THE PREDICTION THEORY OF MULTIVARIATE STOCHASTIC PROCESSES, II

The Linear Predictor

BY

N. WIENER and P. MASANI [1]

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

Contents

1. Introduction

In this paper we shall obtain a linear predictor for a multivariate discrete parameter stationary stochastic process (S.P.) having a spectral density matrix $\mathbf{F}'$, the eigenvalues of which are bounded above and away from zero. To get this we shall de-

(1) This paper, like Part I [12], contains the research we carried out at the Indian Statistical Institute, Calcutta, during 1955–56, along with some simplifications resulting from later work. We would again like to thank the authorities for the excellent facilities placed at our disposal, and Dr. G. Kallianpur for valuable discussions.

Since writing this paper we have learned that some of our results in Part I have been duplicated by H. Helson and D. Lowdenslager, cf. their paper, "Prediction theory and Fourier series in several variables", to be published in this volume of *Acta Mathematica*. We regret that no reference was made to this fact in Part I. In a recent note [*Proc. Nat. Acad. Sci., U.S.A.*, Vol. 43 (1957) pp. 898–992] M. Rosenblatt has derived Theorem 7.10 proved by us in Part I, but his derivation is based on an incorrect lemma. To rectify this one would have to go through the steps followed in our Part I.

velop a coordinate-free algorithm for determining the *generating function* (1) of such a process. In the course of this development we shall obtain an expression for the prediction error matrix $\mathbf{G}$ with lag 1 in terms of $\mathbf{F}'$, thereby clearing up an important lacuna in the theory (cf. [12, Sec. 8]). We shall extensively use the theory of multivariate processes developed in our previous paper [12], and adhere to the notation followed therein. Numerical references prefixed by I are to this paper.

In Sec. 2 we shall enunciate the *prediction problem* for a q-variate stationary process, and show how it can be tackled by the solution of a system of linear equations. This involves matrix inversion. A computationally more efficient approach will be shown to depend on the delicate problem of determining the *generating function* of the process. This is difficult for $q>1$ on account of the non-commutativity of matrix multiplication. In Sec. 3 we shall describe the genesis of our algorithm for accomplishing this from Wiener's original idea of using successive alternating projections in Hilbert space [11]. In Sec. 4 we shall show that if $\mathbf{F}$ is the spectral distribution function of a q-variate, regular, full-rank process $(\mathfrak{f}_n)_{-\infty}^{\infty}$ [I, Sec. 6], then the class $\mathbf{L}_{2,F}$ of $q \times q$ matrix-valued functions, which are square-integrable with respect to the (matricial) spectral measure $\mathbf{F}$ is isomorphic to the space $\mathfrak{M}_\infty$, spanned by the random vector-valued functions $\mathfrak{f}_k$, $-\infty < k < \infty$. In Sec. 5 we shall introduce the *boundedness condition* mentioned in the previous paragraph, and show that the sum of manifolds $\sum_0^\infty \mathfrak{S}(\mathfrak{f}_{-k})$ then becomes topologically closed and therefore identical to the present and past of $\mathfrak{f}_0$, that the reciprocal of the generating function of the process has a Fourier series without negative frequencies, and that the linear prediction with lead ν is given in the time-domain by a unique infinite series $\sum_0^\infty \mathbf{E}_{\nu k}\, \mathfrak{f}_{-k}$ converging in-the-mean, where the matrix coefficients $\mathbf{E}_{\nu k}$ depend on the Fourier coefficients of the generating function and its reciprocal. In Sec. 6 we shall establish (rigorously) the algorithm mentioned in Sec. 3 for getting the generating function and its reciprocal under the boundedness condition, and derive an expression for the *linear predictor* and the *prediction error matrix* in terms of the spectral density; we shall thereby complete the solution of the prediction problem. In Sec. 7 we shall show that the boundedness assumption is fulfilled whenever the spectral density is estimated from correlation matrices, which are themselves computed from time-series

---

(1) By this we mean the function $\boldsymbol{\Phi} = \sum_0^\infty \mathbf{A}_k \mathbf{G}^{\frac{1}{2}} e^{ki\theta}$ of [12, 7.8] in which the coefficients $\mathbf{A}_k$ and $\mathbf{G}$ are as in the Wold Decomposition [12, 6.11]. För a regular, full-rank S.P. see Def. 2.6 below.

observations, i.e. in a large number of practical cases. Finally in Sec. 8 we shall show how the ideas introduced in Secs. 3 and 6 lead to a *general factorization procedure* valid even when the matrix to be factored is not hermitian-valued. We shall also show that in the hermitian case the factorization so obtained is unique up to a constant unitary factor.

The rest of this section will be devoted to recalling some necessary parts of the theory developed in [I] and to introducing supplementary material of an ancillary nature. We shall first explain our notation.

Notation. *As an* [I], *bold face letters* $\mathbf{A}$, $\mathbf{B}$, *etc. will denote* $q \times q$ *matrices with complex entries* a_{ij}, b_{ij}, *etc., and bold face letters* $\mathbf{F}$, $\mathbf{G}$, *etc. will denote functions whose values are such matrices. The symbols* τ, Δ, * *will be reserved for the trace, determinant and adjoint of matrices.* Ω *will stand for a space having a Borel field of subsets over which is defined a probability measure* P. *Bold face letters* $\mathbf{x}$, $\mathbf{y}$ *etc. will refer to* q-*dimensional column vectors with complex components* x_i, y_i, *etc., and bold face letters* $\mathfrak{f}$, $\mathfrak{g}$, *etc. to (random) functions defined over the space* Ω, *whose values are such vectors.* $\mathfrak{L}_2$ *will designate the set of such functions* $\mathfrak{f}$ *with components* $f^{(i)}$ *such that* $\int_\Omega |f^{(i)}(\omega)|^2 dP(\omega) < \infty$, [I, 5.1]. *For* $\mathfrak{f}, \mathfrak{g} \in \mathfrak{L}_2$, $(\mathfrak{f}, \mathfrak{g})$ *will denote the Gramian matrix* $[(f^{(i)}, g^{(j)})]$. $\mathfrak{S}(\boldsymbol{\varphi}_j)_{j\in J}$ *will denote the (closed) subspace spanned by the functions* $\boldsymbol{\varphi}_j \in \mathfrak{L}_2$, *for* $j \in J$, *linear combinations being taken with matrix coefficients* [I, 5.6], *and* $(\mathfrak{f}\,|\,\mathfrak{M})$ *the orthogonal projection of* $\mathfrak{f}$ *on the subspace* $\mathfrak{M}$ [I, 5.9]. *The letters* C, D_+, D_- *will refer to the sets* $|z| = 1$, $|z| < 1$, $1 < |z| \leqslant \infty$ *of the extended complex plane.*

Next, we recall [I, 3.2] that the $q \times q$ matrices with complex entries form a Banach algebra under the usual algebraic operations and either the Banach or Euclidean norms:

1.1
$$\left.\begin{aligned} |\mathbf{A}|_B &= \operatorname*{l.u.b.}_{\mathbf{x} \neq 0} \frac{|\mathbf{A}\mathbf{x}|}{|\mathbf{x}|} \\ |\mathbf{A}|_E &= \{\tau(\mathbf{A}\mathbf{A}^*)\}^{\frac{1}{2}} = \left\{\sum_{i=1}^{q}\sum_{j=1}^{q} |a_{ij}|^2\right\}^{\frac{1}{2}}. \end{aligned}\right\} \tag{1.1}$$

It follows of course that

1.2
$$|\mathbf{A}+\mathbf{B}| \leqslant |\mathbf{A}| + |\mathbf{B}|, \qquad |\mathbf{A}\mathbf{B}| \leqslant |\mathbf{A}|\,|\mathbf{B}| \quad \text{(either norm)} \tag{1.2}$$

1.3
$$|\mathbf{A}^*| = |\mathbf{A}| \quad \text{(either norm)}. \tag{1.3}$$

But we also have the following inequality.

1.4 Lemma. $|\mathbf{A}\,\mathbf{B}|_E \leqslant |\mathbf{A}|_B |\mathbf{B}|_E,\ |\mathbf{A}|_E |\mathbf{B}|_B.$

Proof. Let $\mathbf{a}_1', \ldots, \mathbf{a}_q'$ be the rows of $\mathbf{A}$ and $\mathbf{b}_1, \ldots, \mathbf{b}_q$ the columns of $\mathbf{B}$. Then denoting by $\mathbf{a}_i' \mathbf{b}_j$ the (i, j)th entry of $\mathbf{A}\,\mathbf{B}$, we hawe

$$|\mathbf{A}\,\mathbf{B}|_E^2 = \sum_{i=1}^{q} \sum_{j=1}^{q} |\mathbf{a}_i' \mathbf{b}_j|^2. \tag{1}$$

Now $\mathbf{A}\,\mathbf{b}_j$ is the column vector $(\mathbf{a}_1' \mathbf{b}_j, \ldots, \mathbf{a}_q' \mathbf{b}_j)$. Hence

$$\sum_{i=1}^{q} |\mathbf{a}_1' \mathbf{b}_j|^2 = |\mathbf{A}\,\mathbf{b}_j|^2 \leqslant |\mathbf{A}|_B^2 |\mathbf{b}_j|^2.$$

From (1) we therefore get

$$|\mathbf{A}\,\mathbf{B}|_E^2 \leqslant |\mathbf{A}|_B^2 \sum_{j=1}^{q} |\mathbf{b}_j|^2 = |\mathbf{A}|_B^2 |\mathbf{A}|_E^2,$$

i.e.

$$|\mathbf{A}\,\mathbf{B}|_E \leqslant |\mathbf{A}|_B |\mathbf{B}|_E. \tag{2}$$

Since by (1.3) $|\mathbf{A}\,\mathbf{B}|_E = |\mathbf{B}^* \mathbf{A}^*|_E$, and by (2) and (1.3)

$$|\mathbf{B}^* \mathbf{A}^*|_E \leqslant |\mathbf{B}^*|_B |\mathbf{A}^*|_E = |\mathbf{B}|_B |\mathbf{A}|_E,$$

we get $|\mathbf{A}\,\mathbf{B}|_E \leqslant |\mathbf{A}|_E |\mathbf{B}|_B$. (Q.E.D.)

We shall also need the following simple properties of hermitian matrices, which we shall not, however, prove.

Notation. *If* $\mathbf{A}$, $\mathbf{B}$ *are hermitian, we shall write* $\mathbf{A} \prec \mathbf{B}$ *or* $\mathbf{B} \succ \mathbf{A}$ *to mean that* $\mathbf{B} - \mathbf{A}$ *is non-negative.*

1.5 Lemma. *If* λ, μ *are the smallest and largest eigenvalues of a hermitian matrix* $\mathbf{H}$, *then*

(a) $\lambda \mathbf{I} \prec \mathbf{H} \prec \mu \mathbf{I}$.

(b) $|\mathbf{H}|_B = \max \{|\lambda|, |\mu|\}$.

(c) $\left|\frac{2}{\lambda+\mu}\mathbf{H} - \mathbf{I}\right|_B = \frac{\mu-\lambda}{\mu+\lambda}$, *provided* $\mu + \lambda > 0$.

(d) $\lambda\,\mathbf{A}\,\mathbf{A}^* \prec \mathbf{A}\,\mathbf{H}\,\mathbf{A}^* \prec \mu\,\mathbf{A}\,\mathbf{A}^*$.

To turn to matrix-valued functions we recall [I, 3.4, 3.5] that for $\delta \geqslant 1$ the set $\mathbf{L}_\delta$ of functions $\mathbf{F} = [f_{ij}]$ on C such that each f_{ij} is measurable and $\int_0^{2\pi} |f_{ij}(e^{i\theta})|^\delta \, d\theta < \infty$ is a *Banach space* under the usual operations and the norm

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

$\mathbf{L}_2$ is moreover a *Hilbert space* under these operations and the inner product

1.6
$$((\mathbf{F}, \mathbf{G})) = \frac{1}{2\pi} \int_0^{2\pi} \tau \{\mathbf{F}(e^{i\theta}) \mathbf{G}^*(e^{i\theta})\} \, d\theta, \tag{1.6}$$

the corresponding norm $\| \ \|$ being the same as $| \ |_2$:

1.7
$$\|\mathbf{F}\| = \sqrt{((\mathbf{F}, \mathbf{F}))} = \left\{ \frac{1}{2\pi} \int_0^{2\pi} |\mathbf{F}(e^{i\theta})|_E^2 \, d\theta \right\}^{\frac{1}{2}} = |\mathbf{F}|_2. \tag{1.7}$$

The set $\mathbf{L}_\infty$ of functions $\mathbf{F}$ on C with measurable and essentially bounded entries is a *Banach algebra* under the usual operations and the norm

1.8
$$|\mathbf{F}|_\infty = \underset{0 \leqslant \theta \leqslant 2\pi}{\text{ess. l.u.b.}} \, |\mathbf{F}(e^{i\theta})|_E. \tag{1.8}$$

It remains a Banach algebra, if in the last relation we take the Banach norm instead of the Euclidian.

The *Lebesgue integral* of a matrix-valued function $\mathbf{F}$ is the matrix obtained by integrating each entry of $\mathbf{F}$ [I, 3.6]. Some simple properties of this integral are listed in the next two lemmas, the proofs of which are obvious.

