December, 1966

APPLICATIONS OF MULTIPLICITY THEORY TO N-PLE MARKOV PROCESSES*

V. Mandrekar

Technical Report No. 86

* Research supported in part by U.S. Army Research Office--Durham--
under Grant No. DA-ARO-D-31-124-G562.

1. Preliminaries and notation. Throughout this chapter a q-dimensional second order stochastic process will be denoted by $\{\underline{x}_t\}$ $(-\infty < t < \infty)$ where for each t, $\underline{x}_t$ is a column vector $(x_1(t),\ldots,x_q(t))^*$. Associated with $\{\underline{x}_t\}$ will be the following spaces:

(i) The space of the process up to t, $L_2(\underline{x};t)$ is the subspace $\mathfrak{S}\{x_i(\tau),\ \tau \leq t\}$ of $L_2(\Omega)$ generated by the random variables $\{x_i(\tau)\}$ $(\tau \leq t,\ 1 = 1,2,\ldots,q)$ $L_2(\underline{x};-\infty)$ the intersection of $L_2(\underline{x};t)$ for all real t and $L_2(\underline{x})$ is the smallest subspace of $L_2(\Omega)$ containing all $L_2(\underline{x};t)$ for each t.

(ii) For the processes with mutually orthogonal increments or those which are wide-sense martingales the notation H(;) of Chapter I will be used.

(iii) $P_{\mathcal{M}}$ will denote the projection onto $\mathcal{M}$.

Definitions of deterministic and purely non-deterministic processes are the same as in Chapter I. The following definition of a q-dimensional wide-sense Markov process is due to F. J. Beutler ([1]).

Definition 1.1. A q-dimensional process $\{\underline{x}_t\}$ $(-\infty < t < +\infty)$ is wide-sense Markov if for each i $(i = 1,2,\ldots,q)$ $P_{L_2(\underline{x};s)}\, x_i(t) = P_{\{x_1(s),\ldots,x_q(s)\}}\, x_i(t)$, $(s < t)$.

For our purpose we need the following definition of a q-dimensional wide-sense martingale. The notion of a wide-sense martingale for q = 1 is due to Doob ([3], p. 164).

Definition 1.2. $\underline{u}_t$-process is called a wide-sense martingale if for each k, $(k = 1,2,\ldots,q)$ $P_{H(\underline{u};s)}\, u_k(t) = u_k(s)$ with probability one for $s \leq t$.

The assumption (D) given below will be used throughout this chapter.

(D.1) $\underline{x}_t$-process is continuous in q.m.; i.e., each component process $\{x_i(t)\}$ is continuous in q.m.

(D.2) For all t,s real the covariance matrix function $\Gamma(t,s)$ is non-singular.

The assumption (D.2) and the definition of wide-sense Markov process imply that $P_{L_2(\underline{x};s)}x_i(t) = \sum_{j=1}^{q} a_{ij}(t,s)x_j(s)$, where the matrix $A(t,s) = (a_{ij}(t,s))$ is given by $A(t,s) = \Gamma(t,s)\,\Gamma^{-1}(s,s)$ for $s \leqq t$. It is easily verified that $A(t,s)$ is non-singular for each s,t ($s \leqq t$). The function $A(t,s)$ is called a transition matrix function and is defined only for $s \leqq t$. Beutler [1] has the following theorem which furnishes an operative criterion for verifying the wide-sense Markov property.

Theorem B ([1] Theorem 2). The following statements are equivalent

(1) $\underline{x}_t$ is wide-sense Markov

(2) For $s \leqq t \leqq u$ $A(u,s) = A(u,t)\,A(t,s)$

(3) With $A(t,s) = \Gamma(t,s)\,\Gamma^{-1}(s,s)$ for $s \leqq t \leqq u$ $A(s,u) = A(s,t)A(t,u)$.

In the case of stationary processes $A(t,s) = B(t-s)$ ($s \leqq t$). Hence $B(\cdot)$ can be considered as a function on non-negative real numbers. As will be shown in Theorem 2.2, one can easily characterize wide-sense Markov processes in terms of the transition matrix function $B(\cdot)$. We remark that ($t \geqq 0$) $B(t) = A(t) = \Gamma(t,0)\,\Gamma^{-1}(0,0)$.

2. Characterizations of the wide-sense Markov processes. We first consider the non-stationary processes.

Theorem 2.1. If $\underline{x}_t$ $(-\infty < t < +\infty)$ is q-dimensional stochastic process satisfying (D) then it is wide-sense Markov if and only if $\underline{x}_t = \underline{\Psi}(t)\underline{u}_t$ with probability one, where for every t, $\underline{\Psi}(t)$ is a non-singular $q \times q$ matrix and $\underline{u}_t$ process is a q-dimensional wide-sense martingale with $H(\underline{u};t) = L_2(\underline{x};t)$. Further for all s,t the matrix $J(t,s) = (\mathcal{E}\, u_i(t)\, \overline{u_j(s)})$ is non-singular.

Proof. Sufficiency. Let $\underline{x}_t = \underline{\Psi}(t)\underline{u}_t$ where $\underline{\Psi}(t)$ and $\{\underline{u}_t\}$ are as described above. Then for $s \leq t$ if we donate by $\overline{P}_{L_2(\underline{x};s)}\underline{x}_t$ the column vector $(P_{L_2(\underline{x};s)}\underline{x}_j(t),\ldots,P_{L_2(\underline{x};s)}\underline{x}_q(t))^*$ we have by definition of a wide-sense martingale, with probability one,

$$\overline{P}_{L_2(\underline{x};s)}\, \underline{x}_t = \overline{P}_{L_2(\underline{x};s)}\, \underline{\Psi}(t)\underline{u}_t = \overline{P}_{H(\underline{u};s)}\, \underline{\Psi}(t)\underline{u}_t = \underline{\Psi}(t)\underline{u}_s$$

Since $\underline{u}_s = \underline{\Psi}^{-1}(s)\underline{x}_s$ with probability one, we obtain that the transition matrix function $A(t,s) = \underline{\Psi}(t)\,\underline{\Psi}^{-1}(s)$. The proof of sufficiency is now complete by appealing to Theorem B, (2).

Necessity. Let $\underline{x}_t$-process be wide-sense Markov. Then denoting by $A(t,s)$ the transition matrix function we recall that for $s \leq t$

(2.1) $\quad \overline{P}_{L_2(x;s)}\underline{x}_t = A(t,s)\underline{x}_s \qquad$ with probability one and for $s \leq t \leq u$

(2.2) $\quad A(u,s) = A(u,t)A(t,s)$.

Following Hida, we now define for every real t the function

$$\underline{\Psi}(t) = A(t,s_o) \qquad \text{if } s_o \leq t$$
$$= A^{-1}(s_o,t) \qquad \text{if } t < s_o$$

where s_o is a fixed real number. We shall show that for all s,t $(s < t)$ real

(2.3) $\quad A(t,s) = \underline{\Psi}(t)\underline{\Psi}^{-1}(s)$.

First of all if $s < s_o \leq t$ then (2.3) is a restatement of (2.2) i.e., $A(t,s) = A(t,s_o)\, A(s_o,s)$. Secondly, if $s_o \leq s < t$, from (2.2) we have

$$A(t,s)\, A(s,s_o) = A(t,s_o) \quad \text{i.e.} \quad A(t,s) = A(t,s_o)\, A^{-1}(s,s_o)$$

giving $A(t,s) = \underline{\Psi}(t)\underline{\Psi}^{-1}(s)$. Finally, if $s < t < s_o$ we get $A(s_o,s) = A(s_o,t)A(t,s)$ and hence $A(t,s) = \underline{\Psi}(t)\underline{\Psi}^{-1}(s)$. The fact that $\underline{\Psi}(t)$ is non-singular follows from non-singularity of $A(t,s)$ and the definition of $\underline{\Psi}(t)$. Therefore from (2.1) and (2.3), for $s < t$

(2.4) $\quad \overline{P}_{L_2(\underline{x};s)}\underline{x}_t = \underline{\Psi}(t)\underline{\Psi}^{-1}(s)\underline{x}_s$ with probability one.

