Z. Wahrscheinlichkeitstheorie verw. Gebiete
66, 1-17 (1984)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete


# Measure-Valued Equations for the Optimum Filter in Finitely Additive Nonlinear Filtering Theory*

G. Kallianpur and R.L. Karandikar
Department of Statistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA

## 0. Introduction

The Ito stochastic calculus - in particular, the theory of stochastic differential equations (SDE) and Ito's formula - has proved itself a versatile and powerful technique in the development of nonlinear filtering and control theory. It is convenient to begin with a brief description of the filtering problem.

Let the unobserved signal process $X=(X_u)$ be a Markov process taking values in $\mathbb{R}^d$. It is assumed that its generator is known. The canonical model of the observation process is given by

$$Y_t=\int_0^t h_u(X_u)\,du+W_t, \qquad 0\leqq t\leqq T, \tag{0.1}$$

where $h\colon [0,T]\times\mathbb{R}^d\to\mathbb{R}^m$ is a measurable function such that

$$\int_0^T |h_u(X_u)|^2\,du<\infty \quad \text{a.s.}, \tag{0.2}$$

and $W=(W_t)$ is a standard, $m$-dimensional Wiener process.

The problem is to derive an SDE for the optimal nonlinear filter and to prove uniqueness of its solution under suitable conditions. There have been essentially three types of equations considered in the literature:

(1) Under very general assumptions on the dependence between $X$ and $W$, Fujisaki, Kallianpur and Kunita obtained a gneral SDE for the conditional expectations $\Pi_t(f)=E[f(X_t)|Y_s,\ 0\leqq s\leqq t]$ for a class of $f$'s belonging to the domain of the generator of $X$. This equation will be referred to as the FKK equation [5].

(2) An equivalent and sometimes easier equation to work with is the one for the unnormalized conditional expectation called the Zakai equation [14].

Under suitable conditions, the conditional probability density and hence the unnormalized density exist and satisfy stochastic partial differential equa-

* This research has been supported by AFOSR Grant No. F49620 82 C 0009

tions derived from the FKK and Zakai equations. Most of the interesting recent work on the subject has centered around the study of these stochastic PDE's. (Krylov and Rozovskii [9], Pardoux [11]). Recently, Clark and Davis have used the Kallianpur-Striebel or Bayes formula and the Zakai equation for the unnormalized density to obtain a robust solution to the filtering problem [2, 3].

(3) When the signal and observation noise are independent, Kunita introduced a third equation, a stochastic equation for the conditional probability measure for which he proved the existence of a unique solution [10]. Kunita's equation is not an SDE and analogues of this equation will be called Kunita type equations in this paper. It was left to Szpirglass to show the equivalence of the Kunita and FKK equations in the sense that a solution of one is a solution of the other [13]. It was further shown in [13] that the Zakai equation and the corresponding Kunita equation are also equivalent.

A point of view put forward by Balakrishnan and discussed in his recent papers ([1] and the references in [8]) questions the practical utility of the observation model (0.1) on the ground that the results obtained cannot be instrumented. A theoretical model for Gaussian white noise proposed in [1], that permits us to deal directly with the observed phenomena requires us to use the theory of weak distributions or cylinder measures (first introduced by Segal in connection with problems of physics and later developed further by Gross [12, 6]). The reason is that the space of observations is a Hilbert space of zero Wiener measure.

In our earlier paper [8] the finitely additive white noise approach to nonlinear filtering theory was systematically developed in the important special case when the signal and observation noise are independent. Since many of the difficulties inherent in the calculus of semimartingales simply do not arise in this theory it was possible to obtain the results under less restrictive conditions. The theory naturally lends itself to robust procedures. Moreover, it was shown in [8] that the robust solutions obtained by Davis can be recovered via the white noise approach. A body of results more or less parallel to the existing theory based on stochastic calculus was created in [8].

Our purpose in the present paper is to study the finitely additive white noise theory in a more general framework so as to include applications to signal and observation processes taking values in infinite dimensional separable Hilbert spaces. The chief difficulty here is that we have no conditional density since there is no Lebesgue measure (or any natural measure) in Hilbert space. In place of the partial differential equations of [8] we work with finitely additive analogues of measure-valued equations of FKK, Zakai and Kunita types. These equations (of which the first two are differential equations) are introduced in Sect. 3. The equivalence and the uniqueness of solutions of these equations are established in the four principal results of this paper, Theorems 3.1–3.4. An approximation procedure for obtaining the solution as well as a certain robustness property are also derived. These theorems are based on auxiliary results proved in Sect. 2, on equations governing measures induced by multiplicative transformations of Markov processes.

Before formulating the white noise version of the nonlinear filtering problem in Sect. 1, we summarize properties of weak distributions essential for our purpose and state the finitely additive Bayes formula which is one of the main tools of our investigation and whose proof is given in detail in [8].

A feature worth mentioning both of this paper and its predecessor is that the finitely additive white noise approach stands by itself and is entirely different in spirit from the methods involving approximations to solutions of Ito or Stratonovich stochastic differential equations, discussed, for example, in Ikeda and Watanabe's recent book [7].

## 1. Finitely Additive White Noise Version of the Nonlinear Filtering Model

For the sake of completeness and the convenience of the reader we summarize below the basic concepts of weak distributions (or cylinder measures) on Hilbert space which are essential to the understanding of the finitely additive white noise approach adopted in this paper. The definitions regarding integration with respect to cylinder measures in Hilbert space are taken from Gross [6]. The definition of conditional expectation and the Bayes formula (in the finitely additive set up) are from our earlier paper [8] to which we refer the reader for details.

Let $H$ be a separable Hilbert space and $\mathscr{P}$ the set of orthogonal projections on $H$ having finite dimensional range. For $P\in\mathscr{P}$, let $\mathscr{C}_P=\{P^{-1}B\colon B$ a Borel set in Range $P\}$. Let $\mathscr{C}=\bigcup_{P\in\mathscr{P}}\mathscr{C}_P$. A cylinder measure $\mathbf{n}$ on $H$ is a finitely additive measure on $(H,\mathscr{C})$ such that its restriction to $\mathscr{C}_P$ is countably additive for each $P\in\mathscr{P}$.

Let $L$ be a representative of the weak-distribution corresponding to the cylinder measure $\mathbf{n}$. This means that $L$ is a linear map from $H^*$ (identified with $H$) into $\mathscr{L}(\Omega_1,\mathscr{A}_1,\Pi_1)$ – the space of all random variables on a countably additive probability space $(\Omega_1,\mathscr{A}_1,\Pi_1)$ – such that

$$\mathbf{n}(h\colon((h,h_1),\dots,(h,h_k))\in B)=\Pi_1((L(h_1),L(h_2),\dots,L(h_k))\in B) \tag{1.1}$$

for all Borel sets $B$ in $\mathbb{R}^k$, $h_1,\dots,h_k\in H$, $k\geqq 1$. (Two maps $L$, $L'$ are said to be equivalent if both satisfy (1.1) and the equivalence class of such maps is the weak distribution corresponding to $\mathbf{n}$).

