TOWARD A THEORY OF NONL INEAR STOCHASTIC REALIZATION

Anders Lindquist*, Sanjoy Mitter ${ }^{\dagger}$, and Giorgio Picci ${ }^{\S}$

Feedback and Synthesis of Linear and Nonlinear Systems Proceedings of the Workshop in
Bielefeld, West-Germany, June 22-26, 1981, and Rome, Italy, June 29-July 3, 1981. Edited by A. Isidori and D. Hinrichsen Springer-Verlag Lecture Notes

[^0]§ LADSEB-CNR, Corso Stati Uniti 4, 35100 Padova, Italy.

TOWARD A THEORY OF NONLINEAR STOCHASTIC REALIZATION
Anders Lindquist*, Sanjoy Mitter ${ }^{\dagger}$, and Giorgio Picci ${ }^{\S}$

## 1. INTRODUCTION

The following is a central problem in stochastic systems theory: Given a stationary stochastic process $\{y(t) ; t \in \mathbb{R}\}$, find a (possibly infinite-dimensional) vector Markov process $\{x(t) ; t \in \mathbb{R}\}$, called the state process, and a function $f$ so that $v(t)=$ $f(x(t))$ for all $t \in \mathbb{R}$. Moreover, find a stochastic differential equation driven by a Wiener process and having the state process $x$ as its unique solution. The problem of characterizing the family of all such representations is known as the stochastic reaitzation problem.

There is by now a rather comprehensive theory of stochastic realization for the case that $\{y(t) ; t \in \mathbb{R}\}$ is Goussian [1-3], in which case the representations can be taken to be linecr, i.e. both $f$ and the stochastic differential equation are linear. This linear theory can be applied to non-Goussian processes also, but then we need to give up the requirement that $x$ is Markov and that it is generated by a Wiener process, replacing these concepts by "wide sense Markov" [4] and "orthogonal increment process" respectively. If we are not willing to do so, a noniinear stochastic realization theory is needed. That is the topic of this paper.

In this paper we shall apply Wiener's theory of homogeneous chaos [5,6] to the nonlinear stochastic realization problem. For simplicity and ease of notation we shall assume that the process $y$ is scalar, although the machinary which we develop is sufficient to accomnodate also the vector case. Other assumptions, such as $y$ admitting an innovation representation, are however crucial to our approach. (In this respect, it might be more appropriate to consider a process $y$ with stationary increments, and indeed with minor modifications we could have done so.) In the extension of this work we see the possibility of making contact with nonlinear filtering $[7,8]$ and that is partially a motivation for this work.

## 2. PROBLEM FORMULATION

Let $\{y(t) ; t \in \mathbb{R}\}$ be a non-Gaussian stationary stochastic process which is meansquare continuous, purely nondeterministic, and centered, and let $y$ be the sigma-field

[^1]generated by $y$. Then define $H$ to be the Hilbert space of all centered $Y$-measurable random variables, having inner product $\langle\xi, \eta\rangle=E\{\xi \eta\}$. Since $y$ is stationary there is a strongly continuous group of unitary operators $\left\{U_{t} ; t \in \mathbb{R}\right\}$ on $H$, called the shift, such that $y(t+s)=U_{t} y(s)$ for all $t$ and $s$ [9]. Let $y_{t}^{-}$and $y_{t}^{+}$be the sigma-fields generated by $\{y(s) ; s \leq t\}$ and $\{y(s) ; s \geq t\}$ respectively.

Next assume that $y$ has an innovation process $\{v(t) ; t \in \mathbb{R}\}$, by which we shall here mean a Wiener process such that $\sigma\{\nu(\tau)-\nu(\sigma) ; \tau, \sigma \leq t\}=y_{t}^{-}$. (Here $\sigma\{\cdot\}$ denotes the sigmafield generated by the random variables inside the curly brackets.) Then, by symmetry, it also has a backward innovation process $\{\bar{v}(t) ; t \in \mathbb{R}\}$, i.e. another Wiener process such that $\sigma\{\bar{\nu}(\tau)-\bar{\nu}(\sigma) ; \tau, \sigma \geq t\}=y_{t}^{+}$. Now, since $y=\sigma\{\nu(t) ; t \in \mathbb{R}\}=\sigma\{\bar{\nu}(t) ; \tau \in \mathbb{R}\}$ we can apply Wiener's homogeneous chaos theory [5,6]. Let $H_{1}$ denote the Gaussian space [5] generated by $\{v(t) ; t \in \mathbb{R}\}$ or, which is equivalent, by $\{\bar{v}(t) ; t \in \mathbb{R}\}$. Since $y$ is mean-square continuous, $H_{1}$ is a separable space, and therefore $H_{1}$ has a countable orthonormal basis $\left\{\xi_{i}\right\}_{i=0}^{\infty}$. Now, let $P_{n}$ be the (closed) linear subspace of all polynomials in $\left\{\xi_{i}\right\}_{i=0}^{\infty}$ of degree not exceeding $n$. Next define $H_{n}=P_{n} \ominus P_{n-1}$, i.e. the orthogonal complement of $P_{n-1}$ in $P_{n}$. Then it can be shown $[5,6]$ that

$$
\begin{equation*}
\mathrm{H}=\mathrm{H}_{1} \oplus \mathrm{H}_{2} \oplus \mathrm{H}_{3} \oplus \ldots \tag{1}
\end{equation*}
$$

where denotes orthogonal direct sum. The space $H_{n}$ is called the $n^{t h}$ homogeneous chaos of $H$. Since $y(0) \leqslant H$, there is an orthogonal decomposition

$$
\begin{equation*}
y(0)=y_{1}(0)+y_{2}(0)+y_{3}(0)+\ldots \tag{2}
\end{equation*}
$$

where $y_{n}(0) \in H_{n}$. It is easy to see that each chaos $H_{n}$ is invariant under the shift $U_{t}$, and consequently, for any $t \in \mathbb{R}$, we have a decomposition such as (2) for $y(t)$ in terms of $y_{n}(t):=U_{t} y_{n}(0), n=1,2,3 \ldots$

In order to obtain a state space description we introduce a past space $\mathrm{H}^{-}$and a future space $\mathrm{H}^{+}$as follows. Let $\mathrm{H}^{-}\left(\mathrm{H}^{+}\right)$be the subspace of all centered $\mathrm{y}^{-}$-measurable ( $y^{+}$-measurable) random variables. Then, defining $H_{1}^{-}$and $H_{1}^{+}$to be the Gaussian spaces generated by $\{\nu(\tau)-\nu(\sigma) ; \tau, \sigma \leq 0\}$ and $\{\bar{\nu}(\tau)-\bar{v}(\sigma) ; \tau, \sigma \geq 0\}$ respectively, we obtain the chaos expansions

$$
\left\{\begin{array}{l}
\mathrm{H}^{-}=\mathrm{H}_{1}^{-} \oplus \mathrm{H}_{2}^{-} \oplus \mathrm{H}_{3}^{-} \oplus \ldots  \tag{ja}\\
\mathrm{H}^{+}=\mathrm{H}_{1}^{+} \oplus \mathrm{H}_{2}^{+} \oplus \mathrm{H}_{3}^{+} \oplus \ldots
\end{array}\right.
$$

where clearly $H_{n}^{-} \subset H_{n}$ and $H_{n}^{+} \subset H_{n}$ for all $n$. Note that $H_{1}^{-} \cap H_{1}^{+} \neq \emptyset$ and $H_{1}^{-} \vee H_{1}^{+}=H_{1}$, but $\mathrm{H}^{-} \vee \mathrm{H}^{+} \neq \mathrm{H}$.

Now, if $y$ were a Gaussian process, $y$ would have a component only in the first chaos, i.e. $y=y_{1}$, and consequently state spaces for $y$ could be constructed along the lines of $[1,2]$ by finding the minimal Markovian ( $\mathrm{H}_{1}^{-}, \mathrm{H}_{1}^{+}$)-splitting subspaces in $\mathrm{H}_{1}$. [We recall that, for two subspaces $A$ and $B, X$ is an ( $A, B$ )-splitting subspace if $\left\langle E^{X} \alpha, E_{\beta} X_{\beta}=\langle\alpha, \beta\rangle\right.$ for all $\alpha \in A$ and $\beta \in B$, where $E^{X}$ denotes orthogonal projection on the subspace X.] However, for a non-Gaussian process $y$, there will be some nontrivial component $y_{n}, \mathrm{n}>1$, and consequently the state space construction will have to involve at least those higher chaoses in which $y$ has a component. To this end define the index set
$N:=\left\{n \mid y_{n}(0)=0\right\} \cup\{1\}$. For reasons which will soon be evident, we shall have to always include the first chaos in our analysis. (In particular, see Section 7.)

