ARKIV FÖR MATEMATIK Band 4 nr 19

Read 14 September 1960

On the structure of purely non-deterministic stochastic processes

By Harald Cramér

Introduction

1. The purpose of this paper is to give the proofs of some results recently communicated in a lecture at the Fourth Berkeley Symposium on Mathematical Statistics and Probability [1].

Although these results were expressed in the language of mathematical probability, they may equally well be regarded as concerned with the properties of certain curves in Hilbert space. In this first section of the Introduction we shall briefly state some results of the paper in Hilbert space language, and then in the following sections recur to the "mixed" language which seems convenient when probability questions are treated with the methods of Hilbert space geometry.

Let $\mathfrak{H}$ be a complex Hilbert space, and let, for every real τ, a set of q elements $x_1(\tau), x_2(\tau), \ldots, x_q(\tau)$ of $\mathfrak{H}$ be given. As τ runs through all real values, each element $x_j(\tau)$ describes a "curve" C_j in the space $\mathfrak{H}$. Let $C_j(t)$ denote the "arc" of C_j corresponding to values of $\tau \leqslant t$, and denote by $\mathfrak{H}(\mathbf{x}, t)$ the smallest subspace of $\mathfrak{H}$ containing the arcs $C_1(t), \ldots, C_q(t)$.

As t increases, the $\mathfrak{H}(\mathbf{x}, t)$ form a never decreasing family of subspaces, and the limiting spaces $\mathfrak{H}(\mathbf{x}, +\infty)$ and $\mathfrak{H}(\mathbf{x}, -\infty)$ will exist. It will be assumed that the following two conditions are satisfied:

(A) The strong limits $x_j(t \pm 0)$ exist for $j = 1, \ldots, q$ and for all real t.
(B) The space $\mathfrak{H}(\mathbf{x}, -\infty)$ contains only the zero element of $\mathfrak{H}$.

The projection of an arbitrary element z of $\mathfrak{H}(\mathbf{x}, +\infty)$ on the subspace $\mathfrak{H}(\mathbf{x}, t)$ will be denoted by $P_t z$.

We propose to show that the $x_j(t)$ can be simultaneously and linearly expressed in terms of certain mutually orthogonal elements. For this purpose we shall use considerations closely related to the theory of spectral multiplicity of self-adjoint transformations in a separable Hilbert space (cf. e.g. [7], Chapter VII).

It will be shown that it is possible to find a sequence $z_1, \ldots, z_N$ of elements of $\mathfrak{H}(\mathbf{x}, +\infty)$ such that we have for every $j = 1, \ldots, q$ and for all real t

$$x_j(t) = \sum_{n=1}^{N} \int_{-\infty}^{t} g_{jn}(t, \lambda)\, d z_n(\lambda), \tag{1}$$

where $z_n(\lambda) = P_\lambda z_n$. Here N may be a finite integer or equal to $\aleph_0$, the g_{jn} are complex-valued functions of the real variables t, λ, and the integrals are appropriately defined. Two increments $\Delta z_m(\lambda)$ and $\Delta z_n(\mu)$ are always orthogonal if $m \neq n$, while for $m = n$ they are orthogonal if they correspond to disjoint intervals.

We shall study the properties of the expansion (1), and in particular it will be shown that we have for any $u < t$

$$P_u x_j(t) = \sum_{n=1}^{N} \int_{-\infty}^{u} g_{jn}(t, \lambda)\, d z_n(\lambda). \tag{2}$$

When certain additional conditions are imposed, N is the smallest cardinal number such that a representation of the form (1) holds.

If the elements of $\mathfrak{H}$ are interpreted as random variables, the set of curves $C_1, \ldots, C_q$ will correspond to q simultaneously considered stochastic processes with a continuous time parameter. The above results then yield a representation of such a set of processes in terms of past and present "innovations", as well as an explicit expression for the linear least squares prediction, as will be shown in the sequel.

2. Consider a random variable x defined on a once for all given probability space, and satisfying the relations

$$E x = 0, \quad E|x|^2 < \infty. \tag{3}$$

The set of all random variables defined on the given probability space and satisfying (3) forms a Hilbert space $\mathfrak{H}$, if the inner product and the norm are defined in the usual way:

$$(x, y) = E(x \bar{y}), \quad \|x\|^2 = E|x|^2.$$

If two random variables x, y belonging to $\mathfrak{H}$ are such that

$$\|x - y\|^2 = E|x - y|^2 = 0,$$

they will be considered as identical, and we shall write

$$x = y.$$

In the sequel, any equation between random variables should be interpreted in this sense.

Whenever we are dealing with the *convergence* of a sequence of random variables, it will always be understood that we are concerned with convergence in the topology induced by the norm in $\mathfrak{H}$, which in probabilistic terminology corresponds to convergence in quadratic mean.

A family of complex-valued random variables $x(t)$, where t is a real parameter ranging from $-\infty$ to $+\infty$, and $x(t) \in \mathfrak{H}$ for every t, will be called a *one-dimensional stochastic process* with continuous time parameter t. Further, if

$x_1(t), \ldots, x_q(t)$ are q random variables, each of which is associated with a process of this type, the (column) vector

$$\mathbf{x}(t) = (x_1(t), \ldots, x_q(t)) \tag{4}$$

defines a *q-dimensional stochastic vector process* with continuous time t.

For every fixed t, each component $x_j(t)$ of the vector process (4) is a point in the Hilbert space $\mathfrak{H}$. As t increases from $-\infty$ to $+\infty$, this point describes a curve C_j in $\mathfrak{H}$, and so we are led to consider the set of curves $C_1 \ldots, C_q$ mentioned in the preceding section. The subspace $\mathfrak{H}(\mathbf{x}, t)$ is, from the present point of view, the subspace spanned by the random variables $x_1(\tau), \ldots, x_q(\tau)$ for all $\tau \leqslant t$, and we shall write this

$$\mathfrak{H}(\mathbf{x}, t) = \mathfrak{S}\{x_1(\tau), \ldots, x_q(\tau); \quad \tau \leqslant t\}.$$

$\mathfrak{H}(\mathbf{x}, t)$ may be regarded as the set of all random variables that can be obtained by means of linear operations acting on the components of $\mathbf{x}(\tau)$ for all $\tau \leqslant t$. We evidently have

$$\mathfrak{H}(\mathbf{x}, t_1) \subset \mathfrak{H}(\mathbf{x}, t_2)$$

whenever $t_1 < t_2$. It follows that the limiting spaces $\mathfrak{H}(\mathbf{x}, +\infty)$ and $\mathfrak{H}(\mathbf{x}, -\infty)$ exist, and also that the limiting spaces $\mathfrak{H}(\mathbf{x}, t \pm 0)$ exist for all t.

