Jaroslav Hájek

On linear statistical problems in stochastic processes

Czechoslovak Mathematical Journal, Vol. 12 (1962), No. 3, 404–444

Persistent URL: http://dml.cz/dmlcz/100528

Чехословацкий математический журнал, т. 12 (87) 1962, Прага

ON LINEAR STATISTICAL PROBLEMS IN STOCHASTIC PROCESSES

Jaroslav Hájek, Praha

(Received August 2, 1960)

A unified theoretical basis is developed for the solution of such problems as prediction and filtration (including the unbiased ones), estimation of regression parameters, establishing probability densities for Gaussian processes etc. The results are applied in deriving explicit solutions for stationary processes. The paper is a continuation of [17], [18], [19] and [20].

1. Introduction and summary. Linear statistical problems have been treated in very many papers. Most of them are referred to in extensive monographies [28] and [29]. No doubt, the topic has attracted so much attention because of its great practical and theoretical interest. The aim of the present paper is to contribute to developing a unified theory and to provide explicit results for some particular classes of stationary processes.

In Section 2 we introduce closed linear manifolds generated by random variables x_t and covariances R_{ts}, respectively, and study their interplay. We also discuss the abstract feature of a linear problem and enumerate some possible applications. In Section 4 we analyse conditions under which the solution of a linear problem may be interpreted in terms of individual trajectories (i.e. not only as a limit in the mean). Section 5 contains explicit solutions for a finite segment of a stationary process with a rational spectral density. With respect to previous papers [7], [8] and [9] on this topic, our results exhaust all possibilities, are more explicit, and we indicate when they may be interpreted in terms of individual trajectories. The last two sections are devoted to Gaussian processes. We define strong equivalency of normal (i.e. Gaussian) distributions of a stochastic process and study the determinant and quadratic form defining the probability density of a normal distribution with respect to another one, stongly equivalent to it. In Section 7 we present explicit probability densities for stationary processes with rational spectral densities.

2. Basic concepts and preliminary considerations. Let $\{x_t, t \in T\}$ be an arbitrary stochastic process with finite second moments. Suppose that the mean value $\mathsf{E}x_t$ vanishes, $t \in T$, and that the covariances of x_t and x_s equal $\mathsf{E}x_t\bar{x}_s = R(x_t, x_s) = R_{ts}$, $t \in T$, $s \in T$. Let $\mathscr{X}$ be the closed linear manifold of random variables consisting of finite linear combinations $\sum c_\nu x_{t_\nu}$, $t_\nu \in T$, and of their limits in the mean. Obviously, the covariance of any two random variables $x, y \in \mathscr{X}$, say $R(x, y)$, is uniquely determined by R_{ts}, $t \in T$, $s \in T$, and the mean value of any random variable $x \in \mathscr{X}$ equals 0. The variance of x will be denoted either by $R(x, x)$ or by $R^2(x)$. The closed linear manifold $\mathscr{X}$ is a Hilbert space with the norm $R(x)$ and inner product $R(x, y)$. As usually, random variables x, y such that $R(x - y) = 0$ are considered as identical.

Let U be the following mapping of $\mathscr{X}$ in the space of complex-valued functions, φ_t, $t \in T$:

(2.1) $$(Uv)_t = R(x_t, v) \quad (v \in \mathscr{X};\ t \in T),$$

where $R(x_t, v) = \varphi_t$ is considered as a function of $t \in T$. Let Φ be the set of functions φ such that $\varphi = Uv$ for some $v \in \mathscr{X}$, i.e. such that $\varphi_t = R(x_t, v)$. Obviously $Uv_1 = Uv_2$ implies $v_1 = v_2$, so that there is an inverse operator U^{-1}, $U^{-1}\varphi = v$, $\varphi \in \Phi$. If we introduce in Φ the norm $Q(\varphi) = R(U^{-1}\varphi)$ and the inner product

(2.2) $$Q(\psi, \varphi) = R(U^{-1}\varphi, U^{-1}\psi) \quad (\varphi, \psi \in \Phi),$$

the mapping U will be unitary (i.e. isometric and one-to-one) and Φ also will be a Hilbert space.

If $R^s = R_{ts}$ is considered as a function of t only, and s is fixed, we have $U^{-1}R_s = x_s$, and

(2.3) $$Q(R^s, R^t) = R_{ts} \quad (t, s \in T).$$

Obviously

(2.4) $$\varphi_t = Q(\varphi, R^t) \quad (t \in T),$$

so that $Q(\varphi^n) \to 0$ implies $\varphi_t^n \to 0$ in every point $t \in T$. Clearly, Φ consists of finite linear combinations $\sum c_\nu R^{t_\nu}$ and of their limits in the Q-norm.

Definition 2.1. We shall say that $\mathscr{X}$ and Φ, or, more explicitely, $(\mathscr{X}, R)$ and (Φ^x, Q^x), are closed linear manifolds generated by random variables x_t and covariances R_{ts}, respectively.

To any bounded linear operator A defined on $\mathscr{X}$ there corresponds a bounded linear operator $\bar{A}$ on Φ defined by the following relation:

(2.5) $$(\bar{A}\varphi)_t = R(Ax_t, U^{-1}\varphi) \quad (\varphi \in \Phi),$$

where U^{-1} is the inverse of the unitary mapping (2.1) of $\mathscr{X}$ on Φ. We may also write $\bar{A}\varphi = UA^*U^{-1}\varphi$, where A^* is the adjoint of A. Also conversly, to any linear operator $\bar{A}$ defined on Φ, there corresponds an operator A on $\mathscr{X}$ defined by the relation $Ax = U^{-1}\bar{A}^*Ux$. Obviously, $\overline{AB} = \bar{B}\bar{A}$. In what follows we shall omitt the bar so that any bounded linear operator A will be considered as defined on both spaces Φ and $\mathscr{X}$. We have to bear in mind, of course, that AB in $\mathscr{X}$ must be interpreted as BA in Φ, and vice versa.

Let $\mathscr{X}_0$ be the set of all finite linear combinations $\sum c_\nu x_{t_\nu}$. Let $\mathscr{X}^+$ be the closure of $\mathscr{X}_0$ with respect to a covariance R^+, and A be an operator in $\mathscr{X}^+$. Let R be another covariance. If

$$R(x, y) = R^+(Ax, Ay), \quad (x, y \in \mathscr{X}_0) \tag{2.6}$$

then

$$R(x) \leqq kR^+(x), \tag{2.7}$$

where $k = \|A\|$ and $\|A\|$ is the norm of A. Conversly, (2.7) implies existence of a bounded, positive and symmetric linear operator A in $\mathscr{X}^+$ such that (2.6) holds. Actually, if y is fixed, then $R(y, x)$ represents a linear functional in $\mathscr{X}^+$, and, consequently $R(y, x) = R^+(z, x)$, $x \in \mathscr{X}^+$. On putting $z = By$ and $A = B^{1/2}$, we get the needed result. Obviously

$$\|B\| = \sup_{x \in \mathscr{X}_0} \frac{R(x, x)}{R^+(x, x)}. \tag{2.8}$$

Definition 2.2. If (2.7) is satisfied, we say that the R-norm is dominated by the R^+-norm.

Lemma 2.2. *Let A be a bounded linear operator in the closed liner manifold $\mathscr{Z}$ generated by random variables z_t, $t \in T$, and let $\mathscr{X}$ be the closed linear manifold generated by random variables $x_t = Az_t$, $t \in T$. Let (Φ^z, Q^z) and (Φ^x, Q^x) be closed linear manifolds generated by covariances $R^+_{ts} = R(z_t, z_s)$ and $R_{ts} = R(x_t, x_s)$, respectively.*

Then $\Phi^x \subset \Phi^z$. Moreover, Φ^x consists of functions expressible in the form $\varphi = A\chi$, $\chi \in \Phi^z$, and

$$Q^x(\psi, \varphi) = Q^z(A_x^{-1}\psi, A_x^{-1}\varphi), \tag{2.9}$$

where the function, $A_x^{-1}\varphi$, $\varphi \in \Phi^x$, is uniquely determined by the conditions

$$A(A_x^{-1}\varphi)_t = \varphi_t \quad \text{and} \quad (A_x^{-1}\varphi)_t = R(z_t, v), \quad v \in \mathscr{X}. \tag{2.10}$$

Proof. If $\varphi \in \Phi^x$, then $\varphi_t = R(Az_t, v) = AR(z_t, v) = (A\chi)_t$; and conversly $AR(z_t, v) = R(x_t, v) \in \Phi^x$. If we suppose that $v \in \mathscr{X}$, then v is determined by $R(x_t, v)$ uniquely, and $(A_x^{-1}\varphi)_t = R(z_t, v)$ is a unique solution of (2.10). The proof is accomplished.

Remark 2.1. From Lemma 2.2 it follows that $\Phi^x = \Phi^z$ if R and R^+ dominate each other.

The values φ_t of every $\varphi \in \Phi$ may be considered as values of a linear functional $f(x) = R(x, U^{-1}\varphi)$ in the points $x = x_t$, $t \in T$. On applying the well-known extension theorem ([32], § 35), we get the following result.

Lemma 2.3. *A function φ_t belongs to Φ if and only if there exist a finite constant k such that for any linear combination,*

$$\left|\sum c_v \varphi_{t_v}\right| \leqq kR\left(\sum c_v x_{t_v}\right). \tag{2.11}$$

Example 2.1. Let the family $\{x_t, t \in T\}$ be finite, $\{x_t, t \in T\} = \{x_1, \ldots, x_n\}$, and let the matrix (R_{ij}), $R_{ij} = R(x_i, x_j)$, $1 \leqq i, j \leqq n$, be regular. Then Φ consists of all functions $\varphi = \varphi_i$, $1 \leqq i \leqq n$, and

$$Q(\psi, \varphi) = \sum\sum \psi_i \bar{\varphi}_j Q_{ij}, \tag{2.12}$$

where (Q_{ij}) is the inverse of $(R_{ij}, (Q_{ij}) = (R_{ij})^{-1}$. Moreover

$$U^{-1}\varphi = \sum\sum x_i \bar{\varphi}_j Q_{ij}. \tag{2.13}$$

Now, consider the $x_i's$ as functions of an elementary event $\omega \in \Omega$. Then, for a fixed ω, the sample sequence (trajectory) $x_1(\omega), \ldots, x_n(\omega)$, represents a function of i, $1 \leqq i \leqq n$. If we denote the latter function by x^ω, then (2.13) may be rewritten in a form dual to (2.1)

$$(U^{-1}\varphi)_\omega = Q(x^\omega, \varphi). \tag{2.14}$$

Proof. In view of (2.3) and (2.4), the formulas (2.12) and (2.13) are clearly true for $\varphi = R^t$ and $\psi = R^s$, $1 \leqq t, s \leqq n$, and the general case may be obtained on putting $\varphi = \sum_{\nu=1}^{n} c_\nu R^\nu$, $\psi = \sum_{\nu=1}^{n} d_\nu R^\nu$.

The determination of Φ and $Q(\psi, \varphi)$, and the solution of the equation $Uv = \varphi$ constitute what will be called a linear problem. If T is a finite set, then, as we have seen in Example 2.1, the linear problem is equivalent to inverting the covariance matrix.

Remark 2.2. If $\mathsf{E}x_t$ does not vanish, then $R(x, y) = \mathsf{E}xy - \mathsf{E}x\mathsf{E}y$, and, generally, $R(x) = 0$ is compactible with $\mathsf{E}x \neq 0$. Consequently, $R(x - y) = 0$ does not imply that $x = y$ with probability 1. This difficulty does not arise, if $\varphi_t = \mathsf{E}x_t$, $t \in T$, belongs to Φ, because then $\mathsf{E}x \neq 0$ only if $R(x) > 0$, in view of

$$\mathsf{E}x = R(x, U^{-1}\varphi). \tag{2.15}$$

So, if necessary, the above condition $\mathsf{E}x_t \equiv 0$ may be replaced by a more general condition $\mathsf{E}x_t \in \Phi$. If, and only if, $\mathsf{E}x_t \in \Phi$, $\mathsf{E}x$ represents a linear functional on $(\mathscr{X}, R)$.

Now, let us show in what kinds of applications the linear problems appear.

Application 2.1. Let y be a random variable not belonging to $\mathscr{X}$. The projection of y on $\mathscr{X}$, say Proj y, is given by

$$\text{Proj}\, y = U^{-1}R(x_t, y). \tag{2.16}$$

In particular circumstances, the projection is called prediction, interpolation, filtration, or a regression estimate.

Application 2.2. Let $\varphi_1, \ldots, \varphi_m$ be known linearly independent functions belonging to Φ, and let the mean value of the process depend on an unknown vector $\alpha = (\alpha_1, \ldots, \alpha_m)$, where $\alpha_1, \ldots, \alpha_m$ are arbitrary real or complex numbers, so that

$$\mathsf{E}_\alpha x_t = \sum_{j=1}^{m} \alpha_j \varphi_{jt}. \tag{2.17}$$

Let us choose linear estimates $\hat{\alpha}_j$ of the α_j's, which are best according to an arbitrary criterion with the following property: $\mathsf{E}_\alpha\hat{\Theta} = \mathsf{E}_\alpha\hat{\Theta}'$ for any α and $R(\hat{\Theta}) < R(\hat{\Theta}')$ imply that $\hat{\Theta}$ is a better estimate then $\hat{\Theta}'$ (recall that $R^2(.)$ denotes the variance). Then the vector $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_m)$, where the $\hat{\alpha}_j$'s are the best linear estimates of the α_j's, has the form

(2.18) $$\hat{\alpha} = Cv\,,$$

where $v = (U^{-1}\varphi_1, \ldots, U^{-1}\varphi_m)$ and C is a matrix which depends on other properties of the mentioned criterion. If we postulate that the estimate should be unbiassed and of minimum variance, then $C = B^{-1}$, where $B = (Q(\varphi_j, \varphi_k))$. If we postulate that the mean value of $\mathsf{E}_\alpha|\hat{\alpha}_j - \alpha_j|^2$ with respect to an apriori distribution of $\alpha_1, \ldots, \alpha_m$ should be minimum, then $C = (B + D^{-1})^{-1}$, where $D = (\mathsf{E}_{ap}\alpha_j\bar{\alpha}_k)$ and $\mathsf{E}_{ap}(.)$ denotes the apriori mean value. In the former case B^{-1} also represents the covariance matrix of the vector $(\hat{\alpha}_1, \ldots, \hat{\alpha}_m)$, and in the latter case $(B + D^{-1})^{-1}$ also represents the matrix with elements $\mathsf{E}_{ap}\mathsf{E}_\alpha(\hat{\alpha}_j - \alpha_j)\overline{(\hat{\alpha}_k - \alpha_k)}$. The main idea of the proof is as follows: From (2.15) it follows that $\mathsf{E}_\alpha x = \sum_{j=1}^{m} \alpha_j R(x, U^{-1}\varphi_j)$, so that projection of any estimate $\hat{\Theta}$ on the subspace $\mathscr{X}_m$ spanned by $U^{-1}\varphi, \ldots, U^{-1}\varphi_m$ has the same mean value as $\hat{\Theta}$ for any α, and, if $\hat{\Theta}$ does not belong to $\mathscr{X}_m$, has a smaller variance. See [20] and [28].

Application 2.3. Suppose that (2.17) still holds and that y is a random variable not belonging to $\mathscr{X}$ such that $\mathsf{E}_\alpha y = \sum_{j=1}^{m} \alpha_j c_j$ where c_j are known constants. Let us choose an estimate $\hat{y} \in \mathscr{X}$ of y according to a criterion with the following property: $\mathsf{E}_\alpha\hat{y} = \mathsf{E}_\alpha\hat{y}'$ for any α and $R(\hat{y} - y) < R(\hat{y}' - y)$ implies that $\hat{y}$ is a better estimate then $\hat{y}'$. (Obviously, this situation is a generalisation of one considered in Application 2.2.) Then the best linear estimate of y equals

(2.19) $$\hat{y} = y_0 + \sum_{j=1}^{m} \hat{\alpha}_j(c_j - Q(\varphi_j, \varphi_0))\,,$$

where the $\hat{\alpha}_j$'s have been defined in Application 2.2, and $y_0 = U^{-1}\varphi_0$, where $\varphi_{0t} = R(x_t, y)$, $t \in T$. If we postulate that $\mathsf{E}_\alpha\hat{y} = \mathsf{E}_\alpha y$ for any α and that $R(\hat{y} - y)$ should be minimum, then the best unbiassed linear estimate (2.19) has the following property:

(2.20) $$R^2(\hat{y} - y) = R^2(y_0 - y) + R^2\left(\sum_{j=1}^{m} \hat{\alpha}_j(c_j - Q(\varphi_j, \varphi_0))\right)$$

(see [20] and [28]).

Application 2.4. Let P and P^+ be two normal distributions of a real stochastic process $\{x_t, t \in T\}$ defined by a common covariance $R_{ts} = R_{ts}^+$ and mean values φ_t and $\varphi_t^+ \equiv 0$, respectively. If $\varphi \in \Phi$, then P is absolutely continuous with respect to P^+ and

(2.21) $$\frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\mathsf{P}^+} = \exp\{U^{-1}\varphi - \tfrac{1}{2}Q(\varphi, \varphi)\}\,.$$

If φ does not belong to Φ, then P and P$^+$ are perpendicular (mutually singular). The proof follows from the fact that $U^{-1}\varphi$ is a sufficient statistic for the pair (P, P$^+$), as is shown in [19]. The variance of $U^{-1}\varphi$ equals $Q(\varphi, \varphi)$ with respect to both P and P$^+$. In view of (2.16), the mean value of $U^{-1}\varphi$ equals $R(U^{-1}\varphi, U^{-1}\varphi) = Q(\varphi, \varphi)$, if P holds true, and, obviously, equals 0, if P$^+$ holds true. Now,

$$-\tfrac{1}{2}[U^{-1}\varphi - Q(\varphi, \varphi)]^2/Q(\varphi, \varphi) + \tfrac{1}{2}[U^{-1}\varphi]^2/Q(\varphi, \varphi) = U^{-1}\varphi - \tfrac{1}{2}Q(\varphi, \varphi),$$

which proves (2.21). The case of different covariances will be treated in Sections 6 and 7.

Application 2.5. A somewhat different application is given by the following Lemma: A difference of two covariances $R^+_{ts} - R_{ts}$ represents again a covariance, if and only if R^+ dominates R and the operator B determined by $R(x, y) = R^+(Bx, y)$ has the norm smaller or equal 1.