If we define $\underline{u}_t = \underline{\Psi}^{-1}(t)\underline{x}_t$, then

(2.5) $\quad L_2(\underline{x};t) = H(\underline{u};t)$ for every t.

Thus from (2.4) and (2.5) we get

(2.6) $\quad \overline{P}_{H(\underline{u};s)}\, \underline{u}_t = \underline{u}_s$ (with probability one).

Since $\Gamma(t,s) = \underline{\Psi}(t)J(t,s)\underline{\Psi}^*(s)$ and $\underline{\Psi}(t)$ is non-singular for every t, we have $J(t,s)$ non-singular.

Corollary. If the continuous parameter process $\underline{x}_t$ is continuous in q.m. then so is $\underline{u}_t$ and $\underline{\Psi}(t)$ is a continuous function in the sense that each element of $\underline{\Psi}(t)$ is continuous.

Proof. If $\gamma_{ij}(t,s)$ denote the elements of $\Gamma(t,s)$ then by the continuity in q.m. of the process $\{x_i(t)\}$ we get for every fixed s $\lim_{t \to t_o} \gamma_{ij}(t,s) = \gamma_{ij}(t_o,s)$;

i.e., $\lim_{t \to t_o} \Gamma(t,s) = \Gamma(t_o,s)$. But by Theorem 2.1, $\Gamma(t,s) = \underline{\Psi}(t)J(s,s)\underline{\Psi}^*(s)$ (for $s < t$). Hence $\underline{\Psi}(t) = \Gamma(t,s)\, [J(s,s)\, \underline{\Psi}^*(s)]^{-1}$ as a function of t is continuous (note that s is fixed). To prove continuity in q.m. of $\underline{u}_t$; consider $j_{ii}(t,s)$; $\mathcal{E}|u_i(t) - u_i(t_o)|^2 = j_{ii}(t,t) - J_{ii}(t_o,t_o)$ $(t_o \leq t)$. Now

$J(t,t) = \bar{\psi}^{-1}(t)\Gamma(t,t)[\bar{\psi}^{*}(t)]^{-1}$ and hence we get $\lim_{t \to t_o} J(t,t) = J(t_o,t_o)$. We therefore have $\lim_{t \downarrow t_o} \mathcal{E}|u_i(t) - u_i(t_o)|^2 = 0$. A similar argument gives $\lim_{t \uparrow t_o} \mathcal{E}|u_i(t) - u_i(t_o)|^2 = 0$, thus completing the proof.

We now study stationary wide-sense Markov processes. In this case $(\mathcal{E}\,[x_i(t + h) \cdot \overline{x_j(t)}])$ for any h is a function of h. We denote it by R(h). By Theorem 2.1 and properties of wide-sense martingale it is easy to see that for every $h \geq 0$ and t real

(2.7) $\quad R(h) = \bar{\psi}(t + h)\, J(t,t)\, \bar{\psi}^{*}(t)$.

Let $h = 0$, we get

(2.8) $\quad R(0) = \bar{\psi}(t)\, J(t,t)\, \bar{\psi}^{*}(t)$.

With $t = 0$ in (2.7), one has

(2.9) $\quad R(h) = \bar{\psi}(h)\, J(0,0)\, \psi^{*}(0)$.

Relations (2.7) and (2.9) imply for $h \geq 0$ and $t \geq 0$

(2.10) $\quad R(h) = R(t + h)\, [J(0,0)\, \bar{\psi}^{*}(0)]^{-1}\, J(t,t)\, \bar{\psi}^{*}(t)$.

From (2.10), (2.9) and (2.8) for $t, h \geq 0$,

(2.11) $\quad R(h) = R(t + h)\, R^{-1}(t)\, R(0)$

With $R_1(t) = R(t)\, R^{-1}(0)$ (2.11) reduces to

(2.12) $\quad R_1(t + h) = R_1(t)\, R_1(h)$.

We prove the following theorem.

<u>Theorem 2.2.</u> If $\{\underline{x}_t\}$ $(-\infty < t < +\infty)$ is a q-dimensional stationary process satisfying assumption (D) then it is wide-sense Markov if and only if the transition matrix function $B(t) = e^{tQ}$ for every $t \geq 0$ where Q is a uniquely determined constant $q \times q$ matrix, none of whose eigenvalues has positive real parts.

<u>Proof</u> <u>Necessity</u>. We have already shown that for $R_1(t) = R(t)\, R^{-1}(0)$, equation (2.12) holds. Further from (D.1) it follows that $R_1(t)$ is a continuous function and

therefore $R_1(t) = e^{tQ}$ $(t \geq 0)$ is the solution of (2.12), where Q is a $q \times q$ constant matrix. The assumption (D.2) in addition implies that $R_1(t)$ is non-singular and hence Q is uniquely determined by $R_1(t)$. We recall that $B(t) = R(t)\, R^{-1}(0)$ for $t \geq 0$. Hence $B(t) = e^{tQ}$ $(t \geq 0)$. The statement about the eigenvalues will now be proved. Observe that for any non-negative integer n $B(n) = [B(1)]^n$. Q has an eigenvalue with positive real parts if and only if e^Q $(=B(1))$ has an eigenvalue λ with $|\lambda| > 1$. Suppose that there is an eigenvalue λ with $|\lambda| > 1$. Then

(2.13) $\quad \limsup_{t \to \infty} |\lambda(t)| = \infty$ where $\lambda(t)$ is an eigenvalue of $B(t)$ corresponding to the eigenvalue λ of $B(1)$. But

$$|\lambda(t)| \leq tr(B(t)B^*(t)) \leq tr(R^{-1}(0)[R^{-1}(0)]\ tr(R(t)R^*(t))$$
$$\leq tr(R^{-1}(0)[R^{-1}(0)]^* \left(\sum_1^q |x_i(0)|^2\right)^2.$$

Therefore for all t $|\lambda(t)|$ is bounded contradicting (2.13).

<u>Sufficiency</u>. Clearly $A(t,s) = B(t-s) = e^{(t-s)Q}(s \leq t)$ satisfies Theorem (B) (2). The proof is now complete.

Theorem 2.2 is proved by Doob in his important paper [2] on elementary Gaussian processes. One of the central problems of his paper is to characterize purely non-deterministic stationary Gaussian Markov processes. We shall give an alternative proof of this result (in our notation) based on Theorems I.5.1 and 2.1. First, we state Doob's theorem in its original form for the sake of comparison with our derivation given in Theorem 4.2.

<u>Theorem D</u> (Theorem 4.3 [2]). If $\underline{x}_t$ is a continuous parameter non-degenerate, continuous in q.m., purely non-deterministic, Gaussian Markov process then

(2.14) $\quad \underline{x}_t = \int_{-\infty}^{t} e^{(t-u)Q}\, S d\xi(u)$ where (i) Q, a $q \times q$ matrix, having no positive real parts is uniquely determined by $R(t)$ (ii) $\{\underline{\xi}(u)\}$ is a Gaussian ξ-processes (see [2] p. 263) with covariance matrix $|u-v|\, U$ where U is a diagonal matrix zero and 1 over diagonal (iv) $R(t) = e^{tQ}R(0)$ for $t \geq 0$ $R(-t) = R(0)e^{tQ^*}$ (v) the matrix Q furnished a solution of the prediction problem (vi) the matrix S is uniquely determined and measures the dispersion of $\underline{x}_t$-process from its predicted value i.e., the variance matrix of $\underline{x}_{u+t} - e^{uQ}\underline{x}_t$ is equal to $R(0) - d^{uQ}R(0)e^{uQ^*} \_ uS^2$ as $u \to \infty$.

Clearly the assertions (iv) (v) (vi) of Theorem D follow from (2.14). Hence it suffices to obtain the representation (2.14) by means of our method.

