

Paper Entitled  
"Filtering and Smoothing Equations for the  
Filtering Problems of Benes"

From the Proceedings of  
20th IEEE Conference on  
Decision and Control

pp. 83-88

San Diego, California  
December 1981

## FILTERING AND SMOOTHING EQUATIONS FOR THE FILTERING PROBLEM OF BENEŠ

D. L. Ocone<sup>1</sup>

Mathematics Research Center  
and Dept. of Mathematics  
University of Wisconsin-Madison  
Madison, WI 53706

J. S. Baras<sup>2</sup>

Dept. of Electrical Engineering  
University of Maryland  
College Park, MD 20742

S. I. Marcus<sup>3</sup>

Dept. of Electrical Engineering  
University of Texas at Austin  
Austin, TX 78712

### Summary

Filtering of a diffusion with nonlinear drifts of a type defined by Beneš is studied. Explicit, finite dimensional, recursive filters are shown to exist for moments and polynomial functionals of the diffusion, and a smoothing formula is given. These results are related to current, geometric approaches to filtering.

### 1. Introduction

Let  $f(x)$  be a solution to the Riccati equation

$$f'(x) + f^2(x) = ax^2 + bx + c \quad (1)$$

and assume  $f(x)$  is defined for all  $x \in \mathbb{R}$  and has no singularities. This paper studies the filtering problem

$$dx = f(x)dt + db \quad (2)$$

$$x(0) = x_0 \in \mathbb{R}$$

$$dy = xdt + dw \quad (3)$$

in which  $b(\cdot)$  and  $w(\cdot)$  are independent Wiener processes,  $x(\cdot)$  is the signal, and  $y(\cdot)$  the observation of  $x(\cdot)$ . For this problem, Beneš [1] recently derived an explicit formula for the conditional density of  $x(t)$  given  $F(t) := \sigma$ -algebra generated by  $\{y(s), 0 \leq s \leq t\}$ . This result generated considerable interest because the class

of drifts satisfying (1) includes nonlinear examples, but conditional densities had been computed previously in (2)-(3) only for linear drifts. Aside from conditional densities, however, one also desires to compute explicit filters for various conditional statistics of the signal. Here we shall interpret 'explicit' in terms of finite dimensional computability; given a random process  $\varphi(t)$ , we say that  $\hat{\varphi}(t) := E\{\varphi(t) | F(t)\}$  is finite dimensionally computable (or FDC) if it can be expressed as the output of a finite dimensional system of stochastic differential equations driven by  $y(\cdot)$ .

In this paper, we extend the analysis of (1)-(3) by finding for it conditional statistics that are FDC. The main result, theorem 1 in section 2, lists a number of important FDC statistics. Finite dimensional computability is well known for these statistics when  $f$  is linear, and hence theorem 1 generalizes the linear theory to the full Beneš class (1)-(3). Section 2 of the paper motivates theorem 1 and sets it in the context of current geometric/algebraic theories of filtering. Sections 3 and 4 are devoted to the proof. Section 3 computes expressions for the conditional joint density of  $(x(t), x(s_1), \dots, x(s_n))$ ,  $t > s_1 > \dots > s_n$ , given  $F(t)$  and the conditional joint density of  $(x(s_1), \dots, x(s_n))$  given  $F(t)$  and  $x(t)$ . (These results are of some interest in their own right.) Section 4 formulates the proof of theorem 1 on the basis of the conditional densities and also presents a smoothing formula. We will omit or only sketch some arguments; details may be found in Ocone, Baras and Marcus [11].

1. Supported in part by NSF under Grant No. MCS-7927062
2. Supported by the Office of Naval Research under contract N00014-79-C-0808
3. Supported in part by Joint Services Electronics Program under Contract F49620-77-C-0101

### 2. Estimation Algebras and Finite Dimensional Computability

The FDC statistics listed in theorem 1 were suggested by a study of the linear case ( $f(x)=hx$ ) and its connection to (1)-(3) in the Lie algebraic

approach to filtering. (We assume some familiarity with geometric/algebraic filtering theory in the discussion below; an account of the theory may be found in Brockett [2]. However, the proof of theorem 1 does not require geometric ideas.)

