEOREM 4.1. The optimal filter satisfies the equation

(4.7) $$\hat{z}_t = \hat{z}_0 + \bar{a}_t + \sum_{k=1}^n ((\sigma^k + \hat{\phi}^k) o \mu^k)_t \text{ a.s.}, \quad t \in R_+$$

where, $\bar{a}_t$ is given by (4.4),

(4.8) $$\mu_t^k = x_t^k - x_0^k - \int_0^t \hat{h}_s^k d\langle \mu^k \rangle_s$$

is the innovations of $(x_t^k)$, and

$$\sigma_t^k = E\{(z_t-\hat{z}_t)(h_t^k-\hat{h}_t^k)|\mathcal{F}_t^x\}$$

is the conditional covariance of $z_t$ and $h_t^k$, given $\mathcal{F}_t^x$.

Proof. By Lemma 4.1 and Theorem 3.2 there exist processes $(\psi_t^k) \in L^2(\langle\mu^k\rangle_t, \mathcal{F}_t^x, \mathcal{P})$ such that

$$\hat{z}_t = \hat{z}_0 + \bar{a}_t + \sum_{k=1}^{n} (\psi^k \circ \mu^k)_t \tag{4.9}$$

For convenience denote

$$\tilde{z}_t = z_t - \hat{z}_t, \; \tilde{h}_t = h_t - \hat{h}_t$$

Then

$$\tilde{z}_t = a_t - \bar{a}_t + n_t - \sum_{k=1}^{n} (\psi^k \circ \mu^k)_t$$

and from (4.5), (4.8)

$$\mu_t^k = \int_0^t \tilde{h}_s^k d\langle\mu^k\rangle_s + m_t^k = \eta_t + m_t^k, \text{ say}$$

By the differentiation formula [4] we obtain,

$$\tilde{z}_t\mu_t^k - \tilde{z}_s\mu_s^k = \int_s^t \tilde{z}_\tau \tilde{h}_\tau^k d\langle\mu^k\rangle_\tau + \int_s^t \tilde{z}_s dm_s^k + \int_s^t \mu_\tau^k d(a_\tau - \bar{a}_\tau)$$

$$+ \int_s^t \eta_\tau dn_\tau + \int_s^t m_\tau^k dn_\tau + \int_s^t d\langle n,\mu^k\rangle_\tau - \sum_{j=1}^{n} \int_s^t \mu_\tau^k \psi_\tau^j d\mu_\tau^j - \sum_{j=1}^{n} \int_s^t \psi_j^\tau d\langle\mu^k,\mu^j\rangle_\tau$$

Now we take conditional expectations on both sides with respect to $\mathcal{F}_s^x$. The different terms are as follows.

$$E(\tilde{z}_t\mu_t^k-\tilde{z}_s\mu_s^k|\mathcal{F}_s^x) = 0 \text{ since } E(\tilde{z}_\tau|\mathcal{F}_\tau^x) = 0,$$

$$E(\int_s^t \tilde{z}_\tau\tilde{h}_\tau^k d\langle\mu^k\rangle_\tau|\mathcal{F}_s^x) = E(\int_s^t \sigma_\tau^k d\langle\mu^k\rangle_\tau|\mathcal{F}_s^x).$$

From (4.3), (4.4) it follows that $E(da_\tau-d\bar{a}_\tau|\mathcal{F}_\tau^x) = 0$, hence

$$E(\int_s^t \mu_\tau^k d(a_\tau-\bar{a}_\tau)|\mathcal{F}_s^x) = 0,$$

$$E(\int_s^t \eta_\tau dn_\tau|\mathcal{F}_s^x) = E(\int_s^t m_\tau^k dn_\tau|\mathcal{F}_s^x) = \sum_{j=1}^n E(\int_s^t \mu_\tau^k\psi_\tau^j d\mu_\tau^j|\mathcal{F}_s^x) = 0,$$

by the martingale property. Finally,

$$E(\int_s^t d\langle n,\mu^k\rangle_\tau|\mathcal{F}_s^x) = E\{\int_s^t E(\phi_\tau^k|\mathcal{F}_\tau^x)d\langle m^k\rangle_\tau|\mathcal{F}_s^x\} \text{ by (4.6), and}$$

$$\sum_{j=1}^n E(\int_s^t \psi_\tau^j d\langle\mu^k,\mu^j\rangle|\mathcal{F}_s^x) = E(\int_s^t \psi_\tau^k d\langle\mu^k\rangle_\tau|\mathcal{F}_s^x)$$

Thus,

$$0 = E\{\int_s^t (\sigma_\tau^k + \hat{\phi}_\tau^k - \psi_\tau^k)d\langle\mu^k\rangle_\tau|\mathcal{F}_s^x\},$$

which implies that

$$(\int_0^t (\sigma_s^k + \hat{\phi}_s^k - \psi_s^k)d\langle\mu^k\rangle_s) \in \mathcal{M}_{loc}^1(\mathcal{F}_t^x,P),$$

so that by [3,p.213]

$$\int_0^t (\sigma_s^k + \hat{\phi}_s^k) d\langle \mu^k \rangle_s = \int_0^t \psi_s^k d\langle \mu^k \rangle_s \text{ a.s. for all } t.$$

Using this relation in (4.9) yields (4.7). ◻

Remark 4.2. (i) Suppose $a_t = \int_0^t f_s ds$ and suppose that the unknown process noise and observation noise, $(m_t^k)$ and $(n_t)$, are independent. Then (4.7) becomes

$$\hat{z}_t = \hat{z}_0 + \int_0^t \hat{f}_s ds + \sum_{k=1}^n \int_0^t (\sigma^k + \hat{\phi}^k)_s (dx_s^k - \hat{h}_s^k d\langle \mu^k \rangle_s)$$

which has a striking resemblance with the Kalman-Bucy filter and the formula for nonlinear filters for the Brownian motion case [6,7,9].

