Stochastics, 1981, Vol. 4, pp. 339–348
0090-9491/81/0404-0339 $06.50/0

New Results on the Innovations Problem for Non-Linear Filtering

DEBORAH F. ALLINGER

Mathematics Department, Indiana University, Bloomington, Indiana 47405, U.S.A.

and

SANJOY K. MITTER†

Department of Electrical Engineering and Computer Science and Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Mass. 02139, U.S.A.

(*Accepted for publication September 1, 1980*)

Consider an observed stochastic process consisting of a signal with additive noise. Assume that the signal has finite energy and that the signal and noise are independent. In this paper we show that under the above assumptions the innovations and observations σ-algebra are equal, thereby proving a long-standing conjecture of Kailath.

INTRODUCTION

Let $(\Omega, \mathscr{F}, \mathscr{P})$ be a complete probability space, $F = (\mathscr{F}_t)$, $0 \leqq t \leqq 1$, a non-decreasing family of sub-σ-algebras and $W = (W_t, \mathscr{F}_t)$, $0 \leqq t \leqq 1$, a Wiener process. With a signal process, $\beta = (\beta_t, \mathscr{F}_t)$, and

$$y_t = \int_0^t \beta_s \, ds + W_t \tag{1}$$

as observations, the *innovations problem* is to determine whether $y = (y_t, \mathscr{F}_t)$ is adapted to the innovations process, $(v, \mathscr{F}_t^y)$. This process, whenever it exists (see for example, [1]), is a Wiener process defined by the

†Research supported by the Air Force Office of Scientific Research under Grants AFOSR 77-3281 and 77-3281B.

equation

$$v_t = y_t - \int_0^t \hat{\beta}_s \, ds \tag{2}$$

where $\hat{\beta}_t = E(\beta_t | y_s, \ 0 \leqq s \leqq t)$. The innovations problem, first posed by Kailath in 1967 and subsequently considered by Frost in his thesis [2] can be posed in probabilistic terms; namely, are the σ-algebras generated by these processes the same modulo null sets; i.e. is

$$\sigma\{y_s : s \leqq t\} = \sigma\{v_s, s \leqq t\} \quad (\text{mod } \mathscr{P})?$$

In this paper, we show that in the form conjectured by Kailath [3], namely, if the signal and noise are independent and the signal has finite energy this problem has a positive solution.

Our results generalise all known results on the innovations problem ([4], [3]). In [4] the signal process is assumed to be uniformly bounded. The proof given in [2] is incorrect (see [3]). This problem has also been considered by Beneš [5] and Kallianpur [6] under slightly weaker hypotheses than ours. Their proofs, however, appear to be incorrect. Results similar to ours have been independently obtained by J. M. C. Clark and M. P. Ershov (private communication to S. K. Mitter, late April 1979). The case where the signal and noise are independent and the signal lies in a bounded set in L^2 was proved independently by M. P. Ershov (unpublished) and S. K. Mitter (unpublished). This result was improved by Allinger and Mitter [7] to the case where the signal and noise are independent and the signal satisfies the condition $\exists \alpha > 0$, such that $E \exp(\alpha \int_0^1 \beta_s^2 \, ds) < \infty$, and a proof was sketched by Clark (unpublished) for the case:

$$\int_0^1 E(\beta_t^2 | y_s, 0 \leqq s \leqq t) \, dt < \infty$$

which appears to be incomplete. The proof given in this paper bears similarities to that of Clark, but the martingale techniques introduced are new.

The problem considered here is a subclass of the more general innovations problem for stochastic differential equations ([8], page 260). In this more general form, the innovations problem does not have (in general) a positive solution. A counter example was given by Cirelson ([8], page 150). In Cirelson's example no "filtering" takes place and thus it cannot be considered to be a counter example to the innovations problem for non-linear filtering. Cirelson's example, however, can be modified to obtain examples where filtering does occur (cf. Beneš [9]). The proof presented in this paper utilizes the independence of the signal and noise processes in an essential way. Nevertheless, we feel that the assumption of independence can be removed for a wide class of signal processes. Indeed, using partial differential equation techniques Krylov has recently obtained results in this direction (cf. also earlier work of Beneš [10]).

