THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 55, No. 7, September 1976
Printed in U.S.A.

On Kailath's Innovations Conjecture Hold

By V. E. BENEŠ

(Manuscript received October 30, 1975)

With $z_.$ a signal process, $w_.$ a Brownian motion, and $y_t = \int_0^t z_s ds + w_t$ a noisy observation, the innovations problem is to determine whether $y_.$ is adapted to the innovations process $\nu_.$, which is also a Brownian motion, and is defined using the estimate $\hat{z}_t = E\{z_t | y_s, 0 \leqq s \leqq t\}$ by $y_t = \int_0^t \hat{z}_s ds + \nu_t$. The closely related σ-algebras problem in stochastic DEs is to determine, for a given causal drift α, when a solution of $d\xi = \alpha(t, \xi)dt + dw$ is a causal functional of $w_.$. Previous results on these problems are reviewed and extended. In particular, we broach and answer positively the physically important case of the innovations problem in which the signal satisfies a stochastic DE with drift depending in part on the noisy observations. This case is important because it models a system observed through noise and controlled by feedback of these noisy observations. The last part of the paper shows that the innovations problem has a positive resolution if and only if on some probability space there is a Brownian motion W and a causal solution ξ of $d\xi = \alpha(t, \xi)dt + dW$, where α expresses the estimator $\hat{z}$; that is, α is a causal functional such that $\hat{z}_t = \alpha(t, y)$.

I. INTRODUCTION

Estimation of signals from past observations of them corrupted by noise is a classical problem of filtering theory. The following is a standard mathematical idealization of this problem: The signal $z_.$ is a measurable stochastic process with $E|z_t| < \infty$, the noise $w_.$ is a Brownian motion, and the observations consist of the process

$$y_t = \int_0^t z_s ds + w_t. \tag{1}$$

Define $\hat{z}_t = E\{z_t | y_s, 0 \leqq s \leqq t\}$, the expected value of z_t given the past of the observations up to t. It can be shown[1] that if $\int_0^t z_s^2 ds < \infty$ a.s., then there is a measurable version of $\hat{z}_.$ with $\int_0^t \hat{z}_s^2 ds < \infty$ a.s. The *innovations process* for this setup is defined to be

$$\nu_t = \int_0^t (z_s - \hat{z}_s)ds + w_t,$$

and it is a basic result of Frost[2] and also of Kailath[3] that, under weak conditions, $\nu_.$ is itself a Wiener process with respect to the observations. Thus, (1) is equivalent to the integral equation

$$y_t = \int_0^t \hat{z}_s ds + \nu_t, \tag{2}$$

which reduces the general case (1) to that in which $z_.$ is adapted to $y_.$, a special property useful in questions of absolute continuity in filtering and detection.

Since $\hat{z}_.$ is of necessity adapted to $y_.$, Eq. (2) purports to define $y_.$ in terms of $\nu_.$; the *innovations problem*, first posed by Frost,[2] is precisely to determine whether it really does. Frost asked: Do the innovations contain all the information in the observations? [By (2) they do not contain more.] In the language of probability this is to ask whether the σ-algebras that the processes generate are the same up to null sets; i.e., is

$$\mathcal{Y}_0^t \triangleq \sigma\{y_s, s \leqq t\} = \sigma\{\nu_s, s \leqq t\} \triangleq \mathfrak{N}_0^t \pmod{P}?$$

II. THE σ-ALGEBRAS PROBLEM IN STOCHASTIC DEs

The innovations problem is equivalent to an apparently more general problem from the theory of stochastic DEs, sometimes called the σ-algebras problem: Given a causal drift $a(s, x)$, possibly depending on the past of the function x, and a weak solution of the DE $dx_t = a(t, x)ds + dw_t$, with $w_.$ a Brownian motion, and $x_.$ possibly nonanticipating with respect to $dw_.$, to determine whether

$$\sigma\{x_s, s \leqq t\} = \sigma\{w_s, s \leqq t\} \qquad \pmod{P}.$$

Positive answers to both problems were widely conjectured.

The innovations problem has been outstanding, in both senses of the word, since about 1968, and it has drawn the attention of communications theorists and probabilists alike. The σ-algebras problem has been current in the Soviet Union since the late 1950s; there it has been the object of great effort and a source of stimulus far in excess of its simple origins.[4] Accounts of the innovations problem and its theoretical background are in lecture notes by Meyer[1] and in a paper by Orey.[5]

It is now known that the answer to the general problem is in the negative. B. Cirel'son has given a counterexample[6,17] for the following special case (this case shows, incidentally, that the innovations and σ-algebras problems are in fact the same): Suppose that the signal $z_.$ is a causal functional $a(t, y)$ of the observations; i.e., the signal is entirely determined by feedback from the observations. Then $z = \hat{z}$, $w = \nu$, and the problem reduces to asking whether the observations are "well-defined" in the strong sense of being adapted to the noise;

for in this vestigial or degenerate case, the noise is the only process left. Cirel'son's disturbing example consists of a choice of $a(\cdot, \cdot)$ for which there is just one weak solution $y_.$, which is nonanticipative in that the future increments of $w_.$ are independent of the past of $y_.$, but which cannot be expressed as a functional of $w_.$, causal or not, over any interval.

Prior to this counterexample, several cases of the problems had been settled in the affirmative. J. M. C. Clark[7] proved that if noise and signal are independent and the signal is bounded (uniformly in t and ω), then observations are adapted to innovations. The author[8] extended Clark's method and result to the case where signal and noise are independent and the signal is almost surely (a.s.) square-integrable. The case of gaussian observations turns out affirmatively: here results of Hitsuda[9] imply that $\hat{z}_.$ is a linear functional of the past of $y_.$ and eq. (2) is solvable by a Neumann series. Zvonkin[10] has given an affirmative answer to the σ-algebras problem for the Markov case $a(s, y) = a(y_s)$ bounded and homogeneous in time, using the associated scale function to transform the state space; this result extends to time-dependent bounded $a(s, y_s)$ satisfying Dini's condition.

It should be remarked that although the innovations and σ-algebras problems are mathematically equivalent, they arise in different contexts, involve different emphases, and can be usefully contrasted, as discussed below.

The innovations problem arose in filtering theory, and it focuses especially on the nature of the filter or operator that gives $\hat{z}_t$ as a causal functional $\alpha(t, y)$ on the past of $y_.$; from this point of view, the example of Cirel'son, in which there is no real filtering going on, is a bit wide of the mark; the *real* problem is to find out enough about the filter to be able to settle whether $\sigma\{y\} = \sigma\{\nu\}$ in cases where there is a real signal (determined in part by sources other than the noise, and in part possibly by control or feedback based on the observations), which it is desired to control, transmit, filter, or detect.