(ii) The filter equation (4.7) is not recursive since the quantities $\bar{a}_t$, $\sigma_t^k$, $\hat{\phi}_t^k$ cannot generally be computed as functions of $\hat{z}_t$.

(iii) Equations similar to (4.7) can be obtained for the optimal prediction and smoothing estimates. We state the equation for the latter. For $s > t$,

$$\hat{z}_{t|s} = \hat{z}_t + \int_t^s \sigma_{t|\tau}^k d\mu_\tau^k.$$

where $\sigma_{\tau|t}^k = E\{(z_t - \hat{z}_{t|\tau})(h_t^k - \hat{h}_t^k) | \mathcal{F}_\tau^x\}$.

Appendix: Proof of Lemma 3.2

First of all since $(x_t)$ is continuous it is in $\mathcal{M}^2_{loc}(\mathcal{F}_t, P_1)$ if it is in $\mathcal{M}^1_{loc}(\mathcal{F}_t, P_1)$. To prove the latter it is equivalent to show that $(x_t\Lambda_t) \in \mathcal{M}^1_{loc}(\mathcal{F}_t, P)$. By the differentiation formula

$$x_t\Lambda_t = x_0\Lambda_0 + \int_0^t x_s d\Lambda_s + \int_0^t \Lambda_s dz_s - \int_0^t \Lambda_s d\langle n,z\rangle_s + \langle z,\Lambda\rangle_s$$

By (3.2) $d\langle z,\Lambda\rangle_t = \Lambda_t d\langle z,n\rangle_t$, so that

$$x_t\Lambda_t = x_0\Lambda_0 + \int_0^t x_s d\Lambda_s + \int_0^t \Lambda_s dz_s$$

which is clearly in $\mathcal{M}^1_{loc}(\mathcal{F}_t, P)$. Finally, (3.6) is a well-known property of semimartingales [4].

## REFERENCES


[1] R.S. Bucy, Nonlinear filtering theory, IEEE-T-AC 10 (2), 1965, p. 198.

[2] C. Dellacherie, Capacités et processus stochastiques, Berlin: Springer, 1972.

[3] C. Doléans-Dade, "Quelques applications de la formule de changement de variables pour les semimartingales," Z. für Wahrscheinlichkeits theorie verw. Geb 16 (3), 1970, 181-194.

[4] C. Doléans-Dade and P.A. Meyer, Intégrales stochastiques par rapport aux martingales locales, in Séminaire de Probabilités: IV, Lecture notes in Mathematics #124, Berlin: Springer, 1970, 77-107.

[5] T.E. Duncan, On the absolute continuity of measures, Ann. Math. Stat. 41 (1), 1970, 30-38.

[6] P.A. Frost, T. Kailath, An innovations approach to least squares estimation part III: nonlinear estimation in white Gaussian noise, IEEE-T-AC 16 (3), 1971, 217-226.

[7] M. Fujisaki, G. Kallianpur, H. Kunita, Stochastic differential equations for the nonlinear filtering problem, Osaka J. Math. 9 (1), 1972, 19-40.

[8] I.V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, Theory of Prob. and Appl. 5 (3), 1960, 285-301.

[9] T. Kailath, An innovations approach to least-squares estimation part I: linear filtering in additive white noise, IEEE-T-AC 13 (6), 1968, 646-655.

[10] __________, The innovations approach to detection and estimation theory, Proc. IEEE 58 (5), 1970, 680-695.

[11] T. Kailath and M. Zakai, Absolutely continuity and Radon-Nikodym derivatives for certain measures relative to Wiener measure, Ann. Math. Stat. 42 (1), 1971, 130-140.

[12] G. Kallianpur, C. Striebel, Estimation of stochastic systems: arbitrary system process with additive white noise observation errors, Ann. Math. Stat. 39 (3), 1968, 785-801.

[13] ______________________, Stochastic differential equations occuring in the estimation of continuous parameter stochastic processes. Theory of Prob. and Appl. 14 (4), 1969, 567-594.

[14] R.E. Kalman, R.S. Bucy, New results in linear filtering and prediction theory, Trans. ASME J. Basic Engrg. Ser. 83 (1), 1961, 95-108.

[15] F.B. Knight, A reduction of continuous square-integrable martingales to Brownian motion, in Martingales: a report on a meeting at Oberwolfach, Lecture Notes in Mathematics #190, Berlin: Springer 1971, 19-31.

[16] H. Kunita and S. Watanabe, On square integrable martingales, Nagoya Math. J. 30, 1967, 209-245.

[17] P.A. Meyer, Probabilitiés et potentiel, Paris: Hermann, 1966, (English translation, Probability and Potentials, Waltham, Mass.: Blaisdell, 1966).

[18] P.A. Meyer, Démonstration simplifiée d'un Theoreme de Knight, Séminaire de Probabilites V, Lecture Notes in Mathematics #204, Berlin: Springer, 1971, 191-195.

[19] J.H. Van Schuppen, Estimation theory for continuous time processes, a martingale approach, Ph.D. dissertation, Univ. of Calif., Berkeley,

also Memo-M405, Elec. Res. Lab., Univ. of Calif., Berkeley, 1973.

[20] J.H. Van Schuppen and E. Wong, Transformation of local martingales under a change of law, Memo-M385, Elec. Res. Lab., Univ. of Calif., Berkeley, 1973, to appear in Ann. of Prob.