1.9 Lemma. (a) *If* $\mathbf{F} \in \mathbf{L}_1$, *then*

$$\left| \int_0^{2\pi} \mathbf{F}(e^{i\theta}) \, d\theta \right| \leqslant \int_0^{2\pi} |\mathbf{F}(e^{i\theta})| \, d\theta \quad (\textit{either norm}).$$

(b) *If* $\mathbf{F} \in \mathbf{L}_1$, *and is non-negative hermitian valued a.e., then* $\int_0^{2\pi} \mathbf{F}(e^{i\theta}) \, d\theta$ *is non-negative hermitian.*

(c) $\mathbf{F} \in \mathbf{L}_1$ *implies* $\tau \mathbf{F} \in \mathbf{L}_1$. *The converse holds, provided that the values of* $\mathbf{F}$ *are non-negative hermitian a.e., and its entries are measurable functions.*

1.10 Lemma. (*Schwarz inequality*). *If* $\mathbf{F}, \mathbf{G} \in \mathbf{L}_2$, *then*

(a) $\mathbf{F}\mathbf{G} \in \mathbf{L}_1$

(b) $$\left| \frac{1}{2\pi} \int_0^{2\pi} \mathbf{F}(e^{i\theta}) \mathbf{G}(e^{i\theta}) \, d\theta \right|_E \leqslant |\mathbf{F}\mathbf{G}|_1 \leqslant \frac{1}{2\pi} \int_0^{2\pi} |\mathbf{F}(e^{i\theta})|_E |\mathbf{G}(e^{i\theta})|_E \, d\theta \leqslant \|\mathbf{F}\| \cdot \|\mathbf{G}\|.$$

Proof. (a) follows from I, 3.5(a), and (b) from 1.9(a), (1.2) and the ordinary Schwarz inequality. (Q.E.D.)

A simple application of Lemma 1.4 yields:

1.11 Lemma. *$\mathbf{F} \in \mathbf{L}_2$ and $\mathbf{G} \in \mathbf{L}_\infty$ implies $\mathbf{F}\mathbf{G} \in \mathbf{L}_2$, and*

$$\|\mathbf{F}\mathbf{G}\| \leqslant \|\mathbf{F}\| \cdot M,$$

where $\|\mathbf{F}\|$ is as in (1.7), and $M = \operatorname*{ess.\,l.u.b.}_{0\leqslant\theta\leqslant 2\pi} |\mathbf{F}(e^{i\theta})|$, with either norm.

We recall the *Riesz-Fischer Theorem* I, 3.9 (b), which asserts that

1.12
$$\left.\begin{array}{l} (\mathbf{A}_n)_{-\infty}^{\infty} \textit{ is the sequence of Fourier coefficients of a} \\ \textit{function in } \mathbf{L}_2, \textit{ if and only if } \sum_{-\infty}^{\infty} |\mathbf{A}_n|_E^2 < \infty. \end{array}\right\} \tag{1.12}$$

For $\mathbf{F} \in \mathbf{L}_2$ with Fourier coefficients $\mathbf{A}_k$ we have the *Parseval relations* I, 3.9 (c):

1.13
$$\left.\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} \mathbf{F}(e^{i\theta})\,\mathbf{F}^*(e^{i\theta})\,d\theta &= \sum_{-\infty}^{\infty} \mathbf{A}_k \mathbf{A}_k^*, \\ \|\mathbf{F}\|^2 &= \sum_{-\infty}^{\infty} |\mathbf{A}_k|_E^2. \end{aligned}\right\} \tag{1.13}$$

An important consequence of (1.12) is that if $\mathbf{A}_n$ is the nth Fourier coefficient of a function in $\mathbf{L}_2$, and the sequence $(\mathbf{B}_n)_{-\infty}^{\infty}$ is composed of $\mathbf{A}_n$'s and zeros, then $\mathbf{B}_n$ is also the nth Fourier coefficient of a function in $\mathbf{L}_2$. This suggests a departure from I, 3.10 in the usage of the subscripts $+$, $-$ for functions in $\mathbf{L}_p$ when $p \geqslant 2$:

1.14 Definition. (a) *For $p \geqslant 1$, $\mathbf{L}_p^+$, $\mathbf{L}_p^{0+}$, $\mathbf{L}_p^-$, $\mathbf{L}_p^{0-}$ will denote the subsets of all functions in $\mathbf{L}_p$ whose n-th Fourier coefficients vanish for $n \leqslant 0$, $n < 0$, $n \geqslant 0$, $n > 0$, respectively.*

(b) *If $\mathbf{F} \in \mathbf{L}_p$, where $p \geqslant 2$, and has Fourier coefficients $\mathbf{A}_k$, $-\infty < k < \infty$, then $\mathbf{F}_+$, $\mathbf{F}_{0+}$, $\mathbf{F}_-$, $\mathbf{F}_{0-}$ will denote the functions in $\mathbf{L}_2^+$, $\mathbf{L}_2^{0+}$, $\mathbf{L}_2^-$, $\mathbf{L}_2^{0-}$, whose n-th Fourier coefficients are $\mathbf{A}_n$ for $n > 0$, $n \geqslant 0$, $n < 0$, $n \leqslant 0$, respectively (and zero for the remaining n). $\mathbf{F}_0$ will denote the constant function with value $\mathbf{A}_0$.*

From this definition and the relations (1.12), (1.13) we readily get the following lemma.

1.15 Lemma. (a) *The sets $\mathbf{L}_2^+$, $\mathbf{L}_2^{0+}$, $\mathbf{L}_2^-$, $\mathbf{L}_2^{0-}$ are (closed) subspaces of the Hilbert space $\mathbf{L}_2$, and $\mathbf{L}_2^+ \perp \mathbf{L}_2^{0-}$, $\mathbf{L}_2^{0+} \perp \mathbf{L}_2^-$.*

(b) *If* $\mathbf{F} \in \mathbf{F}_2$, *then*

(1) $\mathbf{F} = \mathbf{F}_- + \mathbf{F}_0 + \mathbf{F}_+ = \mathbf{F}_{0-} + \mathbf{F}_+ = \mathbf{F}_- + \mathbf{F}_{0+}$.

(2) $\|\mathbf{F}\|^2 = \|\mathbf{F}_-\|^2 + \|\mathbf{F}_0\|^2 + \|\mathbf{F}_+\|^2 = \|\mathbf{F}_{0-}\|^2 + \|\mathbf{F}_+\|^2 = \|\mathbf{F}_-\|^2 + \|\mathbf{F}_{0+}\|^2$.

(3) $\|\mathbf{F}_+\|, \|\mathbf{F}_{0+}\|, \|\mathbf{F}_-\|, \|\mathbf{F}_{0-}\| \leqslant \|\mathbf{F}\|$.

(4) $(\mathbf{F}_+)^* = (\mathbf{F}^*)_-, \quad (\mathbf{F}_-)^* = (\mathbf{F}^*)_+$.

Another fact we will require, which is well known, is stated in the next lemma.

1.16 LEMMA. *The bounded linear operators* $\mathfrak{P}$ *on the Banach space* $\mathbf{L}_2$ *into itself form a Banach algebra* $\mathfrak{B}$ *under the usual operations and the Banach norm*

$$|\mathfrak{P}| = \underset{\mathbf{F} \neq 0}{\text{l.u.b.}} \frac{\|\mathfrak{P}(\mathbf{F})\|}{\|\mathbf{F}\|}.$$

Finally we will need the following simple results on the Gramians of random vector-valued functions in $\mathfrak{L}_2$, the proofs of which are immediate from I, 5.8, 5.9.

1.17 LEMMA. (a) *If* $\mathfrak{M}$ *is a subspace of* $\mathfrak{L}_2$ *and* $\hat{\mathfrak{f}} = (\mathfrak{f} \mid \mathfrak{M})$ (*cf.* I, 5.9), *then for* all $\mathfrak{g} \in \mathfrak{M}$

$$(\mathfrak{f} - \mathfrak{g}, \mathfrak{f} - \mathfrak{g}) \succ (\mathfrak{f} - \hat{\mathfrak{f}}, \mathfrak{f} - \hat{\mathfrak{f}}).$$

(b) *If* $\mathfrak{M} = \text{clos.} \bigcup_1^\infty \mathfrak{M}_n$, *where each* $\mathfrak{M}_n$ *is a subsqace and* $\mathfrak{M}_{n+1} \supseteq \mathfrak{M}_n$, *then*

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

2. The prediction problem

Let $(\mathfrak{f}_n)_{-\infty}^{\infty}$ be a q-ple stationary S.P. and let $\mathfrak{M}_0 = \mathfrak{S}(\mathfrak{f}_k)_{-\infty}^0$ be the present and past of $\mathfrak{f}_0$, [I, sec. 6]. Then we define the *linear prediction of* $\mathfrak{f}_n$ *with lead* n *by* $\hat{\mathfrak{f}}_n = (\mathfrak{f}_n \mid \mathfrak{M}_0)$, [I, 5.9]. Since $\hat{\mathfrak{f}}_n \in \mathfrak{M}_0$, it will follow that

2.1 $$\hat{\mathfrak{f}}_n(\omega) = \underset{N \to \infty}{\text{l.i.m.}} \sum_{j=0}^{N} \mathbf{A}_j^{(N)} \mathfrak{f}_{-j}(\omega), \quad n > 0, \tag{2.1}$$

where the $\mathbf{A}_j^{(N)}$ are certain (non-unique) $q \times q$ matrices.

Now for a fixed ω in the probability space Ω the values $\mathbf{x}_j = \mathfrak{f}_j(\omega)$, $-\infty < j < \infty$ constitute a *multiple time series* or in Doob's terminology a *sample function* of the S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$. The components of the past values $\mathbf{x}_{-j}$, $j > 0$, of such a time series can be found from observation. Hence if the matrices $\mathbf{A}_j^{(N)}$ can be determined, we can evaluate the sum occurring in (2.1), and for sufficiently large N, treat it as an approximation to the linear prediction $\hat{\mathbf{x}}_n = \hat{\mathbf{f}}_n(\omega)$ of the value $\mathbf{x}_n$ of this time series at the future time n. Hence an important problem in prediction is to determine the $\mathbf{A}_j^{(N)}$.

In the Wiener-Kolmogorov theory the quantities supposed to be known or given in terms of which the $\mathbf{A}_j^{(N)}$ are to be determined are the correlation matrices $\mathbf{\Gamma}_n = (\mathbf{f}_n, \mathbf{f}_0)$, [I, (6.1)]. This theory has its basis in the case in which the shift operator of the process is generated by a measure-preserving transformation of the probability space Ω onto itself. We shall show in Sec. 7 how under the assumption of ergodicity, the $\mathbf{\Gamma}_n$ may be estimated from time series data. Alternatively, it may be possible in certain cases to hypothesize the values of $\mathbf{\Gamma}_n$ from a theoretical study of the process without recourse to sampling. We may therefore formulate the prediction problem as follows.

2.2 Prediction Problem. *Let $(\mathbf{f}_n)_{-\infty}^{\infty}$ be a q-ple stationary S.P. with a given, known covariance sequence $(\mathbf{\Gamma}_n)_{-\infty}^{\infty}$, and let $\mathfrak{M}_0$ be the present and past of $\mathbf{f}_0$. To determine*

(i) *the $q \times q$ matrices $\mathbf{A}_j^{(N)}$ in the formula* (2.1) *for the prediction $\hat{\mathbf{f}}_n$ of $\mathbf{f}_n$ with lead n,*

(ii) *the prediction error matrix for lead n:*

$$\mathbf{G}_n = (\mathbf{f}_n - \hat{\mathbf{f}}_n, \mathbf{f}_n - \hat{\mathbf{f}}_n).$$

Seemingly the easiest way of solving this problem is by an extension of the method of undetermined coefficients. Since $\hat{\mathbf{f}}_n = (\mathbf{f}_n \mid \mathfrak{M}_0)$, we may choose the $\mathbf{A}_j^{(N)}$ so that

$$\sum_{j=0}^{N} \mathbf{A}_j^{(N)} \mathbf{f}_{-j} = (\mathbf{f}_n \mid \mathfrak{S}(\mathbf{f}_{-k})_0^N), \quad n > 0.$$

Then by 1.17 (b), the $\mathbf{A}_j^{(N)}$ will satisfy (2.1). Also [I, 5.8 (b)]

$$\mathbf{f}_n - \sum_{j=0}^{N} \mathbf{A}_j^{(N)} \mathbf{f}_{-j} \perp \mathbf{f}_0, \mathbf{f}_{-1}, \dots, \mathbf{f}_{-N}.$$

Hence for each $k = 0, \dots, N$,

$$0 = \left(\sum_{j=0}^{N} \mathbf{A}_j^{(N)} \mathbf{f}_{-j} - \mathbf{f}_n, \mathbf{f}_{-k} \right) = \sum_{j=0}^{N} \mathbf{A}_j^{(N)} (\mathbf{f}_{-j}, \mathbf{f}_{-k}) - (\mathbf{f}_n, \mathbf{f}_{-k}),$$

i.e.

$$\sum_{j=0}^{N} \mathbf{A}_j^{(N)} \mathbf{\Gamma}_{k-j} = \mathbf{\Gamma}_{n+k}, \quad k = 0, \dots, N.$$

This system of $N+1$ equations in the $N+1$ unknown matrices $\mathbf{A}_j^{(N)}$, $j=0, \ldots, N$, is equivalent to a single matrix equation, which in block notation may be written

$$[\mathbf{A}_0^{(N)} \ldots \mathbf{A}_N^{(N)}] \begin{bmatrix} \mathbf{\Gamma}_0 & \ldots & \mathbf{\Gamma}_{-N} \\ \vdots & & \vdots \\ \mathbf{\Gamma}_N & \ldots & \mathbf{\Gamma}_0 \end{bmatrix} = [\mathbf{\Gamma}_n \ldots \mathbf{\Gamma}_{n+N}]. \tag{1}$$

In this the first factor on the left and the term on the right are $q \times (N+1)q$ matrices, and the second factor on the left is a $(N+1)q \times (N+1)q$ matrix, which we shall denote by $\mathbf{\Gamma}$. If $\mathbf{\Gamma}$ is invertible, we get

$$[\mathbf{A}_0^{(N)} \ldots \mathbf{A}_N^{(N)}] = [\mathbf{\Gamma}_n \ldots \mathbf{\Gamma}_{n+N}] \begin{bmatrix} \mathbf{\Gamma}_0 & \ldots & \mathbf{\Gamma}_{-N} \\ \vdots & & \vdots \\ \mathbf{\Gamma}_N & & \mathbf{\Gamma}_0 \end{bmatrix}^{-1} \tag{2}$$

from which the unknowns $\mathbf{A}_j^{(N)}$ can be found. We shall now show that $\mathbf{\Gamma}$ is invertible, if the S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$ has full rank, cf. [I, Sec. 6, p. 136].