2. THE INNOVATIONS RESULT

Our main result is the following:

THEOREM 1 *Consider the observation model given by* (1). *Let the signal* $\beta=(\beta_t,\mathscr{F}_t)$ *satisfy*

i) $\beta=(\beta_t,\mathscr{F}_t)$ *and* $W=(W_t,\mathscr{F}_t)$ *are independent*

ii) $E(\int_0^1\beta_s^2\,ds)<\infty$.

Then $\sigma\{y_s|0\leqq s\leqq t\}=\sigma\{v_s|0\leqq s\leqq t\}\bmod\mathscr{P}$, *where* $\sigma\{y_s|0\leqq s\leqq t\}$ (resp. $\sigma\{v_s|0\leqq s\leqq t\}$) *denotes the σ-field generated by* $\{y_s|0\leqq s\leqq t\}$ (resp. $\{v_s|0\leqq s\leqq t\}$).

The proof we give is based on the idea of pathwise uniqueness and uses a result and a remark of Yamada and Watanabe [11]. Since this does not seem to be well known (see, however, the recent book of Stroock and Varadhan [12]), we summarize these results in the Appendix for the reader's convenience.

Proof of Theorem 1 The proof consists of two parts: analyzing the Kallianpur–Striebel functional and then with its aid proving pathwise uniqueness. It is a consequence of the work of Yamada and Watanabe that a unique strong solution of (2) exists. Referring to Eq. (2) we first define $\gamma(t,x)$ as

$$\gamma(t,x)=\frac{\int\limits_\Omega\beta_t(\omega)\exp\left(\int\limits_0^t\beta_s(\omega)\,dx_s-\frac{1}{2}\int\limits_0^t\beta_s^2(\omega)\,ds\right)d\mathscr{P}(\omega)}{\int\limits_\Omega\exp\left(\int\limits_0^t\beta_s(\omega)\,dx_s-\frac{1}{2}\int\limits_0^t\beta_s^2(\omega)\,ds\right)d\mathscr{P}(\omega)}$$

for $t\in[0,1]$, $x\in C[0,1]$. This functional is finite on a set of Wiener measure one; in fact, under (i), (ii), the results of Kallianpur and Striebel [13, Thm. 3 and its corollary] apply to show that $\gamma(t,x)$ is a jointly, measurable, non-anticipative functional. Furthermore, when $\gamma(t,x)$ is evaluated at the observations, $x=y$, we have

$$\gamma(t, y(\omega)) = \dot{\hat{\beta}}_t(\omega), \lambda \times \mathscr{P}\text{-a.s.}$$

(λ denotes Lebesgue measure on $[0,1]$).

Secondly, our hypotheses (i), (ii) guarantee that the innovations can be constructed [1], and so Eq. (2) is satisfied by the observations. To show that any weak solution to (2) is pathwise unique, we need the following lemmas.

LEMMA 1 *Let*

$$\rho(t,x,\omega) = \exp\left(\int_0^t \beta_s(\omega)\,dx_s - \tfrac{1}{2}\int_0^t \beta_s^2(\omega)\,ds\right) \quad x\text{-a.s.}$$

and

$$g(t,x) = \int_\Omega \rho(t,x,\omega)\,d\mathscr{P}(\omega). \quad x\text{-a.s.}$$

Then

a) $$\mu_W\left\{x\colon \sup_{0\leqq t\leqq 1} g(t,x) < \infty\right\} = 1$$

and

b) $$\mu_W\left\{x\colon \inf_{0\leqq t\leqq 1} g(t,x) > 0\right\} = 1$$

where μ_W is Wiener measure on $C[0,1]$.