Hence we call $X$ a state space for $y$ if

$$
\begin{equation*}
X=\underset{n \in N}{\oplus} X_{n}, \tag{4}
\end{equation*}
$$

where $X_{n} \subset H_{n}$ is an $\left(H_{n}^{-}, H_{n}^{+}\right)$-splitting subspace, and $X$ is Markovian in the sense that, if $X:=\sigma\{X\}, X^{-}:=\sigma\left\{V_{t \leq 0} U_{t} X\right\}$ and $X^{+}:=\sigma\left\{V_{t \geq 0} U_{t} X\right\}, X^{-}$and $X^{+}$are conditionally independent given $X$; we shall write this $X^{-} \Perp X^{+} \mid X$. We say that $X$ is minimal if there is no other state space $X^{\prime}$ for which $X^{\prime}:=\sigma\left\{X^{\prime}\right\}$ is properly contained in $X$.

The problem at hand is now to construct all minimal state spaces for $y$ and to obtain a dynamical representation (realization) for each of them.

## 3. THE STRUCTURE OF $H$

According to Itô's Theorem [10]

$$
\begin{equation*}
H_{n}=\left\{I_{n}(f ; v) \mid f \in \hat{L}_{2}\left(\mathbb{R}^{n}\right)\right\} \tag{5}
\end{equation*}
$$

where $I_{n}$ is the multiple Wiener integral

$$
\begin{equation*}
I_{n}(f ; v)=\int_{-\infty}^{\infty} \int_{-\infty}^{t_{1}} \ldots \int_{-\infty}^{t_{n-1}} f\left(t_{1}, t_{2}, \ldots, t_{n}\right) d v\left(t_{1}\right) \ldots d v\left(t_{n}\right) \tag{6}
\end{equation*}
$$

and $\hat{L}_{2}\left(\mathbb{R}^{n}\right)$ are the symmetric functions in $L_{2}\left(\mathbb{R}^{n}\right)$. Although the region of integration is such that (5) remains the same if $\hat{L}_{2}\left(\mathbb{R}^{n}\right)$ is exchanged for $L_{2}\left(\mathbb{R}^{n}\right)$, we prefer the former since we have a one-one correspondence between eiements in $H_{n}$ and $\hat{L}_{2}\left(R^{n}\right)$. In fact, we can establish an isometric isomorphism between these spaces $[5,6,10]$. Now, for $i=1,2, \ldots, n$, let $\eta_{i} \in H_{1}$ be arbitrary. Then there exist unique functions $f_{i} \in L_{2}(R)$ such that $\eta_{i}=\int_{-\infty}^{\infty} \hat{F}_{i}(t) d v(t)$. Next define
where

$$
\begin{equation*}
\eta_{1} * \eta_{2} * \ldots * \eta_{n}=n!I_{n}(f ; \nu) \tag{7}
\end{equation*}
$$

$$
\begin{equation*}
f\left(t_{1}, t_{2}, \ldots, t_{n}\right)=\frac{1}{n!} \sum_{\pi \in G} f_{\pi_{1}}\left(t_{1}\right) f_{\pi_{2}}\left(t_{2}\right) \ldots f_{\pi_{n}}\left(t_{n}\right) \tag{8}
\end{equation*}
$$

$G$ being the symmetric group of permutations of $n$ letters. Since finite linear combinations of functions of type (8) are dense in $\hat{L}_{2}\left(\mathbb{R}^{n}\right)$, Ito's Theorem implies that

$$
\begin{equation*}
H_{n}=\overline{s p}\left\{\eta_{1} * \eta_{2} * \ldots * \eta_{n} \mid \eta_{1}, \eta_{2}, \ldots, \eta_{n} \in H_{1}\right\} \tag{9a}
\end{equation*}
$$

where $\overline{S p}$ denotes closed linear hull. We shall write this as

$$
\begin{equation*}
H_{n}=H_{1} * H_{1} * \ldots * H_{1} \tag{9b}
\end{equation*}
$$

By Itô's formula [11; p.38]

$$
\begin{align*}
\eta_{1} * \eta_{2} * \ldots * \eta_{n} & =\left(\eta_{2} * \eta_{3} * \ldots * \eta_{n}\right) \cdot \eta_{1} \\
& -\sum_{k}\left(\eta_{2} * \ldots * \eta_{k-1} * \eta_{k+1} * \ldots * \eta_{n}\right) \cdot\left\langle\eta_{1}, \eta_{k}\right\rangle \tag{10}
\end{align*}
$$

which can be solved recursively. For example,

$$
\begin{aligned}
& n_{1} * n_{2}=\eta_{1} n_{2}-<n_{1}, n_{2}> \\
& n_{1} * n_{2} * \eta_{3}=n_{1} n_{2} n_{3}-n_{1} \cdot\left\langle n_{2}, n_{3}>-n_{2} \cdot<\eta_{1}, n_{3}>-n_{3} \cdot\left\langle\eta_{1}, n_{2}>\right.\right.
\end{aligned}
$$

The *-operation is obviously commutative. In particular,

$$
\begin{equation*}
\underbrace{\eta * \eta * \ldots * \eta}=h_{n}\left(\eta,\langle\eta, \eta>\rangle^{\frac{1}{2}}\right) \tag{11}
\end{equation*}
$$

n times
(in the sequel we shall write this $\eta^{n *}$ ) where

$$
\begin{equation*}
h_{n}(x, \sigma)=\frac{(-\sigma)^{n}}{n!} \exp \left(\frac{x^{2}}{2 \sigma}\right) \frac{\partial^{n}}{\partial x^{n}} \exp \left(-\frac{x^{2}}{2 \sigma}\right) \tag{12}
\end{equation*}
$$

$\mathrm{n}=0,1,2, \ldots$, are the Hermite polynomials (cf [11; p.37]). Analogously to (9) we have

$$
\left\{\begin{array}{l}
\mathrm{H}_{\mathrm{n}}^{-}=\mathrm{H}_{1}^{-} * \mathrm{H}_{1}^{-} * \ldots * \mathrm{H}_{1}^{-}  \tag{13a}\\
\mathrm{H}_{\mathrm{n}}^{+}=\mathrm{H}_{1}^{+} * \mathrm{H}_{1}^{+} * \ldots * \mathrm{H}_{1}^{+}
\end{array}\right.
$$

Let $H_{1}^{\mathrm{n} \odot}=\mathrm{H}_{1} \odot \mathrm{H}_{1} \odot \ldots \odot \mathrm{H}_{1}$ denote the symmetric tensor-product Hilbert space of $\mathrm{H}_{1}$ by itself taken $n$ times. Then for arbitrary $\xi_{i}, \eta_{i} \in H_{1}, i=1,2, \ldots, n$, with $<\cdot, \cdot>_{\text {no }}$ the inner product in $\mathrm{H}_{1}^{\mathrm{no}}$, we have

$$
\begin{equation*}
\left.<\xi_{1} \circlearrowleft \xi_{2} \odot \ldots \odot \xi_{n}, \eta_{1} \odot \eta_{2} \odot \ldots \odot \eta_{n}\right\rangle=\frac{1}{n!} \sum_{\pi \in G}<\xi_{\pi_{1}}, \eta_{1}><\xi_{\pi_{2}}, \eta_{2}>\ldots<\xi_{\pi_{n}}, \eta_{n}> \tag{14}
\end{equation*}
$$

where $\xi_{1} \odot \xi_{2} \odot \ldots \odot \xi_{n}$ is the symmetric tensor product $[5,6]$. Since finite linear combinations of such tensor products are dense in $H_{I}$, it is now easy to see that $H_{l}^{\mathrm{nO}}$ is isometrically isomorphic to $\hat{L}_{2}\left(\mathbb{R}^{n}\right)$ and hence to $H_{1}^{n *}$. For $n=2$ we can illustrate this by factoring the symmetric bilinear map $\left(\eta_{1}, \eta_{2}\right) \rightarrow \eta_{1} * \eta_{2}$ as follows:

where $\phi_{2}$ is the unique linear map which makes the diagram commute; $\phi_{2}$ is unitary. Similar unitary maps $\phi_{n}$ are defined for $n=3,4, \ldots$

If $A_{1}, A_{2}, \ldots$, are linear operators in $H_{1}$, we define $A_{1} * A_{2} * \ldots * A_{n}: H_{n} \rightarrow H_{n}$ via

$$
\left(A_{1} * A_{2} * \ldots * A_{n}\right)\left(\eta_{1} * \eta_{2} * \ldots * \eta_{n}\right)=\left(A_{1} \eta_{1}\right) *\left(A_{2} n_{2}\right) * \ldots *\left(A_{n} \eta_{n}\right)
$$

on a dense set in $H_{n}$ and then extend it continuously to all of $H_{n}$. We define $A_{1} \odot A_{2} \circ \ldots \because A_{n}: H_{1}^{n ø} \rightarrow H_{1}^{n \odot}$ analogously. For $n=2$ we have then the following picture:

and analogously for $n>2$.