Following Wiener & Masani [8], we shall say that the space $\mathfrak{H}(\mathbf{x}, t)$ represents the *past and present* of the vector process (4), as seen from the point of view of the instant t. The limiting space $\mathfrak{H}(\mathbf{x}, -\infty)$ will be called the *remote past* of the process, while the space $\mathfrak{H}(\mathbf{x}, +\infty)$ will be briefly denoted by $\mathfrak{H}(\mathbf{x})$, and called the *space of the* $\mathbf{x}(t)$ *process*. We then have for any t

$$\mathfrak{H}(\mathbf{x}, -\infty) \subset \mathfrak{H}(\mathbf{x}, t) \subset \mathfrak{H}(\mathbf{x}, +\infty) = \mathfrak{H}(\mathbf{x}) \subset \mathfrak{H}.$$

In the particular case of a vector process $\mathbf{x}(t)$ satisfying

$$\mathfrak{H}(\mathbf{x}, -\infty) = \mathfrak{H}(\mathbf{x}), \tag{5}$$

it will be seen that complete information concerning the process is already contained in the remote past. Accordingly such a process will be called a *deterministic* process.

Any process not satisfying (5) will be called *non-deterministic*. In the extreme case when we have

$$\mathfrak{H}(\mathbf{x}, -\infty) = 0, \tag{6}$$

the remote past does not contain any information at all, and the process is said to be *purely non-deterministic*.

It is known [1] that any vector process (4) can be represented as the sum of a deterministic and a purely non-deterministic component, which are mutually orthogonal. In the present paper, we shall mainly be concerned with the structure of the latter component, and we may then as well assume that the given $\mathbf{x}(t)$ process itself is purely non-deterministic, i.e. that $\mathfrak{H}(\mathbf{x}, -\infty) = 0$.

Before introducing this assumption we shall, however, in sections 4–5 study the properties of the general $\mathbf{x}(t)$ process, without imposing more than the mildly restrictive condition (A), relating to the behaviour of the process in its points of discontinuity.

In section 6 we shall then, besides condition (A), introduce condition (B) which states that the $\mathbf{x}(t)$ process is purely non-deterministic. Throughout the rest of the paper, we shall then be concerned with processes satisfying both conditions (A) and (B).

In the particular case of a *stationary* one-dimensional process $x(t)$ satisfying (A) and (B), it is known that $x(t)$ can be linearly represented in terms of past innovations by an expression of the form[1]

$$x(t)=\int_{-\infty}^{t} g(t-\lambda)\,dz(\lambda),. \tag{7}$$

where $z(\lambda)$ is a process *with orthogonal increments*, which may be called an *innovation process* of $x(t)$.

In sections 8–9, we shall be concerned with representations of a similar kind, but generalized in two directions: the assumption of stationarity will be dropped, and a vector process $\mathbf{x}(t)$ will be considered instead of the one-dimensional $x(t)$. It will be shown that, for any $\mathbf{x}(t)$ process satisfying (A) and (B), we have a representation of the form (1) indicated in section 1 above. The $z_n(\lambda)$ occurring in (1) will now be one-dimensional stochastic processes with orthogonal increments. Accordingly we may say that, in the general case, we are concerned with an innovation process $(z_1(\lambda), \ldots, z_N(\lambda))$ which is multi-dimensional, and possibly even infinite-dimensional.

3. In a series of papers, P. Lévy (cf. [4–6], and further references there given) has investigated the properties of stochastic processes representable, in the notation of the present paper, in the form

$$x(t)=\int_{0}^{t} g(t,\lambda)\,dz(\lambda),$$

where $z(\lambda)$ is a *normal* (i.e Gaussian or Laplacian) process with independent increments. His investigations have been continued in a recent paper by T. Hida [3], who has also considered the more general representation

$$x(t)=\sum_{n=1}^{N}\int_{0}^{t} g_n(t,\lambda)\,dz_n(\lambda)$$

derived from considerations of spectral multiplicity, as well as the case when the lower limits of the integrals are $-\infty$ instead of zero.

[1] Cf. e.g. Doob [2], p. 588, and the references there given. The corresponding representation for a one-dimensional stationary process with *discrete* time parameter was first found by Wold [9], and was generalized to the vector case by Wiener & Masani [8].

In connection with my Berkeley lecture [1], Professor K. Itô kindly drew my attention to the work of Mr. Hida, which was then in the course of being printed. Evidently there are interesting points of contact between Mr. Hida's line of investigation and the one pursued in the present paper.

Discontinuities and innovations

4. Consider a stochastic vector process as defined in section 2,

$$\mathbf{x}(t) = (x_1(t), \ldots, x_q(t)),$$

and let us suppose that the following condition is satisfied:

(A) *The limits* $x_j(t-0)$ *and* $x_j(t+0)$ *exist for* $j = 1, \ldots, q$ *and for all real* t.

We shall then write

$$\mathbf{x}(t-0) = (x_1(t-0), \ldots, x_q(t-0))$$

and similarly for $\mathbf{x}(t+0)$. Any point t such that at least one of the relations

$$\mathbf{x}(t-0) = \mathbf{x}(t) = \mathbf{x}(t+0)$$

is not satisfied, is a *discontinuity point* of the $\mathbf{x}$ process. The point t is a *left* or *right* discontinuity, or both, according as $\mathbf{x}(t-0) \neq \mathbf{x}(t)$, or $\mathbf{x}(t) \neq \mathbf{x}(t+0)$, or both. We shall now prove the following Lemma.

Lemma 1. *For any* $\mathbf{x}(t)$ *process satisfying* (A), *we have*

(*a*) *For* $j = 1, \ldots, q$, *the functions* $E|x_j(t)|^2$ *are bounded throughout every finite* t*-interval.*

(*b*) *The set of all discontinuity points of the* $\mathbf{x}$ *process is at most enumerable.*

(*c*) *The Hilbert space* $\mathfrak{H}(\mathbf{x})$ *is separable.*

In order to prove (*a*), let us suppose that the non-negative function $E|x_j(t)|^2$ were not bounded in a certain finite interval I. Then it would be possible to find a sequence of points $\{t_n\}$ in I, and converging monotonely to a limit t^*, such that $E|x_j(t_n)|^2 \to \infty$. Clearly this is not compatible with condition (A), so that our hypothesis must be wrong, and point (*a*) is proved.

Point (*b*) will be proved if we can show that, for each j, the one-dimensional process $x_j(t)$ has at most an enumerable set of discontinuity points. A discontinuity point of $x_j(t)$ is then, of course, a point t such that at least one of the relations

$$x_j(t-0) = x_j(t) = x_j(t+0)$$

is not satisfied. Clearly it will be enough to show, e.g., that the inequality

$$x_j(t-0) \neq x_j(t) \tag{8}$$

cannot be satisfied in more than an enumerable set of points.

