THE PREDICTION THEORY OF MULTIVARIATE STOCHASTIC PROCESSES

I. The Regularity Condition

BY

N. WIENER and P. MASANI[1]

Cambridge, Mass., U.S.A. and Bombay, India

Contents

1. Introduction

To advance further in the prediction of multivariate (or multiple) stochastic processes, we need the support of a general theory of such processes. It is natural to try to build this theory along the lines of Kolmogorov's important development of the theory of univariate (or simple) processes [5, 6].[2] This work was begun in 1941 by Zasuhin, who was able to announce some important results [18]. But even before Kolmogorov's work, Cramer [1] had obtained a fundamental theorem on the spectrum. Subsequently, Wiener [14–17], Doob [2] and Whittle [13] have studied multiple processes, but a general theory has not as yet been reached. For instance, no spectral

---

[1] This research was carried out at the Indian Statistical Institute, Calcutta, during 1955–56. Our sincere thanks are due to the authorities for the excellent facilities placed at our disposal, and to Dr. G. Kallianpur for valuable discussions and for an English translation of Zasuhin's paper [18].

[2] A similar development, but confined to processes with absolutely continuous spectra, was given independently by Wiener, cf. [14, p. 59].

characterisation of non-determinacy has been given, nor has a relation been established between the prediction error matrix and the spectrum. Also, there has been some doubt regarding the spectral criterion for *regular processes of full rank.*[1] This is the sort of process for which we would expect to have electrically realizable linear filters. Such a criterion was stated by Zasuhin [18, Theorem 3] but without proof. It was rediscovered by Wiener [15], but his derivation [17] is incomplete in that the case in which two components of his multiple process make a zero angle with one another is left out.[2] To cite another lacuna in the theory, although the Wold decomposition announced by Zasuhin [18, Theorem 1] is valid, Doob's derivation of the corresponding decomposition of the spectrum [2, pp. 597–598] seems to be insufficient. Other questions which suggest themselves also remain unanswered, cf. sec. 8.

In part I of this paper we shall complete the theory of multiple stationary stochastic processes in the discrete parameter case in several respects, establishing in particular a spectral condition for full rank. In the course of this proof, we shall find a connection between the prediction error matrix and the spectrum, thus obtaining a determinantal extension of an important identity of Szegö [9, Satz XII]. As corollaries we shall derive the spectral version of the Wold decomposition and the criterion for regularity with full rank, mentioned above.

We shall draw on the work of Cramer, Kolmogorov, Zasuhin, Doob, and the previous work of Wiener, but our treatment will depart from theirs in many ways. We shall make strong use of certain theorems on the boundary values of holomorphic functions of the Hardy class, which are due to Szegö [10] and to Paley and Wiener [7]. These theorems are recapitulated in Sec. 2. We shall then discuss the harmonic analysis of matrix-valued functions, which will be needed in studying multiple spectra, and also establish a determinantal extension of the well known logarithmic inequality, which will play an important role (Sec. 3). In Sec. 4 we will deal with Riemann-Stieltjes integration in which both integrand and integrator are matrix-valued. Sec. 5 will be devoted to the analysis of vector-valued random functions. This is needed, since a multiple S.P. is a one-parameter family of such functions, subject to somewhat unusual concepts of orhogonality and linearity. These preliminaries will occupy a large part of this paper. In Sec. 6 we shall turn to the time-scale (or non-spectral) analysis of multiple processes, and in Sec. 7 take up the spectral analysis, and estab-

[1] Cf. Sec. 6. Roughly speaking, a S.P. (stochastic process) is *regular*, if its "remote past" consists only of the zero-vector; it has *full rank*, if the component random functions of its "innovation" are linearly independent.

[2] WIENER's proof of the factorization of unitary matrix-valued functions, contained in the same paper, is also incomplete.

lish the conditions for non-determinacy and regularity with full rank (Theorems 7.10–7.12). In Sec. 8 we will mention some unsettled points in the theory.

The spectral distribution of a simple S.P. with discrete parameter may be regarded as being defined on the unit circle in the complex plane. The linear predictor for such a process is obtained by factoring the derivative of this distribution into an inner and an outer function (Wiener [14, 15]). In the multiple case the corresponding factorization is of a non-negative hermitian matrix-valued function. The non-commutativity of matrix multiplication makes this factorization much harder. An algorithm for affecting it will be given in part II of this paper, which will appear separately.

2. Boundary values of functions in the Hardy class

Throughout the sequel the symbols C, D_+, D_- will denote the sets $|z|=1$, $|z|<1$, $1<|z|\leq\infty$, respectively, of the extended complex plane.

2.1 Definition. *For $\delta>0$, the classes L_δ and H_δ are defined as follows:*

(a) *L_δ consists of all complex-valued measurable functions f on C for which*

$$\int_0^{2\pi}|f(e^{i\theta})|^\delta\,d\theta<\infty.$$

(b) *The* Hardy class *H_δ will consist of all complex valued holomorphic functions f on D_+ for which there exists a number M such that*

$$\int_0^{2\pi}|f r e^{i\theta})|^\delta\,d\theta\leq M<\infty,\quad 0<r<1.$$

Let $f\in L_1$ on C. Since its nth Fourier coefficient:

$$a_n=\frac{1}{2\pi}\int_0^{2\pi}e^{-ni\theta}f(e^{i\theta})\,d\theta$$

tends to 0 as $n\to\pm\infty$, it follows from the Cauchy-Hadamard formula that $\sum_0^\infty a_n z^n$ converges on D_+ and $\sum_1^\infty a_{-n}z^{-n}$ on D_-. This suggests the following:

2.2 Definition. *If $f\in L_1$ on C and has Fourier coefficients a_n, then we shall call*

$$f_+(z) = \sum_0^\infty a_n z^n, \; z \in D_+, \qquad f_-(z) = \sum_1^\infty a_{-n} z^{-n}, \; z \in D_-,$$

the inner *and* outer functions *determined by* f.

These functions can also be represented by Cauchy integrals:

$$f_+(z) = \frac{1}{2\pi i} \int_C \frac{f(w)}{w-z} dw, \quad z \in D_+$$

$$f_-(z) = \frac{1}{2\pi i} \int_C \frac{f(w)}{w-z} dw, \quad z \in D_-,$$

the sense of integration being counter-clockwise. From this we get the Poisson integral representation

2.3
$$f_+(z) + f_-(1/\bar{z}) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) \cdot P(z, e^{i\theta}) \, d\theta, \quad z \in D_+, \tag{2.3}$$

where $P(z, w) = (|w|^2 - |z|^2)/|w-z|^2$. By expressing the difference

$$f_+(re^{i\theta}) + f_-\left(\frac{1}{r} e^{i\theta}\right) - f(e^{i\theta})$$

as another Poisson integral, and applying the Fejér-Lebesgue technique used in proving the Abel or Cesaro summability of a Fourier series in L_1 [19, ch. III], we can prove the following basic theorem.

2.4 THEOREM. *If* $f \in L_1$ *on* C, *then for almost all* $\theta \in [0, 2\pi]$

$$\lim_{r \to 1-0} \left\{ f_+(re^{i\theta}) + f_-\left(\frac{1}{r} e^{i\theta}\right) \right\} = f(e^{i\theta}).$$

Now suppose that $f \in L_\delta$, where $\delta \geq 1$ and $a_n = 0$ for $n < 0$. Then the L.H.S. of (2.3) reduces to $f_+(z)$. Putting $z = re^{i\theta}$ in this, and noting that since $\delta \geq 1$ the function $x^\delta \, (x \geq 0)$ is convex, we get a uniform upper bound for the integrals

$$\int_0^{2\pi} |f_+(re^{i\theta})|^\delta \, d\theta, \quad 0 < r < 1,$$

by applying Jensen's Inequality [19, p. 68] and Fubini's Theorem. Thus

2.5 COROLLARY. *If* $f \in L_\delta$ *on* C, *where* $\delta \geq 1$, *and its n-th Fourier coefficient vanishes for* $n < 0$, *then* $f_+ \in H_\delta$ *on* D_+.

The following theorem will play a crucial role in the stochastic theory. The proofs of parts (a), (b) due to F. Riesz are given in [19, p. 162]. The proofs of parts (c), (d) are essentially the same as those given by Szegö [10, § 2] for $\delta = 2$.

2.6 Theorem. *Let $f_+ \in H_\delta$ on D_+, where $\delta > 0$, and supppose that f_+ does not vanish identically. Then*

(a) $f(e^{i\theta}) = \lim_{r \to 1-0} f_+(re^{i\theta})$ *exists a.e. on C and $f \in L_\delta$;*

(b) $f_+ \to f$ *in the L_δ-topology, i.e. as $r \to 1-0$,*

$$\int_0^{2\pi} |f_+(re^{i\theta}) - f(e^{i\theta})|^\delta \, d\theta \to 0;$$

(c) $\log |f| \in L_1$ *on C, and*

$$\log |f_+(z)| \leq \frac{1}{2\pi} \int_0^{2\pi} \log |f(e^{i\theta})| \cdot P(z, e^{i\theta}) \, d\theta, \quad z \in D_+; \tag{1}$$

in particular

$$\log |f_+(0)| \leq \frac{1}{2\pi} \int_0^{2\pi} \log |f(e^{i\theta})| \, d\theta; \tag{2}$$

(d) if in (2) we have equality, then f_+ has no zeros in D_+.

The converse of (d) is in general false. For instance, Szegö's function [11, p. 270]

$$f_+(z) = \exp \{(z+1)/(z-1)\}$$

has no zeros on D_+; yet $\log |f_+(0)| = \log (1/e) = -1$, whereas

$$\int_0^{2\pi} \log |f(e^{i\theta})| \, d\theta = 0,$$

since $|f(e^{i\theta})| = 1$, $0 < \theta < 2\pi$. It may be shown that the converse of (d) holds in case the reciprocal $1/f_+$ is itself in a Hardy class.

An obvious corollary of 2.6 is the following result of F. Riesz and Nevanlinna.

2.7 Corollary. *The boundary function of a non-constant function in the class H_δ on D_+, where $\delta > 0$, cannot take on the same value on a subset of C of positive measure.*

The following converse of 2.6 will also play an important part in the stochastic theory.

2.8 THEOREM. *If $\phi \in L_\delta$ on C, where $\delta > 0$, and $\phi \geq 0$ and $\log \phi \in L_1$, then there exists a function $f_+ \in H_\delta$ on D_+ and without zeros on D_+ such that if f is its radial limit,*[1] *then $|f| = \phi$ a.e. on C; and* $f_+(0) = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log \phi(e^{i\theta})\, d\theta\right\}\cdot$

Proof. Our proof is a variation of that given by Szegö [10, § 1] and Wiener [15] for $\delta = 2$. Since the function $g = \log \phi^2 = 2 \log \phi$ is real-valued and in L_1, its Fourier coefficients satisfy the conditions $a_{-n} = \bar{a}_n$, a_0 real. Consequently,[2] $g_-(z) = \bar{g}_+(1/\bar{z})$, whence by 2.4, as $r \to 1-0$,

$$g_+(re^{i\theta}) + \bar{g}_+(re^{i\theta}) \to 2 \log \phi(e^{i\theta}), \text{ a.e.} \tag{1}$$

Letting $f_+ = \exp g_+$, we get

$$f_+(0) = e^{\frac{1}{2}a_0} = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log \phi(e^{i\theta})\, d\theta\right\},$$

and

$$|f_+(re^{i\theta})| \to \phi(e^{i\theta}) \text{ a.e., as } r \to 1-0. \tag{2}$$

Applying Jensen's Inequality [19, p. 68] to the exponential function, we get

$$\begin{aligned}|f_+(re^{i\theta})|^\delta &= \exp\{\delta \cdot \text{real } g_+(re^{i\theta})\}\\ &\leq \frac{1}{2\pi}\int_0^{2\pi} \phi(e^{it}) \cdot P(re^{i\theta}, e^{it})\, dt.\end{aligned}$$

Integrating over $[0, 2\pi]$ and using Fubini's Theorem, we conclude that $f_+ \in H_\delta$. By 2.6 (a), f_+ will have a radial limit f, and by (2) $|f| = \phi$ a.e. Obviously f_+ has no zeros on D_+. (Q.E.D.)

Finally, we will need the following uniqueness theorem.

2.9 THEOREM. *If ϕ is as in 2.8, then there is only one function f_+ with the properties*

$$f_+ \in H_\delta \text{ on } D_+ \tag{1}$$

$$|f_+(re^{i\theta})| \to \phi(e^{i\theta}) \text{ a.e., as } r \to 1-0 \tag{2}$$

$$f_+(0) = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log \phi(e^{i\theta})\, d\theta\right\}\cdot \tag{3}$$

---

[1] The existence of f follows from 2.6 (a); in fact, $f \in L_\delta$.

[2] For reasons of symmetry, we have here taken

$$g_+(z) = \tfrac{1}{2}a_0 + \sum_1^\infty a_n z^n, \qquad g_-(z) = \tfrac{1}{2}a_0 + \sum_1^\infty a_{-n} z^{-n},$$

and to this extent departed from Definition 2.2.

Proof. The function f_+ constructed in the last proof has these properties. Since it has no zeros in D_+, $\log f_+ = u_1 + i v_1$ is holomorphic on D_+. The function $u_1(z) = \log |f_+(z)|$ is therefore harmonic on D_+. Also from (1) of the last proof and (2.3)

$$u_1(z) = \frac{1}{2\pi} \int_0^{2\pi} \log \phi(e^{i\theta}) \cdot P(z, e^{i\theta})\, d\theta, \quad z \in D_+. \tag{4}$$

Suppose that g_+ is another function satisfying (1)–(3). By 2.6 (d), g_+ will have no zeros on D_+, and therefore $\log g_+ = u_2 + i v_2$ is also holomorphic on D_+, and the function $u_2(z) = \log |g_+(z)|$ harmonic on D_+. Also, from 2.6 (c) (i)

$$u_2(z) \leq \frac{1}{2\pi} \int_0^{2\pi} \log \phi(e^{i\theta}) \cdot P(z, e^{i\theta})\, d\theta, \quad z \in D_+. \tag{5}$$

It follows from (4) and (5) that the harmonic function $u(z) = u_1(z) - u_2(z) \geq 0$ on D_+. Moreover, $u(0) = 0$, since both f_+, g_+ satisfy (3). Hence u must vanish throughout D_+, i.e. $u_2 = u_1$. Thus

$$\log g_+(z) = u_1(z) + i v_2(z).$$

But two harmonic functions conjugate to a given harmonic function can differ only by a real constant. Hence $v_2(z) = v_1(z) + \lambda$, and so

$$g_+(z) = f_+(z) \cdot e^{i\lambda}.$$

But since by (3) $g_+(0) = f_+(0)$, we get $e^{i\lambda} = 1$, i.e. $g_+ = f_+$ on D_+. (Q.E.D.)

3. Matrix-valued functions

The analysis of matrix-valued functions carried out in this section will be needed in the study of the spectra of multiple stochastic processes.

Notation. *Bold face letters* **x**, **y**, *etc. will denote q-dimensional column-vectors with complex components* x_i, y_i, *etc. The symbol* $|\mathbf{x}|$ *will denote the Euclidean length of* $\mathbf{x}$:

3.1
$$|\mathbf{x}| = \sqrt{\sum_1^q |x_i|^2}. \tag{3.1}$$

Bold face letters **A**, **B**, *etc. will denote* $q \times q$ *matrices with complex entries* a_{ij}, b_{ij}, *etc.*

and $\mathbf{F}$, $\mathbf{G}$, $\mathbf{\Phi}$, *etc. will denote matrix-valued functions. The symbols* Δ, τ, * *will refer to the determinant, trace and adjoint (i.e. conjugate transpose) of matrices.*

Before dealing with matrix-valued functions, it will be convenient to recall the following basic proposition on the topological and algebraic structure of the space of matrices.

3.2 Theorem. (a) *The space of* $q \times q$ *matrices with complex entries is a Banach algebra under the usual algebraic operations and either of the norms*

$$|\mathbf{A}|_B = \underset{\mathbf{x} \neq 0}{\text{l.u.b.}} \frac{|\mathbf{A}\mathbf{x}|}{|\mathbf{x}|} \quad (\textit{Banach-norm}),$$

$$|\mathbf{A}|_E = \sqrt{\tau(\mathbf{A}\mathbf{A}^*)} = \sqrt{\sum_{i=1}^{q} \sum_{j=1}^{q} |a_{ij}|^2} \quad (\textit{Euclidean norm}).$$

(b) *This space is a Hilbert space under the same algebraic operations and the inner product*

$$(\mathbf{A}, \mathbf{B}) = \tau(A B^*) = \sum_{i=1}^{q} \sum_{j=1}^{q} a_{ij} \bar{b}_{ij}.$$

We refer to Hille [4] for the basic properties of Banach spaces and algebras. The two norms define equivalent topologies in view of the inequalities

$$|\mathbf{A}|_B \le |\mathbf{A}|_E \le \sqrt{q}\, |\mathbf{A}|_B,$$

which in a sense are the best possible. In this topology $\mathbf{A}_n \to \mathbf{A}$ as $n \to \infty$, if and only if each entry of $\mathbf{A}_n$ tends, in the ordinary sense, to the corresponding entry of $\mathbf{A}$. The following lemma will be needed in Sec. 4.