When  $f$  is linear the Kalman-Bucy equations compute the conditional mean  $\hat{x}(t) = E\{x(t)|F(t)\}$  and conditional covariance

$K(t) = E\{x(t) - \hat{x}(t)\}^2 | F(t)$  thus proving that  $\hat{x}(t)$  and  $K(t)$  are FDC. All moments  $x^n(t)$  are also FDC, because the conditional density  $p(x, t | F(t))$  of  $x(t)$  given  $F(t)$  is normal and hence  $E\{(x(t) - \hat{x}(t))^{2n} | F(t)\} = K^n(t)$  and  $E\{(x(t) - \hat{x}(t))^{2n+1} | F(t)\} = 0$  for all  $n$ .

**Remark**

Because  $p(x, t | F_t^Y)$  is normal, it is specified by  $\hat{x}(t)$  and  $K(t)$ , that is, we can write  $p(x, t | F_t^Y) = q(\hat{x}(t), K(t))(x)$ . Let  $\varphi$  be a function such that  $c(\alpha, \beta) = \int \varphi(x) q(\alpha, \beta)(x) dx$  makes sense for all  $\alpha$  and  $\beta$ . Then  $\hat{\varphi}(t) = E\{\varphi(x(t)) | F(t)\} = c(\hat{x}(t), K(t))$  and hence  $\hat{\varphi}$  is FDC. This is another way of showing  $x^n(t)$  is FDC for all  $n$ .

These results may be derived from the Lie algebraic structure of the linear problem. The estimation algebra, which is the Lie algebra generated by the operators in Zakai's equation for the unnormalized conditional density, is the finite dimensional algebra ( $id = identity$ )

$$L_0 = \text{Span}\{1/2 \partial^2/\partial x^2 - 1/2 x^2, x, \partial/\partial x, id\}, \quad (4)$$

$p(x, t | F_t^Y)$  and its parameterization in terms of  $\hat{x}(t)$  and  $K(t)$  can be found from Zakai's equation and the structure of  $L_0$  by a Wei-Norman technique (see Ocone [9]).

Let  $\Lambda$  denote the class of polynomial functionals of  $x(\cdot)$  of the form

$$\eta(t) = \int_0^t \dots \int_0^{s_{r-1}} \gamma(s_1, \dots, s_r) x^{k_1}(s_1) \dots x^{k_r}(s_r) ds_r \dots ds_1$$

where  $r$  is any integer,  $k_1, \dots, k_r$  are non-negative integers, and  $\gamma$  is a bounded, separable function. Marcus and Willsky [7] showed that if  $f$  is linear and  $\eta(\cdot) \in \Lambda$ , then  $\hat{\eta}(t)$  is FDC. Again this result is suggested by a Lie algebraic analysis. In the geometric approach, an FDC statistic is associated with a homomorphism from the estimation algebra to the vector field Lie algebra of the system computing the statistic. Given an ideal  $I$  in the estimation algebra, one may then try to find an FDC statistic inducing a homomorphism with kernel  $I$ ; one methodology for this procedure is suggested in Ocone [10]. As an example, consider the system

$$\begin{aligned} dy &= db & x(0) &= 0 \\ dz &= x^2 dt & z(0) &= 0 \\ dy &= xdt + dw & y(0) &= 0 \end{aligned} \quad (5)$$

The estimation algebra  $\tilde{L}_0 = \{-x^2 \partial/\partial z + 1/2(\partial^2/\partial x^2 - x^2), x\}_{LA}$  has an infinite sequence of ideals

$$I_j = \text{Span}\{id, x \partial^i/\partial z^i, \partial/\partial x \partial^i/\partial z^i, \partial^i/\partial z^i | i \geq j\},$$

and it can be shown that  $\hat{z}^n(t)$  for  $n < j$  are likely candidates for FDC statistics associated to  $I_j$  ([10]). Indeed,  $\hat{z}^n(t)$  are FDC by the Marcus-Willsky theorem, since  $\hat{z}^n(\cdot) \in \Lambda$  for all  $n$ .