Proof: If $\|B\| \leqq 1$, then $I - B$ is a positive operator, so that $R^+_{ts} - R_{ts} = R((I - B)\, x_t, x_s)$ is a covariance. Conversly, if $R^+_{ts} - R_{ts}$ is a covariance, then $R^+(x, x) \geqq R(x, x)$ i.e. R^+ dominates R, and the norm of the operator B defined by $R(x, y) = R(Bx, y)$ is smaller or equal 1, in view of (2.8). Especielly, $R_{ts} - \varphi_t\bar{\varphi}_s$ is a covariance, if and only if $\varphi \in \Phi$ and $Q(\varphi) \leqq 1$. This corollary generalizes a result by A. V. Balakrishnan [21].

3. Stochastic integrals. Let $\mathcal{G}$ be a Borel field of measurable subsets Λ of T, and $\mu = \mu(\Lambda)$ be a σ-finite measure on $\mathcal{G}$. Let $Y_\Lambda = Y(\Lambda)$, be an additive random set function defined on subsets of finite measure and such that $\mathsf{E}Y_\Lambda = 0$ and

$$R(Y_\Lambda, Y_{\Lambda'}) = \mu(\Lambda \cap \Lambda') \quad (\Lambda, \Lambda' \in \mathcal{G}). \tag{3.1}$$

The closed linear manifold $\mathcal{Y}$ generated by random variables Y_Λ consits of random variables expresible in the form

$$v = \int_T h_t\, Y(\mathrm{d}t), \tag{3.2}$$

where h_t is a quadratically integrable function, $h \in \mathcal{L}^2(\mu)$. The stochastic integral (3.2) is defined as a limit in the mean (see J. L. Doob [33]).

The closed linear manifold generated by covariances (3.1) consists of all σ-additive set functions $v(\Lambda)$ such that $v(\lambda) = \int_\Lambda f\, \mathrm{d}\mu$ and

$$\textstyle\int_T |f|^2\, \mathrm{d}\mu < \infty. \tag{3.3}$$

We put $f = \mathrm{d}v/\mathrm{d}\mu$ and denote $\mu(\mathrm{d}t)$ briefly by $\mathrm{d}\mu$. Moreover, we have

$$Q(v_1, v_2) = \int_T \left(\frac{\mathrm{d}v_1}{\mathrm{d}\mu}\right)\overline{\left(\frac{\mathrm{d}v_2}{\mathrm{d}\mu}\right)}\mathrm{d}\mu, \tag{3.4}$$

and the equation $v(\Lambda) = R(Y_\Lambda, v)$ is solved by (3.2), where $h = \mathrm{d}v/\mathrm{d}\mu$.

It often is preferable to introduce the formal derivative $y_t = (\mathrm{d}Y/\mathrm{d}\mu)_t$ (so called white noise) and to write $\int h_t y_t \mu(\mathrm{d}t)$ and $R(y_t, v)$ instead of $\int h_t\, Y(\mathrm{d}t)$ and $[(\mathrm{d}/\mathrm{d}\mu)$

$R(Y_A, v)]$, respectively. Unless the point t has a positive measure, y_t itself has no meaning, but the integrals $\int hy\, d\mu$, $h \in \mathscr{L}^2(\mu)$, and covariances $R(y_t, v)$, $t \in T$, are well-defined in the above sense. We confine ourselves to this remark without entering into the theory of random distributions [12].

$\mathscr{L}^2(\mu)$, with the usual inner product $(h, g) = \int h\bar{g}\, d\mu$, may be considered as the closed linear manifold generated by the covariance function of the white noise y_t (formally, $R(y_t, y_s) = 0$ if $t \neq s$, and $R(y_t, y_t) = 1/\mu(dt)$). The relations

$$R(y_t, v) = h_t \quad \text{and} \quad v = \int \bar{h}_t y_t\, \mu(dt) \tag{3.5}$$

define a unitary transformation $Uv = h\, (U^{-1}h = v)$ of $\mathscr{Y}$ on $\mathscr{L}^2(\mu)$ (of $\mathscr{L}^2(\mu)$ on $\mathscr{Y}$).

Now consider an operator K in $\mathscr{L}^2(\mu)$, generated by a kernel $K(t, s)$, which is measurable on $\mathscr{G} \times \mathscr{G}$ and quadratically integrable w. r. t. $\mu \times \mu$. We have

$$(Kh)_t = \int K(t, s)\, h_s\, \mu(ds)\,. \tag{3.6}$$

This operator may be carried over to $\mathscr{Y}$ according to the formula $Kv = U^{-1}K^*Uv$, where K^* is generated by $K^*(t, s) = \overline{K(s, t)}$. So, if $v = \int \bar{h} y\, d\mu$, then

$$Kv = \int (K^*h)\, y\, d\mu\,. \tag{3.7}$$

The domain of definition of the operator K in $\mathscr{Y}$ may be extended to include the white noise y_t by puting

$$x_t = Ky_t = \int K(t, s)\, y_s\, \mu(ds)\,. \tag{3.8}$$

Then we also may write,

$$Kv = \int \bar{h}(Ky)\, d\mu = \int \bar{h}_t x_t\, d\mu\,, \tag{3.9}$$

where the integral is defined in the weak sense (see [3]), i.e. as a random variable ξ belonging to the closed linear manifold $\mathscr{X}$ spanned by random variables $x_t = Ky_t$, $t \in T$, and such that

$$R(y_\tau, \xi) = \int R(y_\tau, x_t)\, h_t\, d\mu\,.$$

Actually, we have $R(y_\tau, x_t) = \overline{K(t, \tau)} = K^*(\tau, t)$, so that

$$\int R(y_\tau, x_t)\, h_t\, d\mu = (K^*h)_\tau = R(y_\tau, Kv)\,,$$

where the last equality follows from (3.7).

The equality

$$\int \overline{(K^*h)}\, y\, d\mu = \int \bar{h}(Ky)\, d\mu\,, \tag{3.10}$$

obtained from (3.7) and (3.9), will now be generalized for arbitrary bounded operators. Let $X(\Lambda) = AY(\Lambda)$, $\Lambda \in \mathscr{G}$, where $Y(\Lambda)$ is an additive random set function satisfying (3.1) and A is a bounded operator defined in the closed linear manifold $\mathscr{Y}$ generated by $\{Y(\Lambda), \Lambda \in \mathscr{G}\}$. Put, by definition,

$$\int g(t)\, X(\mathrm{d}t) = A \int g(t)\, Y(\mathrm{d}t) \tag{3.11}$$

for any $g \in \mathscr{L}_2(\mu)$. Then

$$\int g(t)\, AY(\mathrm{d}t) = \int (A^* g)_t\, Y(\mathrm{d}t)\,. \tag{3.12}$$

In fact, (3.12) is a direct consequence of the identity $A = U^{-1}A^*U$, where U and U^{-1} are unitary operators defined by (2.5).

Application 3.1. Let $\{x_t, t \in T\}$ be the process expressed by (3.8). Then 1° the closed linear manifold Φ^x generated by covariances $R(x_t, x_s) = R(Ky_t, Ky_s)$, $t \in T$, $s \in T$, consists of functions expressible in the form $\varphi_t = (Kh)_t$, $h \in \mathscr{L}^2(\mu)$, 2° the inner product in Φ^x, say Q^x, is given by

$$Q^x(\psi, \varphi) = \int (K_x^{-1}\psi)_t\, \overline{(K_x^{-1}\varphi)_t}\, \mathrm{d}\mu\,, \tag{3.13}$$

and 3° the equation $\varphi_t = R(x_t, v)$ is solved by

$$v^\varphi = \int (K_x^{-1}\varphi)_t\, Y(\mathrm{d}t)\,. \tag{3.14}$$

The formula (3.14) is based on Lemma 2.2, where $K_x^{-1}\varphi$ is defined as a function such that $KK_x^{-1}\varphi = \varphi$ and $(K_x^{-1}\varphi)_t = R(y_t, x)$, where x belongs to the closed linear manifold $\mathscr{X}$ generated by random variables x_t, $t \in T$. If $\mathscr{X} = \mathscr{Y}$ then $K_x^{-1} = K^{-1}$ is an ordinary inverse of K. Actually, then $0 \equiv \varphi_t = R(x_t, v)$ implies $v \equiv 0$, and hence $(Kh)_t \equiv 0$ if and only if $h_t \equiv 0$.

Application 3.2. The operator K^* is generated by the kernel $K^*(t, s) = \overline{K(s, t)}$, and the operator KK^* by the kernel

$$R(x_t, x_s) = \int K(t, \tau)\, \overline{K(s, \tau)}\, \mu(\mathrm{d}\tau)\,. \tag{3.15}$$

The equality of both sides in (3.15) follows from (3.8). If the function φ is expressible in the form $\varphi = KK^*h$, $h \in \mathscr{L}^2(\mu)$, i.e.

$$\varphi_t = \int R(x_t, x_s)\, h_s\, \mu(\mathrm{d}s)\,, \tag{3.16}$$

then the equation $\varphi_t = R(x_t, v)$ is solved by

$$v = \int \bar{h}_t x_t\, \mathrm{d}\mu = \int \overline{((KK^*)^{-1}\varphi)_t}\, x_t\, \mathrm{d}\mu\,. \tag{3.17}$$

Actually, $v \in \mathscr{X}$ and $\varphi_t = R(x_t, v)$ follows directly from (3.16) and the week definition of the integral $\int \bar{h}_t x_t\, \mathrm{d}\mu$.

Example 3.1. Let $\{x_t, t \leqq t_0\}$ be a semi-infinite segment of a regular stationary process. As is well-known (J. L. Doob [30], Chap. XII, § 5), there exist a uniquely determined quadratically integrable function $c(\tau)$ vanishing for $\tau < 0$, and a process with uncorrelated increments Y_t, $E\,|\mathrm{d}Y_t|^2 = \mathrm{d}t$, such that $x_t = \int_{-\infty}^{t} c(t-s)\,\mathrm{d}Y_s$, and the closed linear manifolds $\mathscr{X}$ and $\mathscr{Y}$ spanned by

$$\{x_t, t \leqq t_0\} \quad \text{and} \quad \left\{y_t = \frac{\mathrm{d}Y_t}{\mathrm{d}t}, \quad t \leqq t_0\right\}$$

respectively, coincide, $\mathscr{X} = \mathscr{Y}$. So we may apply the preceeding results with $K(t, s) = c(t-s)$, $\mu(\mathrm{d}t) = \mathrm{d}t$ and $T = (-\infty, t_0]$. In the present case equation (3.16) is called the Wiener-Hopf equation. The functions given by (3.16), however, do not exhaust the whole set Φ^x of functions φ, for which the linerar problem $\varphi_t = R(x_t, v)$ may be solved.

For a moment, let us consider the whole process x_t, $-\infty < t < \infty$, and denote the usual Fourier-Plancherel transform by F. We know that $F^* = F^{-1}$, which implies, in view of (3.12), that $x_t = \int [Fc(t - .)]_\lambda\,\mathrm{d}(FY)_\lambda$, where, as is well-known, $[Fc(t - .)]_\lambda = e^{it\lambda}a(\lambda)$. On puting $\mathrm{d}Z_\lambda = a(\lambda)\,\mathrm{d}(FY)_\lambda$, we get the well-known spectral representation $x_t = \int e^{it\lambda}\,\mathrm{d}Z_\lambda$, where $\mathsf{E}\,|\mathrm{d}Z_\lambda|^2 = |a(\lambda)|^2\,\mathrm{d}\lambda$, $|a(\lambda)|^2$ being the spectral density.

Example 3.2. Let Y_t be a process with uncorrelated increments such that $E\,|\mathrm{d}Y_t|^2 = \mathrm{d}t$, $-1 \leqq t \leqq t_0$. We are interested in $x_t = Y_t - Y_{t-1}$, $0 \leqq t \leqq t_0$, which is a stationary process with correlation function $R(x_t, x_{t-\tau}) = \max(0, 1 - |\tau|)$. If we add the random variables $x_t = Y_t - Y_{t-1}$, $-1 \leqq t \leqq 0$, we obtain $\mathscr{X} = \mathscr{Y}$. The kernel $K(t, s)$ is apparent from relations

$$\text{(3.18)} \qquad \begin{aligned} x_t &= \int_{-1}^{t} y_s\,\mathrm{d}s\,, \quad \text{if} \quad -1 \leqq t \leqq 0\,, \\ &= \int_{t-1}^{t} y_s\,\mathrm{d}s\,, \quad \text{if} \quad 0 \leqq t \leqq t_0\,. \end{aligned}$$

Let $K_0^{-1}\varphi$ be a function such that $(KK_0^{-1}\varphi)_t = \varphi_t$, $0 \leqq t \leqq t_0$, and that $(K_0^{-1}\varphi)_t = R(y_t, v)$, where v belongs to the closed linear manifold $\mathscr{X}_0$ generated by random variables $\{x_t, 0 \leqq t \leqq t_0\}$. We may find, after some computations, that

$$\text{(3.19)} \qquad (K_0^{-1}\varphi)_t = \varphi'_t + \varphi'_{t-1} + \varphi'_{t-2} + \dots\,,$$

where the sum stops when the index falls into the interval $[-1, 0)$ and φ'_i has been extended so that

(3.20)

$$\begin{aligned} \varphi'_t &= \frac{1}{N+1}\left(c - N\varphi'_{t+1} - (N-1)\,\varphi'_{t+2} - \dots - \varphi'_{t+N}\right) \quad \text{for} \quad t_0 - N - 1 < t < 0\,, \\ &= \frac{1}{N+2}\left(c - (N+1)\,\varphi'_{t+1} - N\varphi'_{t+2} - \dots - \varphi'_{t+N+1}\right) \\ &\qquad \text{for} \quad -1 < t < t_0 - N - 1\,, \end{aligned}$$

where N is the greatest integer not exceeding t_0 and

$$(3.21) \qquad c = \frac{(N+1)(\varphi_0 + \varphi_{t_0}) + \ldots + (\varphi_N + \varphi_{t_0 - N})}{2N - t_0 + 2} .$$

For $\varphi_t \equiv 1$ we could obtain the best linear estimate [15] of a constant mean value.

For $\varphi_t = R(x_{t_0+\tau}, x_t) = \max(0, 1 - |t_0 + \tau - t|)$ and $t_0 = N$,

we would obtain the best predictor [16].

The same method could be applied to processes expressed generally by $x_t = \sum_{k=0}^{m} b_{m-k} Y_{t-k}$. In all cases Φ^x consists of absolutely continuous functions with a quadratically integrable derivative.

4. Solutions expressed in terms of individual trajectories. In what follows we shall assume that the random variables $x_t = x_t(\omega)$ are defined on a probability space $(\Omega, \mathscr{F}, \mathsf{P})$. The covariances and finite-dimensional distributions of random variables $x_t(\omega)$ are not influenced by any adjustment (modification) of every $x_t(\omega)$ on a ω-subset having probability 0. The properties of individual trajectories $x^\omega = x_t(\omega)$, $t \in T$, however, may be changed very substantially.

Theorem 4.1. *The process $x_t = x_t(\omega)$ defined by (3.8) may be adjusted on ω-subsets having probability zero, so that the solution of the equation $\varphi_t = R(x_t, v)$, φ_t expressible as $\varphi_t = \int R(x_t, x_s)\, h_s\, \mathrm{d}\mu$, $h \in \mathscr{L}^2(\mu)$, is given by*

$$(4.1) \qquad v(\omega) = \int_T \bar{h}_t x_t(\omega)\, \mathrm{d}\mu ,$$

where for almost every ω the integral on the right side is to be understood in the usual Lebesgue-Stieltjes sense.

Proof. The compactness of KK^* in $\mathscr{L}^2(\mu)$ follows from (3.6) by standard arguments. Let $\chi_n(t)$ and $\kappa_n (n \geqq 1)$ be the eigen-elements and corresponding eigen-values of KK^*, respectively. The equality of some κ_n's is not excluded. We first show that $\sum \kappa_n < \infty$. As the system $\{\chi_n(t),\ n \geqq 1\}$ is complete and orthonormal, we have

$$(4.2) \qquad \sum_1^\infty \kappa_n = \sum_1^\infty \int |K^*\chi_n|^2\, \mathrm{d}\mu = \int \sum_1^\infty \left| \int K^*(t,s) \chi_n(s)\, \mu(\mathrm{d}s) \right|^2 \mu(\mathrm{d}t) =$$
$$= \iint |K(s,t)|^2\, \mu(\mathrm{d}s)\, \mu(\mathrm{d}t) ,$$

where the last expression is finite according to our assumptions.

Introduce the unitary transforms U and U^{-1} defined by $(Uv)_t = R(y_t, v)$. Obviously, the random variables $v_n = U^{-1}\chi_n$ form a complete orthonormal system in $\mathscr{Y}$. Moreover,

$$R(Kv_m, Kv_n) = R(v_m, K^*Kv_n) = Q(KK^*\chi_n, \chi_m) = \kappa_n Q(\chi_n, \chi_m) ,$$

so that $\{Kv_n, n \geqq 1\}$ is a orthogonal system with $R(Kv_n, Kv_n) = \kappa_n, n \geqq 1$. In view of (3.9), $Kv_n \in \mathscr{X}$, where $\mathscr{X}$ is the closed linear manifold spanned by random variables $x_t = Ky_t, t \in T$. If $R(y, Kv_n) = 0, n \geqq 1$, then $R(K^*y, v_n) = 0, n \geqq 1$, so that $K^*y \equiv 0$ and $(KUy)_t = (UK^*y)_t \equiv 0$. Consequently,

$$R(x_t, y) = \int K(t, s)(Uy)_s \, d\mu = (KUy)_t = 0,$$

$t \in T$, which shows that $y \perp Kv_n, n \geqq 1$, only if $y \perp \mathscr{X}$. We conclude that $\{Kv_n, n \geqq 1\}$, where Kv_n with $\kappa_n = 0$ may be omitted, forms complete system in $\mathscr{X}$.