In concluding this section we point out that the vector-valued stochastic integral $\int F(u)d\underline{\xi}(u)$ where $\underline{\xi}(u)$ is a q-dimensional $\underline{\xi}$-process is defined by Doob ([2], p. 263) for continuous matrix-valued functions F. A complete and rigorous definition of vector-valued stochastic integral is to be found in the recent paper of M. Rosenberg [7]. This definition together with an explanation of the notation employed is given in the next section.

3. <u>Vector Valued Stochastic integrals</u>. If H is a Hilbert-space then $H^{(q)}$ denotes the space of all q x 1 vectors $\underline{h}$ with $h_i \in H$. In $H^{(q)}$ is introduced norm $|||\underline{h}|||^2 = \sum_{i=1}^{q} ||h_i||^2_H$ and an inner produce given by the Gramian matrix $[\underline{h},\underline{h}']$ for any $\underline{h}$, $\underline{h}' \in H^{(q)}$. A linear manifold in $H^{(q)}$ is a non-void subset $\mathcal{M}$ of $H^{(q)}$ such that if $\underline{h}$, $\underline{h}' \in \mathcal{M}$ then $A\underline{h} + B\underline{h}' \in \mathcal{M}$ for all q x q matrices A,B. A subspace of $H^{(q)}$ is a linear manifold closed under the topology $|||\ |||$. For properties of the Gramian and further structural questions we refer the reader to N. Wiener and P. Masani [19].

Let P,Q be any q x M matrix valued functions. Then we say that (P,Q) is integrable with respect to an M x M hermitian matrix valued measure $\underset{\sim}{\rho}$ if the matrix function $P\underset{\sim}{\rho}'Q^*$ is integrable with respect to the $\mathrm{tr}\,\underset{\sim}{\rho}$. We then define $\int P d\underset{\sim}{\rho} Q^* = \int P \underset{\sim}{\rho}' Q^* d\mathrm{tr}$

P is said to be square integrable $[\underset{\sim}{\rho}]$ if $\mathrm{tr}\left(\int P d\underset{\sim}{\rho} P^*\right)$ is finite. If we denote by $\mathcal{L}_2(\underset{\sim}{\rho})$ the class of all measurable P which are square integrable with respect to $\underset{\sim}{\rho}$ where functions P,Q with $\{P(u) - Q(u)\}\underset{\sim}{\rho}'(u) = 0$ a.e. $[\mathrm{tr}\underset{\sim}{\rho}]$ are identified. $\mathcal{L}_2(\underset{\sim}{\rho})$ has the norm $||P||_{\mathcal{L}_2(\ )} = \mathrm{tr}\int P d P^*)$ and gramian

$[P,Q]_{\mathcal{L}_2(\ )} = \int P d\, Q^*$, for all $P,Q \in \mathcal{L}_2(\ )$.

We shall call $\underline{\xi}$ an orthogonally scattered random vector valued measure of dimension M on the real line if for each $B \in \mathcal{B}$, $\underline{\xi}(B) \in L_2^{(M)}(\Omega)$ and for $A,B \in \mathcal{B}$ $[\underline{\xi}(A),\underline{\xi}(B)] = \underset{\sim}{\rho}(A \cap B)$ where $\underset{\sim}{\rho}$ is a hermition matrix valued measure and B the class of Borel sets on the real line. With this set up Rosenberg de-

defines $\int P(u)d\underline{\xi}(u)$ for $P \in \mathcal{L}_2(\underset{\sim}{\rho})$ in the same way as Doob does for $M = q$ (See J. L. Doob [3] p. 596). Further if one denotes by $\mathcal{L}_2(\underline{\xi})$ the subspace of of $L_2^{(q)}(\Omega)$ generated by $\{\underline{\xi}(B), B \in B\}$ with $q \times M$ matrices as coefficients then we have the following [See [7] Theorem 4.6]

<u>Theorem R.</u> The correspondence $P \rightarrow \int P d\, \underline{\xi}$ is an isomorphism from $\mathcal{L}_2(\underset{\sim}{\rho})$ to $\mathcal{L}_2(\underline{\xi})$.

<u>Remark</u> In the above discussion q and M are fixed positive integes with ($M \leq q$) and the space $\mathcal{L}_2(\underset{\sim}{\rho})$ is complete in the norm defined.

4. Purely non-deterministic wide-sense Markov processes. We first prove a representation for the non-stationary case.

Theorem 4.1 If $\underline{x}_t$ is a continuous parameter purely non-deterministic process statisfying assumption (D) then it is wide-sense Markov if and only if

$$x_i(t) = \sum_{k=1}^{q} \sum_{j=1}^{M} \int_{-\infty}^{t} \Psi_{ik}(t)\, h_{kj}(u) dz_j(u) \quad \text{where } \{\Psi_{ik}(t)\}(i,k=1,\ldots q) \text{ are}$$

elements of a non-singular $q \times q$ matrix $\underline{\Psi}(t)$, $h_{kj}(.)$ for each j belong to to $L_2(\rho_j)$ with z_i, ρ_j having the same meaning as in Theorem I.2.2, M is the multiplicity and for ever k, $\sum_{j=1}^{M} \int_{-\infty}^{t} |h_{kj}(u)|^2 dP_j(u)$ is finite. Also $H(\underline{z};t)=H(\underline{x};t)$ for t.

Proof, Necessity. As stated in Theorem 2.1 $\underline{x}_t = \underline{\Psi}(t)\underline{u}_t$ with $L_2(\underline{x};t) = H(\underline{u};t)$. Also from Theorems 1.2.2 and 1.3.1 we have a representation for $\underline{x}_t$-process with $L_2(\underline{x};t) = H(\underline{z};t)$. Since $\underline{u}_t$ is a wide-sense martingate and $H(\underline{z};t) = H(\underline{u};t)$ we have $u_k(t) = \sum_{j=1}^{M} \int_{-\infty}^{t} h_{kj}(u) dz_j(u)$. The result now follows, since

$$x_i(t) = \sum_{k=1}^{q} \Psi_{ik}(t) u_k(t)$$

for all t and $\underline{\Psi}(t)$ is a non-singular $q \times q$ matrix.

Sufficiency. Define $u_k(t) = \sum_{1}^{M} \int_{-\infty}^{t} h_{kj}(u) dz_i(u)$. Clearly $\underline{x}_t = \underline{\Psi}(t)\underline{u}_t$.

Therefore, to complete the proof it suffices to show that $\underline{u}_t$ is a wide-sense martingale. We note that since $\underline{\Psi}(t)$ is non-singular $L_2(\underline{x};t) = H(\underline{u};t)$. As we are given that $L_2(\underline{x};t) = L_2(\underline{z};t)$, we get $L_2(\underline{u};t) = L_2(\underline{z};t)$ for every t. Consider now for $s < t$,

$$P_{H(\underline{u};s)}\,(u_k(t) - u_k(s)) = P_{H(\underline{z};s)}\,\left[\sum_{1}^{M} \int_{s}^{t} h_{kj}(u) dz_i(u)\right] = 0,$$

where the last inequality follows because z_j's are mutually orthogonal processes with orthogonal increments. The proof is now complete.