A function $f$ on $H$ is called a tame function if it is of the form

$$f(y)=\phi((y,h_1),\dots,(y,h_k)) \tag{1.2}$$

for some $k\geqq 1$, $h_1,\dots,h_k\in H$ and a Borel function $\phi\colon\mathbb{R}^k\to\mathbb{R}$. For a tame function $f$ given by (1.2), we associate the random variable $\phi(L(h_1),\dots,L(h_k))$ (on $(\Omega_1,\mathscr{A}_1,\Pi_1)$) and denote it by $f^\sim$. We extend this map $f\to f^\sim$ to a larger class of functions as follows:

*Definition.* Let $\mathscr{L}(H,\mathscr{C},\mathbf{n})$ be the class of continuous functions $f$ on $H$ such that the net $\{(f\circ P)^\sim: P\in\mathscr{P}\}$ (here $P_1<P_2$ if Range $P_1\subseteq$ Range $P_2$) is Cauchy in $\Pi_1$-

measure. Furthermore, for $f \in \mathscr{L}(H, \mathscr{C}, \mathbf{n})$, let

$$f^{\sim} = \lim_{P \in \mathscr{P}} \text{in Probability} \quad (f \circ P)^{\sim}.$$

It can be easily seen that the map $f \to f^{\sim}$ is linear, multiplicative. It is easy to see that for a tame function $f$, the distribution of $f$ and $f^{\sim}$ are identical. Furthermore, the distribution of $f^{\sim}$ for $f \in \mathscr{L}(H, \mathscr{C}, \mathbf{n})$ depends only on $f$ and $n$ and is independent of the representative of the weak distribution. In view of this we make the following definition:

*Definition.* The function $f \in \mathscr{L}(H, \mathscr{C}, \mathbf{n})$ is integrable (with respect to $\mathbf{n}$) if $\int |f^{\sim}| d\Pi_1 < \infty$ and then for $C \in \mathscr{C}$ define the integral of $f$ w.r.t. $\mathbf{n}$ over $C$, denoted by $\int_C f d\mathbf{n}$ by

$$\int_C f d\mathbf{n} = \int_{\Omega_1} (1_C)^{\sim} f^{\sim} d\Pi_1. \tag{1.3}$$

The finitely additive cylinder measure $\mathbf{m}$ on $(H, \mathscr{C})$ such that for all $h \in H$,

$$\mathbf{m}\{y \in H: (y, h) \leqq a\} = \frac{1}{\sqrt{2\pi} \|h\|} \int_{-\infty}^{a} \exp\left(-\frac{1}{2} \frac{x^2}{\|h\|^2}\right) dx \tag{1.4}$$

is called the *canonical Gauss measure* on $H$. The identity map $e$ on $H$, considered as a map from $(H, \mathscr{C}, \mathbf{n})$ into $(H, \mathscr{C})$ is called the Gaussian *white noise.*

The abstract version of the white noise nonlinear filtering model is given by

$$y = \xi + e \tag{1.5}$$

where $\xi$ is an $H$-valued random variable defined on a countably additive probability space $(\Omega, \mathscr{A}, \Pi)$, independent of $e$. To be more precise, let $E = H \times \Omega$ and

$$\mathscr{F} = \bigcup_{P \in \mathscr{P}} \mathscr{C}_P \otimes \mathscr{A}$$

where $\mathscr{C}_P \otimes \mathscr{A}$ is the usual product $\sigma$ field. For $P \in \mathscr{P}$, let $\alpha_P$ be the product measure $(\mathbf{m}|\mathscr{C}_P) \otimes \Pi$. (Observe that the restriction $\mathbf{m}|\mathscr{C}_P$ is countably additive.) It is easily seen that $\alpha_P = \alpha_{P'}$ on $(\mathscr{C}_P \otimes \mathscr{A}) \cap (\mathscr{C}_{P'} \otimes \mathscr{A})$. Thus we can get a unique finitely additive probability measure $\alpha$ on $(E, \mathscr{F})$ such that $\alpha = \alpha_P$ on $\mathscr{C}_P \otimes \mathscr{A}$.

Now, let $e$, $\xi$, $y$ be $H$-valued maps on $E$ defined by

$$\begin{aligned} &e(h, \omega) = h \\ &\xi(h, \omega) = \xi(\omega) \\ &y(h, \omega) = e(h, \omega) + \xi(h, \omega), \qquad (h, \omega) \in H \times \Omega. \end{aligned} \tag{1.6}$$

Now (1.5) is the abstract version of our filtering model on $(E, \mathscr{F}, \alpha)$.

Let $g$ be an integrable function on $(\Omega, \mathscr{A}, \Pi)$. As usual, we are interested in $E(g|y)$. In analogy with the usual definition of conditional expectation, we make the following definition.

*Definition.* If there exists a $v \in \mathscr{L}(H, \mathscr{C}, \mathbf{n})$ such that for all $C \in \mathscr{C}$

$$\int g(\omega)\, 1_C(y(h,\omega))\, d\alpha(h,\omega) = \int_C v(y)\, d\mathbf{n}(y), \tag{1.7}$$

then we define $v$ to be the conditional expectation of $g$ given $y$ and express it as

$$E(g \mid y) = v.$$

*Remark.* It is easy to see that the integrand in (1.7) is $\mathscr{C}_P \otimes \mathscr{A}$ measurable, where $C \in \mathscr{C}_P$, and since the restriction of $\alpha$ to $\mathscr{C}_P \otimes \mathscr{A}$ is countably additive, the integral appearing on the left hand side of (1.7) is well defined.

It was shown in [8] that such a $v$ does exist. We state the related result below.

**Theorem 1.1** (see [8]). *(Bayes Formula in the finitely additive set up).*
*Let $y$, $\xi$ be as in* (1.5). *Let $g$ be an integrable function on $(\Omega, \mathscr{A}, \Pi)$. Then*

$$E(g \mid y) = \frac{\int g(\omega) \exp((y, \xi(\omega)) - \frac{1}{2}\|\xi(\omega)\|^2)\, d\Pi(\omega)}{\int \exp((y, \xi(\omega)) - \frac{1}{2}\|\xi(\omega)\|^2)\, d\Pi(\omega)}.$$

The specific nonlinear filtering model of interest to us can now be formulated in the form (1.5).

Let $(S, \mathscr{S})$ be a measurable space, $S$ being the state space of the signal process $X_t$. $X_t$ is further assumed to be a Markov process defined on a probability space $(\Omega, \mathscr{A}, \Pi)$.