According to our conventions, (8) is equivalent to the relation

$$s(t) = E\,|\,x_j(t-0) - x_j(t)\,|^2 > 0.$$

We are going to show first that, given any positive number h and any finite interval I, there is at most a finite number of points t in I such that $s(t) > h$. This being shown, the desired result follows immediately by allowing first h to tend to zero, and then I to tend the whole real axis.

Suppose, in fact, that I contains an infinite set of points t with $s(t) > h$. We can then find an infinite sequence of points $\{t_n\}$ in I, converging to a limit t^*, and such that

$$s(t_n) = E\,|\,x_j(t_n-0) - x_j(t_n)\,|^2 > h \tag{9}$$

for all n. Evidently we can even find a *monotone* sequence $\{t_n\}$ having these properties. Let us suppose, e.g., that t_n converges *decreasingly* to t^*. (The increasing case can, of course, be treated in the same way.) On account of condition (A) we can then find a sequence of positive numbers $\{\varepsilon_n\}$ tending to zero, such that

$$t_n - \varepsilon_n > t_{n+1},$$

$$E\,|\,x_j(t_n-\varepsilon_n) - x_j(t_n-0)\,|^2 < \frac{c^2 h^2}{16\,K} \tag{10}$$

for all n, where c and K are constants such that $0 < c < 1$ and $E\,|\,x_j(t)\,|^2 < K$ throughout I. The existence of such a constant K follows from point (a) of the Lemma, which has already been established.

Now for any random variables u and v we have

$$\begin{aligned} E\,|\,u+v\,|^2 &= E\,|\,u\,|^2 + E\,|\,v\,|^2 + E(u\,\bar v) + E(\bar u\, v) \\ &\geq E\,|\,u\,|^2 + E\,|\,v\,|^2 - 2\sqrt{E\,|\,u\,|^2\,E\,|\,v\,|^2}. \end{aligned}$$

Taking here

$$\begin{aligned} u &= x_j(t_n-0) - x_j(t_n), \\ v &= x_j(t_n-\varepsilon_n) - x_j(t_n-0), \end{aligned}$$

we obtain from (9) and (10)

$$E\,|\,x_j(t_n-\varepsilon_n) - x_j(t_n)\,|^2 \geq h - 2\sqrt{4\,K\cdot\frac{c^2 h^2}{16\,K}} = (1-c)\,h \tag{11}$$

for all n. On the other hand, the sequences $\{t_n - \varepsilon_n\}$ and $\{t_n\}$ both converge decreasingly to t^*. Consequently by condition (A) we have (convergence, as usual, in the $\mathfrak{H}$ topology, i.e. in quadratic mean)

$$\begin{aligned} x_j(t_n-\varepsilon_n) &\to x_j(t^*+0), \\ x_j(t_n) &\to x_j(t^*+0), \end{aligned}$$

and thus
$$x_j(t_n - \varepsilon_n) - x_j(t_n) \to 0,$$

as n tends to infinity. However, this is incompatible with (11), so that our hypothesis must be wrong, and thus point (*b*) of the Lemma is proved.

The last part of the Lemma now follows immediately, if we consider an enumerable set $\{t_n\}$ including all discontinuity points of $\mathbf{x}(t)$ as well as an everywhere dense set of continuity points. The set of all finite linear combinations of the random variables $x_j(t_n)$ for $j = 1, \ldots, q$ and $n = 1, 2, \ldots$, with coefficients whose real and imaginary parts are both rational, is then an enumerable set dense in $\mathfrak{H}(\mathbf{x})$.

5. Consider now the family of subspaces $\mathfrak{H}(\mathbf{x}, t)$ associated with an $\mathbf{x}(t)$ process satisfying condition (A) of the preceding section.

As observed in section 2, $\mathfrak{H}(\mathbf{x}. t)$ never decreases as t increases. If $\mathfrak{H}(\mathbf{x}, t)$ effectively increases when t ranges over some interval $t_1 < t \leqslant t_2$, so that we have

$$\mathfrak{H}(\mathbf{x}, t_1) \neq \mathfrak{H}(\mathbf{x}, t_2),$$

this means that some new information has entered into the process during that interval. Accordingly we shall then say that the process has received an *innovation* during the interval $t_1 < t \leqslant t_2$, and we shall regard this innovation as being represented by the orthogonal complement

$$\mathfrak{H}(\mathbf{x}, t_2) - \mathfrak{H}(\mathbf{x}, t_1). \tag{12}$$

In fact, if this complement reduces to the zero element, the two spaces are identical, and no innovation has entered during the interval, while in the opposite case (12) is the set of all differences between an element of $\mathfrak{H}(\mathbf{x}, t_2)$ and its projection on $\mathfrak{H}(\mathbf{x}, t_1)$.

If, for a certain value of t, we have

$$\mathfrak{H}(\mathbf{x}, t-h) \neq \mathfrak{H}(\mathbf{x}, t+h)$$

for any $h > 0$, this means that there is a non-vanishing innovation associated with any interval containing t as an interior point. The set of all points t having this property will be called the *innovation spectrum* of the $\mathbf{x}(t)$ process.

$\mathfrak{H}(\mathbf{x}, t)$ being never decreasing as t increases, it follows that the limiting spaces $\mathfrak{H}(\mathbf{x}, t \pm 0)$ will exist for every t. Any point such that at least one of the relations

$$\mathfrak{H}(\mathbf{x}, t-0) = \mathfrak{H}(\mathbf{x}, t) = \mathfrak{H}(\mathbf{x}, t+0)$$

is not satisfied, will certainly belong to the innovation spectrum, and will be called a *discontinuity point* of that spectrum. As in section 4, the terms *left* and *right* discontinuities will be used in the obvious sense.

The set of all discontinuity points constitutes the *discontinuous part* of the innovation spectrum. Let $t_1, t_2, \ldots$, be the points of this set (it will be shown below that the set is at most enumerable), and form the vector sum

$$\mathfrak{G}(\mathbf{x}, t) = \sum_{t_k \leqslant t} \mathfrak{H}(\mathbf{x}, t_k + 0) - \mathfrak{H}(\mathbf{x}, t_k - 0)$$

and the orthogonal complement $\mathfrak{F}(\mathbf{x}, t) = \mathfrak{H}(\mathbf{x}, t) - \mathfrak{G}(\mathbf{x}, t)$. The set of all points t such that $\mathfrak{F}(\mathbf{x}, t-h) \neq \mathfrak{F}(\mathbf{x}, t+h)$ for any $h > 0$ constitutes the *continuous part* of the innovation spectrum of $\mathbf{x}(t)$. If the discontinuous part is not a closed set, its limiting points will belong to the innovation spectrum, without necessarily belonging to any of the two parts here defined.