3.3 Lemma. *If* $\mathbf{H}$ *is hermitian and* $\mathbf{A}_n\mathbf{H} \to \mathbf{L}$ *as* $n \to \infty$, *then there is a matrix* $\mathbf{A}$ *such that* $\mathbf{L} = \mathbf{A}\mathbf{H}$. *(It is not implied that* $\mathbf{A}_n \to \mathbf{A}$.)

Proof. If $\mathbf{H}$ is invertible, then we need only take $\mathbf{A} = \mathbf{L}\mathbf{H}^{-1}$. If $\mathbf{H} = \mathbf{0}$, then $\mathbf{L} = \mathbf{0}$, and we can take any $\mathbf{A}$. We may therefore suppose that $\Delta(\mathbf{H}) = 0 \neq \tau(\mathbf{H})$, i.e. that $\mathbf{H}$ has both a zero and a non-zero eigenvalue.

Let $\lambda_1, \ldots, \lambda_p$ be the non-zero eigenvalues of $\mathbf{H}$. Then there is a unitary matrix $\mathbf{U}$ such that $\mathbf{U}\mathbf{H}\mathbf{U}^* = \mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_p, 0, \ldots, 0$. Let $\mathbf{B}_n = \mathbf{U}\mathbf{A}_n\mathbf{U}^*$. Then

$$\mathbf{B}_n\mathbf{\Lambda} = \mathbf{U}\mathbf{A}_n\mathbf{H}\mathbf{U}^* \to \mathbf{U}\mathbf{L}\mathbf{U}^*, \quad \text{as } n \to \infty.$$

Now all except the first p columns of $\mathbf{B}_n\mathbf{\Lambda}$ vanish. The same must therefore be the case with the limit $\mathbf{U}\mathbf{L}\mathbf{U}^*$. Also if $1 \leq j \leq p$, then since $\lambda_j \neq 0$, the entries in the jth column of $\mathbf{U}\mathbf{L}\mathbf{U}^*$ are expressible in the form $m_{1j}\lambda_j$, $m_{2j}\lambda_j, \ldots, m_{qj}\lambda_j$. It follows that $\mathbf{U}\mathbf{L}\mathbf{U}^* = \mathbf{M}\mathbf{\Lambda}$, where $\mathbf{M}$ is any matrix with the entries m_{ij} in the first p columns. Hence, letting $\mathbf{A} = \mathbf{U}^*\mathbf{M}\mathbf{U}$, we have

$$\mathbf{L} = \mathbf{U}^*\mathbf{M}\mathbf{\Lambda}\mathbf{U} = \mathbf{A}\mathbf{U}^*\mathbf{\Lambda}\mathbf{U} = \mathbf{A}\mathbf{H}.$$

(Q.E.D.)

We shall now turn to the Lebesgue classes and the Lebesgue integral for matrix-valued functions.

3.4 DEFINITION. *The sets* $\mathbf{L}_\delta$, *where* $\delta > 0$, *and* $\mathbf{L}_\infty$ *are defined as follows.*

(a) $\mathbf{L}_\delta$ *consists of all* $q \times q$ *matrix-valued functions* $\mathbf{F} = [f_{ij}]$ *on the unit circle* C *with complex-valued entries* $f_{ij} \in L_\delta$ (*cf.* 2.1).

(b) $\mathbf{L}_\infty$ *consists of such functions* $\mathbf{F}$ *for which each* $f_{ij} \in L_\infty$, *i.e. each* f_{ij} *is essentially bounded.*

Note. In the greater part of this section we shall imagine that the closed interval $[0, 2\pi]$, and not the unit circle C, is the domain of functions in $\mathbf{L}_\delta$ or $\mathbf{L}_\infty$. We shall thus be writing $\mathbf{F}(\theta)$, where strictly speaking we should write $\mathbf{F}(e^{i\theta})$.

The following theorem is provable by essentially classical arguments, cf. Zygmund [19, pp. 73–74], Hille [4, p. 46] and Stone [8, pp. 29–30].

3.5 THEOREM. (a) $\mathbf{F} \in \mathbf{L}_\delta$, *where* $\delta > 0$, *if and only if* $\mathbf{F}$ *has measurable entries and* $|\mathbf{F}|_E \in L_\delta$ *on* $[0, 2\pi]$. $\mathbf{L}_\delta$, $\delta \geq 1$, *is a Banach space under the usual algebraic operations and the norm*

$$|\mathbf{F}|_\delta = \left\{\frac{1}{2\pi}\int_0^{2\pi} |\mathbf{F}(\theta)|_E^\delta \, d\theta\right\}^{1/\delta}$$

(b) $\mathbf{L}_2$ *is a Hilbert space under the same operations and the inner product*

$$((\mathbf{F}, \mathbf{G})) = \frac{1}{2\pi}\int_0^{2\pi} \tau\{\mathbf{F}(\theta)\,\mathbf{G}^*(\theta)\}\, d\theta,$$

the corresponding norm being

$$\|\mathbf{F}\| = \sqrt{((\mathbf{F}, \mathbf{F}))} = |\mathbf{F}|_2.$$

(c) $\mathbf{F} \in \mathbf{L}_\infty$, *if and only if* $\mathbf{F}$ *has measurable entries and* $|\mathbf{F}|_E$ *is essentially bounded.* $\mathbf{L}_\infty$ *is a Banach algebra under the usual algebraic operations and the norm*

$$|\mathbf{F}|_\infty = \underset{0 \leq \theta \leq 2\pi}{\text{ess. l.u.b.}} \, |\mathbf{F}(\theta)|_E.$$

3.6 Definition. *The* Lebesgue integral *of a function* $\mathbf{F} = [f_{ij}] \in \mathbf{L}_\delta$, *where* $\delta \geq 1$, *is defined by*

$$\int_0^{2\pi} F(\theta)\, d\theta = \left[\int_0^{2\pi} f_{ij}(\theta)\, d\theta\right].$$

3.7 Lemma. (a) $\mathbf{F} \in \mathbf{L}_\delta$, $\mathbf{G} \in \mathbf{L}_{\delta'}$, *where* $1/\delta + 1/\delta' = 1$, *implies that* $\mathbf{F} \cdot \mathbf{G} \in L_1$.

(b) $\mathbf{F}_n \to \mathbf{F}$ *in* $\mathbf{L}_\delta$, $\mathbf{G}_n \to \mathbf{G}$ *in* $\mathbf{L}_{\delta'}$, *as* $n \to \infty$, *where* $1/\delta + 1/\delta' = 1$ *implies that*

$$\int_0^{2\pi} \mathbf{F}_n(\theta)\, \mathbf{G}_n(\theta)\, d\theta \to \int_0^{2\pi} \mathbf{F}(\theta)\, \mathbf{G}(\theta)\, d\theta, \quad \textit{as } n \to \infty.$$

(c) $\mathbf{F} \in \mathbf{L}_\delta$, $\delta > 0$, $\mathbf{G} \in \mathbf{L}_\infty$ *implies that* $\mathbf{F}\mathbf{G} \in \mathbf{L}_\delta$.

(d) *If* $\delta' > \delta > 0$, *then* $\mathbf{L}_\infty \subseteq \mathbf{L}_{\delta'} \subseteq \mathbf{L}_\delta$ *and* $|\mathbf{F}|_\infty \geq |\mathbf{F}|_{\delta'} \geq |\mathbf{F}|_\delta$.

(e) *If* $\mathbf{F} \in \mathbf{L}_\delta$, $\delta > 0$, *then* $\Delta F \in L_{\delta/q}$.

Proof. The results (a)–(d) follow readily from the corresponding results for complex-valued functions. As for (e), let $\mathbf{F} = [f_{ij}]$, and consider a term of $\Delta \mathbf{F}(\theta)$:

$$g(\theta) = \pm f_{1j_1}(\theta) \ldots f_{qj_q}(\theta).$$

Since

$$|g(\theta)|^{\delta/q} = \sqrt[q]{|f_{1j_1}(\theta)|^\delta \ldots |f_{qj_q}(\theta)|^\delta} \leq \frac{1}{q} \sum_{i=1}^{q} |f_{ij_i}(\theta)|^\delta,$$

and the integral of the last term is finite, it follows that $g \in L_{\delta/q}$. Since $L_{\delta/q}$ is a vector-space, $\Delta \mathbf{F}$, which is a sum of $q!$ such functions g, will itself belong to $L_{\delta/q}$. (Q.E.D.)

We shall now discuss the harmonic analysis of matrix-valued functions. It follows from 3.7 (c), taking $\mathbf{G}(\theta) = e^{-ni\theta}\mathbf{I}$, that every function $\mathbf{F} \in \mathbf{L}_\delta$, where $\delta \geq 1$, possesses an *n-th Fourier coefficient*

3.8
$$\mathbf{A}_n = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{F}(\theta)\, e^{-ni\theta}\, d\theta. \tag{3.8}$$

If $\mathbf{A}_n = [a_{ij}^{(n)}]$ and $\mathbf{F} = [f_{ij}]$, then $a_{ij}^{(n)}$ will be the nth Fourier coefficient of the function f_{ij}. In the next theorem we state the matricial extensions of some well known results of Fourier analysis.

3.9 Theorem. (a) *If* $\mathbf{A}_n$ *is the n-th Fourier coefficient of* $\mathbf{F} \in \mathbf{L}_\delta$, $\delta \geq 1$, *then* $\mathbf{A}_n \to \mathbf{0}$ *as* $n \to \pm\infty$ *(Riemann-Lebesgue Theorem).*

(b) *If* $\mathbf{A}_n$ *is the n-th Fourier coefficient of* $\mathbf{F} \in \mathbf{L}_2$, *then* $\sum_{-\infty}^{\infty} |\mathbf{A}_n|_E^2 < \infty$; *conversely, if the* $\mathbf{A}_n$ *are such that* $\sum_{-\infty}^{\infty} |\mathbf{A}_n|_E^2 < \infty$, *then there is a Function* $\mathbf{F} \in \mathbf{L}_2$ *whose n-th Fourier coefficient is* $\mathbf{A}_n$ *(Riesz–Fischer Theorem).*

(c) *If* $\mathbf{F}, \mathbf{G} \in \mathbf{L}_2$ *and have n-th Fourier coefficients* $\mathbf{A}_n$, $\mathbf{B}_n$, *then*

$$\frac{1}{2\pi}\int_0^{2\pi} \mathbf{F}(\theta)\,\mathbf{G}^*(\theta)\,d\theta = \sum_{-\infty}^{\infty} \mathbf{A}_n \mathbf{B}_n^* \quad \textit{(Parseval's Identity)}.$$

(d) *With the hypotheses of* (c) *the n-th Fourier coefficient of* $\mathbf{F}\mathbf{G}$ *is* $\sum_{k=-\infty}^{\infty} \mathbf{A}_k \mathbf{B}_{n-k}$ *(Convolution Rule).*

As these results follow readily from the corresponding ones for complex-valued functions, we shall omit the proofs. Now let $\mathbf{F} \in \mathbf{L}_1$ on the unit circle C. Its nth Fourier coefficient $\mathbf{A}_n$ is given by (3.8) in which, however, we must replace $\mathbf{F}(\theta)$ by $\mathbf{F}(e^{i\theta})$. Since by 3.9 (a) $\mathbf{A}_n \to \mathbf{0}$ as $n \to \pm\infty$, it follows that the series $\sum_0^{\infty} \mathbf{A}_n z^n$ converges on D_+ and the series $\sum_1^{\infty} \mathbf{A}_{-n} z^{-n}$ on D_-. We are thus led, as in the scalar case 2.2, to the following definition.

3.10 Definition. *Given* $\mathbf{F} \in \mathbf{L}_1$ *on* C *with n-th Fourier coefficient* A_n, *we shall call*

$$\mathbf{F}_+(z) = \sum_0^{\infty} \mathbf{A}_n z^n, \quad z \in D_+, \qquad \mathbf{F}_-(z) = \sum_1^{\infty} \mathbf{A}_{-n} z^{-n}, \quad z \in D_-,$$

the inner *and* outer functions *determined by* F.

Our last task in this section will be to establish certain determinantal inequalities for hermitian matrices, and to derive a determinantal extension of Jensen's inequality for the logarithmic function [19, pp. 67–68].

3.11 Lemma. *Let* $\mathbf{A}$, $\mathbf{B}$ *be* $q \times q$ *non-negative, hermitian matrices. Then*

(a) $$\sqrt[q]{\Delta(\mathbf{A}+\mathbf{B})} \geq \sqrt[q]{\Delta(\mathbf{A})} + \sqrt[q]{\Delta(\mathbf{B})},$$

(b) $$\frac{\Delta(\mathbf{A}+\mathbf{B})}{\tau(\mathbf{A}+\mathbf{B})} \geq \frac{\Delta(\mathbf{A})}{\tau(\mathbf{A})},$$

(c) $$\Delta(\mathbf{A}+\mathbf{B}) \geq \Delta(\mathbf{A}).$$

Proof. (a) A proof of the first inequality, due to Minkowski, is given in [3, p. 34].

(b) There is no loss of generality in supposing that **A** is diagonal; for we can always diagonalise it by means of a unitary transformation, and this will not affect the trace or determinant of **B**. We thus have with an obvious notation

$$\Delta(\mathbf{A}+\mathbf{B}) = \begin{vmatrix} a_1+b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & a_2+b_{22} & \dots & b_{2q} \\ b_{q1} & b_{q2} & \dots & a_q+b_{qq} \end{vmatrix}$$
$$= \Delta(\mathbf{B}) + \sum_i a_i \Delta(\mathbf{B}_i) + \sum_{i<j} a_i a_j \Delta(\mathbf{B}_{ij}) + \cdots + a_1 a_2 \dots a_q,$$

where $\mathbf{B}_{ijk\dots}$ is the principal minor obtained by deleting the ith, jth, kth, ... rows and columns of **B**. Since **A**, **B** are non-negative, each term in the last expansion is non-negative. Hence retaining only the last two terms, we get

$$\Delta(\mathbf{A}+\mathbf{B}) \geq \sum_k a_1 \dots a_{k-1} \cdot a_{k+1} \dots a_q b_{kk} + \Delta(\mathbf{A}).$$

Consequently,

$$\tau(\mathbf{A}) \cdot \Delta(\mathbf{A}+\mathbf{B}) \geq \sum_k \tau(\mathbf{A}) a_1 \dots a_{k-1} \cdot a_{k+1} \dots a_q b_{kk} + \tau(\mathbf{A}) \Delta(\mathbf{A}). \tag{1}$$

Now $\tau(\mathbf{A}) \geq a_k$, since **A** is non-negative. Hence the first term on the R.H.S. of (1)

$$\geq \sum_k a_1 \dots a_k \dots a_q \cdot b_{kk} = \Delta(\mathbf{A}) \cdot \tau(\mathbf{B});$$

and from (1) we get

$$\tau(\mathbf{A}) \cdot \Delta(\mathbf{A}+\mathbf{B}) \geq \Delta(\mathbf{A}) \{\tau(\mathbf{B}) + \tau(\mathbf{A})\} = \Delta(\mathbf{A}) \cdot \tau(\mathbf{A}+\mathbf{B}).$$

(c) follows trivially from (a) or (b). (Q.E.D.)

3.12 THEOREM. *If* $\mathbf{F} \in \mathbf{L}_1$ *and the values of* **F** *are non-negative hermitian, then*

$$\log \Delta \left(\frac{1}{2\pi} \int_0^{2\pi} \mathbf{F}(\theta)\, d\theta \right) \geq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta \{\mathbf{F}(\theta)\}\, d\theta \geq -\infty.$$

Proof. Extending the inequality 3.11 (a) to a finite number of summands, we get

$$\sqrt[q]{\Delta\left(\sum_1^n c_i \mathbf{A}_i\right)} \geq \sum_1^n c_i \sqrt[q]{\Delta(\mathbf{A}_i)},$$

where each $c_i \geq 0$ and $\mathbf{A}_i$ is non-negative hermitian. It follows that

$$\sqrt[q]{\Delta\left(\sum_1^\infty c_i \mathbf{A}_i\right)} \geq \sum_1^\infty c_i \sqrt[q]{\Delta(\mathbf{A}_i)},$$

provided that the two infinite series converge. Now since $\mathbf{F} \in \mathbf{L}_1$ and therefore by 3.7 (e) $\Delta \mathbf{F} \in \mathbf{L}_{1/q}$, this condition is satisfied, if $c_i = \text{meas. } S_i / 2\pi$, where $S_1, S_2, \ldots$ are measurable sets belonging to a Lebesgue partition of $[0, 2\pi]$, and if $\mathbf{A}_i = \mathbf{F}(\theta_i)$, where θ_i takes almost all values in S_i. Thus

$$\sqrt[q]{\Delta \left\{ \frac{1}{2\pi} \sum_1^\infty \mathbf{F}(\theta_i) \text{ meas. } S_i \right\}} \geq \frac{1}{2\pi} \sum_1^\infty \sqrt[q]{\Delta \{\mathbf{F}(\theta_i)\}} \cdot \text{meas. } S_i. \tag{1}$$

For almost all $\theta_i \in S_i$, the sum on the L.H.S. approaches $\int_0^{2\pi} \mathbf{F}(\theta)\, d\theta$ and that on the R.H.S. approaches $\int_0^{2\pi} \sqrt[q]{\Delta\{\mathbf{F}(\theta)\}}\, d\theta$, as l.u.b. (meas. S_i)$\to 0$. Hence from (1) we infer that

$$\sqrt[q]{\Delta \left\{ \frac{1}{2\pi} \int_0^{2\pi} \mathbf{F}(\theta)\, d\theta \right\}} \geq \frac{1}{2\pi} \int_0^{2\pi} \sqrt[q]{\Delta\{\mathbf{F}(\theta)\}}\, d\theta.$$

Taking logarithms on both sides and applying Jensen's inequality for the log-function on the R.H.S., we get the desired result. (Q.E.D.)