Let $\mathscr{K}$ be a separable Hilbert space and let $H = L^2([0,T], \mathscr{K})$. Let $h = (h_t)$ be a measurable function from $[0,T] \times S$ into $\mathscr{K}$ such that

$$E \int_0^T \|h_s(X_s)\|_{\mathscr{K}}^2\, ds < \infty. \tag{1.8}$$

Let $e = (e_t)$ be $\mathscr{K}$ valued white noise. Consider the nonlinear filtering model

$$y_s = h_s(X_s) + e_s\colon\ 0 \leqq s \leqq T. \tag{1.9}$$

Applying Theorem 1.1 to the model (1.9) with $s \in [0,t]$, $(0 \leqq t \leqq T)$, we have for $y \in H$,

$$E(f(X_t) \mid y_s\colon 0 \leqq s \leqq t) = \frac{1}{\Gamma_t^y(S)} \int f(x)\, d\Gamma_t^y(x) \tag{1.10}$$

where for $B \in \mathscr{S}$

$$\Gamma_t^y(B) = E\, 1_B(X_t) \exp\left(\int_0^t (y_s, h_s(X_s))\, ds - \frac{1}{2}\int_0^t \|h_s(X_s)\|^2\, ds\right). \tag{1.11}$$

In view of (1.10), $\Gamma_t^y$ is called the unnormalized conditional distribution of $X_t$ given $\{y_s\colon 0 \leqq s \leqq t\}$. Let $F_t^y$ be defined by

$$F_t^y(B) = \frac{1}{\Gamma_t^y(S)} \cdot \Gamma_t^y(B), \qquad B \in \mathscr{S}. \tag{1.12}$$

Then we have from (1.10) and (1.12), for $y \in H$

$$E(f(X_t)|y_s\colon 0 \leqq s \leqq t) = \int f(x)\, dF_t^y(x). \tag{1.13}$$

Thus, $F_t^y$ is the conditional distribution of $X_t$ given $\{y_s\colon 0 \leqq s \leqq t\}$.

As mentioned in the Introduction, when $X_t$ is a diffusion Markov process taking values in $\mathbb{R}^d$ and $h\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^m$, then $\mathscr{K} = \mathbb{R}^m$ and we have discussed the situation in [8]. The measure $\Gamma_t^y$ (and also the measure $F_t^y$) then admits a density with respect to Lebesgue measure in $\mathbb{R}^d$ and further satisfies a partial differential equation under certain conditions. This equation has a unique solution and thus $\Gamma_t^y$ can be computed. In some cases (as in Kunita's paper [10]), $S$ is taken to be a compact Hausdorff space in which case $\mathscr{K}$ can again be assumed to be $\mathbb{R}^m$ $(m \geqq 1)$. If $S$ is an infinite-dimensional Hilbert space, it is natural for $\mathscr{K}$ also to be infinite dimensional. As will be seen, the dimensionality of $\mathscr{K}$ plays no role in the proofs of our results. The derivation of the main results of this paper follows a general pattern which has nothing to do with filtering theory as such but is concerned with transformations of distributions of a Markov process by multiplicative functionals of the type (1.12). It is convenient to study the latter problem separately in the next section and apply it to nonlinear filtering theory in Sect. 3.

## 2. Multiplicative Transformations of Markov Processes

Let $(E, \mathscr{E})$ be a measurable space. Let $\mathscr{B}(E)$ be the class of bounded $\mathscr{E}$ measurable functions on $E$. For $f_n$, $f \in \mathscr{B}(E)$, we say that $f_n \to f$ weakly if $f_n \to f$ pointwise and $f_n$ is uniformly bounded.

Let $(Z_t)$ be an $E$-valued homogeneous Markov process with the associated semigroup $(P_t)$ (acting on $\mathscr{B}(E)$). Let

$$\mathscr{B}_0 = \{f \in \mathscr{B}(E)\colon P_t f \to f \text{ weakly as } t \to 0\}.$$

Assume that $(Z_t)$ is such that

(2.1)
- (i) For all $f \in \mathscr{B}(E)$, $\exists f_n \in \mathscr{B}_0$ such that $f_n \to f$ weakly.
- (ii) For all $f \in \mathscr{B}(E)$, $(P_t f)(x)$ is a jointly measurable function of $(t, x)$.
- (iii) For all $f \in \mathscr{B}(E)$, $f(Z_t)$ is a $\mathscr{F}_t$-progressively measurable function where $\mathscr{F}_t = \sigma(Z_s\colon s \leqq t)$.

Let $\mathscr{D}$ be the class of functions of $f$ in $\mathscr{B}_0$ such that there exists $f_0 \in \mathscr{B}_0$ satisfying

$$(P_t f)(x) = f(x) + \int_0^t (P_s f_0)(x)\, ds, \qquad \forall x \in E. \tag{2.2}$$

It is easy to see that (2.2) determines $f_0$ uniquely. Define a map $L$ from $\mathscr{D}$ into $\mathscr{B}_0$ as follows. For $f \in \mathscr{D}$, let $Lf = f_0$ where $f$, $f_0$ are related by (2.2). $L$ is called the extended generator of $P_t$.

For a measure $\mu$ on $(E, \mathscr{E})$ and $f \in \mathscr{B}$, let $\langle f, \mu \rangle$ denote $\int f\, d\mu$.

The lemma proved below has been used by Szpirglas in a similar connection [13].

**Lemma 2.1.** *Let* $\mathscr{D}_2=\{f\in\mathscr{D}\colon Lf\in\mathscr{D}\}$. *Let* $\mu_1$, $\mu_2$ *be finite measures on* $(E,\mathscr{E})$ *such that for all* $f\in\mathscr{D}_2$

$$\langle f,\mu_1\rangle=\langle f,\mu_2\rangle. \tag{2.3}$$

*Then* $\mu_1=\mu_2$.

*Proof.* In view of the assumption (i) (2.1), it suffices to show that (2.3) holds for all $f\in\mathscr{B}_0$.

For $\lambda>0$, let $R_\lambda$ be the resolvent of $P_t$, i.e. for $f\in\mathscr{B}$,

$$(R_\lambda f)(x)=\int_0^\infty e^{-\lambda t}(P_t f)(x)\,dt.$$

Then it is well known that (i) $R_\lambda f\in\mathscr{D}$, $L(R_\lambda f)=\lambda R_\lambda f-f$ and (ii) for $f\in\mathscr{B}_0$, $\lambda R_\lambda f\to f$ weakly as $\lambda\to\infty$ (see [4]).

Now, fix $f\in\mathscr{D}$. Since $L(R_\lambda f)=\lambda R_\lambda f-f$, it follows that $\lambda R_\lambda\in\mathscr{D}_2$ and hence (2.3) holds for $\lambda R_\lambda f$. Since $\lambda R_\lambda f\to f$ weakly, this and the dominated convergence theorem imply that (2.3) holds for $f$.

Now for $f\in\mathscr{B}_0$, $\lambda R_\lambda f\in\mathscr{D}$ and (2.3) holds for $\lambda R_\lambda f$ implies as above that (2.3) holds for $f$. As remarked earlier, this completes the proof.