Let $\mathbf{B}_0, \ldots, \mathbf{B}_N$ be any $q \times q$ matrices and consider the $q \times (N+1)q$ matrix

$$\mathbf{B} = [\mathbf{B}_0, \ldots, \mathbf{B}_N].$$

A simple calculation shows that

$$\left.\begin{aligned} \mathbf{B}\mathbf{\Gamma}\mathbf{B}^* &= [\mathbf{B}_0 \ldots \mathbf{B}_N] \begin{bmatrix} \mathbf{\Gamma}_0 & \ldots & \mathbf{\Gamma}_{-N} \\ \vdots & & \vdots \\ \mathbf{\Gamma}_N & \ldots & \mathbf{\Gamma}_0 \end{bmatrix} \begin{bmatrix} \mathbf{B}_0^* \\ \vdots \\ \mathbf{B}_N^* \end{bmatrix} \\ &= \sum_{j=0}^{N} \sum_{k=0}^{N} \mathbf{B}_j \mathbf{\Gamma}_{j-k} \mathbf{B}_k^* = \left(\sum_0^N \mathbf{B}_j \mathfrak{f}_j, \sum_0^N \mathbf{B}_k \mathfrak{f}_k \right). \end{aligned}\right\} \tag{3}$$

Now take $\mathbf{B}_N$ to be invertible. Then

$$\sum_0^N \mathbf{B}_j \mathfrak{f}_j = \mathbf{B}_N \left(\mathfrak{f}_N + \sum_0^{N-1} \mathbf{B}_N^{-1} \mathbf{B}_j \mathfrak{f}_j \right) = \mathbf{B}_N (\mathfrak{f}_N - \mathfrak{g}),$$

where $\mathfrak{g}$ is in $\mathfrak{M}_{N-1}$, the past of $\mathfrak{f}_N$. Hence from (3)

$$\mathbf{B}\mathbf{\Gamma}\mathbf{B}^* = \mathbf{B}_N (\mathfrak{f}_N - \mathfrak{g}, \mathfrak{f}_N - \mathfrak{g}) \mathbf{B}_N^*. \tag{4}$$

Now by 1.17 (a)

$$(\mathfrak{f}_N - \mathfrak{g}, \mathfrak{f}_N - \mathfrak{g}) \succ (\mathfrak{f}_N - \hat{\mathfrak{f}}_N, \mathfrak{f}_N - \hat{\mathfrak{f}}_N) = \mathbf{G},$$

where $\hat{\mathfrak{f}}_N = (\mathfrak{f}_N \mid \mathfrak{M}_{N-1})$ and $\mathbf{G}$ is the prediction error matrix with lag 1. Since for a full-rank process $(\mathfrak{f}_n)_{-\infty}^{\infty}$, $\Delta(\mathbf{G}) > 0$ by definition, therefore [I, 3.11 (c)]

$$\Delta (\mathfrak{f}_N - \mathfrak{g}, \mathfrak{f}_N - \mathfrak{g}) > 0.$$

Since $\mathbf{B}_N$ is invertible, it follows from (4) that so is $\mathbf{B}\boldsymbol{\Gamma}\mathbf{B}^*$ and therefore also $\boldsymbol{\Gamma}$. For a full-rank process the desired matrix coefficients can therefore be obtained from (2).

This method of solving the prediction problem involves matrix–inversion and is therefore unsuitable as a computational technique except where the matrices are small, i.e. where q is small, very short segments of the past are used, and large prediction errors tolerated. To arrive at a more accurate and efficient computational procedure we have, as often happens, to appeal to more advanced and refined analytical theory; in the present instance to the representability of a relgular S.P. as a one-sided moving average, and the factorizability of its spectral density [I, Sec. 6, 7].

To recall this theory, let $(\mathbf{f}_n)_{-\infty}^{\infty}$ be a q-ple regular, full-rank process, let $(\mathbf{h}_k)_{-\infty}^{\infty}$ be its normalised innovation process [I, 5.12], and $\mathbf{G}$ its prediction error matrix for lag 1. Then by I, 6.12

2.3
$$\mathbf{f}_n = \sum_{k=0}^{\infty} \mathbf{C}_k \mathbf{h}_{n-k}, \quad \mathbf{C}_k = (\mathbf{f}_0, \mathbf{h}_{-k}), \quad \mathbf{C}_0 = \sqrt{\mathbf{G}}. \tag{2.3}$$

By I, 6.13 (b) the remote past $\mathfrak{M}_{-\infty}$ is $\{0\}$, and so by I, 6.10 (b), $\mathfrak{M}_0$ is also the present and past of $\mathbf{h}_0$. Hence from I, 5.11 (c) and I, 5.12 (b)

2.4
$$\left.\begin{aligned} \hat{\mathbf{f}}_n = (\mathbf{f}_n \mid \mathfrak{M}_0) &= \sum_{k=n}^{\infty} \mathbf{C}_k \mathbf{h}_{n-k} \\ \mathbf{f}_n - \hat{\mathbf{f}}_n &= \sum_{k=0}^{n-1} \mathbf{C}_k \mathbf{h}_{n-k} \\ \mathbf{G}_n = (\mathbf{f}_n - \hat{\mathbf{f}}_n, \mathbf{f}_n - \hat{\mathbf{f}}_n) &= \sum_{k=0}^{n-1} \mathbf{C}_k \mathbf{C}_k^*. \end{aligned}\right\} \tag{2.4}$$

To solve the Prediction Problem we have only to determine the matrices $\mathbf{C}_k$ *and the random functions* $\mathbf{h}_k$. Now by I, 7.7 and I, 7.10 (A), if $\mathbf{F}'$ is the spectral density function of the process, then

2.5
$$\left.\begin{aligned} &\mathbf{F}'(e^{i\theta}) = \boldsymbol{\Phi}(e^{i\theta}) \cdot \boldsymbol{\Phi}^*(e^{i\theta}), \quad \text{a.e} \\ &\boldsymbol{\Phi}(e^{i\theta}) = \sum_{k=0}^{\infty} \mathbf{C}_k e^{ki\theta} \in \mathbf{L}_2^{0+} \\ &\boldsymbol{\Phi}_+(0) = \sqrt{\mathbf{G}} \ \text{(which is non-negative hermitian)} \\ &\Delta\{\boldsymbol{\Phi}_+(0)\}^2 = \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log \Delta\{\mathbf{F}'(e^{i\theta})\}\, d\theta\right]. \end{aligned}\right\} \tag{2.5}$$

2.6 Definition. *We shall call* $\mathbf{\Phi} = \sum_{k=0}^{\infty} \mathbf{C}_k e^{ki\theta} \in \mathbf{L}_2^{0+}$, *where the* $\mathbf{C}_k$ *are as in* (2.4), generating function *of the S.P.*

If in some way we can determine the generating function $\mathbf{\Phi}$ or its Fourier coefficients $\mathbf{C}_k$ the only unknowns left in (2.4) would be the random functions $\mathbf{h}_n$. The task before us is therefore to devise a constructive or algorithmic method for finding $\mathbf{\Phi}$ or the coefficients $\mathbf{C}_k$. To see the difficulties in finding this, take $q=1$, i.e. suppose that F' is complex-valued. Its factorization can then be effected as in the proof of I, 2.8. We first obtain the Fourier coefficients a_k of $\log F'(e^{i\theta})$. (1) We then compute the Fourier coefficients C_k of the factor Φ from the equation

$$\sum_0^{\infty} C_k z^k = \exp\left(\tfrac{1}{2} a_0 + \sum_1^{n} a_n z^n\right),$$

by expanding the R.H.S. and equating coefficients of like powers of z. By the Uniqueness Theorem I, 2.9 the C_k so determined will be the desired coefficients. (2) This method will not, however, work for $q>1$, since matrix multiplication is non-commutative and the exponential law

$$\exp(\mathbf{A}+\mathbf{B}) = \exp \mathbf{A} \cdot \exp \mathbf{B}$$

breaks down.

The problem of determining the generating function thus presents fresh difficulties when $q>1$. In Sec. 6 we shall give an algorithmic solution of the problem, valid under a condition of boundedness 5.1, the significance of which will be discussed in Sec. 7. But to get the form of this algorithm we will have to reverse the shift from the time-domain to the frequency-domain made in passing from (2.4) to (2.5), and to find the $\mathbf{C}_k$ from the equations $\mathbf{C}_k = (\mathbf{f}_0, \mathbf{h}_{-k}) = (\mathbf{f}_k, \mathbf{h}_0)$, $k>0$, after expressing $\mathbf{h}_0$ linearly in terms of the $\mathbf{f}_{-k}$ by alternating projections (Sec. 3). In

---

(1) In practice we can find the logarithm from tables or by a cam or analogue computer, and get its Fourier coefficients by a harmonic analyser.

(2) For $q=1$ the method of factoring usually followed in communication engineering, and confined mainly to the continuous parameter case, is to approximate to F' by a rational function, and to determine the zeros and poles of the latter by numerical solution of polynomial equations. The zeros and poles in D_+ are then separated from those in D_-, and the factors Φ, Φ^* isolated. Cf. Wiener [10, 2.03] and Bode and Shannon [2].

This method is motivated by the fact that only filters having rational transfer functions in the frequency domain can be synthesised out of lumped passive elements. As long as we rely on analogue computers to do the prediction, this fact is crucial. But it ceases to be relevant if the computation is to be carried out digitally, as would be more accurate and otherwise more appropriate in the discrete parameter case, since it would obviate the necessity of interpolating. In digital computation it would be more natural to follow our method of factoring than that of rational approximation.

Secs. 5, 6 we shall also show that under the boundedness condition the Fourier coefficients of $\mathbf{\Phi}$ and $\mathbf{\Phi}^{-1}$ can be utilized to get the random functions $\mathbf{h}_{-k}$ of (2.4) as well, and thereby to complete the solution of the Prediction Problem.

3. The alternating process

In this section our approach will be heuristic. We will outline Wiener's idea of using successive alternating projections on Hilbert space in order to derive the components of the innovation function in the 2-ple case, and show how when approached from an operator-theoretical standpoint it suggests a coordinate-free algorithm for determining the generating function.

We shall begin with a lemma on the spectral densities of the component processes of a multiple process.

3.1 Lemma. (a) *Let* $\mathbf{F}=[F_{ij}]$ *be a* $q\times q$ *non-negative hermitian matrix-valued function on* C. *If* $\mathbf{F}\in \mathbf{L}_1$ *and* $\log \Delta\mathbf{F}\in L_1$, *then* $\log F_{ii}\in L_1$ *for* $1\leqslant i\leqslant q$.

(b) *The component processes of a regular, full rank process are regular.*

Proof. (a) Since the values of F are non-negative hermitian, we have $\Delta\mathbf{F}\leqslant \leqslant F_{11}\dots F_{qq}$, whence

$$\log \Delta\mathbf{F}\leqslant \log F_{11}+\cdots+\log F_{qq}\leqslant \log F_{ii}+\tau(\mathbf{F}).$$

Hence
$$\log \Delta\mathbf{F}-\tau(\mathbf{F})\leqslant \log F_{ii}< F_{ii}\quad \text{a.e.}$$

The extreme terms being in L_1, so is the middle term.

(b) If $(\mathbf{f}_n)_{-\infty}^{\infty}$ is regular and has full rank, then by I, 7.12 its spectral distribution $\mathbf{F}$ is absolutely continuous and $\log \Delta\mathbf{F}'\in L_1$. It follows that F_{ii} is absolutely continuous, and by (a) that $\log F'_{ii}\in L_1$. Hence by I, 7.12 the component process $(f_n^{(i)})_{n=-\infty}^{\infty}$ is regular. (Q.E.D.)

Now let $F'=[F'_{ij}]$ be a non-negative hermitian matrix-valued function on C such that $\mathbf{F}'\in L_1$ and $\log \Delta\mathbf{F}'\in L_1$. For notational simplicity we shall suppose that $\mathbf{F}$ is 2×2. By Cramer's Theorem [3, Theorem 5 (b)] and I, 7.12, $\mathbf{F}'$ is the spectral density of a 2-ple regular, full rank process $(\mathbf{f}_n)_{-\infty}^{\infty}$. If $(\mathbf{h}_n)_{-\infty}^{\infty}$ is its normalized innovation process, then, cf. (2.3)–(2.5),

3.2
$$\mathbf{F}'(e^{i\theta})=\mathbf{\Phi}(e^{i\theta})\cdot\mathbf{\Phi}^*(e^{i\theta}),\qquad \mathbf{\Phi}(e^{i\theta})=\sum_0^{\infty}\mathbf{C}_k e^{ki\theta}, \tag{3.2}$$

where Φ is the generating function of the process. Here $\mathbf{G}_k = (\mathbf{f}_0, \mathbf{h}_{-k}) = (\mathbf{f}_k, \mathbf{h}_0)$. Now [I, 6.12] $\mathbf{h}_0 = \sqrt{\mathbf{G}^{-1}} \cdot \mathbf{g}_0$, where $\mathbf{g}_0$ is the innovation function of $(\mathbf{f}_n)_{-\infty}^{\infty}$, and $\mathbf{G} = (\mathbf{g}_0, \mathbf{g}_0)$ is its prediction error matrix. Hence

3.3
$$\mathbf{C}_k = (\mathbf{f}_k, \sqrt{\mathbf{G}^{-1}} \cdot \mathbf{g}_0) = (\mathbf{f}_k, \mathbf{g}_0) \sqrt{\mathbf{G}^{-1}}. \tag{3.3}$$

The matrices $(\mathbf{f}_k, \mathbf{g}_0)$, $(\mathbf{g}_0, \mathbf{g}_0)$ involved in this can be calculated, once we have expressed $\mathbf{g}_0$ linearly in terms of the $\mathbf{f}_{-m}$. If for instance the coefficients $\mathbf{A}_k^{(N)}$ in

3.4
$$\mathbf{g}_0 = \lim_{N\to\infty} \sum_{m=0}^{N} \mathbf{A}_m^{(N)} \mathbf{f}_{-m} \tag{3.4}$$

are known, then we can determine

$$(\mathbf{f}_k, \mathbf{g}_0) = \lim_{N\to\infty} \sum_{m=0}^{\infty} (\mathbf{f}_k, \mathbf{f}_{-m}) \mathbf{A}_m^{(N)*},$$

since the Fourier coefficients $(\mathbf{f}_k, \mathbf{f}_{-m}) = \mathbf{\Gamma}_{k+m}$ of $\mathbf{F}'$ are known beforehand. *Our problem is therefore to express* $\mathbf{g}_0$ *linearly in terms of* $\mathbf{f}_{-m}$, $m \geq 0$.

