ON THE COEFFICIENT OF COHERENCE FOR WEAKLY STATIONARY STOCHASTIC PROCESSES

By L. H. Koopmans

Sandia Laboratory, Albuquerque

1. Summary. The coefficient of coherence is defined for bivariate weakly stationary stochastic processes which have spectral distributions dominated by a fixed Lebesgue-Stieltjes measure. This quantity is shown to possess two of the important properties which make the ordinary correlation coefficient a desirable measure of linear regression for pairs of random variables. This provides a justification for the already common use of the coefficient of coherence as a measure of linear-regression for pairs of stationarily correlated, weakly stationary time series.

2. Introduction. Let $\mathbf{X} = \{X(t) \mid -\infty < t < \infty\}$ be a bivariate, weakly stationary stochastic process. By this we mean that $X(t)$ is a column vector $(X_1(t), X_2(t))'$ of complex valued random variables over a probability space $(\Omega, \mathcal{A}, P)$ such that

(i) $EX_j(t) = 0, j = 1, 2, -\infty < t < \infty$, and

(ii) $EX_j(t + \tau)\overline{X_k(t)}$ is finite in absolute value for $-\infty < t, \tau < \infty$, and depends, functionally, only on τ, for $j, k = 1, 2$.

The restriction to bivariate processes is only a convenience motivated by the bivariate nature of the coefficient of coherence. The component univariate processes, $\mathbf{X}_j = \{X_j(t) \mid -\infty < t < \infty\}$, may be thought of as any pair of coordinate processes from a general q-variate weakly stationary time series.

In order to avoid an unnecessary duplication of notation we will restrict attention to the case in which $\mathbf{X}$ is a continuous time parameter, continuous-in-the-mean stochastic process. The corresponding results when $\mathbf{X}$ is a discrete time parameter process can be obtained from our theory with only minor alterations.

Let $\mathbf{F}(\lambda) = [F_{jk}(\lambda)]$ be the $(2 \times 2$ matrical) spectral distribution function of $\mathbf{X}$ (see e.g. Cramér (1940)). Throughout the paper we will let μ be a fixed Lebesgue-Stieltjes measure on the Borel sets, $\mathcal{B}$, of the real line, $\mathcal{R}$, which dominates the signed measures, μ_{jk}, induced on $(\mathcal{R}, \mathcal{B})$ by $F_{jk}(\lambda), j, k = 1, 2$.

Let $\mathbf{f}(\lambda) = [f_{jk}(\lambda)]$ be the spectral density function (2×2 matrix of Radon-Nikodym derivatives) of $\mathbf{F}(\lambda)$ with respect to μ. It follows (Cramér 1940) that $f_{jj}(\lambda) \geqq 0, j = 1, 2$, and $f_{12}(\lambda) = \bar{f}_{21}(\lambda)$ a.e. (μ). Then the μ-coefficient of coherence or simply the coefficient of coherence of $\mathbf{X}$ is defined by the expression:

$$\rho(\lambda) = |f_{12}(\lambda)|/[f_{11}(\lambda)f_{22}(\lambda)]^{\frac{1}{2}} \qquad \text{on } \bigcap_{j=1}^{2} [f_{jj}(\lambda) > 0]$$
$$= 0 \qquad \text{otherwise.}$$

Received 4 February 1963; revised 5 November 1963.

Since the spectral densities are defined up to μ-equivalence this definition insures that $\rho(\lambda)$ is uniquely defined on $\mathcal{R}$ up to μ-equivalence.

By an immediate extension of a theorem of Cramér (1940, Theorem 1), $\mathbf{f}(\lambda)$ is non-negative definite and, hence, the Coherency Relation,

$$|f_{12}(\lambda)|^2 \leqq f_{11}(\lambda)f_{22}(\lambda) \qquad \text{a.e. } (\mu),$$

holds. Thus, the coefficient of coherence satisfies the inequalities $0 \leqq \rho(\lambda) \leqq 1$, a.e. (μ). That is, the coefficient of coherence assumes the same range of values as does the modulus of the correlation coefficient for pairs of random variables. Note also that, because of the Coherency Relation, $[\rho(\lambda) = 0] = [f_{12}(\lambda) = 0]$.

One of the most useful properties of the correlation coefficient is the existence of an interpretation of its magnitude as a measure of linear regression. We designate this as Property 1 and give the following intuitive version of it.

PROPERTY 1. *If ρ is the modulus of the correlation coefficient of random variables X_1 and X_2, then the proportion of the variance of X_1 attributable to the linear regression of X_1 on X_2 is ρ^2.*

This property has the following corollaries:

1. *The regression of X_1 on X_2 is zero if and only if $\rho = 0$.*
2. *The regression of X_1 on X_2 is complete* (i.e. *X_1 is a linear function of X_2*) *if and only if $\rho = 1$.*
3. *$1 - \rho^2$ is the proportion of the variance of X_1 not attributable to the linear regression of X_1 and X_2 and, thus, is a measure of the error of estimating X_1 by a linear function of X_2.*

Section 4 will be devoted to establishing the analog of this property for the coefficient of coherence.

The coefficient of coherence is the modulus of a spectral parameter first introduced by Wiener (1930) for pairs of functions possessing generalized harmonic decompositions in the case when μ is Lebesgue measure. He stressed that the usefulness of his "coefficient of coherency" and the ordinary correlation coefficient as measures of linear relationship depended to a great extent on the fact that they are both invariant under linear transformation. Actually, the class of linear transformations which leaves the "coefficient of coherency" invariant is too small to be of general physical interest. It does not contain any transformations which "shift phases", for example. Since "phase shifting" is the rule rather than the exception for physical linear devices, a measure of linear relationship which is to be applicable to physical time series should be invariant under these transformations as well. In Section 5 we will show that the coefficient of coherence, as a measure of linear relationship for a certain family of stochastic processes is invariant under a much larger class of linear transformations—a class which contains, in particular, the transformations of physical interest (Theorem 3). Moreover, we will show (Theorem 4) that this invariance property characterizes the coefficient of coherence in the sense that (almost) any function of $\mathbf{f}(\lambda)$ which is invariant under a certain subclass of these linear transformations must be a function of $\rho(\lambda)$. The corresponding property for the correlation coefficient seems to have received little publicity. Consequently, a precise statement and proof of this result, which we call Property 2, will be given in Theorem 2.

Section 6 is devoted to a discussion of the limitations of the class of measures, μ, which can be used to define the coefficient of coherence for a given process. It is shown (Theorem 6) that, in a sense, the coefficient of coherence does not depend on which dominating measure of this class is selected. The section is concluded with a discussion of the special cases which are of interest from the viewpoint of physical applications.

3. Preliminaries. In this section and in Section 4 we will be concerned with a bivariate weakly stationary stochastic process, $\mathbf{X}$, with spectral density $\mathbf{f}(\lambda)$ with respect to μ. The class of all complex valued random variables, X, over $(\Omega, \mathcal{A}, P)$ such that $EX = 0$ and $\|X\|^2 = E|X|^2 < \infty$ constitutes the familiar Hilbert space $\mathfrak{L}_2(P)$ with inner product $\langle X, Y\rangle = EX\bar{Y}$. This space clearly contains all of the component random variables of $\mathbf{X}$.

For an arbitrary bivariate, or univariate weakly stationary stochastic process, $\mathbf{Y}$, let $\mathfrak{M}(\mathbf{Y})$ be the class of all finite linear combinations of the component random variables of $\mathbf{Y}$. This linear manifold will be termed the linear manifold spanned or generated by $\mathbf{Y}$. Its $\mathfrak{L}_2(P)$ closure, $\bar{\mathfrak{M}}(\mathbf{Y})$, will be called the subspace generated by $\mathbf{Y}$. Unless otherwise stated, we will adopt the notation, $\mathfrak{M} = \mathfrak{M}(\mathbf{X})$, $\bar{\mathfrak{M}} = \bar{\mathfrak{M}}(\mathbf{X})$, $\mathfrak{M}_j = \mathfrak{M}(\mathbf{X}_j)$ and $\bar{\mathfrak{M}}_j = \bar{\mathfrak{M}}(\mathbf{X}_j)$, $j = 1, 2$ for the manifolds and subspaces generated by the given process and its components.

Let $\mathbf{U} = \{U_t \mid -\infty < t < \infty\}$ be the group of unitary transformations of $\bar{\mathfrak{M}}$ onto $\bar{\mathfrak{M}}$ determined by the relations

$$U_t X_j(u) = X_j(u + t), \quad j = 1, 2; \qquad -\infty < u, t < \infty.$$

It will be notationally convenient to adopt a matrix representation for linear operators, T, when they are applied to the vector process $\mathbf{X}$. Thus,

$$TX(t) = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} T_{11}X_1(t) + T_{12}X_2(t) \\ T_{21}X_1(t) + T_{22}X_2(t) \end{pmatrix} = \begin{pmatrix} TX_1(t) \\ TX_2(t) \end{pmatrix}.$$

The existence of the operators T_{ij} for T in the class $\mathbf{T}$ to be constructed below will be clear from the construction.

In particular, it is clear from the definition of $\mathbf{U}$ that U_t has the matrix representation $\begin{bmatrix} U_t & 0 \\ 0 & U_t \end{bmatrix}$, and the defining equations for $\mathbf{U}$ may be summarized by $U_t X(u) = X(t + u)$, $-\infty < u, t < \infty$.