The following Grownwall-type inequality will be useful later in this paper.

**Lemma 2.2.** *Let* $\alpha$ *be a positive measurable function such that* $\int_0^T \alpha(s)\,ds<\infty$.

(i) *If* $a(t)$ *is a positive bounded measurable function such that*

$$a(t)\leqq\int_0^t \alpha(s)\,a(s)\,ds, \qquad 0\leqq t\leqq T$$

*then* $a(t)\equiv 0$.

(ii) *If* $a_n(t)$ *is a sequence of bounded measurable functions such that*

$$a_{n+1}(t)\leqq\int_0^t \alpha(s)\,a_n(s)\,ds, \qquad 0\leqq t\leqq T, \qquad n\geqq 1$$

*then* $a_n(t)\to 0$.

*Proof.* Let $\eta(s)=\int_0^t \alpha(u)\,du$, $t_0=\eta(T)+T$ and for $0\leqq t\leqq t_0$,

$$\beta(t)=\inf\{s\geqq 0\colon (s+\eta(s))\geqq t\}.$$

Then $\beta(t)$ is a continuous strictly increasing function and $\beta(t)+\eta(\beta(t))=t$. Thus

$$\frac{d\eta(\beta(t))}{dt}\leqq 1.$$

Now let $b(t)=a(\beta(t))$. Then

$$b(t)=a(\beta(t))\leqq\int_0^{\beta(t)} a(s)\,d\eta(s)$$

$$=\int_0^t a(\beta(s))\,d\eta(\beta(s))=\int_0^t b(s)\,d\eta(\beta(s))\leqq\int_0^t b(s)\,ds.$$

Now, by Grownwall's inequality, it follows that $b(t)\equiv 0$.

Similarly, letting $b_n(t)=a_n(\beta(t))$, we have

$$b_{n+1}(t)\leqq\int_0^t b_n(s)\,ds, \qquad t\leqq t_0$$

and hence by induction, we have

$$b_{n+1}(t)\leqq C\cdot\frac{t^n}{n!}, \qquad 0\leqq t\leqq t_0, \qquad n\geqq 1$$

where $C$ is the bound of $a_1(t)$. Thus $b_n(t)\to 0\ \forall t$.

We now consider transformation of the distribution of $Z_t$ by a multiplicative functional of the type (1.5) and characterize the transformed measure as a unique solution to certain equations.

Let $g$ be a real-valued measurable function on $[0,T]\times E$ such that

(2.4) (i) $|g_s(x)|\leqq\alpha(s)$, where $\int_0^T \alpha(s)\,ds<\infty$.

(ii) $\int_0^t g_s(x)\,ds\leqq C$, $0\leqq C<\infty$.

For $0\leqq t\leqq T$, defien $G_t$, $N_t$ on $\mathscr{E}$ by

$$G_t(A)=E\,1_A(Z_t)\exp\left(\int_0^t g_s(Z_s)\,ds\right)$$

and

$$N_t(A)=\frac{1}{G_t(E)}\cdot G_t(A), \qquad A\in\mathscr{E}.$$

Then it can be easily seen that

(2.5) for $0\leqq t\leqq T$, $G_t$ is a finite positive measure on $(E,\mathscr{E})$ such that for all $A\in\mathscr{E}$, $G_t(A)$ is a bounded Borel function of $t$ and $G_0(A)=E\,1_A(Z_0)$,

and

(2.6) for $0\leqq t\leqq T$, $N_t$ is a probability measure on $(E,\mathscr{E})$ such that for all $A\in\mathscr{E}$, $N_t(A)$ is a Borel function of $t$ and $N_0(A)=E\,1_A(Z_0)$.

**Theorem 2.3.** (a) $G_t$ *satisfies*

(2.7) $$\langle f,G_t\rangle=\langle P_t f,G_0\rangle+\int_0^t\langle g_s(P_{t-s}f),G_s\rangle\,ds.$$

(b) $(G_t)$ *is the unique solution of* (2.7) *in the class of measures* $(K_t)$ *satisfying* (2.5).

(c) *Let* $G_t^n$ *be defined inductively as follows:*

$$G_t^0(A) = E\,1_A(Z_t)$$

*and*

$$G_t^{n+1}(A) = G_t^0(A) + \int_0^t \langle g_s(P_{t-s}1_A), G_s^n\rangle\, ds. \tag{2.8}$$

*Then* $G_t^n$ *converges uniformly in* $t$ *(in the total variation norm) to* $G_t$.

*Proof.* (a) Since

$$f(Z_t)\exp\left(\int_0^t g_s(Z_s)\,ds\right) = f(Z_t)\left[1 + \int_0^t \exp\left(\int_0^r g_s(Z_s)\,ds\right) g_r(Z_r)\,dr\right],$$

we have for $f \in \mathscr{B}(E)$,

$$\begin{aligned}
\langle f, G_t\rangle &= Ef(Z_t) + \int_0^t Ef(Z_t)\exp\left(\int_0^r g_s(Z_s)\,ds\right) g_r(Z_r)\,dr \\
&= E(P_t f)(Z_0) + \int_0^t E\{Ef(Z_t)|Z_s\colon 0 \leqq s \leqq r\}\exp\left(\int_0^r g_s(Z_s)\,ds\right) g_r(Z_r)\,dr \\
&= E(P_t f)(Z_0) + \int_0^t E(P_{t-r}f)(Z_r)\exp\left(\int_0^r g_s(Z_s)\,ds\right) g_r(Z_r)\,dr \\
&= \langle P_t f, G_0\rangle + \int_0^t \langle g_r(P_{t-r}f), G_r\rangle\,dr.
\end{aligned}$$

(b) Let $(K_t)$ satisfy (2.5) and (2.7). Then for $f \in \mathscr{B}(E)$, we have

$$\langle f, G_t - K_t\rangle = \int_0^t \langle g_s(P_{t-s}f),\ G_s - K_s\rangle\, ds. \tag{2.9}$$

Let $a(t) = \sup\limits_{0 \leqq s \leqq t}\ \sup\limits_{\substack{f \in \mathscr{B}(E)\\ \|f\| \leqq 1}} |\langle f, G_s - K_s\rangle|$. ($\|f\| = \sup\limits_{x \in E}|f(x)|$). Then (2.9) gives

$$a(t) \leqq \int_0^t \alpha(s)\,a(s)\,ds$$

and hence, by Lemma 2.2, we have $a(t) \equiv 0$. Thus $G_t = K_t$.

(c) Similarly, if we define

$$a_n(t) = \sup_{0 \leqq s \leqq t}\ \sup_{\substack{f \in \mathscr{B}(E),\\ \|f\| \leqq 1}} |\langle f, G_s^n - G_s\rangle|,$$

then from (2.7) and (2.8), we get for $n \geqq 1$,

$$\langle f, G_t^n - G_t\rangle = \int_0^t \langle g_s P_{t-s} f,\ G_s^{n-1} - G_s\rangle\, ds$$

and hence

$$a_n(t) \leqq \int_0^t \alpha(s)\, a_{n-1}(s)\, ds.$$

Lemma 2.2 now implies that $a_n(t) \to 0$ for all $t$. This is the same as the assertion that $G_t^n$ converges to $G_t$ uniformly in the total variation norm.