For stationary purely non-deterministic processes we recall that M, the multiplicity of the process does not exceed q [See Theorem 1.6.2]. Also from Theorem 1.6.2 and the definition of vector valued stochastic integrals we have

$$(4.1) \qquad \underline{x}_t = \int_{-\infty}^{t} F(t-u)d\underline{\xi}(u)$$

where $F(t-u)$ is a $q \times M$ matrix-valued function and $\underline{\xi}(u)$ is an M-dimensional orthogonally scattered measure. Also we have $L_2(\underline{x};t) = L_2(\underline{\xi};t)$ for each t. Using representations of Theorem 2.1 and an argument similar to that of Theorem 4.1 (Necessity), we obtain that $\underline{u}_t = \int_{-\infty}^{t} H(u)d\underline{\xi}(u)$ where $H(u)$ is a $q \times M$ matrix function and hence

$$(4.2) \qquad \underline{x}_t = \int_{-\infty}^{t} \tilde{\underline{\Psi}}(t)\ H(u)\ d\underline{\xi}(u) \ ,$$

with $L_2(\underline{x};t) = H(\xi;t)$ for each t. We have the following theorem:

<u>Theorem 4.2.</u> Let $\underline{x}_t(-\infty < t < +\infty)$ be a stationary q-dimensional process satisfying assumption (D). Then $\underline{x}_t$ is wide-sense Markov and purely non-deterministic if and only if

$$(4.3) \qquad \underline{x}_t = \int_{-\infty}^{t} e^{(t-u)Q} C d\underline{\xi}(u) \ , \qquad \text{where}$$

(i) Q is a $q \times q$ constant matrix with properties described in Theorem 2.2

(ii) C is a $q \times M$ constant matrix where M equals the rank of the process

(iii) $\underline{\xi}_t$ is an orthogonally scattered random measure such that $[\underline{\xi}(B), \underline{\xi}(B')] = \mu(B \cap B')I$ where B, B' are real Borel sets, μ Lebesgue measure and I is an $M \times M$ identity matrix. Further $L_2(\underline{x};t) = L_2(\underline{\xi};t)$.

<u>Proof</u>. <u>Necessity</u>. From (4.2) and stationarity we have a $q \times M$ measurable

matrix function $G(t-u)$ such that for $u \leq t$

$$G(t-u) = \underline{\Psi}(t)\, H(u)$$

Since $H(u)$ is given almost everywhere, if it is not defined at the origin, completing its definition at zero we obtain for $t \geq 0$

$$G(t) = \underline{\Psi}(t)\, H(0)$$

However since $R(t) = e^{tQ}R(0)$ from (2.8) and (2.9) we get for $t \geq 0$

$$(4.4) \qquad G(t) = e^{tQ}\, \psi(0)\, H(0) \quad \text{i.e.} \quad G(t-u) = e^{(t-u)Q}\, C (u \leq t)$$

where $C = \underline{\psi}(0)\, H(0)$. Hence from (4.2) and (4.4),

$$\underline{x}_t = \int_{-\infty}^{t} e^{(t-u)Q}\, C d\underline{\xi}(u) \quad \text{with } L_2(\underline{x};t) = H(\underline{\xi};t)$$

<u>Sufficiency</u>: If we denote by $\underline{u}_t = \int_{-\infty}^{t} e^{-uQ}\, d\underline{\xi}(u)$. Then obviously $\underline{u}_t$ is a q-dimensional wide-sense martingale and therefore from Theorem 2.1 it follows that $\underline{x}_t$ is wide-sense Markov since $e^{tQ} = R(t)R^{-1}(0)$ is invertible. The proof is complete if we show that $\underline{x}_t$ is purely non-deterministic. But this is obvious from the fact

$$\bigcap_t H(\underline{x};t) = \bigcap_t H(\underline{\xi};t) = \bigcap_t \sum_{i=1}^{M} \oplus H(\xi_i;t) = \{0\}$$

which follows because $\{\xi_i(t)\}$ $(-\infty < t < +\infty)$ $(i = 1,2,\ldots,M)$ are mutually orthogonal processes with stationary orthogonal increments.

Since the Gaussian wide-sense Markov processes are Markov processes, Theorem 4.2 reduces to Theorem D. The $\underline{\xi}_t$-process occuring in the expression (4.3) is an M-dimensional orthogonally scattered measure where M is the rank of $\underline{x}_t$ as defined by E. G. Gladyshev [4]. Its covariance matrix function $\Delta(u,v)$ is of the form $|u-v|I$ where I is the $M \times M$ identity matrix. Therefore Theorem 4.2 renders a more precise form of Theorem D (ii).

N-PLE MARKOV PROCESSES

In the study of representations of N-ple Markov processes we require analytic conditions for proper canonical property.

5. An analytical characterization of a proper canonical representation. Henceforth we shall assume $M \leq q$. Further we denote by $\mathcal{L}_2(\underline{z};t)$ [$\underline{z}(\cdot)$ is a q-dimensional orthogonally scattered vector measure] the subspace of $L^{(q)}(\Omega)$ generated by $\{\underline{z}(B)$, B a Borel subset of $(-\infty, t]\}$ with coefficients $q \times M$ matrices.

Lemma 5.1. $H^{(q)}(\underline{z};t) = \mathcal{L}_2(\underline{z};t) \quad (-\infty < t < +\infty)$.

Proof. A typical element of $H^{(q)}(\underline{z};t)$ is a column vector $(y_1,\ldots,y_q)^* = (y_1,0,\ldots,0)^* + (0,y_2 0,0,\ldots,0)^* + \ldots + (0,\ldots,y_q)^*$ where $y_j \in H(\underline{z};t)$. It suffices therefore to prove that for each i, the vector $(0,0,\ldots,z_i(B),0,\ldots,0)^*$ for each Borel set B belongs to $\mathcal{L}_2(\underline{z};t)$. But this is obviously the case as is seen by taking a diagonal $q \times M$ matrix with unity in the i^{th} place in the diagonal and zero everywhere else. The fact that $\mathcal{L}_2(\underline{z};t) \subset H^{(q)}(\underline{z};t)$ follows by observing that for each Borel set B in $(-\infty,t]$ and $q \times M$ matrix A $A\underline{z}(B) \in H^{(q)}(\underline{z};t)$.

The following is a direct extension of Theorem I.7 of Hida [5], to q-dimensional processes with $M \leq q$. We shall denote a representation for such processes by $\{F(t,u),d\underline{z}(u)\}$ where $F(t,u)$ is a $q \times M$ matrix function and $\underline{z}(B)$ is an M-dimensional orthogonally scattered random vector measure with components $z_i(B)$ $(i = 1,2,\ldots,M)$.

The notion of a proper canonical representation of arbitrary multiplicity M has already been introduced in Chapter I. Under the assumption $M \leq q$ we give necessary and sufficient analytical conditions for a proper canonical representation.

Theorem 5.1. A canonical representation $\{F(t,u),d\underline{z}(u)\}$ is proper if and only if for any real t_0

(5.1) $$\int_{-\infty}^{t} P(u)d\ (u)F^*(t,u) = 0 \text{ for } t \leq t_0 \text{ implies } P(u) = 0 \ \text{a.e. } [\underset{\sim}{\rho}]$$

where is the hermitian $M \times M$ matrix valued measure $\underset{\sim}{\rho}(B) = [\underline{z}(B), \underline{z}(B)]$ and $P(u)$ is a square integrable $q \times M$ matrix-valued function on the real line.

Proof Sufficiency. Let (5.1) hold and let t_o be such that $H(\underline{z};t_o) \neq L_2(\underline{x};t_o)$ and we know that $L_2^{(q)}(\underline{x};t_o) \subset H^{(q)}(\underline{z};t_o)$. Therefore, there is a $\underline{V} \neq \underline{0}$ in $H^{(q)}(\underline{z};t_o)$ such that $[\underline{V},\underline{x}_t] = 0$ for $t \leq t_o$. Consider now $H^{(q)}(\underline{z};t_o) = \mathcal{L}_2(\underline{z};t_o)$. Then by Theorem R of Section 3 we have $\underline{V} = \int_{-\infty}^{t_o} P(u)d\ (u) \neq 0$ such that for all

$t(\leq t_o)$, $\int_{-\infty}^{t} P(u)d\ (u)\ F^*(t,u) = 0$. By (5.1) we get $P(u) = 0$ a.e. $[\underset{\sim}{\rho}]$

contradicting $\underline{V} \neq \underline{0}$.