The usual spectral representations associated with a weakly stationary process are based on the spectral decomposition of the associated group $\mathbf{U}$ of unitary transformations due to Stone (Riesz and Nagy, 1955, p. 383):

$$U_t = \int e^{it\lambda}\, dE_\lambda, \tag{1}$$

where $\{E_\lambda \mid -\infty < \lambda < \infty\}$ is the spectral family of projections determined by **U**. (All integrals will be over the range $(-\infty, \infty)$ unless otherwise specified.) This representation will be extended to the matrix version of U_t by setting

$$E_\lambda = \begin{bmatrix} E_\lambda & 0 \\ 0 & E_\lambda \end{bmatrix}$$

and performing the integration component-wise.

This theorem results in the following spectral representation for the process **X**:

$$X(t) = \int e^{it\lambda}\, dZ(\lambda), \qquad -\infty < t < \infty, \tag{2}$$

where $Z(\lambda) = E_\lambda X(0) = (Z_1(\lambda), Z_2(\lambda))'$.

If $X = (X_1, X_2)'$ and $Y = (Y_1, Y_2)'$ are column vectors of elements of $\mathcal{L}_2(P)$, the Gramian of X and Y, $\langle\langle X, Y\rangle\rangle$, is defined to be the 2×2 matrix whose j, kth element is $\langle X_j, Y_k\rangle$. Then, the spectral representation for the covariance of **X** may be written in the form,

$$\langle\langle X(t+\tau), X(t)\rangle\rangle = \int e^{i\tau\lambda}\mathbf{f}(\lambda)\, d\mu(\lambda), \tag{3}$$

where, by Stone's theorem, $\mathbf{f} = (d/d\mu)\langle\langle E_\lambda X(0), X(0)\rangle\rangle$, i.e., $(\mathbf{g} - \mathbf{h})\mathbf{f}(\mathbf{g} - \mathbf{h})^* = 0$, a.e. (μ).

If $\mathbf{g}(\lambda)$ denotes a measurable function on $\mathcal{R}$ whose values are 2×2 matrices with complex entries, then **G** is defined to be the class of all such functions for which

$$\|\mathbf{g}\|^2 = \operatorname{tr} \int \mathbf{g}\mathbf{f}\mathbf{g}^*\, d\mu < \infty, \tag{4}$$

where * denotes conjugate transpose. The elements of **G** are to be identified by **f** equivalence. That is, $\mathbf{g} \in \mathbf{G}$ will represent any of the elements **h** satisfying Expression 4 for which $\mathbf{h} = \mathbf{g}$, a.e. $(\mathbf{f})\ \lambda$. Then the function $\| \ \|$ defined by Equation 4 is a norm for **G** interpreted as a linear space over the field of complex scalars. By an extension of Riesz's proof of the Riesz-Fischer Theorem (Riesz and Nagy, 1955, p. 59), **G** can be shown to be complete relative to this norm. (This result has also been recently announced by M. Rosenberg (1963).) Thus, with the inner product $(\mathbf{g}, \mathbf{h}) = \operatorname{tr} \int \mathbf{g}\mathbf{f}\mathbf{h}^*\, d\mu$, **G** is a Hilbert space.

We will be interested in the class, **T**, of all linear transformations, T, which map **X** again into continuous-in-the-mean, bivariate, weakly stationary processes, $T\mathbf{X} = \{TX(t) \mid -\infty < t < \infty\}$, relative to the group **U**. That is, the elements of $T\mathbf{X}$ are to satisfy the equations

$$U_t TX(u) = TX(t+u), \qquad -\infty < t, u < \infty. \tag{5}$$

In order for this equation to be meaningful, it is necessary for $TX(t)$ to be in $\overline{\mathfrak{M}} \times \overline{\mathfrak{M}}$ for all t.

This property completely determines each operator when $TX(0)$ is given, since T is then defined on $\mathbf{X}$ by $TX(t) = U_tTX(0)$. T is extended by linearity to $\mathfrak{M}$ and thence uniquely to a domain, $\mathfrak{D}(T)$, in $\bar{\mathfrak{M}}$ to be defined below.

The following construction of the class $\mathbf{T}$ provides a linear isomorphism between $\mathbf{G}$ and $\mathbf{T}$ via the Spectral Representation 1. Let $\mathbf{T}'$ consist of all operators, T, which are defined on $\mathbf{X}$ by an expression of the form $T = \sum_{j=1}^n \mathbf{C}_j U_{t_j}$, for some integer n, real numbers, t_j, and 2×2 matrices of complex constants, $\mathbf{C}_j$. From the Representation (1) for $\mathbf{U}$,

$$T = \int \mathbf{g}(\lambda)\, dE_\lambda, \tag{6}$$

where

$$\mathbf{g}(\lambda) = \sum_{j=1}^n \mathbf{C}_j e^{it_j\lambda}. \tag{7}$$

We will denote this correspondence between T and $\mathbf{g}$ by $T \leftrightarrow \mathbf{g}$.

It follows from Stone's theorem that U_t and T commute. Thus, $TX(t) = TU_tX(0) = \int e^{it\lambda}\mathbf{g}(\lambda)\, dZ(\lambda)$, and $U_tTX(u) = TU_tX(u) = TX(t+u)$, $-\infty < t, u < \infty$. This implies, by the definition of $\mathbf{T}$, that the appropriate linear extension of T is in $\mathbf{T}$.

To see that $\mathbf{g} \in \mathbf{G}$, it suffices to note that there exists a constant M such that the components of $\mathbf{g}$ satisfy the inequality $|g_{ij}| \leqq M$ for all λ. Then it is easy to show that

$$\operatorname{tr} \int \mathbf{g}\mathbf{f}\mathbf{g}^*\, d\mu \leqq 8M^2 \operatorname{tr} \int \mathbf{f}\, d\mu < \infty. \tag{8}$$

Moreover, by virtue of Equation 6 and the discussion following it, every trigonometric polynomial of the form given by Expression 7 corresponds to some $T \in \mathbf{T}'$.

We will now show that $\mathbf{T}$ is the class of operators, T, such that either $T \in \mathbf{T}'$ or there exists a sequence $\{T_n\} \subseteq \mathbf{T}'$ such that $T_n \to T$ in the strong sense on $\mathbf{X}$. That is,

$$\lim_{n\to\infty} \operatorname{tr} \langle\langle (T - T_n)X(t), (T - T_n)X(t)\rangle\rangle = 0, \quad -\infty < t < \infty.$$

If $\mathbf{T}''$ denotes this class of operators, it is clear that $\mathbf{T}'' \subseteq \mathbf{T}$, since the commutability of the elements of $\mathbf{U}$ with those of $\mathbf{T}'$ is preserved in the limit. But if $T \in \mathbf{T}$, then, $TX(0) = Z \in \bar{\mathfrak{M}} \times \bar{\mathfrak{M}}$ and either Z is a finite linear combination of elements of $\mathbf{X}$ (i.e. is in $\mathfrak{M} \times \mathfrak{M}$);

$$Z = \sum_{j=1}^m \mathbf{C}_j X(t_j) = \sum_{j=1}^m \mathbf{C}_j U_{t_j} X(0),$$

or is a limit of a sequence of such linear combinations; $\lim_n \operatorname{tr} \langle\langle Y_n - Z, Y_n - Z\rangle\rangle$

$= 0$. In the first case $T \varepsilon \mathbf{T}'$ by virtue of Equation (5). If we set $T_nX(t) = U_tY_n$ in the second case then, properly extended, $T_n \varepsilon \mathbf{T}'$ and

$$\begin{aligned}\lim_n \operatorname{tr} \langle\langle (T_n - T)X(t), (T_n - T)X(t)\rangle\rangle &= \lim_n \operatorname{tr} \langle\langle U_t(Y_n - Z), U_t(Y_n - Z)\rangle\rangle \\ &= \lim_n \operatorname{tr} \langle\langle Y_n - Z, Y_n - Z\rangle\rangle = 0.\end{aligned}$$

Thus, $T \varepsilon \mathbf{T}''$, and $\mathbf{T} = \mathbf{T}''$.

Now, suppose $T \varepsilon \mathbf{T} \sim \mathbf{T}'$ (where $\sim$ denotes set theoretic difference) and let $\{T_n\} \subseteq \mathbf{T}'$ such that $T_n \to T$ strongly on $\mathbf{X}$ and $T_n \leftrightarrow \mathbf{g}_n$. Then, as a consequence of the Spectral Representation (1),

$$\begin{aligned}\operatorname{tr} \langle\langle (T_n - T_m)X(t), (T_n - T_m)X(t)\rangle\rangle &= \operatorname{tr} \int (\mathbf{g}_n - \mathbf{g}_m)\mathbf{f}(\mathbf{g}_n - \mathbf{g}_m)^* \, d\mu \\ &= \|\mathbf{g}_m - \mathbf{g}_n\|^2 \to 0, \qquad -\infty < t < \infty.\end{aligned} \tag{9}$$

Since $\mathbf{G}$ is complete, there exists $\mathbf{g} \varepsilon \mathbf{G}$ such that $\|\mathbf{g}_n - \mathbf{g}\| \to 0$. We define the relation $\leftrightarrow$ at T by $T \leftrightarrow \mathbf{g}$. It follows from the Representation (1) that

$$TX(t) = \int e^{it\lambda}\mathbf{g}(\lambda) \, dZ(\lambda), \qquad -\infty < t < \infty.$$

It can be shown by applying the inequality $\|\mathbf{g}\| \leqq \sum_i \sum_j \|g_{ij}\|_j$, $\|g\|_j^2 = \int |g|^2 f_{jj} \, d\mu$, to the univariate result, that the trigonometric polynomials of the form given by Expression (7) are dense in $\mathbf{G}$. Thus, if $\mathbf{g} \in \mathbf{G}$, there exists a sequence, $\{\mathbf{g}_n\}$, of these polynomials such that $\|\mathbf{g}_n - \mathbf{g}\| \to 0$. Expression (9) then implies that $\{T_nX(t)\}$ is a Cauchy sequence in $\mathfrak{M} \times \mathfrak{M}$, $-\infty < t < \infty$, where $\mathbf{g}_n \leftrightarrow T_n \varepsilon \mathbf{T}'$. Since $\mathfrak{M}$ is a closed subspace of $\mathfrak{L}_2(P)$ there exists $Y \varepsilon \mathfrak{M} \times \mathfrak{M}$ such that

$$\lim_n T_nX(t) = U_t \lim_n T_nX(0) = U_tY.$$

The assignment $TX(t) = U_tY$ then uniquely determines an element of $\mathbf{T}$ as discussed above, since Equation (5) is satisfied;

$$U_tTX(u) = U_tU_uY = U_{t+u}Y = TX(t + u).$$

We write $T \leftrightarrow \mathbf{g}$ and it follows that $\leftrightarrow$, as a mapping from $\mathbf{T}$ to $\mathbf{G}$, is onto.