**Lemma 2.4.** *Let $K_t$ satisfy* (2.5). *Then* $(K_t)$ *satisfies* (2.13) *iff it satisfies* (2.14).

(2.13) $$\langle f, K_t\rangle = \langle P_t f, K_0\rangle + \int_0^t \langle g_s(P_{t-s} f), K_s\rangle\, ds, \quad f \in \mathscr{B}(E)$$

(2.14) $$\langle f, K_t\rangle = \langle f, K_0\rangle + \int_0^t \langle Lf + g_s f, K_s\rangle\, ds, \quad f \in \mathscr{D}.$$

*Proof.* Let us write $\beta_t(f) = \langle f, K_t\rangle$ and for $f \in \mathscr{D}$, let $\delta_t(f)$ be the difference of the right hand sides of (2.13) and (2.14), i.e.,

(2.15) $$\delta_t(f) = \beta_0(P_t f) + \int_0^t \beta_s(g_s P_{t-s} f)\, ds - \beta_0(f) - \int_0^t \beta_s(Lf + g_s f)\, ds.$$

Since for $f \in \mathscr{D}$,

$$P_t f = f + \int_0^t P_s(Lf)\, ds,$$

we have, using Fubini's theorem

(2.16) $$\beta_0(P_t f) = \beta_0(f) + \int_0^t \beta_0(P_s Lf)\, ds$$

and

(2.17) $$\beta_s(g_s P_{t-s} f) = \beta_s(g_s f) + \int_0^{t-s} \beta_s(g_s P_\tau Lf)\, d\tau.$$

From (2.15), (2.16), and (2.17) we get

(2.18) $$\delta_t(f) = \int_0^t \beta_0(P_s Lf)\, ds + \int_0^t \int_0^{t-s} \beta_s(g_s P_\tau Lf)\, d\tau\, ds - \int_0^t \beta_s(Lf)\, ds.$$

Again, using Fubini's theorem, we get

(2.19) $$\delta_t(f) = \int_0^t \beta_0(P_s Lf)\, ds + \int_0^t \int_0^{t-\tau} \beta_s(g_s P_\tau Lf)\, ds\, d\tau - \int_0^t \beta_s(Lf)\, ds.$$

Now suppose $(K_t)$ satisfies (2.13). Fix $f \in \mathscr{D}$. To show that $(K_t)$ satisfies (2.14), we will show $\delta_t(f) \equiv 0$. Applying (2.13) to $Lf$, we get

$$\beta_u(Lf) = \beta_0(P_u Lf) + \int_0^u \beta_r(g_r P_{u-r} Lf)\, dr$$

and hence

$$\int_0^t \beta_u(Lf)\, du = \int_0^t \beta_0(P_u Lf)\, du + \int_0^t \int_0^u \beta_r(g_r P_{u-r} Lf)\, dr\, du.$$

Substituting $r=s$ and $u-r=\tau$, we have

$$\int_0^t \beta_u(Lf)\,du=\int_0^t \beta_0(P_u Lf)\,du+\int_0^t\int_0^{t-s} \beta_s(g_s P_\tau Lf)\,d\tau\,ds \tag{2.20}$$

and hence from (2.20) and (2.18) we have $\delta_t(f)=0$.

Now, for the other part, assume that $(K_t)$ satisfies (2.14). Fix $f\in\mathscr{D}_2$. By rearrangement of terms in (2.14), we have

$$\int_0^t \beta_s(g_s f)\,ds=\beta_t(f)-\beta_0(f)-\int_0^t \beta_s(Lf)\,ds. \tag{2.21}$$

Applying (2.21) to $P_\tau Lf$, we have

$$\int_0^{t-\tau} \beta_s(g_s P_\tau Lf)\,ds=\beta_{t-\tau}(P_\tau Lf)-\beta_0(P_\tau Lf)-\int_0^{t-\tau} \beta_s(P_\tau L Lf)\,ds. \tag{2.22}$$

Thus from (2.19) and (2.22), we have

$$\begin{aligned}\delta_t(f)&=\int_0^t \beta_{t-\tau}(P_\tau Lf)\,d\tau-\int_0^t\int_0^{t-\tau} \beta_s(P_\tau L Lf)\,ds\,d\tau-\int_0^t \beta_s(Lf)\,ds\\ &=\int_0^t \beta_s(P_{t-s} Lf)\,ds-\int_0^t\int_0^{t-s} \beta_s(P_\tau L Lf)\,d\tau\,ds-\int_0^t \beta_s(Lf)\,ds.\end{aligned} \tag{2.23}$$

Now

$$\beta_s(P_{t-s}Lf)=\beta_s(Lf)+\int_0^{t-s} \beta_s(P_\tau L Lf)\,d\tau. \tag{2.24}$$

From (2.23) and (2.24), it follows that $\delta_t(f)=0$. Thus (2.13) holds for all $f\in\mathscr{D}_2$. In view of Lemma 2.1 this implies that (2.13) holds for all $f\in\mathscr{B}(E)$.

As a consequence of Theorem 2.3 and Lemma 2.4, we have the following theorem.

**Theorem 2.5.** (a) *$G_t$ satisfies*

$$\langle f,G_t\rangle=\langle f,G_0\rangle+\int_0^t \langle Lf+g_s f,G_s\rangle\,ds. \tag{2.25}$$

(b) *$(G_t)$ is the unique solution of* (2.25) *in the class of* $(K_t)$ *satisfying* (2.5).

We will now obtain similar equations for the normalized measures $N_t$.

**Lemma 2.6.** *$N_t$ satisfies the following equation:*

$$\langle f,N_t\rangle=\langle f,N_0\rangle+\int_0^t \langle Lf+g_s f,N_s\rangle\,ds-\int_0^t \langle f,N_s\rangle\langle g_s,N_s\rangle\,ds. \tag{2.26}$$

*Proof.* $\langle f,N_t\rangle=\dfrac{\langle f,G_t\rangle}{\langle 1,G_t\rangle}$ and thus $\langle f,N_t\rangle$ is absolutely continuous in view of

Theorem 2.5. Further,

$$\frac{d}{dt}\langle f, N_t\rangle = \frac{\langle Lf + g_t f, G_t\rangle}{\langle 1, G_t\rangle} + \langle f, G_t\rangle \left[\frac{-1}{\langle 1, G_t\rangle^2}\cdot\langle g_t, G_t\rangle\right]$$
$$= \langle Lf + g_t f, N_t\rangle - \langle f, N_t\rangle\langle g_t, N_t\rangle.$$

Hence (2.26) holds.