When t is a discontinuity point of the innovation spectrum, the number of dimensions of the subspace $\mathfrak{H}(\mathbf{x}, t) - \mathfrak{H}(\mathbf{x}, t-0)$ will be called the *left multiplicity* of the point t. Similarly the *right multiplicity* of t is the number of dimensions of $\mathfrak{H}(\mathbf{x}, t+0) - \mathfrak{H}(\mathbf{x}, t)$. If t is not a left (right) discontinuity, the left (right) multiplicity of t is of course equal to zero. We shall now prove the following Lemma.

Lemma 2. *For any* $\mathbf{x}(t)$ *process satisfying* (A), *we have*

(a) *The set of discontinuity points of the innovation spectrum is at most enumerable.*

(b) *In a left discontinuity point, the left multiplicity is at most equal to* q.

(c) *In a right discontinuity point, the right multiplicity may be any finite integer, or equal to* $\aleph_0$.

(d) *A left discontinuity of the innovation spectrum is always at the same time a left discontinuity of the process. On the other hand, a right discontinuity of the innovation spectrum is not necessarily a discontinuity of the process.*

We first observe that the two subspaces $\mathfrak{H}(\mathbf{x}, t_2) - \mathfrak{H}(\mathbf{x}, t_1)$ and $\mathfrak{H}(\mathbf{x}, u_2) - \mathfrak{H}(\mathbf{x}, u_1)$ corresponding to disjoint time intervals are always orthogonal. It follows that the subspaces

$$\mathfrak{H}(\mathbf{x}, t+0) - \mathfrak{H}(\mathbf{x}, t-0)$$

corresponding to different discontinuity points are orthogonal. By Lemma 1, the space $\mathfrak{H}(\mathbf{x})$ is separable, and cannot include more than an enumerable set of mutually orthogonal subspaces. Hence follows the truth of point (a) of the Lemma.

Point (b) of the Lemma asserts that the orthogonal complement $\mathfrak{H}(\mathbf{x}, t) - \mathfrak{H}(\mathbf{x}, t-0)$ has at most q dimensions. By definition, the space $\mathfrak{H}(\mathbf{x}, t)$ is spanned by the variables $x_j(\tau)$ with $j = 1, \ldots, q$ and $\tau \leqslant t$. Writing

$$x_j(t) = y_j + P_{t-0}\, x_j(t), \quad (j = 1, \ldots, q), \tag{13}$$

where P_{t-0} denotes the projection on $\mathfrak{H}(\mathbf{x}, t-0)$, it will be seen that every element of $\mathfrak{H}(\mathbf{x}, t)$ is the sum of an element of $\mathfrak{H}(\mathbf{x}, t-0)$ and a linear combination of the variables $y_1, \ldots, y_q$, which belong to $\mathfrak{H}(\mathbf{x}, t)$ and are orthogonal to $\mathfrak{H}(\mathbf{x}, t-0)$. Consequently the orthogonal complement $\mathfrak{H}(\mathbf{x}, t) - \mathfrak{H}(\mathbf{x}, t-0)$ is identical with the space spanned by $y_1, \ldots, y_q$, and thus has at most q dimensions. It will also be seen that, choosing the variables $y_1, \ldots, y_q$ in an appropriate way, we can construct examples of processes having, in a given point, a left discontinuity of the innovation spectrum with any given multiplicity not exceeding q. Thus (b) is proved.

We shall now prove (c) by constructing an example of an $\mathbf{x}(t)$ process, the innovation spectrum of which has, in a given point, a right discontinuity of infinite multiplicity. It will then be easily seen how the example may be modified in order to produce a discontinuity of any given finite multiplicity.

Let $z_1, z_2, \ldots$ be an infinite orthonormal sequence of random variables belonging to $\mathfrak{H}$. Denoting by $p_1, p_2, \ldots$ the successive prime numbers ($p_1 = 2$, $p_2 = 3, \ldots$), we define a process $\mathbf{x}(t) = (x_1(t), \ldots, x_q(t))$ by taking

$$x_2(t) = \cdots = x_q(t) = 0 \quad \text{for all } t,$$

$$x_1(t) = 0 \quad \text{for} \quad t \leqslant 0,$$

$$x_1(t) = t\, z_n \quad \text{for} \quad t = p_n^{-k}, \quad k = 1, 2, \ldots$$

$$x_1(t) = t\, z_1 \quad \text{for} \quad t > \tfrac{1}{2}.$$

For values of t in the interval $0 < t < \frac{1}{2}$, which are not of the form p_n^{-k}, we define $x_1(t)$ by linear interpolation:

$$x_1(t) = \frac{(t_2 - t)\, x_1(t_1) + (t - t_1)\, x_1(t_2)}{t_2 - t_1},$$

where $t_1 = p_{n_1}^{-k_1}$ and $t_2 = p_{n_2}^{-k_2}$ are the nearest values below and above t, for which $x_1(t)$ has been defined above. Some easy calculation shows that we have for all $t > 0$

$$E\,|x_1(t)|^2 \leqslant t^2.$$

This shows that the point $t = 0$ is a continuity point for the vector process $\mathbf{x}(t)$. Since every other real t is evidently also a continuity point, the $\mathbf{x}(t)$ process is everywhere continuous, and *a fortiori* satisfies condition (A).

On the other hand, we obviously have $\mathfrak{H}(\mathbf{x}, t) = 0$ for $t \leqslant 0$, while for every $t > 0$ the space $\mathfrak{H}(\mathbf{x}, t)$ will include the infinite orthonormal sequence $z_1, z_2, \ldots$. The point $t = 0$ will thus be a right discontinuity of the innovation spectrum of $\mathbf{x}(t)$, with an infinite right multiplicity. A case with any given finite multiplicity n is obtained if the variables $z_{n+1}, z_{n+2}, \ldots$ are replaced by zero. We have thus proved point (c) of the Lemma.

We observe that, by some further elaboration of the example given above, we may construct a process having the same multiplicity properties in each point of an everywhere dense set of values of t. We can even arrange the example so that the mean square derivative of $\mathbf{x}(t)$ exists for every t.

The last part of the Lemma follows simply from the above proofs of (b) and (c). If t is a left discontinuity of the innovation spectrum, at least one of the variables y_j occurring in (13) must be different from zero. Suppose, e.g., that $y_1 \neq 0$. Then

$$\| x_1(t) - x_1(t-0) \| \geqslant \| x_1(t) - P_{t-0}\, x_1(t) \| = \| y_1 \| > 0,$$

and so t is a left discontinuity of $x_1(t)$, and consequently also of $\mathbf{x}(t)$. On the other hand, the process $\mathbf{x}(t)$ constructed in the proof of (c) provides an example of a process having in $t = 0$ a right discontinuity point of the innovation spectrum (even of infinite multiplicity), which is nevertheless a continuity point of the process. This completes the proof of Lemma 2.