Necessity. Suppose that $H(\underline{z};t) = L_2(\underline{x};t)$ for all t, and let t_o be a real number such that

$$(5.2) \qquad \int_{-\infty}^{t} P(u)d\ (u)\ F^*(t,u) = 0 \text{ for every } t \leq t_o.$$

Observe that since from the proper cannonical property $L_2^{(q)}(\underline{x};t_o) = H^{(q)}(\underline{z};t_o) = \mathcal{L}_2(\underline{z};t_o)$ the vector $\underline{V} = \int_{-\infty}^{t_o} P(u)d\underline{z}(u)$ belongs to $L_2^{(q)}(\underline{x};t_o)$. But (5.2) implies that $[\underline{V}, \underline{x}_t] = 0$ for all $t \leq t_o$. Hence $\underline{V} = \underline{0}$ giving $P(u) = 0$ a.e. $[\underset{\sim}{\rho}]$. This proves the theorem.

The above criterion will be useful in our discussion of N-ple Markov processes.

6. Finite dimensional wide-sense N-ple Markov processes. In the definition of vector valued wide-sense N-ple Markov processes we require the concept of the projection on a subspace of $L_2^{(q)}(\underline{x})$. We recall here a lemma due to N. Wiener and P. Masani [10], which proves the existence of the projection of an element $\underline{h}$ and gives its structure. The notation used is that of Section 3.

Lemma WM. (Lemma 5.8 [10]). (a). If $\mathfrak{M}$ is a subspace of $H^{(q)}$ there exists a subspace $\mathcal{M}$ of H such that $\mathfrak{M} = \mathcal{M}^{(q)}$, where $\mathcal{M}^{(q)}$ denotes the Cartesian product $\mathcal{M} \otimes \ldots \otimes \mathcal{M}$ with q-factors. $\mathcal{M}$ is a set of all components of all elements in $\mathfrak{M}$. (b). If $\mathfrak{M}$ is a subspace of $H^{(q)}$ and $\underline{h} \in H^{(q)}$, then there exists a unique $\underline{h}' \in \mathfrak{M}$ such that $||\underline{h} - \underline{h}'||_{H^{(q)}} \leq ||\underline{h} - \underline{g}||_{H^{(q)}}$ for all $\underline{g} \in \mathfrak{M}$. For this $\underline{h}'$, $h_i' = P_{\mathcal{M}} h_i$, $\mathcal{M}$ being as in (a). An element $\underline{h}'$ satisfies the preceding condition if and only if $\underline{h} - \underline{h}' \perp \mathfrak{M}$ where orthogonality is in the sense of the Gramian. (c). If $\mathfrak{M}, \mathfrak{N}$ are subspaces of $H^{(q)}$ and $\mathfrak{M} \subset \mathfrak{N}$, then there exists a unique subspace $\mathfrak{M}' \subset \mathfrak{N}$ such that $\mathfrak{N} = \mathfrak{M} \oplus \mathfrak{M}'$ and $\mathfrak{M}$ is orthogonal to $\mathfrak{M}'$. Parts (d) and (e) of Lemma 5.8 of [10] are not given here because they will not be referred to. Following Wiener and Masani we give

Definition 6.1 The unique element $\underline{h}'$ of Lemma WM (b) is called the orthogonal projection of $\underline{h}$ onto $\mathfrak{M}$ and is denoted by $(\underline{h}|\mathfrak{M})$.

Extending usual idea of linar independence, we give following definition of linearly independent vectors $\underline{h}_1, \underline{h}_2 \ldots \underline{h}_N \in H^{(q)}$.

Definition 6.2. The vectors $\underline{h}_i \in H^{(q)}$ (i=1,2..N) are linearly independent in $H^{(q)}$ if for any q x q matrices $A_1, \ldots A_N$, $\sum A_i \underline{h}_i = \underline{0}$ and $A_i \underline{h}_i$ is different from the zero element of $H^{(q)}$ for at last one i implies that A_i are zero matrices.

Now we define a q-dimensional real continuous parameter widesense N-ple Markov process. For one-dimensional continuous parameter Gaussion processes the definition is due to Hida [5] and for discrete Gaussion processes the definition goes back to Doob [2].

Definition 6.3. We say that a q-dimensional continuous parameter process is wide-sense N-ple Markov if for any sequence $\{t_i\}$ of N-real numbers $(t_1 < t_2 <.. < t_N)$ and for $t_0 \leq t_1$, the vectors $(\underline{x}_{t_i} \mid L_2^{(q)}(\underline{x};\ t_0))$ are linearly independent in $L_2^{(q)}(\underline{x};\ t_0)$, and the vectors $(\underline{x}_{t_i} | L_2^{(q)}(\underline{x};\ t_0))$ are linearly dependent if i=1, 2... N +1 and $t_{N+1} > t_N$.

We now proceed to the extension of Theorem II.2 of Hida, to obtain a representation for a q-dimensional (not necessarily stationary) wide sense N-ple Markov process using the theory of Chapter I.

Lemma 6.1 Let t and s(s < t) be any real numbers. If $\Gamma(t,s)$ is non-singular, then the vector $(\underline{x}_t | L_2^{(q)}(\underline{x};s))$ is non-degenerate, i.e., its covariance matrix is non-singular.

Proof. From Lemma WM with $\mathcal{M} = L_2^{(q)}(\underline{x};\ s)$ we get $(\underline{x}_t | L_2^{(q)}(\underline{x};s))$ is the column vector $(P_{L_2(\underline{x};s)}x_1(t),\ldots,P_{L_2(\underline{x};s)}x_q(t))^*$. First we observe that none of the elements $P_{L_2(\underline{x};\ s)}x_i(t)$ $(i=1,2,\ldots,q)$ can be zero; for otherwise $\Gamma_{ij}(t,s) = \mathcal{E}(x_i(t)\ x_j(s)) = \mathcal{E}(x_i(s)\ P_{L_2(\underline{x};s)}x_j(t)) = 0$ for all $j=1,2,\ldots,q$ contradicting the non-singularity of $\Gamma(t,s)$. If the vector is degenerate then for some i, $P_{L_2(\underline{x};s)}x_i(t) = \sum a_{ij}\ P_{L_2(\underline{x};s)}x_j(t)$, Also $P_{L_2(\underline{x};s)}x_i(t) \neq 0$. Hence there is at least one j such that $a_{ij} \neq 0$. Now $\mathcal{E}(x_i(t)x_k(s) =$ $\sum a_{ij}\ \mathcal{E}\left[P_{L_2(\underline{x};s)}x_j(t)x_k(s)\right]$ Thus $\Gamma_{ik}(t,s) = \sum_{i\neq j} a_{ij}\ \mathcal{E}\left[P_{L_2(\underline{x};s)}x_j(t)x_k(s)\right] = \sum_{i\neq j} a_{ij}\Gamma_{jk}(t,s)$, $(k=1,2,\ldots,q)$. This contradicts the non-singularity of $\Gamma(t,s)$. and the lemma is proved.

From the definition of wide-sense N-ple Markov processes it follows that if $\{s_i\}$ $(s_1 < s_2 < ,\ldots, < s_N)$ is a given sequence and $\tau > s_N$, then for each $s_0 \leqq s_1$, there exist q x q matrices $A_j(\tau; s_1,\ldots,s_N)$ such that $(\underline{x}_\tau | L_2^{(q)}(\underline{x}; s_0)) = \sum_{k=1}^{N} A_k(\tau; s_1,\ldots,s_N)(\underline{x}_{s_k} | L_2^{(q)}(\underline{x}; s_0))$. Taking a sequence $\{t_j\}$ $(t_N > t_{N-1},\ldots, > t_1 > s_N)$ we have

$$(6.1) \qquad (\underline{x}_{t_j} | L_2^{(q)}(\underline{x}; s_0)) = \sum_{k=1}^{N} A_k(t_j; s_1,\ldots,s_N)(\underline{x}_{s_k} | L_2^{(q)}(\underline{x}; s_0)).$$

Denote by $\hat{A}(\underset{\sim}{t},\underset{\sim}{s})$ the qN x qN matrix having $A_k(t_j; s_1,\ldots,s_N)$ as its $(k,j)^{th}$ (q x q) block matrix, $(k,j = 1,2,\ldots,N)$. Then we have the following lemma.