The σ-algebras problem arises in stochastic functional DEs, and is therefore more general in scope, since the drift functionals considered need no longer be filters or conditional expectations like $\hat{z}_t$; the emphasis is on dynamics, causality, and nonanticipation, with no admixture of estimation. If the drift functional to be considered is a filter, it may have special properties that are useful in the investigation. (See the method of Clark.[7])

III. SUMMARY

Relevant notions from stochastic DEs are defined in Section IV: causal functionals, weak solutions, causal solutions, and nonanticipative solutions. A n.a.s. martingale-type condition for validity of the innovations conjecture is given in Section V; as an application, this condition yields (Section VI) the (known) conjecture for the case of gaussian observations. Section VIII describes some of Zvonkin's positive results for the Markov case and an extension. In Section VIII, we investigate the problem of calculating the estimate $\hat{z}_.$ and give various relationships based on absolute continuity of measures. Section IX is devoted to the physically important case of signals $z_.$ that solve stochastic DEs with drift based on feedback of observations; we show that if this drift is Lip in the feedback of linear growth in the signal uniformly in the observations, then observations are adapted to innovations. In Sections X and XI, finally, we show that validity of the innovations conjecture is equivalent to the causal solvability, on some probability space, of the equation $d\xi = \alpha(t, \xi)dt + dW$, with W Brownian and α the (a?) functional, such that $\hat{z}_t = \alpha(t, y)$.

IV. CAUSAL SOLUTIONS OF STOCHASTIC DEs

A measurable functional $\gamma: [0, \infty) \times C[0, \infty) \to R$ is called *causal* if for each $t \in [0, \infty)$, $x_s = y_s$ for $s \leqq t$ implies

$$\gamma(s, x) = \gamma(s, y) \qquad x, y \in C[0, \infty).$$

The idea expressed by this definition is the physical one that $\gamma(t, \cdot)$ cannot depend functionally on any more than the past of its argument up to t; thus, it has the same value at t for two functions that agree for $s \leqq t$. In spite of the presence of the word "depend" in the previous sentence, causality of a functional is expressible as a measurability property, and has no immediate relation to any probability measure.

Let α be a causal functional. A *weak* solution of the stochastic DE,

$$dx = \alpha(t, x)dt + dW, \qquad \text{W. Brownian}, \tag{3}$$

is a process $\xi_.$ such that

$$(T\xi)_t = \xi_t - \int_0^t \alpha(s, \xi)ds = \nu_t$$

is a Brownian motion on its own past. If $\xi_.$ is adapted to $\nu_.$; i.e., if for each t, ξ_t is measurable with respect to $\sigma\{\nu_s, s \leqq t\}$, then $\xi_.$ is called a *causal* solution, and there is a causal functional φ such that $\xi_t = \varphi(t, \nu)$ at each t with probability one. A solution $\xi_.$ is called *non-anticipative* if (roughly) the future increments of $\nu_.$ are independent of the past of $\xi_.$; i.e., for each t

$$\sigma\{\nu_u - \nu_t, u \geqq t\} \perp\!\!\!\perp \sigma\{\xi_s, s \leqq t\}.$$

This is a probabilistic property, and it is equivalent to ν's being a martingale on the larger algebras $\sigma\{\xi_s, s \leqq t\}$. It can be seen that for (3), causal $\Rightarrow$ nonanticipative, but the converse is false: it is known that there are drift functionals α for which there exist nonanticipative weak solutions, but no causal solutions.[17]

V. A SEMI-MARTINGALE CONDITION

Theorem 1: $\sigma\{\nu_s, s \leqq t\} = \sigma\{y_s, s \leqq t\}$ *(mod P) for each* $t \geqq 0$ *iff there is a* $\mathcal{Y}_0^{\cdot}$*-adapted martingale* x_t *and a causal functional* ψ: $[0, \infty) \times C[0, \infty) \rightarrow R$ *such that the observations are representable (as a semi-martingale on their own past) by*

$$y_t = x_t + \int_0^t \psi(s, x)ds.$$

Proof: The hypothesis on $y.$ and the equation $dy = \hat{z}dt + d\nu$ imply that

$$x_t - \nu_t = \int_0^t [\hat{z}_s - \psi(s, x)]ds. \tag{4}$$

The innovations process is a Wiener process with respect to the observations; thus, the left side of (4) is a continuous $\mathcal{Y}_0^{\cdot}$-martingale. Since the right side is absolutely continuous in t, it follows that both sides vanish identically, so that $x.$ and $\nu.$ are indistinguishable processes, and

$$\int_0^t \hat{z}_s ds = \int_0^t \psi(s, x)ds = \int_0^t \psi(s, \nu)ds.$$

But the right-hand side is ν-adapted because ψ is causal. The theorem follows from $dy = \hat{z}dt + d\nu$. For the converse, we argue thus: if the σ-algebras coincide, there is a causal functional φ such that $y_t = \varphi(t, \nu)$. The innovations theorem makes $\nu.$ a $\mathcal{Y}_0^{\cdot}$-martingale with

$$y_t = \nu_t + \int_0^t \hat{z} \circ \varphi ds;$$

then, take $x = \nu$ and $\psi = \hat{z} \circ \varphi$.

VI. APPLICATION TO GAUSSIAN OBSERVATIONS

The theorem just proved affords us a simple demonstration of the validity of the innovations conjecture for gaussian observations. Suppose that the signal $z.$ is square-integrable almost surely. Then, a theorem of Kailath and Zakai[11] implies that the measure induced by $y.$ is absolutely continuous with respect to Wiener measure. The class of gaussian processes absolutely continuous with respect to Wiener measure has been characterized in a causal way by Hitsuda[9]: a process $y.$ belongs to this class iff there is a Wiener process $W.$ adapted to $y.$

and a Volterra kernel $m(\cdot, \cdot) \in L_2[0, 1]^2$ such that

$$y_t = W_t + \int_0^t \int_0^s m(s, u) dW_u ds.$$

Thus, $W_.$ is a martingale on the past of $y_.$ and in Theorem 1 we can set $x = W$, $\psi(s, W) = \int_0^s m(s, u) dW_u$, to conclude that ν is W and $\hat{z}$ is ψ. Iteration of the relation between y and W gives

$$\nu_t = y_t - \int_0^t \left(\int_0^s m(s, u) dy_u - \int_0^s m(s, u) \cdot \int_0^u m(u, v) du dy_v + \cdots \right) ds,$$

so that letting

$$l(t, s) = m(t, s) - \int_0^t m(t, u) m(u, s) du + \cdots$$

be the Neumann series or the resolvent of $m(\cdot, \cdot)$, we see that

$$y_t = \nu_t + \int_0^t \int_0^s l(s, u) dy_u ds$$

$$\hat{z}_t = \int_0^t l(t, u) dy_u.$$

Thus, the map $\hat{z}$ is linear in $y_.$ when $y_.$ is gaussian, as was expected.

VII. RESULTS OF ZVONKIN FOR THE MARKOV CASE

For a stochastic DE of the form $dy = a(y_t)dt + dw_t$, Zvonkin[10] has shown that if $a(\cdot)$ is bounded, then there is a causal solution $y_.$. His procedure[12] is to look at the scale function

$$u(y) = \int_0^y \exp - 2 \int_0^z a(s) ds dz = \int_0^y \beta(z) dz \tag{5}$$

and to note that $\sigma\{y_s, s \leqq t\} = \sigma\{u(y_s), s \leqq t\}$ because $u(\cdot)$ is monotone. Then to show that $u(y_s)$ can be got causally from $w.$, he uses Ito's rule on $z_t = u(y_t)$ to get

$$\begin{aligned} dz_t &= \beta(y_t) dy_t - \beta(y_t) a(y_t) dw_t \\ &= \beta[u^{-1}(z_t)] dw_t. \end{aligned}$$

By calculus he finds $a(\cdot)$ bounded $\Rightarrow \beta(u^{-1}) \in \text{Lip}$; hence, $z_.$ is a causal functional of $w_.$ and so is $y_.$.