**Lemma 2.7.** *Let* $(K_t)$ *satisfy* (2.6). *Then* $(K_t)$ *satisfies* (2.27) *iff* $(K_t)$ *satisfies* (2.28):

$$\langle f, K_t\rangle = \langle f, K_0\rangle + \int_0^t \langle Lf + g_s f, K_s\rangle\, ds - \int_0^t \langle f, K_s\rangle\langle g_s, K_s\rangle\, ds \tag{2.27}$$

$$\langle f, K_t\rangle = \langle P_t f, K_0\rangle + \int_0^t \langle g_s P_{t-s} f, K_s\rangle\, ds - \int_0^t \langle g_s, K_s\rangle\langle P_{t-s} f, K_s\rangle\, ds. \tag{2.28}$$

*Proof.* The proof of this lemma is similar to that of Lemma 2.4. Let $\beta_t(f) = \langle f, K_t\rangle$ and let $\delta_t(f)$ be the difference in the RHS of (2.27) and (2.28). As in Lemma 2.3, using the identity $P_t f = f + \int_0^t P_s Lf$ and Fubini's theorem, we have

$$\begin{aligned}\delta_t(f) &= \int_0^t \beta_0(P_s Lf)\, ds + \int_0^t \int_0^{t-s} \beta_s(g_s P_\tau Lf)\, d\tau\, ds \\ &\quad - \int_0^t \int_0^{t-s} \beta_s(g_s)\, \beta_s(P_\tau Lf)\, d\tau\, ds - \int_0^t \beta_s(Lf)\, ds \\ &= \int_0^t \beta_0(P_s Lf)\, ds + \int_0^t \int_0^{t-\tau} \beta_s(g_s P_\tau Lf)\, ds\, d\tau \\ &\quad - \int_0^t \int_0^{t-\tau} \beta_s(g_s)\, \beta_s(P_\tau Lf)\, ds\, d\tau - \int_0^t \beta_s(Lf)\, ds.\end{aligned} \tag{2.29}$$

Now suppose $(K_t)$ satisfies (2.28). Then for $f \in \mathscr{D}$, apply (2.28) to the function $Lf$ and integrating over $[0, t]$, we get that the RHS of (2.29) is zero, so that $\delta_t(f) = 0$ and hence $(K_t)$ satisfies (2.27).

On the other hand, if $(K_t)$ satisfies (2.27), we will show that $\delta_t(f) = 0$ for $f \in \mathscr{D}_2$, to complete the proof as in Lemma 2.4. So fix $f \in \mathscr{D}_2$. Apply (2.27) to $P_\tau Lf$ and rearrange terms to get

$$\begin{aligned}\int_0^{t-\tau} \beta_s(g_s P_\tau Lf)\, ds &= \beta_{t-\tau}(P_\tau Lf) - \beta_0(P_\tau Lf) - \int_0^{t-\tau} \beta_s(P_\tau L Lf)\, ds \\ &\quad + \int_0^{t-\tau} \beta_s(g_s)\, \beta_s(P_\tau Lf)\, ds.\end{aligned} \tag{2.30}$$

Also, using the identity $P_{t-s} Lf = Lf + \int_0^t P_\tau L Lf\, d\tau$, we have

$$(2.31)\qquad \int_0^t \beta_{t-\tau}(P_\tau Lf)\,d\tau = \int_0^t \beta_s(P_{t-s}Lf)\,ds$$
$$= \int_0^t \beta_s(Lf)\,ds + \int_0^t \int_0^{t-s} \beta_s(P_\tau L Lf)\,d\tau\,ds$$
$$= \int_0^t \beta_s(Lf)\,ds + \int_0^t \int_0^{t-\tau} \beta_s(P_\tau L Lf)\,ds\,d\tau.$$

Now, (2.29), (2.30) and (2.31) imply that $\delta_t(f)=0$. As remarked earlier, this implies that $(K_t)$ satisfies (2.28).

**Theorem 2.8.** (a) $(N_t)$ *satisfies*

$$(2.32)\qquad \langle f, N_t\rangle = \langle f, N_0\rangle + \int_0^t \langle Lf + g_s f, N_s\rangle\,ds - \int_0^t \langle f, N_s\rangle \langle g_s, N_s\rangle\,ds, \qquad \forall f \in \mathscr{D}.$$

(b) $(N_t)$ *satisfies*

$$(2.33)\qquad \langle f, N_t\rangle = \langle P_t f, N_0\rangle + \int_0^t \langle g_s P_{t-s} f, N_s\rangle\,ds - \int_0^t \langle g_s, N_s\rangle \langle P_{t-s} f, N_s\rangle\,ds, \qquad \forall f \in \mathscr{B}(E).$$

(c) (i) $(N_t)$ *is the unique solution of* (2.32) *in the class of* $(K_t)$ *satisfying* (2.6).
(ii) $(N_t)$ *is the unique solution of* (2.33) *in the class of* $(K_t)$ *satisfying* (2.6).
(d) *Define* $N_t^n$ *inductively by* $\langle f, N_t^0\rangle = Ef(Z_t)$ *and for* $n \geqq 0$

$$\langle f, N_t^{n+1}\rangle = \langle P_t f, N_0^0\rangle + \int_0^t \langle g_s P_{t-s} f, N_s^n\rangle\,ds - \int_0^t \langle g_s, N_s^n\rangle \langle P_{t-s} f, N_s^n\rangle\,ds$$

*Then* $N_t^n \to N_t$ *uniformly in the total variation norm.*

*Proof.* (a) and (b) follow from Lemmas 2.6 and 2.7. In (c) we will prove (ii). Then (i) follows from Lemma 2.7 and (ii).

Let $(K_t)$ satisfy (2.6) and (2.33). Then

$$(2.35)\qquad \langle f, N_t - K_t\rangle = \int_0^t \langle g_s P_{t-s} f, N_s - K_s\rangle\,ds - \int_0^t \langle g_s, N_s - K_s\rangle \cdot \langle P_{t-s} f, N_s\rangle\,ds + \int_0^t \langle g_s, K_s\rangle \langle P_{t-s} f, N_s - K_s\rangle\,ds.$$

Thus if $a(t) = \sup_{s \leqq t} \sup_{\substack{f \in \mathscr{B}(E) \\ \|f\| \leqq 1}} |\langle f, N_s - K_s\rangle|$, then from (2.35) we have, for a suitable constant $C_1$, $a(t) \leqq C_1 \int_0^t \alpha(s)\, a(s)\,ds$ and hence, from Lemma 2.2, $a(t)=0$. This proves (c). (d) is proved by proceeding as in the uniqueness proof above and in Theorem 2.3.