Innovation processes

6. From now on, it will be assumed that we are dealing with a vector process $\mathbf{x}(t)$ satisfying not only condition (A) of section 4, but also the following condition:

(B) $\mathbf{x}(t)$ *is a purely non-deterministic process, i.e. we have* $\mathfrak{H}(\mathbf{x}, -\infty) = 0$.

In the Hilbert space $\mathfrak{H}(\mathbf{x}) = \mathfrak{H}(\mathbf{x}, +\infty)$ of the process, we shall in general denote by $P_{\mathfrak{M}}$ the projection operator whose range is the subspace $\mathfrak{M}$. However, when $\mathfrak{M}$ is the particular subspace $\mathfrak{H}(\mathbf{x}, t)$, we write simply P_t instead of $P_{\mathfrak{H}(\mathbf{x}, t)}$.

As t increases from $-\infty$ to $+\infty$, the P_t form a never decreasing family of projections, with

$$P_{-\infty} = 0, \quad P_{+\infty} = I.$$

P_{t+0} is the projection on $\mathfrak{H}(\mathbf{x}, t+0)$, and similarly for P_{t-0}. The difference $P_{t_2} - P_{t_1}$, where $t_1 < t_2$, denotes the projection on the orthogonal complement $\mathfrak{H}(\mathbf{x}, t_2) - \mathfrak{H}(\mathbf{x}, t_1)$. It follows, in particular, that the projections $P_{t_2} - P_{t_1}$ and $P_{u_2} - P_{u_1}$ will be mutually orthogonal, as soon as the corresponding time intervals are disjoint.

Further, the points t of the innovation spectrum are characterized by the property

$$P_{t+h} - P_{t-h} > 0$$

for any $h > 0$, while the discontinuity points of that spectrum are characterized by the relation

$$P_{t+0} - P_{t-0} > 0.$$

Consider now any element z of the Hilbert space $\mathfrak{H}(\mathbf{x})$, and let us define a stochastic process $z(\lambda)$ by writing for any real λ

$$z(\lambda) = P_\lambda z. \tag{14}$$

It then follows from the above that $z(\lambda)$ is a process *with orthogonal increments*, such that

$$z(-\infty) = 0, \quad z(+\infty) = z.$$

We have $E\, z(\lambda) = 0$, and if we write

$$E\,|z(\lambda)|^2 = F_z(\lambda),$$

$F_z(\lambda)$ will be a never decreasing function of the real variable λ, such that

$$F_z(-\infty) = 0, \quad F_z(+\infty) = E|z|^2.$$

The points of increase of $z(\lambda)$, i.e. the points λ such that the increment $z(\lambda+h) - z(\lambda-h)$ does not reduce to zero for any $h > 0$, are identical with the points of increase of $F_z(\lambda)$, and form a subset of the innovation spectrum of the $\mathbf{x}(t)$ process. Similarly the left (right) discontinuities of $z(\lambda)$ are identical with the left (right) discontinuities of $F_z(\lambda)$, and form a subset of the set of all left (right) discontinuities of the innovation spectrum of $\mathbf{x}(t)$. Any increment $dz(\lambda)$ belongs to the subspace $d_\lambda \mathfrak{H}(\mathbf{x}, \lambda)$, and is thus built up by a certain part of the elements which enter as innovations into the $\mathbf{x}(t)$ process between the time points $t=\lambda$ and $t=\lambda+d\lambda$.

On account of these facts, we shall denote the (one-dimensional) $z(\lambda)$ process as a *partial innovation process* associated with the given vector process $\mathbf{x}(t)$.

7. For any $t \leqslant +\infty$, we shall denote by $\mathfrak{H}(z, t)$ the Hilbert space spanned by the random variables $z(\lambda)$ for all $\lambda \leqslant t$:

$$\mathfrak{H}(z, t) = \mathfrak{S}\{z(\lambda); \quad \lambda \leqslant t\}.$$

It follows from (14) that $z(\lambda)$ is always an element of $\mathfrak{H}(\mathbf{x}, \lambda)$, and consequently $\mathfrak{H}(z, t)$, is a subspace of $\mathfrak{H}(\mathbf{x}, t)$:

$$\mathfrak{H}(z, t) \subset \mathfrak{H}(\mathbf{x}, t).$$

Instead of $\mathfrak{H}(z, +\infty)$ we shall write briefly $\mathfrak{H}(z)$. Evidently $\mathfrak{H}(z, t)$ is the projection of $\mathfrak{H}(z)$ on $\mathfrak{H}(\mathbf{x}, t)$.

If no λ is at the same time a left and a right discontinuity of $z(\lambda)$, then $\mathfrak{H}(z, t)$ is, for every $t \leqslant +\infty$, identical[1] with the set $\mathfrak{H}^*(z, t)$ of all random variables y representable in the form

$$y = \int_{-\infty}^{t} g(\lambda)\, dz(\lambda) \tag{15}$$

with an F_z-measurable g such that

$$E|y|^2 = \int_{-\infty}^{t} |g(\lambda)|^2\, dF_z(\lambda) < \infty.$$

On the other hand, if in a certain point $\lambda < t$ the left and the right jumps of $z(\lambda)$, say u and v respectively, are both different from zero, all random variables $Au + Bv$ with constant A and B will belong to $\mathfrak{H}(z, t)$ while, with the usual definition of the integral (15), the discontinuity at λ will only provide the variables $A(u+v)$ as elements of $\mathfrak{H}^*(z, t)$.

We shall require the following Lemma, which is only a restatement of familiar facts concerning Hilbert space.

Lemma 3. *If y and z are elements of $\mathfrak{H}(\mathbf{x})$ such that $y \perp \mathfrak{H}(z)$, then $\mathfrak{H}(y) \perp \mathfrak{H}(z)$.*

Since $\mathfrak{H}(z)$ is the space spanned by all $z(\lambda) = P_\lambda z$, the relation $y \perp \mathfrak{H}(z)$ is equivalent to $y \perp P_\lambda z$ for all real λ. By the same argument, the assertion of Lemma 3 is equivalent to the relation

$$P_\lambda y \perp P_\mu z$$

[1] Cf., e.g., [2], p. 425–428. The integral in (15) should be so defined that, if the upper limit t is a discontinuity of $z(\lambda)$, a left jump of $z(\lambda)$ is included in the value of the integral, but not a right jump.

for all λ and μ. Now $y = P_\lambda y + w$, where $w \perp \mathfrak{H}(\mathrm{x}, \lambda)$, and thus in particular $w \perp P_\lambda z$. Hence for any λ

$$P_\lambda y = y - w \perp P_\lambda z.$$

Suppose now first $\lambda > \mu$. Then $P_\mu P_\lambda = P_\mu$, and thus

$$P_\lambda y = P_\mu y + w',$$

where $w' \perp \mathfrak{H}(\mathrm{x}, \mu)$. Hence in particular $w' \perp P_\mu z$. Since we have just proved that $P_\mu y \perp P_\mu z$, it follows that $P_\lambda y \perp P_\mu z$.