Lemma 6.2. If $\underline{x}_t(-\infty < t < +\infty)$ is a q-dimensional wide-sense N-ple Markov process satisfying assumption (D.2) then $\hat{A}(\underset{\sim}{t},\underset{\sim}{s})$ is non-singular.

Proof. We first prove that for any sequence $\{t_i\}(t_N > t_{N-1} > \ldots > t_1 > s_0)$ the set

(6.2) $\{P_{L_2(\underline{x};\, s_0)} x_i(t_j)\}$ $i=1,2,\ldots,q$, $j=1,2,\ldots,N$, is linearly independent in $L_2(\underline{x})$. If not, then there exist a_{ij} not all zero such that $\sum_{i,j} a_{ij}\, y_i(t_j) = 0$ where we write $y_i(t_j) = P_{L_2(\underline{x};\, s_0)} x_i(t_j)$, ($s_0$ being fixed throughout the argument). Since from Lemma 6.1, for no pair i,j $y_i(t_j) = 0$ letting $a_{ij} \neq 0$, we have

$$(6.3) \qquad y_i(t_j) = \sum_{k,m}^{*} b_{km}\, y_k(t_m) \qquad \text{where } \sum_{k,m}^{*} \text{ denotes the}$$

summation over all k,m $(k=1,\ldots,q;\ m=1,\ldots,N)$ such that no pair $(k,m) = (i,j)$; though b_{km} depends on (i,j) we do not indicate it here in order to keep the notation simple. Also since $y_i(t_j) \neq 0$ (Lemma 6.1) there is at least one

$(k,m) \neq (i,j)$ such that $b_{km} \neq 0$. We now consider the following two possibilities.

Case I. Suppose $b_{kj} = 0$ for all $k(\neq i)$.

Then (6.3) has the form

(6.4) $$y_i(t_j) = \sum_{\substack{k,m \\ (m \neq j)}}^{*} b_{km}\, y_k(t_m).$$ Consider now q x q matrices

$A_\ell (\ell=1,2,\ldots,N)$ such that $A_j = ((a_{np}^{(j)}))$, $a_{1i}^{(j)}=1$ and $a_{np}^{(j)} = 0$ otherwise; for $\ell \neq j$ $A_\ell = ((a_{np}^{(\ell)}))$ with $a_{1p}^{(\ell)} = -b_{p\ell}$ for $p = 1,2,\ldots,q$ and $a_{np}^{(\ell)} = 0$ otherwise.

Then from (6.4) we have $\sum_{\ell=1}^{N} A_\ell \underline{y}_{t_\ell} = \underline{0}$, $A_j \underline{y}_{t_j} \neq \underline{0}$ and A_j is not a zero matrix; i.e., the vectors $(\underline{x}_{t_\ell} | L_2^{(q)}(\underline{x};s_0))$ $(\ell=1,2,\ldots,N)$ are linearly dependent. This contradicts the definition of the wide-sense N-ple Markov process.

Case II. There is a non-void subset $J \subset \{1,2,\ldots,q\}$ such that $b_{kj} \neq 0$ $k \in J$ $(i \notin J)$.

If the element

(6.5) $$y_i(t_j) - \sum_{k \in J} b_{kj}\, y_k(t_j)$$

is zero then for $\nu = 1,2,\ldots,q$ we have

$$[y_i(t_j)\, y_\nu(t_j)] = \sum_{k \in J} b_{kj}\, (y_k(t_j)\, y_\nu(t_j)).$$ But this contradicts Lemma 6.1. Hence the element given by (6.5) is not zero. We now rewrite (6.3) as

(6.6) $$y_i(t_j) - \sum_{k \in J} b_{k_j}\, y_k(t_j) = \sum_{k,m}^{*} b_{km}\, y_k(t_m)$$

Now introduce the matrices $A_\ell = ((a_{np}^{(\ell)}))$ where

(i) $\ell = j$, $a_{1p}^{(j)} = -b_{pj}$ $(p \in J)$, $a_{1i}^{(j)} = 1$ and $a_{np}^{(j)} = 0$ otherwise;

(ii) $\ell \neq j$ $a_{1p}^{(\ell)} = -b_{p\ell}$ $(p = 1,2,\ldots,q)$ and $a_{np}^{(\ell)} = 0$ otherwise. Then (6.6) becomes

(6.7) $$\sum_{\ell=1}^{N} A_\ell \underline{y}_{t_\ell} = \underline{0}.$$

(. .)). Further $A_j \underline{y}_{t_j} \neq \underline{0}$ since the element in (6.5) has been shown to be non-zero. As in the concluding part of Case I, these facts imply a contradiction of the N-ple Markov property. Thus we have established the linear independence (in $L_2(\underline{x})$) of the set (6.2). By a similar argument the set $\{y_i(s_k),\ i = 1,2,\ldots,q,\ k = 1,2,\ldots,N\}$ is linearly independent in $L_2(\underline{x})$. Also we can write (6.1) as

$$(6.8)\quad \begin{cases} (y_1(t_1),\ y_2(t_1),\ldots,\ y_q(t_1),\ldots,\ y_1(t_N),\ldots,\ y_q(t_N))^* \\ = \hat{A}(\underset{\sim}{t},\underset{\sim}{s})\ (y_1(s_1),\ y_2(s_1),\ldots,\ y_q(s_1),\ldots,\ y_1(s_N),\ldots,\ y_q(s_N))^*. \end{cases}$$

Hence $\hat{A}(\underset{\sim}{t},\underset{\sim}{s})$ is non-singular. This completes the proof of Lemma 6.2.

We now state the main result of this section.

Theorem 6.1 Let $\{\underline{x}_t\}$ be a real continuous parameter purely non-deterministic q-dimensional wide-sense N-ple Markov process with multiplicity $M \leq q$ and satisfying the assumption (D). Then

$$(6.9)\qquad \underline{x}_t = \sum_{i=1}^{N} \int_{-\infty}^{t} \underline{\Psi}_i(t)\, G_i(u)\, d\underline{z}(u) \qquad \text{where for each } i,$$

$\underline{\Psi}_i(\cdot)$ is a q x q matrix-valued function such that for any N points $\{t_i\}$ $(t_1 < t_2 <,\ldots, < t_N)$ the qN x qN matrix with $(i,j)^{th}$ q x q block matrix $\{\underline{\Psi}_i(t_j)\}$ is non-singular and $G_i(u)$ is a q x M matrix valued function in $\mathcal{L}_2(\underset{\sim}{\rho})$ $(\underset{\sim}{\rho}(B) = [\underline{z}(B),\ \underline{z}(B)]_{L_2^{(M)}(\underline{x})})$. The functions $\{G_i(u)\}$ are linearly independent in $\mathcal{L}_2(\underset{\sim}{\rho};(-\infty,t])$ i.e. for each t, and for any q x q matrices A_i, $\sum_{i=1}^{N} A_i\, G_i(u) = 0$ ($G_i(.)$ restricted to $(-\infty,t]$) and $A_i\, G_i(u) \neq 0$ for at least one i implies $A_i = 0$ for all i.