Consequently, for every $t \in T$, we have

(4.3) $$x_t = \sum_1^\infty Kv_n \frac{R(x_t, Kv_n)}{\kappa_n} = \sum_1^\infty Kv_n \frac{R(y_t, K^*Kv_n)}{\kappa_n} = \sum_1^\infty Kv_n \cdot \chi_n(t),$$

where the sum converges in the mean. Further, the sequence $\{\sum_1^N \chi_n(t)\, Kv_n(\omega), N \geqq 1\}$ converges in the $(\mu \times \mathsf{P})$ – mean on $T \times \Omega$ because

(4.4) $$\int_T \int_\Omega |\sum_{N+1}^{N+p} \chi_n(t)\, Kv_n(\omega)|^2 \, d\mathsf{P}\, d\mu = \sum_{N+1}^{N+p} \kappa_n$$

and $\sum_1^\infty \kappa_n < \infty$. Consequently, we may draw a subsequence such that

(4.5) $$\lim_{k \to \infty} \sum_1^{N_k} \chi_n(t)\, Kv_n(\omega) = \bar{x}_t(\omega)$$

exist for almost all (t, ω) with respect to $\mu \times \mathsf{P}$. Let T_0 be a subset of T such that $\mu(T - T_0) = 0$ and that for $t \in T_0$ the limit (4.5) exists with probability one. On puting

(4.6) $$\begin{aligned} \tilde{x}_t(\omega) &= \bar{x}_t(\omega), \quad \text{for} \quad t \in T_0, \\ &= x_t(\omega), \quad \text{for} \quad t \in T - T_0, \end{aligned}$$

and recalling that the limits (4.3) and (4.5) coincide with probability 1 for $t \in T_0$, we conclude that $\{\tilde{x}_t(\omega), t \in T\}$ is an equivalent modification of the process $\{x_t(\omega), t \in T\}$.

Now $\int(\sum_1^\infty |Kv_n|^2)\, d\mathsf{P} = \sum_1^\infty \kappa_n < \infty$, so that

(4.7) $$\sum_1^\infty |Kv_n(\omega)|^2 < \infty$$

with probability one. Without any loss of generality, we may assume that (4.7) holds true for any $\omega \in \Omega$. Then $\sum_1^\infty \chi_n(t)\, Kv_n(\omega)$ for every fixed ω represents an orthonormal expansion of $\tilde{x}_t(\omega)$ in $\mathscr{L}^2(\mu)$. So, if $v(\omega)$ is given by (4.1), where $x_t(\omega) \equiv \tilde{x}_t(\omega)$, we get

(4.8) $$\begin{aligned} v(\omega) = \sum_1^\infty Kv_n(\omega) \int_T \bar{h}_t \chi_n(t)\, d\mu = \sum_1^\infty \kappa_n^{-1} Kv_n(\omega)\, . \\ . \int_T \bar{h}_t (KK^*\chi_n)_t \, d\mu = \sum_1^\infty \kappa_n^{-1} Kv_n(\omega) \int_T \bar{\varphi}_t \chi_n(t)\, d\mu \end{aligned}$$

and, in view of (4.3),

$$R(x_t, v) = \sum_1^\infty \chi_n(t) \int_T \bar{\varphi}_t \chi_n(t) \, \mathrm{d}\mu = \varphi_t , \tag{4.9}$$

which concludes the proof.

Remark 4.1. Let x_{t_1}, $t_1 \in T_1$, be arbitrary random variables from the closed linear manifold $\mathscr{X}$ spanned by random variables x_t, $t \in T$. Clearly, the extended process $\{x_t, t \in T \cup T_1\}$ spans the same closed linear manifold as the process $\{x_t, t \in T\}$. However, the class of random variables v expressible in the form

$$v = \int_{T \cup T_1} \bar{h}_t x_t \, \mathrm{d}\mu$$

never is smaller that the class of random variables $v = \int_T \bar{h}_t x_t \, \mathrm{d}\mu$, and usually will be larger. By this device, we may enlarge considerably the class of functions φ expressibles as $\varphi_t = \int R(x_t, x_s) \, \mu(\mathrm{d}s)$, permitting the solution to be written in the form (4.1) For example, if the process is originaly defined on a segment $0 \leqq t \leqq T$, the added random variables may be derivatives in the endpoints (see the next section). In the extreme case we could include in the set $\{x_t, t \in T \cup T_1\}$ every random variable from $\mathscr{X}$, i.e. put $\{x_t, t \in T \cup T_1\} = \mathscr{X}$.

Another way to enlarge the class of functions φ for which the solution is expressible in terms of individual trajectories consists in replacing the white noise y_t by an ordinary random process z_t, $t \in T$. We shall suppose that z is the closed linear manofold spanned by the z_t's and K is a compact linear operator in $\mathscr{Z}$. Let Φ^z, Φ^x and Φ° be closed linear manifolds generated by covariances $R(z_t, z_s)$, $R(Kz_t, Kz_s)$ and $R(K^*Kz_t, K^*Kz_s)$, respectively. From Lemma 2.2. it follows that

$$\Phi^\circ \subset \Phi^x \subset \Phi^z . \tag{4.10}$$

Let Q^z and Q^x be the inner product in Φ^z and Φ^x, respectively.

Lemma 4.1. *If $\varphi \in \Phi^\circ$ and $\psi \in \Phi^x$, then*

$$Q^x(\psi, \varphi) = Q^z(\psi, (KK^*)^{-1} \varphi) , \tag{4.11}$$

where $(KK^*)^{-1} \varphi$ is any function such that $KK^*(KK^*)^{-1} \varphi = \varphi$.

Proof. Clearly, $Q^z(\psi, (KK^*)^{-1}\varphi) = Q(K^{-1}\psi, K^*(KK^*)^{-1}\varphi)$, where $K^{-1}\psi$ is any solution if the equation $K\chi = \psi$. In particular, we may take the solution $K_x^{-1}\psi$ considered in Lemma 2.2. Now if $(KK^*)^{-1}\varphi = R(z_t, v)$, then $K^*(KK^*)^{-1}\varphi = R(z_t, Kv)$, where Kv belongs to the closed linear manifold $\mathscr{X}$ spanned by the random variables $x_t = Kz_t$, $t \in T$. Consequently, $K^*(KK^*)^{-1}\varphi = K_x^{-1}\varphi$ where K_x^{-1} is defined by (2.10). So, we have $Q^z(\psi, (KK^*)^{-1}\varphi) = Q^z(K_x^{-1}\psi, K_x^{-1}\varphi)$, which gives (4.11), in accordance with (2.9). The proof is finished.

Note that the right side of (4.11) is an extention of Q^x, originally defined on $\Phi^x \times \Phi^x$, on $\Phi^z \times \Phi^\circ$. Also observe that in Lemma 2.2. we need a particular choice of K^{-1}, $K^{-1} = K_x^{-1}$, whereas $(KK^*)^{-1}$ may be chosen in any way, if non-unique.

Now we shall extend the formula (4.1) to the case when the white noise y_t is replaced by an ordinary process z_t, $t \in T$.

Theorem 4.2. *Suppose that the above-considered compact operator K in $\mathcal{X}$ is such that the eigen-values κ_n of K^*K form a convergent series, $\sum_1^\infty \kappa_n < \infty$. Then the process $x_t = Kz_t$, $t \in T$, may be adjusted so that almost every trajectory $x^\omega = x_t(\omega)$ belongs to Φ^z, and the solution of the equation $\varphi_t = R(x_t, v)$, $\varphi \in \Phi^0$, equals*

$$v(\omega) = Q^z(x^\omega, (KK^*)^{-1}\varphi) . \tag{4.12}$$

Proof. We may proceed quite similarly as in proving Theorem 4.1, with the only exception that the limit

$$\sum_1^\infty Kv_n(\omega)\, \chi_n(t) = \sum_1^\infty Kv_n(\omega)\, R(z_t, v_n) = R(z_t, \sum_1^\infty v_n Kv_n(\omega)) = \bar{x}_t(\omega) \tag{4.13}$$

exist for every t and each ω satisfying (4.7). Other arguments need no changes.

The formula (4.12) presumes that almost every trajectory $x^\omega = x_t(\omega)$ belongs to Φ^z. The following theorem shows that in the Gaussian case x^ω cannot belong to Φ^z almost surely, unless the operator K has the property assumed in Theorem 4.2, so that this theorem cannot be strengthened.

Theorem 4.3. *Let $x_t(\omega)$ and $z_t(\omega)$ be Gaussian processes related by a bounded linear operator K, $x_t = kz_t$, $t \in T$, and Φ^z be the closed linear manifold generated by covariances $R(z_t, z_s)$. Then, for an equivalent modification of the process $x_t(\omega)$,*

$$P(x^\omega \in \Phi^z) = 1 , \tag{4.14}$$

if and only if the operator K^*K is compact and the series of its eigen-values κ_n is convergent, $\sum_1^\infty \kappa_n < \infty$.

Proof. Sufficiency follows from Theorem 4.2. Necessity. If (4.14) holds true, then there exist a function $w(\omega, \omega')$ representing a random variable from $\mathcal{X}$ for ω fixed, and such that

$$x_t(\omega) = \int z_t(\omega')\, w(\omega, \omega')\, P(d\omega') = Kz_t(\omega) . \tag{4.15}$$

If $\{z_n\}$ converges weakly to 0, then

$$\lim_{n\to\infty} x_n(\omega) = \lim_{n\to\infty} \int z_n(\omega')\, w(\omega, \omega')\, P(d\omega') = 0$$

for almost every ω, and, concequently, $x_n \to 0$ in probability. If the $x_n's$ are Gaussian, then

$$P(|x_n| \geqq \varepsilon) \geqq 2 \int_{\varepsilon/R(x_n)}^\infty (2\pi)^{-1/2} e^{-\frac{1}{2}\tau^2}\, d\tau ,$$

so that convergence in probability implies convergence in the mean, $R(x_n) \to 0$. It means that operator K transforms weakly convergent sequences in strongly convergent ones, and consequently is compact (see [32], § 85). Obviously, the operator K^*K also is compact. Let κ_n be the non-zero eigen-values of K^*K, and let $v_n(\omega)$ be the corresponding eigen-elements. We have

$$(4.16) \qquad \sum_1^\infty \kappa_n = \sum_1^\infty R(Kv_n, Kv_n) = \mathsf{E}\{\sum_1^\infty |Kv_n(\omega)|^2\}$$

and, for almost all ω,

$$(4.17) \qquad \sum_1^\infty |Kv_n(\omega)|^2 = \sum_1^\infty |R(v_n, w^\omega)|^2 = R(w^\omega, w^\omega) < \infty ,$$

where $w^\omega = w(\omega, .)$. The random variables Kv_n are uncorrelated, because $R(Kv_n, Kv_m) = R(K^*Kv_n, v_m) = \kappa_n R(v_n, v_m)$. If we suppose that they are Gaussian, then they are independent. It remains to show that finiteness of (4.17) for almost all ω implies finiteness of (4.16), or, equivalently, that infiniteness of (4.16) would imply infiniteness of (4.17) with positive probability. If $\sum_1^\infty \kappa_n = \infty$, we may form an infinite sequence of partial sums

$$\xi_k = \sum_{n_k+1}^{n_{k+1}} Kv_n$$

such that each partial sum consist either (i) of a unique elements Kv_n such that $\kappa_n \geqq 1$, or (ii) of several elements such that $\kappa_n \leqq 1$ and $\sum_{n_k+1}^{n_k} \kappa_n \geqq 4$. The ξ_k's have mean values $\mathsf{E}\xi_k = \sum_{n_k+1}^{n_{k+1}} \kappa_n$ and variances $R^2(\xi_k) = 2\sum_{n_k+1}^{n_{k+1}} \kappa_n^2$ so that

$$\mathsf{P}(\xi_k > \tfrac{1}{4}) = 2\int_{1/2}^{\infty} (2\pi)^{-1/2} e^{-\frac{1}{2}\tau^2}\, d\tau > \tfrac{1}{3}$$

in the case (i) and

$$\mathsf{P}(\xi_k > \tfrac{1}{4}) \geqq 1 - \frac{2\sum \kappa_n^2}{|\sum \kappa_n - \frac{1}{4}|} \geqq 1 - \frac{2\sum \kappa_n}{|\sum \kappa_n - \frac{1}{4}|^2} \geqq \tfrac{1}{3} \quad (n_k < n \leqq n_{k+1})$$

in the case (ii). So $\sum_1^\infty \mathsf{P}(\xi_k > \frac{1}{4}) = \infty$ where the ξ_k's are independent. On applying to the wellknown lemma (Borel-Cantelli), we get that with probability 1 the event $\xi_k > \frac{1}{4}$ takes place for an infinite number of indices k. It means that $\sum \xi_k = \sum Kv_n = \infty$ with probability one. So (4.17) cannot hold unless $\sum \kappa_n < \infty$. The proof is finished.

5. Application to stationary processes with rational spectral density. Let us consider a finite segment $\{x_t, 0 \leqq t \leqq T\}$ of a second-order stationary process with spectral density of form

$$(5.1) \qquad f(\lambda) = \frac{1}{2\pi} \left|\sum_{k=0}^{n} a_{n-k}(i\lambda)^k\right|^{-2} \quad (-\infty < \lambda < \infty),$$

where the constants a_{n-k} are real and chosen so that all roots of $\sum a_{n-k}\lambda^k = 0$ have negative real parts. Put

$$(5.2) \qquad X_t = \int_0^t x_s \, \mathrm{d}s \quad (0 \leqq t \leqq T),$$

and

$$(5.3) \qquad Y_t = \sum_{k=0}^{n} a_{n-k} X_t^{(k)} \quad (0 \leqq t \leqq T),$$

where $X_t^{(k)} = (\mathrm{d}^k/\mathrm{d}t^k)\, X_t$. Of course, $X_t^{(k)} = x_t^{(k-1)}$ for $k > 0$.

Lemma 5.1. *The process Y_t, defined by (5.3), has uncorrelated increments and $E|\mathrm{d}Y_t|^2 = \mathrm{d}t$. $Y_t - Y_0$ is uncorrelated with random variables $x_0, x_0', \ldots, x_0^{(n-1)}$, and the closed linear manifold $\mathscr{Y}$ generated by $\{Y_t - Y_0, 0 \leqq t \leqq T, x_0, \ldots, x_0^{(n-1)}\}$ equals the closed linear manifold $\mathscr{X}$ generated by $\{x_t, 0 \leqq t \leqq T\}$, $\mathscr{Y} = \mathscr{X}$.*

Proof. On making use of the well-known unitary mapping W, defined by $Wx_t = e^{it\lambda}$, and remembering that $R(x, y) = \int Wx\, Wy\, f(\lambda)\, \mathrm{d}\lambda$ we get from (5.2) and (5.3) that

$$(5.4) \qquad W(Y_t - Y_0) = \frac{e^{it\lambda} - 1}{i\lambda} \sum_{k=0}^{n} a_{n-k}(i\lambda)^k$$

and $Wx_0^{(j)} = (i\lambda)^j$, $0 \leqq j \leqq n - 1$. Consequently,

$$(5.5) \quad R(Y_t - Y_0, Y_s - Y_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(e^{it\lambda} - 1)(e^{-is\lambda} - 1)}{\lambda^2} \, \mathrm{d}\lambda = \min(t, s),$$

which proves that Y_t has uncorrelated increments and $\mathrm{E}|\mathrm{d}Y_t|^2 = \mathrm{d}t$. Further, we have

$$(5.6) \qquad R(Y_t - Y_0, x_0^{(j)}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(e^{it\lambda} - 1)(-i\lambda)^{j-1}}{\sum a_{n-k}(-i\lambda)^k} \, \mathrm{d}\lambda = 0,$$

because, according to our assumption, the function under the integral sign is analytic in the upper hyperplane, including the real line, and is of order $|\lambda|^{j-1-n}$. Finally, if $v \in \mathscr{X}$ and $v \perp y_t$, $0 \leqq t \leqq T$, and $v \perp x_0^{(j)}$, $0 \leqq j \leqq n - 1$, then the function $\varphi_t = R(x_t, v)$ satisfies the relations $(L\varphi)_t = 0$, $0 \leqq t \leqq T$, $\varphi_0^{(j)} = 0$, $0 \leqq j \leqq n - 1$, i.e. $\varphi_t = 0$, $0 \leqq t \leqq T$ and $v \equiv 0$. The proof is concluded.

Let $\mathscr{L}_n^2$ denote the set of complex-value functions of $t \in [0, T]$ possessing quadratically integrable derivatives up to the order n. Introduce in $\mathscr{L}_n^2$ the linear operators

$$(5.7) \qquad (L\varphi)_t = \sum_{k=0}^{n} a_{n-k} \varphi_t^{(k)},$$

$$(5.8) \qquad (L^*\varphi)_t = \sum_{k=0}^{n} (-1)^k a_{n-k} \varphi_t^{(k)},$$

where a_{n-k} are taken from (5.1). Introduce in $\mathscr{L}_{2n}^2$ the operator

$$(5.9) \quad (L^*L\varphi)_t = \sum_{k=0}^{n} \sum_{j=0}^{n} (-1)^k a_{n-k} a_{n-j} \varphi_t^{(j+k)} = (-1)^n a_0^2 \varphi_t^{(2n)} + \sum_{k=1}^{n} (-1)^k A_{n-k} \varphi_t^{(2k)}.$$

Theorem 5.1. *Let $\{x_t, t \in T\}$ be a finite segment of stationary process with spectral density (5.1). Let (Φ^x, Q^x) be closed linear manifold generated by covariances $R(x_t, x_s)$, $0 \leqq t, s \leqq T$.*

Then, in above notation, $\Phi^x = \mathscr{L}_n^2$ and

$$(5.10) \qquad Q^x(\psi, \varphi) = = \int_0^T (L\psi)_t (L\bar{\varphi}_t)\, dt + \sum_{\substack{0 \leqq j,\; k \leqq n-1 \\ j+k \text{ even}}} \psi_0^{(j)} \bar{\varphi}_0^{(k)} 2 \sum_{i=\max(0, j+k+1-n)}^{\min(j,k)} (-1)^{j-i} a_{n-i} a_{n+i-j-k-1},$$

or, equivalently,

$$(5.11) \qquad Q^x(\psi, \varphi) = a_0^2 \int_0^T \psi_t^{(n)} \bar{\varphi}_t^{(n)}\, dt + \sum_{k=1}^{n} A_{n-k} \int_0^T \psi_t^{(k)} \bar{\varphi}_t^{(k)}\, dt + \sum_{0 \leqq j,\; k \leqq n-1} [\psi_T^{(j)} \bar{\varphi}_T^{(k)} + \psi_0^{(j)} \bar{\varphi}_0^{(k)}] \sum_{i=\max(0, j+k+1-n)}^{\min(j,k)} (-1)^{j-i} a_{n-i} a_{n+i-j-k-1}.$$

If $\varphi \in \mathscr{L}_{2n}^2$, we may write

$$(5.12) \qquad Q^x(\psi, \varphi) = = \int_0^T \psi_t (L^* L\bar{\varphi})_t\, dt + \sum_{j=0}^{n-1} \psi_T^{(j)} \sum_{k=0}^{n-1-j} (-1)^k (L\bar{\varphi})_T^{(k)} a_{n-j-k-1} + \sum_{j=0}^{n-1} \psi_0^{(j)} \sum_{k=0}^{n-1-j} (-1)^j (L^* \bar{\varphi})_0^{(k)} a_{n-k-j-1}.$$