## 4. STATE SPACE CONSTRUCTION

THEOREM 1. The subspace $\mathrm{X} \in \mathrm{H}$ is a minimal state space for y if and only if

$$
\begin{equation*}
x=\operatorname{m}_{n \in N} x_{n} \tag{15a}
\end{equation*}
$$

where $X_{1}$ is a minimal Markovian $\left(H_{1}^{-}, H_{1}^{+}\right)$-splititing subspace and

$$
\begin{equation*}
\mathrm{x}_{\mathrm{n}}=\mathrm{x}_{1}^{\mathrm{n} *}:=\mathrm{x}_{1} * \mathrm{x}_{1} * \ldots * \mathrm{x}_{1} \tag{15b}
\end{equation*}
$$

( n times). Then each $\mathrm{X}_{\mathrm{n}}$ is a minimal $\left(\mathrm{H}_{\mathrm{n}}^{-}, \mathrm{H}_{\mathrm{n}}^{+}\right.$) -splitting subspace.
The proof of this theorem is based on the following lemmas.
LEMMA 1. Let $n=n_{1} \odot n_{2} \circlearrowleft \ldots \odot n_{n}$ where $n_{i} \in H_{1}$ for $i=1,2, \ldots, n$. Let $X$ be a subspace of $\mathrm{H}_{1}$. Then

$$
E^{X o x} \ldots \otimes x_{n}=\left(E_{n_{1}}\right) \bullet\left(E_{n_{2}}\right) \bullet \ldots \odot\left(E_{\eta_{2}}^{X_{n}}\right) .
$$

PROOF. Let $\hat{\eta}_{i}:=\mathrm{E}_{\eta_{i}}$, and let $\xi=\xi_{1} \odot \xi_{2} \odot \ldots \odot \xi_{\mathrm{n}}$ where $\xi_{1}, \xi_{2}, \ldots, \xi_{\mathrm{n}}$ are arbitrary elements in $X$. Then, by (14),

$$
\begin{aligned}
& \left\langle n_{1} \oplus n_{2} \oplus \ldots \odot n_{n}-\hat{n}_{1} \oplus \hat{n}_{2} \oplus \ldots \theta \hat{n}_{n}{ }_{n \theta}=\right. \\
& \left\langle\left(n_{1}-\hat{n}_{1}\right) \oplus n_{2} \oplus \ldots \odot n_{n}, \xi\right\rangle+\left\langle\hat{n}_{1} \otimes\left(n_{2}-\hat{n}_{2}\right) \oplus \ldots \odot n_{n}, \xi\right\rangle+\ldots+\left\langle\hat{n}_{1} \otimes \hat{n}_{2} \otimes \ldots \odot\left(n_{n}-\hat{n}_{n}\right), \xi\right\rangle= \\
& \frac{1}{n!} \sum_{\pi \in G}\left\{<n_{1}-\hat{n}_{1}, \xi_{\pi_{1}}><n_{2}, \xi_{\pi_{2}}>\ldots<n_{n} \xi_{\pi_{n}}>+\left\langle n_{1}, \xi_{\pi_{1}}><n_{2}-\hat{n}_{2}, \xi_{\pi_{2}}>\ldots<n_{n}, \xi_{\pi_{n}}>+\right.\right. \\
& \ldots+\left\langle n_{1}, \xi_{\pi_{1}}><\eta_{2}, \xi_{\pi_{n}}>\ldots<n_{n}-\hat{\eta}_{n}, \xi_{\pi_{n}}>\right\},
\end{aligned}
$$

which equals zero since $\eta_{i}-\hat{n}_{i} \perp X$.
LEMMA 2. Let $X_{1}$ be a subspace of $H_{1}$, and let $X_{n}$ be defined by (15b). Then $X_{1}$ is an $\left(\mathrm{H}_{1}^{-}, \mathrm{H}_{1}^{+}\right)$-splitting subspace if and only if $\mathrm{X}_{\mathrm{n}}$ is an $\left(\mathrm{H}_{\mathrm{n}}^{-}, \mathrm{H}_{\mathrm{n}}^{+}\right)$-splitting subspace.
PROOF. Due to isomorphism, we can identify $H_{n}^{-}, H_{n}^{+}$and $X_{n}$ with $\left(H_{1}^{-}\right)^{n \odot},\left(H_{1}^{+}\right)^{n \theta}$ and $X_{1}^{n \Theta}$ respectively. But Lemma 1 , (14), and the definition of splitting subspace imply that $\mathrm{H}_{1}^{-} \perp \mathrm{H}_{1}^{+} \mid \mathrm{X}$ if and only if $\left(\mathrm{H}_{1}^{-}\right)^{\mathrm{n} \Theta} \perp\left(\mathrm{H}_{1}^{+}\right)^{\mathrm{n} \Theta} \mid \mathrm{X}_{1}^{\mathrm{n} \Theta}$. Hence the lemma follows.

LEMMA 3. Let $X_{1}$ and $X_{n}$ be the splitting subspaces of Lemma 2. Then $X_{n}$ is minimal if and only if $X_{1}$ is minimat.

PROOF. By Proposition 1 in [12] it suffices to show that the condition $E^{-X_{1}} H_{1}^{-}=X_{1}$ is equivalent to $E^{X_{n}} H_{n}^{-}=X_{n}$ (bar over $E$ stands for closure) and that $\vec{E}_{\mathrm{X}_{1}} H_{1}^{+}=X_{1}$ is equivalent to $E^{X_{n_{H}}^{+}}=X_{n}$. By isomorphism $E^{X_{n_{4}}^{-}}=X_{n}$ can be identified with $E^{X_{1}^{n \Theta}}\left(H_{1}^{-}\right)^{n \Theta}=X_{1}^{n \theta}$, which, by Lemma 1 , holds if and only if $E{ }^{X_{1}} H_{1}^{-}=X$. A similar argument establishes the other equival ence.
PROOF OF THEOREM 1. Let the $X$ described in the theorem be denoted $\hat{x}$, and set $\hat{x}:=\sigma\left(\hat{X}_{1}\right)$. Then $\hat{X}$ is the space of all centered $\hat{X}$-measurable random variables in $H[5,6]$, and
$\sigma(\hat{X})=\hat{X}$. Moreover, if $X^{-}:=V \quad t \leq 0_{t} X$, it is not hard to see that $\hat{X}^{-}:=\sigma\left(\hat{X}^{-}\right)=\sigma\left(\hat{X}_{1}^{-}\right)$, since $U_{t} \hat{X}_{n}=\left(U_{t} X_{1}\right)^{n *}$. An analogous relation holds for summation over the future.
(if): Show that $\hat{X}$ is a minimal state space. Since $\hat{X}_{1}$ is Markovian, so is $\hat{X}$. Hence, in view of Lemma $2, \hat{X}$ is a state space for $y$. Now assume that $\hat{X}$ is not minimal. Then there is a state space $X$ such that $X:=\sigma(X)$ is properly contained in $\hat{X}$. Then, since all $\xi \in X$ are $\hat{X}$-measurable and $\hat{X}$ is the space of all centered $\hat{X}$-measurable random variables, $X \subset \hat{X}$. Therefore $X_{1}$ must be a proper subset of $\hat{X}_{1}$, or else $X=\sigma\left(X_{1}\right)=\sigma\left(\hat{X}_{1}\right)=\hat{X}$. This contradicts the minimality of $\hat{X}_{1}$.
(only if): Let $X$ be a minimal state space for $y$. First let us assume that $X_{1}$ is not minimal. Then there is another $\left(\mathrm{H}_{1}^{-}, \mathrm{H}_{1}^{+}\right)$-splitting subspace $\hat{X}_{1}$ which is a proper subspace of $X_{1}$. Let $\hat{X}=\sigma\left(\hat{X}_{1}\right)$, and let $\hat{X}$ be the space of all $\hat{X}$-measurable elements in $H$. Clearly $\hat{X} \subset X:=\sigma(X)$. We want to show that this inclusion is proper, contradicting minimality of $X$. But this is the case, for there is a $\xi \in X_{1}$ such that $\xi \perp \hat{X}_{1}$. Consequently, by the Gaussian property, $\sigma\{\xi\}$ and $\hat{X}$ are independent, while both are subfields of $x$. Hence $X_{1}$ must be minimal, and $X_{1}=\hat{X}_{1}$. Next assume that $X_{n}$ is not of the form (15b); i.e. $X_{n} \neq \hat{X}_{n}$. Then since $\hat{X}_{n}$ is minimal (Lemma 3 ), $X_{n} \neq \hat{X}_{n}$, i.e. there is a $\xi \in X$ which does not belong to $\hat{X}$ and consequently is not $\hat{X}$-measurable. Hence $\hat{X}$ is a proper subfield of $X$ contradicting minimality of $X$. Therefore $X=\hat{X}$. Finally $\hat{X}$ is Markovian only if $\hat{X}_{1}$ is Markovian." The last statement of the theorem follows from Lemma 3 . $\square$