Now this argument depends in part on the fact that u satisfies $\frac{1}{2}u'' + u'a = 0$, and at once suggests extensions to the inhomogeneous case $a(t, y) = a(t, y_t)$. We give an example based on Zvonkin's paper:[10]

Theorem 2: Suppose that for some strictly increasing function $k(\cdot)$ there exists a solution $u(t, y)$ of the Cauchy problem

$$u(0, y) = k(y)$$
$$u_1 + \tfrac{1}{2}u_{22} + a(t, y)u_2 = \mathfrak{L}u = 0^{\dagger}$$

such that $(\log u_2)_2$ is bounded. Then, the stochastic DE

$$dy_t = a(t, y_t)dt + dw_t \tag{6}$$

has a causal solution.

Proof: Since $k\uparrow$, it is clear that $u_y > 0$ and that the transformation $z_t = u(t, y_t)$ is bijective. Using Ito's rule, we calculate the stochastic differential of $z.$ as

$$\begin{aligned} dz_t &= (\mathfrak{L}u)(t, y_t)dt + u_2(t, y_t)dw_t \\ &= u_2[t, u^{-1}(t, z_t)]dw_t. \end{aligned} \tag{7}$$

Now

$$\begin{aligned} \frac{\partial}{\partial z} u_2[t, u^{-1}(t, z)] &= \frac{u_{22}[t, u^{-1}(t, z)]}{u_2[t, u^{-1}(t, z)]} \\ &= (\log u_2)_2 \big|_{y=u^{-1}(t,z)} \text{ bounded.} \end{aligned}$$

Hence, by the usual Ito's theory of stochastic DEs, the martingale eq. (7) has a unique causal solution $z.$ which is a bijection of $y.$ pointwise in time. Hence, $y.$ is a causal functional of $w.$ too. The homogeneous case follows if we take $u(t, y) \equiv$ the scale function (5).

In a similar vein we can show this result:

Theorem 3: If $a(\cdot, \cdot)$ is bounded and such that

$$\frac{\partial}{\partial t} \int_0^y a(t, z)dz$$

exists and is bounded, then (6) has a causal solution.
Proof: Let

$$u(t, y) = \int_0^y \exp - 2 \int_0^z a(t, x)dxdz$$

so that $u_2 > 0$, and $z_t = u(t, y_t)$ is bijective. We have

$$dz_t = u_1[t, u^{-1}(t, z_t)]dt + u_2[t, u^{-1}(t, z_t)]dw_t. \tag{8}$$

By calculus obtain formally

$$\frac{\partial}{\partial z} u_1[t, u^{-1}(t, z)] = -2a(t, z)$$
$$\frac{\partial}{\partial z} u_2[t, u^{-1}(t, z)] = -2\frac{\partial}{\partial t} \int_0^z a(t, x)dx,$$

† Numerical indexes show which variable is differentiated and how many times.

both bounded by hypothesis. Hence, (8), and so (6), has a causal solution, by Ito's theory.

Note that Zvonkin's Markov case is not relevant to Kailath's innovations problem, because the filter giving $\hat{z}_t$ will almost invariably depend with advantage on the whole past of the observation process.

VIII. CALCULATION OF THE ESTIMATOR $\hat{z}$

Let (Ω, B, P) be a probability space on which are defined the signal process $z_.$, the noise $w_.$, the estimate $\hat{z}_.$, and the innovations process $\nu_.$, related by

$$y_t = \int_0^t z_s ds + w_t = \int_0^t \hat{z}_s ds + \nu_t.$$

Under some mild technical assumptions, this setup has implicit in it a rich structure that allows us to give a "formula" for $\hat{z}$, *inter alia*. To penetrate deeper into the situation, it is convenient to introduce an absolutely continuous change of measure which makes the observation process Brownian. We shall restrict attention to the interval $0 \leqq t \leqq 1$, and assume that

(i) $$\int_0^1 z_s^2 ds < \infty \quad \text{a.s.}$$

(ii) There is a system of increasing σ-algebras $\mathfrak{F}_t$, $0 \leqq t \leqq 1$, to which $z_.$ and $w_.$ are adapted, and $w_.$ is a Wiener process with respect to $(P, \mathfrak{F}_.)$.

(ii) $$E \exp\left\{-\int_0^1 z_s dw_s - \frac{1}{2}\int_0^1 z_s^2 ds\right\} = 1.$$

Then by Girsanov's theorem,[13] the observation process $y_.$ is a Wiener process on $\mathfrak{F}_.$ (and thus on $\mathcal{Y}_0^{\cdot} = \sigma\{y_s, s \leqq \cdot\}$) under the transformed measure P_0 defined on $\mathfrak{F}_1$ by

$$\frac{dP_0}{dP} = \exp\left\{-\int_0^1 z_s dw_s - \frac{1}{2}\int_0^1 z_s^2 ds\right\}.$$

It is convenient to use the functional notation[5]

$$q(f, g)_t = \exp\left\{\int_0^t f_s dg_s - \frac{1}{2}\int_0^t f_s^2 ds\right\}.$$

Then $(dP_0/dP)^{-1} = q(z, y)_1 > 0$ a.s. and for $A \in \mathfrak{F}_1$

$$P(A) = \int_A q(z, y)_1 \frac{dP_0}{dP} dP, \qquad \frac{dP}{dP_0} = q(z, y)_1$$

so that $P \ll P_0$ and so $P \sim P_0$. The following formula for $\hat{z}$ is then readily justified: with $\mathcal{Y}_0^t = \sigma\{y_s, 0 \leqq s \leqq t\}$,

$$\hat{z}_t = \frac{E_0\{z_t q(z, y)_1 | \mathcal{Y}_0^t\}}{E_0\{q(z, y)_1 | \mathcal{Y}_0^t\}}.$$

For let $A \in \mathcal{Y}_0^t$, so that

$$\begin{aligned}\int_A z_t dP &= \int_A z_t \frac{dP}{dP_0} dP_0 = \int_A E_0\{z_t q(z, y)_1 | \mathcal{Y}_0^t\} dP_0 \\ &= \int_A \frac{E_0\{z_t q(z, y)_1 | \mathcal{Y}_0^t\}}{E_0\{q(z, y)_1 | \mathcal{Y}_0^t\}} E_0\{q(z, y)_1 | \mathcal{Y}_0^t\} dP_0 \\ &= \int_A \frac{E_0\{z_t q(z, y)_1 | \mathcal{Y}_0^t\}}{E_0\{q(z, y)_1 | \mathcal{Y}_0^t\}} dP, \end{aligned} \tag{9}$$

since for $A \in \mathcal{Y}_0^t$ and a $\mathcal{Y}_0^t$-measurable function $\mathfrak{N}$ integrable

$$\int_A \mathfrak{N} \frac{dP}{dP_0} dP_0 = \int_A \mathfrak{N} E_0\left\{\frac{dP}{dP_0} | \mathcal{Y}_0^t\right\} dP_0.$$