The solution $v^\varphi = U^{-1}\varphi$ of the equation $\varphi_t = R(x_t, v)$, $\varphi \in \Phi^x$, is given by (5.10) or (5.11) on substituting $dX_t^{(k)}$ for $\psi_t^{(k)}\, dt$, $0 \leqq k \leqq n$, and defining the integrals in the sense of Section 3.1. Random variables $x_t(\omega)$ may be adjusted on zero-probabality ω-subsets so that the trajectories $x^\omega = x_t(\omega)$ belong to $\mathscr{L}_{n-1}^2$ with probality 1 and for $\varphi \in \mathscr{L}_{n+1}^2$, $v^\varphi(\omega)$ is given by (5.10) or (5.11) on substituting $x_t^{(k)}(\omega)$ for $\psi_t^{(k)}$, $0 \leqq k \leqq n - 1$, and

$$x_T^{(n-1)}(\omega)\, \bar{\varphi}_T^{(j)} - x_0^{(n-1)}(\omega)\, \bar{\varphi}_0^{(j)} - \int_0^T x_t^{(n-1)}(\omega)\, \bar{\varphi}_t^{(j+1)}\, dt \quad \text{for} \quad \int_0^T x_t^{(n)} \varphi_t^{(j)}\, dt.$$

If $\varphi \in \mathscr{L}_{2n}^2$, $v^\varphi(\omega)$ is given by

$$(5.13) \quad v^\varphi(\omega) = \int_0^T x_t(\omega) (L^* L\bar{\varphi})_t\, dt + \sum_{j=0}^{n-1} x_T^{(j)}(\omega) \sum_{k=0}^{n-1-j} (-1)^k (L\bar{\varphi})_T^{(k)} a_{n-j-k-1} + \sum_{j=0}^{n-1} x_0^{(j)}(\omega) \sum_{k=0}^{n-1-j} (-1)^j (L\bar{\varphi})_0^{(k)} a_{n-j-k-1}.$$

Proof. From Lemma 5.1 it follows that every element $v \in \mathscr{X}$ is of form

$$(5.14) \qquad v = v_1 + v_2 = \sum_{j=0}^{n-1} c_j x_0^{(j)} + \int_0^T g(t)\, dY_t$$

where $\int |g(t)|^2 \, dt < \infty$, and the random variable $v_1 = \sum c_j x_0^{(j)}$ is uncorrelated with dY_t, $0 \leqq t \leqq T$, and $v_2 = \int g_t \, dY_t$ is uncorrelated with $x_0^{(j)}$, $0 \leqq j \leqq n-1$. If $\varphi_t = R(x_t, v)$, then

$$(5.15) \qquad \varphi_0^{(j)} = R(x_0^{(j)}, v) = R(x_0^{(j)}, v_1) \quad (0 \leqq j \leqq n-1),$$

and, in view of (5.3),

$$(5.15') \qquad (L\varphi)_t = \frac{d}{dt} R(Y_t, v) = R(y_t, v_2) \quad (0 \leqq t \leqq T),$$

where $y_t = dY_t/dt$ is the white noise and $L\varphi \in \mathcal{L}^2$. Consequently, if $\varphi \in \Phi^x$, then $\varphi \in \mathcal{L}_n^2$. Now, let $\varphi \in \mathcal{L}_n^2$, and chose the constants $c_0, \ldots, c_{n-1}$ and the function g_t such that (5.15) and (5.15′) hold true for v given by (5.14). A direct determination of $c_0, \ldots, c_{n-1}$ would insolve cumbersome inversion of the matrix $(R(x_0^{(j)}, x_0^{(k)}))_{j,k-1}^n$. For this reason, we first establish g_t and, subsequently, we find $c_0, \ldots, c_{n-1}$ by an indirect method. In accordance with (3.5), (5.15′) is solved by

$$(5.16) \qquad v_2 = \int_0^T (L\overline{\varphi})_t \, dY_t = \int_0^T (L\overline{\varphi})_t \, d(LX)_t \, .$$

The last form is based on (5.3). On inserting (5.16) in (5.14), we obtain

$$(5.17) \qquad v = \sum_{j=0}^{n-1} c_j x_0^{(j)} + \int_0^T (L\overline{\varphi})_t \, d(LX)_t \, .$$

Now consider a unitary and selfadjoint operator V defined in $\mathcal{X}$ by $Vx_t = x_{T-t}$. On carrying over V to Φ^x, we obtain $(V\varphi)_t = R(Vx_t, v^\varphi) = R(x_{T-t}, v^\varphi) = \varphi_{T-t}$. If $\varphi_{T-t} = \varphi_t$, $0 \leqq t \leqq T$, we have $V\varphi = \varphi$ and $v = V^*v = Vv$ for $v = U_x^{-1}\varphi$ (i.e. for v such that $\varphi_t = R(x_t, v)$). Consequently, if $\varphi_{T-t} = \varphi_t$, the formula (5.17) must be invariant with respect to the substitution of x_{T-t} for x_t, i.e. of $(L^*X)_{T-t}$ for $(LX)_t$ and $(-1)^j x_T^{(j)}$ for $x_0^{(j)}$. This means that, for $\varphi_{T-t} = \varphi_t$, we have

(5.18)

$$v = \int_0^T (L\overline{\varphi})_t \, d(L^*X)_{T-t} + \sum_{j=0}^{n-1} c_j (-1)^j x_T^{(j)} = \int_0^T (L^*\varphi)_t \, d(L^*X)_t + \sum_{j=0}^{n-1} c_j (-1)^j x_T^{(j)} ,$$

where the last expression is obtained from the preceeding one on substituting $(L\overline{\varphi})_t = (L^*\overline{\varphi})_{T-t}$, which is justified because $\varphi_{T-t} = \varphi_t$. If is easy to show that

$$(5.19) \qquad \int_0^T (L\overline{\varphi})_t \, d(LX_t) =$$

$$= \int_0^T (L^*\overline{\varphi})_t \, d(L^*X)_t + \sum_{\substack{j+k \text{ odd} \\ k>j}}^{n-1} \sum^{n-1} a_{n-k} a_{n-j} \sum_{v=0}^{k-j-1} (-1)^v [\overline{\varphi}_T^{(j+v)} x_T^{(k-j-v)} - \overline{\varphi}_0^{(j+v)} x_0^{(k-j-v)}] =$$

$$= \int_0^T (L^*\overline{\varphi})_t \, d(L^*X)_t + \sum_{j+k \text{ even}}^{n-1} \sum^{n-1} (x_t^{(j)} \overline{\varphi}_t^{(k)} -$$

$$- x_0^{(j)} \overline{\varphi}_0^{(k)}) \sum_{i=\max(0,j+k+1-n)}^{\min(j,k)} (-1)^{j-i} a_{n-i} a_{n+i-j-k-1} \, .$$

On comparing (5.17), (5.18) and (5.19), we can easily see that

$$(5.20)\quad v_1 = \sum_{j=0}^{n-1} c_j x_0^{(j)} = \sum_{j+k \text{ even}}^{n-1} \sum^{n-1} x_0^{(j)} \overline{\varphi}_0^{(k)} 2 \sum_{i=\max(0,j+k+1-n)}^{\min(j,k)} (-1)^{j-i} a_{n-i} a_{n+i-j-k-1} ,$$

which inserted in (5.17) gives the needed result. Since the condition. $\varphi_{T-t} = \varphi_t$ imposes no restrictions on $\varphi_0, \ldots, \varphi_0^{(n-1)}$, (5.20) represents a general solution of (5.15).

As we know, $f(x) = R(x, v^\varphi)$ is linear functional attaining the value φ_t at the point $x = x_t$. Consequently, $Q(\psi, \varphi) = R(v^\varphi, v^\psi)$ is obtained from (5.17) simply by replacing x_t by ψ_t. Bearing in mind (5.20), we thus obtain (5.10). Formulas (5.11) and (5.12) are obtained from (5.10) simply by integration per partes, the details of which we shall not reproduce here. In deriving (5.12) we may apply Green's formula and the principle of symetry, assuming that $\varphi_{T-t} = \varphi_t$.

In order to show that (5.13) may be considered as a special case of (4.1), let us extend the parameter set by adding parameter-values $01, \ldots, 0n$ and $T1, \ldots, Tn$, and putting $x_{0i} = x_0^{(i-1)}$ and $x_{Ti} = x_T^{(i-1)}$, $1 \leqq i \leqq n$. The white noise y_t will be extended so that y_{0i} are any orthonormal linear combinations of $x_{0i} = x_0^{(i-1)}$, and $y_{T1}, \ldots, y_{Tn}$ are any orthonormal random variables uncorrelated with x_t, $0 \leqq t \leqq T$ (or with y_t, $0 \leqq t \leqq T$, and $y_{01}, \ldots, y_{0n}$). Defining the linear operator K by $x_t = Ky_t$, $0 \leqq t \leqq T$, $t = 01, \ldots, \text{on}, T1, \ldots, Tn$, we can see that for $0 \leqq t \leqq T$, L coincides with K_x^{-1} defined by (2.10). In our case $K_x^{-1}\varphi$ simply is such a solution of $K(.) = \varphi$ that vanishes for $t = T1, \ldots, Tn$. After this extention of the set of parameter-values, (5.13) is equivalent to (3.17), where μ is defined as Lebesque measure for $0 \leqq t \leqq T$, and $\mu(0i) = \mu(Ti) = 1$, $1 \leqq i \leqq n$. Moreover, (5.13) is equivalent (4.1). The only purpose of adding points $T1, \ldots, Tn$ was to enlarge the range of the operator KK^* so that it contains all functions from $\mathscr{L}_{2n}^2$ (see Remark 4.1).

Finally, if $n > 1$ then the derivative $x_t' = z_t$ exist. Considering the operator K defined by

$$x_t = Kz_t = x_0 + \int_0^t z_s \, ds ,$$

and applying Theorem 4.2, we readily see that $x_t(\omega)$ may be adjusted so that $x^\omega \in \mathscr{L}_{n-1}^2$ with probability 1, and that, for $\varphi \in \mathscr{L}_{n+1}^2$, the solution of $\varphi_t = R(x_t, v)$ has the form mentioned in the theorem. The proof is finished.

Remark 5.1. From (5.20) it follows that the inverse of $(R(x_0^{(j)}, x_0^{(k)}))_{j,k=0}^{n-1}$ consists of elements

$$(5.21)\quad D_{jk} = 2 \sum_{i=\max(0,j+k+1-n)}^{\min(j,k)} (-1)^{j-i} a_{n-i} a_{n+i-j-k-1} \quad \text{for } j+k \text{ even},$$

$$= 0 \quad \text{for } j+k \text{ odd}.$$

Now, we shall consider a general rational spectral density

$$(5.22)\qquad g(\lambda) = \frac{1}{2\pi}\left|\frac{\sum_{k=0}^{m} b_{m-k}(i\lambda)^k}{\sum_{k=0}^{n} a_{n-k}(i\lambda)^k}\right|^2 \qquad (-\infty < \lambda < \infty;\ m < n),$$

where the a's and b's are real and all roots of equations $\sum a_{n-k}\lambda^k = 0$ and $\sum b_{m-k}\lambda^k = 0$ have negative real parts. As is well-known, if x_t is a process with spectral density (5.1), then the process

$$(5.23)\qquad z_t = \sum_{k=0}^{m} b_{m-k} x_t^{(k)} \quad (m < n)$$

has the spectral density the (5.22).

Let us add $2m$ parameter-values $01, \ldots, 0m, T1, \ldots, Tm$ and define $z_{0i} = x_{0i} = x_0^{(i-1)}$, $z_{Ti} = x_{Ti} = x_T^{(i-1)}$, $1 \leqq i \leqq m$. Then the operator H defined by $x_t = Hz_t$, $0 \leqq t \leqq T$, $t = 01, \ldots, 0m, T1, \ldots, Tm$ is bounded and may be expressed by

$$(5.24)\qquad x_t = \sum_{i=1}^{m} x_0^{(i-1)} h_j(t) + \int_0^T H(t, s)\, z_s\, \mathrm{d}s \quad (0 \leqq t \leqq T),$$

where $h_j(t)$ are m linearly independent solutions of the equation $\sum b_{m-k}h(t) = 0$, and $H(t, s)$ may be established by well-known methods [31] of the theory of linear differential equations. The differential operator involved in (5.23) will be denoted shortly by M,

$$(5.25)\qquad M = \sum_{k=0}^{m} b_{m-k} \frac{\mathrm{d}^k}{\mathrm{d}t^k}.$$

Further, we put

$$M^* = \sum_{k=1}^{m} b_{m-k}(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}t^k} \quad \text{and} \quad MM^* = \sum_{k=1}^{m}\sum_{j=1}^{m} b_{m-k}b_{m-j}(-1)^k \frac{\mathrm{d}^{k+j}}{\mathrm{d}t^{k+j}}.$$

Theorem 5.2. *Let $\{z_t, 0 \leqq t \leqq T\}$ be a finite segment of stationary process with spectral density (5.22). Let (Φ^z, Q^z) the closed linear manifold generated by covariances $R(z_t, z_s)$.*

Then $\Phi^z = \mathscr{L}^2_{n-m}$ and the solution $v^\chi = U_z^{-1}\chi$ of the equation $\chi_t = R(z_t, v)$ is the same as the solution of the equation $\psi_t = R(x_t, v)$, where x_t is given by (5.24) and ψ_t is uniquely determined by

$$(5.26)\qquad (M\psi)_t = \chi_t \quad \textit{and} \quad Q^x(\psi, h_j) = 0 \quad (1 \leqq j \leqq m),$$

where M is given by (5.25), $h_1, \ldots, h_m$ are m independent solutions of $\sum b_{m-k}h(t) = 0$ and Q^x is given by (5.10) or (5.11). The inner product Q^z in Φ^z is given by $Q^z(v, \chi) = Q^x(\xi, \psi)$ where Q^x is given by (5.10), ψ is determined by (5.26), and ξ also is determined by (5.26) on substituting v for χ.

Random variables $z_t(\omega)$ may be adjusted on zero-probability ω-subsets so that the trajectories $z^\omega = R_t(\omega)$ belong to $\mathscr{L}^2_{n-m-1}$ with probability 1, and that the in-

volved integral and differential operators may be applied to individual trajectories for $\chi \in \mathscr{L}^2_{n-m+1}$. *If* $\chi \in \mathscr{L}^2_{2n-2m}$, *then the solution is given by*

$$(5.27)\quad v^{\chi}(\omega) = \int_0^T z_t(\omega)\,(L^*L\varphi)_t\,dt + \sum_{j=0}^{n-m-1} z_T^{(j)}(\omega) \sum_{k=0}^{n-1-j} (-1)^z (L\varphi)_T^{(k)} a_{n-j-k-1} + \\ + \sum_{j=0}^{n-m-1} z_0^{(j)}(\omega) \sum_{k=0}^{n-1-j} (-1)^j (L^*\varphi)_0^{(k)} a_{n-j-k-1}\,,$$

where φ_t *is uniquely determined by*

$$(5.28)\qquad (MM^*\varphi)_t = \chi_t\,,$$

and

$$(5.29)\qquad \sum_{k=0}^{n-1-j} (-1)^k (L\varphi)_T^{(k)} a_{n-j-k-1} = 0 \quad (n-m \leqq j \leqq n-1)\,,$$

$$(5.30)\qquad \sum_{k=0}^{n-1-j} (-1)^j (L^*\varphi)_T^{(k)} a_{n-j-k-1} = 0 \quad (n-m \leqq j \leqq n-1)\,.$$

Proof. Variances generated by spectral density (5.22) and variances generated by the spectral density (5.1), where $n \equiv n - m$, dominate each other, so that the closed linear manifolds generated by respective covariances coincide (cf. Remark 2.1). So, in view of Theorem 5.1, $\Phi^z = \mathscr{L}^2_{n-m}$.

Now, if $R(x_t, v) = \varphi_t$, and φ_t satisfies (5.26), then obviously $R(z_t, v) = \chi_t$. Further, it is easy to verify on the basis of (5.24) that $Q^x(\varphi, h_j) = 0$, $1 \leqq j \leqq m$, guarantees that the solution v belongs to the closed linear manifold $\mathscr{Z}$ generated by random variables z_t, $0 \leqq t \leqq T$.

Further, if $\chi \in \mathscr{L}^2_{2n-2m}$, and the conditions (5.29) and (5.30) are satisfied, then the operator M may be applied to the right side of (5.13) term by term under the integral sign which shows that $Mv^{\varphi} = v^{\chi}$, where v^{χ} is given by (5.27). Consequently $R(z_t, v^{\chi}) = MM^*R(x_t, v\varphi) = (MM^*\varphi)_t = \chi_t$, $0 \leqq t \leqq T$, in accordance with (5.27).

The assertions concerning individual trajectories are implied by Theorems 4.1 and 4.2 as in the preceeding theorem.

Remark 5.2. If $h_1, \ldots, h_m$ are the solutions of $Mh = 0$, then the solutions of $MM^*h = 0$ are $h_1, \ldots, h_1, h_1^*, \ldots, h_m^*$ where $h^*(t) = h(T-t)$. Let φ_0 be particular solution of $MM^*\varphi = \chi$. Then

$$\varphi = \varphi_0 + \sum_1^m (c_j h_j + d_j h_j^*), \quad \text{where} \quad c_1, \ldots, c_m, d_1, \ldots, d_m$$

are determined by (5.29) and (5.30). If $\chi_t = \chi_{T-t}$, then $c_j = d_j$, and if $\chi_t = -\chi_{T-t}$, then $c_j = -d_j$. In both the cases the number of unknown constants is m, and not $2m$, and they are determined by (5.29) or (5.30). Denoting $\chi_t^* = \chi_{T-t}$, we have

$$(5.31)\qquad \varphi = \varphi^\circ + \tfrac{1}{2} \sum_{j-1}^m \{c_j(\chi + \chi^*)\,[h_j + h_j^*] + c_j(\chi - \chi^*)\,[h_j - h_j^*]\}\,,$$

where $c_j(v)$ denote the constants corresponding to a function v.