## 3. Equations of Nonlinear Filtering; Existence and Uniqueness of Solutions and Robust Filtering

We now use the results of the previous section to obtain equations for the measures $\Gamma_t^y$ and $F_t^y$ related to the nonlinear filtering problem introduced in Sect. 1. Recall that $F_t^y$ is the conditional distribution of $X_t$ given $\{y_s: 0 \leqq s \leqq t\}$ in the model (1.9). ($\Gamma_t^y$ is the unnormalized conditional distribution.) We will show that for all $y \in H$, $\Gamma_t^y$ and $F_t^y$ can be obtained by the method of successive approximation. Finally, we will show that the maps $y \to \Gamma_t^y$ and $y \to F_t^y$ are continuous (with the total variation norm on the range). Thus, $\Gamma_t^y$ and $F_t^y$ are robust filters.

Let $E = [0, T] \times S$, $\mathscr{E} = \mathscr{B}_{[0,T]} \otimes \mathscr{S}$. Let $Z_t = (t, X_t)$. Then $Z_t$ is a homogeneous Markov process (see Dynkin [4], p. 167). Assume that $(X_t)$ is such that the process $Z_t$ satisfies the condition (2.1) imposed in Sect. 2.

Let $Q_t^s$ be the two-parameter non-homogeneous semigroup on $\mathscr{B}(S)$ defined formally by the relation

$$(Q_t^s f)(x) = E(f(X_t) | X_s = x).$$

If $P_t$ denotes, as in Section 2, the semigroup associated with $(Z_t)$, then we have

$$(3.1) \qquad (P_t f)(s, x) = [Q_{s+t}^s f(s+t, \cdot)](x).$$

Assume that $h$ satisfies

$$(3.2) \qquad \|h_s(x)\| \leqq q(s) \quad \forall x \in \mathscr{S}, \quad \int_0^T q^2(s)\, ds < \infty.$$

Condition (3.2) will be assumed to hold throughout the rest of this section.

**Theorem 3.1.** (a) *$\Gamma_t^y$ satisfies the following equation*

$$(3.4) \qquad \langle f, \Gamma_t^y \rangle = \langle Q_t^0 f, \Gamma_0^y \rangle + \int_0^t \langle [(h_s, y_s) - \tfrac{1}{2} \|h_s\|^2] Q_t^s f, \Gamma_s^y \rangle \, ds.$$

(b) *$\Gamma_t^y$ is the unique solution of* (3.4) *in the class of measures* $(K_t)$ *satisfying* (3.5).

(3.5) *For $B \in \mathscr{S}$, $K_0(B) = E\, 1_B(X_0)$, $K_t(B)$ is a bounded Borel measurable function of $t$.*

(c) *Define $\Gamma_{t,n}^y$ inductively as follows: $\Gamma_{t,0}^y(B) = E\, 1_B(X_t)$, and for $n \geqq 0$, $\langle f, \Gamma_{t,n+1}^y \rangle$ is defined by the right hand side of* (3.4), *with $\Gamma_{t,n}^y$ replacing $\Gamma_t^y$. Then $\Gamma_{t,n}^y$ converges uniformly (in $t$) to $\Gamma_t^y$ in the total variation norm.*

(d) *If $y_n \to y$ in $H$, then $\Gamma_t^{y_n} \to \Gamma_t^y$ in the total variation norm.*

*Proof.* Fix $y \in H$. Define $g$: $[0, T] \times E \to \mathbb{R}$ by

$$(3.6) \qquad g_s(t, x) = (h_s(x), y_s) - \tfrac{1}{2} \|h_s(x)\|^2.$$

We will also write $g_s(x) = g_s(s, x)$.

Observe that

$$|g_s(t,x)| \leqq (q(s)\cdot \|y_s\| + \tfrac{1}{2}q^2(s)) = \alpha(s), \qquad \int_0^T \alpha(s)\,ds < \infty$$

and

$$\int_0^u g_s(t,x)\,ds \leqq \tfrac{1}{2}\int_0^u \|y_s\|^2\,ds \leqq \|y\|_H^2.$$

Now for $f \in \mathscr{B}(E)$,

$$\begin{aligned}\langle f, G_t\rangle &= E\left(f(Z_t)\exp\int_0^t g_s(Z_s)\,ds\right)\\ &= Ef(t,X_t)\exp\left(\int_0^t [(h_s(X_s), y_s) - \tfrac{1}{2}\|h_s(X_s)\|^2]\,ds\right)\\ &= \langle f(t,\cdot), \Gamma_t^y\rangle,\end{aligned}$$

and hence we have

$$\text{(3.7)} \qquad \langle f, G_t\rangle = \langle f(t,\cdot), \Gamma_t^y\rangle \quad \text{for } f \in \mathscr{B}(E).$$

In view of (3.1) and (3.7), the Eq. (2.7) reduces to

$$\text{(3.8)} \qquad \begin{aligned}\langle f(t,\cdot), \Gamma_t^y\rangle &= \langle Q_t^0 f(t,\cdot), \Gamma_0^y\rangle\\ &\quad + \int_0^t \langle g_s(x)(Q_t^s f(t,\cdot)(x), \Gamma_s^y\rangle\,ds, \qquad f \in \mathscr{B}(E).\end{aligned}$$

Now, (3.8) holds iff (3.4) holds and thus (a) and (b) follows from Theorem 2.3. It is easily seen that

$$\langle f, G_t^n\rangle = \langle f(t,\cdot), \Gamma_{t,n}^y\rangle, \qquad f \in \mathscr{B}(E)$$

and thus (c) also follows from Theorem 2.3.

For (d), let $g_n$ be defined by (3.6) for $y = y_n$. Then we have, since $y_n \to y$ in $H$,

$$\int_0^t g_{n,s}(s, X_s)\,ds \to \int_0^t g_s(s, X_s)\,ds \qquad \text{pointwise}$$

and for some constant $C$,

$$\int_0^t g_{n,s}(s, X_s)\,ds \leqq \tfrac{1}{2}\int_0^t \|y_n(s)\|^2\,ds \leqq \|y_n\|_H^2 \leqq C.$$

Thus

$$\exp\left(\int_0^t g_{n,s}(s, X_s)\,ds\right) \to \exp\left(\int_0^t g_s(s, X_s)\,ds\right) \qquad \text{in } L^1(\Omega, \mathscr{A}, \Pi)$$

and hence $\Gamma_t^{y_n} \to \Gamma_t^y$ in total variation norm.

**Theorem 3.2.** (a) $\Gamma_t^y$ *satisfies the following equation*

$$(3.9)\qquad \langle f(t,\cdot),\Gamma_t^y\rangle=\langle f(t,\cdot),\Gamma_0^y\rangle+\int_0^t \langle (Lf)(s,\cdot),\Gamma_s^y\rangle\,ds$$
$$+\int_0^t \langle [(h_s,y_s)-\tfrac{1}{2}\|h_s\|^2]\,f(s,\cdot),\Gamma_s^y\rangle\,ds,\qquad f\in\mathscr{D}.$$

(b) $(\Gamma_t^y)$ *is the unique solution of* (3.9) *in the class of measures* $(K_t)$ *satisfying* (3.5).