3.13 THEOREM. *If* $\mathbf{F} \in \mathbf{L}_\delta$ *on the unit circle* C, *where* $\delta \geq 1$, *and its n-th Fourier coefficient vanishes for* $n > 0$, *then*

(a) $\Delta \mathbf{F}_+ \in H_{\delta/q}$ *on* D_+, $\Delta \mathbf{F} \in L_{\delta/q}$ *on* C,

(b) *either* $\Delta \mathbf{F}_+$ *vanishes identically, or* $\log \Delta \mathbf{F} \in L_1$ *on* C *and*

$$\log |\Delta\{\mathbf{F}_+(0)\}| \leq \frac{1}{2\pi} \int_0^{2\pi} \log |\Delta\{\mathbf{F}(e^{i\theta})\}|\, d\theta.$$

Proof. (a) By 3.7 (e), $\Delta \mathbf{F} \in L_{\delta/q}$ on C. Next, each entry f_{ij} of $\mathbf{F}$ is in the class L_δ and its nth Fourier coefficient vanishes for $n < 0$. Hence by 2.5 $f_{ij+} \in H_\delta$ on D_+. Consider a term of $\Delta \mathbf{F}_+(z)$:

$$g(z) = \pm f_{1j_1+}(z) \ldots f_{qj_q+}(z).$$

As in the proof of 3.7 (e),

$$|g(z)|^{\delta/q} \leq \frac{1}{q} \sum_{i=1}^q |f_{ij_i+}(z)|^\delta, \quad z \in D_+.$$

By taking $z = re^{i\theta}$, $0 \leq r < 1$, and integrating from 0 to 2π we see that $g \in H_{\delta/q}$. Hence $\Delta \mathbf{F}_+$, which is the sum of $q!$ such terms, is itself in $H_{\delta/q}$. (b) now follows from 2.6 (c). (Q.E.D.)

4. Matricial Riemann-Stieltjes integration

We must first consider matrix-valued functions of bounded variation.

4.1 DEFINITION. *Let* $\mathbf{F}$ *be a* $q \times q$ *matrix-valued function on* $[a, b]$. *We say that*

(a) $\mathbf{F}$ *is of* bounded variation, *if and only if the set of variations*

$$\sum_{k=1}^{n} |\mathbf{F}(x_k) - \mathbf{F}(x_{k-1})|_E$$

of $\mathbf{F}$ *over different finite partitions* $\{x_0, \ldots, x_n\}$ *of* $[a, b]$ *is bounded above;*

(b) $\mathbf{F}$ *is* non-decreasing, *if and only if its values are hermitian, and the difference* $\mathbf{F}(x') - \mathbf{F}(x)$ *is non-negative hermitian whenever* $x' > x$.

The following lemma is easily verified.

4.2 LEMMA. (a) $\mathbf{F} = [f_{ij}]$ *is of bounded variation on* $[a, b]$, *if and only if each entry* f_{ij} *is of this type.*

(b) *If* $\mathbf{F} = [f_{ij}]$ *is non-decreasing and bounded on* $[a, b]$, *then each* f_{ii} *is real-valued, non-decreasing and bounded on* $[a, b]$, *and each* $f_{ij} (i \neq j)$ *is a function of bounded variation, in general complex-valued.*

From 4.2 and the well known properties of complex-valued functions of bounded variation it follows that if $\mathbf{F}$ is of bounded variation, it has at most denumerably many points of discontinuity, all of them simple, and that $\mathbf{F}'$ exists a.e. and is in $\mathbf{L}_1$ on $[a, b]$. But in general $\mathbf{F}$ will not be *absolutely continuous,* i.e. we will not have

$$\mathbf{F}(x) = \mathbf{F}(a) + \int_a^x \mathbf{F}'(t)\,dt, \quad a \leq x \leq b.$$

For any function $\mathbf{F}$ of bounded variation on $[a, b]$, we define the functions, $\mathbf{F}^{(a)}$, $\mathbf{F}^{(d)}$, $\mathbf{F}^{(s)}$ by

$$\left.\begin{aligned} \mathbf{F}^{(a)}(x) &= \mathbf{F}(a) + \int_a^x \mathbf{F}'(t)\,dt, \\ \mathbf{F}^{(d)}(x) &= \sum_{a \leq t \leq x} \{\mathbf{F}(t+0) - \mathbf{F}(t-0)\}, \\ \mathbf{F}^{(s)}(x) &= \mathbf{F}(x) - \mathbf{F}^{(a)}(x) - \mathbf{F}^{(d)}(x), \end{aligned}\right\} \tag{4.3}$$

4.3

and call these the *absolutely continuous, discontinuous, singular parts* of $\mathbf{F}$, respectively. We see at once that $\mathbf{F}^{(s)}$ is continuous on $[a, b]$. Also, since $\mathbf{F}^{(a)\prime} = \mathbf{F}'$ a.e., and $\mathbf{F}^{(d)\prime} = 0$, except at the points of discontinuity of $\mathbf{F}$, it follows that $\mathbf{F}^{(s)\prime} = 0$, a.e. Adopting the same superscript notation for the entries f_{ij}, we readily obtain

4.4 $$\mathbf{F}^{(a)} = [f_{ij}^{(a)}], \qquad \mathbf{F}^{(d)} = [f_{ij}^{(d)}], \qquad \mathbf{F}^{(s)} = [F_{ij}^{(s)}]. \tag{4.4}$$

It is an important fact that if a real-valued function f is non-decreasing, then so are $f^{(a)}$, $f^{(d)}$, $f^{(s)}$. We owe to Cramer [1, Theorem 2] the corresponding result for matrix-valued functions:

4.5 THEOREM. *If* $\mathbf{F}$ *is non-decreasing and bounded on* $[a, b]$, *then so are the parts* $\mathbf{F}^{(a)}$, $\mathbf{F}^{(d)}$, $\mathbf{F}^{(s)}$.

An immediate consequence of this is the following corollary.

4.6 COROLLARY. *If* $\mathbf{F}$ *is non-decreasing and bounded on* $[a, b]$ *then*

$$\mathbf{F}(x') - \mathbf{F}(x) - \int_x^{x'} \mathbf{F}'(t)\,dt$$

is non-negative hermitian for $a \le x \le x' \le b$.

We now turn to RS-integration.

4.7 DEFINITION. *Let* $\mathbf{F}, \mathbf{G}$ *be* $q \times q$ *matrix-valued functions on* $[a, b]$;
$\pi = \{x_0, \ldots, x_n\}$ *be a partition of* $[a, b]$;
$|\pi| = \max(x_k - x_{k-1})$, $k = 1, \ldots, n$;
and $\pi^* = \{t_1, \ldots, t_n\}$, *where* $x_{k-1} \le t_k \le x_k$.

(a) *If as* $|\pi| \to 0$,

$$S(\mathbf{F}, \mathbf{G}, \pi, \pi^*) = \sum_{k=1}^{n} \mathbf{F}(t_k)\{\mathbf{G}(x_k) - \mathbf{G}(x_{k-1})\}$$

tends to a limit $\mathbf{L}$, *then we call* $\mathbf{L}$ *the* left RS-integral *of* $\mathbf{F}$ *w.r.t.* $\mathbf{G}$ *from* a *to* b, *and denote it by* $\int_a^b \mathbf{F}(x) \cdot d\mathbf{G}(x)$.

(b) *We similarly define the* right RS-integral $\int_a^b d\mathbf{G}(x) \cdot \mathbf{F}(x)$.

(c) *If as* $|\pi| \to 0$,

$$S'(\mathbf{F}, \mathbf{G}, \pi, \pi^*) = \sum_{k=1}^{n} \mathbf{F}(t_k)\{\mathbf{G}(x_k) - \mathbf{G}(x_{k-1})\} F^*(t_k)$$

tends to a limit $\mathbf{L}$, *then we call* $\mathbf{L}$ *the* bilateral RS-integral *of* $\mathbf{F}, \mathbf{F}^*$ *w.r.t.* $\mathbf{G}$ *from* a *to* b, *and denote it by* $\int_a^b \mathbf{F}(x) \cdot d\mathbf{G}(x) \cdot \mathbf{F}^*(x)$.

A simple calculation shows that if $\mathbf{F} = [f_{ij}]$, $\mathbf{G} = [g_{ij}]$, then

$$S(\mathbf{F}, \mathbf{G}, \pi, \pi^*) = \left[\sum_{\lambda=1}^{q} S(f_{i\lambda}, g_{\lambda j}, \pi, \pi^*) \right],$$

$$S'(\mathbf{F}, \mathbf{G}, \pi, \pi^*) = \left[\sum_{\lambda=1}^{q} \sum_{\mu=1}^{q} S(f_{i\lambda} \cdot \bar{f}_{j\mu}, g_{\lambda\mu}, \pi, \pi^*) \right],$$

from which we at once infer the following result.

4.8 Lemma. *Let* $\mathbf{F} = [f_{ij}]$, $\mathbf{G} = [g_{ij}]$ *be functions on* $[a, b]$.

(a) *The integral* $\int_a^b \mathbf{F}(x)\, d\mathbf{G}(x)$ *will exist, if and only if all the ordinary RS-integrals* $\int_a^b f_{i\lambda}(x)\, d g_{\lambda j}(x)$ *exist, and in this case*

$$\int_a^b \mathbf{F}(x) \cdot d\mathbf{G}(x) = \left[\sum_{\lambda=1}^{q} \int_a^b f_{i\lambda}(x)\, d g_{\lambda j}(x) \right].$$

(*b*) *Analogous results hold for the integrals*

$$\int_a^b d\mathbf{G}(x) \cdot \mathbf{F}(x), \quad \int_a^b \mathbf{F}(x) \cdot d\mathbf{G}(x) \cdot \mathbf{F}^*(x).$$

This lemma along with 4.2 permits the immediate extension to the matrix-case of many results of the classical Riemann-Stieltjes theory. Thus

4.9 Theorem. (a) *If* $\mathbf{F}$ *is continuous and* $\mathbf{G}$ *of bounded variation on* $[a, b]$ *then the integrals*

$$\int_a^b \mathbf{F}(\theta) \cdot d\mathbf{G}(\theta), \quad \int_a^b d\mathbf{G}(\theta) \cdot \mathbf{F}(\theta), \quad \int_a^b \mathbf{F}(\theta) \cdot d\mathbf{G}(\theta) \cdot \mathbf{F}^*(\theta)$$

exist.

(b) $$\int_a^b \mathbf{F}(\theta)\, d\mathbf{G}_1(\theta) \pm \int_a^b \mathbf{F}(\theta)\, d\mathbf{G}_2(\theta) = \int_a^b \mathbf{F}(\theta)\, d\{\mathbf{G}_1(\theta) \pm \mathbf{G}_2(\theta\}.$$

(c) $$\int_a^b \mathbf{F}(x) \cdot d\mathbf{G}(x) + \int_a^b \mathbf{G}(x) \cdot d\mathbf{F}(x) = \mathbf{F}(b)\, \mathbf{G}(b) - \mathbf{F}(a)\, \mathbf{G}(a).$$

(d) *If* $\mathbf{G}$ *is absolutely continuous then* $\int_a^b \mathbf{F}(\theta) \cdot d\mathbf{G}(\theta) = \int_a^b \mathbf{F}(\theta)\, \mathbf{G}'(\theta)\, d\theta.$

For non-decreasing integrator-functions we have the following useful result.

4.10 THEOREM. *If* $\mathbf{F}$ *is continuous and* $\mathbf{G}$ *is non-decreasing and bounded on* $[a, b]$, *then* $\mathbf{F}\mathbf{G}'\mathbf{F}^* \in \mathbf{L}_1$ *on* $[a, b]$, *and*

$$\int_a^b \mathbf{F}(x)\, d\mathbf{G}(x)\cdot\mathbf{F}^*(x) - \int_a^b \mathbf{F}(x)\,\mathbf{G}'(x)\,\mathbf{F}^*(x)\,dx$$

is non-negative hermitian.

Proof. As remarked after 4.2, $\mathbf{G}' \in \mathbf{L}_1$. Also $\mathbf{F}, \mathbf{F}^* \in \mathbf{L}_\infty$. Hence from 3.7 (c) $\mathbf{F}\mathbf{G}'\mathbf{F}^* \in \mathbf{L}_1$.

Next, writing $\mathbf{A} \geq \mathbf{B}$ to mean that $\mathbf{A} - \mathbf{B}$ is non-negative hermitian, and using the notation of 4.7 (c), we have from 4.6

$$\begin{aligned} S'(\mathbf{F}, \mathbf{G}, \pi, \pi^*) &= \sum_{k-1}^{n} \mathbf{F}(t_k)\,\{\mathbf{G}(x_k) - \mathbf{G}(x_{k-1}\}\,\mathbf{F}^*(t_k) \\ &\geq \sum_{k=1}^{n} \mathbf{F}(t_k)\cdot \int_{x_{k-1}}^{x_k} \mathbf{G}'(t)\,dt\cdot\mathbf{F}^*(t_k) \\ &\geq \int_a^b \mathbf{H}_{\pi\pi^*}(t)\,dt, \end{aligned} \tag{1}$$

where the functions $\mathbf{H}_{\pi\pi^*}$ have the same domain S as $\mathbf{G}'$, and

$$\mathbf{H}_{\pi\pi^*}(t) = \mathbf{F}(t_k)\,\mathbf{G}'(t)\,\mathbf{F}^*(t_k), \quad \text{if } t \in S \text{ and } x_{k-1} \leq t \leq x_k.$$

Now let $t \in S$ and let $k(\pi)$ be the integer such shat $t \in [x_{k(\pi)-1}, x_{k(\pi)}]$. Obviously as $|\pi| \to 0$. we have $t_{k(\pi)} \to t$, whence from the continuity of $\mathbf{F}$,

$$\mathbf{H}_{\pi\pi^*}(t) \to \mathbf{F}(t)\,\mathbf{G}'(t)\,\mathbf{F}^*(t), \quad t \in S. \tag{2}$$

Also, if M is an upper bound of $\mathbf{F}$, then

$$|\mathbf{H}_{\pi\pi^*}(t)|_E \leq M^2\,|\mathbf{G}'(t)|_E, \quad t \in S. \tag{3}$$

Since $\mathbf{G}' \in \mathbf{L}_1$, it follows from (2), (3) and Lebesgue's Theorem on Dominated Convergence [4, p. 48] that

$$\int_a^b \mathbf{H}_{\pi\pi^*}(t)\,dt \to \int_a^b \mathbf{F}(t)\cdot\mathbf{G}'(t)\cdot\mathbf{F}^*(t)\,dt, \text{ as } |\pi| \to 0.$$

The desired relation thus follows from (1) on letting $|\pi| \to 0$. (Q.E.D.)