Example 5.1. If $M\varphi = \varphi' + \beta\varphi$, $\beta > 0$, then the solution of (5.28) is given by

$$(5.31')\quad \varphi_t = \frac{1}{2\beta}\int_0^T e^{-|t-s|\beta}\chi_s\,ds + \tfrac{1}{2}(c_1 + c_2)\,e^{-t\beta} + \tfrac{1}{2}(c_1 - c_2)\,e^{-(T-t)\beta},$$

where

$$(5.32)\quad c_1 = \frac{\sum_{1\leqq j\leqq n/2} \beta^{2j-2}\sum_{k=0}^{n-2j} a_{n-2j-k}(\chi_0^{(k)} + \chi_T^{(k)}) - L(-\beta)\int_0^T (e^{-t\beta} + e^{-(T-t)\beta})\,\chi_t\,dt}{L(\beta) + L(-\beta)\,e^{-T\beta}},$$

and

$$(5.33)\quad c_2 = \frac{\sum_{1\leqq j\leqq n/2} \beta^{2j-2}\sum_{k=0}^{n-2j} a_{n-2j-k}(\chi_0^{(k)} - (-1)^k\chi_T^{(k)}) - L(-\beta)\int_0^T (e^{-t\beta} - e^{-(T-t)\beta})\,\chi_t\,dt}{L(\beta) - L(-\beta)\,e^{-T\beta}},$$

where

$$L(\beta) = \sum_{k=0}^{n} a_{n-k}\beta^k .$$

Moreover,

$$(5.34)\qquad (L^*L\varphi)_t = \varphi_t L^*L(\beta) + \frac{L^*L\left(\frac{d}{dt}\right) - L^*L(\beta)}{\beta^2 - \left(\frac{d}{dt}\right)^2}\,\chi_t,$$

where $L^*L(\beta) = L(-\beta)\,L(\beta)$ and d/dt denotes the differential operator. We omit the details.

Integral-valued parameter t. The spectral density of an n-th order Markovian process with integral-valued t is given by

$$(5.35)\qquad f(\lambda) = \frac{1}{2\pi}\left|\sum_{k=0}^{n} a_{n-k}e^{i\lambda k}\right|^{-2} \quad (-\pi \leqq \lambda \leqq \pi),$$

where a_{n-k} are real and such that all roots of $\sum_{k=0}^{n} a_{n-k}\lambda^k = 0$ are greater that 1 in absolute value.

It may be easily shown that the random variables

$$(5.36)\qquad y_t = \sum_{k=0}^{n} a_{n-k}x_{t-k} \quad (n \leqq t \leqq N)$$

are uncorrelated mutually as well as with $x_0, \ldots, x_{n-1}$, and have unit variance, $R(y_t) = 1$. If we consider a finite segment, of the process, $\{x_t, 0 \leqq t \leqq N\}$, $N \geqq 2n$, then Φ^x consist of all complex-valued functions φ_t, $0 \leqq t \leqq N$, and it may by shown by the previous method that

$$Q^x(\psi, \varphi) = \sum_{t=0}^{N}\sum_{s=0}^{N}\psi_t\bar{\varphi}_s Q^x_{ts}, \quad \text{where for } |t - s| > n \ Q^x_{ts} = 0,$$

and for $|t - s| \leqq n$

$$(5.37) \quad Q^x_{ts} = \sum_{i=0}^{\min[N-t, N-s, n-|t-s|]} a_{n-i}a_{n-i-|t-s|}\,, \quad \text{if} \quad \max(t, s) > N - n\,,$$

$$= \sum_{i=0}^{n-|t-s|} a_{n-i}a_{n-i-|t-s|}\,, \quad \text{if} \quad n \leqq t,\ s \leqq N - n\,,$$

$$= \sum_{i=0}^{\min[t,s,n-|t-s|]} a_{n-i}a_{n-i-|t-s|}\,, \quad \text{if} \quad \min(t, s) < n\,.$$

The solution of $\varphi_t = R(x_t, v)$ is given by

$$(5.38) \qquad v^\varphi = \sum_{t=0}^{N} \sum_{s=0}^{N} x_t \overline{\varphi}_s Q^x_{ts}\,,$$

which may be put in a form analogous to (5.13)

$$(5.39) \quad v^\varphi = \sum_{t=n}^{N-n} x_t(L^*L\varphi)_t + \sum_{t=N-n+1}^{N} x_t \sum_{k=0}^{N-t} (L\varphi)_{t+k} a_{n-k} + \sum_{t=0}^{n-1} x_t \sum_{k=0}^{t} (L^*\varphi)_{t-k} a_{n-k}\,,$$

where

$$(L\varphi)_t = \sum_{k=0}^{n} a_{n-k}\varphi_{t-k} \quad \text{and} \quad (L^*\varphi)_t = \sum_{k=0}^{n} a_{n-k}\varphi_{t+k}\,.$$

We also find that the inverted covariance matrix $(R(x_j, x_k))^{n-1}_{j,k=0}$ consists of the following elements

(5.40)

$$D_{jk} = \sum_{i=0}^{\min[j,k,n-|j-k|]} a_{n-i}a_{n-i-|j-k|} - \sum_{i=n-\max(j,k)}^{n-|j-k|} a_{n-i}a_{n-i+|j-k|} \quad (0 \leqq j,\ k \leqq n - 1)\,.$$

If we have a process $\{z_t, 0 \leqq t \leqq N\}$ possessing a general rational spectral density

$$(5.41) \qquad g(\lambda) = \frac{1}{2\pi} \left| \frac{\sum_{k=0}^{m} b_{m-k} e^{i\lambda k}}{\sum_{k=0}^{m} a_{n-k} e^{i\lambda k}} \right|^2 \qquad (-\pi \leqq \lambda \leqq \pi)\,,$$

where both the a's and b's are real and such that the roots of $\sum_{k=0}^{m} a_{n-k}\lambda^k = 0$ and $\sum_{k=0}^{m} b_{m-k}\lambda^k = 0$ lie outside of the unit circle, we may proceed as we did in the continuous-parameter case.

First, we may consider, the process $\{x_t, -m \leqq t \leqq N\}$ having the spectral density (5.35) related to z_t by the difference equation

$$(5.42) \qquad z_t = \sum_{k=0}^{m} b_{m-k} x_{t-k} = (Mx)_t \quad (0 \leqq t \leqq N)\,.$$

Then the equation $\chi_t = R(z_t, v)$ is solved by

$$v^\varphi = \sum_{t=-m}^{N} \sum_{s=-m}^{N} x_t \overline{\psi}_t Q^{ts}_x$$

where ψ is uniquely determined by $(M\psi)_t = \sum_{k=0}^{m} b_{m-k}\,\psi_{t-k} = \chi_t,\ 0 \leqq t \leqq N$, and $Q^x(\psi, h_j) = 0$ for m arbitrary linearly independent solutions of $(Mh)_t = 0, 0 \leqq t \leqq N$. Similarly $Q^z(v, \chi) = Q^x(\xi, \psi)$, where ξ is an arbitrary solution of the equation $M\xi = v$.

Second, we may consider the process $\{x_t, -m \leqq t \leqq N + m\}$ and the adjoint operator

$$(M^*\varphi)_t = \sum_{k=0}^{m} b_{m-k}\varphi_{t+k}\,.$$

Then we obtain the following analogues of equation (5.27):

$$v = \sum_{t=0}^{N} z_t(L^*L\varphi)_t \quad (m \geqq n)\,, \tag{5.43}$$

and

$$v = \sum_{t=n-m}^{N-n+m} z_t(L^*L\varphi)_t + \sum_{t=N-n+m}^{N} z_t \sum_{k=0}^{N-t+m} (L\varphi)_{t+k}a_{n-k} + \sum_{t=0}^{n-m-1} z_t \sum_{k=0}^{t+m} (L^*\varphi)_{t-k}a_{n-k} \tag{5.44}$$
$$(m < n)\,,$$

where φ_t is uniquely determined by $(MM^*\varphi)_t = \chi_t,\ 0 \leqq t \leqq N$, and

$$\sum_{k=0}^{r} (L^*\varphi)_{r-m-k}a_{n-k} = 0 \quad (r = 0, \ldots, m-1)\,, \tag{5.45}$$

and

$$\sum_{k=0}^{r} (L\varphi)_{N+m-r+k}a_{n-k} = 0 \quad (r = 0, \ldots, m-1)\,. \tag{5.46}$$

Cf. (5.29) and (5.30). Remark (5.2) is also valid for the present case. If $n = 0$, we get results for the moving-average scheme.

6. Strong equivalency of normal distributions. Let us first consider two normal distributions P and P^+ of a random sequence $\{v_n, n \geqq 1\}$ defined by vanishing mean values $\mathsf{E}v_n = \mathsf{E}^+v_n = 0,\ n \geqq 1$, and covariances

$$\begin{aligned} &R(v_n, v_m) = 0\,, \quad R^+(v_n, v_m) = 0\,, \quad \text{if } n \neq m\,, \\ &R(v_n, v_n) = 1\,, \quad R^+(v_n, v_n) = \frac{1}{\lambda_n} \qquad (n \geqq 1)\,. \end{aligned} \tag{6.1}$$

The J-divergence of P a P^+ restricted to the vector $\{v_1, \ldots, v_n\}$, say J_n, equals (see [18])

$$J_n = \tfrac{1}{2}\sum_{i=1}^{n} \frac{(1-\lambda_i)^2}{\lambda_i}\,, \tag{6.2}$$

and, consequently, $J_\infty = \lim_{n\to\infty} J_n < \infty$, if and only if

$$\sum_{n=1}^{\infty} (1-\lambda_n)^2 < \infty\,. \tag{6.3}$$

So, according to [18], P and P^+ defined on the Borel field generated by $\{v_n, n \geqq 1\}$

are equivalent, $P \sim P^+$, if and only if (6.3) is true. If $P \sim P^+$, then it follows from theory of martingales ([30], Th. 4.3, Ch. VIII) that

(6.4) $$\frac{dP}{dP^+} = \exp\{-\tfrac{1}{2}\sum_1^\infty [v_n^2(1-\lambda_n) + \log \lambda_n]\},$$

where the sum converges with probability 1. In general, however, we cannot write

(6.5) $$\frac{dP}{dP^+} = (\prod_1^\infty \lambda_n)^{-1/2} \exp\{-\tfrac{1}{2}\sum_1^\infty v_n^2(1-\lambda_n)\},$$

unless the product $\prod_1^\infty \lambda_n$ is absolutely convergent. The well-known necessary and sufficient condition for absolute convergence of $\prod_1^\infty \lambda_n$ is $\lambda_n \neq 0$ and

(6.6) $$\sum_1^\infty |1-\lambda_n| < \infty,$$

which is stronger then (6.3). If (6.6) is satisfied, then almost surely

(6.7) $$\sum_1^\infty v_n^2(\omega)\,|1-\lambda_n| < \infty,$$

because $E\{\sum_1^\infty v_n^2(\omega)\,|1-\lambda_n|\} \leqq \sum_1^\infty |1-\lambda_n| < \infty$.

Thus (6.6) is a necessary and sufficient condition for dP/dP^+ being of the form (6.5), where the product $\prod_1^\infty \lambda_n$ is absolutely convergent and the quadratic form $\sum_1^\infty v_n^2(1-\lambda_n)$ is absolutely convergent with probability 1. This leads us to the notion of strongly equivalent normal distributions introduced in the following.

Definition 6.1. Two covariances R and R^+ will be called strongly equivalent, if they dominate each other (Definition 2.2), the operator B defined by $R(x, y) = R^+(Bx, y)$ has purely point spectrum, and the non-unity eigen-values λ_n of B satisfy the condition (6.6). Two normal distribution P and P^+ defined by strongly equivalent covariances R a R^+ and by vanishing mean values will be called strongly equivalent.

Lemma 6.1. *The Probability density of a normal distribution* P *with respect to another normal distribution* P^+, *strongly equivalent to* P, *is given by* (6.5) *where* v_n *and* λ_n *denote the eigen-elements and eigen-values of the operator* B *mentioned in Definition* 6.1. *The* v_n'*s are normed by* $R(v_n, v_n) = 1$, $n \geqq 1$.

Proof. Clear.

Unfortunately, the right side of (6.5) scarcely may be considered as an ultimate expression for dP/dP^+, because the eigen-values and eigen-elements are difficult to establish even if dP/dP^+ may be found explicitely (see Section 7). This leads us to derive a theory of the product $\prod_1^\infty \lambda_n$ and of the quadratic form $\sum_1^\infty v_n^2(1-\lambda_n)$.

We begin with the following:

Definition 6.2. Let H be a compact symmetric operator, and κ_n be the non-zero eigen-values of H. If $\sum_1^\infty |\kappa_n| < \infty$, we say that H has a finite trace, and put $\operatorname{tr} H = \sum_1^\infty \kappa_n$.

Definition 6.3. Let $B - I$ be a compact symmetric operator and λ_n be the non-unity eigen-values of B. If the product $\prod_1^\infty \lambda_n$ is absolutely convergent, we say that B has a determinant, and put $\det B = \prod_1^\infty \lambda_n$.

Obviously, B has a determinant if and only if the eigen-values of B are different from 0 and $B - I$ has a finite trace. If $\{v_n\}$ is the orthonormal system of eigen-elements of B, corresponding to the non-zero eigen-values, and $\{x_n\}$ is another orthonormal system in $(\mathscr{X}, R^+)$, then

$$\sum_1^\infty |(Hx_n, x_n)| = \sum_{n=1}^\infty \Big| \sum_{j=1}^\infty (Hx_n, v_j)(x_n, v_j)\Big| = \sum_{n=1}^\infty \Big| \sum_{j=1}^\infty \kappa_j |(x_n, v_j)|^2\Big| \leqq$$
$$\leqq \sum_{j=1}^\infty |\kappa_j| \sum_{n=1}^\infty |(x_n, v_j)|^2 ,$$

i.e.

$$\sum_1^\infty |(Hx_n, x_n)| \leqq \sum_1^\infty |\kappa_j| , \tag{6.8}$$

where $(.,.) = R^+(.,.)$ and $H = B - I$.

Denote by $\log B$ the operator which has the same eigen-elements as the operator B but eigen-values $\log \lambda_n$ instead of λ_n. Denoting the eigen-elements of B by v_n, we have

$$(Bx, x) = \sum_1^\infty \lambda_n |(x, v_n)|^2 \quad \text{and} \quad (\log Bx, x) = \sum_1^\infty \log \lambda_n |(x, v_n)|^2 .$$

On making use of Jensen inequality, we get for $(x, x) = 1$

$$\log (Bx, x) \geqq (\log Bx, x) \quad [(x, x) = 1] . \tag{6.9}$$

Let $\mathscr{X}_B$ be the subspace of $\mathscr{X}$ spanned by eigen-elements of B corresponding to non-unity eigen-values. $\mathscr{X}_B$ is separable even if $\mathscr{X}$ is not separable. For any orthonormal system $\{x_n\}$ complete in $\mathscr{X}_B$ we have

$$\log \det B = \operatorname{tr} \log B = \sum_1^\infty (\log Bx_n, x_n) \leqq \sum_1^\infty \log (Bx_n, x_n) ,$$

i.e.

$$\det B \leqq \prod_{n=1}^\infty (Bx_n, x_n) , \tag{6.10}$$

where the absolute convergence on the right side is guaranted by (6.8).

Let $\mathscr{X}_n$ be a n-dimensional subspace of $\mathscr{X}$, and let I_n^+ denote the projection on $\mathscr{X}_n$. The contraction of B on $\mathscr{X}_n$ say B_n, will be defined as follows.

$$B_n = I + I_n^+(B - I) I_n^+ . \tag{6.11}$$

Obviously $(B_n x, x) = (Bx, x)$, if $x \in \mathscr{X}_n$, and $(B_n x, x) = (x, x)$, if $x \perp \mathscr{X}_n$. The operator B_n is well-defined on $\mathscr{X}$ as well as on $\mathscr{X}_n$ and in both cases the determinant Det B_n is the same. If $\mathscr{X}_n$ is spanned by linearly independent elements $x_1, \ldots, x_n$, then

$$\det B_n = \frac{|(Bx_i, x_j)|}{|(x_i, x_j)|} \quad (i, j = 1, \ldots, n) , \tag{6.12}$$

where $|a_{ij}|$ is the ordinary determinant of a matrice (a_{ij}). Relation (6.12) will be clear, if we take for $x_1, \ldots, x_n$ the eigen-elements of B_n.

Theorem 6.1. *Let $\mathscr{X}_1 \subset \mathscr{X}_2 \subset \ldots$ be a sequence of finite-dimensional subspaces of $\mathscr{X}$ such that the smallest subspace containing $\bigcup_1^\infty \mathscr{X}_n$ contains all eigen-elements of B corresponding to non-unity eigen-values. Let B_n be defined by* (6.11). *Then*

$$\det B = \lim_{n \to \infty} \det B_n . \tag{6.13}$$

Proof. Let v_j and $\lambda_j \neq 1$ be the eigen-elements and eigen-values of B, respectively. For any $\varepsilon > 0$ we may chose l so that $\sum_{l+1}^\infty |\lambda_j - 1| < \varepsilon$ and then N so that in $\mathscr{X}_N$ there exists an orthonormal system $x_1, \ldots, x_l$ such that

$$\sum_1^l |((B - I) x_j, x_j)| \geqq \sum_1^l |((B - I) v_j, v_j)| - \varepsilon = \sum_1^l |\lambda_j - 1| - \varepsilon . \tag{6.14}$$

In view of (6.8) we have for any orthonormal system $\{z_n\}$, such that $z_1 = x_1, \ldots, z_l = x_l$,

$$\sum_{l+1}^\infty |((B - I) z_n, z_n)| \leqq \sum_1^\infty |\lambda_n - 1| - \sum_1^l |((B - I) z_n, z_n)| < 2\varepsilon . \tag{6.15}$$

Now let $n > N$, and $c^{-1} = \min_{1 \leqq j \leqq l} (Bx_j, x_j)$. Let $\{v_{nj}\}$ and $\{v_j\}$ be the eigen-elements of B_n and B, respectively. From (6.10), (6.14) and (6.15) it follows that

$$\det B_n = \prod_{j=1}^n (Bv_{nj}, v_{nj}) \geqq (1 - 2\varepsilon) \det B \tag{6.16}$$

and, in view of $\sum_{l+1}^\infty |\lambda_j - 1| < \varepsilon$,

$$\begin{aligned} \det B &\geqq (1 - \varepsilon) \prod_1^l (Bv_j, v_j) \geqq (1 - \varepsilon)(1 - c\varepsilon) \prod_1^l (Bx_j, x_j) \geqq \\ &\geqq (1 - \varepsilon)(1 - 2\varepsilon)(1 - c\varepsilon) \prod_1^n (Bx_j, x_j) \geqq (1 - 3\varepsilon - c\varepsilon) \det B_n , \end{aligned} \tag{6.17}$$

where $x_1, \ldots, x_n$ is an orthonormal system from $\mathscr{X}_n$ such that $x_1, \ldots, x_l$ satisfy (6.14).