Now by I, (6.8) $\mathbf{g}_0 = \mathbf{f}_0 - \hat{\mathbf{f}}_0$, where $\hat{\mathbf{f}}_0 = (\mathbf{f}_0 \mid \mathfrak{M}_{-1})$, and $\mathfrak{M}_{-1}$ is the past of $\mathbf{f}_0$. Letting $\mathfrak{M}_{-1}^{(i)}$ be the past of $f_0^{(i)}$, we have by I, 5.8 (b) (e) and I, 6.5

$$\hat{f}_0^{(i)} = \{f_0^{(i)} \mid \text{clos.}\ (\mathfrak{M}_{-1}^{(1)} + \mathfrak{M}_{-1}^{(2)})\}$$

and therefore

3.5
$$g_0^{(i)} = f_0^{(i)} - \hat{f}_0^{(i)} = (f_0^{(i)} \mid \{\text{clos.}\ (\mathfrak{M}_{-1}^{(1)} + \mathfrak{M}_{-1}^{(2)})\}^{\perp}). \tag{3.5}$$

Now let $(h_n^{(j)})_{n=-\infty}^{\infty}$ be the normalized innovation process of the simple process $(f_n^{(j)})_{n=-\infty}^{\infty}$. Since by 3.1 (b) the latter process is regular, $\{h_n^{(j)}\}_{n=-\infty}^{-1}$ will be an orthonormal basis for the subspace $\mathfrak{M}_{-1}^{(j)}$. By determining the generating function and thence solving the Prediction Problem 2.2 for the simple processes $(f_n^{(j)})_{n=-\infty}^{\infty}$, $j = 1, 2$, we can determine the random functions $h_n^{(j)}$, and from these obtain $(f_0^{(i)} \mid \mathfrak{M}_{-1}^{(j)})$, $j = 1, 2$. The problem before us is therefore *to determine the projection of* $f_0^{(i)}$ *on the space* $\{\text{clos.}\ (\mathfrak{M}_{-1}^{(1)} + \mathfrak{M}_{-1}^{(2)})\}^{\perp}$, *given its projections on the spaces* $\mathfrak{M}_{-1}^{(1)}$ *and* $\mathfrak{M}_{-1}^{(2)}$. So formulated the problem is seen to rest on the following theorem (von Neumann [9, p. 55]). (1)

3.6 THEOREM. (*Alternating projections*). *Let* P_1, P_2, P *be projection operators on a Hilbert space* $\mathfrak{H}$ *onto the subspaces* $\mathfrak{M}_1$, $\mathfrak{M}_2$, $\mathfrak{M}_1 \cap \mathfrak{M}_2$. *If* A_n *is the n-th term of either of the sequences*

---

(1) In [11] WIENER proved this theorem, unaware that it was already known to VON NEUMANN.

$$P_1,\ P_2P_1,\ P_1P_2P_1,\ P_2P_1P_2P_1,\ \ldots$$
$$P_2,\ P_1P_2,\ P_2P_1P_2,\ P_1P_2P_1P_2,\ \ldots.$$

then $A_n \to P$ *strongly* (1), *as* $n \to \infty$.

This at once yields the following corollary, which is really what we need.

3.7 COROLLARY. *With the notation of* 3.6, *if* Q *is the projection operator on* $\mathfrak{H}$ *onto the subspace* $\{\text{clos.}\ (\mathfrak{M}_1 + \mathfrak{M}_2)\}^\perp$, *then*

(a) $$Q = I - P_1 - P_2 + P_1P_2 + P_2P_1 - P_1P_2P_1 - P_2P_1P_2 + \cdots$$

the convergence being in the strong sense;

(b) *for all* $f \in \mathfrak{M}_i^\perp$,

$$Q(f) = f - P_j(f) + P_iP_j(f) - P_jP_iP_j(f) + \cdots, \quad j \neq i$$

the convergence being with respect to the norm in $\mathfrak{H}$.

We shall apply this corollary, taking $\mathfrak{H} = \mathfrak{L}_2$ and $\mathfrak{M}_i = \mathfrak{M}_{-1}^{(i)}$. We will be able to use the formula given in (b), which is simpler than that in (a), if $f_0^{(i)} \perp \mathfrak{M}_{-1}^{(i)}$. This condition, which by the stationarity property will imply that the process $\{f_n^{(i)}\}_{n=-\infty}^{\infty}$ is orthogonal, can be secured by an initial factorization of the diagonal entries of $\mathbf{F}'$, as we shall now show.

Since $\mathbf{F}' \in \mathbf{L}_1$, $\log \Delta \mathbf{F}' \in L_1$ and therefore by 3.1, F'_{ii}, $\log F'_{ii} \in L_1$, it follows by I, 2.8 that

$$F'_{ii} = |\phi_i|^2, \text{ where } \phi_i \in L_2^{0+}, \quad i = 1, 2. \tag{1}$$

The Fourier coefficients of ϕ_i can be found by the method explained in Sec. 2, so that we may regard ϕ_1, ϕ_2 as known. Now

$$\mathbf{F}' = \begin{bmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{bmatrix} \begin{bmatrix} 1 & F'_{12}/\phi_1\bar{\phi}_2 \\ \bar{F}'_{12}/\bar{\phi}_1\phi_2 & 1 \end{bmatrix} \begin{bmatrix} \bar{\phi}_1 & 0 \\ 0 & \bar{\phi}_2 \end{bmatrix}. \tag{2}$$

In this the first and third matrices on the right are in $\mathbf{L}_2^{0+}$ and $\mathbf{L}_2^{0-}$, and the one in the middle, which we shall denote by $\hat{\mathbf{F}}$, is well defined a.e. on C, since by (1) the functions ϕ_i can vanish almost nowhere on C. If $\hat{\mathbf{F}}$ can be factored, then from (2) we would get a factorization for $\mathbf{F}'$. Now $\hat{\mathbf{F}}$ does fulfill the conditions of factorizability, [I, 7.13], viz. $\hat{\mathbf{F}} \in \mathbf{L}_1$ and $\log \Delta \hat{\mathbf{F}} \in L_1$. For since $|F'_{12}|^2 \leqslant F'_{11}F'_{22} = |\phi_1|^2 |\phi_2|^2$, therefore $|F'_{ij}/\phi_i\bar{\phi}_j| \leqslant 1$. Thus $\hat{\mathbf{F}}$ is in $\mathbf{L}_\infty$ and therefore in $\mathbf{L}_1$. Also since,

(1) i.e. for each $f \in \mathfrak{H}$, $|A_n(f) - P(f)| \to 0$, as $n \to \infty$, $|\ |$ being the norm in $\mathfrak{H}$.

$$\log \Delta \mathbf{F}' = \log |\phi_1|^2 + \log |\phi_2|^2 + \log \Delta \hat{\mathbf{F}},$$

therefore
$$\log \Delta \hat{\mathbf{F}} = \log \Delta \mathbf{F}' - (\log F'_{11} + \log F'_{22}),$$

which is in L_1 by 3.1 (a). Hence there is no loss of generality in assuming, to begin with, that the diagonal entries of $\mathbf{F}'$ are 1.

With this assumption, the component processes $(f_n^{(i)})_{n=-\infty}^{\infty}$ are orthonormal. Hence $f_0^{(i)} \perp \mathfrak{M}_{-1}^{(i)}$, and the formula 3.7 (b) can be used to get

$$g_0^{(i)} = f_0^{(i)} - P_j (f_0^{(i)} + P_i P_j (f_0^{(i)}) - P_j P_i P_j (f_0^{(i)}) + \cdots \quad j \neq i, \tag{3}$$

where P_j is the projection operator on $\mathfrak{L}_2$ onto the subspace $\mathfrak{M}_{-1}^{(j)}$. Since $(f_{-m}^{(j)})_{m=1}^{\infty}$ is an orthogonal basis of this subspace,

$$\left.\begin{aligned} P_j (f_0^{(i)}) &= \sum_{m=1}^{\infty} (f_0^{(i)}, f_{-m}^{(j)})\, f_{-m}^{(j)} \\ P_i P_j (f_0^{(i)}) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (f_0^{(i)}, f_{-m}^{(j)})\, (f_{-m}^{(j)}, f_{-n}^{(i)})\, f_{-n}^{(i)} \end{aligned}\right\} \tag{4}$$

and so on. The coefficients involved, viz.

$$a_m = (f_m^{(1)}, f_0^{(2)}), \quad b_m = (f_m^{(2)}, f_0^{(1)}) = \bar{a}_{-m},$$

are the Fourier coefficients of the non-diagonal entries of $\mathbf{F}$, and are therefore known. From (3) and (4) we get

$$\left.\begin{aligned} g_0^{(1)} = f_0^{(1)} &- \sum_m f_{-m}^{(2)} a_m + \sum_m \sum_n f_{-m}^{(1)} b_{m-n} a_n - \sum_m \sum_n \sum_p f_{-m}^{(2)} a_{m-n} b_{n-p} a_p \\ &+ \sum_m \sum_n \sum_p \sum_q f_{-m}^{(1)} b_{m-n} a_{n-p} b_{p-q} a_q - \cdots, \\ g_0^{(2)} = f_0^{(2)} &- \sum_m f_{-m}^{(1)} b_m + \sum_m \sum_n f_{-m}^{(2)} a_{m-n} b_n - \sum_m \sum_n \sum_p f_{-m}^{(1)} b_{m-n} a_{n-p} b_p \\ &+ \sum_m \sum_n \sum_p \sum_q f_{-m}^{(2)} a_{m-n} b_{n-p} a_{p-q} b_q - \cdots, \end{aligned}\right\} \tag{5}$$

where all subscripts run from 1 to ∞.

It would not be permissible to separate the terms in $f_{-m}^{(1)}$ from those in $f_{-m}^{(2)}$ in these series. We may do so, however, in their partial sums $g_{0,N}^{(i)}$ consisting of N terms, so that

$$\left.\begin{aligned} g^{(1)}_{0,N} &= f^{(1)}_0 + \sum_m f^{(1)}_{-m}\{\sum_n b_{m-n}a_n + \sum_n\sum_p\sum_q b_{m-n}a_{n-p}b_{p-q}a_q + \cdots\} \\ &\quad - \sum_m f^{(2)}_{-m}\{a_m + \sum_n\sum_p a_{m-n}b_{n-p}a_p + \cdots\} \\ g^{(2)}_{0,N} &= -\sum_m f^{(1)}_{-m}\{b_m + \sum_n\sum_p b_{m-n}a_{n-p}b_p + \cdots\} + f^{(2)}_0 \\ &\quad + \sum_m f^{(2)}_{-m}\{\sum_n a_{m-n}b_n + \sum_n\sum_p\sum_q a_{m-n}b_{n-p}a_{p-q}b_q + \cdots\}, \end{aligned}\right\} \tag{6}$$

or in matrix notation,

$$\begin{aligned} \mathbf{g}_{0,N} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\mathbf{f}_0 - \sum_m \Bigg\{ \begin{bmatrix}0 & a_m\\ b_m & 0\end{bmatrix} + \sum_n \begin{bmatrix} b_{m-n}a_n & \\ 0 & a_{m-n}b_n\end{bmatrix} \\ - \sum_n\sum_p \begin{bmatrix}0 & a_{m-n}b_{n-p}a_p \\ b_{m-n}a_{n-p}b_p & 0\end{bmatrix} + \cdots \Bigg\}\mathbf{f}_{-m}, \end{aligned} \tag{7}$$

where there are a finite number of terms, depending on N, between each pair of braces { }, and all subscripts run from 1 to ∞.

The matrix coefficient of $\mathbf{f}_{-m}$ in the last expression is what we have denoted by $\mathbf{A}^{(N)}_m$ in (3.4). These coefficients are thus determined. The desired coefficients $\mathbf{C}_k$ can be gotten from the $\mathbf{A}^{(N)}_m$ as explained earlier, cf. (3.3) and ensuing remarks. (1) For $q > 2$ an analogous method based on q projections can be worked out, but the expressions for the coefficients $\mathbf{C}_k$ will appear different for different q, and will be hard to handle for large q. As it stands, this approach is therefore unsuitable as a computational algorithm.

We shall now indicate how a reinterpretation of the idea underlying this solution leads to a procedure for finding the generating function which is valid for any $q \geqslant 2$. This is obtained when we try to derive a sequence of operations on the space of matrix-valued functions of the type used by Masani [6] from the sequence of alternating projections of $\mathfrak{L}_2(\Omega)$ discussed above. We first note the following lemma.