Finite dimensional systems for  $\hat{z}^n(t)$  and associated Lie algebra homomorphisms from  $\tilde{L}_0$  to  $I_j$ ,  $j > n$ , have been constructed by Liu and Marcus [5] for low order  $j$  and  $n$ .

This analysis has important implications for the case in which  $f$  can be any, possibly nonlinear, solution of (1), because then the estimation algebra of (2)-(3), defined as  $L_1 = \text{Lie algebra generated by } 1/2(\partial^2/\partial x^2 - x^2) - \partial/\partial x f(x)$  and  $x$ , is isomorphic to the linear estimation algebra  $L$  at (4) by the mapping

$$\phi(A) = e^{F(x)} A e^{-F(x)} : L_1 \rightarrow L, \quad F(x) = \int^x f(s) ds.$$

The filtering problem (1)-(3) is therefore equivalent in structure to the linear problem (Brockett [3], Mitter [8]). Likewise, consider the analogue of (5):

$$\begin{aligned} dx &= f(x)dt + db \\ dz &= x^2 dt \\ dy &= xdt + dw \\ x(0) &= x, y(0) = z(0) = 0 \end{aligned} \quad (6)$$

and let  $\tilde{L}_1$  denote its estimation algebra. Again, if  $f$  solves (1)  $\phi$  establishes an isomorphism between  $\tilde{L}_1$  and  $L_1$  of (5), thus implying that  $\tilde{L}_1$  is identical in structure to  $L_1$ , and, in particular, possesses the sequence of ideals  $\{\tilde{I}_j = \phi^{-1}(I_j) | j > 0\}$  which should function like  $\{I_j | j > 0\}$ . For the problem (1)-(3) one then expects, just as in the linear case, that a) the conditional density  $p(x, t | F(t))$  can be computed explicitly in terms of a finite number of recursively generated statistics; and b)  $\hat{x}^n(t)$ ,  $\hat{z}^n(t)$ , (in (6)), and, more generally,  $\hat{\eta}(t)$ ,  $\eta \in \Lambda$ , are all FDC. Statement a) is just the previously mentioned result of Beneš [1] and can be recovered easily from the Gaussian, linear solution by using  $\phi$  in Zakai's equation (Mitter [8]). Statement b) has not yet been treated. We will show here that it is in fact true and so restate it as

Theorem 1

If the drift  $f$  in (2)-(3) satisfies (1)

- i)  $\hat{x}^n(t)$  is FDC for every positive integer  $n$
- ii)  $\hat{\eta}(t)$  is FDC for every  $\eta(\cdot) \in \Lambda$ .

Theorem 1 thus generalizes linear filtering results and shows that the Lie algebraic interpretation of linear filtering extends successfully to (1)-(3). However, because of the generality of (ii), convenience and rigor dictate a proof based not upon geometric techniques but upon calculating joint conditional densities. This is sketched in sections 3 and 4.

3. Conditional Joint Densities

We assume given the filtering problem (1)-(3). Let  $t = s_0 > s_1 > \dots > s_n \geq 0$ ,  $\bar{s} = (s_0, s_1, \dots, s_n)$ , and  $\bar{z} = (z_0, z_1, \dots, z_n)^T$ . Further, let  $p(\bar{z}, \bar{s} | F(t))$  denote the conditional joint density of  $(x(t), s(s_1), \dots, x(s_n))$  conditioned on  $F(t)$  for any bounded, Borel  $\psi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ ,

$$E\{\psi(x(t), \dots, x(s_n)) | F(t)\} = \int \psi(\bar{z}) p(\bar{z}, \bar{s} | F(t)) d\bar{z}.$$

Our object is to compute  $p(\bar{z}, \bar{s} | F(t))$ .