Now we may let $\varepsilon \to 0$ and $c^{-1} \geqq c_0^{-1} > 0$, where c_0^{-1} is independent of ε. The proof is completed.

Now we shall study the ratio $\det B/\det B_n$, where B_n is given by (6.11), or more generally, $\det B/\det B_c$, where $B_c = I + I_c^+(B - I)I_c^+$, I_c^+ denoting the projection on a subspace $\mathscr{X}_c$ (not necessarily finite-dimensional) of $\mathscr{X}$. More precisely, I_c^+ denotes the projection on $\mathscr{X}_c$ with respect to the covariance R^+. The projection with respect to the covariance $R(x, y) = R^+(Bx, y)$, say I_c, will be generally different from I_c^+. Now, let us introduce the following "conditional" covariances

(6.18) $$R^+(x, y \mid \mathscr{X}_c) = R^+(x - I_c^+x, y - I_c^+y),$$

(6.19) $$R(x, y \mid \mathscr{X}_c) = R(x - I_cx, y - I_cy).$$

If $R^+(. \mid \mathscr{X}_c)$ dominates $R(. \mid \mathscr{X}_c)$, denote by B^c the operator defined by

(6.20) $$R(x, y \mid \mathscr{X}_c) = R^+(B^cx, y \mid \mathscr{X}_c).$$

Theorem 6.2. *If* $\det B$ *exists, then the determinants* $\det B_c$ *and* $\det B^c$ *also exist, and*

(6.21) $$\det B = \det B_c \det B^c.$$

Proof. First suppose that we have $n + m$ random variables $x_1, \ldots, x_n, x_{n+1}, \ldots, \ldots, x_{n+m}$ which are independent with unit variances, if P^+ is true, and have an arbitrary non-singular normal distribution with covariance matrix $R = (R_{ij})$, if P is true. Let $R_n = (R_{ij})_{i,j=1}^n$ and $R^n = (R_{ij}^n)_{i,j=n+1}^{n+m}$, where R_{ij}^n is the conditional (partial) covariance of x_i and x_j $(i, j > n)$ for given $x_1, \ldots, x_n$. In this case (6.21) is equivalent to

(6.22) $$|R| = |R_n|\,|R^n|,$$

where $|\cdot|$ denotes the determinant of the corresponding matrix. Equation (6.22) may be proved as follows: We introduce random variables $z_i = x_i$ $(1 \leqq i \leqq n)$ and $z_i = x_i - I_nx_i$ $(n + 1 \leqq i \leqq n + m)$, where I_n is the projection on the subspace spanned by $x_1, \ldots, x_n$. In matrix notation $z = Ax$, $A = \{a_{ij}\}$, where $|A| = 1$, because $a_{ij} = 0\,(i < j)$ and $a_{ii} = 1\,(1 \leqq i \leqq n+m)$. The covariance matrix of random variables z_i equals

(6.23) $$ARA^* = \begin{pmatrix} R_n & O \\ O & R^n \end{pmatrix},$$

from which it follows that $|R| = |A|\,|R|\,|A^*| = |ARA^*| = |R_n|\,|R^n|$.

The general case will be obtained by Theorem 6.1. We take random variables $y_1, \ldots, y_m$ from $\mathscr{X} \ominus \mathscr{X}_c$ and n random variables $x_1, \ldots, x_n$ from $\mathscr{X}_c$ so that the closed linear manifold spanned by $x_1, \ldots, x_n$ contains the projections I_cy_i and $I_c^+y_i$ of y_i on $\mathscr{X}_c$. Then we let $n \to \infty$, and subsequently $m \to \infty$ so that the closed linear manifold spanned by $\{y_1, y_2, \ldots, x_1, x_2, \ldots\}$ contains all eigen-elements of B with non-unity eigen-values. The proof is finished.

Now we shall study the quadratic form $\sum_1^\infty v_n^2(1 - \lambda_n)$ appearing in (6.5).

Theorem 6.3. *Let* P *and* P^+ *be two normal distributions of a stochastic process* $\{x_t, t \in T\}$. *Let* $\int x_t \, d\mathsf{P} = \int x_t \, d\mathsf{P}^+ = 0$, $t \in T$. *Assume that* $x_t = Kz_t$, $t \in T$, *where* z_t *is a stochastic process or a white noise, and* K *is a compact operator such that* K^*K *has a finite trace. Let* (Φ, Q), (Φ^+, Q^+), (Φ^z, Q^z) *be closed linear manifolds generated by covariances* $R(x_t, x_s)$, $R^+(x_t, x_s)$ *and* $R(z_t, z_s)$, *respectively. Let* B *be the linear operator defined by* $R(x, y) = R^+(Bx, y)$. *Assume that* $\Phi^+ = \Phi$ *and for some constant* C

$$|Q((I - B)\varphi, \psi)| \leqq C Q^z(\varphi) \, Q^z(\psi) \quad (\varphi, \psi \in \Phi^x) . \tag{6.24}$$

Then P *and* P^+ *are strongly equivalent. Moreover, there exist a unique extention* $\widehat{Q - Q^+}$ *of* $Q - Q^+$ *from* $\Phi \times \Phi$ *on* $\Phi^z \times \Phi^z$, *continuous in the* Q^z*-norm, and we have*

$$\frac{d\mathsf{P}}{d\mathsf{P}^+} = (\det B)^{-1/2} \, e^{-\frac{1}{2}\widehat{Q - Q^+}(x^\omega, x^\omega)} . \tag{6.25}$$

where $x^\omega = x_t(\omega)$ *is the trajectory of the process modified according to Theorem* 4.2 *or* 4.1.

Proof. $I - B$ has a finite trace because K^*K has a finite trace and because (6.24) holds, where $Q^z(\varphi) = Q(K\varphi)$. Since $\Phi^+ = \Phi$, the covariances R a R^+ dominate each the other and B has all eigen-values different from 0. Hence B has a determinant, and P a P^+ are strongly equivalent.

In Φ the operator B is defined by $Q^+(\psi, \varphi) = Q(B\psi, \varphi)$, in view of Lemma 2.2. Let $\chi_n(t)$ be the eigen-elements of B corresponding to non-unity eigenvalues, $\lambda_n \neq 1$. We first show that $\chi_n = KK^*h_n$, where $h_n \in \Phi^z$. In view of (6.24), we have

$$|Q(\psi, \chi_n)| = \frac{|Q(\psi, (I - B)\chi_n)|}{|1 - \lambda_n|} \leqq C \frac{Q^z(\chi_n)}{|1 - \lambda_n|} Q^z(\psi) , \tag{6.26}$$

which shows that $Q(\psi, \chi_n)$ is a linear functional on (Φ^z, Q^z). So $Q(\psi, \chi_n) = Q^z(\psi, h_n)$ for some $h_n \in \Phi^z$, which implies, in accordance with (4.11), that $\chi_n = KK^*h_n$ $(Q^x \equiv Q)$.

Now from (6.24) it follows that $(Q - Q^+)(\varphi, \varphi) = Q^z(A\varphi, \varphi)$, where A is a bounded operator in Φ^z. Consequently, if $\varphi_n \to \varphi$, where $\varphi_n \in \Phi$ and $\varphi \in \Phi^z$, there exist a limit

$$\widehat{Q - Q^+}(\varphi, \varphi) = \lim_{n \to \infty} (Q - Q^+)(\varphi_n, \varphi_n) = Q^z(A\varphi, \varphi) , \tag{6.27}$$

which represents a unique continuous extention of $Q - Q^+$ on $\Phi^z \times \Phi^z$. Because

$$(Q - Q^+)(\chi_n, \chi_n) = Q((1 - \lambda_n)\chi_n, \chi_n) = Q^z((1 - \lambda_n)\chi_n, h_n) ,$$

we conclude that $A\chi_n = (1 - \lambda_n) h_n$. So, on developing $Q - Q^+(\varphi, \varphi)$ into a series, we get

$$\widehat{Q - Q^+}(\varphi, \varphi) = \sum_1^\infty |Q(\varphi, \chi_n)|^2 (1 - \lambda_n) = \tag{6.28}$$

$$= \sum_1^\infty \left| \widehat{Q - Q^+}\left(\varphi, \frac{\chi_n}{1 - \lambda_n}\right)\right|^2 (1 - \lambda_n) = \sum_1^\infty |Q^z(\varphi, h_n)|^2 (1 - \lambda_n) .$$

Random variables v_n satisfying $\chi_n(t) = R(x_t, v_n)$ are eigen-elements of B in $\mathscr{X}$. Because $\chi_n = KK^*h_n$, from Theorem 4.2 or 4.1 we have

$$v_n(\omega) = Q^z(x^\omega, h_n)\,. \tag{6.29}$$

On comparing (6.25) and (6.29), we get $\widehat{Q - Q^+}(x^\omega, x^\omega) = \sum_1^\infty v_n^2(\omega)(1 - \lambda_n)$. Returning to the equation (6.5) and noting that $\det B = \prod_1^\infty \lambda_n$, we see that the proof is completed.

7. Probability densities for stationary Gaussian processes. We begin with the case of an integral-valued parameter $t = 0, 1, \ldots, N$, which is easier. The probability densities may be taken with respect to Lebesgue measure, because the number of random variables is finite.

Theorem 7.1. (i) *The probability density of a finite part $\{x_t, 0 \leqq t \leqq N\}$ of a Gaussian stationary process with vanishing mean values and covariances generated by the spectral density* (5.35) *is given by*

$$p(x_0, \ldots, x_N) = |Q_{ts}|^{-1/2} \exp\left(-\tfrac{1}{2}\sum_{t=0}^{N}\sum_{t=0}^{N} Q_{ts}x_t x_s\right) = \tag{7.1}$$
$$= |D_{jk}|^{1/2} a_0^{N-n+1} \exp\left[-\tfrac{1}{2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} D_{jk}x_j x_k - \tfrac{1}{2}\sum_{t=n}^{N}\left(\sum_{s=0}^{n} a_k x_{t-k}\right)^2\right],$$

where Q_{ts}, $0 \leqq t, s \leqq N$, and D_{jk}, $0 \leqq j, k \leqq n - 1$ are given by (5.37) *and* (5.40), *respectively, and $|.|$ denotes the determinant.*

(ii) *The probability density of a finite part $\{z_t, 0 \leqq t \leqq N\}$ of a stationary Gaussian process with vanishing mean values and covariances generated by the spectral density* (5.41) *is given by*

$$p(z_0, \ldots, z_N) = |D_{jk}|^{1/2}\, |R(x_i, x_h | z_0, \ldots, z_N)|^{1/2}\, a_0^{N+m+1-n} b_0^{-N-1}\,. \tag{7.2}$$
$$.\exp\left[-\tfrac{1}{2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} D_{jk}\check{x}_{j-m}\check{x}_{k-m} - \tfrac{1}{2}\sum_{t=n-m}^{N}\left(\sum_{k=0}^{n} a_k \check{x}_{t-k}\right)^2\right],$$

where $x_{-m}, \ldots, x_N$ is a process considered in (i) *and such that $z_t = \sum b_k x_{t-k}$, and $R(x_i, x_h \mid z_0, \ldots, z_N)$, $-m \leqq i, h \leqq -1$, are conditional (partial) covariances of x_i, x_h, when $z_0, \ldots, z_N$ are fixed. Moreover, $\check{x}_{-m}, \ldots, \check{x}_N$ is the solution of equations $z_t = \sum b_k \psi_{t-k}$ satisfying the conditions*

$$\sum_{-m}^{N}\sum_{-m}^{N} Q_{ts}\psi_t h_s = 0$$

for some m linearly independet solutions of $\sum b_k h_{t-k} = 0$.

Proof. (i) As Q_{ts} given by (5.37) represent elements of inverted covariance matrix $R = (R_{ts})$, the first expression for $p(x_0, \ldots, x_N)$ in (7.1) is clear. If we put $R_n = (R(x_i, x_j))_{i,j=0}^{n-1}$, then, in accordance with Theorem 6.2,

$$|Q_{ts}| = |R_n|^{-1}\left[\prod_{t=n}^{N} R^2(x_t - I_{t-1}x_t)\right]^{-1}, \tag{7.3}$$

where $I_{t-1}x_t$ is the projection of x_t on the subspace spanned by the random variables $x_0, \ldots, x_{t-1}$. However, we have $|R_n|^{-1} = |D_{jk}|$ and $R^2(x_t - I_{t-1}x_t) = a_0^{-2}$, as $y_t = a_0(x_t - I_{t-1}x_t)$, where y_t is given by (5.36). Thus we obtain the determinant of the second expression for $p(x_0, \ldots, x_N)$ in (7.1). The quadratic form of the second expression is a mere transcription of $\sum\sum Q_{ts}x_t x_s$, and is based on the fact that the y_t's given by (5.36) are independent of each other and of $x_0, \ldots, x_{n-1}$.

(ii) First we derive the determinant. Put $z_t = x_t$, if $-m \leqq t \leqq -1$, and $z_t = \sum b_k x_{t-k}$, if $0 \leqq t \leqq N$. This transformation may be denoted in the matrix form as $z = Cx$, where the matrix $C = \{c_{ij}\}$ is such that $c_{ij} = 0$, $i < j$, and $c_{ii} = 1$, if $-m \leqq i \leqq 1$, and $c_{ii} = b_0$, if $0 \leqq i \leqq N$. Consequently $|C| = b_0^{N+1}$ and

$$|R(z_t, z_s)| = b_0^{2N+2}|R(x_t, x_s)|\,, \quad -m \leqq t,\ s \leqq N\,.$$

Now we know from (i) that

$$|R(x_t, x_s)| = |D_{jk}|^{-1}\, a_0^{-2(N+m+1-n)}$$

which gives

$$(7.4) \qquad |R(z_t, z_s)| = |D_{jk}|^{-1}\, a_0^{-2(N+m+1-n)}\, b_0^{2(N+1)} \quad (-m \leqq t,\ s \leqq N)\,.$$

Finally, according to Theorem 6.2,

$$(7.5) \qquad |R(z_t, z_s)| = |R(z_{t'}, z_{s'})|\, |R(x_i, x_h \mid z_0, \ldots, z_N)|\,. \\ .\,(-m \leqq t,\ s \leqq N;\ 0 \leqq t', s' \leqq N;\ -m \leqq i, h \leqq -1)\,,$$

where $x_i = z_i$, if $-m \leqq i \leqq -1$. On combining (7.4) and (7.5), we get for $|R(z_{t'}, z_{s'})|$, $0 \leqq t', s' \leqq N$, the expression appearing in (7.2). The quadratic form in (7.2) follows from the form of $Q^z(v, \chi)$ described below equation (5.42).

Now we proceed to the case of a continuous t, and first consider the n-th order Markovian processes. The probability density cannot be taken with respect to Lebesgue measure, because the system of random variables is infinite. The dominating distribution $\mathsf{P}^+ = \mathsf{P}^+_{n,a}$ of $\{x_t, 0 \leqq t \leqq T\}$ used in the next theorem will be defined by the following conditions:

(7.6) the vector $(x_0, x_0', \ldots, x_0^{(n-1)})$ is distributed according to n-dimensional Lebesgue measure; $x_t^{(n-1)}$ is a Gaussian process with independent increments such that

$$\mathsf{E}\,|\mathrm{d}x_t^{(n-1)}|^2 = a^{-2}\,\mathrm{d}t\,;$$

$x_t^{(n-1)} - x_0^{(n-1)}$ is independent of $(x_0, \ldots, x_0^{(n-1)})$; the mean values vanish.

Theorem 13.1. *Let $\{x_t, 0 \leqq t \leqq T\}$ be a finite segment of a stationary Gaussian process with vanishing mean values and with covariances generated by spectral density (5.1). Then the distribution of $\{x_t, 0 \leqq t \leqq T\}$, say P, is strongly equivalent to $\mathsf{P}^+ = \mathsf{P}^+_{n,a_0}$ defined by (7.6), and*

(7.7)
$$\frac{\mathrm{dP}}{\mathrm{dP}^+} = |D_{jk}|^{1/2} \exp\left\{\frac{1}{2}\frac{a_1}{a_0} T - \frac{1}{2}\sum_{k=0}^{n-1} A_{n-k} \int_0^T |x_t^{(k)}(\omega)|^2 \,\mathrm{d}t - \right.$$
$$\left. - \frac{1}{4}\sum_{j+k}^{n-1} \sum_{\text{even}}^{n-1} [x_T^{(j)} x_T^{(k)} + x_0^{(j)} x_0^{(k)}]\, D_{jk}\right\},$$

where $D_{j,k}$, $0 \leqq j, k \leqq n - 1$ are given by (5.21) and A_{n-k} by (5.9).

Proof. Let (Φ^+, Q^+) be the closed linear manifold generated by covariances $R^+(x_t, x_s)$ corresponding to the P^+_{n,a_0} – distribution. It is easy to see that $\Phi^+ = \mathscr{L}^2_n$ and

(7.8)
$$Q^+(\psi, \varphi) = Q^+_{n,a_0}(\psi, \varphi) = a_0^2 \int_0^T \psi_t^{(n)} \bar{\varphi}_t^{(n)} \,\mathrm{d}t \quad (\psi, \varphi \in \mathscr{L}^2_n).$$

Consequently, for Q given by (5.11), where we have made use of (5.21), we get

(7.9)
$$(Q - Q^+)(\psi, \psi) = \sum_{k=0}^{n-1} A_{n-k} \int_0^T |\psi_t^{(k)}|^2 \,\mathrm{d}t +$$
$$+ \frac{1}{2}\sum_{j=0}^{n-1} \sum_{k=0}^{n-1} [\psi_T^{(j)} \bar{\psi}_T^{(k)} + \psi_0^{(j)} \bar{\psi}_0^{(k)}]\, D_{jk}\,.$$

Now let α_1 be a root of $L(\lambda) = \sum a_{n-k}\lambda^k = 0$, and let $L_1(\lambda) = L(\lambda)/(\lambda - \alpha_1)$. Then $z_t = x' + \alpha_1 x_t$ is a $(n - 1)$-th order Markovian process, if $n > 1$, and a white noise, if $n = 1$. We have

$$x_t = x_0 e^{\alpha_1 t} + e^{\alpha_1 t} \int_0^t e^{-\alpha_1 t} z_s \,\mathrm{d}s = K z_t\,,$$

where K^*K has a finite spure. Obviously Q^z satisfies (6.24). Consequently, according to Theorem 6.3, P and $\mathrm{P}^+ = \mathrm{P}^+_{n,a_0}$ are strongly equivalent. Since $\Phi^z = \mathscr{L}^2_{n-1}$, $Q - Q^+$ may be extended to $\mathscr{L}^2_{n-1}$. The right side of (7.9) is, however, adjusted so that it directly represents the extention of $Q - Q^+$ to $\mathscr{L}^2_{n-1}$. Now, in view of (6.25), we only have to substitute x_t for ψ_t in (7.9), which yields he quadratic form of (7.7).