We turn finally to the Fourier–Stieltjes analysis of matrix-valued functions. Let $\mathbf{F}$ be of bounded variation on $[0, 2\pi]$. Then from 4.9 (a), $\mathbf{F}$ will always possess a *n-th Fourier–Stieltjes coefficient*

4.11
$$\mathbf{A}_n = \frac{1}{2\pi}\int_0^{2\pi} e^{-ni\theta}\, d\mathbf{F}(\theta) = \left[\frac{1}{2\pi}\int_0^{2\pi} e^{-ni\theta}\, d f_{ij}(\theta)\right]. \qquad (4.11)$$

Since the function $e^{-ni\theta}\mathbf{I}$ is absolutely continuous, we get using 4.9 (c) (d)

$$\mathbf{A}_n = \mathbf{A}_0 + \frac{n i}{2\pi}\int_0^{2\pi} e^{-ni\theta}\,\mathbf{F}(\theta)\, d\theta,$$

i.e.
$$\frac{1}{2\pi}\int_0^{2\pi} e^{-ni\theta}\, d\mathbf{F}(\theta) = \frac{1}{ni}(\mathbf{A}_n - \mathbf{A}_0), \quad n \neq 0.$$

Since
$$\frac{1}{2\pi}\int_0^{2\pi} e^{-ni\theta}\,\theta\, d\theta = -1/ni, \quad n \neq 0,$$

it follows that $\mathbf{A}_n/ni$ is the nth Fourier coefficient of $\mathbf{F}(\theta) - \mathbf{A}_0\theta$ for $n \neq 0$. This being of bounded variation, we have as in the classical case [19, p. 25].

$$\tfrac{1}{2}\{\mathbf{F}(\theta+) + \mathbf{F}(\theta-)\} - \mathbf{A}_0\theta = \mathbf{B}_0 + \sum_{n\neq 0}\frac{\mathbf{A}_n}{ni}e^{ni\theta}, \quad 0 < \theta < 2\pi,$$

$$\tfrac{1}{2}\{\mathbf{F}(0+) + \mathbf{F}(2\pi-)\} = \mathbf{B}_0 + \sum_{n\neq 0}\frac{\mathbf{A}_n}{ni},$$

where $\mathbf{B}_0$ is the 0th Fourier coefficient of $\mathbf{F}(\theta) - \mathbf{A}_0\theta$. Obviously, $\mathbf{B}_0$ cannot be determined from the Fourier-Stieltjes coefficients $\mathbf{A}_n$, since functions of bounded variation differing by constants will have the same $\mathbf{A}_n$. Eliminating $\mathbf{B}_0$ between the last two relations, we get

4.12
$$\frac{\mathbf{F}(\theta+) + \mathbf{F}(\theta-)}{2} - \frac{\mathbf{F}(0+) + \mathbf{F}(2\pi-)}{2} = \mathbf{A}_0(\theta - \pi) + \sum_{n\neq 0}\frac{\mathbf{A}_n}{ni}(e^{ni\theta} - 1), \quad \theta < \theta < 2\pi. \qquad (4.12)$$

4.13 Lemma. (a) *If for all* n, $\int_0^{2\pi} e^{ni\theta}\, d\mathbf{F}(\theta) = \mathbf{0}$, *then* $\mathbf{F}$ *is constant-valued.*

(b) *If for all* n, $\int_0^{2\pi} e^{ni\theta}\, d\mathbf{F}(\theta) = \int_0^{2\pi} e^{ni\theta}\, d\mathbf{G}(\theta)$, *then* $\mathbf{F}$ *and* $\mathbf{G}$ *differ by a constant matrix.*

5. Vector-valued random functions

A multiple stochastic process is a one-parameter family of vector-valued random functions. The basic analytic properties of such functions must therefore be studied before we can effectively deal with multiple processes.

Let Ω be a space possessing a Borel field $\mathfrak{F}$ of subsets over which is defined a probability measure P. Let $\mathfrak{L}_2 = L_2(\Omega)$ be the set of all complex-valued P-measurable functions f on Ω for which $\int_\Omega |f(\omega)|^2 \, dP(\omega) < \infty$. Then, cf. Stone [8, pp. 208–209], $\mathfrak{L}_2$ is a Hilbert space with the usual operations and the inner product

$$(f, g) = \int_\Omega f(\omega) \cdot \overline{g(\omega)} \, dP(\omega).$$

The corresponding norm is

$$|f| = \sqrt{(f, f)} = \sqrt{\int_\Omega |f(\omega)|^2 \, dP(\omega)}.$$

For $X, Y \subseteq \mathfrak{L}_2$ we shall denote by $X + Y$ the set of all functions $f + g$, with $f \in X$, $g \in Y$. $\mathfrak{S}(f_j)_{j \in J}$ will denote the *subspace*, i.e. closed linear manifold, spanned by the functions f_j, for each j in the index-set J. Finally, $(f | \mathfrak{M})$ will denote the *orthogonal projection* of f on the subspace $\mathfrak{M}$ of $\mathfrak{L}_2$.

We shall now define appropriate analogues of these concepts for q-dimensional (column) vector-valued functions on Ω. We shall denote such functions by the bold face letters $\mathbf{f}$, $\mathbf{g}$, $\boldsymbol{\varphi}$, $\boldsymbol{\psi}$, etc. The components of $\mathbf{f}$, $\mathbf{g}$, etc. will be denoted by $f^{(j)}$, $g^{(j)}$, etc. $j = 1, \ldots q$.

5.1 Definition. *We define the set* $\mathbf{L}_2$ *as consisting of all q-dimensional (column) vector-valued functions* $\mathbf{f}$ *on* Ω, *with complex valued-components* $f^{(j)} \in \mathfrak{L}_2$.

As in 3.5, $\mathbf{f} \in \mathbf{L}_2$, if and only if the components $f^{(j)}$ are measurable and

$$\int_\Omega |\mathbf{f}(\omega)|^2 \, dP(\omega) < \infty;$$

where $|\ |$ denotes the Euclidean length, (3.1); moreover, cf. Stone [8, pp. 29–30], $\mathbf{L}_2$ is a Hilbert space under the usual operations and the *inner product*

5.2
$$((\mathbf{f}, \mathbf{g})) = \int_\Omega \sum_{j=1}^{q} \mathbf{f}^{(j)}(\omega) \overline{\mathbf{g}^{(j)}(\omega)} \, dP(\omega). \tag{5.2}$$

This inner product generates the *norm*

5.3
$$\|\mathfrak{f}\| = \sqrt{((\mathfrak{f}, \mathfrak{f}))} = \sqrt{\int_{\Omega} |\mathfrak{f}(\omega)|^2 \, dP(\omega)} \qquad (5.3)$$

which in turn induces a *topology* in the space: if $\mathfrak{f}_n, \mathfrak{f} \in \mathfrak{L}_2$ we shall say that $\mathfrak{f}_n \to \mathfrak{f}$, as $n \to \infty$, if and only if $\|\mathfrak{f}_n - \mathfrak{f}\| \to 0$ as $n \to \infty$. This is equivalent to saying that for each $j = 1, \ldots, q$, $f_n^{(j)} \to f^{(j)}$ in $\mathfrak{L}_2$, i.e.

$$\int_{\Omega} |f_n^{(j)}(\omega) - f^{(j)}(\omega)|^2 \, dP(\omega) \to 0, \text{ as } n \to \infty.$$

Convergence will always be understood in this sense, e.g. the equation $\mathfrak{f} = \sum_{-\infty}^{\infty} \mathfrak{f}_n$ will mean that $\|\mathfrak{f} - \sum_{-N}^{N} \mathfrak{f}_n\| \to 0$, as $N \to \infty$.

The inner produkt (5.2) does not play any significant role in the stochastic theory, although the corresponding norm (5.3) and the topology it induces do, and has to be replaced by Gramian matrices:

5.4 DEFINITION. *If* $\mathfrak{f}, \mathfrak{g} \in \mathfrak{L}_2$ *then the matrix*

$$(\mathfrak{f}, \mathfrak{g}) = [(f^{(j)}, g^{(k)})] = \left[\int_{\Omega} f^{(j)}(\omega) \overline{g^{(k)}(\omega)} \, dP(\omega)\right]$$

is called the Gramian *of the pair* $\mathfrak{f}, \mathfrak{g}$.

We see from (5.2)-(5.4) that

$$((\mathfrak{f}, \mathfrak{g})) = \tau(\mathfrak{f}, \mathfrak{g}), \qquad \|\mathfrak{f}\| = \sqrt{\tau(\mathfrak{f}, \mathfrak{f})}.$$

The next two definitions differ from the usual ones in that Gramians replace inner products, and matrix coefficients replace complex coefficients in linear combinations.

5.5 DEFINITION. *We say that*

(a) $\mathfrak{f} \perp \mathfrak{g}$ *if and only if* $(\mathfrak{f}, \mathfrak{g}) = 0$;

(b) $\mathfrak{f}$ *is a* normal *vector if and only if* $(\mathfrak{f}, \mathfrak{f}) = \mathbf{I}$;

(c) *the sequence* $(\mathfrak{f}_n)_{-\infty}^{\infty}$ *is* orthonormal *if and only if* $(\mathfrak{f}_m, \mathfrak{f}_n) = \delta_{mn} \mathbf{I}$.

5.6 DEFINITION. (a) *A* linear manifold *in* $\mathfrak{L}_2$ *is a non-void subset* $\mathfrak{M}$ *such that if* $\mathfrak{f}, \mathfrak{g} \in \mathfrak{M}$, *then* $\mathbf{A}\mathfrak{f} + \mathbf{B}\mathfrak{g} \in \mathfrak{M}$, *for all* $q \times q$ *matrices* $\mathbf{A}, \mathbf{B}$.

(b) *A* subspace *of* $\mathfrak{L}_2$ *is a linear manifold, which is closed in the topology of the norm* $\| \ \|$, (*cf.* (5.3)).

(c) *The* subspace (linear manifold) spanned by a subset $\mathfrak{M}$ of $\mathfrak{L}_2$ *is the intersection of all subspaces (linear manifolds) containing* $\mathfrak{M}$.[1] *The subspace spanned by the indexed set* $\{\mathfrak{f}_j\}_{j \in J}$ *will be denoted by* $\mathfrak{S}(\mathfrak{f}_j)_{j \in J}$.

(d) *If* $\mathfrak{M}_j \subseteq \mathfrak{L}_2$ *for* $j \in J$, *then by* $\sum_{j \in J} \mathfrak{M}_j$, *we mean the set of all sums* $\sum_{j \in J} \mathfrak{f}_j$, *with* $\mathfrak{f}_j \in \mathfrak{M}_j$, *which converge in the topology of the norm* $\| \ \|$.

In the next three lemmas we sum up some basic facts governing the notions just introduced.

5.7 Lemma. (a) $(\mathfrak{g}, \mathfrak{f}) = (\mathfrak{f}, \mathfrak{g})^*$, $(\mathfrak{f}, \mathfrak{f})$ *is non-negative hermitian.*

(b) *If* $\mathfrak{f}_n \to \mathfrak{f}$, $\mathfrak{g}_n \to \mathfrak{g}$, *as* $n \to \infty$, *then* $(\mathfrak{f}_n, \mathfrak{g}_n) \to (\mathfrak{f}, \mathfrak{g})$.

(c)
$$\left(\sum_{j=1}^{m} \mathbf{A}_j \mathfrak{f}_j, \quad \sum_{k=1}^{n} \mathbf{B}_k \mathfrak{g}_k \right) = \sum_{j=1}^{m} \sum_{k=1}^{n} \mathbf{A}_j (\mathfrak{f}_j, \mathfrak{f}_k) \mathbf{B}_k^*.$$

(d) $\mathfrak{f} \perp \mathfrak{g}$, *if and only if* $f^{(i)} \perp g^{(j)}$ *for* $i, j = 1, \dots q$; *this implies* $\|\mathfrak{f} + \mathfrak{g}\|^2 = \|\mathfrak{f}\|^2 + \|\mathfrak{g}\|^2$.

(e) $\mathfrak{f}$ *is normal, if and only if its components form an orthonormal set in* $\mathfrak{L}_2$.

(f) *The set* $\{\mathfrak{f}_j\}_{j \in J}$ *is orthonormal in* $\mathfrak{L}_2$, *if and only if the components* $f_j^{(i)}$, *where* $j \in J$, $i = 1, \dots, q$, *form an orthonormal set in* $\mathfrak{L}_2$.

(g) *If* $(\mathfrak{f}_m, \mathfrak{f}_n) = \delta_{mm} \mathbf{K}$, *where* $\mathbf{K}$ *is invertible, and* $\mathfrak{g}_n = \sqrt{\mathbf{K}^{-1}} \cdot \mathfrak{f}_n$, *then* $(\mathfrak{g}_n)_{-\infty}^{\infty}$ *is orthonormal.*

The proofs of these obvious results are omitted. Somewhat less obvious are the following results regarding the new concept of subspace. But the proof being of a routine nature is omitted.

5.8 Lemma. (a) $\mathfrak{M}$ *is a subspace of* $\mathfrak{L}_2$, *if and only if there is a subspace* $\mathfrak{M}$ *of* $\mathfrak{L}_2$ *such that* $\mathfrak{M} = \mathfrak{M}^q$, *where* $\mathfrak{M}^q$ *denotes the Cartesian product* $\mathfrak{M} \otimes \cdots \otimes \mathfrak{M}$ *with* q *factors.* $\mathfrak{M}$ *is the set of all components of all functions in* $\mathfrak{M}$.

(b) *If* $\mathfrak{M}$ *is a subspace of* $\mathfrak{L}_2$ *and* $\mathfrak{f} \in \mathfrak{L}_2$, *then there exists a unique* $\mathfrak{g}$ *such that*

$$\mathfrak{g} \in \mathfrak{M}; \ \|\mathfrak{f} - \mathfrak{g}\| \le \|\mathfrak{f} - \mathbf{h}\|, \quad \textit{for all } \mathbf{h} \in \mathfrak{M}. \tag{1}$$

For this $\mathfrak{g}$, $g^{(i)} = (f^{(i)} | \mathfrak{M})$, $\mathfrak{M}$ *being as in* (a). *A function* $\mathfrak{g} \in \mathfrak{M}$ *satisfies* (1), *if and only if* $\mathfrak{f} - \mathfrak{g} \perp \mathfrak{M}$.

(c) *If* $\mathfrak{M}, \mathfrak{N}$ *are subspaces of of* $\mathfrak{L}_2$, *and* $\mathfrak{M} \subseteq \mathfrak{N}$, *then there exists a unique subspace* $\mathfrak{M}' \subseteq \mathfrak{N}$ *such that*

$$\mathfrak{N} = \mathfrak{M} + \mathfrak{M}', \ \mathfrak{M} \perp \mathfrak{M}'.$$

[1] As in the classical case it is easily seen that the intersection of any family of subspaccs (manifolds) is a subspace (manifold).

(d) $\mathfrak{g} \in \mathfrak{S}(\mathfrak{f}_j)_{j\in J}$, *if and only if* $\mathfrak{g} = \lim_{n\to\infty} \mathfrak{g}_n$, *where* $\mathfrak{g}_n$ *is a linear combination of a finite number of* $\mathfrak{f}_j$, $j \in J$, *with matrix coefficients.*

(e) *If* $\mathfrak{M} = \mathfrak{S}(\mathfrak{f}_j)_{j\in J}$, $\mathfrak{M}^{(i)} = \mathfrak{S}(\mathfrak{f}_j^{(i)})_{j\in J}$, $i = 1, \dots, q$, *and* $\mathfrak{M}$ *is defined as in* (a), *then* $\mathfrak{M} = \text{clos.} \sum_{i=1}^{q} \mathfrak{M}^{(i)}$.[1]

5.9 Definition. *The unique function* $\mathfrak{g}$ *of* 5.8 (b) *is called the* orthogonal projection *of* $\mathfrak{f}$ *on* $\mathfrak{M}$, *and will be denoted by* $(\mathfrak{f}\,|\,\mathfrak{M})$.

The following can now be proved almost verbatim as in Hilbert space theory.

5.10 Lemma. (a) *If* $\mathfrak{M}$, $\mathfrak{N}$ *are orthogonal subspaces of* $\mathfrak{L}_2$, *then* $\mathfrak{M} + \mathfrak{N}$ *is a (closed) subspace, and for any* $\mathfrak{f}$,

$$(\mathfrak{f}\,|\,\mathfrak{M} + \mathfrak{N}) = (\mathfrak{f}\,|\,\mathfrak{M}) + (\mathfrak{f}\,|\,\mathfrak{N}).$$

(b) *If* $\mathfrak{M}, \mathfrak{N}$ *are subspaces such that* $\mathfrak{M} \subseteq \mathfrak{N}$, *then* $\|(\mathfrak{f}\,|\,\mathfrak{M})\| \le \|(\mathfrak{f}\,|\,\mathfrak{N})\|$.

(c) *If* $\mathfrak{M} = \bigcap_{-\infty}^{\infty} \mathfrak{M}_n$, *where* $\mathfrak{M}_n$ *are subspaces such that* $\mathfrak{M}_n \subseteq \mathfrak{M}_{n+1}$, *then*[2]

$$(\mathfrak{f}\,|\,\mathfrak{M}) = \lim_{n\to-\infty} (\mathfrak{f}\,|\,\mathfrak{M}_n).$$

To any non-zero vector in Hilbert space corresponds a unit vector which is a scalar multiple of the original. But to a non-zero vector $\mathfrak{f} \in \mathfrak{L}_2$ does not always correspond a normal vector $\mathfrak{g}$ (Definition 5.5 (b)), which is a matrix multiple of $\mathfrak{f}$. This is because the Gramian $(\mathfrak{f}, \mathfrak{f})$ need not be invertible; if it were we could take $\mathfrak{g} = \sqrt{(\mathfrak{f}, \mathfrak{f})^{-1}} \cdot \mathfrak{f}$. Accordingly, to cover the degenerate cases arising in the theory of multiple processes, we have to take Fourier expansions with respect to sequences of orthogonal but not necessarily orthonormal vectors. The following is a basic proposition regarding such expansions.

5.11 Theorem. *Let* $(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_n) = \delta_{mn}\,\mathbf{K}, \quad \mathbf{K} \neq \mathbf{0}.$

(a) *If*[3]

$$\mathfrak{f} = \sum_{-\infty}^{\infty} \mathbf{A}_n \boldsymbol{\varphi}_n, \quad \mathfrak{g} = \sum_{-\infty}^{\infty} \mathbf{B}_n \boldsymbol{\varphi}_n,$$

then

$$\left.\begin{aligned} (\mathfrak{f}, \mathfrak{g}) = \sum_{-\infty}^{\infty} \mathbf{A}_n \mathbf{K} \mathbf{B}_n^*, \qquad (\mathfrak{f}, \mathfrak{f}) = \sum_{-\infty}^{\infty} \mathbf{A}_n \mathbf{K} \mathbf{A}_n^*; \\ \|\mathfrak{f}\|^2 = \sum_{-\infty}^{\infty} |\mathbf{A}_n \mathbf{K}^{\frac{1}{2}}|_E^2 < \infty; \\ (\mathfrak{f}, \boldsymbol{\varphi}_n) = \mathbf{A}_n \mathbf{K}. \end{aligned}\right\} \tag{1}$$

[1] We have to take the closure, for as with other infinite-dimensional spaces, the topological closure of subsets X, Y does not imply that of $X + Y$.