The formula is stated in terms of an auxiliary process  $\xi(t)$ , evolving in  $\mathbb{R}^3$  and defined by

$$d\xi = A(y)\xi dt + [1 \ y(t) \ 0]^T dB$$

$$\xi(0) = (x, 0, 0)^T$$

where

$$A(y)(t) = \begin{bmatrix} -k & 0 & 0 \\ 0 & & \\ ky(t) - 1/2b & 0 & 0 \end{bmatrix}$$

$$k = (a+1)^{1/2}$$

and  $B$  is a Brownian motion independent of the signal and observation noises  $b(\cdot)$  and  $w(\cdot)$ . Let

$$\Xi_0 = (\xi^T(t), \dots, \xi^T(s_n))^T$$

$$\Xi = (\xi(t), \xi_1(s), \dots, \xi_1(s_n))^T$$

We need the following statistics:

$$m(t) = E\{\xi(t) | F(t)\}$$

$$R(t, s) = \text{Cov}(\xi(t), \xi(s) | F(t))$$

$$R(t) = R(t, t)$$

$$M(t, s_1, \dots, s_n) = E\{\Xi | F(t)\}$$

$$Q(t, s_1, \dots, s_n) = \text{Var}\{\Xi | F(t)\}$$

$$P(t, s_1, \dots, s_n) = \text{Cov}(\Xi, \Xi_0 | F(t))$$

$$= E\{(\Xi - M)(\Xi_0 - E\Xi_0 | F(t))^T | F(t)\}$$

(To simplify later expressions, we often drop the  $(t, s_1, \dots, s_n)$  dependence and write only  $M, Q$  or  $P$ .) These statistics are properly viewed as non-anticipating functionals on  $C[0, \infty)$  evaluated at  $y(\cdot)$ . As final pieces of notation, set  $v = (0, 1, -1, 0, \dots, 0)^T \in \mathbb{R}^{3(n+1)}$  and  $F(z_0) = \int^z f(s) ds$ .

Theorem 2

For (1)-(3),

$$p(\bar{z}, \bar{s} | F(t)) = (1/N) \exp\{F(z_0) + z_0 y(t) + k/2 z_0^2\}$$

$$\times \exp\{-1/2(\bar{z} - M + Pv, Q^{-1}(\bar{z} - M + Pv))\} \quad (7)$$

where  $N = N(t, s_1, \dots, s_n, x_0)$  is a normalization factor not depending on  $\bar{z}$ .

The demonstration of this result is analogous to Beneš's calculation in [1] of  $p(z_0, t | F(t))$ .

One performs a series of Girsanov transformations in the Kallianpur-Striebel formula of conditional estimation and obtains thereby an expression for  $p(\bar{z}, \bar{s} | F(t))$  as the conditional (on  $F(t)$ ) expectation of a function of  $\Xi$ . Since  $\xi(s), s \leq t$ , is a Gaussian process when conditioned on  $F(t)$ , this expectation is easy to evaluate. For a full proof, see [11].

From theorem 2, we see that the joint conditional density consists of a Gaussian factor multiplied by  $\exp F(z_0)$ . If we further condition on  $x(t)$  we remove the  $\exp F(z_0)$  and arrive at a normal density. To this end, let

$$\Xi^{(2)} = (\xi_1(s_1), \dots, \xi_n(s_n))^T$$

$$Q = \begin{bmatrix} r_{11}(t) & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$$

$$Pv = [(Pv)_1 \quad (Pv)^{(2)T}]^T$$

$$M = [m_1(t) \quad M^{(2)T}]^T$$

Note that  $Q_{22} = \text{Var}(\Xi^{(2)} | F(t))$ .

Corollary 3

The conditional law of  $(x(s_1), \dots, x(s_n))$

given  $F(t)$  and  $x(t)$  is normal with mean  $M^{(2)} - Pv^{(2)} + r_{11}^{-1}(t)Q_{21}(x(t)-m_1(t)+(Pv)_1)$  and variance  $Q_{22} - r_{11}^{-1}(t)Q_{21}Q_{12}$ .

The conditional normality revealed in corollary 3 is of key importance in our proof of theorem 1.