That $\leftrightarrow$ is one-to-one follows easily from the fact that the elements of $\mathbf{T}$ are uniquely determined by their values on $\mathbf{X}$. The linearity of $\leftrightarrow$ follows from the integral representation of elements of $\mathbf{T}$ on $\mathbf{X}$ and the fact that if T_1, $T_2 \varepsilon \mathbf{T}$ and α, β are any complex numbers, then $\alpha T_1 + \beta T_2 \varepsilon \mathbf{T}$ with $\mathfrak{D}(\alpha T_1 + \beta T_2) \supseteq \mathfrak{D}(T_1) \cap \mathfrak{D}(T_2)$; i.e., $\mathbf{T}$ is a linear space over the complex field. This establishes that $\leftrightarrow$ is a linear isomorphism between $\mathbf{G}$ and $\mathbf{T}$.

A representation of the domains of the elements of $\mathbf{T}$ is provided by the linear isomorphism between $\mathfrak{M}$ and the class $\mathbf{H}$ of all row vectors, $x(\lambda)$, of measurable complex valued functions (identified by equivalence as above) such that $\int x\mathbf{f}x^* \, d\mu < \infty$.

This isomorphism, which we also denote by $\leftrightarrow$, can be established by an argument paralleling the above for $\mathbf{T}$ and $\mathbf{G}$. It follows that if $X \leftrightarrow x(\lambda)$, $X = \int x(\lambda) \, dZ(\lambda)$ and $\|X\|^2 = \int x\mathbf{f}x^* \, d\mu$. Now, if $T \leftrightarrow \mathbf{g}$, $X \varepsilon \mathfrak{D}(T)$ if and only if $X \leftrightarrow x(\lambda)$ and $\int x\mathbf{gfg}^*x^* \, d\mu < \infty$. Also, $TX = \int x(\lambda) \, \mathbf{g}(\lambda) \, dZ(\lambda)$.

These results and others needed in the sequel are summarized in Theorem A below. This proof follows closely that for the corresponding univariate result (which we will call Theorem A_1) given, for a special case, in (Doob, 1953, p. 534). The remainder of the proof is straightforward and is omitted.

THEOREM A.

1. *The correspondence,* $\leftrightarrow$, *between* $\mathbf{T}$ *and* $\mathbf{G}$ *is a linear isomorphism.*

2. *If* $T \leftrightarrow \mathbf{g}$, *the bivariate, weakly stationary stochastic process* $T\mathbf{X}$ *has the spectral representation,*

$$TX(t) = \int e^{it\lambda}\mathbf{g}(\lambda) \, dZ(\lambda), \qquad -\infty < t < \infty. \tag{10}$$

The spectral distribution of $T\mathbf{X}$ *is absolutely continuous with respect to* μ *and has spectral density* $\mathbf{gfg}^*$.

3. *Let* $T \leftrightarrow \mathbf{g}$ *and* $R \leftrightarrow \mathbf{h}$. *Then if* $x \leftrightarrow X \varepsilon \mathfrak{D}(T)$, $y \leftrightarrow Y \varepsilon \mathfrak{D}(R)$,

$$\langle TX, RY \rangle = \int x\mathbf{gfh}^*y^* \, d\mu. \tag{11}$$

In particular,

$$\langle\langle TX(t), RX(s) \rangle\rangle = \int e^{i(t-s)\lambda}\mathbf{g}(\lambda)\mathbf{f}(\lambda)\mathbf{h}^*(\lambda) \, d\mu(\lambda), \qquad -\infty < t,s < \infty. \tag{12}$$

4. *Further, if the matrix product* $\mathbf{gh} \varepsilon \mathbf{G}$ *and* $S \leftrightarrow \mathbf{gh}$, *then* $\mathfrak{M} \subseteq \mathfrak{D}(RT) \subseteq \mathfrak{D}(S) \subseteq \bar{\mathfrak{M}}$ *and* $S = RT$ *on* $\mathfrak{D}(RT)$, *where* RT *denotes the composition of* R *and* T. *Thus, if* $\mathfrak{D}(RT) = \bar{\mathfrak{M}}$, $RT \leftrightarrow \mathbf{gh}$.

The set of continuous orthogonal projections, $\mathbf{T}^\pi$, in $\mathbf{T}$ will be of particular importance to us. The following corollary to Theorem A provides a characterization of this subclass in terms of a subclass of the elements of $\mathbf{G}$.

COROLLARY A.

1. *Let* $\mathbf{G}^\pi$ *be the collection of* $\mathbf{g} \varepsilon \mathbf{G}$ *for which* (a.e. (f)), (i) $\mathbf{g}^2 = \mathbf{g}$, (ii) $\mathbf{gf} = \mathbf{fg}^*$, *and such that* (iii) $\int x\mathbf{gfg}^* x^* \, d\mu \leqq \int x\mathbf{f}x^* \, d\mu$ *for every* $x \varepsilon \mathbf{H}$. *Then* $T \varepsilon \mathbf{T}^\pi$ *if and only if* $T \leftrightarrow \mathbf{g} \varepsilon \mathbf{G}^\pi$.

2. *Let* I *denote the identity operator on* $\bar{\mathfrak{M}}$. *If* $\mathbf{g} \leftrightarrow T \varepsilon \mathbf{T}^\pi$, *then the complementary orthogonal projection,* $I - T$, *is in* $\mathbf{T}^\pi$ *and* $I - T \leftrightarrow \mathbf{1} - \mathbf{g}$, *where* $\mathbf{1}$ *is the* 2×2 *identity matrix.*

PROOF. Necessary and sufficient conditions that T be a continuous orthogonal projection are $T^2 = T$ and $T = T^*$ on $\bar{\mathfrak{M}}$, where * denotes adjoint (Riesz and Nagy, 1955, p. 267). Furthermore, such an operator satisfies $\|TX\| \leqq \|X\|$ for all $X \varepsilon \bar{\mathfrak{M}}$.

Let $T \varepsilon \mathbf{T}^{\pi}$ and let $T \leftrightarrow \mathbf{g} \varepsilon \mathbf{G}$. Then Condition (iii) follows from Equation 11 and the inequality $\|TX\|^2 \leqq \|X\|^2$.

If $T = T^*$, by Equation(12),

$$(13)\quad \langle\langle T(t+\tau), X(t)\rangle\rangle = \int e^{i\tau\lambda}\mathbf{gf1}\,d\mu = \int e^{i\tau\lambda}\mathbf{1fg}^*\,d\mu = \langle\langle X(t+\tau), TX(t)\rangle\rangle$$

for $-\infty < t, \tau < \infty$. Condition (ii) is then a consequence of the one-to-one property of the Fourier transform on $\mathcal{L}_1(\mu)$.

We now show that[1] $T^2 = T$ implies $\mathbf{g} = \mathbf{g}^2 \varepsilon \mathbf{G}$. Let $\mathbf{g} = [g_{jk}]$, $S_n = \{\lambda \mid |g_{jk}(\lambda)| \leqq n,\ j, k = 1, 2\}$, $n = 1, 2, \cdots$, and let I_n be the set characteristic function of S_n. Let $\mathbf{g}_n = I_n \mathbf{1}$. Clearly, $\mathbf{g}_n \varepsilon \mathbf{G}$, $\mathbf{g}_n^2 = \mathbf{g}_n$ and $\mathbf{g}_n\mathbf{g} = \mathbf{g}\mathbf{g}_n \varepsilon \mathbf{G}$. Moreover,

$$\int x\mathbf{g}_n\mathbf{gfg}^*\mathbf{g}_n x^*\,d\mu \leqq \int x\mathbf{gfg}^* x^*\,d\mu \leqq \int x\mathbf{f}x^*\,d\mu.$$

Thus, if $T_n \leftrightarrow \mathbf{g}_n$, $\mathfrak{D}(T_nT) = \mathfrak{D}(T) = \mathfrak{M}$, $T_n^2 = T_n$ and $T_nT = TT_n \leftrightarrow \mathbf{g}_n\mathbf{g}$. Now, by Inequality (8),

$$\operatorname{tr}\int (\mathbf{gg}_n)^2\mathbf{f}((\mathbf{gg}_n)^2)^*\,d\mu \leqq 8n^4 \operatorname{tr}\int \mathbf{f}\,d\mu < \infty.$$

Hence, $(\mathbf{gg}_n)^2 \varepsilon \mathbf{G}$, and, since $\mathfrak{D}((TT_n)^2) = \mathfrak{D}(TT_n) = \mathfrak{M}$, $(TT_n)^2 \leftrightarrow (\mathbf{gg}_n)^2$ by Theorem A. But $(TT_n)^2 = T^2T_n^2 = TT_n$. Thus,

$$I_n\mathbf{g}^2 = (\mathbf{gg}_n)^2 = \mathbf{gg}_n = I_n\mathbf{g} \quad \text{a.e.}(\mathbf{f}) \quad \text{for all } n.$$

This implies that $\mathbf{g} = \mathbf{g}^2 \varepsilon \mathbf{G}$.