3.8 LEMMA. *If* $\mathbf{g}_k = \sum_{n=0}^{N} \mathbf{A}_n \mathbf{f}_{k-n}$ $\quad -\infty < k < \infty$, *and* $\mathbf{F}'$ *is the spectral density of the* S.P. $(\mathbf{f}_n)^{\infty}_{-\infty}$, *then the spectral density* $\mathbf{G}$ *of the process* $(\mathbf{g}_k)^{\infty}_{-\infty}$ *is given by*

(1) We know that these $\mathbf{C}_k$ will lead to the factorization of $\mathbf{F}'$, only because we were able to derive such a factorization beforehand in I, 7.13, by treating $\mathbf{F}'$ as the spectral density of a full rank process. To prove I, 7.13 we had to make use of the spectral criterion for regularity with full rank given in I, 7.12. In [11] Wiener attempted to derive this criterion from the expressions (5). Such a derivation does not seem to be possible. Wiener's proof is in fact incomplete: convergence difficulties appear, which become pronounced when the pasts $\mathfrak{M}^{(1)}_{-1}$, $\mathfrak{M}^{(2)}_{-1}$ of the component processes are inclined at a zero angle.

$$\mathbf{G}(e^{i\theta}) = \left(\sum_{n=0}^{N} \mathbf{A}_n e^{ni\theta}\right) \mathbf{F}'(e^{i\theta}) \left(\sum_{n=0}^{N} \mathbf{A}_n e^{ni\theta}\right)^*.$$

Proof. First we have

$$\left(\sum_{-N}^{N} \mathbf{A}_n \mathbf{f}_{-n}, \sum_{-M}^{M} \mathbf{B}_m \mathbf{f}_m\right) = \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{-N}^{N} \mathbf{A}_n e^{ni\theta}\right) \mathbf{F}'(e^{i\theta}) \left(\sum_{-M}^{M} \mathbf{B}_m e^{ni\theta}\right)^* d\theta.$$

This can be proved in exactly the same way as I, 7.9 (a). It follows that for all integers k

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-ki\theta}\, \mathbf{G}(e^{i\theta})\, d\theta = (\mathbf{g}_k, \mathbf{g}_0)$$

$$= \left(\sum_{n=0}^{N} \mathbf{A}_n \mathbf{f}_{k-n}, \sum_{n=0}^{N} \mathbf{A}_n \mathbf{f}_{-n}\right)$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{n=0}^{N} \mathbf{A}_n e^{(n-k)i\theta}\right) \mathbf{F}'(e^{i\theta}) \left(\sum_{n=0}^{N} \mathbf{A}_n e^{ni\theta}\right)^* d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} e^{-ki\theta} \left(\sum_{n=0}^{N} \mathbf{A}_n e^{ni\theta}\right) \mathbf{F}'(e^{i\theta}) \left(\sum_{n=0}^{N} \mathbf{A}_n e^{ni\theta}\right)^* d\theta.$$

From this the result follows. (Q.E.D.)

Proceeding heuristically, let us suppose that the expressions between the four braces { } of (6) converge separately as $N \to \infty$, so that we can replace $g_{0,N}^{(i)}$ by $g_0^{(i)}$, and take infinitely many terms in each { }. The corresponding matrix version (7) will then contain $\mathbf{g}_0$ instead of $\mathbf{g}_{0,N}$ and there will be infinitely many terms between the braces { }, which give the matrix coefficient of $\mathbf{f}_{-m}$. Denoting this matrix by by $\mathbf{A}_m$, $m \geq 1$, and letting $\mathbf{A}_0 = \mathbf{I}$, we get

$$\mathbf{g}_0 = \sum_{m=0}^{\infty} \mathbf{A}_m \mathbf{f}_{-m}. \tag{8}$$

Now since the innovation process $(\mathbf{g}_n)_{-\infty}^{\infty}$ of $(\mathbf{f}_n)_{-\infty}^{\infty}$ is orthogonal [I, 6.9], therefore its spectral density has the constant value $\mathbf{G} = (\mathbf{g}_0, \mathbf{g}_0)$. A heuristic extension of the last lemma thus suggests that

$$\mathbf{G} = (\mathbf{g}_0, \mathbf{g}_0) = \mathbf{\Psi}(e^{i\theta})\, \mathbf{F}'(e^{i\theta})\, \mathbf{\Psi}^*(e^{i\theta}), \quad \mathbf{\Psi}(e^{i\theta}) = \sum_{m=0}^{\infty} \mathbf{A}_m e^{mi\theta}. \tag{9}$$

Now let $\mathbf{\Phi}$ be the generating function of our S.P. Then by (2.5)

$$\mathbf{F}' = \mathbf{\Phi}\,\mathbf{\Phi}^* = (\mathbf{\Phi}\sqrt{\mathbf{G}^{-1}})\,\mathbf{G}\,(\mathbf{\Phi}\sqrt{\mathbf{G}^{-1}})^*.$$

Since $|\Delta(\mathbf{\Phi})| = \sqrt{\Delta(\mathbf{F}')} \neq 0$, a.e. we see that $\mathbf{\Phi}^{-1}$ exists a.e. and

$$\mathbf{G} = (\sqrt{\mathbf{G}}\,\mathbf{\Phi}^{-1})\,\mathbf{F}'\,(\sqrt{\mathbf{G}}\,\mathbf{\Phi}^{-1})^*. \tag{10}$$

Now assume that $\mathbf{\Phi}^{-1} \in \mathbf{L}_2^{0+}$. Then from (9), (10) and known uniqueness theorems (e.g. I, 2.9), it would appear that $\sqrt{\mathbf{G}}\,\mathbf{\Phi}^{-1}$ and $\mathbf{\Psi}$ are equal.

This suggests a further study of the function $\mathbf{\Psi}$. Letting

$$\mathbf{B}_m = \begin{bmatrix} 0 & a_m \\ b_m & 0 \end{bmatrix},$$

we find that

3.9
$$\mathbf{A}_m = -\mathbf{B}_m + \sum_n \mathbf{B}_n \mathbf{B}_{m-n} - \sum_n \sum_p \mathbf{B}_p \mathbf{B}_{n-p} \mathbf{B}_{m-n} + \cdots, \tag{3.9}$$

where all subscripts run from 1 to ∞. Hence from (9)

$$\begin{aligned}\mathbf{\Psi}(e^{i\theta}) &= \mathbf{I} - \sum_{m=1}^{\infty} \mathbf{B}_m e^{mi\theta} + \sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} \mathbf{B}_n \mathbf{B}_{m-n} \right) e^{mi\theta} \\ &\quad - \sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} \sum_{p=1}^{\infty} \mathbf{B}_p \mathbf{B}_{n-p} \mathbf{B}_{m-n} \right) e^{mi\theta} + \cdots \\ &= \mathbf{I} - \mathbf{\Psi}_1(e^{i\theta}) + \mathbf{\Psi}_2(e^{i\theta}) - \mathbf{\Psi}_3(e^{i\theta}) + \cdots, \text{ say.}\end{aligned}$$

Now let $\mathbf{M} = \mathbf{F}' - \mathbf{I}$. Then $\mathbf{B}_m$ will be the mth Fourier coefficient of M for $m > 0$, and a straightforward calculation shows that with the subscript notation of 1.14 (b),

3.10
$$\mathbf{\Psi}_1 = \mathbf{M}_+,\ \mathbf{\Psi}_2 = (\mathbf{M}_+ \mathbf{M})_+,\ \mathbf{\Psi}_3 = \{(\mathbf{M}_+ \mathbf{M})_+ \mathbf{M}\}_+, \cdots. \tag{3.10}$$

Hence

3.11
$$\mathbf{\Psi} = \mathbf{I} - \mathbf{M}_+ + (\mathbf{M}_+ \mathbf{M})_+ - \{(\mathbf{M}_+ \mathbf{M})_+ \mathbf{M}\}_+ + \cdots. \tag{3.11}$$

Thus $\mathbf{\Psi}$ can be derived from the (known) spectral density. From (9) we get $\mathbf{G}$ and thence $\mathbf{\Phi} = \mathbf{\Psi}^{-1}\sqrt{\mathbf{G}}$.

To put this derivation on a sound footing we would have to justify the change in the order of summation made in going from (5) to (8), show that (9) is correct, that $\mathbf{\Phi}^{-1} \in \mathbf{L}_2^{0+}$ that $\sqrt{\mathbf{G}}\,\mathbf{\Phi}^{-1} = \mathbf{\Psi}$ and that the series (3.11) converges in the mean. We will follow a different approach. The crucial point that $\mathbf{\Phi}^{-1} \in \mathbf{L}_2^{0+}$ will be settled in Sec. 5 under the Boundedness Condition 5.1 on $\mathbf{F}'$. For this we will need the isomorphism between $\mathfrak{M}_\infty$ and the $\mathbf{L}_2$-class under spectral weighting, which is established in the next section. The other unsettled points will then either be circumvented or disposed of by means of the Boundedness Condition.

4. The $\mathbf{L}_2$-class with respect to spectral measure

In this section we shall study the class $\mathbf{L}_{2,F}$ of matrix-valued functions which are square-integrable with respect to the spectral distribution $\mathbf{F}$ of a regular full-rank S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$, and show that it is isomorphic to $\mathfrak{M}_\infty$, the subspace of $\mathfrak{L}_2$ spanned by the vector-valued random functions $\mathbf{f}_n$, $-\infty < n < \infty$ (cf. I, 5.6). This isomorphism will be needed in the next two sections. Throughout this section we shall assume that $\mathbf{F}$ is the spectral distribution of a q-ple S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ of this type. By I, 7.12 $\mathbf{F}$ will then be absolutely continuous. We shall therefore define $\mathbf{L}_{2,F}$ as *the class of all* $q \times q$ *matrix-valued functions* $\mathbf{\Phi}$ *on the unit circle* C *such that* $\mathbf{\Phi}\mathbf{F}'\mathbf{\Phi}^* \in \mathbf{L}_1$ *on* C. For brevity, however, it will be convenient to sometimes write θ instead of $e^{i\theta}$ for the arguments of functions in $\mathbf{L}_{2,F}$, i.e. to imagine that the domain of these functions is the closed interval $[0, 2\pi]$ and not the circle C.

From this definition of $\mathbf{L}_{2,F}$ we readily get the following lemmas, cf. (1.6) and (1.7).

4.1 LEMMA. (a) $\mathbf{\Phi} \in \mathbf{L}_{2,F}$, *if and only if* $\mathbf{\Phi}\sqrt{\mathbf{F}'} \in \mathbf{L}_2$.

(b) *If* $\mathbf{\Phi}, \mathbf{\Psi} \in \mathbf{L}_{2,F}$, *then* $\mathbf{\Phi}\mathbf{F}'\mathbf{\Psi}^* \in \mathbf{L}_1$, *and*

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

$$\frac{1}{2\pi}\int_0^{2\pi} \tau\{\mathbf{\Phi}(\theta)\mathbf{F}'(\theta)\mathbf{\Phi}^*(\theta)\}\,d\theta = \|\mathbf{\Phi}\sqrt{\mathbf{F}'}\|^2.$$

(c) *If* $\mathbf{\Phi} \in L_\infty$ and $\mathbf{\Psi} \in \mathbf{L}_{2,F}$ *then* $\mathbf{\Phi}\mathbf{\Psi} \in \mathbf{L}_{2,F}$; *in particular every Laurent polynomial in* $e^{i\theta}$ *with matrix coefficients is in* $L_{2,F}$.

4.2 Lemma. *$\mathbf{L}_{2,F}$ is a Hilbert space under the usual algebraic operations and the inner product*

$$((\mathbf{\Phi}, \mathbf{\Psi}))_F = ((\mathbf{\Phi}\sqrt{\mathbf{F}'}, \mathbf{\Psi}\sqrt{\mathbf{F}'})), \tag{1}$$

the corresponding norm being

$$\|\mathbf{\Phi}\|_F = \sqrt{((\mathbf{\Phi}, \mathbf{\Phi}))_F} = \|\mathbf{\Phi}\sqrt{\mathbf{F}'}\|. \tag{2}$$

Proof. Since $\mathbf{L}_2$ is a vector space, it follows at once from 4.1 (a) that so is $\mathbf{L}_{2,F}$. Also, by (1) $((\ ,\ ))_F$ has all the properties of an inner product, and by (2) $\|\ \|_F$ all the properties of a norm.

To show that $\mathbf{L}_{2,F}$ is complete, we let $(\mathbf{\Phi}_n)_1^\infty$ be a Cauchy sequence in $\mathbf{L}_{2,F}$. The equality

$$\|\mathbf{\Phi}_m - \mathbf{\Phi}_n\|_F = \|\mathbf{\Phi}_m\sqrt{\mathbf{F}'} - \mathbf{\Phi}_n\sqrt{\mathbf{F}'}\|$$

then shows that $(\mathbf{\Phi}_n\sqrt{\mathbf{F}'})_1^\infty$ is a Cauchy sequence in $\mathbf{L}_2$, and therefore has a limit in $\mathbf{L}_2$. Since by I, 7.12 $\log \Delta \mathbf{F}' \in L_1$ and therefore $\mathbf{F}'$ is invertible a.e., it easily follows that this limit is of the form $\mathbf{\Phi}\sqrt{\mathbf{F}'}$. By 4.1 (a) $\mathbf{\Phi} \in \mathbf{L}_{2,F}$. Since, as $n \to \infty$,

$$\|\mathbf{\Phi}_n - \mathbf{\Phi}\|_F = \|\mathbf{\Phi}_n\sqrt{\mathbf{F}'} - \mathbf{\Phi}\sqrt{\mathbf{F}'}\| \to 0,$$

we conclude that $\mathbf{\Phi}_n \to \mathbf{\Phi}$ in $\mathbf{L}_{2,F}$. (Q.E.D.)