## 5. THE STATE SPACE COMPONENT OF THE FIRST CHAOS

Thus it remains to determine the minimal Markovian ( $\mathrm{H}_{1}^{-}, \mathrm{H}_{1}^{+}$)-splitting subspaces $X_{1}$. This is almost the problem solved in $[1-3]$. To explain how it differs, let $\zeta \in H_{1}^{-} \cap H_{1}^{+}$ be defined in the following manner. If $y_{1}(0) \neq 0$, set $\zeta:=y_{1}(0)$, otherwise let it be arbitrary. (Remember that $H_{1}^{-} \cap H_{1}^{+} \neq \emptyset$.) Next define the process $z(t):=U_{t} \zeta$. Then $z(t) \in H_{1}$ for all $t$. Moreover,

$$
\left\{\begin{array}{l}
\mathrm{H}_{1}^{-}(z) \subset \mathrm{H}_{1}^{-}  \tag{16a}\\
\mathrm{H}_{1}^{+}(z) \subset \mathrm{H}_{1}^{+}
\end{array}\right.
$$

where $H_{1}^{-}(z)$ and $H_{1}^{+}(z)$ are the closed linear hulls of the random variables $\{z(t) ; t \leq 0\}$ and $\{z(t) ; t \geq 0\}$ respectively. Since $y$ is purely nondeterministic and mean-square continuous, so is $z$. Therefore $z$ has a spectral density $\Phi(i \omega)$. A scalar solution $W$ of the equation

$$
\begin{equation*}
W(s) W(-s)=\Phi(s) \tag{17}
\end{equation*}
$$

will be called a (full-rank) spectral factor of $z$. Now, if $y$ is Gaussian as assumed in $[1,2], z=y$ and we have equality in each of relations (16). Then there is a procedure in $[1,2]$ to determine $X_{1}$ from a certain pair ( $W, \bar{W}$ ) of spectral factors. However, in the non-Gaussian case, $z \neq y$, and we cannot assume that relations (16) hold with equality, not even when $z=y_{1}$. Hence there is a "mismatch" between the process $z$ and the geometry in $H_{1}$, and consequently the procedure of $[1,2]$ will have to be modified.

Fortunately the basic results of [1,2] depend in no crucial way on the spectral factor construction. The following result found in [1,2] is a consequence of the geometry in $H_{1}$ only. The theorem requires some new notation: For any Wiener process $\{u(t) ; t \in R\} \in H_{1}$, let $H_{1}^{-}(d u)$ and $H_{1}^{+}(d u)$ be the Gaussian spaces generated by the increments $\{u(\tau)-u(\sigma) ; \tau, \sigma \leq 0\}$ and $\{u(\tau)-u(\sigma) ; \tau, \sigma \geq 0\}$ respectively. In particular, we have $H_{1}^{-}(d v)=H_{1}^{-}$and $H_{1}^{+}(d \bar{v})=H_{1}^{+}$. Here and in the sequel, when we talk of a "Wiener process," we shall always mean a centered Gaussian process defined on the whole real line by a spectral representation

$$
\begin{equation*}
u(t)=\int_{-\infty}^{\infty} \frac{e^{i \omega t}-1}{i \omega} d \hat{u}(i \omega), \tag{18}
\end{equation*}
$$

where d $\hat{u}$ is a Gaussian orthogonal stochastic measure such that $E|d \hat{u}|^{2}=\frac{1}{2 \pi} d \omega$.
THEOREM 2. A subspace $X_{1} \in H_{1}$ is a minimal Markovian $\left(H_{1}^{-}, H_{1}^{+}\right)$-splitting subspace if and onty if

$$
\begin{equation*}
x=H_{1}^{-}(d u) \ominus H_{1}^{-}(d \bar{u}) \tag{19}
\end{equation*}
$$

for some pair $(u, \bar{u})$ of wiener processes in $H_{1}$ such that

$$
\begin{align*}
& H_{1}^{-}(d \bar{u}) \subset H_{1}^{-}(d u)  \tag{20a}\\
& H_{1}^{-} \subset H_{1}^{-}(d u)  \tag{20b}\\
& H_{1}^{+} \subset H_{1}^{+}(d \bar{u})  \tag{20c}\\
& H_{1}^{+}(d \bar{u})=H_{1}^{+} \vee H_{1}^{+}(d u)  \tag{20d}\\
& H_{1}^{-}(d u)=H_{1}^{-} \vee H_{1}^{-}(d \bar{u}) . \tag{20e}
\end{align*}
$$

The processes $u$ and $\bar{u}$ (which are essentially unique) are called respectively the forward and the backward generating processes of $X$. (Condition (20a) is equivalent to $\mathrm{H}_{1}^{-}(\mathrm{du})$ and $\mathrm{H}_{1}^{+}(\mathrm{du})$ intersecting perpendicularly. Moreover, (20d) is an observability and (20e) a constructibility condition [1,2].)

The Gaussian space of any Wiener process $u$ in $H_{1}$ coincides with $H_{1}$ [9], and consequently any $\eta \in H_{1}$ can be written

$$
\begin{equation*}
n=\int_{-\infty}^{\infty} f(-t) d u(t) \tag{21a}
\end{equation*}
$$

where $f \in L_{2}(\mathbb{R})$, or equivalently,

$$
\begin{equation*}
\eta=\int_{-\infty}^{\infty} \hat{\mathbf{f}}(i \omega) d \hat{u}(i \omega) \tag{21b}
\end{equation*}
$$

where $\omega \rightarrow \hat{f}(i \omega)$ is the Fourier transform [9]. [We shall refer also to the function $\hat{f}$ as the Fourier-transform, although it properly should be called the (double-sided) Laplace-transform.] Relations (22) establishes an isometric isomorphism between $H_{1}$ and $L_{2}$ (II), where III is the imaginary axis. Let $T_{u}: H_{1} \rightarrow L_{2}$ (II) be the map $T_{u} \eta=\hat{f}$. Then it can be seen that $T_{u}$ is unitary. Let $T_{u}^{*}$ denote the adjoint, i.e. $\eta=T_{u}^{*} \hat{f}$, which is relation (2lb). The shift $U_{t}$ corresponds to $e^{i \omega t}$ under the isomorphism $T_{u}$.

LEMMA 4. There is a one-one comespondence between Wiener processes $u$ in $H_{1}$ and. spectral factors $W$ of $z$ described by the following rule. For each $u$, $W:=T_{u}$ is a spectral factor. For each spectral factor $W$, u defined by (18) and $d \hat{u}=W^{-1} G d \hat{v}$, where $G:=T v^{\zeta}$, is a wiener process.