Let $\mathbf{g} \varepsilon \mathbf{G}$ satisfy Conditions (i)–(iii), and let $\mathbf{g} \leftrightarrow T$. Then $\|TX\| \leqq \|X\|$ for all $X \varepsilon \mathfrak{M}$ by Condition (iii).

It is easily shown that Condition (ii) is equivalent to $\mathbf{gf} = \mathbf{fg}^*$ a.e. (μ). Thus, since T is linear and continuous and inner product is continuous, Equation (8) yields $\langle TX, Y\rangle = \langle X, TY\rangle$ for all $X, Y \varepsilon \mathfrak{M}$. Hence $T = T^*$.

Finally, $\mathbf{g} = \mathbf{g}^2$ and $\mathfrak{D}(T) = \mathfrak{M}$ yields $T^2 \leftrightarrow \mathbf{g}^2 = \mathbf{g} \leftrightarrow T$, i.e., $T^2 = T$ on $\mathfrak{M}$.

The proof of Part 2 is clear.

An important subclass of $\mathbf{T}^{\pi}$ is isolated in the following immediate consequence of Corollary A.

Corollary B. *Let I_S be the set characteristic function of a Borel set S and let $\mathbf{g}_S = I_S\mathbf{1}$. Then $\mathbf{g}_S \varepsilon \mathbf{G}^{\pi}$ for all $S \varepsilon \mathfrak{B}$. If $T_S \leftrightarrow \mathbf{g}_S$, then $I - T_S = T_{S^c} \leftrightarrow \mathbf{g}_{S^c}$, where S^c is the complement of S. The spectral density of $T_S\mathbf{X}$ is $I_S\mathbf{f}$.*

If $\mathbf{X}$ is a bivariate process or a univariate component of a bivariate process we will use the notation $\mathbf{X} = \mathbf{Y} + \mathbf{Z}$ to mean that there exists $T \varepsilon \mathbf{T}^{\pi}$ such that $\mathbf{Y} = T\mathbf{X}$ and $\mathbf{Z} = (I - T)\mathbf{X}$. If $\mathfrak{M}$, $\mathfrak{M}_Y$ and $\mathfrak{M}_Z$ are the subspaces generated by these processes it follows that

$$\mathfrak{M} = \mathfrak{M}_Y \oplus \mathfrak{M}_Z,\ \mathfrak{M}_Y \perp \mathfrak{M}_Z,$$

[1] This part of the proof is due to M. J. Norris.

where $\oplus$ denotes direct (orthogonal) sum, thus

$$\mathfrak{M}_Y = \mathfrak{M} \ominus \mathfrak{M}_Z \quad \text{and} \quad \mathfrak{M}_Z = \mathfrak{M} \ominus \mathfrak{M}_Y,$$

the orthogonal complements of $\mathfrak{M}_Z$ and $\mathfrak{M}_Y$ in $\mathfrak{M}$.

Moreover, $X(t) = Y(t) + Z(t)$ is the unique representation of $X(t)$ as a sum of elements in $\mathfrak{M}_Y$ and $\mathfrak{M}_Z$, $-\infty < t < \infty$.

4. The analog of Property 1. Let X_1 and X_2 be random variables in $\mathcal{L}_2(P)$. The modulus of the correlation coefficient of X_1 and X_2 is

$$\rho = |\langle X_1, X_2\rangle| / \|X_1\| \, \|X_2\| \quad \text{for} \quad \|X_j\| > 0, \qquad j = 1, 2.$$

Let $\mathfrak{M}_j$ be the linear subspace of $\mathcal{L}_2(P)$ spanned by X_j (i.e. $\mathfrak{M}_j = \{cX_j \mid c$ complex$\}$), and let $\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2$. Then a precise formulation of Property 1 can be given as follows:

PROPERTY 1. *There exist unique random variables* $X_1^{(j)}$, $j = 1, 2$ *such that* (i) $X_1 = X_1^{(1)} + X_1^{(2)}$, (ii) $X_1^{(1)} \varepsilon \mathfrak{M}_2$ *and* (iii) $X_1^{(2)} \varepsilon \mathfrak{M} \ominus \mathfrak{M}_2$. *Moreover,* (iv) $\|X_1^{(1)}\| = \rho \|X_1\|$, (v) $\|X_1^{(2)}\| = (1 - \rho^2)^{\frac{1}{2}} \|X_1\|$.

Immediate corollaries are:

COROLLARY 1. $\rho = 0$ *if and only if* $\mathfrak{M}_1 \perp \mathfrak{M}_2$.

COROLLARY 2. $\rho = 1$ *if and only if* $\mathfrak{M}_1 = \mathfrak{M}_2$.

Clearly $X_1^{(1)}$ is the orthogonal projection of X_1 on $\mathfrak{M}_2$. The intuitive interpretation of this property, given in the Introduction, easily follows from this result by adopting the variance ratio, $\|X_1^{(1)}\|^2 / \|X_1\|^2$, as a measure of the degree of linear regression of X_1 on X_2.

The usual proof of this result consists of the explicit determination of the projection by a minimization procedure. In Theorem 1 a spectral parameter closely resembling this projection is shown to lead to the appropriate element of $\mathbf{T}^\pi$ to establish the theorem.

If the definition of the modulus of the correlation coefficient is extended by setting $\rho = 0$ in the (trivial) case $\|X_1\| \cdot \|X_2\| = 0$, Property 1 and Corollary 1 remain the same, since $\mathfrak{M}_j = \{0\}$ is orthogonal to every element of $\mathfrak{M}$. In this case Corollary 2 becomes:

COROLLARY 2′. $\rho = 1$ *if and only if* $\mathfrak{M}_j \neq \{0\}, j = 1, 2$, *and* $\mathfrak{M}_1 = \mathfrak{M}_2$.

This version of Property 1 has the following analog for the coefficient of coherence.

THEOREM 1. *Let* $\mathbf{X}$ *be a bivariate weakly stationary stochastic process with component univariate processes* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *and let* $\mathfrak{M}$, $\mathfrak{M}_1$ *and* $\mathfrak{M}_2$ *be the linear subspaces generated by these processes. Then there exist unique univariate weakly stationary process* $\mathbf{X}_1^{(j)}$ $j = 1, 2$ *such that* (i) $\mathbf{X}_1 = \mathbf{X}_1^{(1)} + \mathbf{X}_1^{(2)}$, (ii) $\mathbf{X}_1^{(1)} \subseteq \mathfrak{M}_2$, *and* (iii) $\mathbf{X}_1^{(2)} \subseteq \mathfrak{M} \ominus \mathfrak{M}_2$.

If $\mathbf{f} = [f_{jk}]$ *is the spectral density and* $\rho(\lambda)$ *is the coefficient of coherence of* $\mathbf{X}$, *then, for* $j = 1, 2$, *the spectral density,* $f^{(j)}$, *of* $\mathbf{X}_1^{(j)}$ *is, to* $\mathbf{f}$ *equivalence,*

$$f^{(1)}(\lambda) = \rho^2(\lambda) f_{11}(\lambda),$$

and

$$f^{(2)}(\lambda) = (1 - \rho^2(\lambda)) f_{11}(\lambda).$$

PROOF. Let

$$\mathbf{g}(\lambda) = \begin{bmatrix} 0 & h(\lambda) \\ 0 & 1 \end{bmatrix},$$

where $h = f_{12}/(f_{22} + I_2)$ and I_2 is the set characteristic function of $[f_{22} = 0]$. Then, since $h f_{22} = f_{12}$, a.e. (μ), an easy computation establishes that

$$\mathbf{gfg}^* = \begin{bmatrix} \rho^2 f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \qquad \text{a.e.}(\mu). \tag{14}$$

Thus, $\operatorname{tr} \int \mathbf{gfg}^* \, d\mu \leqq \operatorname{tr} \int \mathbf{f} \, d\mu < \infty$, since $0 \leqq \rho \leqq 1$.

It is easily verified that $\mathbf{g}^2 = \mathbf{g}$ and $\mathbf{gf} = \mathbf{fg}^*$, a.e. (μ). Also, if $u(\lambda) = (u_1(\lambda), u_2(\lambda))$, $u\mathbf{f}u^* - u\mathbf{gfg}^*u^* = (1 - \rho^2) |u_1|^2 f_{11} \geqq 0$, a.e. (μ), and Property (iii) of Corollary A is satisfied. Thus, $\mathbf{g} \leftrightarrow T \varepsilon \mathbf{T}^\pi$. Let $\mathcal{R}_T$ denote the range of T. We will now establish that $\mathcal{R}_T = \overline{\mathfrak{M}}_2$.

From Expression (10) and the definition of $\mathbf{g}$,

$$TX_1(t) = \int e^{it\lambda} h \, dZ_2,$$

and

$$TX_2(t) = X_2(t), \qquad -\infty < t < \infty. \tag{15}$$

By Theorem A_1, $TX_1(t) \varepsilon \overline{\mathfrak{M}}_2$ for all t, thus $\mathcal{R}_T \subseteq \overline{\mathfrak{M}}_2$.