[2] Cf. the footnote to 5.6 (c).

[3] Cf. the remarks on convergence after (4.3).

(b) *The linear manifold* $\sum_{n=-\infty}^{\infty} \mathfrak{S}(\boldsymbol{\varphi}_n)$ *is closed and identical to the subspace* $\mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}$.

(c) *For every* $\mathbf{f} \in \mathfrak{L}_2$, *there exist matrices* $\mathbf{A}_n$ *such that*

$$(\mathbf{f} \mid \mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}) = \sum_{-\infty}^{\infty} \mathbf{A}_n \boldsymbol{\varphi}_n, \quad \mathbf{A}_n \mathbf{K} = (\mathbf{f}, \boldsymbol{\varphi}_n).$$

Proof. (a) By 5.7 (c)

$$\left(\sum_{-N}^{N} \mathbf{A}_n \boldsymbol{\varphi}_n, \sum_{-N}^{N} \mathbf{B}_n \boldsymbol{\varphi}_n \right) = \sum_{m=-N}^{N} \sum_{n=-N}^{N} \mathbf{A}_n (\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_n) \mathbf{B}_n^* = \sum_{m=-N}^{N} \mathbf{A}_m \mathbf{K} \mathbf{B}_m^*, \tag{2}$$

since $(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_n) = \boldsymbol{\delta}_{mn} \mathbf{K}$. By 5.7 (b) the L.H.S. of (2) tends to $(\mathbf{f}, \mathbf{g})$. Hence from (2) the infinite series $\sum_{-\infty}^{\infty} \mathbf{A}_m \mathbf{K} \mathbf{B}_m^*$ converges to $(\mathbf{f}, \mathbf{g})$. Taking $\mathbf{g} = \mathbf{f}$, we get the corresponding result for $(\mathbf{f}, \mathbf{f})$. The formula for $\|\mathbf{f}\|^2$ follows, since $\|\mathbf{f}\|^2 = \tau(\mathbf{f}, \mathbf{f})$. Finally taking $\mathbf{B}_m = \delta_{mn} \mathbf{I}$ in (1)

$$(\mathbf{f}, \boldsymbol{\varphi}_n) = \left(\sum_{-\infty}^{\infty} \mathbf{A}_m \boldsymbol{\varphi}_m, \boldsymbol{\varphi}_n \right) = \mathbf{A}_n \mathbf{K}.$$

(b) Let $\mathfrak{M} = \sum_{n=-\infty}^{\infty} \mathfrak{S}(\boldsymbol{\varphi}_n)$. Obviously $\mathfrak{M} \subseteq \mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}$. Also $\mathfrak{M}$ is a manifold containing every $\boldsymbol{\varphi}_n$. Hence if $\mathfrak{M}$ is closed, then $\mathfrak{M}$ is a subspace containing every $\boldsymbol{\varphi}_n$ and so $\mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty} \subseteq \mathfrak{M}$. We have therefore only to show that $\mathfrak{M}$ is closed.

Suppose $\mathbf{g}_k \in \mathfrak{M}$ and $\mathbf{g}_k \to \mathbf{g}$. Let

$$\mathbf{g}_k = \sum_{n=-\infty}^{\infty} \mathbf{A}_{kn} \boldsymbol{\varphi}_n.$$

Then cf. (a)

$$\|\mathbf{g}_j - \mathbf{g}_k\|^2 = \sum_{n=-\infty}^{\infty} |(\mathbf{A}_{jn} - \mathbf{A}_{kn}) K^{\frac{1}{2}}|_E^2. \tag{3}$$

Since by the Cauchy condition, the L.H.S. $\to 0$, as $j, k \to \infty$, therefore for all n

$$|(\mathbf{A}_{jn} - \mathbf{A}_{kn}) \mathbf{K}^{\frac{1}{2}}|_E \to 0 \quad \text{as } j, k \to \infty.$$

Hence for each n, the sequence $(\mathbf{A}_{jn} \mathbf{K}^{\frac{1}{2}})_{j=1}^{\infty}$ converges. By 3.3 its limit must be of the form $\mathbf{B}_n \mathbf{K}^{\frac{1}{2}}$. Letting $k \to \infty$ in (3) we thus get

$$\begin{aligned} \|\mathbf{g}_j - \mathbf{g}\|^2 &= \sum_{n=-\infty}^{\infty} |(\mathbf{A}_{jn} - \mathbf{B}_n) \mathbf{K}^{\frac{1}{2}}|_E^2 \\ &= \Big\| \sum_{n=-\infty}^{\infty} (\mathbf{A}_{jn} - \mathbf{B}_n) \boldsymbol{\varphi}_n \Big\|^2 \\ &= \Big\| \mathbf{g}_j - \sum_{n=-\infty}^{\infty} \mathbf{B}_n \boldsymbol{\varphi}_n \Big\|^2. \end{aligned}$$

Letting $j \to \infty$, we get $\mathbf{g} = \sum_{n=-\infty}^{\infty} \mathbf{B}_n \boldsymbol{\varphi}_n$, i.e. $\mathbf{g} \in \mathfrak{M}$. Thus $\mathfrak{M}$ is closed.

(c) Since $\mathbf{g} = (\mathbf{f} \mid \mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}) \in \mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty} = \sum_{-\infty}^{\infty} \mathfrak{S}(\boldsymbol{\varphi}_n)$,

there exist matrices $\mathbf{A}_n$ such that

$$\mathbf{g} = \sum_{-\infty}^{\infty} \mathbf{A}_n \boldsymbol{\varphi}_n.$$

By (a), $\mathbf{A}_n \mathbf{K} = (\mathbf{g}, \boldsymbol{\varphi}_n)$. But $\mathbf{f} - \mathbf{g} \perp \boldsymbol{\varphi}_n$, and therefore $(\mathbf{g}, \boldsymbol{\varphi}_n) = (\mathbf{f}, \boldsymbol{\varphi}_n)$. (Q.E.D.)

Putting $\mathbf{K} = \mathbf{I}$ in the last Theorem we get the following specialisation for Fourier expansions with respect to an orthonormal system (Doob [2, p. 595]).

5.12 COROLLARY. *Let* $(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}$ *be orthonormal. Then*

(a) *the linear manifold* $\sum_{-\infty}^{\infty} \mathfrak{S}(\boldsymbol{\varphi}_n)$ *is closed and identical with the subspace* $\mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}$.

(b) *For any* $\mathbf{f}, \mathbf{g} \in \mathfrak{S}(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}$,

$$\mathbf{f} = \sum_{-\infty}^{\infty} (\mathbf{f}, \boldsymbol{\varphi}_n)\, \boldsymbol{\varphi}_n, \qquad \|\mathbf{f}\|^2 = \sum_{-\infty}^{\infty} |(\mathbf{f}, \boldsymbol{\varphi}_n)|_E^2 < \infty,$$

$$(\mathbf{f}, \mathbf{g}) = \sum_{-\infty}^{\infty} (\mathbf{f}, \boldsymbol{\varphi}_n)(\boldsymbol{\varphi}_n, \mathbf{g}).$$

(c) *If* $\mathbf{f} = \sum_{-\infty}^{\infty} \mathbf{A}_n \boldsymbol{\varphi}_n$, *then* $\mathbf{A}_n = (\mathbf{f}, \boldsymbol{\varphi}_n)$.

6. Time-domain analysis of multivariate processes

By a *q-variate* (or *q-ple*) *stationary stochastic process* we shall mean a sequence $(\mathbf{f}_n)_{-\infty}^{\infty}$ of vector-valued functions $\mathbf{f}_n \in \mathfrak{L}_2$ (cf. Definition 5.1) such that the Gramian matrix

6.1 $$(\mathbf{f}_m, \mathbf{f}_n) = \boldsymbol{\Gamma}_{m-n} = [\gamma_{ij}^{(m-n)}] \tag{6.1}$$

depends only on the difference $m-n$ and not on m and n separately. $\boldsymbol{\Gamma}_n$ is called the *covariance or correlation matrix for lead* n, and the sequence $(\boldsymbol{\Gamma}_n)_{-\infty}^{\infty}$ is called *covariance-sequence* of the S.P.

If $(\mathbf{f}_n)_{-\infty}^{\infty}$ is a q-ple S.P. and $f_n^{(1)}, \ldots, f_n^{(q)}$ are the components of $\mathbf{f}_n$, then each $f_n^{(i)} \in \mathfrak{L}_2$ and by (6.1)

6.2 $$(f_m^{(i)}, f_n^{(i)}) = \gamma_{ii}^{(m-n)} \tag{6.2}$$

depends only on the difference $m-n$. Thus the q-ple process $(\mathbf{f}_n)_{-\infty}^{\infty}$ has associated with it q simple process $(f_n^{(i)})_{n=-\infty}^{\infty}$, $i = 1, \ldots, q$, which are stationary in the wide sense (Doob [2, p. 95]).

It follows from the theory of Hilbert spaces (Doob [2, pp. 461–462]) that we can associate with any simple S.P. $(f_n^{(i)})_{n=-\infty}^{\infty}$ a *shift operator* U_i on $\mathfrak{L}_2$ into itself, which is unitary and such that $U_i^n f_0^{(i)} = f_n^{(i)}$. This operator is not necessarily unique. If the simple processes $(f_n^{(i)})_{n=-\infty}^{\infty}$, where $i = 1, \ldots, q$, are the components of q-ple process $(\mathfrak{f}_n)_{-\infty}^{\infty}$ satisfying the stationary condition (6.1), then we can take $U_i = U_j$, as was shown by Kolmogorov [5, Theorem 1], so that a single unitary operator U on $\mathfrak{L}_2$ exists such that

6.3
$$U^n (f_0^{(i)}) = f_n^{(i)}, \quad -\infty < n < \infty, \quad i = 1, \ldots, q, \tag{6.3}$$

Obviously, any two unitary operators satisfying (6.3) will agree on the subspace of $\mathfrak{L}_2$ spanned by the components $f_n^{(i)}$, $-\infty < n < \infty$, $i = 1, \ldots, q$, so that U may be considered "unique", as far as its applicability in the stochastic theory is concerned. We shall call U the *shift operator* of the q-ple process $(\mathfrak{f}_n)_{-\infty}^{\infty}$, and write

6.3′
$$U^n \mathfrak{f}_0 = \mathfrak{f}_n \tag{6.3'}$$

as an abbreviation of (6.3).[1] The same convention will be applied to other transformations T on $\mathfrak{L}_2$ into itself:

6.4
$$\textit{If} \quad \mathfrak{g} = (g^{(1)}, \ldots, g^{(q)}), \quad \textit{then} \;\; T\mathfrak{g} = (Tg^{(1)}, \ldots, Tg^{(q)}). \tag{6.4}$$

We shall denote by $\mathfrak{M}_n$ the space $\mathfrak{S}(\mathfrak{f}_k)_{-\infty}^{n}$ (Definition 5.6 (c)). This is called the *present and past of f_n*. Obviously $\mathfrak{M}_n \subseteq \mathfrak{M}_{n+1}$. The space $\mathfrak{M}_{-\infty} = \bigcap_{-\infty}^{\infty} \mathfrak{M}_n$ will be called the *remote past of the process*, and $\mathfrak{M}_{\infty} = \text{clos.} \bigcup_{-\infty}^{\infty} \mathfrak{M}_n = \mathfrak{S}(\mathfrak{f}_k)_{-\infty}^{\infty}$ the *space spanned by it*. These are all subspaces of $\mathbf{L}_2$. The present and past of the component $f_n^{(i)}$, i.e. the subspace $\mathfrak{S}(f_k^{(i)})_{k=-\infty}^{n}$ of $\mathfrak{L}_2$, will be denoted by $\mathfrak{M}_n^{(i)}$ and we shall similarly define $\mathfrak{M}_{-\infty}^{(i)}$ and $\mathfrak{M}_{\infty}^{(i)}$. It follows from 5.8 (e) that

6.5
$$\mathfrak{g} = (g^{(1)}, \ldots, g^{(q)}) \in \mathfrak{M}_n, \quad \textit{if and only if each} \;\; g^{(i)} \in \text{clos.} \sum_{j=1}^{q} \mathfrak{M}_n^{(j)}. \tag{6.5}$$

It is also obvious that if U is the shift operator of the process $(\mathfrak{f}_n)_{-\infty}^{\infty}$, then

6.6
$$U^n (\mathfrak{f}_j \,|\, \mathfrak{M}_k) = (\mathfrak{f}_{j+n} \,|\, \mathfrak{M}_{k+n}), \tag{6.6}$$

and (with the usual conventions)

6.7
$$U^n (\mathfrak{M}_k) = \mathfrak{M}_{k+n}. \tag{6.7}$$

[1] i.e. we shall regard U as acting on the space $\mathbf{L}_2$, rather than on $\mathfrak{L}_2$.

We shall say that the S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$ is *non-deterministic*,[1] if and only if for some n, $\mathfrak{f}_n \notin \mathfrak{M}_{n-1}$. From the stationarity property (6.1) it follows that the last relation holds for a single n only if it holds for all n. Hence for any non-deterministic process $(\mathfrak{f}_n)_{-\infty}^{\infty}$,

$$\mathfrak{g}_n = \mathfrak{f}_n - (\mathfrak{f}_n \mid \mathfrak{M}_{n-1}) \neq 0, \quad -\infty < n < \infty. \tag{6.8}$$

6.8

We may look upon the vector-functions $\mathfrak{g}_n$ as the *innovations* out of which the $\mathfrak{f}_n$-process is built. Accordingly, we shall call $(\mathfrak{g}_n)_{-\infty}^{\infty}$ the *innovation-process* associated with $(\mathfrak{f}_n)_{-\infty}^{\infty}$. It plays an important part in the theory on account of its simple structure, as shown by the following obvious lemma:

6.9 LEMMA. *If* $(\mathfrak{g}_n)_{-\infty}^{\infty}$ *is the innovation-process of a non-deterministic S.P.* $(\mathfrak{f}_n)_{-\infty}^{\infty}$ *with shift operator* U, *then*

(a) $$\mathfrak{g}_n = U^n \mathfrak{g}_0$$

(b) $$(\mathfrak{g}_m, \mathfrak{g}_n) = \delta_{mn} G, \quad \textit{where } G = (\mathfrak{g}_0, \mathfrak{g}_0) = (\mathfrak{f}_n, \mathfrak{g}_n).$$

We shall call the Gramian $G = (\mathfrak{g}_0, \mathfrak{g}_0)$ of 5.9 (b) the *prediction-error matrix with lag* 1 of the S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$, and following Zasuhin [18] refer to its rank ϱ as the *rank* of this process. We shall say that the S.P. has *full rank*, if $\varrho = q$. Obviously, the S.P. will be non-deterministic if $\varrho \geq 1$, and vice versa, and in this case by 6.9 (b) the innovation process $(\mathfrak{g}_n)_{-\infty}^{\infty}$ will be *orthogonal*. But unless G is invertible, i.e. $\varrho = q$ and the S.P. has full rank, we cannot derive from $(\mathfrak{g}_n)_{-\infty}^{\infty}$ an *orthonormal* process (cf. remarks preceding 5.11). Questions of rank thus render the multiple theory more difficult than the simple theory.

We shall now establish the Wold decomposition of a multiple process, which was announced by Zasuhin [18, Theorem 1] but without proof. We need the following lemma.

6.10 LEMMA. *If* $\mathfrak{f}_n, \mathfrak{g}_n$ *are as in* 6.9, *and* $\mathfrak{M}_n$, $\mathfrak{N}_n$ *are the present and past of* $\mathfrak{f}_n, \mathfrak{g}_n$, *respectively, then*

(a) *for* $m < n$, $$\mathfrak{M}_n = \mathfrak{M}_m + \mathfrak{S}(\mathfrak{g}_k)_{m+1}^{n}, \qquad \mathfrak{M}_m \perp \mathfrak{S}(\mathfrak{g}_k)_{m+1}^{n}$$

(b) $$\mathfrak{M}_n = \mathfrak{M}_{-\infty} + \mathfrak{N}_n, \qquad \mathfrak{M}_{-\infty} \perp \mathfrak{N}_{\infty}.$$

Proof. (a) Since $\mathfrak{M}_{n-1} \subseteq \mathfrak{M}_n$ and $\mathfrak{g}_n \in \mathfrak{M}_n$, therefore

$$\mathfrak{M}_{n-1} + \mathfrak{S}(\mathfrak{g}_n) \subseteq \mathfrak{M}_n. \tag{1}$$

[1] This term is preferable to the term *non-singular* used by KOLMOGOROV [5], and the term *regular* used by DOOB [2].