In [*I*] we saw that although the $\mathfrak{L}_2$-norm $\|\ \|$ is important in the stochastic theory, the corresponding inner product $((\ ,\ ))$ is not and has to be replaced by the Gramian $(\ ,\ )$, cf. I, (5.2)–5.4. The situation is the same with regard to the norm $\|\ \|_F$ and the inner product $((\ ,\ ))_F$ of 4.2. What takes the place of the latter in the stochastic theory is a matrix, analogous to the Gramian, which we shall now define.

4.3 Definition. *For $\mathbf{\Phi}, \mathbf{\Psi} \in \mathbf{L}_{2,F}$ we define the matrix $(\mathbf{\Phi}, \mathbf{\Psi})_F$ by*

$$(\mathbf{\Phi}, \mathbf{\Psi})_F = \frac{1}{2\pi}\int_0^{2\pi} \mathbf{\Phi}(\theta)\, \mathbf{F}'(\theta)\, \mathbf{\Psi}^*(\theta)\, d\theta.$$

The relation between this and the inner product and norm of 4.2 is given by

4.4 $$((\mathbf{\Phi}, \mathbf{\Psi}))_F = \tau(\mathbf{\Phi}, \mathbf{\Psi})_F, \quad \|\mathbf{\Phi}\|_F = \sqrt{\tau(\mathbf{\Phi}, \mathbf{\Phi})_F}. \tag{4.4}$$

In view of 4.3 and 4.2 (1), (2) we at once get the following form of the Schwarz inequalities for $\mathbf{L}_{2,F}$ from the corresponding inequalities for $\mathbf{L}_2$ given in 1.10.

4.5 Lemma (*Schwarz inequality*). *If* $\mathbf{\Phi}, \mathbf{\Psi} \in \mathbf{L}_{2,F}$ *then*

$$|(\mathbf{\Phi}, \mathbf{\Psi})_F|_E \leqslant |\mathbf{\Phi}\mathbf{F}'\mathbf{\Psi}^*|_1 \leqslant \|\mathbf{\Phi}\|_F\|\mathbf{\Psi}\|_F.$$

We shall now turn to the isomorphism between $\mathbf{L}_{2,F}$ and $\mathfrak{M}_\infty = \mathfrak{S}(\mathbf{f}_k)_{-\infty}^{\infty}$. We shall use the following notation.

Notation. (a) *Finite linear combinations* $\sum_{-n}^{m} \mathbf{A}_k \mathbf{f}_{-k}$ *with matrix coefficients* $\mathbf{A}_k$ *will be denoted by the symbols* $\mathbf{P}(\mathbf{f})$, $\mathbf{Q}(\mathbf{f})$, *etc.*

(b) *If* $\mathbf{P}(\mathbf{f}) = \sum_{-n}^{n} \mathbf{A}_k \mathbf{f}_{-k}$, *then we shall write* $\hat{\mathbf{P}}(e^{i\theta}) = \sum_{-n}^{n} \mathbf{A}_k e^{ki\theta}$.

We shall show that the correspondence so defined between finite linear combinations of the $\mathbf{f}_{-k}$ with matrix coefficients, and Laurent polynomials in $e^{i\theta}$ with the same coefficients can be extended to all random functions in $\mathfrak{M}_\infty$ and all matrix-valued functions in $\mathbf{L}_{2,F}$ to yield an isomorphism between these spaces. We need the following lemmas.

4.6 Lemma. $(\mathbf{P}(f), \mathbf{Q}(f)) = (\hat{\mathbf{P}}, \hat{\mathbf{Q}})_F$.

Proof. We have to show that

$$\left(\sum_{-m}^{m} \mathbf{A}_j \mathbf{f}_{-j}, \sum_{-n}^{n} \mathbf{B}_k \mathbf{f}_{-k}\right) = \frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{-m}^{m} \mathbf{A}_j e^{ji\theta}\right) \mathbf{F}'(e^{i\theta}) \left(\sum_{-n}^{n} \mathbf{B}_k e^{ki\theta}\right)^* d\theta.$$

This can be done exactly as in our proof of I, 7.9 (a), if we note that since $\mathbf{F}$ is absolutely continuous

$$(\mathbf{f}_{-j}, \mathbf{f}_{-k}) = \frac{1}{2\pi}\int_0^{2\pi} e^{i(j-k)\theta}\,\mathbf{F}'(e^{i\theta})\,d\theta$$

(Q.E.D.)

4.7 Lemma. (a) $\mathbf{P}_n(\mathbf{f}) \to \boldsymbol{\varphi}$ *in* $\mathfrak{M}_\infty$, *if and only if there exists a function* $\mathbf{\Phi} \in \mathbf{L}_{2,F}$ *such that* $\hat{\mathbf{P}}_n \to \mathbf{\Phi}$ *in* $\mathbf{L}_{2,F}$.

(b) *If* $\mathbf{P}_n(\mathbf{f}) \to \boldsymbol{\varphi}$, $\mathbf{Q}_n(\mathbf{f}) \to \boldsymbol{\psi}$ *in* $\mathfrak{M}_\infty$, *and* $\mathbf{\Phi}, \mathbf{\Psi}$ correspond to $\boldsymbol{\varphi}, \boldsymbol{\psi}$ as in (a), *then*

$$(\mathbf{\Phi}, \mathbf{\Psi})_F = (\boldsymbol{\varphi}, \boldsymbol{\psi}).$$

Proof. (a) By 4.6

$$(\mathbf{P}_m(\mathfrak{f})-\mathbf{P}_n(\mathfrak{f}),\ \mathbf{P}_m(\mathfrak{f})-\mathbf{P}_n(\mathfrak{f}))=(\hat{\mathbf{P}}_m-\hat{\mathbf{P}}_n,\ \hat{\mathbf{P}}_m-\hat{\mathbf{P}}_n)_F.$$

Taking the square root of the trace on each side,

$$\|\mathbf{P}_m(\mathfrak{f})-\mathbf{P}_n(\mathfrak{f})\|=\|\hat{\mathbf{P}}_m-\hat{\mathbf{P}}_n\|_F.$$

Now if $\mathbf{P}_n(\mathfrak{f})\to\mathbf{\Phi}$, the L.H.S. approaches 0 as $m, n\to\infty$. Hence $(\hat{\mathbf{P}}_n)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbf{L}_{2,F}$, which by 4.2 is a complete space. The sequence therefore has a limit $\mathbf{\Phi}\in\mathbf{L}_{2,F}$. Working backwards we get the converse.

(b) By 4.6 $$(\mathbf{P}_n(\mathfrak{f}),\ \mathbf{Q}_n(\mathfrak{f}))=(\hat{\mathbf{P}}_n,\ \hat{\mathbf{Q}}_n)_F,$$

and by I, 5.7 (b), $(\mathbf{P}_n(\mathfrak{f}),\ \mathbf{Q}_n(\mathfrak{f}))\to(\boldsymbol{\varphi},\ \boldsymbol{\psi})$, as $n\to\infty$. To prove (b) we need only show that

$$(\hat{\mathbf{P}}_n,\ \hat{\mathbf{Q}}_n)_F\to(\mathbf{\Phi},\ \mathbf{\Psi})_F,\ \text{as } n\to\infty. \tag{1}$$

Now by 4.5

$$|(\hat{\mathbf{P}}_n,\ \hat{\mathbf{Q}}_n)_F-(\mathbf{\Phi},\ \mathbf{\Psi})_F|_E=|(\hat{\mathbf{P}}_n-\mathbf{\Phi},\ \hat{\mathbf{Q}}_n)_F+(\mathbf{\Phi},\ \hat{\mathbf{Q}}_n-\mathbf{\Psi})_F|_E$$
$$\leqslant\|\hat{\mathbf{P}}_n-\mathbf{\Phi}\|_F\|\hat{\mathbf{Q}}_n\|_F+\|\mathbf{\Phi}\|_F\|\hat{\mathbf{Q}}_n-\mathbf{\Psi}\|_F.$$

Since $\hat{\mathbf{P}}_n\to\mathbf{\Phi}$, $\hat{\mathbf{Q}}_n\to\mathbf{\Psi}$ in $\mathbf{L}_{2,F}$, the R.H.S.$\to 0$, as $n\to\infty$. Thus (1) is established. (Q.E.D.)

4.8 DEFINITION. *Let $(\mathfrak{f}_n)_{-\infty}^{\infty}$ be a regular full-rank S.P. with spectral distribution* $\mathbf{F}$. *Then for each $\boldsymbol{\varphi}\in\mathfrak{M}_\infty$ we define the* corresponding member $\mathbf{\Phi}\in\mathbf{L}_{2,F}$ *as follows:*

(i) *if* $\boldsymbol{\varphi}=\mathbf{P}(\mathfrak{f})$, *then* $\mathbf{\Phi}=\hat{\mathbf{P}}$,

(ii) *if* $\boldsymbol{\varphi}=\lim_{n\to\infty}\mathbf{P}_n(\mathfrak{f})$ *in* $\mathfrak{M}_\infty$, *then* (1) $\mathbf{\Phi}=\lim_{n\to\infty}\hat{\mathbf{P}}_n$ *in* $\mathbf{L}_{2,F}$.

We note that to the function $\mathfrak{f}_{-k}$ in $\mathfrak{M}_\infty$ corresponds the function $e^{ki\theta}\mathbf{I}$ in $\mathbf{L}_{2,F}$. It also readily follows that if $\mathbf{A}$ is a $q\times q$ matrix, then to the functions $\mathbf{A}\cdot\mathbf{P}(\mathfrak{f})$, $\mathbf{P}(\mathfrak{f})+\mathbf{Q}(\mathfrak{f})$ in $\mathfrak{M}_\infty$, correspond the functions $\mathbf{A}\cdot\hat{\mathbf{P}}$, $\hat{\mathbf{P}}+\hat{\mathbf{Q}}$ in $\mathbf{L}_{2,F}$. Also by 4.6 $(\mathbf{P}(\mathfrak{f}),\ \mathbf{Q}(\mathfrak{f}))=(\hat{\mathbf{P}},\ \hat{\mathbf{Q}})_F$. By a limiting argument these results can be extended to all functions in $\mathfrak{M}_\infty$, so that if to $\boldsymbol{\varphi}$, $\boldsymbol{\psi}\in\mathfrak{M}_\infty$ correspond the functions $\mathbf{\Phi}$, $\mathbf{\Psi}\in\mathbf{L}_{2,F}$, then to $\mathbf{A}\cdot\boldsymbol{\varphi}$, $\boldsymbol{\varphi}+\boldsymbol{\psi}$ correspond $\mathbf{A}\cdot\mathbf{\Phi}$, $\mathbf{\Phi}+\mathbf{\Psi}$, and $(\boldsymbol{\varphi},\ \boldsymbol{\psi})=(\mathbf{\Phi},\ \mathbf{\Psi})_F$. Furthermore, if as is natural we identify functions $\mathbf{\Phi}$, $\mathbf{\Psi}\in\mathbf{L}_{2,F}$ which differ only on subsets of C of zero $\mathbf{F}$-measure,

(1) As just shown in 4.7 the limit on the right will exist if that on the left exists.

i.e. for which $\|\mathbf{\Phi}-\mathbf{\Psi}\|_F=0$, then the correspondence from $\mathfrak{M}_\infty$ into $\mathbf{L}_{2,F}$ is one-one. We shall now show that this correspondence is *onto* $\mathbf{L}_{2,F}$. We need the following lemma.

4.9 LEMMA. *Let* $(\mathbf{f}_n)_{-\infty}^{\infty}$ *and* $\mathbf{F}$ *be as in* 4.8, $(\mathbf{h}_n)_{-\infty}^{\infty}$ *be the normalised innovation process of* $(\mathbf{f}_n)_{-\infty}^{\infty}$, *and* $\mathbf{\Phi}$ *its generating function. Then*

(a) *the function* $e^{-ni\theta}\mathbf{\Phi}^{-1}\in\mathbf{L}_{2,F}$, *and corresponds to the function* $\mathbf{h}_n\in\mathfrak{M}_\infty$;

(b) *for any* $\mathbf{\Psi}\in\mathbf{L}_{2,F}$, $\mathbf{\Psi}\mathbf{\Phi}\in\mathbf{L}_2$;

(c) *for any* $\mathbf{\Psi}\in\mathbf{L}_{2,F}$, *if* $\mathbf{A}_k$ *is k-th Fourier coefficient of* $\mathbf{\Psi}\mathbf{\Phi}$, *then as* $n\to\infty$ $\left(\sum_{-n}^{n}\mathbf{A}_k e^{ki\theta}\right)\mathbf{\Phi}^{-1}\to\mathbf{\Psi}$ *in the* $\mathbf{L}_{2,F}$*-norm.*

Proof. Since $\mathbf{h}_n\in\mathfrak{M}_\infty$, there is a corresponding member in $\mathbf{L}_{2,F}$, say $\mathbf{\Psi}$. As remarked in the preceding para,

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-ki\theta}\mathbf{F}'(e^{i\theta})\mathbf{\Psi}^*(e^{i\theta})\,d\theta = (e^{-ki\theta}\mathbf{I},\mathbf{\Psi})_F = (\mathbf{f}_k,\mathbf{h}_n) = (\mathbf{f}_0,\mathbf{h}_{n-k}).$$

The last term is $\mathbf{0}$ for $k<n$, since in this case $\mathbf{h}_{n-k}\perp\mathfrak{M}_0$. Thus for each k, the function $\mathbf{F}'\mathbf{\Psi}^*$ has the same kth Fourier coefficient as the function

$$\sum_{k=n}^{\infty}(\mathbf{f}_0,\mathbf{h}_{n-k})e^{ki\theta} = e^{ni\theta}\sum_{j=0}^{\infty}(\mathbf{f}_0,\mathbf{h}_{-j})e^{ji\theta} = e^{ni\theta}\mathbf{\Phi}(e^{i\theta}).$$

Hence $\mathbf{F}'\mathbf{\Psi}^*=e^{ni\theta}\mathbf{\Phi}$, a.e. But by (2.5), $\mathbf{F}'=\mathbf{\Phi}\mathbf{\Phi}^*$ a.e., and since $\mathbf{F}'$ is invertible a.e., so is $\mathbf{\Phi}$. It readily follows that $\mathbf{\Psi}=e^{-ni\theta}\mathbf{\Phi}^{-1}$, which shows that $e^{-ni\theta}\mathbf{\Phi}^{-1}\in\mathbf{L}_{2,F}$, and corresponds to $\mathbf{h}_n$.