But, from Expression (15), $\mathcal{R}_T$ contains the generators of $\mathfrak{M}_2$. Thus, since $\mathcal{R}_T$ is closed, $\overline{\mathfrak{M}}_2 \subseteq \mathcal{R}_T$ and we have shown that $\overline{\mathfrak{M}}_2 = \mathcal{R}_T$.

The proof of the theorem now follows from the assignment: $\mathbf{X}_1^{(1)} = T\mathbf{X}_1$ and $\mathbf{X}_1^{(2)} = (I - T)\mathbf{X}_1$. The spectral densities $f^{(1)}$ and $f^{(2)}$ are the elements in the first row and first column of the matrices $\mathbf{gfg}^*$ and $(\mathbf{1} - \mathbf{g})\, \mathbf{f}(\mathbf{1} - \mathbf{g})^* = \mathbf{f} - \mathbf{gfg}^*$, which can be read directly from Expression (14).

COROLLARY 1.

A. *Let* $T_0 \varepsilon \mathbf{T}^\pi$ *be the projection determined by setting* $S = [\rho = 0]$ *in Corollary B. If* $\overline{\mathfrak{M}}_j^0$ *is the linear subspace generated by* $T_0\mathbf{X}_j$, $j = 1, 2$, *then* $\overline{\mathfrak{M}}_1^0 \perp \overline{\mathfrak{M}}_2^0$.

B. *Moreover, if* $\overline{\mathfrak{M}}_1 \perp \overline{\mathfrak{M}}_2$ *for a process,* $\mathbf{X}$, *then* $\rho(\lambda) = 0$ a.e. (μ).

PROOF.

A. An application of Theorem 1 to $\mathbf{X}$ yields

$$T_0\mathbf{X}_1 = T_0\mathbf{X}_1^{(1)} + T_0\mathbf{X}_1^{(2)},$$

where $T_0\mathbf{X}_1^{(1)} \subseteq \overline{\mathfrak{M}}_2^0$ and $T_0\mathbf{X}_1^{(2)} \subseteq \mathcal{R}_{T_0} \ominus \mathfrak{M}_2^0$. Moreover, it is easily seen that the spectral density of $T_0\mathbf{X}_1^{(1)}$ is

$$f_0^{(1)} = I_0\rho^2 f_{11},$$

where ρ and f_{11} are the spectral parameters of $\mathbf{X}$ defined above.

But $I_0\rho^2 = 0$ a.e.(μ) by definition of ρ and I_0. Thus,

$$\|T_0 X_1^{(1)}(t)\|^2 = \int f_0^{(1)}\, d\mu = 0, \qquad -\infty < t < \infty,$$

which implies $T_0\mathbf{X}_1 = T_0\mathbf{X}_1^{(2)}$. Thus, $\mathfrak{M}_1^0 \subseteq \mathfrak{R}_{T_0} \ominus \mathfrak{M}_2^0$, or $\mathfrak{M}_1^0 \perp \mathfrak{M}_2^0$.

B. If $\mathfrak{M}_1 \perp \mathfrak{M}_2$, then $\langle X_1(s), X_2(t)\rangle = 0$ for $-\infty < t, s < \infty$. By virtue of Equation (12) it follows that $\int \exp(i\tau\lambda) f_{12}\, d\mu = 0$ for $-\infty < \tau < \infty$. The one-to-one property of the Fourier transform on $\mathfrak{L}_1(\mu)$ then implies that $f_{12} = 0$ and, thus, $\rho = 0$ a.e. (μ).

COROLLARY 2.

A. *Let $T_1 \varepsilon \mathbf{T}^\pi$ be the projection determined by setting $S = [\rho = 1]$ in Corollary B, and let $\mathfrak{M}_j^1$ be the linear subspace of $\mathfrak{M}$ generated by $T_1\mathbf{X}_j$, $j = 1, 2$. Then $\mathfrak{M}_1^1 = \mathfrak{M}_2^1$, and $\mathfrak{M}_j^1 \neq \{0\}$, $j = 1, 2$, if and only if $\mu\{[\rho = 1]\} > 0$.*

B. *If $\mathbf{X}$ is a bivariate process for which $\mathfrak{M}_1 \subseteq \mathfrak{M}_2$ or $\mathfrak{M}_2 \subseteq \mathfrak{M}_1$, then*

$$\rho(\lambda) = 1 \quad \textit{whenever} \quad \rho(\lambda) \neq 0 \quad \text{a.e. } (\mu).$$

PROOF.

A. Let I_1 denote the set characteristic function of $[\rho = 1]$. Then, applying Theorem 1 to $T_1\mathbf{X}$, $T_1\mathbf{X}_1 = T_1\mathbf{X}_1^{(1)} + T_1\mathbf{X}_1^{(2)}$ where $T_1\mathbf{X}_1^{(1)} \subseteq \mathfrak{M}_2^1$. The spectral density of $T_1\mathbf{X}_1^{(2)}$ is $I_1(1-\rho^2) f_{11} = 0$ a.e. (μ) which implies $T_1\mathbf{X}_1^{(2)} = \{0\}$. Thus $T_1\mathbf{X}_1 \subseteq \mathfrak{M}_2^1$ which yields $\mathfrak{M}_1^1 \subseteq \mathfrak{M}_2^1$. The reverse inclusion is obtained in the same way by exchanging the component processes of $T_1\mathbf{X}$ before applying Theorem 1. Moreover, since $[\rho = 1] \subseteq \bigcap_{j=1}^2 [f_{jj} > 0]$ and $\|T_1 X_j(t)\|^2 = \int I_1 f_{jj}\, d\mu$, it is clear that $\mathfrak{M}_j^1 \neq \{0\}$, $j = 1, 2$, if and only if $\mu\{[\rho = 1]\} > 0$.

B. It suffices to consider the case $\mathfrak{M}_1 \subseteq \mathfrak{M}_2$. In the notation of Theorem 1, $\mathbf{X}_1 = \mathbf{X}_1^{(1)} + \mathbf{X}_1^{(2)}$. Then $\mathfrak{M}_1 \subseteq \mathfrak{M}_2$ implies $\mathbf{X}_1^{(2)} = \{0\}$, thus $\int f^{(2)}\, d\mu = \int (1 - \rho^2) f_{11}\, d\mu = 0$. But this implies $[\rho = 1] \supseteq [f_{11} > 0] \supseteq \bigcap_{j=1}^2 [f_{jj} > 0] \supseteq [\rho \neq 0]$ a.e. (μ) as was to be shown.

Results paralleling those of Corollaries 1 and 2 were first obtained by Kolmogorov (1941) in the discrete time parameter case. In his terminology, $\mathbf{X}_i$ is said to be subordinate to $\mathbf{X}_j$ if $\mathfrak{M}_i \subseteq \mathfrak{M}_j$.

5. The analog of Property 2. Let $\mathfrak{L}_2^+(P) = \mathfrak{L}_2(P) \sim \{0\}$. The modulus of the correlation coefficient, $\rho(X)$, is then a measure of linear dependence defined for all $X = (X_1, X_2)\, \varepsilon\, \mathfrak{X} = [\mathfrak{L}_2^+(P)]^2$. Let T be a linear transformation from $\mathfrak{L}_2(P)$ to $\mathfrak{L}_2(P)$ and let $TX = (TX_1, TX_2)$. For a given $X \varepsilon \mathfrak{X}$, let

$$\mathbf{T}_I^X = \{T \mid \rho(TX) = \rho(X)\},$$

the class of linear transformations for which ρ is invariant at X. This class is quite large. It contains, for example, all operators whose restrictions to the plane $\mathfrak{M}^X = \{c_1X_1 + c_2X_2 \mid c_1, c_2 \text{ complex}\}$ are rotations and reflections in $\mathfrak{M}^X$. It also contains all operators, T, such that

$$TX_j = a_jX_j, \qquad a_j \neq 0, j = 1, 2. \tag{16}$$

We will denote the subclass of $\mathbf{T}_I^X$ whose elements satisfy Expression (16) by $\mathbf{T}_0^X$. If we let $\mathfrak{M}_j^X = \{cX_j \mid c \text{ complex}\}$, $j = 1, 2$, it is easily seen that another characterization of this subclass (which we will need later in the time series context) is

$$\mathbf{T}_0^X = \{T \mid T\mathfrak{M}_j^X \subseteq \mathfrak{M}_j^X \quad \text{and} \quad \mathfrak{N}(T) \cap \mathfrak{M}_j^X = \{0\}, j = 1, 2\},$$

where $\mathfrak{N}(T)$ denotes the null space of T.

In the construction of a function, $H(X)$, to measure the degree of linear relationship between the components of X, it is natural to require that the function at least be invariant at X with respect to the class $\mathbf{T}_0^X$ for all $X \varepsilon \mathfrak{X}$. Physically, this means that each random variable can be independently subjected to a series of possibly unknown linear transformations, and the linear relationship of the original pair can be recovered by evaluating H at the resultant pair of random variables. Put in another way, from the projective interpretation of linear regression given by Property 1, a measure of linear relationship should depend on the angular separation of $\mathfrak{M}_1$ and $\mathfrak{M}_2$ and not on the magnitudes of their individual elements. Thus, we will require that

$$H(TX) = H(X) \quad \text{for all} \quad T \varepsilon \mathbf{T}_0^X.$$

Another requirement which seems natural to impose on $H(X)$ is that it depends upon X through the intrinsic inner product structure of the Hilbert space $\mathfrak{L}_2(P)$ restricted to $\mathfrak{M}$. That is,

$$H(X) = V(\|X_1\|, \|X_2\|, \langle X_1, X_2\rangle),$$

for some function V.