It is now necessary to find the determinant. The determinant corresponding to the vector $(x_0, \ldots, x_0^{(n-1)})$ equals $|D_{jk}|$, D_{jk} given by (5.21). This determinant is to be multiplied by the determinant of the operator B such that

(7.10)
$$R(x_t^{(n-1)}, x_s^{(n-1)} \mid x_0, \ldots, x_0^{(n-1)}) = R^+(Bx_t^{(n-1)}, x_s^{(n-1)} \mid x_0, \ldots, x_0^{(n-1)})\,.$$

Let B_N be the restriction of B ot the subspace spanned by random variables

(7.11)
$$u_i = x_{iT/N}^{(n-1)}\,, \quad (i = 1, \ldots, N)\,.$$

According to Theorem 6.2, we have

(7.12)
$$\det B_N = \prod_{i=1}^{N} \frac{R(u_i - u_i^0, u_i - u_i^0)}{R^+(u_i - u_i^+, u_i - u_i^+)},$$

where u_i^0 and u_i^+ are the projections of u_i on the subspace spanned by $(x_0, \ldots, x_0^{(n-1)},$

$u_1, \ldots, u_{i-1})$ with respect to the R-covariance and R^+-covariance, respectively. According to the assumptions (7.6), $u_i^+ = u_{i-1}$ and

$$R^+(u_i - u_i^+, u_i - u_i^+) = \frac{T}{N} a_0^{-2}. \tag{7.13}$$

If R-covariance holds true, the situation is more complicated. First we find the projection of u_i, say $\bar{u}_i$, on the subspace spanned by $\{x_t, 0 \leqq t \leqq [(i-1)/N]\, T\}$. Because the process $\{x_{T-t}, 0 \leqq t \leqq T\}$ has the same distribution as $\{x_t, 0 \leqq t \leqq T\}$, we may write (5.10) in the following equivalent form:

in the following equivalent form:

$$Q(\psi, \varphi) = \int_0^T (L^*\bar{\varphi})_t\, (L^*\psi)_t\, \mathrm{d}t + \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \psi_T^{(j)} \bar{\varphi}_T^{(k)} D_{jk}\,, \tag{7.14}$$

where D_{jk} is given by (5.21). When looking for $\bar{u}_i$ we have to put $T \equiv [(i-1)/N] \,.\, T$ and $\varphi_t = R(x_t, u_i)$. However,

$$(L^*\varphi)_t = L^*R(x_t, u_i) = R(L^*x_t, u_i) = 0\,, \quad 0 \leqq t \leqq [(i-1)/N] \,.\, T\,, \tag{7.15}$$

because L^*x_t is a white noise independent of x_s, $s > t$, similarly as Lx_t was a white noise independent of x_s, $s < t$. This means that for $\varphi_t = R(x_t, u_i)$

$$Q(\psi, \varphi) = \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \psi_{(i-1)T/N}^{(j)} \bar{\varphi}_{(i-1)T/N}^{(k)} D_{jk}\,.$$

If we replace ψ_t by x_t and substitute

$$\varphi_t^{(k)} = \frac{\partial^k}{\partial t^k} R(x_t, x_{iT/N}^{(n-1)}) = (-1)^k R_{t-iT/N}^{(n-1+k)}\,,$$

where $R_{t-s} = R_{ts}$, we get

$$\begin{aligned} \bar{u}_i &= \sum_{j=0}^{n-1} x_{(i-1)T/N}^{(j)} \sum_{k=0}^{n-1} (-1)^k R_{T/N}^{(n-1+k)} D_{jk} = \\ &= \sum_{j=0}^{n-1} x_{(i-1)T/N}^{(j)} \sum_{k=0}^{n-1} (-1)^k \left[R_0^{(n-1+k)} + \frac{T}{N} R_{0+}^{(n+k)} + O(N^{-2}) \right] D_{jk}\,. \end{aligned} \tag{7.16}$$

Now

$$\begin{aligned} \sum_{k=0}^{n-1} (-1)^k R_0^{(n-1+k)} D_{jk} &= 1\,, \quad \text{if} \quad j = n-1\,, \\ &= 0\,, \quad \text{if} \quad h < n-1\,, \end{aligned} \tag{7.17}$$

because $((-1)^k R_0^{(j+k)})$ is the covariance matrix of the vector $(x_t, \ldots, x_t^{(n-1)})$, and (D_{jk}) is its inverse. Moreover, in view of $L^*{}_t R_{s-t}^{(k)} = 0$, $t < s$, we have

$$R_{0+}^{(n+k)} = -\frac{1}{a_0} \sum_{h=0}^{n-1} a_{n-k} R_{0+}^{(h+k)} \quad (0 \leqq k \leqq n) \tag{7.18}$$

which, in connection with (7.17) gives

(7.19) $$\sum_{k=0}^{n-1}(-1)^k R_{0+}^{(n+k)} D_{jk} = -\frac{a_{n-j}}{a_0}, \quad 0 \leqq j \leqq n-1\,.$$

If we insert (7.17) and (7.19) into (7.16), we get

(7.20) $$\bar{u}_i = u_{i-1} - \frac{T}{N}\sum_{j=0}^{n-1}\frac{a_{n-j}}{a_0}\, x_{(i-1)T/N}^{(j)} + 0(N^{-2})\,.$$

From (7.20) it follows that

(7.21) $$R(u_i - \bar{u}_i,\, u_i - \bar{u}_i) =$$

$$= 2(1-(-1)^n R_{T/N}^{(2n-2)}) + \frac{2T}{N}\sum_{j=0}^{n-1}\frac{a_{n-j}}{a_0}\,.\,(-1)^j(R_{T/N}^{(n-1+j)} - R_0^{(n-1+j)}) +$$
$$+ \frac{T^2}{N^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\frac{a_{n-j}a_{n-k}}{a_0^2}(-1)^j R_{0+}^{(j+k)} + 0(N^{-3}) =$$
$$= 2(-1)^n\frac{T}{N}R_{0+}^{(2n-1)} + (-1)^n\frac{T^2}{N^2}R_{0+}^{(2n)} + \frac{2T^2}{N^2}\sum_{j=0}^{n-1}\frac{a_{n-j}}{a_0}(-1)^j R_{0+}^{(n+j)} +$$
$$+ \frac{T^2}{N^2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1}\frac{a_{n-j}a_{n-k}}{a_0^2}(-1)^j R_{0+}^{(j+k)} + 0(N^{-3}) =$$
$$= 2(-1)^n\frac{T}{N}R_{0+}^{(2n-1)} + \frac{T^2}{N^2}\sum_{j=0}^{n-1}\frac{a_{n-j}}{a_0}(-1)^j R_{0+}^{(n+j)} + 0(N^{-3}) =$$
$$= 2(-1)^n\frac{T}{N}R_{0+}^{(2n-1)} - 2(-1)^n\frac{T^2}{N^2}\frac{a_1}{a_0}R_{0+}^{(2n-1)} + 0(N^{-3}) =$$
$$= \frac{T}{N}a_0^{-2}\left(1 - \frac{T}{N}\frac{a_1}{a_0}\right) + 0(N^{-3}),$$

where we have made use of (7.18), and of the adjoint relation

$$(-1)^n R_{0-}^{(n+k)} = -\sum_{k=0}^{n-1}(-1)^h\frac{a_{n-h}}{a_0}R_{0-}^{(h+k)},$$

together with $R_{0\pm}^{(2j+1)} = 0$, $j < n-1$, $R_{0+}^{(2j)} = R_{0-}^{(2j)} = R_0^{(2j)}$, $0 \leqq j \leqq n-1$, and

$$R_{0+}^{(2n-1)} = -R_{0-}^{(2n-1)} = (-1)^n\tfrac{1}{2}a_0^{-2}\,.$$

If we chose the projection u_i^0 of u_i on the subspace spanned by $(x_0, \ldots, x_0^{(n-1)}, u_1, \ldots$ $\ldots, u_{i-1})$, we cannot make use of random variables $x_{(i-1)T/N}^{(j)}$, $0 \leqq j \leqq n-2$, appearing in (7.20). However, we may approximate them by sums

$$\hat{x}_{(i-1)T/N}^{(j)} = \sum_{j=0}^{n-1}x_0^{(j)}c_j + \sum_{k=1}^{i-1}u_k\, d_k,$$

so that

$$R^2(x_{(i-1)T/N}^{(j)} - \hat{x}_{(i-1)T/N}^{(j)}) = 0(N^{-1})\,.$$

For example we may put

$$\hat{x}^{(n-2)}_{(i-1)T/N} = x_0^{(n-2)} + \sum_{k=1}^{i-1} u_k \frac{T}{N}$$

etc. Consequently, we have

$$(7.22) \qquad R(u_i - u_i^0, u_i - u_i^0) = \frac{T}{N} a_0^{-2}\left(1 - \frac{T}{N}\frac{a_1}{a_0}\right) + 0(N^{-3}).$$

So, in accordance with Theorem 6.1 and with (7.13) and (7.22),

$$(7.23) \quad \det B = \lim_{\substack{n\to\infty \\ N=2^n}} \det B_N = \lim_{\substack{n\to\infty \\ N=2^n}} \prod_{i=1}^{N} \frac{\frac{T}{N} a_0^{-2}\left(1 - \frac{T}{N}\frac{a_1}{a_0}\right) + 0(N^{-3})}{\frac{T}{N} a_0^{-2}} = e^{-a_1/a_0 T},$$

which concludes the proof.

In the case of a general rational spectral density the leading term of $Q^z(\chi, \chi)$ is

$$a_0^2 b_0^{-2} \int_0^T |\chi_t^{(n-m)}|^2 \, \mathrm{d}t = Q^+(\chi, \chi),$$

$$(7.24) \qquad Q^z(\chi, \chi) = a_0^2 b_0^{-2} \int_0^T |\chi_t^{(n-m)}|^2 \, \mathrm{d}t + \dots$$

and $Q^z(\chi, \chi) - Q^+(\chi, \chi)$ is dominated in the sense of (6.29) by $\bar{Q}(\chi, \chi)$ corresponding to any rational spectral density with $\bar{n} - \bar{m} = n - m - 1$. Without entering into details let us present the following

Theorem 7.3. *Let $\{z_t, 0 \leqq t \leqq T\}$ be a finite segment of a stationary Gaussian process with vanishing mean values and with covariances generated by spectral density (5.22). Then the distribution of $\{z_t, 0 \leqq t \leqq T\}$, say P, is strongly equivalent to $\mathsf{P}^+ = \mathsf{P}^+_{n-m, a_0/b_0}$ and*

$$(7.25) \qquad \frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\mathsf{P}^+} = |D_{jk}|^{1/2} b_0^{m-n} |R(x_0^{(i)}, x_0^{(h)} | z_t, 0 \leqq t \leqq 1)|^{1/2} \, .$$

$$. \exp\left\{\frac{1}{2}\left(\frac{a_1}{a_0} - \frac{b_1}{b_0}\right) T - \tfrac{1}{2}\widehat{Q^z - Q^+(x^\omega, x^\omega)}\right\}.$$

where $Q^+ = Q^+_{n-m, a_0/b_0}$ is given by (7.8), and the D_{jk}'s, $0 \leqq j, k \leqq n-1$, are given by (5.21). Further, x_t is a process considered in Theorem 7.2 such that $z_t = \sum_{k=0}^{m} b_{m-k} x_t^{(k)}$, and $R(x_0^{(i)}, x_0^{(h)} \mid z_t, 0 \leqq t \leqq T)$, $0 \leqq i, h \leqq m - 1$, are conditional (partial) covariances of $x_0^{(i)}$ and $x_0^{(h)}$, when $\{z_t, 0 \leqq t \leqq T\}$ is fixed.

Proof. The assertions concerning the quadratic form follow from Theorems 5.2

and 6.3 and from the above discussion. As for the determinant, let us go back to formula (7.7). If the dominating distribution P^+ were modified so that

$$\sum_{k=0}^{m} b_{m-k} x_t^{(m-m-1+k)} = Y_t$$

has independent increments and $E^+|dY_t|^2 = a_0^2 b_0^{-2}$, then the only change in the determinant would consist in replacing

$$\exp\left\{\frac{1}{2}\frac{a_1}{a_0} T\right\} \quad \text{by} \quad \exp\left\{\frac{1}{2}\left(\frac{a_1}{a_0} - \frac{b_1}{b_0}\right) T\right\}.$$

This could be proved by arguments similar to those used in the proof of Theorem 7.2. Now, if we put $z_t = \sum_{k=0}^{m} b_{m-k} x_t^{(k)}$, we can easily see that P^+ – distribution of z_t is the one used in (7.25). Further the transformation $(z_0^{(-m)}, \ldots, z^{(n-m-1)}) = A(x_0, \ldots, x_0^{(n-1)})$, defined by

(7.26)
$$z_0^{(j)} = \sum_{k=0}^{m} b_{m-k} x_0^{(j+k)}, \quad \text{if} \quad 0 \leqq j \leqq n-m-1,$$
$$= x_0^{(j+m)} \qquad \text{if} \quad -m \leqq j \leqq -1$$

has the determinant $|A| = b_0^{n-m}$. Now replacing x_0 by z_0 amounts to multiplying the determinant by b_0^{n-m}, and excluding random variables $x_0^{(j+m)} = z_0^{(j)}$, $-m \leqq j \leqq -1$, amounts to dividing the determinant by the factor $|R(x_0^{(i)}, x_0^{(h)} \mid z_t, 0 \leqq t \leqq T)|$, which follows from Theorem 6.2. The proof is finished.

Remark 7.1. Obviously, the following limit exists:

(7.27)
$$\lim_{T\to\infty} |R(x_0^{(i)}, x_0^{(h)} \mid z_t, 0 \leqq t \leqq T)| = |R(x_0^{(i)}, x_0^{(h)} \mid z_t, 0 \leqq t < \infty)|.$$

Remark 7.2. Consider a stationary Gaussian process $\{x_t, 0 \leqq t \leqq T\}$ with correlation function $R(\tau) = \max(0, 1 - |\tau|)$. From Example 3.2 it follows, after some computations, that, for $0 < T < 1$,

(7.28)
$$Q(\varphi, \varphi) = \frac{1}{2}\frac{(\varphi_0 + \varphi_T)^2}{2 - T} + \frac{1}{2}\int_0^T |\varphi_t'|^2 \, dt.$$

Consequently, the respective distribution, say P, is strongly equivalent to $P^+ = P_{1,2}^+$ defined by (7.6), and

(7.29)
$$\frac{dP}{dP^+} = \text{const} \exp\left[-\frac{1}{4}\frac{(x_0 + x_T)^2}{2 - T}\right].$$

If, however, $1 < T < 2$, we get

(7.30)
$$Q(\varphi, \varphi) = \frac{1}{6}\frac{(2\varphi_0 + 2\varphi_T + \varphi_1 + \varphi_{T-1})^2}{4 - T} + \frac{1}{2}\int_{T-1}^{1} |\varphi_t'|^2 \, dt +$$
$$+ \frac{2}{3}\int_1^T (|\varphi_t'|^2 + |\varphi_{t-1}'|^2 + \varphi_t'\varphi_{t-1}')^2 \, dt,$$

which shows that the distribution is not equivalent to any distribution we have met with.

If we introduce two parameters by puting $R(\tau) = d^2.\ \max(0, 1 - |\tau|/a)$, then for $0 < T < a$ and $P^+ = P^+_{1,2d^2/a}$,

$$\frac{dP}{dP^+} = \left(\frac{2a - T}{2a}\right)^{1/2} . \exp\left[-\frac{1}{4}\frac{(x_0 + x_T)^2}{d^2(2 - T/a)}\right]. \tag{7.31}$$

The determinant was established as follows:

If we put $\varphi_t = R(s + \Delta - t) = d^2(1 - (s + \Delta - t)/a)$, $0 \leqq t \leqq s$, then

$$Q(\varphi, \varphi) = d^2(1 - 2\Delta a^{-1} + 2\Delta^2 a^{-1}(2a - s)^{-1})$$

gives the variance of the projection of $x_{s+\Delta}$ on the subspace spanned by $\{x_t, 0 \leqq t \leqq s\}$, so that the residual variance equals $\Delta 2d^2 a^{-1}(1 - \Delta(2a - s)^{-1})$. Then we may proceed with a development similar to that used in evaluating of (7.23). Note that $x_0 + x_T$ is a sufficient statistic for estimating a. As is well-known

$$2d^2 a^{-1} = \underset{n\to\infty}{\text{l. i. m.}} \sum_{i=1}^{n} (x_{(i-1)T/n} - X_{iT/n})^2 .$$

Remark 7.3. From (7.7) it follows that the vector

$$\left\{\int_0^T |x_t|^2 \, dt, \ldots, \int_0^T |x_t^{(n-1)}|^2 \, dt, x_0, \ldots, x_0^{(n-1)}, x_T, \ldots, x_T^{(n-1)}\right\}$$

represents a sufficient statistics for all n-th order Markovian processes with fixed a_0. In the case of a general rational spectral density with $m > 0$ apparently no sufficient statistic exists, which would not be equivalent to the whole process $\{z_t, 0 \leqq t \leqq 1\}$. See Example 7.3.

Remark 7.4. We have proved, by the way, that distributions P_1 and P_2 corresponding to two rational spectral densities are strongly equivalent, if $n_1 - m_1 = n_2 - m_2$ and $a_{01}b_{02} = a_{02}b_{01}$, and perpendicular in other cases. This result (with equivalence instead of strong equivalence) has been announced by V. F. Pisarenko [24].

Example 7.1. If $n = 1$, we get

$$R(\tau) = (2a_0 a_1)^{-1} e^{-(a_1/a_0)|\tau|},$$

and

$$\frac{dP}{dP^+} = (2a_0 a_1)^{1/2} \exp\left\{\frac{1}{2}\frac{a_1}{a_0} T - \frac{1}{2}a_1^2 \int_0^T |x_t|^2 \, dt\right\}, \tag{7.32}$$

where $x_t = x_t(\omega)$. See also [25].