PROOF. Let $W:=T_{u} \zeta$. Then

$$
\begin{equation*}
z(t)=\int_{-\infty}^{\infty} e^{i \omega t_{W} d \hat{u}}, \tag{22}
\end{equation*}
$$

from which it is easy to see that the inverse Fourier transform of $E\{z(t) z(0)$ ' $\}$ is $W(i \omega) W(-i \omega)$, establishing $W$ as a spectral factor. In particular $G$ is a spectral factor. Therefore, for any spectral factor $W$, $d \hat{u}:=W^{-1} G d \hat{v}$ is a Gaussian orthogonal stochastic measure such that $E|d \hat{u}|^{2}=\frac{1}{2 \pi} \hat{d} \omega$, for $d \hat{v}$ is. Hence (18) is a Wiener process.

Next we introduce the Hardy spaces $H_{2}^{+}$and $H_{2}^{-}$: Let $H_{2}^{+}\left(\mathrm{H}_{2}^{-}\right)$be the subspace of $\mathrm{L}_{2}$ (II) of functions whose inverse Fourier-transforms vanish on the negative (positive) real line. From (21) it follows that $T_{u} H^{-}(d u)=H_{2}^{+}$and $T_{u} H^{+}(d u)=H_{2}^{-}$. A function $K$ which is bounded and analytic in the open left half-plane and has modulus one on the imaginary axis is called inner. Define $K^{*}(i \omega):=K(-i \omega)$; $K *$ is the inverse of $K$. If $f \in H_{2}^{+}$and K is inner, $f \mathrm{~K} \in \mathrm{H}_{2}^{+}$, and $H_{2}^{+} \mathrm{K}$ is a subspace of $H_{2}^{+}$. Let $H(K)$ denote the orthogonal complement of $\mathrm{H}_{2}^{+} \mathrm{K}$ in $\mathrm{H}_{2}^{+}$.

THEOREM 3. Let $\zeta \in H_{1}^{-} \cap H_{1}^{+}$be arbitramy, and set $G:=T_{V} \zeta$ and $\bar{G}:=T_{\bar{v}} \zeta$. Let $\Gamma:=G / \bar{G}$. Then $\mathrm{X}_{1}$ is a minimal Markovian $\left(\mathrm{H}_{1}^{-}, \mathrm{H}_{1}^{+}\right)$-splitting subspace if and only if there is a pair of inner functions $\left(Q, \bar{Q}^{*}\right)$ such that $K:=\Gamma Q \bar{Q}^{*}$ is also inner, $K$ and $Q$ are coprime, K and $\mathrm{Q}^{*}$ are coprime, and

$$
\begin{equation*}
X_{1}=T_{u}^{*} H(K), \tag{23}
\end{equation*}
$$

where $u$ is the Wiener process (18) with $d \hat{u}=Q * d \hat{v}$.
PROOF. We present an appropriately modified version of the proof in [1,2]. The idea is to translate conditions (20) to the Hardy space setting and apply Beurling's Theorem [13]. To this end, first note that if $u_{1}$ and $u_{2}$ are two Wiener processes in $H_{1}$, and $W_{1}$ and $W_{2}$ are their corresponding spectral factors,

$$
\begin{equation*}
T_{u_{2}} \eta=\left(T_{u_{1}} n\right)\left(W_{2} / W_{1}\right) \tag{24}
\end{equation*}
$$

for any $n \in H_{1}$, as is easily seen from Lemma 4. Then, if $W$ and $\bar{W}$ are the spectral factors corresponding to $u$ and $\bar{u}$ respectively, applying the map $T_{u}$ to (20a), (20b) and (20e) and $\mathrm{T}_{\overline{\mathrm{u}}}$ to (20c) and (20d) yields

$$
\begin{array}{lll}
H_{2}^{+} \mathrm{K} \subset \mathrm{H}_{2}^{+} & \text {where } & \mathrm{K}:=\mathrm{W} / \overline{\mathrm{W}} \\
\mathrm{H}_{2}^{+} \mathrm{Q} \subset \mathrm{H}_{2}^{+} & \text {where } & \mathrm{Q}:=\mathrm{W} / \mathrm{G} \\
\mathrm{H}_{2}^{-} \mathrm{Q} \subset \mathrm{H}_{2}^{-} & \text {where } & \overline{\mathrm{Q}}:=\overline{\mathrm{W}} / \overline{\mathrm{G}} \\
\mathrm{H}_{2}^{-}=\left(\mathrm{H}_{2}^{-} \mathrm{Q}\right) & \vee\left(\mathrm{H}_{2}^{-} \mathrm{K}^{*}\right) & \\
\mathrm{H}_{2}^{+}=\left(\mathrm{H}_{2}^{+} \mathrm{Q}\right) \vee\left(\mathrm{H}_{2}^{+} \mathrm{K}\right) . & \tag{25e}
\end{array}
$$

Now, since $H_{1}^{-}(d \bar{u})$ is invariant under the $\operatorname{shift}\left\{U_{t} ; t \leq 0\right\}, H_{2}^{+} K$ is invariant under $\left\{e^{i \omega t} ; t \leq 0\right\}$. Therefore, by Beurling's Theorem [13], (25a) holds if and only if $K$ is inner. In the same way we see that (25b) is equivalent to $Q$ being inner and (25c) to $\bar{Q}$ being inner with respect to $H_{2}^{-}$, i.e. $\bar{Q}^{*}$ inner. Moreover, (25d) and (25e) are valid if and only if the stated coprimeness conditions hold [13], and (19) is equivalent to $T_{u} X=H_{2}^{+} \ominus\left(H_{2}^{+} K\right)=: H(K)$, i.e. (23). The statement about u follows from Lemma 4.

REMARK. Let us pinpoint in what way this theorem differs from the corresponding result in [1,2]. In the case studied in [1,2], the pairs ( $W, \bar{W}$ ) which generate splitting subspaces are precisely those for which $W \in H_{2}^{+}, \bar{W} \in H_{2}^{-}$, and $K$ is inner. In the present setting these three conditions must also hold, but in addition we must have $W=G Q$ and $\bar{W}=\bar{G} \bar{Q}$. These factorizations correspond to the inner-outer factorizations of [1, 2], but the difference is that now $G$ and $\bar{G}$ are not outer. Consequently, some of the pairs $(W, \bar{W})$ mentioned above will be excluded. Note that the innovation process does not correspond to an outer spectral factor of $z$, since $\overline{\mathrm{E}}^{\mathrm{H}^{-}} \mathrm{H}^{+}$is not the predictor space of $z$.
6. THE STATE PROCESS

We recall from Section 2 that

$$
\begin{equation*}
y(t)=\sum_{n \in N} y_{n}(t) \tag{26a}
\end{equation*}
$$

where

$$
\begin{equation*}
y_{n}(t)=\int_{-\infty}^{t} \int_{-\infty}^{t_{1}} \ldots \int_{-\infty}^{t_{n-1}} g_{n}\left(t-t_{1}, t-t_{2}, \ldots, t-t_{n}\right) d v\left(t_{1}\right) d v\left(t_{2}\right) \ldots d v\left(t_{n}\right) \tag{26b}
\end{equation*}
$$

for some $g_{n} \in L_{2}\left(R^{n}\right)$. Let us assume that this innovation representation is given, i.e. that the functions $\left\{\mathrm{g}_{\mathrm{n}} ; \mathrm{n} \in \mathrm{N}\right\}$ are known.