#### 4. Proof of Theorem 1

i) From theorem 2 one can recover the formula of Beneš [1]:

$$p(z, t | F(t)) = \frac{1}{N(t, x)} \exp\{F(z) - (z - \mu(t))^2 / 2\sigma(t)\}$$

in which  $N(t, x)$  is a normalization factor, and

$$\dot{\sigma} = 1 - k^2 \sigma^2 \quad \sigma(0) = 0 \quad (8)$$

$$d\mu = [-k^2 \sigma \mu - 1/2b\sigma] dt + \sigma dy \quad \mu(0) = x \quad (9)$$

( $\dot{\sigma}$  denotes the time derivative of  $\sigma$ .) It can be shown that  $\hat{x}^n(t)$  is well defined for all  $n$ . The finite dimensionally computable  $\mu$  and  $\sigma$  characterize  $p(z, t | F(t))$ , and a repetition of the argument in the remark of section 2 then establishes that  $\hat{x}^n(t)$  is FDC for every  $n$ .

Let us give a second proof that yields the filtering equations explicitly.

#### Lemma 4

For  $n \geq 0$

$$(\sigma^{-1} - a) \hat{x}^{n+2} = (b + 2\mu\sigma^{-2}) \hat{x}^{n+1} + (c + ((2n+1) - \mu^2) \sigma^{-1}) \hat{x}^n + 2\mu\sigma^{-1} \hat{x}^{n-1} - n(n-1) \hat{x}^{n-2}$$

#### Proof

Integrate

$$\int dz \exp F(z) \frac{d^2}{dz^2} (z^n \exp[-(z-\mu)^2 / 2\sigma])$$

by parts and use (1).

It may be shown from (8) that  $\sigma^{-1}(t) - a > 0$  for all  $t$ . Thus from lemma 4 it is sufficient to show that  $\hat{x}$  is FDC in order to show  $\hat{x}^n$  is FDC for any  $n$ . However,  $\hat{x}$  satisfies the stochastic differential equation

$$d\hat{x} = \hat{f}(\hat{x})dt + (\hat{x}^2 - \hat{x}^2)(dy - \hat{x}dt) \quad (10)$$

(Fujisaki, Kallianpur, Kunita [4]). Another integration by parts argument shows that

$$\hat{f}(\hat{x}) = (\mu - \hat{x})\sigma^{-1} \quad (11)$$

By using lemma 4 for  $n = 0$ , and (11), the right hand side of (10) may be written solely in terms  $\mu$ ,  $\sigma$ , and  $\hat{x}$ . Call such a version of (10), (10)'. (8), (9), and (10)' then constitute a finite dimensional system with state  $(\hat{x}, \mu, \sigma)$ .

ii) For simplicity we consider only the case

$$\eta(t) = \int_0^t \dots \int_0^{s_{n-1}} \gamma(s_1, \dots, s_n) x(s_1) \dots x(s_n) ds_1 \dots ds_{n-1}$$

The following identity, stated here in the form it appears in Marcus and Willsky [7], is fundamental.

#### Lemma 5

Let  $(u_1, \dots, u_m)$  be a normal random vector with  $e_j = Eu_j$  and  $V_{ij} = \text{cov}(u_i, u_j)$ . Then

$$E[u_1 \dots u_m] = \prod_{j=1}^m e_j + \sum_{j_1, j_2} V_{j_1 j_2} e_{j_1} \dots e_{j_2} + \dots + \sum_{j_1, j_2, j_3, j_4} V_{j_1 j_2} V_{j_3 j_4} e_{j_1} \dots e_{j_4} + \dots$$

The sums are taken over all possible combinations of pairs of indices.

#### Lemma 6

Let  $t \geq s_1 > s_2 > \dots > s_n$ . There exist functions  $a_j(t, s_1, \dots, s_n; y(\cdot))$ ,  $0 \leq j \leq n$ , such that

$$E\{x(s_1) \dots x(s_n) | F(t), x(t)\} = \sum_{j=0}^n x^j(t) a_j(t, s_1, \dots, s_n)$$

and each  $a_j$  has the form

$$a_j(t, s_1, \dots, s_n) = \sum_i \beta_{j,1}^i(t) \dots \beta_{n,n}^i(s_n)$$

in which each  $\beta_{j,m}^i$ ,  $0 \leq m \leq n$  is either deterministic or a non-anticipating FDC functional of  $y(\cdot)$ .