Proof. By Theorem I.2.2 and Theorem I.3.1, $\underline{x}_t$ has a proper comonical representation of multiplicity M. Since $M \leq q$ this representation can be expressed as $\{F(t,u), d\underline{z}\}$ where $F(t,\cdot)$ is a $q \times M$ matrix-valued function in $\mathcal{L}_2(\underset{\sim}{\rho})$. Let $\{t_i\}$ be a sequence of distinct points with $t_N > t_{N-1} >, \ldots, > t_1$ and $\tau > t_N$. Then by the wide-sense N-ple Markov property for all $\sigma \leq t_1$ there exist $q \times q$ matrices $\{A_i(\tau; t_1 \ldots, t_N)\}_{j=1,2,\ldots,N}$. not all zero such that

$$\underline{x}_\tau - \sum_{j=1}^{N} A_j(\tau; t_1, \ldots, t_N)\, \underline{x}_{t_j} \perp L_2^{(q)}(\underline{x}; \sigma) \quad (\sigma \leq t_1)$$

where orthogonality is in the Gramian sense. Hence for all $\sigma \leq t_1$, we obtain

$$0 = [\underline{x}_\tau - \sum_{j=1}^{N} A_j(\tau; t_1 \ldots t_N)\underline{x}_{t_j},\ \underline{x}_\sigma]_{L_2^{(q)}(\underline{x})} = \int_{-\infty}^{\sigma} [F(\tau,u) - \sum_{j=1}^{N} A_j(\tau; t, \ldots, t_N) F(t_j, u)]$$

$d\underset{\sim}{\rho}(u)\, F^*(\sigma,u)$. Hence by Theorem 5.1,

$$(6.10) \qquad F(\tau,u) = \sum_{j=1}^{N} A_j(\tau; t_1 \ldots t_N)\, F(t_j,u)\ [\underset{\sim}{\rho}; (-\infty, t_1]], \qquad \text{since}$$

the representation $\{F(t,u), d\underline{z}(u)\}$ is proper canonical. (In (6.10) $[\underset{\sim}{\rho}; (-\infty, t_1]]$ means almost everywhere $[\underset{\sim}{\rho}]$ on the interval $(-\infty, t_1]$.). If we have another sequence $\{s_k\}$ $(t_1 > s_N > \ldots > s_1)$ then from (6.10) we obtain

$$(6.11) \qquad F(t_j,u) = \sum_{k=1}^{N} A_k(t_j; s_1 \ldots s_N F(s_k,u) \quad [\underset{\sim}{\rho}; (-\infty, s_1]].$$

Now from the definition of $A_k(t_j; s_1 \ldots s_N)$ $(k, j=1,2,\ldots,N)$ and Lemma 6.2 the matrix $\hat{A}(\underset{\sim}{t},\underset{\sim}{s})$ defined there is nonsingular. Let $\hat{B}(\underset{\sim}{s},\underset{\sim}{t}) = \hat{A}^{-1}(\underset{\sim}{t},\underset{\sim}{s})$. From (6.10) and (6.11) we deduce

$$(6.12) \qquad F(\tau;u) = \sum_{j,k} A_j(\tau; t, \ldots, t_N)\, A_k(t_j; s_1 \ldots s_N)\, F(s_k, u)$$

$$= \sum_{k} A_k(\tau; s_1 \ldots s_N)\, F(s_k, u) \quad [\underset{\sim}{\rho}; (-\infty, s_1)].$$

Now (6.12) implies that

$$\sum_k \Big(\sum_j A_j(\tau;t_1 \ldots t_N)\, A_k(t_j;s_1 \ldots s_N) - A_k(\tau;s_1 \ldots s_N)\Big) F(s_k, u) = 0 \; [\underset{\sim}{\rho};(-\infty,s_1]],$$

which can be rewritten as (with sequence $\{t_i\}$ $\{s_i\}$ and number τ fixed)

$$(6.13) \qquad \sum_k C_k F(s_k;u) = 0 \qquad [\underset{\sim}{\rho};(-\infty,s_1)].$$

Consider $(\underline{x}_{s_k} \mid L_2^{(q)}(\underline{x};s_1))$. Since by the canonical property

$$P_{L_2(\underline{x};s_1)} x_i(s_k) = \sum_{i=1}^{q} \sum_{j=1}^{M} \int_{-\infty}^{s_1} f_{ij}(s_k,u)\,dz_j(u), \text{ we get}$$

$(\underline{x}_{s_k} \mid L_2^{(q)}(\underline{x};s_1)) = \int_{-\infty}^{s_1} F(s_k,u)\,d\underline{z}(u)$. Now if in (6.13) $C_k\, F(s_k,u) = 0\; [\underset{\sim}{\rho};(-\infty,s_1]]$

and $C_k \neq 0$ then we get $C_k(\underline{x}_{s_k} | L_2{}^{(q)}(\underline{x};\, s_1)) = \underline{0}$. This contradicts Lemma 6.1.

Hence C_k is a zero matrix each k by the wide sense N-ple Markov property and (6.13)

Thus

$$(6.14) \qquad A_k(\tau;\, s_1 \ldots s_N) = \sum_{j=1}^{N} A_j(\tau;\, t_1 \ldots t_N)\, A_k(t_j;\, s_1 \ldots s_N)\,.$$

If $\mathcal{A}(\tau;\, \underset{\sim}{s})$ denotes q x qN matrix with q x q block matrices $A_k(\tau;\, s_1 \ldots s_N)$,

Viz , $\mathcal{A}(\tau;\underset{\sim}{s}) = \{ A_1(\tau s_1 \ldots s_N), \ldots, A_N(\tau;\, s_1 \ldots s_N)\}$ then (6.14) can be

expressed as

$$(6.15) \qquad \mathcal{A}(\tau;\, \underset{\sim}{s}) = \mathcal{A}(\tau;\, \underset{\sim}{t})\ \hat{A}(\underset{\sim}{t},\underset{\sim}{s}).$$

Recalling that $\hat{B}(\underset{\sim}{s},\underset{\sim}{t}) = \hat{A}^{-1}(\underset{\sim}{t},\underset{\sim}{s})$ we define

$$(6.16) \qquad \hat{\underline{\Psi}}_{\underset{\sim}{s}}(\tau) = \mathcal{A}(\tau;\underset{\sim}{s})\ \hat{B}(\underset{\sim}{s},\underset{\sim}{t}).$$

If $s_1' < s_2' < \ldots < s_N' - s_1 \ldots < s_N < t_1 < \ldots < t_N < \tau_1$ then we get

$\hat{\underline{\Psi}}_{\underset{\sim}{s}'}(\tau) = \mathcal{A}(\tau;\underset{\sim}{s}')\,\hat{B}(\underset{\sim}{s};\, \underset{\sim}{t}) = \mathcal{A}(\tau;\underset{\sim}{s}')\,\hat{B}(\underset{\sim}{s}',\underset{\sim}{s})\,\hat{B}(\underset{\sim}{s},\underset{\sim}{t}) = \mathcal{A}(t,\underset{\sim}{s})\,\hat{B}(\underset{\sim}{s},\underset{\sim}{t})$ since

$\hat{A}(\underset{\sim}{t},\underset{\sim}{s}') = \hat{A}(\underset{\sim}{t},\underset{\sim}{s})\,\hat{A}(\underset{\sim}{s},\underset{\sim}{s}')$ from (6.15). Hence (6.15) and (6.16) give

$\hat{\underline{\Psi}}_{\underset{\sim}{s}'}(\tau) = \hat{\underline{\Psi}}_{\underset{\sim}{s}}(\tau)$. Let $\mathcal{S}$ be the set of all sequence $\underset{\sim}{s} = \{s_i\}$ where