The modulus of the correlation coefficient clearly satisfies both of these requirements. It is the purpose of Theorem 2 to demonstrate that, in a sense to be made precise, it is the only measure of linear relationship on $\mathfrak{X}$ which does satisfy both requirements. This is the characterization of $\rho(X)$ we have called Property 2.

Let $\mathbf{V}$ be the class of all ordered pairs (V, α) where V is a function such that $\mathfrak{D}_V$, the domain of V, contains the subset $\mathfrak{D} = \{(x, y, z) \mid |z| \leqq xy, x \geqq 0, y \geqq 0\}$ of complex Euclidean 3 space, and α is an equivalence relation on $V(\mathfrak{D})$. Similarly, let $\mathbf{W}$ be the class of all ordered pairs (W, β) where W is a function such that $\mathfrak{D}_W$ contains the real (closed) interval $\mathfrak{g} = [0, 1]$ and β is an equivalence relation on $W(\mathfrak{g})$.

THEOREM 2. *If $H(X)$ is any function defined on $\mathfrak{X}$ which satisfies the conditions,* (i) $H(X) \,\alpha\, V(\|X_1\|, \|X_2\|, \langle X_1, X_2\rangle)$, *and* (ii) $H(TX) \,\alpha\, H(X)$ *for all* $T \varepsilon \mathbf{T}_0^X$, *then there exists $(W, \beta) \varepsilon \mathbf{W}$ such that $W(\mathfrak{g}) \subseteq V(\mathfrak{D})$, $\beta = \alpha$ restricted to $W(\mathfrak{g})$, and*

$$H(X) \,\alpha\, W(\rho(X)).$$

PROOF. For each $X \varepsilon \mathfrak{X}$ let

$$a_1(X) = \exp(-i \arg \langle X_1, X_2\rangle)/\|X_1\|$$
$$a_2(X) = 1/\|X_2\|,$$

where $-\pi < \arg z \leqq \pi$, $\arg 0 = 0$, and let T be defined on $\bar{\mathfrak{M}}$ by $TX_j = a_j(X)X_j$, $j = 1, 2$. Then, Conditions (i) and (ii) imply

$$\begin{aligned} H(X) &\propto H(TX) \\ &\propto V(\|a_1(X)X_1\|, \|a_2(X)X_2\|, \langle a_1(X)X_1, a_2(X)X_2\rangle) \\ &\propto V(1, 1, |\langle X_1, X_2\rangle|/\|X_1\| \, \|X_2\|). \end{aligned}$$

Thus,

$$W(\rho) \propto V(1, 1, \rho),$$

and the theorem is established.

The class of functions which would make reasonable measures of linear relationship is much smaller than the one described, in that $V(\mathfrak{D})$ and $W(\mathfrak{I})$ should, at least, possess complete linear orders. However, such a requirement is not needed in the proof of the Theorem.

In order to obtain the analogous result for the coefficient of coherence we must augment our previous notation slightly to permit the discussion of a class of stochastic processes. As before, μ will be a fixed Lebesgue-Stieltjes measure.

Let $\mathfrak{X}_\mu$ be the class of all bivariate, weakly stationary stochastic processes whose spectral distributions are absolutely continuous with respect to μ. With the usual identification of Radon-Nikodym derivatives according to μ-equivalence, the class $\mathfrak{X}_\mu$ is in one-to-one correspondence with the set of spectral densities, $\mathbf{f}^X$, of processes $\mathbf{X} \in \mathfrak{X}_\mu$. Now, each $\mathbf{X} \in \mathfrak{X}_\mu$ determines a group of unitary operators, $\mathbf{U}^X = \{U_t^X \mid -\infty < t < \infty\}$, on is span, $\bar{\mathfrak{M}}^X$, and the classes $\mathbf{G}^X$ and $\mathbf{T}^X$ and the spaces $\mathfrak{M}_j^X$ and $\bar{\mathfrak{M}}_j^X$ can be defined as in Section 3.

Let $H(\mathbf{X})$ be a function defined for all $\mathbf{X} \in \mathfrak{X}_\mu$. As in the case of the correlation coefficient, in order that $H(\mathbf{X})$ qualify as an admissible measure of linear relationship for the component processes of $\mathbf{X}$, it will be required to satisfy two conditions. First, for each $\mathbf{X}$, $H(\mathbf{X})$ must depend on $\mathbf{X}$ through the intrinsic $\mathfrak{L}_2(P)$ inner product structure in the subspace $\bar{\mathfrak{M}}^X$. By Expression (3), this requirement will be met if, for each $\mathbf{X} \in \mathfrak{X}_\mu$, $H(\mathbf{X})$ is a function, $H(\mathbf{X}, \lambda)$, of the spectral density of $\mathbf{X}$. The precise form to be assumed for this dependence will be given shortly.

Second, $H(\mathbf{X})$ will be required to be invariant under the class of non-trivial operators in $\mathbf{T}^X$ which do not "intermix" the component processes of $\mathbf{X}$. That is, $H(\mathbf{X})$ is to be invariant for all T in the class

$$\mathbf{T}_0^X = \{T \in \mathbf{T}^X \mid T\mathfrak{M}_j^X \subseteq \bar{\mathfrak{M}}_j^X \quad \text{and} \quad \mathfrak{N}(T) \cap \bar{\mathfrak{M}}_j^X = \{0\}, j = 1, 2\}.$$

Physically, this means that the components of a bivariate time series can be independently passed through a series of possibly unknown linear filters and the linear relationship of the original components can be recovered by a measurement on the final time series. The first requirement implies that this measurement can be obtained from an analysis of the spectrum of the bivariate time series. This is an appealing property from a practical standpoint as time series analysts are generally well equipped to estimate power spectra.

The coefficient of coherence clearly satisfies the first requirement. The following lemma will be needed to establish that it also satisfies the second.

LEMMA 1. *Let* $\mathbf{X}$ *be an element of* $\mathfrak{X}_\mu$ *with spectral density* $\mathbf{f}^X = [f_{jk}^X]$. *If* $\mathbf{g} = [g_{jk}] \leftrightarrow T \,\varepsilon\, \mathbf{T}^X$, *then* $T \,\varepsilon\, \mathbf{T}_0^X$ *if and only if*

(i) $g_{12} = g_{21} = 0$ a.e. $(\mathbf{f}^X)$

(ii) $g_{jj} \neq 0$ a.e. (f_{jj}^X), $j = 1, 2$.

PROOF. We delete the superscript X in the proof. It will be shown that Condition (i) is equivalent to $T\mathfrak{M}_j \subseteq \overline{\mathfrak{M}}_j$, $j = 1, 2$, and Condition (ii) is equivalent to $\mathfrak{N}(T) \cap \overline{\mathfrak{M}}_j = \{0\}, j = 1, 2$.

1. Let $T \,\varepsilon\, \mathbf{T}$ be such that $T\mathfrak{M}_j \subseteq \overline{\mathfrak{M}}_j$, $j = 1, 2$, and let $T \leftrightarrow \mathbf{g}$. Let T_j be the restriction of T to $\mathfrak{M}_j$. Then by Theorem A_1 there exist functions, h_j, such that $\int |h_j|^2 f_{jj}\, d\mu < \infty$ and $T_jX_j(t) = \int \exp(it\lambda)\, h_j\, dZ_j$, $j = 1, 2$.

Let

$$\mathbf{g}' = \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix}.$$

Then $\operatorname{tr} \int \mathbf{g}'\mathbf{f}\mathbf{g}'^*\, d\mu = \sum_{j=1}^2 \int |h_j|^2 f_{jj}\, d\mu < \infty$, which implies $\mathbf{g}' \,\varepsilon\, \mathbf{G}$. Thus, there exists $T' \,\varepsilon\, \mathbf{T}$ such that $T' \leftrightarrow \mathbf{g}'$ and $T_j' = T_j$, $j = 1, 2$. But, by definition of $\mathfrak{M}$, every $T \,\varepsilon\, \mathbf{T}$ is uniquely determined on $\mathfrak{M}$ by its restrictions to $\mathfrak{M}_1$ and $\mathfrak{M}_2$, and uniquely determines $\mathbf{g}$ up to $\mathbf{f}$ equivalence by its values on $\mathfrak{M}$. Thus, $\mathbf{g} = \mathbf{g}'$. The converse result is obtained by an easy application of Theorems A and A_1.

2. Let R_j be the projection on $\overline{\mathfrak{M}}$ determined by setting $S = [g_{jj} = 0]$ in Corollary B, and let $\mathfrak{M}_j^0$ be the linear manifold generated by $\mathbf{X}_j^0 = R_j\mathbf{X}_j$, $j = 1, 2$. By Corollary B, $\|X_j^0(t)\|^2 = \int I_j f_{jj}\, d\mu$, where I_j is the set characteristic function of $[g_{jj} = 0]$. Thus, $\mathfrak{M}_j^0 = \overline{\mathfrak{M}}_j^0 = \{0\}$ if and only if $g_{jj} \neq 0$ a.e. (f_{jj}). The proof of the lemma will be completed by showing that

$$\mathfrak{M}_j^0 \subseteq \mathfrak{N}(T) \cap \overline{\mathfrak{M}}_j \subseteq \overline{\mathfrak{M}}_j^0 \quad \text{for} \quad j = 1, 2.$$

By Theorem A_1, $\mathfrak{M}_j^0 \subseteq \mathfrak{D}(T)$ and $\|TX_j^0(t)\|^2 = \int |g_{jj}|^2 I_j f_{jj}\, d\mu$. But, by definition of I_j, $|g_{jj}|^2 I_j f_{jj} = 0$ a.e. (μ). Thus $T\mathfrak{M}_j^0 = \{0\}$, which implies $\mathfrak{M}_j^0 \subseteq \mathfrak{N}(T)$. By the one dimensional version of Corollary B, $\mathbf{X}_j^0 \subseteq \overline{\mathfrak{M}}_j$. Thus, $\mathfrak{M}_j^0 \subseteq \mathfrak{N}(T) \cap \overline{\mathfrak{M}}_j$, $j = 1, 2$. Let $Y_j \,\varepsilon\, \mathfrak{N}(T) \cap \overline{\mathfrak{M}}$. Then, as indicated in Section 3 for the bivariate case, there exists a measurable, complex valued function y_j on $\mathfrak{R}$ such that $\int |y_j|^2 f_{jj}\, d\mu < \infty$ and $Y_j = \int y_j dZ_j$.