Let us now consider a minimal state space $X$ with forward generating process $u$. Then, since $H_{1}^{-} \in H_{1}^{-}(d u)$,

$$
\begin{equation*}
y_{n}(0) \in H_{n}^{-}(d u):=H_{1}^{-}(d u) * H_{1}^{-}(d u) * \ldots * H_{1}^{-}(d u) \tag{27}
\end{equation*}
$$

( $n$ times) and consequently there is a representation

$$
\begin{equation*}
y_{n}(0)=\int_{-\infty}^{0} \int_{-\infty}^{t_{1}} \ldots \int_{-\infty}^{t_{n-1}} w_{n}\left(-t_{1},-t_{2}, \ldots,-t_{n}\right) d u\left(t_{1}\right) d u\left(t_{2}\right) \ldots d u\left(t_{n}\right) \tag{28}
\end{equation*}
$$

for some $L_{2}\left(R^{n}\right)$. Defining $w_{n}$ to be zero whenever some argument is zero, we may write this $y_{n}(0)=I_{n}\left(w_{n} ; u\right)$. By the same recipe we write $y_{n}(0)=I_{n}\left(g_{n} ; v\right)$. We need to determine $w_{n}$ from $g_{n}$. To this end, let $\hat{f} \leqslant L_{2}\left(I^{n}\right)$ be the $n$-fold Fourier-transform

$$
\hat{f}\left(i \omega_{1}, \ldots, i \omega_{n}\right)=\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} e^{-j \omega_{1} t_{1}-\ldots-i \omega_{n} t_{n_{n}}}{ }_{f\left(t_{1}, \ldots, t_{n}\right) d t_{1} \ldots d t_{n} . . . . . . . .}
$$

Let $F_{n}: L_{2}\left(R^{n}\right) \rightarrow L_{2}\left(I^{n}\right)$ be the operator defined by $\hat{f}=F_{n} f$. The following is a multidimensional version of (21).

LEMMA 5. Let $f \in L_{2}\left(\mathbb{R}^{n}\right)$ and set $\hat{i}:=F_{n} f$. Let u be a wiener process (18). Then

$$
\begin{equation*}
I_{n}(f ; u)=I_{n}(\hat{f} ; \hat{u}) \tag{29}
\end{equation*}
$$

PROOF. First let $f$ be of the form (8). Then $\hat{f}$ has this form too, and $\hat{f}_{i}=F_{1} f_{i}$. From (21) we have

$$
\eta_{i}=\int_{-\infty}^{\infty} f_{i}(-t) d u(t)=\int_{-\infty}^{\infty} \hat{f}_{i}(i \omega) d \hat{u}(i \omega)
$$

Then, by Itô's formula, i.e. (11) with $v$ exchanged for $u$ or $\hat{u}$, each member of (29) can be reduced to the same expression in $\eta_{1}, \eta_{2}, \ldots, \eta_{n}$. Hence (29) holds for functions of type (8). Then, since finite linear combinations of functions of type (8) are dense in $L_{2}\left(\mathbb{R}^{n}\right)$ or $L_{2}\left(\mathbb{I}^{n}\right)$, (20) holds in general.

Consequently, defining $W_{n}:=F_{n} W_{n}$ and $G_{n}:=F_{n} g_{n}, W_{n}$ can be determined from $g_{n}$ via the relation

$$
\begin{equation*}
W_{n}\left(i \omega_{1}, \ldots, i \omega_{n}\right)=G_{n}\left(i \omega_{1}, \ldots, i \omega_{n}\right) Q\left(i \omega_{1}\right) \ldots Q\left(i \omega_{n}\right), \tag{30}
\end{equation*}
$$

for $d \hat{v}=Q d \hat{u}$ (Theorem 3).
It is well-known [6], and we have already used this fact in Section 2, that $\eta=I_{n}(\hat{f} ; \hat{u})$ defines an isomorphism between $H_{n}$ and $L_{2}\left(I^{n}\right)$. More precisely $T_{u}^{(n)}: H_{n} \rightarrow L_{2}\left(\mathbb{I I}^{n}\right)$, defined by $T_{u}^{(n)} \eta=\hat{f}$, is a map with the property that $(n!)^{\frac{1}{2}} T_{u}^{(n)}$ is unitary. The space of Fourier-transforms of functions (such as $w_{n}$ ) in $L_{2}\left(\mathbb{R}^{n}\right)$ which vanish whenever an argument is negative, can be identified with $\left(H_{2}^{+}\right)^{n \Theta}$ so that $T_{u}^{(n)} H_{n}^{-}(d u)=\left(H_{2}^{+}\right)^{n \Theta}$. In the sequel we shall use precisely this realization of the tensor-product Hilbert space $\left(H_{2}^{+}\right)^{n ๑}$. Then the tensor product $f_{1} \odot f_{2} \odot \ldots \odot f_{n}$ is given by (8). Also, for subspaces $A_{1}, A_{2}, \ldots, A_{n}$ in $H_{n}$,

$$
\begin{equation*}
T_{n}^{(n)}\left\{A_{1} * A_{2} * \ldots * A_{n}\right\}=\left(T_{u} A_{1}\right) \odot\left(T_{u} A_{2}\right) \odot \ldots \odot\left(T_{u} A_{n}\right) \tag{31}
\end{equation*}
$$

so that in particular

$$
\begin{equation*}
\mathrm{T}_{\mathrm{u}}^{(\mathrm{n})} \mathrm{X}_{\mathrm{n}}=H(K) \odot H(K) \odot \ldots \cdot H(K) \tag{32}
\end{equation*}
$$

( $n$ times). Then, since $y_{n}(0) \in X_{n}, W_{n} \in H(K){ }^{n \Theta}$.
Following [1,2] we say that $X$ is regular if $H(K)$ contains only Fourier-transforms of continuous functions. All $X$ with $\operatorname{dim} X_{1}<\infty$ are clearly regular. It can be shown $[1,2]$ that if X is regular the functional

$$
\begin{equation*}
v \hat{f}=\frac{1}{(2 \pi)^{n}} \int_{\mathbb{R}^{n}} \hat{f}\left(i \omega_{1}, \ldots, i \omega_{n}\right) d \omega_{1} \ldots d \omega_{n} \tag{33}
\end{equation*}
$$

is bounded on $H(K)^{n \ominus}$. Hence, since $V \hat{f}=f(0)$ where $f=F_{n}^{-1} \hat{f}$, there is a $B_{n} \in H(K)^{n \ominus}$ such that

$$
\begin{equation*}
f(0)=\left\langle\hat{f}, B_{n}\right\rangle_{H(K)}^{n \odot} \tag{34}
\end{equation*}
$$

(Riesz Theorem). Next, as in $[1,2]$, we define a strongly continuous semigroup $\left\{e^{A t} ; t \geq 0\right\}$ on $H(K)$ by

$$
\begin{equation*}
e^{A t} \hat{\mathrm{f}}=\mathrm{P}^{H(K)} e^{-\mathrm{i} \omega \mathrm{t}} \hat{\mathrm{f}} \tag{35}
\end{equation*}
$$

where $P^{H(K)}$ denotes the orthogonal projection on the subspace $H(K)$. Moreover define the linear bounded operator $C_{n}: H(K){ }^{n \Theta} \rightarrow R$ given by

$$
\begin{equation*}
C_{n} \hat{f}=\left\langle W_{n}, \hat{f}\right\rangle_{H(K)}{ }^{n \odot} . \tag{36}
\end{equation*}
$$

Then the following lemma is a multilinear version of the construction in $[14,15]$ which is being used in $[1,2]$.

LEMMA 6. The integrand in (28) admits the factorization

$$
\begin{equation*}
w_{n}\left(t_{1}, t_{2}, \ldots, t_{n}\right)=C_{n}\left(e^{A t_{1}} \odot e^{A t_{2}} \oplus \ldots \odot e^{A t_{n}}\right) B_{n} \tag{37}
\end{equation*}
$$

for $t_{1}, t_{2}, \ldots, t_{n} \geq 0$.
PROOF. In view of (34)

$$
w_{n}\left(t_{1}, t_{2}, \ldots, t_{n}\right)=\left\langle e^{\left.i \omega_{1} t_{1}+i \omega_{2} t_{2}+\ldots+i \omega_{n} t_{n_{w_{n}}}, B_{n}\right\rangle .}\right.
$$

Since $W_{n} \in H(K)^{n \odot}$, we have

$$
w_{n}\left(t_{1}, t_{2}, \ldots, t_{n}\right)=\left\langle w_{n}, p^{H(K)} e^{n \Theta} e^{-i \omega_{1} t_{1}-i \omega_{2} t_{2}-\cdots-i \omega_{n} t_{B_{n}}}\right\rangle
$$

which is the required result.
Consequently

$$
\begin{equation*}
y_{n}(t)=c_{n} x_{n}(t) \tag{38}
\end{equation*}
$$

where $x_{n}(t)$ is the $H(K)^{\text {no }}$-valued process

$$
\begin{equation*}
\left.x_{n}(t)=\int_{-\infty}^{t} \int_{-\infty}^{t_{1}} \ldots \int_{-\infty}^{t_{n-1}\left[e^{A\left(t-t_{1}\right)}\right.} 0 \ldots \theta e^{A\left(t-t_{n}\right)}\right] B_{n} d u\left(t_{1}\right) \ldots d u\left(t_{n}\right) . \tag{39}
\end{equation*}
$$

If $H(K)$ is infinite dimensional, $\left\{x_{n}(t) ; t \in R\right\}$ is not an ordinary stochastic process but must be defined in a weak sense [16]. Then the state process $\{x(t) ; t \in \mathbb{R}\}$ is defined as the (possibly weakly defined) $\oplus_{n \in N} H(K){ }^{n \Theta}$-valued process with components $x_{n}$; $\mathrm{n} \in \mathrm{N}$. This terminology is motivated by the following result developed along the lines in $[1,2]$.