Proof Essentially, Lemma 1 is demonstrated by P. A. Meyer in his paper, "Sur un problème de filtration," *Springer-Verlag Notes* 321. The main idea is that $\{g(t, W(\omega)), \mathscr{F}_t^W\}$ is a Brownian martingale which has a *right continuous* modification [Corollary to Thm. 3.1, 8]. Thus, we obtain (a) since

$$\mathscr{P}\left\{\omega\colon \sup_{0\leqq t\leqq 1} g(t, W(\omega)) > \lambda\right\} \leqq 1/\lambda$$

for $\lambda > 0$; (b) is obtained analogously.

LEMMA 2 *Let*

$$\alpha(t,x)=\int_{\Omega}\left[\int_0^1 \beta_s^2(\omega)\,ds\right]\cdot \rho(t,x,\omega)\,d\mathscr{P}(\omega).$$

Then

$$\mu_W\left\{x:\sup_{0\leq t\leq 1}\alpha(t,x)<\infty\right\}=1.$$

Proof The process $\{\alpha(t,W(\omega)),\mathscr{F}_t^W\}$ is a right continuous martingale and for $0\leqq t\leqq 1$, $E_{\mathscr{P}}(\alpha(t,W))=\int_\Omega(\int_0^1\beta_s^2(\omega)\,ds)\,d\mathscr{P}(\omega)$.

LEMMA 3 *Let*

$$m(t,x)=\int_{\Omega}|\beta_t(\omega)|^2\rho(t,x,\omega)\,d\mathscr{P}(\omega).$$

Then

$$\mu_W\left\{x:\int_0^1 m(t,x)\,dt<\infty\right\}=1.$$

Proof Observe that

$$\int_{\Omega}\left[\int_0^1 m(t,W(\omega))\,dt\right]d\mathscr{P}(\omega)=\int_0^1\left[\int_{\Omega}|\beta_t(\omega)|^2\,d\mathscr{P}(\omega)\right]dt<\infty.$$

We return to the problem of comparing two weak solutions ξ_0, ξ_1 of (2) assuming that ξ_0, ξ_1 are both defined on the space $(\Omega,\mathscr{F},\mathscr{P})$.

Moreover, we may assume that

$$\mathscr{P}\left\{\int_0^1\gamma(t,\xi_i)^2\,dt<\infty\right\}=1$$

for $i=0,1$, by restricting attention to the class of all such solutions. Thus, we conclude from Lemmas 1–3 that

$$\sup_{0\leq t\leq 1} g(t,\xi_i(\omega))<\infty$$

$$\inf_{0\leq t\leq 1} g(t,\xi_i(\omega))>0$$

$$\sup_{0\leq t\leq 1} \alpha(t,\xi_i(\omega))<\infty$$

and

$$\int_0^1 m(t, \xi_i(\omega))\,dt < \infty - \omega \text{ a.s.}$$

Completion of the proof of Theorem 1. To complete the proof of the theorem we show that if ξ_0, ξ_1 are weak solutions of (2), then

$$\sup_{0 \leqq t \leqq 1} |\xi_0(t, \omega) - \xi_1(t, \omega)| = 0 \ \mathscr{P}\text{-a.s.}$$

Proof On $[0,1] \times \Omega$ define

$$L(t, \omega) = |\gamma(t, \xi_0) - \gamma(t, \xi_1)| = \left|\frac{d(\xi_0 - \xi_1)}{dt}\right|$$

$$= \left|\frac{f(t, \xi_0) - f(t, \xi_1)}{g(t, \xi_0)} + f(t, \xi_1)\left(\frac{1}{g(t, \xi_0)} - \frac{1}{g(t, \xi_1)}\right)\right|$$

where $f(t, x) = \int_\Omega \beta_t(\omega)\rho(t, x, \omega)\,d\mathscr{P}(\omega)$
Then

$$(L(t, \omega))^2 \leqq K(\omega)\{(f(t, \xi_0) - f(t, \xi_1))^2 + (f(t, \xi_1))^2[g(t, \xi_0) - g(t, \xi_1)]^2\} \tag{3}$$

where $K(\omega)$ is a constant greater than

$$\max\left(\frac{1}{\inf\limits_{0 \leqq t \leqq 1} (g(t, \xi_0))^2}, \frac{1}{\inf\limits_{0 \leqq t \leqq 1} (g(t, \xi_0))^2 \cdot \inf\limits_{0 \leqq t \leqq 1} (g(t, \xi_1))^2}\right) < \infty \omega\text{-a.s.}$$