Thus, the ratio (integrand) in eq. (9) is a version of $\hat{z}_t$. The process

$$E_0\left\{\frac{dP}{dP_0} | \mathcal{Y}_0^t\right\}$$

is a positive martingale on the past of the Brownian (under P_0) motion $y_.$ and can therefore be expected to have a special form. It can in fact be shown by arguments of Shiryaev and Liptser[14] that

$$E\left\{\frac{dP_0}{dP} | \mathcal{Y}_0^t\right\} = q^{-1}(\hat{z}, y)_t.$$

From this it follows that

$$E_0\left\{\frac{dP}{dP_0} | \mathcal{Y}_0^t\right\} = q(\hat{z}, y)_t.$$

It is obvious intuitively that $q(z, y)_1$ in eq. (9) can be changed to $q(z, y)_t$: since $q(z, y)_.$ is an $\mathfrak{F}_.$-martingale, we have a.s.

$$\begin{aligned} E_0\left\{\frac{dP}{dP_0} q^{-1}(z, y)_t | \mathfrak{F}_t\right\} &\equiv 1 \\ E_0\left\{\frac{dP}{dP_0} | \mathfrak{F}_t\right\} &= q(z, y)_t \\ E_0\left\{z_t \frac{dP}{dP_0} | \mathfrak{F}_t\right\} &= E_0\{z_t q(z, y)_t | \mathfrak{F}_t\}. \end{aligned}$$

Hence, $\mathcal{Y}_0^t \subseteq \mathfrak{F}_t$ gives

$$E_0\left\{z_t \frac{dP}{dP_0} | \mathcal{Y}_0^t\right\} = E_0\{z_t q(z, y)_t | \mathcal{Y}_0^t\}, \quad \text{a.s.}$$

Thus, the filtering $\hat{z}_.$ can be represented as

$$\hat{z}_t = \frac{E_0\{z_t q(z, y)_t | \mathcal{Y}_0^t\}}{q(\hat{z}, y)_t}$$
$$= \frac{d/dt\langle y, q(\hat{z}, y)\rangle_t}{q(\hat{z}, y)_t}.$$

The last equation, it can be verified, is an identity valid for any $\mathcal{Y}_0^{\cdot}$-adapted, a.s. square-integrable functional, not just $\hat{z}$. Thus, the meat of the formula is the numerator, as could be expected intuitively since the denominator is basically a normalizer.

IX. SIGNALS SOLVING ITO DEs WITH DRIFT BASED ON FEEDBACK OF OBSERVATIONS

The positive results of Clark,[7] based on the assumption of independence between signal and noise, seem adequate for many practical purposes of one-way communication or detection. For the more general physical applications to estimation and control, involving feedback of observations to control the signal, it would be pleasant to be able to weaken this assumption and allow some physically reasonable dependence between signal and noise. A natural setup to investigate is a generalization of the usual Kalman filter situation, in which the signal and the observation each solves a stochastic DE, with independent driving white noises, and with the drift for the signal equation depending on the observations. Thus, we let the signal $z_.$ and observation $y_.$ respectively solve

$$dz_t = b(t, z, y)dt + dW_t \tag{10}$$

$$dy_t = z_t dt + dw_t. \tag{11}$$

The second equation is, of course, eq. (1) differentiated; the functional b, causal in both z and y simultaneously, represents the deterministic dynamics of a system described by $z_.$ and depending on the past of both signal and observation.

As noted at the end of Section I, the only way to make headway is to find out something about the form of the filter that gives $\hat{z}_.$; we shall show that in the case of eqs. (10) and (11) an analog of the Kallianpur-Striebel[15] formula for $\hat{z}$ provides enough structure on which to hang a proof similar to Clark's.[7]

Let us assume, as is physically reasonable, that $b(t, z, y)$ grows at most linearly with $\sup_{0 \leq s \leq t}|z_s|$, uniformly in y. Then

$$Eq[b(W, w), W]_t q(W, w)_t = 1,$$

and we can "solve" (10) and (11) by Girsanov's theorem in such a way that the joint solution process (z_t, y_t) is absolutely continuous with respect to the two-dimensional Brownian motion, here (W_t, w_t), with derivative $q[b(W, w), W]_1 q(W, w)_1$. It is then easily seen that $\hat{z}_t$ should have the form

$$\hat{z}_t = \frac{\int M(dW) q[b(W, y), W]_t q(W, y)_t W_t}{\int M(dW) q[b(W, y), W]_t q(W, y)_t} = \alpha(t, y), \tag{12}$$

where M is the Wiener measure for W alone. Verification is left to the reader; either of two more or less equivalent methods will do: direct integration over $\mathcal{Y}_0^t$ sets using the absolute continuity or introduction of P_1 by $dP_1 = q(-z, w)q[-b(z, y), W]dP$, and a use of it similar to that of P_0 in Section VIII. Note that P_1 makes (z, y) a 2-dimensional Brownian motion.

Theorem 4: If $z_.$ and $y_.$ are nonanticipating solutions of eqs. (10) and (11), and there exists a constant K such that for x, ξ, and $\eta \in C[0, \infty)$

$$|b(t, x_1, \eta)| \leqq K[1 + \sup_{0 \leqq s \leqq t} |x_1(s)|], \tag{13}$$

$$|b(t, x, \eta) - b(t, x, \xi)| \leqq K \sup_{0 \leqq s \leqq t} |y - \xi|$$

then $\sigma\{y_s, s \leqq t\} = \sigma\{\nu_s, s \leqq t\}$ (mod P).

Remark: Eq. (13) $\not\Rightarrow$ (10) and (11) have a unique causal solution. In fact, our argument will not devolve on whether eqs. (10) and (11) have a strong solution at all; the unique (in law) non-anticipative solution *is* the Girsanov solution, with derivative $q[b(W, w), W]q(W, w)$, which determines $\hat{z}$ via eq. (12).