#### Proof

The conditional distribution of  $(x(s_1), \dots, x(s_n))$  given  $x(t)$  and  $F(t)$  is normal by corollary 3. Let  $(l_1, \dots, l_j)^T$  be the conditional mean given in corollary 3 and  $\hat{Q}$  the conditional variance. Then from lemma 5

$$\begin{aligned}
E\{x(s_1) \dots x(s_n) | F(t), x(t)\} &= \ell_1 \dots \ell_n \\
&+ \Sigma \tilde{Q}_{j_1 j_2} \ell_{j_3} \dots \ell_{j_n} \quad (12) \\
&+ \Sigma \Sigma \tilde{Q}_{j_1 j_2} \tilde{Q}_{j_3 j_4} \ell_{j_5} \dots \ell_{j_n} + \dots
\end{aligned}$$

Since  $(\ell_1, \dots, \ell_j)$  is a linear function of  $x(t)$ , (12) is a polynomial in  $x(t)$  of order  $n$ . Moreover, it is clear that the coefficient  $a_j$  of  $x^j(t)$  will be separable if  $Q(t, s_1, \dots, s_n)$  and hence  $\tilde{Q}$  are separable. But for each  $i, j$

$$Q_{ij} = \text{Var}(\xi_1(s_{i-1}), \xi_1(s_{j-1})),$$

which is separable since  $\xi_1$  solves a linear equation. Note also that  $Q_{ij}$  is deterministic for all  $i, j$ . Thus the  $y$ -dependent contributions in  $a_j$  come from the  $y$ -dependent terms  $(Pv)^{(2)}_j$  and  $(Pv)_1$  in  $\ell_i$   $1 \leq i \leq n$ . Now a typical element of  $Pv$  is of the form

$$\begin{aligned}
\text{cov}(\xi_1(s), \xi_2(t) - \xi_3(t)) \\
= \text{cov}(\xi_1(s), \xi_2(s) - \xi_3(s)) \\
= k^{-1} \sinh ks \int_0^t e^{-ku} g(u) du \\
- k^{-1} \sinh ks \int_0^s e^{-ku} g(u) du
\end{aligned}$$

in which  $s < t$  and  $g(u) = ky(u) - (1/2)b$ . Since the individual terms in this expression are FDC, we may conclude, after piecing everything together, that the  $y$ -dependent components  $\beta_{j,m}^i$  are non-anticipating, FDC functionals of  $y(\cdot)$ .

We are now ready to prove that  $\hat{\eta}(t)$  is FDC. By using lemma 6

$$\begin{aligned}
\hat{\eta}(t) &= E\{E\{\eta(t) | F(t), x(t)\} | F(t)\} \\
&= \sum_{j=0}^n \hat{x}^j(t) \\
&\times \int_0^t \dots \int_0^{s_{n-1}} \gamma(s_1, \dots, s_n) a_j(t, s_1, \dots, s_n) ds_n \dots ds_1 \quad (13)
\end{aligned}$$

Since  $\{\hat{x}^j(t)\}_{j=1}^n$  are FDC by part i), the proof will be complete if we show that the coefficients of  $\hat{x}^j(t)$  in (13) are FDC. However, from lemma 6, each coefficient is a sum of terms of the form

$$u_{n+1}(t) = \alpha_0(t) \int_0^t \dots \int_0^{s_n} \alpha_1(s_1) \dots \alpha_n(s_n) ds_n \dots ds_1$$

for which  $\alpha_i(s)$  is either deterministic on a deterministic function times a  $y$ -dependent  $\beta_{j,m}^i$ . Thus each  $\alpha_i$  may be assumed to be FDC. Since  $u_{n+1}(t)$  is computed by the system

$$\begin{aligned}
\dot{u}_1(t) &= \alpha_n(t) \\
\dot{u}_2(t) &= \alpha_{n-1}(t) u_1(t) \\
&\vdots \\
&\vdots \\
\dot{u}_{n+1}(t) &= \alpha_0(t) u_n(t)
\end{aligned}$$

$u_{n+1}(t)$  is FDC also. This completes the proof.