For $0 \leqq u \leqq 1$, we write

$$\int_0^u L^2(t, \omega)\,dt \leqq K(\omega)\left\{\int_0^u \left[\int_\Omega |\beta_t(\hat{\omega})||\rho(t, \xi_0, \hat{\omega}) - \rho(t, \xi_1, \hat{\omega})|\,d\mathscr{P}(\hat{\omega})\right]^2 dt\right.$$

$$\left. + \int_0^u (f(t, \xi_1))^2 \cdot \left[\int_\Omega |\rho(t, \xi_0, \hat{\omega}) - \rho(t, \xi_1, \hat{\omega})|^2 d\mathscr{P}(\hat{\omega})\right] dt\right\}.$$

Since $|e^x - e^y| \leqq \frac{1}{2}(e^x + e^y)|x-y|$ for any values x, y it follows for all $\hat{\omega}$, t that

$$|\rho(t,\xi_0,\hat{\omega}) - \rho(t,\xi_1,\hat{\omega})|$$

$$\leqq \tfrac{1}{2}|\rho(t,\xi_0,\hat{\omega}) + \rho(t,\xi_1,\hat{\omega})| \cdot \left|\int_0^t \beta_s(\hat{\omega})\, d(\xi_0(\omega) - \xi_1(\omega))\right|$$

$$= \tfrac{1}{2}|\rho(t,\xi_0,\hat{\omega}) + \rho(t,\xi_1,\hat{\omega})| \cdot \int_0^t |\beta_s(\hat{\omega}) L(s,\omega)|\, ds. \tag{4}$$

Applying Hölder's inequality to the last integral term in (4), and bringing out $(\int_0^t (\mathscr{L}(s,\omega))^2\, ds)^{1/2}$, yields

$$\int_0^u (L(t,\omega))^2\, dt \leqq K(\omega) \int_0^u \left[\int_0^t (L(s,\omega))^2\, ds\right] \cdot \psi(t,\omega)\, dt \quad \omega\text{-a.s.} \tag{5}$$

where

$$\psi(t,\omega) = \left[\int_\Omega |\beta_t(\hat{\omega})| \frac{|(\rho(t,\xi_0,\hat{\omega}) + \rho(t,\xi_1,\hat{\omega})|}{2} \cdot \left(\int_0^t \beta_s^2(\hat{\omega})\, ds\right)^{1/2} d\mathscr{P}(\hat{\omega})\right]^2$$

$$+ \left[\int_\Omega \beta_t(\hat{\omega}) \rho(t,\xi_1,\hat{\omega})\, d\mathscr{P}(\hat{\omega})\right]^2$$

$$\cdot \left[\int_\Omega \left(\frac{\rho(t,\xi_0,\hat{\omega}) + \rho(t,\xi_1,\hat{\omega})}{2}\right)\left(\int_0^t \beta_s^2(\hat{\omega})\, ds\right)^{1/2} d\mathscr{P}(\hat{\omega})\right]^2.$$

By showing that $\psi(t,\omega)$ is an integrable function of t, ω-a.s., one may then iterate (5) and conclude that for $0 \leqq u \leqq 1$,

$$\int_0^u (\mathscr{L}(t,\omega))^2\, dt = 0 \quad \omega\text{-a.s.}$$

Hence

$$\sup_{0 \leqq t \leqq 1} |\xi_0(t,\omega) - \xi_1(t,\omega)| \leqq \int_0^1 (\mathscr{L}(t,\omega))^2\, dt = 0 \quad \omega\text{-a.s.}$$

and we have established path-wise uniqueness for weak solutions of (2).