Proof: With the explicit form eq. (12) for $\hat{z}$ available, a form of argument previously used by the author[8] (and generalized from that of Clark[7]) can be used: we exhibit a sequence of $\nu_.$-adapted processes converging to $\hat{z}_.$; the result then follows from eq. (2): Let

$$S(m, t) = \{\sup_{0 \leqq s \leqq t} |W_s| \leqq m\}, \quad m = 1, 2, \cdots.$$

It can be seen that the approximations

$$\hat{z}_m(t) = \frac{\int_{S(m,t)} M(dW) q[b(W, y), W]_t q(W, y)_t W_t}{\int_{S(m,t)} M(dW) q[b(W, y), W]_t q(W, y)_t} = \hat{z}_m(y)_t, \tag{14}$$

approach $\hat{z}_t$ as $m \to \infty$, and that each one is adapted to $y_.$; therefore, it is enough to prove that each one is adapted to $\nu_.$ itself. Now set

$$\hat{z}_t^{m,0} \equiv 0, \qquad m = 1, 2, \cdots$$

$$y_t^{m,n} = \int_0^t \hat{z}_s^{m,n} ds + \nu_t, \qquad m \wedge n \geqq 1$$

$$\hat{z}_t^{m,n+1} = \hat{z}_m(y^{m,n})_t,$$

$$y_t^m = \int_0^t \hat{z}_m(s) ds + \nu_t,$$

and note that $d(y^{m,n} - y^m) = (\hat{z}^{m,n} - \hat{z}_m)dt$ and

$$q(f, y^{m,n})_t = q(f, y^m) \exp \int_0^t f_s(\hat{z}^{m,n} - \hat{z}_m)_s ds.$$

With $\hat{z}^{m,n} - \hat{z}_m = \psi^{m,n}$ for short, we find from eq. (14) that

$$\psi_t^{m,n+1} = \int_{S(m,t)} M(dW) W_t \left\{ \frac{q[b(W, y^{m,n}), W]_t q(W, y^{m,n})_t}{A_t^m(y^{m,n})} - \frac{q[b(W, y^m), W]_t q(W, y^m)_t}{A_t^m(y^m)} \right\},$$

where

$$A_t^m(f) = \int_{S(m,t)} M(dw) q[b(w, f), w]_t q(w, f)_t.$$

Subtracting the two fractions on the right of $\psi_t^{m,n+1}$ above, and using the further abbreviation $p(f, g)_t$ for $q[b(f, g), f]_t$, we find

$$\psi_t^{m,n+1} = \int_{S(m,t)^2} M(dW) M(dw) w_t \cdot \frac{p(w, y^m)_t q(w, y^m)_t p(W, y^{m,n})_t q(W, y^{m,n})_t - p(w, y^{m,n})_t q(w, y^{m,n})_t p(W, y^m)_t q(W, y^m)_t}{A_t^m(y^{m,n}) A_t^m(y^m)}.$$

The numerator in the integrand is just

$$p(w, y^m)_t q(w, y^m)_t p(W, y^m)_t q(W, y^m)_t \cdot \Bigg[\exp \left\{ \int_0^t [b(W, y^{m,n}) - b(W, y^m)] dw - \frac{1}{2} \int_0^t [b^2(W, y^{m,n}) - b^2(W, y^m)] ds + \int_0^t W_s \psi_s^{m,n} ds \right\} - \exp \left\{ \int_0^t [b(w, y^{m,n}) - b(w, y^m)] dw - \frac{1}{2} \int_0^t [b^2(w, y^{m,n}) - b^2(w, y^m)] ds + \int_0^t w_s \psi_s^{m,n} ds \right\} \Bigg],$$

so we can use the inequality $|e^A - e^B| \leqq \frac{1}{2}(e^A + e^B)|A - B|$ to find that

$$|\psi_t^{m,n+1}| \leqq \frac{M}{2}\int_{S(m,t)^2} M(dW)M(dw)$$

$$\cdot \frac{p(w, y^m)_t q(w, y^m)_t p(W, y^{m,n})_t q(W, y^{m,n})_t + p(w, y^{m,n})_t q(w, y^{m,n})_t p(W, y^m) q(W, y^m)_t}{A_t^m(y^{m,n}) A_t^m(y^m)}$$

$$\cdot \Bigg| \int_0^t [b(W, y^{m,n}) - b(W, y^m)]dW - \frac{1}{2}\int_0^t [b^2(W, y^{m,n}) - b^2(W, y^m)]ds + \int_0^t W_s \psi_s^{m,n} ds$$

$$- \int_0^t [b(w, y^{m,n}) - b(w, y^m)]dw + \frac{1}{2}\int_0^t [b^2(w, y^{m,n}) - b^2(w, y^m)]ds - \int_0^t w_s \psi_s^{m,n} ds \Bigg| .$$

The Lipschitz and growth conditions on b imply that on the range of integration, with $x = W$ or w

$$|b^2(x, y^{m,n}) - b^2(x, y^m)| \leqq 2K^2(1 + m) \sup_{0 \leqq u \leqq s} \left| \int_0^u \psi_\tau^{m,n} d\tau \right|$$

$$\leqq 2K^2(1 + m) \int_0^s |\psi_u^{m,n}| du,$$

where we have used $d(y^{m,n} - y^m) = \psi^{m,n} dt$. Hence, with

$$\theta_{mn}(W) = \int_0^t [b(W, y^{m,n}) - b(W, y^m)]dW,$$

we find

$$|\psi_t^{m,n+1}| \leqq m^2 \int_0^t |\psi_s^{m,n}| ds + 2K^2(1 + m) \int_0^t \int_0^s |\psi_u^{m,n}| du ds$$

$$+ 2 \int_{S(m,t)} M(dW)\theta_{mn}(W) \left[\frac{p(W, y^{m,n})_t q(W, y^{m,n})_t}{A_t^m(y^{m,n})} + \frac{p(W, y^m)_t q(W, y^m)_t}{A_t^m(y^m)} \right].$$

Since θ_{mn} is a stochastic integral, Schwarz's inequality implies that

$$\int_{S(m,t)} M(dW)\theta_{mn}(W) p(W, y^{m,n})_t q(W, y^{m,n})_t$$

$$\leqq K \left[\int_0^t \left(\int_0^s |\psi_u^{m,n}| du \right)^2 ds \right]^{\frac{1}{2}} \cdot \left[\int_{S(m,t)} M(dW) p^2(W, y^{m,n})_t q^2(W, y^{m,n})_t \right]^{\frac{1}{2}} . \quad (15)$$

To bound the second factor on the right, we use

$$q^2(f, g)_t = q(2f, g)_t \exp \int_0^t |f|^2 ds$$

and the relations

$$y_t^m = \int_0^t \hat{z}^m ds + \nu_t$$

$$\int_0^t W_s d\nu = W_t \nu_t - \int_0^t \nu_s dW$$

to find that on $S(m, t)$

$$\begin{aligned} p^2(W, y^m)_t q^2(W, y^m)_t \\ &= q[2b(W, y^m), W]_t q(2W, y^m)_t \exp \left\{ \int_0^t b^2(W, y^m) ds + \int_0^t W_s^2 ds \right\} \\ &= q[2b(W, y^m) - 2\nu, W]_t \exp \left\{ \int_0^t [b(W, y^m)_s - 2\nu_s]^2 ds \right. \\ &\quad \left. + 2\nu_t W_t + 2 \int_0^t W_s \hat{z}_s^m ds + \int_0^t W_s^2 \right\} \\ &\leqq q[2b(W, y^m) - 2\nu, W]_t \exp \left\{ 3m^2 t + 2m|\nu_t| \right. \\ &\quad \left. + \int_0^t [K(1 + m) + |\nu_s|]^2 ds \right\}. \end{aligned}$$