(b) Let $\mathbf{\Psi}\in\mathbf{L}_{2,F}$. Then by 4.1, $\mathbf{\Psi}\sqrt{\mathbf{F}'}\in\mathbf{L}_2$, and therefore $|\mathbf{\Psi}\sqrt{\mathbf{F}'}|_E^2\in L_1$. Now since $\mathbf{F}'=\mathbf{\Phi}\mathbf{\Phi}^*$, a.e., we have

$$|\mathbf{\Psi}\sqrt{\mathbf{F}'}|_E^2=\tau(\mathbf{\Psi}\mathbf{F}'\mathbf{\Psi}^*)=\tau(\mathbf{\Psi}\mathbf{\Phi}\mathbf{\Phi}^*\mathbf{\Psi}^*)=|\mathbf{\Psi}\mathbf{\Phi}|_E^2. \tag{2}$$

Thus $|\mathbf{\Psi}\mathbf{\Phi}|_E^2\in L_1$ and therefore $\mathbf{\Psi}\mathbf{\Phi}\in\mathbf{L}_2$, cf. I, 3.5 (a).

(c) Since by (a) $\mathbf{\Phi}^{-1}\in\mathbf{L}_{2,F}$ and $\sum_{-n}^{n}\mathbf{A}_k e^{ki\theta}$ is bounded, therefore by 4.1 (c)

$$\sum_{-n}^{n}\mathbf{A}_k e^{ki\theta})\,\mathbf{\Phi}^{-1}(e^{i\theta})\in\mathbf{L}_{2,F}.$$

Next, by 4.2 and (2)

$$\left\| \left(\sum_{-n}^{n} \mathbf{A}_k e^{ki\theta} \right) \mathbf{\Phi}^{-1} - \mathbf{\Psi} \right\|_F = \left\| \left\{ \sum_{-n}^{n} \mathbf{A}_k e^{ki\theta} - \mathbf{\Psi}\mathbf{\Phi} \right\} \mathbf{\Phi}^{-1} \sqrt{\mathbf{F}'} \right\|$$

$$= \left\| \left\{ \sum_{-n}^{n} \mathbf{A}_k e^{ki\theta} - \mathbf{\Psi}\mathbf{\Phi} \right\} \mathbf{\Phi}^{-1} \mathbf{\Phi} \right\|$$

$$= \left\| \sum_{-n}^{n} \mathbf{A}_k e^{ki\theta} - \mathbf{\Psi}\mathbf{\Phi} \right\|$$

Now since $\mathbf{A}_k$ is the kth Fourier coefficients of $\mathbf{\Psi}\mathbf{\Phi}$, the R.H.S.$\to 0$, as $n \to \infty$. Hence $\left(\sum_{-n}^{n} \mathbf{A}_k e^{ki\theta} \right) \mathbf{\Phi}^{-1} \to \mathbf{\Psi}$ in the norm $\| \ \|_F$. (Q.E.D.)

We are now ready to show that our correspondence is *onto* $\mathbf{L}_{2,F}$. Let $\mathbf{\Psi} \in \mathbf{L}_{2,F}$ and let $\mathbf{A}_k$ be the kth Fourier coefficient of the function $\mathbf{\Psi}\mathbf{\Phi} \in \mathbf{L}_2$, where $\mathbf{\Phi}$ is as in 4.9. Then since $\sum_{-\infty}^{\infty} |\mathbf{A}_k|_E^2 < \infty$ and the process $(\mathbf{h}_k)_{-\infty}^{\infty}$ is orthonormal,

$$\sum_{-n}^{n} \mathbf{A}_k \mathbf{h}_{-k} \to \textit{ some } \mathbf{g} \in \mathfrak{M}_\infty, \textit{ as } n \to \infty. \tag{3}$$

By 4.9 (a) and the fact that the correspondence preserves addition, and multiplication by matrices, it follows that to the function on the left of (3) corresponds the function

$$\sum_{-n}^{n} \mathbf{A}_k e^{ki\theta} \mathbf{\Phi}^{-1} = \left(\sum_{-n}^{n} \mathbf{A}_k e^{ki\theta} \right) \mathbf{\Phi}^{-1}$$

in $\mathbf{L}_{2,F}$. By our Definition 4.8 (ii) its limit in $\mathbf{L}_{2,F}$, as $n \to \infty$, corresponds to $\mathbf{g}$. But by 4.9 (c) this limit is $\mathbf{\Psi}$. Thus $\mathbf{\Psi}$ corresponds to $\mathbf{g} \in \mathfrak{M}_\infty$. To sum up, we have the following theorem.

4.10 Theorem. *If $(\mathbf{f}_n)_{-\infty}^{\infty}$ is a regular full-rank process with spectral distribution $\mathbf{F}$, then the correspondence defined in 4.8 is an isomorphism on $\mathfrak{M}_\infty$ onto $\mathbf{L}_{2,F}$, on the understanding that we identify members of $\mathbf{L}_{2,F}$, which differ on sets of zero $\mathbf{F}$-measure. More fully, if to $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathfrak{M}_\infty$ correspond $\mathbf{\Phi}, \mathbf{\Psi} \in \mathbf{L}_{2,F}$, then to $\boldsymbol{\varphi} + \boldsymbol{\psi}$, $\mathbf{A}\boldsymbol{\varphi}$ correspond $\mathbf{\Phi} + \mathbf{\Psi}$, $\mathbf{A}\mathbf{\Phi}$, and*

$$(\boldsymbol{\varphi}, \boldsymbol{\psi}) = (\mathbf{\Phi}, \mathbf{\Psi})_F = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{\Phi}(e^{i\theta}) \mathbf{F}'(e^{i\theta}) \mathbf{\Psi}^*(e^{i\theta}) \, d\theta$$

$$\|\boldsymbol{\varphi}\|^2 = \|\mathbf{\Phi}\|_F^2 = \frac{1}{2\pi} \int_0^{2\pi} |\mathbf{\Phi}(e^{i\theta}) \sqrt{\mathbf{F}'(e^{i\theta})}|_E^2 \, d\theta.$$

An important application of this theorem is the following

4.11 COROLLARY. *Let* $(\mathbf{f}_n)_{-\infty}^{\infty}$ *be a regular full-rank process with spectral distribution* $\mathbf{F}$ *and generating function* $\mathbf{\Phi}$, *and let* (1)

$$\mathbf{Y}_\nu(e^{i\theta}) = [e^{-\nu i\theta}\,\mathbf{\Phi}(e^{i\theta})]_{0+}\,\mathbf{\Phi}^{-1}(e^{i\theta}), \quad \nu > 0.$$

Then $\mathbf{Y}_\nu \in \mathbf{L}_{2,F}$ *and corresponds to the linear predictor* $\hat{\mathbf{f}}_\nu = (\mathbf{f}_\nu \mid \mathfrak{M}_0)$ *under the isomorphism defined in* 4.8.

Proof. Since, cf (2.5), $\mathbf{F}' = \mathbf{\Phi}\mathbf{\Phi}^*$ a.e., therefore

$$\mathbf{Y}_\nu(e^{i\theta})\,\mathbf{F}'(e^{i\theta})\,\mathbf{Y}_\nu^*(e^{i\theta}) = [e^{-\nu i\theta}\,\mathbf{\Phi}(e^{i\theta})]_{0+}\,[e^{-\nu i\theta}\,\mathbf{\Phi}(e^{i\theta})]_{0+}^*.$$

This is in L_1, since $[e^{-\nu i\theta}\,\mathbf{\Phi}(e^{i\theta})]_{0+} \in \mathbf{L}_2$. Thus $\mathbf{Y}_\nu \in \mathbf{L}_{2,F}$.

Now let $\mathbf{C}_k$ be the kth Fourier coefficient of $\mathbf{\Phi}$. Then

$$\mathbf{Y}_\nu(e^{i\theta})\,\mathbf{\Phi}(e^{i\theta}) = [e^{-\nu i\theta}\,\mathbf{\Phi}(e^{i\theta})]_{0+} = \sum_{k=0}^{\infty} \mathbf{C}_{\nu+k}\, e^{ki\theta} \in \mathbf{L}_2.$$

Hence by 4.9 (c), as $N \to \infty$

$$\left(\sum_{k=0}^{N} \mathbf{C}_{\nu+k}\, e^{ki\theta}\right)\mathbf{\Phi}^{-1}(e^{i\theta}) \to \mathbf{Y}_\nu(e^{i\theta}) \text{ in the } \mathbf{L}_{2,F}\text{-norm.}$$

It follows that if $\boldsymbol{\varphi}$ is the random function in $\mathfrak{M}_\infty$ corresponding to $\mathbf{Y}_\nu$ in $\mathbf{L}_{2,F}$, then (cf. 4.9 (a)),

$$\sum_{k=0}^{N} \mathbf{C}_{\nu+k}\,\mathbf{h}_{-k} \to \boldsymbol{\varphi}, \text{ as } N \to \infty$$

But, cf. (2.4), the last sum tends to $\hat{\mathbf{f}}_\nu$ as $N \to \infty$. Hence $\boldsymbol{\varphi} = \hat{\mathbf{f}}_\nu$. (Q.E.D.)

5. The Boundedness Condition

To progress further we have to assume that the eigenvalues of our spectral density matrix are essentially bounded above and away from zero. By 1.5 (a) this assumption may be stated as follows.

5.1. Boundedness Condition. *Our* q-ple *regular, full-rank* S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ *has a spectral density* $\mathbf{F}'$ *such that*

$$\lambda\,\mathbf{I} \prec \mathbf{F}'(e^{i\theta}) \prec \lambda'\,\mathbf{I}, \quad 0 < \lambda \leqslant \lambda' < \infty.$$

(1) Since $\Delta\,\mathbf{\Phi}$ vanishes almost nowhere on C, $\mathbf{\Phi}^{-1}(e^{i\theta})$ is defined a.e. on C.

We shall show in this section that this condition entails the following consequences:

(i) $\mathbf{L}_2$ and $\mathbf{L}_{2,F}$ become identical topological spaces.

(ii) The sum of (one-dimensional) manifolds $\sum_{k=0}^{\infty} \mathfrak{S}(\mathbf{f}_{-k})$ becomes topologically closed (cf. I, 5.6 (d)), and therefore identical to $\mathfrak{M}_0$ the present and past of $\mathbf{f}_0$.

(iii) The innovation function $\mathbf{h}_0$ is expressible as a mean-convergent infinite series $\sum_{k=0}^{\infty} \mathbf{D}_k \mathbf{f}_{-k}$.

(iv) If $\mathbf{\Phi}$ is the generating function of the S.P., then $\mathbf{\Phi}^{-1} \in \mathbf{L}_{\infty}^{0-}$.

(v) The linear predictor for any lag ν is expressible as the sum of a mean-convergent infinite series $\sum_{k=0}^{\infty} \mathbf{E}_{\nu k} \mathbf{f}_{-k}$, where $\mathbf{E}_{\nu k}$ is a finite sum of products of the Fourier coefficients of $\mathbf{\Phi}$ and $\mathbf{\Phi}^{-1}$.

5.2 Lemma. *If* $\mathbf{F}'$ *satifies the condition* 5.1, *then*

(a) $\mathbf{L}_{2,F} = \mathbf{L}_2$;

(b) *for all* $\mathbf{\Phi} \in \mathbf{L}_2$,

$$\lambda \int_0^{2\pi} \mathbf{\Phi}(e^{i\theta})\, \mathbf{\Phi}^*(e^{i\theta})\, d\theta \prec 2\pi (\mathbf{\Phi}, \mathbf{\Phi})_F \prec \lambda' \int_0^{2\pi} \mathbf{\Phi}(e^{i\theta})\, \mathbf{\Phi}^*(e^{i\theta})\, d\theta \tag{1}$$

$$\lambda \| \mathbf{\Phi} \|^2 \leqslant \| \mathbf{\Phi} \|_F^2 \leqslant \lambda' \| \mathbf{\Phi} \|^2; \tag{2}$$

(c) $\mathbf{L}_{2,F}$*-convergence and* $\mathbf{L}_2$*-convergence are equivalent.*

Proof. (a) From 5.1

$$\sqrt{\lambda}\,\mathbf{I} \prec \sqrt{\mathbf{F}'} \prec \sqrt{\lambda'}\,\mathbf{I}, \quad \frac{1}{\sqrt{\lambda'}}\,\mathbf{I} \prec (\sqrt{\mathbf{F}'})^{-1} \prec \frac{1}{\sqrt{\lambda}}\,\mathbf{I}, \quad \text{a.e.}$$

Hence $\sqrt{\mathbf{F}'}$, $(\sqrt{\mathbf{F}'})^{-1} \in \mathbf{L}_\infty$. Hence if $\mathbf{\Phi} \in \mathbf{L}_2$, then $\mathbf{\Phi}\sqrt{\mathbf{F}'} \in \mathbf{L}_2$ and therefore by 4.1 (a) $\mathbf{\Phi} \in \mathbf{L}_{2,F}$. Next, if $\mathbf{\Phi} \in \mathbf{L}_{2,F}$, then by 4.1 (a) $\mathbf{\Phi}\sqrt{\mathbf{F}'} \in \mathbf{L}_2$, and therefore since $(\sqrt{\mathbf{F}'})^{-1} \in \mathbf{L}_\infty$, $\mathbf{\Phi} = \mathbf{\Phi}\sqrt{\mathbf{F}'}\,(\sqrt{\mathbf{F}'})^{-1} \in \mathbf{L}_2$. Thus $\mathbf{L}_2 = \mathbf{L}_{2,F}$.