*Remark.* Recall that $L$ is the extended generator of $(t,X_t)$ and $\mathscr{D}$ is its domain.

*Proof.* In view of (3.7), the Eq. (2.25) reduces to (3.9) and hence (a) and (b) follow from Theorem 2.5.

The following theorems on the conditional distribution $F_t^y$ can be deduced from Theorem 2.7 and

$$(3.10)\qquad \langle f,N_t\rangle=\langle f(t,\cdot),F_t^y\rangle,$$

(where for a fixed $y$, $g$ is defined by (3.6)) in the same way as Theorems 3.1 and 3.2 were deduced from Theorems 2.3 and 2.5. (Part (d) of Theorem 3.3 follows from (d), Theorem 3.1.)

**Theorem 3.3.** (a) $F_t^y$ *satisfies*

$$(3.11)\qquad \langle f,F_t^y\rangle=\langle Q_t^0 f,F_0^y\rangle+\int_0^t \langle [(h_s,y_s)-\tfrac{1}{2}\|h_s\|^2]\,Q_t^s f,F_s^y\rangle\,ds$$
$$-\int_0^t \langle (h_s,y_s)-\tfrac{1}{2}\|h_s\|^2,\,F_s^y\rangle\langle Q_t^s f,F_s^y\rangle\,ds,\qquad f\in\mathscr{B}(S)$$

(b) $F_t^y$ *is the unique solution of* (3.11) *in the class of probability measures* $(K_t)$ *satisfying* (3.12):

(3.12) $K_0(B)=E\,1_B(X_0)$ *and* $K_t(B)$ *is a Borel measurable function of* $t$ *for all* $B\in\mathscr{S}$.

(c) *Define* $F_{t,n}^y$ *inductively by* $F_{t,0}^y(B)=E\,1_B(X_t)$ *and for* $n\geqq 0$, $\langle f,F_{t,n+1}^y\rangle$ *is defined by the right hand side of* (3.11), *with* $F_{t,n}^y$ *replacing* $F_t^y$. *Then* $F_{t,n}^y$ *converges uniformly (in* $t$*) to* $F_t^y$ *in the total variation norm.*

(d) *If* $y_n\to y$ *in* $H$, *then* $F_t^{y_n}\to F_t^y$ *in total variation norm.*

**Theorem 3.4.** (a) $F_t^y$ satisfies

$$(3.13)\qquad \langle f(t,\cdot),F_t^y\rangle=\langle f(0,\cdot),F_0^y\rangle+\int_0^t \langle (Lf)(s,\cdot),F_s^y\rangle\,ds$$
$$+\int_0^t \langle [(h_s,y_s)-\tfrac{1}{2}\|h_s\|^2]\,f(s,\cdot),F_s^y\rangle\,ds$$
$$-\int_0^t \langle [(h_s,y_s)-\tfrac{1}{2}\|h_s\|^2],\,F_s^y\rangle\langle f(s,\cdot),F_s^y\rangle\,ds,\qquad f\in\mathscr{D}.$$

(b) $F_t^y$ is the unique solution of (3.13) in the class of probability measures $(K_t)$ satisfying (3.12).

*Remark 1.* Theorem 3.2 is concerned with the Zakai equation of the unnormalized conditional measure $\Gamma_t^y$ while Theorem 3.1 is the corresponding Kunita type equation. The FKK equation for the conditional probability measure $F_t^y$ and its Kunita equation are studied in Theorems 3.4 and 3.3 respectively. All four of these equations are, of course, in the white noise set up.

*Remark 2.* If $(X_t)$ is a homogenous Markov process, it suffices to consider functions of '$x$' only (i.e., functions on $S$) in Theorems 3.2 and 3.4. Then $L$ will be the generator of $(X_t)$. In this case, in Theorems 3.1 and 3.3, $Q_t^s$ is replaced by $Q_{t-s}$, $Q_r$ being the one-parameter semigroup associated with $(X_t)$.

*Remark 3.* If $(X_t)$ is an $\mathbb{R}^d$-valued diffusion, then $L=\left(\frac{\partial}{\partial t}+\mathscr{L}_t\right)$ on $C^{1,2}([0,T] \times \mathbb{R}^d)$, where for each $t$, $\mathscr{L}_t$ is a second order differential operator of the form $\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2}{\partial x^i \partial x^j}+\sum_{i=1}^{d} b_i(t,x)\frac{\partial}{\partial x^i}$. Further, in this case, it suffices to demand that (3.9) and (3.13) hold for functions of $x$ only (i.e., $f\in C^2(\mathbb{R}^d)$).

## References


1. Balakrishnan, A.V.: Non-linear white noise theory, Multivariate Analysis V, P.R. Krishnaiah, Ed. Amsterdam: North Holland 1980
2. Clark, J.M.C.: The design of robust approximations to the stochastic differential equations in nonlinear filtering, in: Communications Systems and Random Process Theory. NATO Advanced Study Institute Series, Alphen aan den Rijn: Sijthoff and Noordhoff 1978
3. Davis, M.H.A.: Pathwise solutions and multiplicative functionals in non-linear filtering. 18th IEEE Conference on Decision and Control, Fort Lauderdale, Florida (1979)
4. Dynkin, E.B.: Markov processes, Vol. 1. Berlin-Heidelberg-New York: Springer 1965
5. Fujisaki, M., Kallianpur, G., Kunita, H.: Stochastic differential equations for the nonlinear filtering problem. Osaka J. Math., **1**, 19–40 (1972)
6. Gross, L.: Measurable functions on Hilbert space. Trans. Amer. Math. Soc., **105**, 372–390 (1962)
7. Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes. Amsterdam: North Holland 1981
8. Kallianpur, G., Karandikar, R.L.: A finitely additive white noise approach to nonlinear filtering. Appl. Math. Optim., **10**, 159–185 (1983)
9. Krylov, N.V., Rozovskii, B.L.: Stochastic evolution equations. J. Soviet Math. **16**, 1233–1276 (1981). (English translation)
10. Kunita, H.: Asymptotic behaviour of the nonlinear filtering errors of Markov processes. J. Multivariate. Anal. **1**, 365–393 (1971)
11. Pardoux, E.: Stochastic partial differential equations and filtering of diffusion processes. Stochastics, **2**, 127–168 (1979)
12. Segal, I.E.: Tensor algebras over Hilbert spaces. Trans. Amer. Math. Soc., **81**, 106–134 (1956)
13. Szpirglass, J.: Sur l'equivalence d'equations differentielles stochastiques à valeurs measures intervenant dans le filtrage Markovien non linèaire. Ann. Inst. Henri Poincare, Section B, **XIV**, 33–59 (1978)
14. Zakai, M.: On the optimal filtering of diffusion processes. Z. Wahrscheinlichkeitstheorie verw. Geb., **11**, 230–243 (1969)


Received January 15, 1983