PROPOSITION 1. Let x be a regular state space and let $\mathrm{x}_{\mathrm{n}}$ be given by (39). Then

Moreover, for each $\mathrm{n} \in \mathbb{N}$,

$$
\begin{equation*}
\left\{\left\langle\hat{f}, x_{n}(0)\right\rangle_{H(K)}{ }^{n 0} \mid \hat{f} \in H(K)^{n \theta_{0}}\right\}=x_{n} . \tag{40}
\end{equation*}
$$

$$
\begin{equation*}
X_{n} \perp H_{n}^{+}(d u) \tag{41}
\end{equation*}
$$

PROOF. Let $\xi \in X_{n}$ be arbitrary, and let $\hat{f}:=T_{u}^{(n)} \xi$ and $f:=F_{n}^{-1} \hat{f}$. Then, by Lemma 5,

$$
\begin{equation*}
\xi=\int_{-\infty}^{0} \int_{-\infty}^{t_{1}} \cdots \int_{-\infty}^{t_{n-1}} f\left(t_{1}, \ldots, t_{n}\right) d u\left(t_{1}\right) \ldots d u\left(t_{n}\right) . \tag{42}
\end{equation*}
$$

By exchanging $w_{n}$ for $f$ in the proof of Lemma 6, we obtain

$$
\begin{equation*}
f\left(t_{1}, \ldots, t_{n}\right)=\left\langle\hat{f},\left(e^{-A t_{1}} \cdot \ldots e^{-A t_{n}}\right)_{B_{n}}\right\rangle_{H(K)}{ }_{n \Theta} . \tag{43}
\end{equation*}
$$

Then (42) and (43) together yield (40). In view of (19), $X_{1} \subset H_{1}^{-}(d u) \perp H_{1}^{+}(d u)$, from which (41) follows.

In partular, for $n=1$ we can write (38) and (39) in the following suggestive form

$$
\left\{\begin{align*}
d x_{1} & =A x_{1} d t+B_{1} d u  \tag{44}\\
y_{1} & =C_{1} x_{1}
\end{align*}\right.
$$

The higher-chaos subsystems are nonlinear. In the next section we shall illustrate this with an example.

Note that a backward realization for X generated by $\overline{\mathrm{u}}$ is obtained by developing the above analysis in $\left(\mathrm{H}_{2}^{-}\right)^{\text {no }}$ rather than in $\left(\mathrm{H}_{2}^{+}\right)^{\text {n® }}$. Whereas the forward property is characterized by (41), the backward one is determined by $X_{n} \perp H_{n}^{-}(d \bar{u})$ for each $n \in N$.

Finally, in the case that $X$ is not regular, other constructions involving rigged Hilbert spaces are possible [19].

## 7. THE FINITE-DIMENSIONAL BILINEAR CASE

To illustrate our point let us consider the simplest possible nonlinear problem. Let the process $y$ have the innovation representation

$$
\begin{equation*}
y(t)=\int_{-\infty}^{t} g_{1}(t-\sigma) d v(\sigma)+\int_{-\infty}^{t} \int_{-\infty}^{\tau} g_{2}(t-\tau, t-\sigma) d v(\sigma) d v(\tau) \tag{45a}
\end{equation*}
$$

and the backward innovation representation

$$
\begin{equation*}
y(t)=\int_{t}^{\infty} \bar{g}_{1}(t-\sigma) d \bar{v}(\sigma)+\int_{t}^{\infty} \int_{\tau}^{\infty} \bar{g}_{2}(t-\tau, t-\sigma) d \bar{v}(\sigma) d \bar{v}(\tau) . \tag{45b}
\end{equation*}
$$

Assume that $G_{1}:=F_{1} g_{1}$ is a rational function which is not identically zero. Then $\bar{G}_{1}:=F_{1} \bar{g}_{1}$ has the same properties, and $y_{1} \neq 0$. Moreover $y_{1}$ has a rational spectral density, namely $\Phi(s):=G_{1}(s) G_{1}(-s)$.

Now, setting $\Gamma:=G_{1} / \bar{G}_{1}$, find all pairs ( $Q, \bar{Q}^{*}$ ) of inner functions such that $K:=\Gamma Q \bar{Q}^{*}$ is inner and coprime with $Q$ and $\bar{Q}^{*}$. For each such solution form

$$
\begin{equation*}
X_{1}=\int_{-\infty}^{\infty} H(K) Q^{*} d v . \tag{46}
\end{equation*}
$$

Theorem 3 states that the $X_{1}$-spaces obtained in this way are precisely the minimal Markovian ( $\mathrm{H}_{1}^{-}, \mathrm{H}_{1}^{+}$) -splitting subspaces. In particular, $\mathrm{Q}_{1}=1$ yields $\mathrm{X}_{1}=\mathrm{E}^{\mathrm{H}_{1}^{-}} \mathrm{H}_{1}^{+}$, and $\bar{Q}_{1}=1$ yields $X_{1}=E^{\mathrm{H}_{1}^{+}} \mathrm{H}_{1}^{-}$. Since $\Gamma$ is rational, it can be shown that $K$ must be rational, and consequently $X_{1}$ is finite-dimensional [17]. In fact, all $X_{1}$ have the same dimension $n[1,2]$. By using the procedure described in Section 7 of [18] we can determine an $n \times n$-matrix $A_{1}$ and an $n \times 1$-matrix $B_{1}$ from $K$ and a $l \times n$-matrix $C_{1}$ from $W:=G Q$ so that

$$
\left\{\begin{align*}
d x_{1} & =A_{1} x_{1} d t+B_{1} d u  \tag{47}\\
y_{1} & =C_{1} x_{1}
\end{align*}\right.
$$

where $\operatorname{sp}\left\{x_{1}(t), \ldots, x_{n}(t)\right\}=U_{t} x_{1}, H_{1}^{+}(d u) \perp x_{1}$ and

$$
\begin{equation*}
u(t)=\int_{-\infty}^{\infty} \frac{e^{i \omega}-1}{i \omega} Q^{*}(i \omega) d \hat{\nu} \tag{48}
\end{equation*}
$$

To each $X_{1}$ there corresponds a minimal state space, namely

$$
\begin{equation*}
x=x_{1} \ominus x_{2}, \tag{49a}
\end{equation*}
$$

where

$$
\begin{equation*}
X_{2}=X_{1} * X_{1} \tag{49b}
\end{equation*}
$$

Hence, for each $t$, the $\frac{1}{2} n(n+1)$ random variables $\left\{x_{1}^{i j}:=x_{1}^{i}(t) * x_{1}^{j}(t) ; j \leq i\right\}$ span $U_{t} X_{2}$. (Remember that $x_{1}^{i j}=x_{1}^{j i}$.) Let $\left\{x_{2}(t) ; t \in \mathbb{R}\right\}$ be the $\frac{1}{2} n(n+1)$-dimensional stationary vector process with components $x^{i j}(t)$. Applying Ito's differentiation rule [6] to

$$
x_{1}^{i j}(t)=x_{1}^{i}(t) x_{1}^{j}(t)-E\left\{x_{1}^{i}(t) x_{1}^{j}(t)\right\}
$$

we obtain

$$
d x_{l}^{i j}(t)=\sum_{k=1}^{n}\left[a_{i k} x_{l}^{k j}(t)+a_{j k} x_{l}^{i k}(t)\right] d t+\left(b_{i} x_{l}^{j}(t)+b_{j} x_{1}^{i}(t)\right) d u
$$

where $a_{i k}$ and $b_{i}$ are the components of $A_{1}$ and $B_{1}$ respectively. Defining the $\frac{1}{2} n(n+1) \times \frac{1}{2} n(n+1)$-matrix $A_{2}$ and the $\frac{1}{2} n(n+1) \times n$-matrix $B_{2}$ appropriately, this can be written

$$
\begin{equation*}
d x_{2}=A_{2} x_{2} d t+B_{2} x_{1} d u \tag{50}
\end{equation*}
$$