Example 7.2. If

$$f(\lambda) = \frac{1}{2\pi}\frac{a_0^{-2}}{(\lambda^2 + \alpha_1^2)(\lambda^2 + a_2^2)} = \frac{1}{2\pi}\frac{a_0^{-2}}{|(i\lambda)^2 + (i\lambda)(\alpha_1 + \alpha_2) + \alpha_1\alpha_2|^2}, \tag{7.33}$$

$$(\alpha_2 > \alpha_1 > 0)$$

then $a_1/a_0 = \alpha_1 + \alpha_2$, $a_2/a_0 = \alpha_1\alpha_2$. Consequently,

$$R(\tau) = \frac{1}{2a_0^2}\left(\frac{1}{\alpha_1}e^{-\alpha_1|\tau|} - \frac{1}{\alpha_2}e^{-\alpha_2|\tau|}\right), \tag{7.34}$$

$D_{12} = 0$, $D_{00} = 2a_2a_1 = 2(\alpha_1 + \alpha_2)\,\alpha_1\alpha_2a_0^2$, $D_{11} = 2a_0a_1 = 2(\alpha_1 + \alpha_2)\,a_0^2$, and

$$\frac{\mathrm{dP}}{\mathrm{dP}^+} = 2(\alpha_1 + \alpha_2)\,(\alpha_1\alpha_2)^{1/2}\,a_0^2 \exp\left[\tfrac{1}{2}(\alpha_1 + \alpha_2)\,T)\right]. \tag{7.35}$$

$$. \exp\{\tfrac{1}{2}(\alpha_1^2 + \alpha_2^2)\,a_0^2\int_0^T |x_t'|^2\,dt - \tfrac{1}{2}\alpha_1^2\alpha_2^2a_0^2\int_0^T |x_t|^2\,dt +$$

$$+ \tfrac{1}{2}(x_0'^2 + x_T'^2)\,(\alpha_1 + \alpha_2)\,a_0^2 - \tfrac{1}{2}(x_0^2 + x_T^2)\,(\alpha_1 + \alpha_2)\,\alpha_1\alpha_2a_0^2\}\,.$$

Example 7.3. Consider a general spectral density with $n = 2$ and $m = 1$,

$$g(\lambda) = \frac{b_0^2(\lambda^2 + \beta^2)}{|a_0(i\lambda)^2 + a_1(i\lambda) + a_2|^2} \tag{7.36}$$

and put $L(\beta) = a_0\beta^2 + a_1\beta + a_2$, $L^*(\beta) = a_0\beta^2 - a_1\beta + a_2$ and $LL^*(\beta) = L(\beta)L^*(\varepsilon) = a_0^2\beta^4 + (2a_0a_2 - a_1^2)\,\beta^2 + a_2^2$, etc. In order to find $Q^z(\chi, \chi)$, let us first suppose that $\chi \in \mathscr{L}_2^2$, and use the form of $Q^z(\chi, \chi)$ resulting from the right side of (5.27) after substituting χ_t for $z_t(\omega)$. On obtaining φ_t from formula (5.31), we get

$$\varphi_t = \frac{1}{2\beta}\int_0^T e^{-|t-s|\beta}\,\chi_s\,\mathrm{d}s + c_1e^{-t\beta} + c_2e^{(\tau-t)\beta}\,, \tag{7.37}$$

where

(7.38)

$$c_1 = \frac{a_0L(\beta)\,(\chi_0 - L^*(\beta)\,\frac{1}{2\beta}\int_0^T e^{-t\beta}\chi_t\,\mathrm{d}t) - L^*(\beta)\,e^{-T\beta}(\chi_T - L^*(\beta)\,\frac{1}{2\beta}\int_0^T e^{-(T-t)\beta}\chi_t\,\mathrm{d}t)}{LL(\beta) - L^*L^*(\beta)\,e^{-2T\beta}}$$

and

(7.39) $c_2 =$

$$= \frac{a_0L(\beta)\,(\chi_T - L^*(\beta)\frac{1}{2\beta}\int_0^T e^{-(T-t)\beta}\chi_t\,\mathrm{d}t) - L^*(\beta)\,e^{-T\beta}(\chi_0 - L^*(\beta)\frac{1}{2\beta}\int_0^T e^{-t\beta}\,\chi_t\,\mathrm{d}t)}{LL(\beta) - L^*L^*(\beta)\,e^{-2T\beta}}.$$

Now in view of (5.29) and (5.30) $(L\varphi)_T = (L^*\varphi)_0 = 0$, so that

$$Q^z(\chi, \chi) = \int_0^T \chi_t(LL^*\varphi)_t\,\mathrm{d}t + \chi_Ta_0[a_0\varphi_T''' + a_1\varphi_T'' + a_2\varphi_T'] +$$

$$+ \chi_0a_0[a_0\varphi_0''' - a_1\varphi_0'' + a_2\varphi_0']\,.$$

On using the relations $\varphi_t'' = -\chi_t + \beta\varphi_t$ and (7.37), we get, after some transformations, that

$$(7.40)\quad Q^z(\chi,\chi) = \left(\frac{a_0}{b_0}\right)^2 \int_0^T |\chi_t'|^2\,dt + (a_1^2 - 2a_0a_2 - \beta^2 a_0^2)\, b_0^{-2} \int_0^T |\chi_t|^2\,dt +$$
$$+ \left(\frac{a_0}{b_0}\right)^2 LL^*(\beta) \int_0^T \int_0^T e^{-|t-s|\beta}\chi_t\chi_s\,dt\,ds +$$
$$+ \left(|\chi_0|^2 + |\chi_T|^2\right) b_0^{-2}\left(a_0a_1 - \beta a_0^2 \frac{LL(\beta) + L^*L^*(\beta)\,e^{-2T\beta}}{LL(\beta) - L^*L^*(\beta)\,e^{-2T\beta}}\right) -$$
$$- \tfrac{1}{2}\beta\left(\frac{a_0}{b_0}\right)^2\left(\left|\frac{1}{\beta}\int_0^T e^{-t\beta}\chi_t\,dt\right|^2 + \left|\frac{1}{\beta}\int_0^T e^{-(T-t)\beta}\chi_t\,dt\right|^2\right)\frac{LLL^*L^*(\beta)}{LL(\beta) - L^*L^*(\beta)\,e^{-2T\beta}} +$$
$$+ 2\left(\frac{a_0}{b_0}\right)^2\left(\chi_T\int_0^T e^{-(T-t)\beta}\chi_t\,dt + \chi_0\int_0^T e^{-t\beta}\chi_t\,dt\right)\frac{LLL^*(\beta)}{LL(\beta) - L^*L^*(\beta)\,e^{-2T\beta}} +$$
$$+ 4\beta\left(\frac{a_0}{b_0}\right)^2\left(\chi_0\chi_T + L^*L^*(\beta)\frac{1}{2\beta}\int_0^T e^{-t\beta}\chi_t\,dt\,\frac{1}{2\beta}\int_0^T e^{-(T-t)\beta}\chi_t\,dt\right).$$
$$\frac{LL^*(\beta)\,e^{-T\beta}}{LL(\beta) - L^*L^*(\beta)\,e^{-2T\beta}} - 2\left(\frac{a_0}{b_0}\right)^2\chi_0\int_0^T e^{-(T-t)\beta}\chi_t\,dt +$$
$$+ \chi_T\int_0^T e^{-t\beta}\chi_t\,dt\bigg)\frac{LL^*L^*(\beta)\,e^{-T\beta}}{LL(\beta) - L^*L^*(\beta)\,e^{-2T\beta}}.$$

If $T \to \infty$ the terms involving $e^{-T\beta}$ become negligible. The quadratic form (7.40), obviously, is well-defined for any $\chi \in \mathscr{L}_1^2$. On putting

$$(7.41)\qquad \widehat{Q^z - Q^+}(\chi,\chi) = Q^z(\chi,\chi) - \left(\frac{a_0}{b_0}\right)^2 \int_0^T |\chi_t'|^2\,dt\,,$$

$\widehat{Q^z - Q^+}$ will be well-defined for any $\chi \in \mathscr{L}_0^2$. The probability density will equal

$$(7.42)\quad \frac{dP}{dP^+} = 2b_0^{-1}a_1(a_0a_2)^{1/2}\,R(x_0\quad z_t,\ 0 \leqq t \leqq T)\,.\,\exp\left[\frac{1}{2}\left(\frac{a_1}{a_0} - \beta\right)T\right].$$
$$.\,\exp\left\{-\tfrac{1}{2}\widehat{Q^z - Q^+}(x^\omega, x^\omega)\right\},$$

where $\widehat{Q^z - Q^+}$ is given by (7.41) and (7.40). The conditional variance of x_0 equals

$$(7.43)\qquad R^2(x_0 \mid z_t,\ 0 \leqq t \leqq T) = R^2(x_0) - Q^z(v, v)\,,$$

where $R^2(x_0)$ is the absolute variance and $v_t = R(x_0, z_t) = b_0R(x_0, x_t' + \beta x_t) \equiv \equiv b_0[R'(t) + \beta R(t)]$. In the special case considered in Example 7.2, we have

$$(7.44)\quad R^2(x_0) = \frac{1}{2}\frac{|\alpha_1 - \alpha_2|}{a_0^2\alpha_1\alpha_2}\,,\qquad v_t = \frac{1}{2a_0^2}\left(e^{-\alpha_1 t}\frac{\beta - \alpha_1}{\alpha_1} - e^{-\alpha_2 t}\frac{\beta - \alpha_2}{\alpha_2}\right).$$

[1] *А. Н. Колмогоров:* Стационарные последовательности в гильбертовском пространстве. Бюлл. МГУ, *2*, № 6, 1941, 1—40.

[2] *N. Wiener:* Extrapolation, interpolation and smoothing of stationary time series. Cambridge-New York, 1949.

[3] *Kari Karhunen:* Über die lineare Methoden in der Wahrscheinlichkeitsrechnung. Ann. Acad. Sci. Fennicae, Ser. A, I. Math. Phys., *37*, 1947, 3—79.

[4] *U. Grenander:* Stochastic processes and statistical inference. Arkiv för Matematik, *1*, 1949, 195—277.

[5] *U. Grenander:* On empirical spectral analysis of stochastic processes. Arkiv för Matematik, *1*, 1952, 503—531.

[6] *U. Grenander* nad *M. Rosenblatt:* Statistical analysis of stationary time series. Stockholm, 1956.

[7] *L. A. Zadeh* and *R. Ragazzini:* Extension of Wiener's theory of prediction. Journ. Appl. Phys. *21*, 1950, 645—655.

[8] *C. L. Dolph* and *M. A. Woodbury:* On the relations between Green's functions and covariance of certain stochastic processes. TAMS, *72*, 1952, 519—550.

[9] *А. М. Яглом:* Экстраполирование, интерполирование и фильтрация стационарных случайных процессов с рациональной плотностю. Труды Моск. Мат. Общ., *4*, 1955, 333—374.

[10] *А. М. Яглом:* Явные формулы для экстраполяции, фильтрации и вычисления количества информации в теории гауссовскик стохастических процессов. Trans. Sec. Prague Conf. Inf. Th. etc., Praha 1960, 251—262.

[11] *И. М. Гельфанд* и *А. М. Яглом:* О вычислении количества информации о случайной функции, содержащейся в другой такой функции. Успехи мат. наук, *XII*, 1957, 3—52.

[12] *И. М. Гельфанд:* Обобщеные случайные процессы. ДАН 100, № *5* (1955), 853—856.

[13] *А. В. Скороход:* О дифференцируемости мер соответствующих случайным процессам. I. Процессы с независимими Приращениями. Теория вероятностей, *II*, 1957, 417—442. II. Марковские процессы, Теория вероятностей, *V*, 1960, 45—53.

[14] *А. В. Скороход:* Про одну задачу статистики гауссiвских процесiв. Доповiдi Академиiї наук УРСР, *9*, 1960, 1167—69.

[15] *Я. Гаек* (J. Hájek): Линейная оценка средней стационарного случайного процесса с выпуклой корреляционной функцией. Czech. Math. Journal, *6*, 1956, 94—117.

[16] *J. Hájek:* Predicting a stationary process when the correlation function is convex. Czech. Math. Journ., *8*, 1958, 150—154.

[17] *J. Hájek:* A property of *J*-divergences of marginal probability distributions. Czech. Math. Journ., *8*, 1958, 460—463.

[18] *Я. Гаек* (J. Hájek): Об одном свойстве нормальных распределений произвольного случайного процесса. Czech. Math. Journal, *8*, 1958, 610—618.

[19] *J. Hájek:* On a simple linear model in Gaussian processes. Trans. Second. Prague Conf. Inf. Th. etc., 1960, 185—197.

[20] *J. Hájek:* On linear estimation theory for an infinite number of observations. Теория вероятностей, *VI*, 1961, 182—193.

[21] *A. V. Balakrishnan:* On a charakterization of covariances. Ann. Math. Stat., *30*, 1959, 670—675.

[22] *G. Kallianpur:* A problem in optimum filtering with finite date. Ann. Math. Stat., *30*, 1959, 659—669.

[23] *Ли-Хен-Вон:* К теории отпимальной фильтрации сигнала при наличии внутренных шумов. Теория вероятностей, *IV*, 1959, 458—464.

[24] *В. Ф. Писаренко:* Об абсолютной непрерывности мер, соответствующих гауссовским процессам с рациональной спектральной плотностью. Теория вероятностей, *IV*, 1959, 481—481.

[25] *Charlotte T. Striebel:* Densities for stochastic processes. Ann. Math. Stat., *30*, 549—567, 1959.

[26] *М. С. Пинскер:* Энтропия, скорость создания энтропии и энтропийная устойчивость гауссовских случайных величин и процессов. ДАН СССР, *133*, № 3, 1960, 531—533.

[27] *В. С. Пугачев:* Эффективный метод нахождения Бейесова решения. Trans. Sec. Prague Conf. Inf. Th. etc., Praha, Nakl. ČSAV, 1959, 531—540.

[28] *В. С. Пугачев:* Теория случайных функций и ее применения к задачам автоматического управления, второе изд. Москва, 1960.

[29] *D. Middleton,* Introduction to statistical communication theory. London, 1960.

[30] *J. L. Doob:* Stochastic processes. New York-Toronto 1953, (Russian translation). Moskva, 1956.

[31] *E. A. Coddington* and *N. Levinson:* Theory of ordinary diferential equations. New York-Toronto-London 1955, Moskva (Russian translation) 1958.

[32] *F. Riesz* et *B. Sz.-Nagy:* Lecons d'analyse fonctionnelle. Budapest 1953, Moskva (Russian translation) 1954.

Резюме

О ЛИНЕЙНЫХ СТАТИСТИЧЕСКИХ ПРОБЛЕМАХ В СТОХАСТИЧЕСКИХ ПРОЦЕССАХ

ЯРОСЛАВ ГАЕК (Jaroslav Hájek), Прага

В работе развита объединенная теория линейных статистических проблем, в том числе предсказывания, фильтрации, оценок коэффициентов регрессии и определения плотности вероятности одной гауссовской меры по другой.

Пусть $\{x_t, t \in T\}$ — какой-нибудь стохастический процесс такой, что $E|x_t|^2 < \infty$ и $Ex_t = 0$, $t \in T$. Пусть $\mathscr{X}$ — пространство Гильберта, элементами которого являются линейные комбинации случайных величин x_t и их пределы по норме $\|x\| = [E|x|^2]^{1/2}$. Пространству $\mathscr{X}$ можно поставить в соответствие пространство Φ, элементами которого являются комплексные функции $\varphi_t(t \in T)$ такие, что для определенного элемента $v \in \mathfrak{X}$

$$(2.1) \qquad \varphi_t = Ex_t\bar{v} \quad (t \in T)\,.$$

Притом норма $[Q(\varphi, \varphi)]^{1/2}$ в пространстве Φ дана соотношением $Q(\varphi, \varphi) = E|v|^2$, где v — элемент $\mathscr{X}$, соответствующий функции φ в смысле уравнения (2.1). Связанные с процессом $\{x_t\}$ линейные проблемы состоят, вообще, в отыскании Φ и $Q(\varphi, \varphi)$ и в решении уравненич (2.1) относительно v.

В § 2 определяются основные свойства пространств $\mathscr{X}$ и Φ а также здесь формулируются и решаются в общем виде основные типы линейных проблем. § 4 посвящается паре процессов $\{x_t\}$ и $\{y_t\}$, связанных соотношением $x_t = Ky_t$,

где K – линейный оператор. Доказывается, что для гауссовских процессов $\{x_t\}$ можно осуществить в функциях вида $\psi_t = Ey_t\bar{v}$, $v \in y$, тогда и только тогда, если оператор K^*K – ядерный. В § 5 дается в явном виде решение линейных проблем для стационарных процессов с рациональной спектральной плотностью на конечном интервале.

В § 6 две гауссовские меры P и P^+ названы сильно эквивалентными, если линейный оператор, переводящий одну ковариантную функцию в другую, является ядерным (для эквивалентности достаточно, чтобы оператор был типа, Гильберта-Шмидта). Здесь выводятся две теоремы о поведении определителя этого оператора и достаточное условие для того, чтобы $\mathrm{d}P/\mathrm{d}P^+$ было функцией квадратичной формы, определенной на выборочных функциях. В § 7 эти общие теоремы применяются к стационарным процессам с рациональной спектральной плотностью или с функцией корреляции $R(\tau) = \max(0, 1 - |\tau|)$. Например, выводится, что для гауссовской меры P, индуцированной спектральной плотностью

$$f(\lambda) = \frac{1}{2\pi}\left|\sum_{k=0}^{n} a_{n-k}(i\lambda)^k\right|^{-2} = \frac{1}{2\pi}\left|\sum_{k=0}^{n} A_{n-k}\lambda^{2k}\right|^{-1}$$

имеет место равенство

$$\frac{\mathrm{d}P}{\mathrm{d}P^+} = |D_{jk}|^{1/2} \exp\left\{\frac{1}{2}\frac{a_1}{a_0}T - \frac{1}{2}\sum_{k=0}^{n-1} A_{n-k}\int_0^T |x_t^{(k)}(\omega)|^2\,\mathrm{d}t - \right.$$
$$\left. - \frac{1}{4}\sum_{j+k}^{n-1}\ \sum_{\text{четные}}^{n-1} [x_T^{(j)}x_T^{(k)} + x_0^{(j)}x_0^{(k)}]\, D_{jk}\right\},$$

где D_{jk} даны уравнением (5.21), $|D_{jk}|$ означает определитель, а распределение $P^+ = P^+_{n,a_0}$ определяется на стр. 433.