Now, $Y_j \,\varepsilon\, \mathfrak{N}(T)$ implies $TY_j = \int g_{jj} y_j dZ_j = 0$, where $T \leftrightarrow \mathbf{g} = [g_{jk}]$. But, $Y_j = R_jY_j + (1 - R_j)\, Y_j$, where R_j is the above defined projection. Thus, $T(1 - R_j)Y_j = 0$, and $\|T(1 - R_j)\, Y_j\|^2 = \int |g_{jj}|^2\, (1 - I_j)\, |y_j|^2 f_{jj}\, d\mu = 0$, which implies that $|g_{jj}|^2\, (1 - I_j)\, |y_j| = 0$ a.e. (f_{jj}). But $|g_{jj}|^2\, (1 - I_j) > 0$ on $[g_{jj} \neq 0]$ by definition of I_j, thus $y_j(1 - I_j) = 0$ a.e. (f_{jj}). This implies $(1 - R_j)\, Y_j = 0$; hence, $Y_j = R_jY_j \,\varepsilon\, \overline{\mathfrak{M}}_j^0$. From this we obtain the inclusion $\mathfrak{N}(T) \cap \overline{\mathfrak{M}}_j \subseteq \overline{\mathfrak{M}}_j^0$, $j = 1, 2$, and the lemma is proved.

THEOREM 3. *Let the coefficient of coherence for* $\mathbf{X} \,\varepsilon\, \mathfrak{X}_\mu$ *be denoted by* $\rho(\mathbf{X}, \lambda)$. *Then*

$$\rho(T\mathbf{X}, \lambda) = \rho(\mathbf{X}, \lambda) \quad \text{a.e. } (\mu) \quad \textit{for all} \quad T \,\varepsilon\, \mathbf{T}_0^X.$$

PROOF. Let $T \varepsilon \mathbf{T}_0^X$ and $\mathbf{g} \leftrightarrow T$. Then, by Lemma 1 and Theorem A_1 the spectral density of $T\mathbf{X}$ is

$$\mathbf{gfg}^* = \begin{bmatrix} |g_{11}|^2 f_{11} & g_{11}\bar{g}_{22} f_{12} \\ \bar{g}_{11} g_{22} f_{21} & |g_{22}|^2 f_{22} \end{bmatrix},$$

where $|g_{jj}| > 0$ a.e. (μ) on $[f_{jj} > 0], j = 1, 2$.

Thus,

$$\rho^2(T\mathbf{X}, \cdot) = |g_{11}\bar{g}_{22} f_{12}|^2 / |g_{11}|^2 f_{11} \, |g_{22}|^2 f_{22} = |f_{12}|^2 / f_{11} f_{22} = \rho^2(\mathbf{X}, \cdot)$$

a.e. (μ) on $\bigcap_{j=1}^2 [f_{jj} > 0]$.

On $\bigcup_{j=1}^2 [f_{jj} = 0]$, $\rho(T\mathbf{X}, \lambda) = 0 = \rho(\mathbf{X}, \lambda)$ a.e. (μ) independent of the values of g_{jj}.

We now establish the analog of Property 2, as given by Theorem 2, for the coefficient of coherence. In order to insure that the functions $H(\mathbf{X}, \cdot)$ will be measurable and to make allowance for the possible zeros of the spectral densities f_{jj}^X, we will modify the classes $\mathbf{V}$ and $\mathbf{W}$ defined above in the following manner.

Let $\mathbf{V}'$ be the class of all ordered triples $(V, \alpha, \mathcal{B}_V)$ such that $\mathcal{B}_V$ is a sigma field of subsets of $V(\mathfrak{D})$ satisfying the following conditions:

(i) $\mathcal{B}_V$ contains the elements of the partition of $V(\mathfrak{D})$ induced by α,

(ii) V is measurable with respect to $\mathcal{B}_V$ and the restriction of the Borel sets of complex Euclidean 3 space to $\mathfrak{D}$.

(iii) $V(0, 0, 0) \,\alpha\, V(1, 0, 0) \,\alpha\, V(0, 1, 0) \,\alpha\, V(1, 1, 0)$.

Let $\mathbf{W}'$ be the class of all ordered triples $(W, \beta, \mathcal{B}_W)$ such that $\mathcal{B}_W$ is a sigma field of subsets of $W(\mathcal{I})$ satisfying the conditions:

(i) $\mathcal{B}_W$ contains the elements of the partition of $W(\mathcal{I})$ induced by β,

(ii) W is measurable with respect to $\mathcal{B}_W$ and the restriction of the Borel sets of $\mathcal{R}$ to $\mathcal{I}$.

THEOREM 4. *If $H(\mathbf{X}, \lambda)$ is a function on $\mathfrak{X}_\mu \times \mathcal{R}$ such that there exists $(V, \alpha, \mathcal{B}_V) \,\varepsilon\, \mathbf{V}'$ which satisfy the conditions:*

(i) $H(\mathbf{X}, \lambda) \,\alpha\, V(f_{11}^X(\lambda), f_{22}^X(\lambda), f_{12}^X(\lambda))$ a.e. (μ)

and

(ii) $H(T\mathbf{X}, \lambda) \,\alpha\, H(\mathbf{X}, \lambda)$ a.e. (μ) *for all* $T \varepsilon \mathbf{T}_0^X$,

then there exists $(W, \beta, \mathcal{B}_W) \,\varepsilon\, \mathbf{W}'$ such that $W(\mathcal{I}) \subseteq V(\mathfrak{D})$, $\beta = \alpha$ restricted to $W(\mathcal{I})$, $\mathcal{B}_W \subseteq \mathcal{B}_V$ and

$$H(\mathbf{X}, \lambda) \,\alpha\, W(\rho(\mathbf{X}, \lambda)) \quad \text{a.e. } (\mu).$$

PROOF. Fix $\mathbf{X} \varepsilon \mathfrak{X}_\mu$. Define the bi-sequence of 2×2 matrix valued functions $\{\mathbf{g}_n | -\infty < n < \infty\}$ by

$$\mathbf{g}_n = \begin{bmatrix} h_{1,n} & 0 \\ 0 & h_{2,n} \end{bmatrix},$$

where

$$h_{1,n} = I_{1,n} \exp(-i \arg f_{12}^X)/[(f_{11}^X)^{\frac{1}{2}} + (1 - I_{1,n})] + (1 - I_{1,n}),$$

$$h_{2,n} = I_{2,n}/[(f_{22}^X)^{\frac{1}{2}} + (1 - I_{2,n})] + (1 - I_{2,n}), \quad \text{a.e. } (\mu),$$

and $I_{j,n}$ is the set characteristic function of $[f_{jj} > 0] \cap [n < \lambda \leqq n + 1]$. Now, it is easily verified that

$$\operatorname{tr} \int \mathbf{g}_n \mathbf{f}^X \mathbf{g}_n^* \, d\mu = \sum_{j=1}^{2} \int |h_{j,n}|^2 f_{jj}^X \, d\mu \leqq 2 + \operatorname{tr} \int \mathbf{f}^X \, d\mu < \infty.$$

Since $h_{j,n} \neq 0$ a.e. (μ), $\mathbf{g}_n \leftrightarrow T_n \varepsilon \mathbf{T}_0^X$, by Lemma 1. Thus, the application of Conditions (i) and (ii) of the theorem statement yields, by a straightforward computation,

$$H(\mathbf{X}, \cdot) \,\alpha\, H(T_n \mathbf{X}, \cdot) \,\alpha\, V(1, 1, |f_{12}^X|/(f_{11}^X f_{22}^X)^{\frac{1}{2}}), \quad \text{a.e. } (\mu)$$

on $\bigcap_{j=1}^2 [f_{jj}^X > 0] \cap [n < \lambda \leqq n + 1]$, for $n = 0, \pm 1, \pm 2, \cdots$. Thus

$$H(\mathbf{X}, \lambda) \,\alpha\, W(\rho(\mathbf{X}, \lambda)) \quad \text{a.e. } (\mu)$$

on $\bigcap_{j=1}^2 [f_{jj}^X > 0]$, where $W(\rho) \,\alpha\, V(1, 1, \rho)$.

Now, $\bigcup_{j=1}^2 [f_{jj}^X = 0] \subseteq [\rho(\mathbf{X}, \cdot) = 0]$. But Condition (iii) on the elements of $\mathbf{V}'$ guarantees that

$$H(\mathbf{X}, \lambda) \,\alpha\, W(0) \quad \text{a.e. } (\mu) \quad \text{on} \quad [\rho(\mathbf{X}, \cdot) = 0].$$

Thus, the theorem is proved.