This proof is similar in its use of Gaussian moment identities to the one undertaken by Marcus and Willsky [7] to treat the  $f =$  linear function case. When  $f$  is linear, conditional normality obtains without the further conditioning on  $x(t)$ , i.e.  $(x(s_1), \dots, x(s_n))$  is conditionally normal given  $F(t)$  alone. For this reason, Marcus and Willsky are able to use the filtering equation

$$d\hat{\eta} = E\{x(t)\eta_1(t) | F(t)\} dt + [\hat{\eta}\hat{x} - \hat{\eta}\hat{x}] [dy - \hat{x}dt]$$

$$\begin{aligned}
\eta_1(t) &= \\
&= \int_0^t \dots \int_0^{s_{n-1}} \gamma(t, s_1, \dots, s_{n-1}) x(s_1) \dots x(s_{n-1}) ds_{n-1} \dots ds_1
\end{aligned}$$

and a simpler moment identity to construct a proof by induction on the order of the integral in  $\eta(t)$ . Marcus, Mitter, and Ocone [6] give a second proof of the linear case using homogeneous chaos theory. Such a proof might be possible in the general case by first conditioning on  $x(t)$ , but this appears to be more complicated and has not been tried.

Smoothing formulae for (1)-(3) may also be derived using the results of section 3.

#### Theorem 7

Let  $s < t$ . Then

$$\begin{aligned}
E\{x(s) | F(t)\} &= \frac{\sinh ks}{\sinh kt} [\hat{x}(t) - x_0 e^{-kt} + (P(t,s)v)_1] \\
&+ x_0 e^{-ks} - (P(t,s)v)_2.
\end{aligned}$$

#### Proof

This is an immediate consequence of corollary 3 since  $r_{11}^{-1}(t) Q_{21} = \sinh ks / \sinh kt$  and

$$m_1(t) = x_0 e^{-kt}.$$

### References

1. V. E. Benes, Exact finite dimensional filters for certain diffusions with nonlinear drift, Stochastics, 5 (1981), 65-92.
2. R. W. Brockett, Geometrical methods in stochastic control and estimation, Proceedings of the NATO ASI on Stochastic systems, Reidel Press, to appear.
3. R. W. Brockett, Classification and Equivalence in Estimation Theory, 1971 Conf. on Decision and Control, IEEE, N.Y., 1979.
4. M. Fujisaki, G. Kallianpur, and H. Kunita, Stochastic differential equations for the nonlinear filtering problem, Osaka J. Math. 9 (1972), 19-40.
5. C-H, Liu and S. I. Marcus, The Lie algebraic structure of a class of finite dimensional non linear filters, in Lectures in Applied Mathematics 18 (1980), 277-297.
6. S. I. Marcus, S. K. Mitter, and D. Ocone, Finite dimensional nonlinear estimation for a class of systems in continuous and discrete time, in Analysis and Optimisation of Stochastic Systems (Proc. Internat. Conf. Univ. Oxford, Oxford, 1978), 387-406, Academic Press, London, 1980.
7. S. I. Marcus and A. S. Willsky, Algebraic structure and finite dimensional non linear estimation, SIAM J. Math. Anal. 9 (1978), 312-327.
8. S. K. Mitter, On the Analogy between Mathematical Problems of Non-linear Filtering and Quantum Physics, to appear in Ricerche di Automatica, 1981.
9. D. Ocone, Topics in Non linear Filtering Theory, Doctoral Dissertation, Mathematics Department, MIT, 1980.
10. D. Ocone, Finite dimensionally computable statistics and estimation algebras in non-linear filtering, Proceedings, Inter. Symp. on Math. Theory of Networks and Systems. August 1981, Santa Monica, CA.
11. D. Ocone, J. S. Baras, S. I. Marcus, Explicit filters for diffusions with special non-linear drifts, in preparation.