Integrating this bilinear equation we get an expression of type

$$
x_{2}(t)=\int_{-\infty}^{t} \int_{-\infty}^{\tau} f(t-\tau, t-\sigma) d u(\sigma) d u(\tau),
$$

where $f$ is a vector-valued function. Moreover

$$
y_{2}(t)=\int_{-\infty}^{t} \int_{-\infty}^{\tau} w_{2}(t-\tau, t-\sigma) d u(\sigma) d u(\tau)
$$

where $w_{2}$ is obtained from $g_{2}$ via formula (30). Now, since $y_{2}(0) \in X_{2}$, there are real numbers $\left\{c_{k} ; k=1,2, \ldots, \frac{1}{2} n(n+1)\right\}$ such that

$$
w_{2}(\tau, \sigma)=\sum_{k} c_{k} f_{k}(\tau, \sigma)
$$

and these numbers can be determined by known methods. Let $C_{2}$ be the $\frac{1}{2} n(n+1)$-dimensional row vector with components $c_{k}$. Then

$$
\begin{equation*}
y_{2}(t)=c_{2} x_{2}(t) \tag{51}
\end{equation*}
$$

Since $y=y_{1}+y_{2}$,

$$
\left\{\begin{align*}
d x_{1} & =A_{1} x_{1} d t+B_{1} d u  \tag{52}\\
d x_{2} & =A_{2} x_{2} d t+B_{2} x_{1} d u \\
y & =C_{1} x_{1}+C_{2} x_{2}
\end{align*}\right.
$$

is a realization of $y$, for $x=\left[\begin{array}{l}x_{1} \\ x_{2}\end{array}\right]$ is a Markov process. Note that even if $y_{1}$ were $z e r o$ we would need to include $x_{1}$ is the state process $x$, for $x_{2}$ by itself is not Markov.

Let $\hat{x}$ be the state process corresponding to $X_{1}=E_{H_{1}^{-}}^{H_{1}^{+}}$(in the coordinate-system of (52)). It is shown in [1-3] that, for any $X_{1}, E^{H_{1}^{-}} x_{1}(0)=\hat{x}_{1}(0)$. Therefore, in view of the definition of $x_{2}$ and the fact that $E^{H_{2}^{-}}=E^{H_{1}^{-}} * E^{H_{1}^{-}}, \quad E^{H_{2}^{-}} x_{2}(0)=\hat{x}_{2}(0)$. ( $E^{\mathrm{H}_{1}^{-}}$ and $\mathrm{E}^{\mathrm{H}_{2}^{-}}$applied to a vector means that the projection is performed componentwise.) Consequently the conditional expectation of $x(t)$ given $y_{t}^{-}$is

$$
\begin{equation*}
E^{t^{-}} x(t)=\hat{x}(t) \tag{53}
\end{equation*}
$$

for any realization (52). For this reason, remembering that the forward generating process of $E^{H_{1}^{-}} H_{1}^{+}$is the innovation $v$, we may call the system

$$
\left\{\begin{array}{l}
d \hat{x}_{1}=A_{1} \hat{x}_{1} d t+B_{1} d v \\
d \hat{x}_{2}=A_{2} \hat{x}_{2} d t+B_{2} \hat{x}_{1} d v  \tag{54}\\
y=C_{1} \hat{x}_{1}+C_{2} \hat{x}_{2}
\end{array}\right.
$$

the steady state non-linear filter of (52), and we have shown that this filter is invariant over the class (52) of minimal realizations. A similar result can be obtained for backward realizations in terms of $\bar{v}$.

## 8. CONCLUDING REMARKS

The purpose of this paper is to investigate the structural aspects of the nonlinear stochastic realization problem and to clarify basis concepts. This is a first step toward a nonlinear realization theory. Hence we have not concerned ourselves with algorithmic aspects of the problem, and our analysis is based on the availability of an innovation representation, the actual determination of which is a nontrivial problem in itself (see [20]).

The question of state space construction needs to be further studied. It cuuld be argued that condition (4) is too restrictive since there could well be $\left(\oplus_{n \in N^{\prime}} H_{n}^{-}, \oplus_{n \in N_{n}} H_{n}^{+}\right)$-splitting subspaces which are not of the form (4), having a nonzero angle with some (or even all) $H_{n}$. Hence, if we can do without realizations of individual $y_{n}$ but only need their sum $y$, it is possible that we are missing state spaces of smaller size.

Our interest in the nonlinear realization problem emanates from its potential value as a conceptual framework for certain classes of nonlinear filtering problems. This will be the topic of a future study.

## REFERENCES

1. A. Lindquist and G. Picci, State space models for Gaussian stochastic processes, Stochastic Sustems: The Mathematics of Fittering and Identification and Applications, M. Hazewinkel and J.C. Willems, Eds., Reidel Publ. Co., 1981.
2. A. Lindquist and G. Picci, Realization theory for multivariate Gaussian processes (to appear).
3. G. Ruckebusch, Theorie géometrique de la représentation markovienne, Thèse de doctorat d'état, Univ. Paris VI, 1980.
4. J.L. Doob, Stochastic Processes, John Wiley, New York, 1953.
5. J. Neven, Processus Aléatoire Gcussiens, Presses de L'Université de Montréal, 1968.
6. G. Kallianpur, Stochastic Filtering Theory, Springer-Verlag, 1980.
7. D. Ocone, Topics in Nonlinear Filtering Theory, Ph.D. thesis, M.I.T., December 1980.
8. S.K. Mitter and D. Ocone, Multiple integral expansions for nonlinear filtering, Proc. 18th Conf. Decision and Control, Ft. Lauderdale, Florida, December 1979.
9. Yu. A. Rozanov, Stationary Random Processes, Holden-Day, 1967.
10. K. Ito, Multiple Wiener integrals, J. Math. Soc. Japan 3, 1951, 157-169.
11. H.P. McKean, Stochastic Integrals, Academic Press, 1969.
12. A. Lindquist, G. Picci and G. Ruckebusch, On minimal splitting subspaces and Markovian representations, Math. Systems Theory 12, 271-279 (1979).
13. H. Helson, Lecture Notes on Invariant Subspaces, Academic Press, 1964.
14. J.S. Baras and R.W. Brockett, $\mathrm{H}^{2}$ functions and infinite dimensional realization theory, SIAM J. ControI 13(1975),. 221-241.
15. J.S. Baras and P. Dewilde, Invariant subspace methods in linear multivariabledistributed systems and lumped-distributed network synthesis, Proc. IEEE 64 (1976), 160-178.
16. A.V. Balakrishnan, Applied Functional Analysis, Springer-Verlag, 1976.
17. P.A. Fuhrmann, Linear Systems and Operators in Hilbert Space, McGraw-Hill, 1981.
18. A. Lindquist and G. Picci, Realization theory for multivariate Gaussian processes II: State space theory revisited and dynamical representations of finitedimensional state spaces, Proc. 2nd Interm. Conf. On Information Sciences and Systems, Patras, Greece, July 1979, Reidel Publ. Co., 108-129.
19. J.W. Helton, Systems with infinite dimensional state space: The Hilbert space approach, Proc. IEEE 64 (January 1976), 145-160.
20. H.P. McKean, Wiener's theory of nonlinear noise, Stochastic Differential Equations, SIAM-AMS Proceedings, American Math. Society, 1973, 191-209.

[^0]:    * Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506. This research was supported partially by the National Science Foundation under grant ECS7903731 and partially by the Air Force Office of Scientific Research under grant AFOSR 78-3519.
    + Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 . This research was supported by the Air force Office of Scientific Research under grant AFOSR 7T-3281D.

[^1]:    * Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506. This research was supported partially by the National Science Foundation under grant ECS7903731 and partially by the Air Force Office of Scientific Research under grant AFOSR 78-3519.
    $\dagger$ Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. This research was supported by the Air Force Office of Scientific Research under grant AFOSR 77-3281D.
    § LADSEB-CNR, Corso Stati Uniti 4, 35100 Padova, Italy.