On the other hand, if $\lambda < \mu$ we have $P_\mu z = P_\lambda z + w''$, where $w'' \perp \mathfrak{H}(\mathrm{x}, \lambda)$, and thus $w'' \perp P_\lambda y$. Hence we obtain as before $P_\mu z \perp P_\lambda y$, and so Lemma 3 is proved.

Representation of $\mathrm{x}(t)$

8. We begin by proving the following Lemma, assuming as before that we are dealing with a given vector process $\mathrm{x}(t)$ satisfying conditions (A) and (B).

Lemma 4. *It is possible to find a finite or infinite sequence $z_1, z_2, \ldots$ of non-vanishing elements of $\mathfrak{H}(\mathrm{x})$ such that we have for every $t \leqslant +\infty$*

$$\mathfrak{H}(z_j, t) \perp \mathfrak{H}(z_k, t) \quad \text{for} \quad j \neq k, \tag{16}$$

$$\mathfrak{H}(\mathrm{x}, t) = \mathfrak{H}(z_1, t) + \mathfrak{H}(z_2, t) + \ldots, \tag{17}$$

where the second member of (17) *denotes the vector sum of the mutually orthogonal spaces involved.*

We first observe that it is sufficient to show that we can find a z_n sequence such that (16) and (17) hold for $t = +\infty$, since their validity for any finite t then easily follows.

By Lemma 1, the space, $\mathfrak{H}(\mathrm{x})$ is separable. Neglecting trivial cases, we may assume that $\mathfrak{H}(\mathrm{x})$ is infinite-dimensional. Thus a complete orthonormal system in $\mathfrak{H}(\mathrm{x})$ will form an infinite sequence, say $z_1^*, z_2^*, \ldots$. Starting from the sequence of the z_n^*, we shall now construct an infinite sequence $z_1, z_2, \ldots$ satisfying (16) and (17) for $t = +\infty$. Discarding any z_n which reduces to zero, we then obtain a finite or enumerable sequence of non-vanishing elements having the same properties, and so the Lemma will be proved.

We define the z_n sequence by the relations

$$\begin{aligned} z_1 &= z_1^*, \\ z_2 &= z_2^* - P_{\mathfrak{M}_1} z_2^*, \\ &\cdots\cdots \\ z_n &= z_n^* - P_{\mathfrak{M}_{n-1}} z_n^*, \\ &\cdots\cdots \end{aligned} \tag{18}$$

where $\mathfrak{M}_n$ denotes the vector sum

$$\mathfrak{M}_n = \mathfrak{H}(z_1) + \cdots + \mathfrak{H}(z_n).$$

Then for any $n > 1$ we have $z_n \perp \mathfrak{M}_{n-1}$, and consequently $z_n \perp \mathfrak{H}(z_j)$ for $j = 1, \ldots, n-1$. By Lemma 3 it then follows that we have $\mathfrak{H}(z_j) \perp \mathfrak{H}(z_k)$ for $j \neq k$, so that (16) is satisfied.

We further have

$$z_n^* = z_n + P_{\mathfrak{M}_{n-1}} z_n^*,$$

and since $z_n \in \mathfrak{H}(z_n)$, this shows that $z_n^* \in \mathfrak{H}(z_n) + \mathfrak{M}_{n-1} = \mathfrak{M}_n$. Now $\mathfrak{M}_n$ is a subset of the infinite vector sum $\mathfrak{H}(z_1) + \mathfrak{H}(z_2) + \ldots$. This sum, which is a subspace of $\mathfrak{H}(\mathbf{x})$, will thus contain the complete orthonormal system $z_1^*, z_2^*, \ldots$, and will consequently be identical with $\mathfrak{H}(\mathbf{x})$, so that (17) is also satisfied, and the Lemma is proved.

9. The sequence $z_1, z_2, \ldots$ considered in Lemma 4 is not uniquely determined, and we now proceed to show that it can be chosen in a way which will suit our purpose.

By Lemma 2, the discontinuities of the innovation spectrum of $\mathbf{x}(t)$ form an at most enumerable set. Let them be denoted by $\lambda_1, \lambda_2, \ldots$, and consider the subspaces

$$\mathfrak{U}_k = \mathfrak{H}(\mathbf{x}, \lambda_k) - \mathfrak{H}(\mathbf{x}, \lambda_k - 0),$$

$$\mathfrak{V}_k = \mathfrak{H}(\mathbf{x}, \lambda_k + 0) - \mathfrak{H}(\mathbf{x}, \lambda_k),$$

for $k = 1, 2, \ldots$. If h_k and j_k are the numbers of dimensions of $\mathfrak{U}_k$ and $\mathfrak{V}_k$ respectively, then by Lemma 2 we have $0 \leqslant h_k \leqslant q$, while j_k may be any finite non-negative integer or $\aleph_0$. According to the terminology of section 5, h_k is the left multiplicity of the point λ_k, while j_k is the right multiplicity. The number

$$N' = \sup_k (h_k + j_k)$$

will be called the *multiplicity of the discontinuous part* of the innovation spectrum. Let

$$u_{k1}, \ldots, u_{kh_k},$$

$$v_{k1}, \ldots, v_{kj_k}$$

be complete orthonormal systems in $\mathfrak{U}_k$ and $\mathfrak{V}_k$ respectively. If $\mathfrak{U}_k$ or $\mathfrak{V}_k$ reduces to zero, the corresponding system does of course not occur; however, when λ_k is a discontinuity, h_k and j_k cannot both be equal to zero, so that at least one of the u and v systems must contain a non-vanishing number of terms.

If z denotes any of the u_{kh} or v_{kj}, it will be seen that we have $P_\lambda z = z$ for $\lambda > \lambda_k$, and $P_\lambda z = 0$ for $\lambda < \lambda_k$. It follows that the space $\mathfrak{H}(z)$ will then be one-dimensional, and consist of all constant multiples of z. If y is any variable in $\mathfrak{H}(\mathbf{x})$ such that $y \perp z$, it then follows from Lemma 3 that $\mathfrak{H}(y) \perp \mathfrak{H}(z)$.

Suppose now that one of the orthonormal variables from which we started our proof of Lemma 4, say z_n^*, is identical with one of the u_{kh} or v_{kj}. By means of the remark just made, it then follows from the relations (18) that z_n^*

is orthogonal to $z_1, \ldots, z_{n-1}$, and thus also orthogonal to $\mathfrak{M}_{n-1}$. Consequently by (18) we obtain $z_n = z_n^*$.