$s_1 < s_2 < \dots < s_N < \tau$, τ being fixed throughout. For any two sequences $\underset{\sim}{s}$, $\underset{\sim}{s}'$ in S define the relation $\prec$ as follows: $s \prec s$ if $s_N' < s_1$. it is easy to see that $\prec$ is a direction on the set $\mathcal{S}$ of all such sequences. Further for each τ the limit of the net $\{\hat{\underline{\bar{\Psi}}}_{\underset{\sim}{s}}(\tau), s \in \ \}$ exist from the fact, proved above, that for $\underset{\sim}{s}' < \underset{\sim}{s} < \tau$ $\hat{\underline{\bar{\Psi}}}_{\underset{\sim}{s}}(\tau) = \hat{\underline{\bar{\Psi}}}_{\underset{\sim}{s}'}(\tau)$. Denoting this limit by $\hat{\underline{\bar{\Psi}}}(\tau)$ we find from (6.16), (6.15) and the non-singularily of $\hat{A}(\underset{\sim}{t},\underset{\sim}{s})$ that the qN x qN matrix $\{\underline{\bar{\Psi}}_i(t_j)\}$ of the theorem is non-singular where $\underline{\bar{\Psi}}_i(\tau)$ denotes the ith block q x q matrix of $\hat{\underline{\bar{\Psi}}}(\tau)$. We write equation (6.10) as

$$(6.17) \qquad F(\tau,u) = \mathcal{A}(\tau;\underset{\sim}{t})\,\mathcal{F}(\underset{\sim}{t};u), \ [\underset{\sim}{\rho};(-\infty,t_1]]$$

where $\mathcal{F}(\underset{\sim}{t},u)$ denotes the qN x M matrix $(F(t_1,u),\dots,F(t_N,u))^*$ Let $\hat{G}(u,\underset{\sim}{s},\underset{\sim}{t})$ be the qN x M matrix $\hat{B}^{-1}(\underset{\sim}{s},\underset{\sim}{t})\,\mathcal{F}(\underset{\sim}{t},u)$. Then (6.17) takes the form

$$(6.18) \qquad F(\tau,u) = \hat{\underline{\bar{\Psi}}}(\tau)\,\hat{G}(u,\underset{\sim}{s},\underset{\sim}{t}), \quad a.e.[\underset{\sim}{\rho};(-\infty,t_1]].$$

Let $\{t_i\}$ $(i=1,2,\dots,N)$ and $\{s_j'\}$ $(j=1,2,\dots,N)$ be sequences in with $\underset{\sim}{s}' < \underset{\sim}{t}'$ then

$$(6.19) \qquad F(\tau,u) = \hat{\underline{\bar{\Psi}}}(\tau)\,\hat{G}(u;\underset{\sim}{s}',\underset{\sim}{t}') \qquad [\underset{\sim}{\rho};(-\infty,t_1]].$$

Now from equations (6.18) and (6.19) and the non-singularity of $\{\underline{\bar{\Psi}}_i(t_j)\}$ we obtain

$$\hat{G}(u,\underset{\sim}{s},\underset{\sim}{t}) = \hat{G}(u,\underset{\sim}{s}',\underset{\sim}{t}'). \quad [\underset{\sim}{\rho};-\infty,t_1]$$

Hence we may set

$$(6.20) \qquad \hat{G}(u,\underset{\sim}{s}',\underset{\sim}{t}') = \hat{G}(u); \text{ say, for all } \underset{\sim}{s}', \underset{\sim}{t}' \in \mathcal{S}.$$

Hence from (6.18) and (6.20)

$$F(\tau,u) = \sum_{i=1}^{N} \underline{\bar{\Psi}}_i(\tau)\,\bar{G}_i(u) \qquad [\underset{\sim}{\rho};(-\infty,\tau]]$$

for each $t_1 < \tau$. Also $\lim_{t_1 \to \tau} ||F(t_1,u) - F(\tau,u)||_{\alpha_2(\underset{\sim}{\rho})} = 0.$

Therefore

$$F(\tau,u) = \sum_{i=1}^{N} \bar{\Psi}_i(\tau)\, G_i(u)\ , \quad [\rho;(-\infty,\tau]].$$

Thus (with τ replaced by t) we get

$$\underline{x}_t = \sum_{i=1}^{N} \int_{-\infty}^{t} \bar{\Psi}_i(t)\, G_i(u)\, d\underline{z}(u).$$

To complete the proof we observe that for $(u \leq t)$ $F(t,u) = \sum_{i=1}^{N} \bar{\Psi}_i(t) G_i(u)$; $\{F(t_j,u)\}$ are linearly independent in $\mathcal{L}_2(\rho)$ for $t_j > t$ $(j=1,2,\ldots,N)$ and that the matrix $\{\bar{\Psi}_i(t_j)\}$ invertible. This imples that $\{G_i(u)\}$ restricted to $(-\infty,t]$ are linearly independent in $\mathcal{L}_2(\rho)$ for each t.

Remarks.1. If we define for each i

(6.21) $$\underline{u}_t^{(i)} = \int_{-\infty}^{t} G_i(u)\, d\underline{z}(u)$$ then $\underline{u}_t^{(i)} - \underline{u}_s^{(i)} \perp L_2^{(q)}(\underline{u}^{(i)};s)\ (s < t)$.

(orthogonality again in the Gramian sense). Hence $\underline{u}_t^{(i)}$ is for each i is a wide-sense q-dimensional martingale and

(6.22) $$\underline{x}_t = \sum_{i=1}^{N} \bar{\Psi}_i(t)\, \underline{u}_t^{(i)}.$$

Furthermore since $L_2(\underline{x};t) \subset \mathfrak{S}\{\bigcup_{i=1}^{q} H(\underline{u}^{(i)};t)\} \subset L_2(\underline{z};t) = L_2(\underline{x};t)$ from (6.21), (6.22) and the proper canonical property, we get

(6.23) $$H(\underline{x};t) = \mathfrak{S}\{\bigcup_{i=1}^{N} H(\underline{u}^{(i)};t)\}.$$

If N = 1, this reduces to the representation of Theorem 2.1. However, the result here is obtained for purely non-deterministic processes.

2. The assumption $M \leq q$ is not very restrictive since it is satisfied for stationary processes.

7. Stationary Wide-sense N-ple Markov processes. From (6.22), (6.23) and Theorem I.5.1 the corresponding representation for stationary purely non-deterministic N-ple Markov processes satisfying (D) is given by

$$(7.1) \qquad \underline{x}_t = \sum_{i=1}^{N} \int_{-\infty}^{t} \underline{\bar{\Psi}}_i(t)\, H_i(u)\, d\underline{\xi}(u).$$

Here $\sum_{i=1}^{N} \underline{\bar{\Psi}}_i(t)\, H_i(u)$ is a function of $t-u$. In fact it is $\sum_{i=1}^{N} \underline{\bar{\Psi}}_i(t-u)\, H_i(0)$ $(u \leq t)$ where $\underline{\bar{\Psi}}_i(\cdot)$ is zero on the negative real line or

$\sum_{i=1}^{N} \underline{\bar{\Psi}}_i(0)\, H_i(u-t)$ $(u \leq t)$ where $H_i(\cdot)$ is zero on the positive real line.

The further determination of the kernel $\sum_{i=1}^{N} \underline{\bar{\Psi}}_i(t)\, H_i(u)$ leads under certain conditions to a vector generalization of continuous parameter Ornstein-Uhlenbeck processes. These purely non-deterministic processes also have rational spectral density matrices and are of importance in multidimensional prediction problems (See A. M. Yaglom [9]). It is proposed to study these questions in detail at a later time.

Chapter II.

[1]. Beutler, F. J. "Multivariate wide-sense Markov processes and prediction theory," Ann. Math. Stat. 34 1963.

[2]. Doob, J. L. "The elementary Gaussian processes," Ann. Math. Stat. 15 1944, 229-282.

[3]. Doob, J. L. Stochastic Processes, Wiley, New York, 1953.

[4]. Gladyshev, E. G. "On multidimensional stationary random processes," Theory of Probability and Its Applications III, 1958.

[5]. Hida, T. "Canonical representations of Gaussian processes and their applications," Mem. Coll. Sci. Kyoto A33, 1960.

[6]. Ince, E. L. Ordinary differential equations, Dover, 1956.

[7]. Rosenberg, M. "Square integrability of matrix-valued functions with respect to a non-negative definite hermitiom measure," Duke Math. J. 31, No. 2, 1964.

[8]. Schwarz, L. Theorie des distributions 1, 2, Hermann, Paris.

[9]. Yaglom, A. M. "Effective solutions of the linear approximation problems for multivariate stationary processes with a rational spectrum," Theory of Probability and Its Applications II, 1957.

[10]. Wiener, N. and Masani, P. "The prediction theory of multivariate stochastic processes," Acta Mathematica 98, 1957, 111-149.