Since the q factor on the right integrates to 1 with respect to $M(dW)$, it can be seen that the square root of

$$\int_{S(m,t)} M(dW) p^2(W, y^m)_t q^2(W, y^m)_t$$

is bounded by a t-integrable function depending on m and ν. Since $|\hat{z}^{mn}| \leqq m$, the same result holds with y^{mn} for y^m in p and q. Also, by Jensen's inequality, with $S = S(m, t)$

$$\begin{aligned} A_t^m(y^m) &= \int_S M(dW) q[b(W, y^m), W]_t q(W, y^m)_t \\ &\geqq M\{S\} \exp \left\{ M^{-1}\{S\} \int_S \left[\int_0^t b(W, y^m) dW - \frac{1}{2} \int_0^t b^2(W, y^m) ds \right. \right. \\ &\quad \left. \left. + \int_0^t W_s \hat{z}_s^m ds + \iint_0^t W_s d\nu_s - \frac{1}{2} \int_0^t W_s^2 ds \right] \right\} \\ &\geqq M\{S\} \exp \left\{ -\tfrac{1}{2} K^2 (1 + m)^2 t - \tfrac{3}{2} m^2 t - m|\nu_t| \right. \\ &\quad \left. \cdot \int_S M(dW) \int_0^t [b(W, y^m)_s - \nu_s] dW_s \right\}. \end{aligned}$$

If $T = \inf s\colon |W_s| = m$, the integral in the exponent is bounded by the square root of

$$\int M(dW) \int_0^t \chi_{T>s}[b(W, y^m) - \nu]^2 ds \leqq \int_0^t [K(1 + m) + |\nu_s|]^2 ds.$$

The same argument applies to $A_t^m(y^{m,n})$, since $|\hat{z}^{m,n}| \leqq m$.

It follows that

$$|\psi_t^{m,n+1}| \leqq m^2 \int_0^t |\psi_s^{m,n}| ds + 2K^2(1 + m) \int_0^t \int_0^s |\psi_u^{m,n}| duds$$
$$+ F(t, m, \nu)\left[\int_0^t \left(\int_0^s |\psi_u^{m,n}| du\right)^2 ds\right]^{\frac{1}{2}},$$

where F is a t-integrable function depending only on m and ν. Thus, by arguments similar to those for Gronwall's inequality, it is seen that $\psi_t^{m,n}$ converge to zero. It follows that $\hat{z}_{m.}$ are $\nu_.$-adapted, and so is $y.$ by eq. (2).

Remark: The reader is invited to speculate on how the above proof would be carried out if it were postulated that the dependence of $b(t, z, y)$ on y in eq. (10) came only through the estimate $\hat{z}$, as, for example, $b(t, z, y) = \beta(t, z, \hat{z})$. In this case, there is no longer a formula for $\hat{z}$, but only a functional equation.

X. DISCUSSION OF THE GENERAL PROBLEM

There is a general result of measure theory to the effect that a function x is measurable on the σ-algebra induced by another function y, iff it is representable by an explicit composition with y, that is, as a function of y: $x = \varphi \circ y$. This might be called the "explicit" function theorem, as opposed to the "implicit" function theorems, like Filippov's lemma. For here x and y are given and φ is to be found, while in Filippov's lemma x and φ are given and y is to be found. We suggest that the σ-algebras and innovations problems are very close in spirit to the ideas around the explicit function theorem. This suggestion only provides what we think is a *hilfsaussichtspunkt;* without more information (about $\hat{z}$ or $a(\cdot, \cdot)$, and more insight and work, it does not help settle any particular case. What it helps do, though, is place the problems and concepts into the general framework of stochastic equations, especially into the circle of ideas developed by M. P. Yershov.[16] See also Ref. 18.

Our final results on the innovations problem will clarify the role of the integral equation relating $y.$ and $\nu.$. Thus, for each t, $\hat{z}_t$ is measurable with respect to $\sigma\{y_s, s \leqq t\}$; hence, there is a causal functional α such that $\hat{z}_t = \alpha(t, y)$, or more precisely, such that some version of $\hat{z}_t$ is indistinguishable from $\alpha(t, y)$; then the relation between $y.$ and $\nu.$ is essentially

$$y_t = \int_0^t \alpha(s, y)ds + \nu_t. \tag{16}$$

Thus, it is apparent intuitively, and can be proved, that if $y_.$ is adapted to $\nu_.$, then there is a causal solution to (4), namely $y_.$ itself, expressible as $\varphi(\cdot, \nu)$ with φ causal. What we shall show is the converse, that causal solvability somewhere of the stochastic DE (16) implies a positive answer to the Frost-Kailath conjecture. In particular, we shall prove that $\sigma\{y_s, s \leqq t\} = \sigma\{\nu_s, s \leqq t\}$ (mod P) for each t iff on some probability space there is a Brownian motion $W_.$ and a causal solution ξ of

$$\xi_t = \int_0^t \alpha(s, \xi)ds + W_t, \tag{17}$$

which induces the same measure as $y_.$ does.

This result gives a necessary and sufficient condition for Frost and Kailath's innovations conjecture to hold, and it embodies the sense in which the innovations problem resembles the explicit function theorem. The direct part or necessity is obvious. For the sufficiency, we argue that if eq. (17) has a causal solution on some probability space, then it is expressible as a causal functional φ of a Brownian motion defined there. This functional can be "exported"; i.e., it can be applied to any other Brownian motion on any other space to give a causal solution. In particular, applying it to the innovations process $\nu_.$ gives a causal solution $(\varphi\nu)_t = \varphi(t, \nu)$, which under weak conditions induces the same measure as $y_.$ does. This, along with the properties

$$(\varphi\nu)_t - \int_0^t \alpha(s, \varphi\nu)ds \triangleq (T\varphi\nu)_t = \nu_t, \quad \text{a.s.} \tag{18}$$

$$\sigma\{\nu_s, s \leqq t\} \subseteq \sigma\{y_s, s \leqq t\}, \tag{19}$$

allows us to prove the basic property that for any integrable causal functional $\beta(t, y)$

$$E\{\beta(t, y) \,|\, \nu_s, s \leqq t\} = \beta(t, \varphi\nu) \quad \text{a.s.}$$

This result can be applied in several ways to give the desired final result that $y_.$ and $\varphi\nu_.$ are modifications of each other. We describe two—one a digressive application of martingales and the other short and direct.

XI. MARTINGALE ARGUMENTS USING P_0

In this section, we use some of the properties of the measure P_0 defined in Section VIII. We assume that α is a causal functional such that $\hat{z}_t = \alpha(t, y)$; i.e., really such that $\hat{z}_t$ and $\alpha(t, y)$ are indistinguishable processes. We let $\mathcal{C}_t$ be the Borel σ-subalgebra of $C[0, \infty)$ generated by sets of the form $\{x : x_s \in B\}$, $s \leqq t$, B Borel, and $\mathfrak{N}_0^t$ is $\sigma\{\nu_s, 0 \leqq s \leqq t\}$, the σ-algebra generated by the innovations.

Theorem 5: If there is, on some probability space, a Brownian motion W and a causal solution ξ of

$$d\xi = \alpha(t, \xi)ds + W_t$$

given by a causal functional φ as $\xi_t = \varphi(t, W)$, and if $\varphi\nu$ and y are identical in law, then they are modifications of each other, and $\mathfrak{N}_0^t = \mathcal{Y}_0^t$ (mod P).

The proof is the sequence of lemmas which follow.