Choosing the orthonormal system $z_1^*, z_2^*, \ldots$, so that all the variables u_{kh} and v_{kj} $(k=1,2,\ldots, h=1,\ldots,h_k, j=1,\ldots,j_k)$ occur in it, we thus see that the same variables will also occur in the sequence $z_1, z_2, \ldots$ constructed according to (18), and satisfying the conditions of Lemma 4. Besides the u_{kh} and v_{kj}, there may be other elements in the z_n sequence. Let $w_1, w_2, \ldots$ be those elements, if any, which are different from all the u_{kh} and v_{kj}. We are going to show that, if

$$w(\lambda) = P_\lambda w$$

s the partial innovation process corresponding to any of the w_n, then $w(\lambda)$ *has no discontinuities.*

In fact, by a remark made in section 6, any discontinuity of $w(\lambda)$ would be a discontinuity of the innovation spectrum of $\mathbf{x}(t)$, i.e. equal to one of the λ_k. The corresponding jump of $w(\lambda)$, say w^*, would then be an element of the space $\mathfrak{H}(\mathbf{x}, \lambda_k+0) - \mathfrak{H}(\mathbf{x}, \lambda_k-0)$. At the same time, w^* would belong to the space $\mathfrak{H}(w)$, and by Lemma 4 would thus be orthogonal to all the u_{kh} and v_{kj}. However, the latter variables form a complete orthonormal system in $\mathfrak{H}(\mathbf{x}, \lambda_k+0) - \mathfrak{H}(\mathbf{x}, \lambda_k-0)$ so that we must have $w^*=0$, and it follows that $w(\lambda)$ has no discontinuities.

Let now the sequence of non-vanishing elements $z_1, z_2, \ldots$ considered in Lemma 4 be chosen in all possible ways that are consistent with the requirement that all the u_{kh} and v_{kj} should occur in it. Let, in each case, M denote the cardinal number of the corresponding sequence $w_1, w_2, \ldots$, formed by those elements which are different from all the u_{kh} and v_{kj}. The numbers M will then have a non-negative lower bound:

$$N'' = \inf M,$$

which we shall call the *multiplicity of the continuous part* of the innovation spectrum. Finally

$$N = \max(N', N'') \tag{19}$$

will be called the *spectral multiplicity* of the $\mathbf{x}(t)$ process. As soon as $\mathbf{x}(t)$ is not identically zero, N will be a finite positive integer, or equal to $\aleph_0$.

It follows from the definition of the multiplicity N'' of the continuous part that it is possible to find a sequence $z_1, z_2, \ldots$ satisfying the above requirements, and such that the corresponding set $w_1, w_2, \ldots$ will have precisely the cardinal number N''. We shall then say that these w_n form a *minimal w sequence.*[1]

In the sequel, $w_1, w_2, \ldots$ will denote the elements of a fixed minimal w sequence. We now propose to construct, by means of this given w sequence, a particular sequence $z_1, z_2, \ldots$ satisfying the conditions of Lemma 4, which will then be used for the proof of our representation theorem for $\mathbf{x}(t)$.

[1] By an adaptation of the proofs of theorem 7.5 and 7.6 of Stone [7] to the case considered here, it can be shown that a minimal w sequence can be chosen in such a way that the set of all points of increase of $w_1(\lambda) = P_\lambda w_1$ is identical with the continuous part of the innovation spectrum of $\mathbf{x}(t)$, and includes the corresponding set of any $w_n(\lambda)$ with $n>1$ as a subset. As this property is not indispensable for the proof of the representation theorem given below, we shall restrict ourselves here to this remark.

We first observe that, owing to the way in which the w_j have been chosen, we have by Lemma 4 for every $t \leqslant +\infty$

$$\mathfrak{H}(\mathbf{x}, t) = \sum_{k,h} \mathfrak{H}(u_{kh}, t) + \sum_{k,j} \mathfrak{H}(v_{kj}, t) + \sum_{j} \mathfrak{H}(w_j, t), \tag{20}$$

where the sums denote vector addition, and all the $\mathfrak{H}$ spaces appearing in the second member are mutually orthogonal.

For any discontinuity point λ_k we now arrange the corresponding variables u_{kh} and v_{kj} into a single sequence

$$s_{k1}, \; s_{k2}, \; \ldots,$$

where the number of terms will be $h_k + j_k$. The new sequence $z_1, z_2, \ldots$ which we have in view is then defined by taking

$$z_n = w_n + \sum_k s_{kn}, \tag{21}$$

where the w_n are the elements of our fixed minimal w sequence. The summation is extended over all discontinuity points λ_k, and we take $s_{kn} = 0$ whenever $n > h_k + j_k$, and $w_n = 0$ whenever $n > N''$. The number of non-vanishing terms in the z_n sequence defined in this way will evidently be equal to the spectral multiplicity N of the $\mathbf{x}(t)$ process as defined by (19).

According to a remark made above, any space $\mathfrak{H}(s_{kn})$ is one-dimensional, and consists of all constant multiples of the variable s_{kn}. Hence it easily follows that we have for $t \leqslant +\infty$

$$\mathfrak{H}(z_n, t) = \mathfrak{H}(w_n, t) + \sum_k \mathfrak{H}(s_{kn}, t),$$

and further according to (20)

$$\begin{gathered} \mathfrak{H}(\mathbf{x}, t) = \mathfrak{H}(z_1, t) + \mathfrak{H}(z_2, t) + \ldots, \\ \mathfrak{H}(z_j\, t) \perp \mathfrak{H}(z_k, t) \quad \text{for} \quad j \neq k, \end{gathered} \tag{22}$$

so that the z_n defined by (21) satisfy the conditions of Lemma 4.

We observe that it follows from (21) that no $z_n(\lambda)$ can have a left and a right discontinuity in the same point λ. In fact, any discontinuity of $z_n(\lambda)$ will be either a left discontinuity with a jump u_{kh}, or a right discontinuity with a jump v_{kj}. By a remark made in section 6, the space $\mathfrak{H}(z_n, t)$ will then be identical with the set of all random variables y representable in the form (15).

In our given q-dimensional (column) vector process

$$\mathbf{x}(t) = (x_1(t), \ldots, x_q(t))$$

every component $x_j(t)$ is a random variable belonging to $\mathfrak{H}(\mathbf{x}, t)$. It then follows from (15) and (22) that we have for all real t

$$x_j(t) = \sum_{n=1}^{N} \int_{-\infty}^{t} g_{jn}(t, \lambda)\, d z_n(\lambda). \tag{23}$$

If the multiplicity N is infinite, the series in the second member will converge in the usual sense, so that we have

$$\sum_{n=1}^{N} \int_{-\infty}^{t} |g_{jn}(t, \lambda)|^2\, d F_{z_n}(\lambda) < \infty. \tag{24}$$

Introducing the $q \times N$ order matrix function

$$\mathfrak{G}(t, \lambda) = \{g_{jn}(t, \lambda)\}, \tag{25}$$

$(j = 1, \ldots, q;\, n = 1, \ldots, N)$, and the N-dimensional (column) vector process

$$\mathbf{z}(\lambda) = (z_1(\lambda), \ldots, z_N(\lambda)), \tag{26}$$

it is seen that (23) may be written

$$\mathbf{x}(t) = \int_{-\infty}^{t} \mathfrak{G}(t, \lambda)\, d\, \mathbf{z}(\lambda). \tag{27}$$

The vector process $\mathbf{z}(\lambda)$ defined by (26) has *orthogonal increments*, in the sense that two increments $\Delta z_j(\lambda)$ and $\Delta z_k(\mu)$ are always orthogonal if $j \neq k$, while for $j = k$ they are orthogonal if the corresponding time intervals are disjoint.