Now by (6.8) $\qquad \mathbf{f}_n = \mathbf{g}_n + (\mathbf{f}_n \mid \mathfrak{M}_{n-1}) \in \mathfrak{S}(\mathbf{g}_n) + \mathfrak{M}_{n-1}.$

Also, for $k < n$, $\mathbf{f}_k \in \mathfrak{M}_{n-1} \subseteq \mathfrak{S}(\mathbf{g}_n) + \mathfrak{M}_{n-1}$. It follows that (1) holds with $\supseteq$ replacing $\subseteq$, and therefore with equality. By iteration of this equality we get

$$\begin{aligned}\mathfrak{M}_n &= \mathfrak{M}_{n-1} + \mathfrak{S}(\mathbf{g}_n)\\ &= \mathfrak{M}_{n-2} + \mathfrak{S}(\mathbf{g}_n, \mathbf{g}_{n-1})\\ &\vdots\\ &= \mathfrak{M}_m + \mathfrak{S}(\mathbf{g}_n, \mathbf{g}_{n-1}, \ldots, \mathbf{g}_{m+1}).\end{aligned}$$

Since by (6.8) $\mathbf{g}_{m+1}, \mathbf{g}_{m+2}, \ldots \perp \mathfrak{M}_m, \mathfrak{M}_{m+1}, \ldots$, respectively, and $\mathfrak{M}_m$ is contained in all these subspaces, it follows that $\mathfrak{S}(\mathbf{g}_k)_{m+1}^{\infty} \perp \mathfrak{M}_m$.

(b) If $\mathbf{h} \in \mathfrak{M}_{-\infty}$, then for each n, $\mathbf{h} \in \mathfrak{M}_{n-1} \perp \mathbf{g}_n$. Hence $\mathfrak{M}_{-\infty} \perp \mathfrak{N}_{\infty}$. Next, since $\mathbf{g}_n \in \mathfrak{M}_n$, therefore $\mathfrak{N}_n \subseteq \mathfrak{M}_n$. Also $\mathfrak{M}_{-\infty} \subseteq \mathfrak{M}_n$. Hence

$$\mathfrak{M}_{-\infty} + \mathfrak{N}_n \subseteq \mathfrak{M}_n. \tag{2}$$

Now let $\mathbf{h} \in \mathfrak{M}_n$. Then since $\mathfrak{M}_{-\infty} \perp \mathfrak{N}_n$ and by (a) $\mathfrak{M}_m \perp \mathfrak{S}(\mathbf{g}_k)_{m+1}^{n}$, we have (cf. 5.10)

$$\begin{aligned}(\mathbf{h} \mid \mathfrak{M}_{-\infty} + \mathfrak{N}_n) &= (\mathbf{h} \mid \mathfrak{M}_{-\infty}) + (\mathbf{h} \mid \mathfrak{N}_n)\\ &= \lim_{m\to-\infty} (\mathbf{h} \mid \mathfrak{M}_m) + \lim_{m\to-\infty} (\mathbf{h} \mid \mathfrak{S}(\mathbf{g}_k)_{m+1}^{n})\\ &= \lim_{m\to-\infty} (\mathbf{h} \mid \mathfrak{M}_m + \mathfrak{S}(\mathbf{g}_k)_{m+1}^{n})\\ &= \lim_{m\to-\infty} (\mathbf{h} \mid \mathfrak{M}_n) = (\mathbf{h} \mid \mathfrak{M}_n) = \mathbf{h},\end{aligned}$$

in view of (a). Hence $\mathbf{h} \in \mathfrak{M}_{-\infty} + \mathfrak{N}_n$. Thus (2) holds with $\supseteq$ replacing $\subseteq$, and therefore with equality. (Q.E.D.)

6.11 THEOREM. (*Wold decomposition.*) *If* $(\mathbf{g}_n)_{-\infty}^{\infty}$ *is the innovation process of a non-deterministic q-ple* S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$, *and* $\mathfrak{M}_n$, $\mathfrak{N}_n$ *are the present and past of* $\mathbf{f}_n$, $\mathbf{g}_n$, *respectively, then*

(a) $\mathbf{f}_n = \mathbf{u}_n + \mathbf{v}_n$, *where* $\mathbf{u}_n = (\mathbf{f}_n \mid \mathfrak{N}_n) \perp \mathbf{v}_n = (\mathbf{f}_n \mid \mathfrak{M}_{-\infty})$;

(b) *the* S.P. $(\mathbf{u}_n)_{-\infty}^{\infty}$ *is a one-sided moving average:*

$$\mathbf{u}_n = \sum_{k=0}^{\infty} \mathbf{A}_k \mathbf{g}_{n-k}, \quad \|\mathbf{u}_n\|^2 = \sum_{k=0}^{\infty} |\mathbf{A}_k \mathbf{G}^{\frac{1}{2}}|_E^2 < \infty,$$

where $\mathbf{G} = (\mathbf{g}_0, \mathbf{g}_0)$ *and the* $\mathbf{A}_k$ *are any matrices such that*

$$\mathbf{A}_k \mathbf{G} = (\mathbf{u}_0, \mathbf{g}_{-k}) = (\mathbf{f}_0, \mathbf{g}_{-k}), \quad \mathbf{A}_0 \mathbf{G} = \mathbf{G} = \mathbf{G}\mathbf{A}_0^*;$$

(c) *the* S.P. $(\mathbf{v}_k)_{-\infty}^{\infty}$ *is deterministic, and for each* n, $\mathfrak{S}(\mathbf{v}_k)_{-\infty}^{n} = \mathfrak{M}_{-\infty}$.

Proof. (a) Since $\mathbf{f}_n \in \mathfrak{M}_n$, we have $\mathbf{f}_n = (\mathbf{f}_n \mid \mathfrak{M}_n)$. Hence from 6.10 (b) and 5.10 (a)

$$\mathbf{f}_n = (\mathbf{f}_n \mid \mathfrak{M}_{-\infty}) + (\mathbf{f}_n \mid \mathfrak{N}_n) = \mathbf{v}_n + \mathbf{u}_n.$$

Obviously $\mathbf{v}_n \perp \mathbf{u}_n$.

(b) Since $\mathbf{v}_n \perp$ each $\mathbf{g}_j$, and by 6.9 (a) $\mathbf{g}_{n-k} = U^n \mathbf{g}_{-k}$, therefore

$$(\mathbf{u}_n, \mathbf{g}_{n-k}) = (\mathbf{u}_n + \mathbf{v}_n, \mathbf{g}_{n-k}) = (\mathbf{f}_n, \mathbf{g}_{n-k}) = (U^n \mathbf{f}_0, U^n \mathbf{g}_{-k}) = (\mathbf{f}_0, \mathbf{g}_{-k}). \tag{1}$$

Now $\mathbf{u}_n \in \mathfrak{N}_n = \mathfrak{S}(\mathbf{g}_k)_{-\infty}^n$, and by 6.9 (b) $(\mathbf{g}_j, \mathbf{g}_k) = \delta_{jk}\mathbf{G}$. Hence by 5.11 (c) and (a)

$$\mathbf{u}_m = \sum_{k=0}^{\infty} \mathbf{A}_{nk} \cdot \mathbf{g}_{n-k},$$

where $$\mathbf{A}_{nk}\mathbf{G} = (\mathbf{u}_n, \mathbf{g}_{n-k}), \qquad \|\mathbf{u}_n\|^2 = \sum_{k=0}^{\infty} |\mathbf{A}_{nk}\mathbf{G}^{\frac{1}{2}}|_E^2 < \infty.$$

From (1) we see, however, that $\mathbf{A}_{nk}$ is independent of n so that we may write $\mathbf{A}_k$ instead of $\mathbf{A}_{nk}$. Finally by 6.9 (b)

$$\mathbf{A}_0 \mathbf{G} = (\mathbf{f}_0, \mathbf{g}_0) = (\mathbf{g}_0, \mathbf{g}_0) = \mathbf{G}.$$

Taking adjoints we get $\mathbf{G} = \mathbf{G}\mathbf{A}_0^*$.

(c) We first note that

$$\mathfrak{M}_n = \mathfrak{N}_n + \mathfrak{S}(\mathbf{v}_k)_{-\infty}^n, \qquad \mathfrak{N}_n \perp \mathfrak{S}(\mathbf{v}_k)_{-\infty}^n. \tag{2}$$

This is obvious since on the one hand $\mathfrak{N}_n \subseteq \mathfrak{M}_n$ and $\mathbf{v}_k \in \mathfrak{M}_{-\infty} \subseteq \mathfrak{M}_n$, and on the other for all $m \le n$

$$\mathbf{f}_m = \mathbf{u}_m + \mathbf{v}_m \in \mathfrak{N}_n + \mathfrak{S}(\mathbf{v}_k)_{-\infty}^n.$$

Comparing (2) with 6.10 (b), we conclude from 5.8 (c) that $\mathfrak{S}(\mathbf{v}_k)_{-\infty}^n = \mathfrak{M}_{-\infty}$. (Q.E.D.)

From the last theorem we easily get the following corollary for processes of full rank given in Doob [2, p. 597] (cf. 5.7 (g)).

6.12 Corollary. *If $\mathbf{g}_0$ is the innovation function of a full-rank process $(\mathbf{f}_n)_{-\infty}^{\infty}$, then with the notation of the last theorem,*

$$\mathbf{u}_n = \sum_{k=0}^{\infty} \mathbf{C}_k \mathbf{h}_{n-k}, \qquad \|\mathbf{u}_n\|^2 = \sum_{0}^{\infty} |\mathbf{C}_k|_E^2 < \infty,$$

where $(\mathbf{h}_k)_{-\infty}^{\infty}$ is the normalised innovation process of $(\mathbf{f}_n)_{-\infty}^{\infty}$:

$$\mathbf{h}_0 = \sqrt{\mathbf{G}^{-1}} \cdot \mathbf{g}_0, \qquad \mathbf{h}_k = U^k \mathbf{h}_0, \qquad \mathbf{G} = (\mathbf{g}_0, \mathbf{g}_0),$$

and $$\mathbf{C}_0 = \sqrt{\mathbf{G}}, \qquad \mathbf{C}_k = (\mathbf{u}_0, \mathbf{h}_{-k}) = (\mathbf{f}_0, \mathbf{h}_{-k}).$$

We shall call a S.P. *regular*, if $(\mathbf{f}_0 \mid \mathfrak{M}_{-n}) \to 0$, as $n \to \infty$.[1] Such a process is ob-

[1] This usage of the term *regular* is the same as Kolmogorov's [5]. Doob [2] uses the word in a different sense.

viously non-deterministic and hence the last theorem is applicable to it. Conditions equivalent to regularity are stated in the next theorem, which is an extension of Kolmogorov [5, Theorem 19].

6.13 Theorem. *Each of the following conditions is equivalent to the regularity of the stationary* S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$:

(a) $(\mathfrak{f}_n)_{-\infty}^{\infty}$ *is a one-sided moving average:*

$$\mathfrak{f}_n = \sum_{k=0}^{\infty} \mathbf{A}_k \boldsymbol{\varphi}_{n-k}, \qquad (\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_n) = \delta_{mn}\mathbf{K}; \tag{1}$$

(b) $\mathfrak{M}_{-\infty} = \{0\}$.

Proof. (a) Let the S.P. be regular. Then by 6.10 we have with the usual notation

$$\mathfrak{M}_0 = \mathfrak{M}_{-n} + \mathfrak{S}(\mathfrak{g}_n)_{-n+1}^{0}, \qquad \mathfrak{M}_{-n} \perp \mathfrak{S}(\mathfrak{g}_k)_{-n+1}^{0}.$$

Since $(\mathfrak{g}_j, \mathfrak{g}_k) = \delta_{jk}\mathbf{G}$, it follows from 5.10 (a) and 5.11 (c) that

$$\mathfrak{f}_0 = (\mathfrak{f}_0 \mid \mathfrak{M}_0) = (\mathfrak{f}_0 \mid \mathfrak{M}_{-n}) + \sum_{0}^{n+1} \mathbf{A}_k \mathfrak{g}_{-k}.$$

Since by hypothesis $(\mathfrak{f}_0 \mid \mathfrak{M}_{-n}) \to 0$, we see that

$$\mathfrak{f}_0 = \sum_{0}^{\infty} \mathbf{A}_k \mathfrak{g}_{-k}.$$

U^n upplied to both sides yields (1) with $\boldsymbol{\varphi}_n = \mathfrak{g}_n$ and $\mathbf{K} = \mathbf{G}$. Thus $(\mathfrak{f}_n)_{-\infty}^{\infty}$ is a one-sided moving average.

Conversely, suppose that (1) holds. Then by 5.11 (a)

$$\|\mathfrak{f}_0\|^2 = \sum_{0}^{\infty} |\mathbf{A}_k \mathbf{K}^{\frac{1}{2}}|_E^2 < \infty. \tag{2}$$

Letting $\mathfrak{N}_n = \mathfrak{S}(\boldsymbol{\varphi}_k)_{-\infty}^{n}$, it follows by 5.11 (c) that

$$(\mathfrak{f}_0 \mid \mathfrak{N}_{-n}) = \sum_{k=n}^{\infty} \mathbf{A}_k \boldsymbol{\varphi}_{-k}, \qquad \|(\mathfrak{f}_0 \mid \mathfrak{N}_{-n})\|^2 = \sum_{k=n}^{\infty} |\mathbf{A}_k \mathbf{K}^{\frac{1}{2}}|_E^2,$$

whence from (2) $\|(\mathfrak{f}_0 \mid \mathfrak{N}_{-n})\| \to 0$, as $n \to \infty$. But obviously from (1)

$$\mathfrak{M}_{-n} \subseteq \mathfrak{N}_{-n} \quad \text{and} \quad \|(\mathfrak{f}_0 \mid \mathfrak{M}_{-n})\| \le \|(\mathfrak{f}_0 \mid \mathfrak{N}_{-n})\|. \tag{3}$$

Hence the process is regular.

(b) Let the S.P. be regular. Then (1)–(3) hold. From the first relation in (3) it follows that $\mathfrak{M}_{-\infty} \subseteq \mathfrak{N}_{-\infty}$. So we need only to show that $\mathfrak{N}_{-\infty} = \{0\}$. Suppose $\boldsymbol{\psi} \in \mathfrak{N}_{-\infty}$. Then for all n, $\boldsymbol{\psi} \in \mathfrak{N}_{-n}$, and therefore by 5.11 (c)

$$\boldsymbol{\psi} = \sum_n^\infty \mathbf{B}_k \boldsymbol{\varphi}_{-k}, \qquad \|\boldsymbol{\psi}\|^2 = \sum_n^\infty |\mathbf{B}_k \mathbf{K}^{\frac{1}{2}}|_E^2,$$

where $\mathbf{B}_k \mathbf{K} = (\boldsymbol{\psi}, \boldsymbol{\varphi}_k)$. Since $\sum_0^\infty |\mathbf{B}_k \mathbf{K}^{\frac{1}{2}}|_E^2 < \infty$, we have $\sum_{k=n}^\infty |\mathbf{B}_k \mathbf{K}^{\frac{1}{2}}|_E^2 \to 0$, as $n \to \infty$. Hence $\boldsymbol{\psi} = \mathbf{0}$. Thus $\mathfrak{N}_{-\infty} = \{0\}$.

Conversely, if $\mathfrak{M}_{-\infty} = \{0\}$, then by 5.10 (c)

$$\lim_{n\to\infty} (\mathbf{f}_0 \,|\, \mathfrak{M}_{-n}) = (\mathbf{f}_0 \,|\, \mathfrak{M}_{-\infty}) = \mathbf{0},$$

i.e. the process is regular. (Q.E.D.)

7. Spectral analysis of multivariate processes

As in the simple case so in multiple, the chief advantage of turning from the time-domain analysis of stochastic processes to the *frequency-domain* or *spectral analysis* is the possibility of using the powerful methods of harmonic analysis.

Let the shift operator U of a q-ple stationary S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ have the spectral resolution

$$U = \int_0^{2\pi} e^{-i\theta}\, dE_\theta.$$

Then from (6.3) and the well-known properties of the spectral resolution, cf. e.g. Doob [2, p. 636],

$$(f_n^{(i)}, f_0^{(j)}) = (U^n f_0^{(i)}, f_0^{(j)}) = \int_0^{2\pi} e^{-ni\theta}\, d\,(E_\theta f_0^{(i)}, f_0^{(j)}). \tag{1}$$

Letting $\qquad F_{ij}(\theta) = 2\pi\,(E_\theta f_0^{(i)}, f_0^{(j)}), \qquad \mathbf{F}(\theta) = [F_{ij}(\theta)],$

we see that the covariance matrix $\mathbf{\Gamma}_n$ is itself given by

7.1 $$\mathbf{\Gamma}_n = (\mathbf{f}_n, \mathbf{f}_0) = \frac{1}{2\pi}\int_0^{2\pi} e^{-ni\theta}\, d\mathbf{F}(\theta). \tag{7.1}$$

Following convention (6.4), let us put $E_\theta \mathbf{f}_0 = (E_\theta f_0^{(1)}, \ldots, E_\theta f_0^{(q)})$. Then by 5.8 (b) $E_\theta \mathbf{f}_0 = (\mathbf{f}_0 \,|\, \mathfrak{M}_\theta)$, where $\mathfrak{M}_\theta$ is the subspace of $\mathfrak{L}_2$ consisting of vector-functions all of whose components lie in the range $\mathfrak{M}_\theta$ ($\subseteq \mathfrak{L}_2$) of the projection operator E_θ. Hence, as is easily verified,

7.2 $$\mathbf{F}(\theta) = 2\pi\,(E_\theta \mathbf{f}_0, \mathbf{f}_0) = 2\pi(\mathbf{E}_\theta \mathbf{f}_0, E_\theta \mathbf{f}_0). \tag{7.2}$$

Since the projection operators E_θ, $0 \le \theta \le 2\pi$, constitute a resolution of the identity,[1] it readily follows [5, Theorem 4] that

7.3 F *is bounded, non-decreasing and right-continuous on* $[0, 2\pi]$, *and* $\mathbf{F}(0) = \mathbf{0}$. (7.3)

In fact, from (4.12) we get

7.4
$$\begin{cases} \mathbf{F}(\theta) = \mathbf{F}_1(\theta+) \\ \mathbf{F}_1(\theta) - \mathbf{F}_1(0) = \mathbf{\Gamma}_0(\theta - \pi) + \sum_{n \neq 0} \frac{\mathbf{\Gamma}_n}{n i}(e^{ni\theta} - 1), \end{cases} \tag{7.4}$$

which shows that F is uniquely determined by the sequence $(\mathbf{\Gamma}_n)_{-\infty}^{\infty}$.