To see that $\psi(t,\omega)$ is integrable in t, ω-a.s., apply Hölder's inequality to the first term to obtain,

$$\left[\int_\Omega |\beta_t(\hat\omega)| \frac{|\rho(t,\xi_0,\hat\omega)+\rho(t,\xi_1,\hat\omega)|}{2}\cdot\left(\int_0^t \beta_s^2(\hat\omega)\,ds\right)^{1/2} d\mathscr{P}(\hat\omega)\right]^2$$

$$\leqq \left(\int_\Omega |\beta_t^2(\hat\omega)| \frac{|\rho(t,\xi_0,\hat\omega)+\rho(t,\xi_1,\hat\omega)|}{2}\, d\mathscr{P}(\hat\omega)\right)$$

$$\cdot\left(\int_\Omega \left[\int_0^t \beta_s^2(\hat\omega)\,ds\right] \frac{(\rho(t,\xi_0,\hat\omega)+\rho(t,\xi_1,\hat\omega))}{2}\, d\mathscr{P}(\hat\omega)\right)$$

$$\leqq \int_\Omega |\beta_t^2(\hat\omega)| \frac{(\rho(t,\xi_0,\hat\omega)+\rho(t,\xi_1,\hat\omega))}{2}\, d\mathscr{P}(\hat\omega)$$

$$\cdot \sup_{0\leqq t\leqq 1} \int_\Omega \left[\int_0^1 \beta_s^2(\hat\omega)\,ds\right]\cdot\left(\frac{(\rho(t,\xi_0,\hat\omega)+\rho(t,\xi_1,\hat\omega))}{2}\right) d\mathscr{P}(\hat\omega).$$

Integrating this product over t gives

$$\sup_{0\leqq t\leqq 1} \int_\Omega \left[\int_0^1 \beta_s^2(\hat\omega)\,ds\right]\cdot\left(\frac{(\rho(t,\xi_0,\hat\omega)+\rho(t,\xi_1,\hat\omega))}{2}\right) d\mathscr{P}(\hat\omega)$$

$$\cdot\int_0^u \left[\int_\Omega |\beta_t^2(\hat\omega)| \frac{|\rho(t,\xi_0,\hat\omega)+\rho(t,\xi_1,\hat\omega)|}{2}\, d\mathscr{P}(\hat\omega)\right] dt < \infty \qquad \omega\text{-a.s.}$$

by Lemmas 2 and 3.

The second term of $\psi(t,\omega)$ is handled analogously using Lemmas 1, 2 and 3. This completes the proof.

Thus, the observations process $\{y_t\}$, $0\leqq t\leqq 1$, is the (unique) strong solution satisfying (2) under the restriction that

$$\mathscr{P}\left(\int_0^1 \gamma(t,\xi)^2\,dt < \infty\right) = 1.$$

FINAL REMARKS

Let us rewrite Eq. (2) as

$$v = (I - \mathscr{N})y,$$

where $\mathcal{N}$ is a non-linear operator from $C[0,1;\mu_y]$ into $C[0,1;\mu_v]$. Under assumptions (a) and (b) we have shown that an inverse operator $(I+\hat{\mathcal{N}})$ exists such that $\mathcal{P}$-a.s.

$$y=(I+\hat{\mathcal{N}})v.$$

Moreover, if $\pi_t:C[0,1]\to C[0,1]$ denotes the truncation operator defined by

$$(\pi_t x)(s)=\begin{cases} x_s, & 0\leqq s\leqq t\\ 0, & \text{otherwise,}\end{cases}$$

then

$$(I+\hat{\mathcal{N}})(\pi_t v)=\pi_t y,$$

that is, the operator $(I+\hat{\mathcal{N}})$ is *causal*.

Our results in this paper suggest the investigation of causal-invertibility of non-linear causal operators on abstract Wiener Spaces (in the sense of Gross) using methods of stochastic integration and martingales. Such an investigation would also be of importance in the theory of stochastic stability of feedback systems.