If we denote by $\mathfrak{H}(\mathbf{z}, t)$ the Hilbert space spanned by the variables $z_1(\lambda), \ldots, z_N(\lambda)$ for all $\lambda \leqslant t$:

$$\mathfrak{H}(\mathbf{z}, t) = \mathfrak{S}\{z_1(\lambda), \ldots, z_N(\lambda);\ \lambda \leqslant t\},$$

it follows from (22) that we have for every $t \leqslant +\infty$

$$\mathfrak{H}(\mathbf{x}, t) = \mathfrak{H}(\mathbf{z}, t). \tag{28}$$

According to the representation formula (27) and the property expressed by (28), it seems appropriate to call $\mathbf{z}(\lambda)$ a *total innovation process* associated with the given $\mathbf{x}(t)$. While $\mathbf{z}(\lambda)$ is not uniquely determined, its dimensionality N is uniquely determined by (19) as the spectral multiplicity of $\mathbf{x}(t)$. It is also seen that N is the smallest cardinal number for which there exists a representation of the form (27), with the properties specified by (22)–(28).

Summing up our results, we now have the following representation theorem.

Theorem 1. *Any stochastic vector process* $\mathbf{x}(t)$ *satisfying conditions* (A) *and* (B) *can be represented in the form* (27), *where* $\mathfrak{G}(t, \lambda)$ *and* $\mathbf{z}(\lambda)$ *are defined by* (25) *and* (26). *N is the spectral multiplicity of the* $\mathbf{x}(t)$ *process. If N is infinite, the expansions* (23) *formally obtained for the components* $x_j(t)$ *are convergent in quadratic mean, as shown by* (24). $z_1, \ldots, z_N$ *are random variables in* $\mathfrak{H}(\mathbf{x})$ *satisfying* (22), *and such that no* $z_j(\lambda) = P_\lambda z_j$ *has a left and a right discontinuity in the same point* λ. *The vector process* $\mathbf{z}(\lambda)$ *has orthogonal increments and satisfies* (28).

No representation with these properties holds for any smaller value of N.

If $\mathrm{x}(t) = (x_1(t), \ldots, x_q(t))$ is a vector process satisfying (A) and (B), each component $x_j(t)$, regarded as a one-dimensional process, has a certain spectral multiplicity N_j, and thus by Theorem 1 may be represented in the following form

$$x_j(t) = \sum_{n=1}^{N_j} \int_{-\infty}^{t} g_{jn}(t, \lambda)\, d z_{jn}(\lambda).$$

It can then be shown (although the proof is slightly more involved than may possibly be expected) that the spectral multiplicity N of $\mathrm{x}(t)$ satisfies the inequality

$$N \leqslant \sum_{j=1}^{q} N_j. \tag{29}$$

Consider, in particular, the case of a *stationary* vector process $\mathrm{x}(t)$, i.e. a process such that every second order covariance moment of the components is a function of the corresponding time difference:

$$E(x_j(t)\,\overline{x_k(u)}) = R_{jk}(t-u).$$

This process will satisfy (A) and (B) if and only if (*a*) the functions $R_{11}(t), \ldots, R_{qq}(t)$ are continuous at $t=0$, and (*b*) each component $x_j(t)$ is a purely non-deterministic stationary process. When these conditions are satisfied, each $x_j(t)$ has a representation of the form (7), and accordingly the spectral multiplicity of $x_j(t)$ is equal to one. It then follows from (29) that we have in this case $N \leqslant q$.

On the other hand, the process $\mathrm{x}(t)$ constructed in connection with the proof of Lemma 2, point (*c*), evidently provides an example of a vector process satisfying (A) and (B), and having an infinite spectral multiplicity N. As already observed, this example may be easily modified so as to yield a process with any given finite multiplicity.

If, in the relation (27), all components of the vectors on both sides are projected on the space $H(\mathrm{x}, u)$, where $u < t$, we finally obtain the following theorem, denoting by $P_u \mathrm{x}(t)$ the vector with the components $P_u x_j(t)$, for $j = 1, \ldots, q$.

Theorem 2. *The best linear (least squares) prediction of* $\mathrm{x}(t)$ *in terms of all variables* $x_j(\tau)$ *with* $j = 1, \ldots, q$ *and* $\tau \leqslant u$ *is given by the expression*

$$P_u \mathrm{x}(t) = \int_{-\infty}^{u} \mathrm{G}(t, \lambda)\, d\mathrm{z}(\lambda).$$

The square of the corresponding error of prediction for any component $x_j(t)$ *is*

$$E\,|x_j(t) - P_u x_j(t)|^2 = \sum_{n=1}^{N} \int_{u}^{t} |g_{jn}(t, \lambda)|^2\, dF_{z_n}(\lambda).$$

REFERENCES

1. CRAMÉR, H., On some classes of non-stationary stochastic processes. To appear in: Proc. Fourth Berkeley Symp. on Mathematical Statistics and Probability.
2. DOOB, J. L., Stochastic processes. Wiley, New York, 1953.
3. HIDA, T., Canonical representations of Gaussian processes and their applications. Mem. Coll. Sci. Kyoto, A33 (1960).
4. LÉVY, P., A special problem of Brownian motion, and a general theory of Gaussian random functions. Proc. Third Berkeley Symp. on Mathematical Statistics and Probability, *2*, 133 (1956).
5. ——, Sur une classe de courbes de l'espace de Hilbert et sur une équation intégrale non linéaire. Ann. Ec. Norm. Sup. *73*, 121 (1956).
6. ——, Fonctions aléatoires à corrélation linéaire. Illinois J. Math. *1*, 217 (1957).
7. STONE, M. H., Linear transformations in Hilbert space. American Math. Soc. colloquium publication. New York, 1932.
8. WIENER, N., and MASANI, P., The prediction theory of multivariate stochastic processes. Acta Math. *98*, 111 (1957); and *99*, 93 (1958).
9. WOLD, H., A Study in the Analysis of Stationary Time Series. Thesis, University of Stockholm. Uppsala, 1938.

Tryckt den 22 februari 1961

Uppsala 1961. Almqvist & Wiksells Boktryckeri AB