The function F satisfying (7.1) and (7.3) is called the *spectral distribution function* of the S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$. Every multiple stationary process thus possesses one and only one such function. Conversely, as was shown by Cramer [1, Theorem 5 (b)], every function F satisfying (7.3) is the spectral distribution of some multiple stationary S.P. (cf. also [5, Theorem 4]). From (7.3) it follows, as remarked after 4.2, that

7.5 F *has a derivative a.e., which has non-negative hermitian values and belongs to* $\mathbf{L}_1$. (7.5)

But F will not always be absolutely continuous, i.e. in general we will not have

$$\mathbf{F}(\theta) = \int_0^\theta \mathbf{F}'(t)\, dt, \quad 0 \le \theta \le 2\pi.$$

If F is absolutely continuous, so that by 4.9 (d)

7.6
$$\mathbf{\Gamma}_n = (\mathbf{f}_n, \mathbf{f}_0) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta}\, \mathbf{F}'(\theta)\, d\theta, \tag{7.6}$$

then F′ is called the *spectral density function* of the S.P.

From here on, it will be convenient to regard F, F′ as functions not on the interval $[0, 2\pi]$ but on the unit circle C in the complex plane. This will amount to writing $e^{i\theta}$ for the argument, where previously we have written θ. The next theorem gives a sufficient condition for the absolute continuity of F.

7.7 THEOREM. (a) *The moving-average process* $(\mathbf{f}_n)_{-\infty}^{\infty}$:

$$\mathbf{f}_n = \sum_{k=-\infty}^{\infty} \mathbf{A}_k \mathbf{g}_{n-k}, \qquad (\mathbf{g}_i, \mathbf{g}_j) = \delta_{ij}\mathbf{G}, \qquad \sum_{-\infty}^{\infty} |\mathbf{A}_k \mathbf{G}^{\frac{1}{2}}|_E^2 < \infty,$$

has an absolutely continuous spectral distribution F *such that*

---

[1] i.e. $E_0 = 0$, $E_{2\pi} = I$, $E_{\theta+0} = E_\theta$, and $\theta > \theta'$ implies $E_\theta - E_{\theta'}$ is non-negative hermitian.

$$\mathbf{F}'(e^{i\theta}) = \mathbf{\Phi}(e^{i\theta}) \cdot \mathbf{\Phi}^*(e^{i\theta}), \qquad \mathbf{\Phi}(e^{i\theta}) = \sum_{k=-\infty}^{\infty} \mathbf{A}_k \mathbf{G}^{\frac{1}{2}} e^{ki\theta}, \quad \text{a.e.}$$

(b) *If for this process,* $\mathbf{A}_k = \mathbf{0}$ *for* $k < 0$, *then*

$$\mathbf{\Phi}(e^{i\theta}) = \sum_{k=0}^{\infty} \mathbf{A}_k \mathbf{G}^{\frac{1}{2}} e^{ki\theta},$$

and either $\Delta \mathbf{\Phi}_+$ *vanishes identically, or* $\log \Delta \mathbf{F}' \in L_1$ *on* C *and*

$$\log \Delta(\mathbf{A}_0 \mathbf{G} \mathbf{A}_0^*) \le \frac{1}{2\pi} \int_0^{2\pi} \log \Delta \{\mathbf{F}'(e^{i\theta})\} \, d\theta.$$

Proof. (a) By 5.11 (a)

$$\mathbf{\Gamma}_n = (\mathbf{f}_n, \mathbf{f}_0) = \left(\sum_{k=-\infty}^{\infty} \mathbf{A}_{n+k} \mathbf{g}_{-k}, \sum_{k=-\infty}^{\infty} \mathbf{A}_k \mathbf{g}_{-k} \right) = \sum_{k=-\infty}^{\infty} \mathbf{A}_{n+k} \mathbf{G} \mathbf{A}_k^*.$$

Next, since $\sum_{-\infty}^{\infty} |\mathbf{A}_k \mathbf{G}^{\frac{1}{2}}|_E^2 < \infty$, therefore by 3.9 (b) the functions $\mathbf{\Phi}, \mathbf{\Phi}^* \in \mathbf{L}_2$, and so $\mathbf{\Phi}\mathbf{\Phi}^* \in \mathbf{L}_1$. Since $\mathbf{A}_k \mathbf{G}^{\frac{1}{2}}$, $\mathbf{G}^{\frac{1}{2}} \mathbf{A}_{-k}^*$ are the kth Fourier coefficients of $\mathbf{\Phi}$, $\mathbf{\Phi}^*$, respectively, it follows from 3.9 (d) that the nth Fourier coefficient of $\mathbf{\Phi}\mathbf{\Phi}^*$ is

$$\sum_{k=-\infty}^{\infty} \mathbf{A}_k \mathbf{G}^{\frac{1}{2}} \mathbf{G}^{\frac{1}{2}} \mathbf{A}_{k-n}^* = \sum_{k=-\infty}^{\infty} \mathbf{A}_k \mathbf{G} \mathbf{A}_{k-n}^* = \mathbf{\Gamma}_n.$$

Thus

$$\int_0^{2\pi} e^{-ni\theta} \mathbf{\Phi}(e^{i\theta}) \cdot \mathbf{\Phi}^*(e^{i\theta}) \, d\theta = 2\pi \mathbf{\Gamma}_n = \int_0^{2\pi} e^{-ni\theta} d\mathbf{F}(e^{i\theta}),$$

where $\mathbf{F}$ is the spectral distribution. We easily conclude from this (cf. 4.9 (d) and 4.13 (b)) that

$$\mathbf{F}(\theta) = \int_0^{\theta} \mathbf{\Phi}(e^{it}) \mathbf{\Phi}^*(e^{it}) \, dt,$$

i.e. $\mathbf{F}$ is absolutely continuous, and $\mathbf{F}' = \mathbf{\Phi}\mathbf{\Phi}^*$, a.e.

(b) The new expression for $\mathbf{\Phi}$ is obvious. Since $\mathbf{\Phi} \in \mathbf{L}_2$ and its nth Fourier coefficient vanishes for $n < 0$, it follows by 3.13 (b) that either $\Delta \mathbf{\Phi}_+$ vanishes identically or $\log \Delta \mathbf{\Phi} \in L_1$ on C and

$$\log |\Delta(\mathbf{A}_0 \mathbf{G}^{\frac{1}{2}})| = \log |\Delta\{\mathbf{\Phi}_+(0)\}| \le \frac{1}{2\pi} \int_0^{2\pi} \log |\Delta\{\mathbf{\Phi}(e^{i\theta})\}| \, d\theta.$$

The desired result follows, since

$$\Delta(\mathbf{A}_0 \mathbf{G} \mathbf{A}_0^*) = |\Delta(\mathbf{A}_0 \mathbf{G}^{\frac{1}{2}})|^2, \qquad \Delta \mathbf{F}' = |\Delta \mathbf{\Phi}|^2.$$

(Q.E.D.)

On combining this theorem with Theorem 6.11 on the Wold decomposition we get the following useful result.

7.8 Main Lemma I. *Let* $(\mathbf{f}_n)_{-\infty}^{\infty}$ *be a non-deterministic* S.P., *let* $\mathbf{u}_n$, $\mathbf{v}_n$, $\mathbf{A}_k$, $\mathbf{G}$ *be as in* 6.11, *and let* $\mathbf{F}$, $\mathbf{F}_u$, $\mathbf{F}_v$ *be the spectral distributions of the* $\mathbf{f}_n-$, $\mathbf{u}_n-$, $\mathbf{v}_n-$ *processes. Then*

(a) $$\mathbf{F} = \mathbf{F}_u + \mathbf{F}_v;$$

(b) $\mathbf{F}_u$ *is absolutely continuous, and*

$$\mathbf{F}'_u(e^{i\theta}) = \mathbf{\Phi}(e^{i\theta})\,\mathbf{\Phi}^*(e^{i\theta}), \qquad \mathbf{\Phi}(e^{i\theta}) = \sum_{k=0}^{\infty} \mathbf{A}_k \mathbf{G}^{\frac{1}{2}} e^{ki\theta}, \qquad \mathbf{A}_k\mathbf{G} = (\mathbf{f}_0, \mathbf{g}_{-k});$$

(c) *if* $(\mathbf{f}_n)_{-\infty}^{\infty}$ *has full rank then* $\log \Delta \mathbf{F}'_u \in L_1$ *on* C, *and*

$$\log \Delta(\mathbf{G}) \leq \frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'_u(e^{i\theta})\}\, d\theta.$$

Proof. (a) By 6.11 (a), $\mathbf{u}_m \perp \mathbf{v}_n$, and therefore

$$(\mathbf{f}_n, \mathbf{f}_0) = (\mathbf{u}_n + \mathbf{v}_n, \mathbf{u}_0 + \mathbf{v}_0) = (\mathbf{u}_n, \mathbf{u}_0) + (\mathbf{v}_n, \mathbf{v}_0).$$

In other words, $$\frac{1}{2\pi}\int_0^{2\pi} e^{-ni\theta}\, d\mathbf{F} = \frac{1}{2\pi}\int_0^{2\pi} e^{-ni\theta}\, d(\mathbf{F}_u + \mathbf{F}_v),$$

whence by 4.13 (b), $\mathbf{F} = \mathbf{F}_u + \mathbf{F}_v + \text{const}$. The constant is zero, since all three functions vanish when $\theta = 0$.

(b) follows from the last theorem, since by 6.11 (b)

$$\mathbf{u}_n = \sum_{k=0}^{\infty} \mathbf{A}_k \mathbf{g}_{n-k}, \qquad (\mathbf{g}_i, \mathbf{g}_j) = \delta_{ij}\mathbf{G}, \qquad \sum_0^{\infty} |\mathbf{A}_k \mathbf{G}^{\frac{1}{2}}|_E^2 < \infty.$$

(c) If the S.P. has full rank, then $\mathbf{G}$ is invertible. The relation $\mathbf{A}_0\mathbf{G} = \mathbf{G}$ of 6.11 (b) now shows that $\mathbf{A}_0 = \mathbf{I}$. Hence $\Delta\{\mathbf{\Phi}_+(0)\} = \Delta(\mathbf{G}^{\frac{1}{2}}) \neq 0$. The integrability of $\log \Delta \mathbf{F}'_u$ and the desired inequality now follow from 7.7 (b). (Q.E.D.)

7.9 Main Lemma II. *Let* $\mathbf{F}$ *be the spectral distribution of the* S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$, *and let*

$$\mathbf{P}(z) = \sum_0^N \mathbf{A}_n z^n, \qquad \mathbf{P}(\mathbf{f}) = \sum_0^N \mathbf{A}_n \mathbf{f}_{-n}.$$

Then

(a) $$(\mathbf{P}(\mathbf{f}), \mathbf{P}(\mathbf{f})) = \frac{1}{2\pi}\int_0^{2\pi} \mathbf{P}(e^{i\theta}) \cdot d\mathbf{F}(e^{i\theta}) \cdot \mathbf{P}^*(e^{i\theta})$$

(b) $$\log \Delta(\mathbf{P}(\mathbf{f}), \mathbf{P}(\mathbf{f})) \geq \frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta + \log|\Delta \mathbf{A}_0|^2,$$

where the integral on the right may equal $-\infty$.

Proof. (a) Since $\mathbf{P}$ is continuous and $\mathbf{F}$ is of bounded variation, therefore by 4.9 (a) the RS-integral on the R.H.S. of (a) exists. Also by 5.7 (c) and (7.1)

$$\begin{aligned}\left(\sum_{m=0}^{N} \mathbf{A}_m \mathbf{f}_{-m}, \sum_{n=0}^{N} \mathbf{A}_n \mathbf{f}_{-n}\right) &= \sum_{m=0}^{N}\sum_{n=0}^{N} \mathbf{A}_m (\mathbf{f}_{-m}, \mathbf{f}_{-n}) \mathbf{A}_n^* \\ &= \sum_{m=0}^{N}\sum_{n=0}^{N} \mathbf{A}_m \cdot \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)\theta}\, d\mathbf{F}(e^{i\theta}) \cdot \mathbf{A}_n^* \\ &= \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{m=0}^{N} \mathbf{A}_m e^{mi\theta}\right) \cdot d\mathbf{F}(e^{i\theta}) \cdot \left(\sum_{m=0}^{N} \mathbf{A}_n e^{ni\theta}\right)^*\end{aligned}$$

from which (a) follows.

(b) From 4.10 and 3.11 (c)

$$\Delta\left\{\frac{1}{2\pi}\int_0^{2\pi} \mathbf{P}(e^{i\theta}) \cdot d\mathbf{F}(e^{i\theta}) \cdot \mathbf{P}^*(e^{i\theta})\right\} \geq \Delta\left\{\frac{1}{2\pi}\int_0^{2\pi} \mathbf{P}(e^{i\theta})\, \mathbf{F}'(e^{i\theta})\, \mathbf{P}^*(e^{i\theta})\, d\theta\right\}. \tag{1}$$

But since the values of $\mathbf{P}\mathbf{F}'\mathbf{P}^*$ are non-negative hermitian, therefore by 3.12

$$\begin{aligned}\log \Delta&\left\{\frac{1}{2\pi}\int_0^{2\pi} \mathbf{P}(e^{i\theta})\, \mathbf{F}'(e^{i\theta})\, \mathbf{P}^*(e^{i\theta})\, d\theta\right\} \\ &\geq \frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{P}(e^{i\theta})\, \mathbf{F}'(e^{i\theta})\, \mathbf{P}^*(e^{i\theta})\}\, d\theta \\ &\geq \frac{1}{2\pi}\int_0^{2\pi} \log|\Delta\{\mathbf{P}(e^{i\theta})\}|^2\, d\theta + \frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta.\end{aligned} \tag{2}$$

Now since each entry of $\mathbf{P}(z)$ is a polynomial in z, therefore $\Delta\{\mathbf{P}(z)\}$, being a sum of products of such entries, is itself a polynomial in z. Hence by Jensen's Theorem, the first term on the R.H.S. of (2) is

$$\geq \log|\Delta\{\mathbf{P}(0)\}|^2 = \log|\Delta(\mathbf{A}_0)|^2.$$

Also by (a), the L.H.S. of (1) is $\Delta(\mathbf{P}(f), \mathbf{P}(f))$. The desired inequality thus follows from (1) and (2). (Q.E.D.)

7.10 Main Theorem I. *The q-ple* S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$ *is of full rank if and only if it has a spectral distribution* $\mathbf{F}$ *such that* $\log \Delta \mathbf{F}' \in L_1$ on C. *In this case, if* $\mathbf{G}$ *is the prediction error matrix with lag* 1, *then*

$$\text{(A)} \qquad \Delta(\mathbf{G}) = \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta\right].$$

Proof. Let the S.P. be of full rank. Then by 6.11, we have the Wold decomposition $\mathfrak{f}_n = \mathfrak{u}_n + \mathfrak{v}_n$. Denoting the spectral distribution of the $\mathfrak{u}_n$- and $\mathfrak{v}_n$-processes by $\mathbf{F}_u$, $\mathbf{F}_v$, we have from 7.8 (c) that $\log \Delta F'_u \in L_1$ and

$$\log \Delta(\mathbf{G}) \le \frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'_u(e^{i\theta})\}\, d\theta. \tag{1}$$

Also by 7.8 (a), $\mathbf{F}' = \mathbf{F}'_u + \mathbf{F}'_v$. Since $\mathbf{F}'_v$ has non-negative hermitian values, it follows from 3.11 (c) that

$$\Delta\{\mathbf{F}'_u(e^{i\theta})\} \le \Delta\{\mathbf{F}'(e^{i\theta})\}.$$

Taking logarithms, integrating over $[0, 2\pi]$, and applying (1), we get

$$\log \Delta(\mathbf{G}) \le \frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta. \tag{2}$$

The S.P. being of full rank, $\Delta(\mathbf{G}) > 0$. The last integral cannot therefore be equal to $-\infty$. But neither can it be equal to ∞; for by 3.12 it is dominated by

$$\log \Delta\left\{\frac{1}{2\pi}\int_0^{2\pi} \mathbf{F}'(e^{i\theta})\, d\theta\right\},$$

which is finite since $\mathbf{F}' \in \mathbf{L}_1$ and $\mathbf{F}'$ has non-negative hermitian values (cf. (7.5)). Thus $\log \Delta \mathbf{F}' \in L_1$.