Acknowledgements

It is a pleasure for the second author to thank Daniel Ocone of the Mathematics Department, M.I.T., for first noticing that there was a gap in the proof of the innovations conjecture in the paper by Beneš; S. R. S. Varadhan for the conversations about the Yamada–Watanabe theorem and other suggestions; V. Beneš, J. M. C. Clark and T. Kailath for discussions on the innovations problem.

References

[1] M. Fujisaki, G. Kallianpur and H. Kunita, Stochastic differential equations for the nonlinear filtering problems, *Osaka J. Math.* **9** (1972), 19–40.

[2] P. Frost, *Estimation in Continuous Time Nonlinear Systems*, Dissertation, Stanford University, Stanford, California, June 1968.

[3] P. A. Frost and T. Kailath, An innovations approach to least-squares estimation—Part III: nonlinear estimation in white gaussian noise, *IEEE Trans. on Automatic Control* **AC-16** 3 (1971), 217–226.

[4] J. M. C. Clark, Conditions for the one-to-one correspondence between an observation process and its innovations, *Tech. Report No. 1*, Imperial College, London, England, 1969.

[5] V. E. Beneš, Extension of Clark's innovations equivalence theorem to the case of signal independent of noise, with $\int_0^T z_s^2\, ds < \infty$ a.s., *Mathematical Programming Study* **5** (1976), 2–7.

[6] G. Kallianpur, A linear stochastic system with discontinuous control, *Proc. of International Symposium on Stochastic Differential Equations, Kyoto 1972*, edited by K. Ito, Wiley Interscience, 1978, pp. 127–140.

[7] D. F. Allinger and S. K. Mitter, New results on the innovations problem, *Technical Memo LIDS*, October 1979. Presented at the 18th IEEE Conference on Decision and Control, Florida, December 1979.

[8] R. Liptser and A. Shiryayev, *Statistics of Random Processes, I, General Theory*, Springer-Verlag, New York, 1977.

[9] V. E. Beneš, Nonexistence of strong nonanticipating solutions to stochastic DEs: implications for functional DEs, filtering and control, *Stochastic Processes and Their Applications* **5**, no. 3 (1977), 243–263.

[10] V. E. Beneš, On Kailath's innovations conjecture. *Bell System Technical Journal*, **55**, no. 7, September 1976, pp. 981–1001.

[11] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations, *J. Math. Kyoto Univ.* **11** (1971), 155–167.

[12] D. W. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin, New York, 1979.

[13] G. Kallianpur and C. Striebel, Estimation of stochastic processes: arbitrary system processes with additive white noise error, *Ann. Math. Stat.* **39** (1968), 785–801.

Appendix

Consider the stochastic differential equation on $0 \leqq t \leqq 1$

$$d\xi_t = \gamma(t, \xi)\, dt + dw_t,$$

where w_t is a Wiener process and $\gamma(\cdot, \cdot)$ is a jointly measurable functional on $[0,1] \times \mathscr{C}(0,1;\mu_w)$ and we assume the equation has a weak solution.

We use the definition of weak and strong solutions in the sense of Liptser–Shiryayev (cf. [8], Defns. 8, 9, 10, 11, pp. 127–128).

Then according to Remark 2 and Corollary 1 of Yamada–Watanabe [11] we have:

If any two weak solutions given on the same probability space coincide pathwise, then a unique non-anticipating functional $\phi(t,x)$, $t \in [0,1]$, $x \in \mathscr{C}(0,1;\mu_w)$ exists such that for any solution ξ the representation $\xi_{(\cdot)} = \phi(\cdot, v)$ holds a.s. Moreover, this result holds when the class of solutions is restricted to some subset $\mathscr{Y} \subset \mathscr{C}(0,1;\mu_w)$.

Note added in proof

Additional Reference: N. V. Krylov; On the equivalence of ∇-algebras in the filtering problem of diffusion processes, *Theory of Probability and its Applications*, Vol. XXIV, No. 4 (1979), 772–781.