Lemma 1: $E\{(dP_0/dP) \mid \mathfrak{N}_0^t\} = q^{-1}(\alpha\varphi\nu, \varphi\nu)_t$.

Proof: Let $A \in \mathfrak{N}_0^t$, so that by the integral equation, A differs from a set of the form $\{Ty \in B\}$, $B \in \mathcal{C}_t$, by at most a null set. Then,

$$\int_A dP_0 \int E\left\{\frac{dP_0}{dP} \mid \mathcal{Y}_0^t\right\} dP = \int_{y \in T^{-1}B} q^{-1}(\hat{z}, y)_t dP$$

$$= \int_{\varphi\nu \in T^{-1}B} q^{-1}(\alpha\varphi\nu, \varphi\nu)_t dP = \int_{\nu \in B} q^{-1}(\alpha\varphi\nu, \varphi\nu) dP.$$

Thus, $q^{-1}(\alpha\varphi\nu, \varphi\nu)_t$ is measurable on $\mathfrak{N}_0^t$ and has the same integrals over $\mathfrak{N}_0^t$ sets as dP_0/dP; thus, it is a version of $E\{(dP_0/dP) \mid \mathfrak{N}_0^t\}$.

Lemma 2: *$\varphi\nu_t$ is a $\mathfrak{N}_0^t$-martingale under P_0.*

Proof: $y.$ is a (Brownian) martingale under P_0; since $\mathfrak{N}_0^t \subseteq \mathcal{Y}_0^t$, then

$$\int_A y_t dP_0 = \int_A y_s dP_0 \quad \text{if} \quad A \in \mathfrak{N}_0^s \quad \text{and} \quad s < t.$$

For $A \in \mathfrak{N}_0^s$, eq. (2) implies that there exists $B \in \mathcal{C}_s$ with $A = \{Ty \in B\}$. Since $y \sim \varphi\nu$ in law, there follows

$$\int_A y_t dP_0 = \int_{Ty \in B} y_t q^{-1}(z, y) = \int_{Ty \in B} y_t E\{q^{-1}(z, y)_t \mid \mathcal{Y}_0^t\} dP$$

$$= \int_{Ty \in B} y_t q^{-1}(\hat{z}, y)_t dP = \int_{T\varphi\nu \in B} (\varphi\nu)_t q^{-1}(\alpha\varphi\nu, \varphi\nu)_t dP$$

$$= \int_{\nu \in B} (\varphi\nu)_t E\left\{\frac{dP_0}{dP} \mid \mathfrak{N}_0^t\right\} dP = \int_A (\varphi\nu)_t dP_0.$$

Similarly, for $s < t$ and $A \in \mathfrak{N}_0^s$,

$$\int_A y_s dP_0 = \int_A (\varphi\nu)_s dP_0.$$

Hence, $A \in \mathfrak{N}_0^s$ implies $\int_A [\varphi\nu_t - \varphi\nu_s]dP_0 = 0$, and so $\varphi\nu.$ is a $(\mathfrak{N}_0^{\cdot}, P_0)$-martingale, being adapted to $\mathfrak{N}_0^{\cdot}$.

Lemma 3: $E_0\{y_t | \mathfrak{N}_0^t\} = \varphi\nu_t$ *a.s.*

Proof: This is shown in the same way as Lemmas 1 and 2, by integrating over $A \in \mathfrak{N}_0^t$, and using the adaptedness of ν to y, and the property $T\varphi\nu = \nu$ a.s.

Lemma 4: $\varphi\nu.$ *is a Brownian motion under* $(P_0, \mathfrak{N}_0^{\cdot})$.

Proof: Let $\Delta_s = \hat{z}_s - \alpha(s, \varphi\nu)$

$$y_t - \varphi\nu_t = \int_0^t \Delta_s ds$$

$$(y_t - \varphi\nu_t)^2 = 2\int_0^t \Delta_s \int_0^s \Delta_u du ds.$$

However, since $y.$ and $\varphi\nu.$ are P_0-martingales on $\mathcal{Y}_0^{\cdot}$ and $\mathfrak{N}_0^{\cdot}$, respectively, the change of variables formula applied to each separately gives

$$y_t^2 = 2\int_0^t y_s dy_s + t \tag{20}$$

$$\varphi\nu_t^2 = 2\int_0^t \varphi\nu_s d\varphi\nu_s + \langle\varphi\nu\rangle_t. \tag{21}$$

Also

$$y_t\varphi\nu_t = y_t^2 - y_t\int_0^t \Delta_s ds.$$

The right-hand side is a product of semi-martingales on $\mathcal{Y}_0^{\cdot}$, and change of variables gives

$$y_t\varphi\nu_t = 2\int_0^t y_s dy_s + t - \int_0^t \Delta_s dy_s - \int_0^t y_s\Delta_s ds.$$

Using eqs. (20) and (21), we find that

$$(y_t - \varphi\nu_t)^2 = \langle\varphi\nu\rangle_t - t + 2\int_0^t\int_0^s \Delta du ds$$
$$- 2\int_0^t \left[y_s - \int_0^s \Delta du - \varphi\nu_s\right] dy_s$$

$$\langle\varphi\nu\rangle_t = t.$$

Thus, $\varphi\nu.$ is a continuous martingale with quadratic variation t, and so a Brownian motion, on $(P_0, \mathfrak{N}_0^{\cdot})$.

Lemma 5: $E_0(y_t - \varphi\nu_t)^2 = 0$.

Proof: The processes $y.$ and $\varphi v.$ are Brownian under P_0 with respect to $\mathcal{Y}_0^{\cdot}$ and $\mathfrak{N}_0^{\cdot}$, respectively, so $E_0 y_t^2 = E_0\varphi\nu_t^2 = t$. Thus, by Lemma 3

$$
\begin{aligned}
E_0(y_t - \varphi\nu_t)^2 &= 2(t - E_0 y_t \varphi\nu_t) \\
&= 2(t - E_0 E_0\{y_t | \mathfrak{N}_0^t\} \varphi\nu_t) \\
&= 2(t - E\langle\varphi\nu\rangle_t) \\
&= 0.
\end{aligned}
$$

The indicated expectation exists because both $y_.$ and $\varphi\nu_.$ are Brownian under P_0, on respective algebras $\mathfrak{Y}_0^{\cdot}$ and $\mathfrak{N}_0^{\cdot}$. This lemma shows that y and $\varphi\nu$ are modifications of each other under P_0 (and so under P), and completes the proof of Theorem 5.

XII. DIRECT PROOF OF THEOREM 5

It is possible to give a short proof of Theorem 5 not depending on the auxiliary measure P_0 or the representation for $\hat{z}$ given in Section VIII. This proof depends only on eqs. (18) and (19), the causality of φ, and the fact that $y \sim \varphi\nu$ in law; otherwise, it is just an exercise in integration.