Next, suppose that $(\mathfrak{f}_n)_{-\infty}^{\infty}$ is such that $\log \Delta \mathbf{F}' \in L_1$. Consider the innovation

$$\mathfrak{g} = \mathfrak{f}_0 - (\mathfrak{f}_0 \,|\, \mathfrak{M}_{-1}),$$

where $\mathfrak{M}_{-1}$ is the past of $\mathfrak{f}_0$. Since $(\mathfrak{f}_0 \,|\, \mathfrak{M}_{-1}) \in \mathfrak{M}_{-1}$, we have $\mathfrak{g} = \lim_{N\to\infty} \mathfrak{g}_N$, where

$$\mathfrak{g}_N = \mathfrak{f}_0 - \sum_{n=1}^{N} \mathbf{B}_n^{(N)} \mathfrak{f}_{-n}.$$

Since the coefficient of $\mathfrak{f}_0$ is $\mathbf{I}$, it follows from 7.9 (b) that

$$\log \Delta(\mathfrak{g}_N, \mathfrak{g}_N) \ge \frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta.$$

Now let $N \to \infty$. Then by 5.7 (b), $(\mathbf{g}_N, \mathbf{g}_N) \to (\mathbf{g}, \mathbf{g}) = \mathbf{G}$, and the last inequality reduces to

$$\log \Delta(\mathbf{G}) \geq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta \{\mathbf{F}'(e^{i\theta})\}\, d\theta. \tag{3}$$

Since by assumption the last integral is finite, we conclude that $\Delta(\mathbf{G}) > 0$, i.e. the S.P. has full rank.

Finally, suppose once more that the S.P. has full rank. As just shown this entails (2) and the L-integrability of $\log \Delta \mathbf{F}'$. The last fact in turn entails (3). The inequalities (2) and (3) together yield the equality (A). (Q.E.D.)

The equality 7.10 (A) may be rephrased as a result on matrix-polynominal approximation in a "mean-square" sense with respect to a given matricial measure or weighting $\mathbf{F}$; viz.

$$\exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta\right]$$
$$= \lim_{N\to\infty} \Delta\left\{\frac{1}{2\pi}\int_0^{2\pi}\left(\mathbf{I} - \sum_{n=1}^{N} \mathbf{B}_n^{(N)} e^{ni\theta}\right)\cdot d\mathbf{F}(e^{i\theta})\cdot\left(\mathbf{I} - \sum_{n=1}^{N} \mathbf{B}_n^{(N)} e^{ni\theta}\right)^*\right\},$$

where for each N, the matrix coefficients $\mathbf{B}_1^{(N)}, \ldots, \mathbf{B}_N^{(N)}$ are chosen so as to minimise

$$\tau\left\{\frac{1}{2\pi}\int_0^{2\pi}\left(\mathbf{I} - \sum_{n=1}^{N} \mathbf{B}_n e^{ni\theta}\right) d\mathbf{F}(e^{i\theta}) \left(\mathbf{I} - \sum_{n=1}^{N} \mathbf{B}_n e^{ni\theta}\right)^*\right\}.$$

So interpreted, 7.10 (A) is seen to be a matricial extension of an identity first obtained by Szegö [9, Satz XII]; cf. Doob [2, pp. 638–639].

As a corollary of 6.10 we get the following spectral version of the Wold decomposition. This has been given by Doob [2, pp. 597–598], but his justification, based on an extension of the stochastic integrals of Cramer, seems to us to be inadequate, cf. sec. 1.

7.11 MAIN THEOREM II. *For any full rank* S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ *the spectral distributions of the* $\mathbf{u}_n$*- and* $\mathbf{v}_n$*-processes of its Wold decomposition* (*Theorem* 6.11) *are the absolutely continuous and discontinuous plus singular parts of its spectral distribution. In symbols, cf.* (4.3),

$$\mathbf{F}_u = \mathbf{F}^{(a)}, \quad \mathbf{F}_v = \mathbf{F}^{(d)} + \mathbf{F}^{(s)}.$$

Proof. As we already know from 7.8 that $\mathbf{F} = \mathbf{F}_u + \mathbf{F}_v$, and that $\mathbf{F}_u$ is absolutely continuous, it will suffice to show that $\mathbf{F}_v' = 0$ a.e. on C. Now from 7.8

$$\mathbf{F}' = \mathbf{F}'_u + \mathbf{F}'_v = \mathbf{\Phi} \cdot \mathbf{\Phi}^* + \mathbf{F}'_v, \quad \mathbf{\Phi} \in \mathbf{L}_2.$$

Since the values of $\mathbf{F}'_v$ are non-negative hermitian, we get from 3.11 (b)

$$\frac{\Delta\{\mathbf{F}'(e^{i\theta})\}}{\tau\{\mathbf{F}'(e^{i\theta})\}} \geq \frac{\Delta\{\mathbf{F}'_u(e^{i\theta})\}}{\tau\{\mathbf{F}'_u(e^{i\theta})\}} = \frac{\Delta\{\mathbf{F}'_u(e^{i\theta})\}}{|\mathbf{\Phi}(e^{i\theta})|_E^2};$$

whence
$$\Delta\{\mathbf{F}'(e^{i\theta})\} \geq \Delta\{F'_u(e^{i\theta})\} \cdot \left[1 + \frac{\tau\{\mathbf{F}'_v(e^{i\theta})\}}{|\mathbf{\Phi}(e^{i\theta})|_E^2}\right].$$

On taking logarithms and integrating over $[0, 2\pi]$ we get

$$\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta$$
$$\geq \int_0^{2\pi} \log \Delta\{\mathbf{F}'_u(e^{i\theta})\}\, d\theta + \int_0^{2\pi} \log\left[1 + \frac{\tau\{\mathbf{F}'_v(e^{i\theta})\}}{|\mathbf{\Phi}(e^{i\theta})|_E^2}\right] d\theta$$
$$\geq 2\pi \log \Delta(\mathbf{G}) + \int_0^{2\pi} \log\left[1 + \frac{\tau\{\mathbf{F}'_v(e^{i\theta})\}}{|\mathbf{\Phi}(e^{i\theta})|_E^2}\right] d\theta, \tag{1}$$

the last step being a consequence of the inequality in 7.8 (c), which is available since the S.P. has full rank. But by 7.10 (A), the integral on the L.H.S. of (1) equals $2\pi \log \Delta(\mathbf{G})$. Since the integrand on the R.H.S. of (1) is non-negative, it follows that

$$\log\left[1 + \frac{\tau\{\mathbf{F}'_v(e^{i\theta})\}}{|\mathbf{\Phi}(e^{i\theta})|_E^2}\right] = 0 \quad \text{a.e.},$$

i.e.
$$\tau\{\mathbf{F}'_v(e^{i\theta})\} / |\mathbf{\Phi}(e^{i\theta})|_E^2 = 0, \quad \text{a.e.}$$

Now since $\mathbf{\Phi} \in \mathbf{L}_2$, the denominator can become infinite only on a set of zero measure. The numerator $\tau\{\mathbf{F}'_v(e^{i\theta})\}$ must therefore vanish a.e. Since $\mathbf{F}'_v(e^{i\theta})$ is non-negative hermitian, we conclude that $\mathbf{F}'_v$ vanishes a.e. (Q.E.D.)

As another corollary we obtain the spectral criterion for regularity with full rank, announced by Zasuhin [18, Theorem 3], but not so far fully established (cf. sec. 1):

7.12 Main Theorem III. *The* S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$ *is regular and of full rank, if and only if it has an absolutely continuous spectral distribution* $\mathbf{F}$ *such that* $\log \Delta \mathbf{F}' \in L_1$. *In this case we have, of course, the equality* 7.10 (A).

Proof. Let the S.P. be regular and of full rank. Since it has full rank, therefore by 7.10, $\log \Delta \mathbf{F}' \in L_1$. Also,

$$\mathfrak{f}_n = \mathfrak{u}_n + \mathfrak{v}_n, \tag{1}$$

where $(\mathfrak{u}_n)_{-\infty}^{\infty}$ is a one-sided moving average, and $\mathfrak{v}_n = (\mathfrak{f}_n | \mathfrak{M}_{-\infty})$, $\mathfrak{M}_{-\infty}$ being the remote past of $(\mathfrak{f}_n)_{-\infty}^{\infty}$. Now since $(\mathfrak{f}_n)_{-\infty}^{\infty}$ is regular, $\mathfrak{M}_{-\infty} = \{\mathfrak{0}\}$ by 6.13 (b), and therefore $\mathbf{v}_n = \mathbf{0}$. Hence by (1), $\mathbf{f}_n = \mathbf{u}_n$. The S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ is thus a one-sided moving average. Hence by 7.7 its spectral distribution is absolutely continuous.

Next suppose $(\mathbf{f}_n)_{-\infty}^{\infty}$ has an absolutely continuous spectral distribution $\mathbf{F}$ such that $\log \Delta \mathbf{F}' \in L_1$. Since $\log \Delta \mathbf{F}' \in L_1$, therefore by 7.10, the S.P. has full rank. Hence by 6.11 we have the Wold decomposition (1) and by 7.11, $\mathbf{F}_v = \mathbf{F}^{(d)} + \mathbf{F}^{(s)}$. But since $\mathbf{F}$ is absolutely continuous $\mathbf{F}^{(d)} = \mathbf{0} = \mathbf{F}^{(s)}$. Hence $\mathbf{F}_v = \mathbf{0}$. It follows that $\mathbf{v}_n = \mathbf{0}$, and hence from (1) that $(\mathbf{f}_n)_{-\infty}^{\infty}$ is a one-sided moving average process. By 6.13 (a), it must be regular. (Q.E.D.)

We shall now derive a factorization theorem for matrix-valued functions, which in a sense is a matricial extension of 2.8. This had been forseen by Wiener in [17], but his proof was incomplete in that the required spectral condition for regularity had not been duly established.

Let $\mathbf{F}$ be a non-negative hermitian matrix-valued function on $\mathbf{C}$ such that $\mathbf{F} \in \mathbf{L}_1$ and $\log \Delta \mathbf{F} \in L_1$. Letting

$$\hat{\mathbf{F}}(\theta) = \int_0^{\theta} \mathbf{F}(t)\, dt, \quad 0 \leq \theta \leq 2\pi,$$

it follows from Cramer's Theorem [1, Theorem 5 (b)] that there exists a q-ple stationary process $(\mathbf{f}_n)_{-\infty}^{\infty}$ with spectral distribution $\hat{\mathbf{F}}$. Since $\hat{\mathbf{F}}$ is absolutely continuous and $\log \Delta \hat{\mathbf{F}}' \in L_1$, we see from 7.12 that this process is regular and of full rank. By 6.13 its remote past $\mathfrak{M}_{-\infty} = \{\mathbf{0}\}$, and hence by the Wold decomposition 6.11 and 6.12, it is a one-sided moving average:

$$\mathbf{f}_n = \sum_{k=0}^{\infty} \mathbf{C}_k \mathbf{h}_{n-k}, \qquad (\mathbf{h}_i, \mathbf{h}_j) = \delta_{ij} \quad \mathbf{I}, \sum_0^{\infty} |\mathbf{C}_k|_E^2 < \infty.$$

In this $\mathbf{C}_0 = \sqrt{\mathbf{G}}$, where $\mathbf{G}$ is the prediction error matrix of the process. By 7.7 it follows that

$$\mathbf{F} = \hat{\mathbf{F}}' = \mathbf{\Phi}\mathbf{\Phi}^* \text{ a.e.}, \qquad \mathbf{\Phi}(e^{i\theta}) = \sum_0^{\infty} \mathbf{C}_k e^{ki\theta}.$$

We note that $\mathbf{\Phi}_+(0) = \mathbf{C}_0 = \sqrt{\mathbf{G}}$ is non-negative hermitian. Moreover by 7.10 (A)

$$\Delta \{\mathbf{\Phi}_+(0)\}^2 = \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log \Delta \{\mathbf{F}(e^{i\theta})\}\, d\theta\right].$$

We have thus proved the following theorem.

7.13 Theorem. *Given a non-negative hermitian matrix-valued function* $\mathbf{F}$ *on* C *such that* $\mathbf{F} \in \mathbf{L}_1$ *and* $\log \Delta \mathbf{F} \in L_1$, *there exists a function* $\mathbf{\Phi} \in \mathbf{L}_2$ *on* C, *the* n-*th Fourier coefficient of which vanishes for* $n < 0$, *such that*

$$\mathbf{F}(e^{i\theta}) = \mathbf{\Phi}(e^{i\theta}) \cdot \mathbf{\Phi}^*(e^{i\theta}) \quad \text{a.e.},$$

$\mathbf{\Phi}_+(0)$ *is non-negative hermitian, and*

$$\Delta\{\mathbf{\Phi}_+(0)\}^2 = \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}(e^{i\theta})\}\, d\theta\right].$$

An important problem in multiple prediction is given such a function $\mathbf{F}$, to find the Fourier coefficients of the factor $\mathbf{\Phi}$. As it stands, our proof of 7.13 gives no clue to determining these coefficients. We shall show in Part II, however, that if we combine the arguments used in this proof with Wiener's idea of resorting to successive alternating projections, then we do get a method of finding these coefficients.

8. Some unsettled questions

The relationship between the prediction error matrix $\mathbf{G}$ and the spectral distribution $\mathbf{F}$ needs further investigation. In 7.10 (A) we have only been able to relate the determinant of G with that of $\mathbf{F}'$. But it is clear that all stochastic processes with the same spectral distribution $\mathbf{F}$ would have the same $\mathbf{G}$, and that a stronger relation exists between $\mathbf{G}$ and $\mathbf{F}$.

No spectral characterisation has yet been given of a S.P. having a rank $\varrho < q$, and in particular of a non-deterministic process, $\varrho \geq 1$. This question is tied up with the one mentioned in the previous para: a full fledged relation between $\mathbf{G}$ and $\mathbf{F}$ would yield the desired characterisations. But it may be possible to tackle questions of rank even without the knowledge of such a relation.

We might also mention a lacuna, which will confront us in Part II. This concerns the necessary and sufficient condition that the linear prediction with lead ν be expressible in terms of the past elements $\mathfrak{f}_k$, $k \leq 0$, by a single infinite series converging in-the-mean. This question remains open even in the case of a simple process. We shall show in Part II that a sufficient condition is that $\mathbf{F}$ be absolutely continuous and the eigenvalues of $\mathbf{F}'$ be essentially bounded above and away from zero.

References

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

[2]. J. L. DOOB, *Stochastic Processes*. New York, 1953.

[3]. G. H. HARDY, J. E. LITTLEWOOD & G. POLYA, *Inequalities*. Cambridge, 1934.

[4]. E. HILLE, *Functional Analysis and Semi-Groups*. New York, 1948.

[5]. A. KOLMOGOROV, Stationary sequences in Hilbert space. (Russian.) *Bull. Math. Univ., Moscou*, 2, No. 6 (1941), 40 pp. (English translation by NATASHA ARTIN.)

[6]. ——, Interpolation und Extrapolation von stationären zufälligen Folgen. *Bull. Acad. Sci. (Nauk) U.R.S.S. Ser. Math.*, 5 (1941), 3–14.

9† – 573805. *Acta mathematica*. 98. Imprimé le 21 novembre 1957.

[7]. R. E. A. C. PALEY & N. WIENER, *Fourier Transforms in the Complex Domain.* New York, 1933.
[8]. M. H. STONE, *Linear Transformations in Hilbert Space and their Application to Analysis.* New York, 1932.
[9]. G. SZEGÖ, Beiträge zur Theorie der Toeplitzchen Formen. *Math. Z.*, 6 (1920), 167–202.
[10]. ——, Über die Randwerte analytischer Funktionen. *Math. Ann.*, 84 (1921), 188–212.
[11]. ——, *Orthogonal Polynomials.* New York, 1939.
[12]. E. C. TITCHMARSH, *The Theory of Functions.* Oxford, 1939.
[13]. P. WHITTLE, The analysis of multiple stationary time series. *J. Roy. Statist. Soc., Ser. B.*, 15 (1953), 125–139.
[14]. N. WIENER, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series.* New York, 1950.
[15]. ——, Comprehensive view of prediction theory. *Proc. Int. Congress Math., Harvard*, Vol. (ii), (1950), 308–21.
[16]. ——, Mathematical problems of communication theory. *M.I.T.* (1953) (Mimeographed).
[17]. ——, On the factorization of matrices. *Comment. Math. Helv.*, 29 (1955), 97–111.
[18]. V. ZASUHIN, On the theory of multidimensional stationary random processes. (Russian.) *C. R. (Doklady) Acad. Sci. U.R.S.S.*, 33 (1941), 435.
[19]. A. ZYGMUND, *Trigonometrical Series.* Warsaw, 1935.