6. Special cases. The class of possible dominating measures, μ, for a given stochastic process is determined by an extension of Lebesgue's decomposition theorem to matrix valued functions due to Cramér (1940, Theorem 2). We quote a version of this theorem due to Masani (1959).

THEOREM B. *Let* $\mathbf{F} = [F_{ij}]$ *be a* $q \times q$ *matrix valued function on* $[a, b]$ *which is monotone increasing,* i.e. $\mathbf{F}(\lambda) - \mathbf{F}(\lambda')$ *is non-negative definite and hermitian for* $\lambda > \lambda'$, *and let*

$$\mathbf{F}^{(a)} = [F_{ij}^{(a)}],\ \mathbf{F}^{(d)} = [F_{ij}^{(d)}],\ \mathbf{F}^{(s)} = [F_{ij}^{(s)}],$$

where $F_{ij}^{(a)}$, $F_{ij}^{(d)}$, $F_{ij}^{(s)}$ *are the absolutely continuous, discrete and singular parts of the function* F_{ij} *(which is necessarily of bounded variation). Then* $\mathbf{F}^{(a)}$, $\mathbf{F}^{(d)}$ *and* $\mathbf{F}^{(s)}$ *are themselves monotone increasing on* $[a, b]$.

This theorem with Corollary B provides an easy proof to a corresponding decomposition of q-variate, weakly stationary stochastic processes originally established by Rozanov (1958) in the case of discrete time. We give the proof for the bivariate case.

THEOREM 5. *Let* $\mathbf{X}$ *be a bivariate, weakly stationary stochastic process with spectral distribution function* $\mathbf{F}$, *and let* $\mathbf{F}^{(a)}$, $\mathbf{F}^{(d)}$ *and* $\mathbf{F}^{(s)}$ *be the components of* $\mathbf{F}$ *given in Theorem B. Then*

$$\mathbf{X} = \mathbf{X}^{(a)} + \mathbf{X}^{(d)} + \mathbf{X}^{(s)},$$

where $\mathbf{X}^{(x)}$ *is a bivariate, weakly stationary process with spectral distribution function* $\mathbf{F}^{(x)}$, $x = a, d, s$.

PROOF. Let $\mu_{ij}^{(x)}$ be the signed measure induced on $(\mathcal{R}, \mathcal{B})$ by the component $F_{ij}^{(x)}$ of $\mathbf{F}^{(x)}$, and let $\mu^{(x)} = \mu_{11}^{(x)} + \mu_{22}^{(x)}$, $x = a, d, s$. Then, it follows from Theorem

B that $\mu^{(x)}$ dominates $\mu_{ij}^{(x)}$ for $i, j = 1, 2$, and that there exists a measurable partition $\{S^{(a)}, S^{(d)}, S^{(s)}\}$ of $\mathfrak{R}$ such that $\mu^{(y)}(S^{(x)}) = 0$ for $x \neq y$.

Let $\mu = \mu^{(a)} + \mu^{(d)} + \mu^{(s)}$. Then $\mathbf{F}$ is absolutely continuous with respect to μ and possesses a μ-spectral density function $\mathbf{f}$. We are now in a situation where Theorem A applies to $\mathbf{X}$ and $\mathbf{f}$.

Two applications of Corollary B with $S = S^{(a)}$, and $S = S^{(d)}$ yield

$$\mathbf{X} = T^{(a)}\mathbf{X} + T^{(d)}\mathbf{X} + T^{(s)}\mathbf{X},$$

where $T^{(x)}$ is the projection determined by $S^{(x)}$. This is the desired decomposition with $\mathbf{X}^{(x)} = T^{(x)}\mathbf{X}$.

The spectral density of $\mathbf{X}^{(x)}$ with respect to μ, is $\mathbf{f}^{(x)} = I^{(x)}\mathbf{f}$, where $I^{(x)}$ is the set characteristic function of $S^{(x)}$. Thus, the spectral distribution function of $\mathbf{X}^{(x)}$ is

$$\mathbf{G}^{(x)}(\lambda) = \int_{-\infty}^{\lambda} I^{(x)}\mathbf{f}\, d\mu = \int_{-\infty}^{\lambda} I^{(x)}\, d\mathbf{F} = \int_{-\infty}^{\lambda} d\mathbf{F}^{(x)} = \mathbf{F}^{(x)}(\lambda),$$

since $\mu^{(y)}(S^{(x)}) = 0$ implies that $\int_a^b I^{(x)}\, d\mathbf{F}^{(y)} = 0$ for all $a \leqq b$. The theorem is proved.

Let μ_X denote the measure, defined in Theorem 5, which dominates the spectral distribution $\mathbf{F}^X$ of a bivariate process $\mathbf{X}$. It is easily seen that a Lebesgue-Stieltjes measure μ dominates $\mathbf{F}^X$ if and only if it dominates μ_X. Thus, the class of measures for which the μ-coefficient of coherence for $\mathbf{X}$ is defined, is $\mathbf{M}^X = \{\mu \mid \mu_X \ll \mu\}$.

We now show that the coefficient of coherence is, in a sense, independent of the particular measure $\mu \varepsilon \mathbf{M}^X$ chosen to define it.

THEOREM 6. *If* $\rho_\mu(\lambda)$ *denotes the* μ*-coefficient of coherence of* $\mathbf{X}$, *then*

$$\rho_\mu(\lambda) = \rho_{\mu_X}(\lambda) \quad \text{a.e. } (\mu_X)$$

for all $\mu \varepsilon \mathbf{M}^X$.

PROOF. Fix $\mu \varepsilon \mathbf{M}^X$. Let $\mathbf{f}^X$ be the spectral density of $\mathbf{F}^X$ with respect to μ_X, $\mathbf{f}$ the spectral density of $\mathbf{F}^X$ with respect to μ, and let $a(\lambda)$ be the Radon-Nikodym derivative of μ_X with respect to μ. Then, the chain rule for these derivatives yields

$$\mathbf{f}(\lambda) = a(\lambda)\mathbf{f}^X(\lambda) \quad \text{a.e. } (\mu).$$

Now, a straightforward argument paralleling that of the proof of Theorem 3 implies that, except for a μ_X null set, $[\rho_\mu \neq \rho_{\mu_X}] \subseteq A = [a = 0]$. But $\mu_X(A) = \int I_A\, a\, d\mu = 0$, where I_A is the set characteristic function of A. Thus, $\mu_X([\rho_\mu \neq \rho_{\mu_X}]) = 0$ as was to be shown.

The special cases of particular interest from the viewpoint of physical applications are the following:

1. $\mu_1 =$ Lebesgue measure. $\mathfrak{X}_{\mu_1}$ contains all processes of the form $\mathbf{X} = \mathbf{X}^{(a)}$ (Theorem 5).

2. μ_2 is point measure on a countable set S (i.e. $\mu_2(\{x\}) = 1$ for all $x \varepsilon S$). Then $\mathfrak{X}_{\mu_2}$ contains all processes $\mathbf{X} = \mathbf{X}^{(d)}$ such that if S_j^X is the set of discontinuity points of F_{jj}^X, $j = 1, 2$, then

$$S_1^X \cup S_2^X \subseteq S.$$

3. $\mu_3 = \mu_1 + \mu_2$. $\mathfrak{X}_{\mu_3}$ contains all processes $\mathbf{X} = \mathbf{X}^{(a)} + \mathbf{X}^{(d)}$ where $\mathbf{X}^{(a)} \varepsilon \mathfrak{X}_{\mu_1}$, $\mathbf{X}^{(d)} \varepsilon \mathfrak{X}_{\mu_2}$.

If ρ_j is the μ_j-coefficient of coherence of $\mathbf{X} \varepsilon \mathfrak{X}_{\mu_3}$, then it can be shown that

$$\rho_3(\lambda) = \rho_1(\lambda) + \rho_2(\lambda) \quad \text{a.e. } (\mu_X^{(a)} + \mu_X^{(d)}).$$

Several applications in which the coefficient of coherence occurs as an important parameter are given in (Amos and Koopmans. 1963).

Acknowledgment. I am grateful to Dr. M. J. Norris for many helpful suggestions for the improvement of this paper. I would also like to thank the referee for his careful scrutiny of the manuscript. The work was carried out under the auspices of the Atomic Energy Commission.

REFERENCES

AMOS, D. E. and KOOPMANS, L. H. (1963). Tables of the distribution of the coefficient of coherence for stationary bivariate Gaussian processes. *Sandia Corporation Monograph* SCR-483.

CRAMÉR, H. (1940). On the theory of stationary random processes. *Ann. of Math.* **41** 215–230.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

KOLMOGOROV, A. N. (1941). Stationary sequences in Hilbert space. *Bull. Math. Univ. Moscow* **2** No. 6 40 pp. (English translation by Natasha A. Brunswick, New York Univ.)

MASANI, P. (1959). Matrix-valued functions and the Wold decomposition. *Probability and Statistics*. (U. Grenander, ed.) 175–189.

RIESZ, F. and NAGY B. Sz. (1955). *Functional Analysis*. Unger, New York.

ROZANOV, YU A. (1958). Spectral theory of multi-dimensional stationary random processes with discrete time. *Uspehi Mat. Nauk* (*N. S.*) **13** No. 2 (80), 93–142. (English translation in *Selected Translations in Mathematical Statistics and Probability*. **1**, Amer. Math. Soc., Providence.)

WIENER, N. (1930). Generalized harmonic analysis. *Acta Mathematica* **55** 117–258.

ROSENBERG, M. (1963). The square integrability of matrix-valued functions with respect to non-negative hermetian measure. *Amer. Math. Soc. Notices*, **10** No. 3, Abstract #599-3.