By hypothesis there exists a causal functional φ, such that $T\varphi x = x$ for almost all x with respect to Wiener measure. Thus, the process $(\varphi\nu)_t = \varphi(t, \nu)$ (defined on the same probability space as $y_.$ and $\nu_.$) is identical in law to $y_.$ such that $T\varphi\nu = \nu$ with probability one. Let β be a causal functional such that $E|\beta(t, y)| < \infty$ for each t. The next step is to prove that

$$E\{\beta(t, y) | \nu_s, 0 \leqq s \leqq t\} = \beta(t, \varphi\nu), \quad \text{a.s.}$$

Let then $A \in \sigma\{\nu_s, 0 \leqq s \leqq t\}$. A has the form $\{\omega : \nu \in B\}$ with B a Borel set of $C[0, t]$, so by the integral equation it differs from $\{\omega : y \in T^{-1}B\}$ by at most a null set. Then, since $\varphi\nu$ and y are identical in law, and $\varphi \subseteq T^{-1}$ on a set of Wiener measure one, we find

$$
\begin{aligned}
\int_A \beta(t, y) dP &= \int_{\{y \in T^{-1}B\}} \beta(t, y) dP = \int_{\{\varphi\nu \in T^{-1}B\}} \beta(t, \varphi\nu) dP \\
&= \int_{\{T\varphi\nu \in B\}} \beta(t, \varphi\nu) dP = \int_{\{\nu \in B\}} \beta(t, \varphi\nu) dP \\
&= \int_A \beta(t, \varphi\nu) dP.
\end{aligned}
$$

Thus, $\beta(t, \varphi\nu)$ has the same integrals as $\beta(t, y)$ over sets defined by $\nu_.$ over $[0, t]$, and (since φ is causal) is measurable on $\sigma\{\nu_s, 0 \leqq s \leqq t\}$. Hence, it is a version of $E\{\beta(t, y) | \nu_s, 0 \leqq s \leqq t\}$. To complete the proof, let $\beta(t, x) = \alpha^+(t, x)^{\frac{1}{2}}$, where $a^+ = \max\{0, a\}$, to find, since $y \sim \varphi\nu$ in law, that

$$
\begin{aligned}
&E|\alpha^+(t, y)^{\frac{1}{2}} - \alpha^+(t, \varphi\nu)^{\frac{1}{2}}|^2 \\
&\qquad = 2E\alpha^+(t, y) - 2E\alpha^+(t, \varphi\nu)^{\frac{1}{2}} E\{\alpha^+(t, y)^{\frac{1}{2}} | \nu_s, 0 \leqq s \leqq t\} = 0.
\end{aligned}
$$

Similarly, $E|\alpha^{-}(t, y)^{\frac{1}{2}} - \alpha^{-}(t, \varphi\nu)^{\frac{1}{2}}|^2 = 0$. Hence,

$$\alpha(\cdot, y) = \alpha(\cdot, \varphi\nu) \quad \text{a.s.} \quad \lambda \times P \quad (\lambda = \text{Lebesgue measure}),$$

so that for each t, by the integral equations,

$$y_t - (\varphi\nu)_t = \int_0^t [\alpha(s, y) - \alpha(s, \varphi\nu)]ds = 0 \quad \text{a.s.}$$

Since φ is causal, it follows that y_t is equal almost surely to a function measurable on $\sigma\{\nu_s, s \leqq t\}$. Since this is true for each t, it follows that for each t the algebras $\sigma\{\nu_s, s \leqq t\}$ and $\sigma\{y_s, s \leqq t\}$ are equal (mod P).

Remark 1: Since $\int_0^t \hat{z}^2 ds < \infty$ a.s., then if also

$$\int_0^t \alpha(s, \varphi\nu)^2 ds < \infty \quad \text{a.s.},$$

a theorem of Kailath and Zakai will imply that the respective measures induced by y and $\varphi\nu$ are each absolutely continuous with respect to Wiener measure with the same Radon–Nikodým derivative $q[\alpha(x), x]$. Hence, $y \sim \varphi\nu$ in law, as desired for the hypothesis of Theorem 5. Thus, the condition that y and $\varphi\nu$ induce the same measure is easily met, in comparison with the difficulty of finding a causal solution.

Remark 2: The condition in Theorem 5 that ξ be causal can be replaced, if we are content to work only over a finite interval $[0, T]$, by the conditions that ξ be a strong solution over $[0, T]$ in the sense of Ref. 16, and that it be nonanticipative. For it has been remarked by Yershov[6] that a strong (over $[0, T]$) nonanticipative solution is necessarily causal.

XIII. ACKNOWLEDGMENTS

The author is indebted to P. Frost, T. T. Kadota, T. Kailath, H. P. McKean, Jr., L. A. Shepp, and M. P. Yershov for useful discussions regarding the innovations and σ-algebras problems, and to the manuscript reviewer for several valuable comments and suggestions.

REFERENCES

1. P. A. Meyer, "Sur un problème de filtration, Seminaire de Probabilités VII," Springer Lecture Notes 321 (1973), pp. 223–247.
2. P. Frost, "Estimation in Continuous-Time Nonlinear Systems," dissertation, Stanford University, Stanford, Calif., June 1968.
3. T. Kailath, "Some Extensions of the Innovations Theorem," B.S.T.J., *50* (April 1971), pp. 1487–1494.
4. Y. Rozanov, personal communication.
5. S. Orey, "Radon-Nikodým Derivatives of Probability Measures: Martingale Methods." Notes from lectures at Tokyo University of Education, Spring 1974.
6. M. P. Yershov, personal communication.

7. J. M. C. Clark, "Conditions for the One-to-One Correspondence Between an Observation Process and Its Innovation." Tech. Report 1 (1969), Imperial College, London, England.
8. V. E. Beneš, "Extension of Clark's Innovations Equivalence Theorem to the Case of Signal z. Independent of Noise, with $\int_0^t z_s^2 ds < \infty$ a.s.," unpublished work.
9. M. Hitsuda, "Representation of Gaussian Processes Equivalent to Wiener Process," Osaka J. Math., *5*, No. 2 (December 1968), pp. 299–312.
10. A. K. Zvonkin, "Transformation of the Phase Space of a Diffusion Process, Annihilating the Drift" (in Russian), Matematicheskii Sbornik, *93*, No. 1 (January 1974), pp. 129–149.
11. T. Kailath and M. Zakai, "Absolute Continuity and Randon-Nikodým Derivatives for Certain Measures Relative to Wiener Measure," Ann. Math. Stat., *42* (February 1971), pp. 130–140.
12. S. Watanabe, personal communication.
13. Cf. e.g. Orey, loc. cit., p. 7, Theorem 1.
14. R. Sh. Liptser and A. N. Shiryaev, "Statistics of Random Processes" (in Russian), Izdat, "Nauka," Moscow, 1974.
15. G. Kallianpur and C. Striebel, "Estimation of Stochastic Processes: Arbitrary System Processes With Additive White Noise Errors," Ann. Math. Stat., *39* (June 1968), pp. 785–801.
16. M. P. Yershov, "Extension of measures and stochastic equations" (in Russian), Teoriya Veroiatnostei i ee Prim., *19*, No. 3 (July, August, September 1974), pp. 457–471.
17. B. S. Cirel'son, "An Example of a Stochastic Differential Equation Not Possessing a Strong Solution" (in Russian), Teoriya Veroiatnostei i ee Prim., *20* (March–April 1975), pp. 427–430.
18. V. E. Beneš, "Weak and Strong Solutions for Stochastic Equations," unpublished work.