(b) By 5.1 and 1.5 (d)

$$\lambda\, \mathbf{\Phi}\mathbf{\Phi}^* \prec \mathbf{\Phi}\mathbf{F}'\mathbf{\Phi}^* \prec \lambda'\, \mathbf{\Phi}\mathbf{\Phi}^*, \quad \text{a.e.}$$

Hence, cf. 1.9 (b), their integrals must bear the same relations, i.e. we have (1). Dividing by 2π and taking traces, we get (2).

(c) Let $\mathbf{\Phi}_n$, $\mathbf{\Phi} \in \mathbf{L}_{2,F} = \mathbf{L}_2$. The inequalities

$$\sqrt{\lambda}\cdot\|\mathbf{\Phi}_n - \mathbf{\Phi}\| \leqslant \|\mathbf{\Phi}_n - \mathbf{\Phi}\|_F \leqslant \sqrt{\lambda'}\cdot\|\mathbf{\Phi}_n - \mathbf{\Phi}\|$$

show that $\mathbf{\Phi}_n \to \mathbf{\Phi}$ in $\mathbf{L}_{2,F}$ if and only if $\mathbf{\Phi}_n \to \mathbf{\Phi}$ in $\mathbf{L}_2$. (Q.E.D.)

5.3 THEOREM. *Let the spectral density* $\mathbf{F}'$ *of a S.P.* $(\mathbf{f}_n)_{-\infty}^{\infty}$ *satisfy the condition* 5.1. *Then*

(a) *for all matrices* $\mathbf{A}_0, \ldots, \mathbf{A}_n$,

$$\lambda \sum_0^n \mathbf{A}_k \mathbf{A}_k^* \prec \left(\sum_0^n \mathbf{A}_k \mathbf{f}_{-k}, \sum_0^n \mathbf{A}_k \mathbf{f}_{-k}\right) \prec \lambda' \sum_0^n \mathbf{A}_k \mathbf{A}_k^*;$$

(b) *if* $\mathfrak{M}_0$ *is the present and past of* $\mathbf{f}_0$, *then* $\mathfrak{M}_0 = \sum_0^\infty \mathfrak{S}(\mathbf{f}_{-k})$;

(c) *if* $\mathbf{g} = \sum_0^\infty \mathbf{B}_k \mathbf{f}_{-k} \in \mathfrak{M}_0$, *then*

$$\lambda \sum_0^\infty |\mathbf{B}_k|_E^2 \leqslant \|\mathbf{g}\|^2 \leqslant \lambda' \sum_0^\infty |\mathbf{B}_k|_E^2.$$

Proof. (a) Let $\mathbf{\Phi}(e^{i\theta}) = \sum_0^n \mathbf{A}_k e^{ki\theta}$. Then by 5.2 (b)

$$\lambda \int_0^{2\pi} \mathbf{\Phi}(e^{i\theta})\mathbf{\Phi}^*(e^{i\theta})\,d\theta \prec 2\pi(\mathbf{\Phi}, \mathbf{\Phi})_F \prec \lambda' \int_0^{2\pi} \mathbf{\Phi}(e^{i\theta})\mathbf{\Phi}^*(e^{i\theta})\,d\theta.$$

Now by the Parseval relation (1.13), the integral in the border terms equals $2\pi \sum_0^n \mathbf{A}_k \mathbf{A}_k^*$. Also by 4.10

$$(\mathbf{\Phi}, \mathbf{\Phi})_F = \left(\sum_0^n \mathbf{A}_k \mathbf{f}_{-k}, \sum_0^n \mathbf{A}_k \mathbf{f}_{-k}\right).$$

(b) Taking the trace of each term in the inequalities (a) we get

$$\lambda \sum_0^n |\mathbf{A}_k|_E^2 \leqslant \left\|\sum_0^n \mathbf{A}_k \mathbf{f}_{-k}\right\|^2 \leqslant \lambda' \sum_0^n |\mathbf{A}_k|_E^2. \tag{3}$$

Now obviously $\sum_0^\infty \mathfrak{S}(\mathbf{f}_{-k}) \subseteq \mathfrak{M}_0$. Hence we have only to prove the reverse inclusion, i.e. show that given any $\mathbf{g} \in \mathfrak{M}_0$, there exist matrices $\mathbf{B}_k$ such that

$$\mathbf{g} = \sum_0^\infty \mathbf{B}_k \mathbf{f}_{-k}, \tag{I}$$

the last series being convergent in the $\mathfrak{L}_2$-norm $\|\ \|$ of I, (5.3). Let $\mathbf{g} \in \mathfrak{M}_0$. Then

$$\mathbf{g} = \lim_{n\to\infty} \mathbf{g}_k, \quad \textit{where } \mathbf{g}_k = \sum_{k=0}^n \mathbf{A}_k^{(n)} \mathbf{f}_{-n}.$$

For convenience we define $\mathbf{A}_k^{(n)}=0$ for $k>n$. Then by (3) for all $n\geqslant m$,

$$\|\mathbf{g}_m-\mathbf{g}_n\|^2=\|\sum_{k=0}^{n}(\mathbf{A}_k^{(m)}-\mathbf{A}_k^{(n)})\,\mathbf{f}_{-k}\|^2\geqslant\lambda\sum_{k=0}^{n}|\mathbf{A}_k^{(m)}-\mathbf{A}_k^{(n)}|_E^2.$$

It follows that for $0\leqslant j\leqslant m\leqslant n$,

$$\|\mathbf{g}_m-\mathbf{g}_n\|^2\geqslant\lambda\sum_{k=0}^{m}|\mathbf{A}_k^{(m)}-\mathbf{A}_k^{(n)}|_E^2\geqslant\lambda\,|\mathbf{A}_j^{(m)}-\mathbf{A}_j^{(n)}|_E^2. \tag{4}$$

Now the left member of (4) tends to 0, as $m, n\to\infty$. The same must therefore be true of the right member. Since the space of matrices is complete under the Euclidean norm, we infer that

$$\mathbf{A}_j^{(n)}\to\mathbf{B}_j,\quad\text{as } n\to\infty,\qquad 0\leqslant j<\infty. \tag{5}$$

Next, let $n\to\infty$ in (4). Then since the series in the middle has only a finite number of terms, it follows from (5) and (3) that

$$\begin{aligned}\|\mathbf{g}_m-\mathbf{g}\|^2&\geqslant\lambda\sum_{k=0}^{m}|\mathbf{A}_k^{(m)}-\mathbf{B}_k|_E^2\\&\geqslant\frac{\lambda}{\lambda'}\|\sum_{k=0}^{m}(\mathbf{A}_k^{(m)}-\mathbf{B}_n)\,\mathbf{f}_{-k}\|^2\\&\geqslant\frac{\lambda}{\lambda'}\|\mathbf{g}_m-\sum_{k=0}^{m}\mathbf{B}_k\,\mathbf{f}_{-k}\|^2\end{aligned}$$

Since $\mathbf{g}_m\to\mathbf{g}$ as $m\to\infty$, we conclude that $\sum_0^m\mathbf{B}_k\,\mathbf{f}_{-k}\to\mathbf{g}$ as $m\to\infty$. Thus I.

(c) From (3) we have

$$\lambda\sum_0^n|\mathbf{B}_k|_E^2\leqslant\|\sum_0^n\mathbf{B}_k\,\mathbf{f}_{-k}\|^2\leqslant\lambda'\sum_0^n|\mathbf{B}_k|_E^2.$$

Since the sum in the middle approaches $\mathbf{g}$ as $n\to\infty$, it follows that $\sum_0^\infty|\mathbf{B}_k|_E^2<\infty$, and the inequalities given in (c) hold. (Q.E.D.)

Now let $(\mathbf{f}_n)_{-\infty}^{\infty}$ be as in 5.1 and let $(\mathbf{h}_n)_{-\infty}^{\infty}$ be its normalised innovation process. Since $\mathbf{h}_0\in\mathfrak{M}_0$, it follows by 5.3 (b) (c) that

$$\mathbf{h}_0=\sum_0^\infty\mathbf{D}_k\,\mathbf{f}_{-k},\quad\sum_0^\infty|\mathbf{D}_k|_E^2<\infty.$$

Since [I, 6.12] $\mathbf{h}_n=U^n\,\mathbf{h}_0$, where U is the shift operator of the process $(\mathbf{f}_n)_{-\infty}^{\infty}$, we get the following result.

5.4 COROLLARY. *If* $(\mathbf{h}_n)_{-\infty}^{\infty}$ *is the normalised innovation process of a S.P.* $(\mathbf{f}_n)_{-\infty}^{\infty}$ *having a spectral density which satisfies the condition* 5.1, *then there exist matrices* $\mathbf{D}_k$ *such that*

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

Now let $\mathbf{\Phi}$ be the generating function of such a process. We know that $\mathbf{\Phi} \in \mathbf{L}_2^{0+}$, cf. (2.5). Under the boundedness condition, the equality $\mathbf{\Phi}\mathbf{\Phi}^* = \mathbf{F}'$ shows that $\mathbf{\Phi} \in \mathbf{L}_\infty^{0+}$. Next, the equality $(\mathbf{\Phi}^{-1})^* \mathbf{\Phi}^{-1} = (\mathbf{F}')^{-1}$ shows that $\mathbf{\Phi}^{-1} \in \mathbf{L}_\infty$. We shall now show that $\mathbf{\Phi}^{-1} \in \mathbf{L}_\infty^{0+}$.

Since the series for $\mathbf{h}_0$ given in 5.4 converges in the $\mathfrak{L}_2$-sense [I, (5.3)], it follows by Theorem 4.10 that $\sum_0^n \mathbf{D}_k e^{ki\theta}$ tends to the function in $\mathbf{L}_{2,F}$ corresponding to $\mathbf{h}_0 \in \mathfrak{M}_\infty$. By 4.9 (a) this function is $\mathbf{\Phi}^{-1}$. Since by 5.2 the $\mathbf{L}_{2,F}$- and $\mathbf{L}_2$-topologies are equivalent we see that

$$\sum_0^n \mathbf{D}_k e^{ki\theta} \to \mathbf{\Phi}^{-1} \text{ in the } \mathbf{L}_2\text{-norm } \|\ \|.$$

Thus $\sum_0^\infty \mathbf{D}_k e^{ki\theta}$ is the Fourier series of $\mathbf{\Phi}^{-1}$, i.e. $\mathbf{\Phi}^{-1} \in \mathbf{L}_2^{0+}$. Since, as already remarked, $\mathbf{\Phi}^{-1} \in \mathbf{L}_\infty$, we conclude that $\mathbf{\Phi}^{-1} \in \mathbf{L}_\infty^{0+}$.

We may sum up these results as follows.

5.5 THEOREM. *If* (i) $(\mathbf{f}_n)_{-\infty}^{\infty}$ *is a q-ple S.P. with a spectral density satisfying the Boundedness Condition* 5.1,

(ii) $(\mathbf{h}_n)_{-\infty}^{\infty}$ *is its normalised innovation process,*

(iii) $\mathbf{\Phi}$ *is its generating function,*

then

(a) $\mathbf{\Phi}, \mathbf{\Phi}^{-1} \in \mathbf{L}_\infty^{0+}$,

(b) $\mathbf{h}_n = \sum_0^\infty \mathbf{D}_k \mathbf{f}_{n-k}$, *where* $\mathbf{D}_k$ *is the k-th Fourier coefficient of* $\mathbf{\Phi}^{-1}$.

We now turn to the linear predictor. From (2.4) and 5.5 (b),

5.6
$$\hat{\mathbf{f}}_\nu = \sum_{n=0}^{\infty} \mathbf{C}_{\nu+n} \mathbf{h}_{-n}, \qquad \mathbf{h}_n = \sum_{j=0}^{\infty} \mathbf{D}_j \mathbf{f}_{n-j}, \quad \nu > 0. \tag{5.6}$$

Substituting from the second equation of (5.6) into the first, and heuristically interchanging the order of summation, we get

$$\hat{\mathbf{f}}_\nu = \sum_{n=0}^{\infty} \mathbf{C}_{\nu+n} \sum_{k=n}^{\infty} \mathbf{D}_{-n+k} \mathbf{f}_{-k} = \sum_{k=0}^{\infty} \left(\sum_{n=0}^{k} \mathbf{C}_{\nu+n} \mathbf{D}_{k-n} \right) \mathbf{f}_{-k}.$$

Rather than justify this change in the order of summation, we shall show directly that the series on the right converges in $\mathfrak{L}_2$ to the function $\hat{\mathbf{f}}_\nu$:

5.7 THEOREM. *If the process* $(\mathbf{f}_n)_{-\infty}^{\infty}$ *has a spectral density* $\mathbf{F}'$ *satisfying the Boundedness Condition* 5.1 *and the generating function* $\mathbf{\Phi}$, *then for* $\nu > 0$

$$\hat{\mathbf{f}}_\nu = \sum_{k=0}^{\infty} \mathbf{E}_{\nu k}\,\mathbf{f}_{-k}, \quad \textit{where} \quad \mathbf{E}_{\nu k} = \sum_{n=0}^{k} \mathbf{C}_{\nu+n}\,\mathbf{D}_{k-n},$$

$\mathbf{C}_j$, $\mathbf{D}_j$ *being the* j-*th Fourier coefficients of* $\mathbf{\Phi}$, $\mathbf{\Phi}^{-1}$. *The prediction-error matrix for lag* ν *is given, cf.* (2.4), *by*

$$\mathbf{G}_\nu = \sum_{n=0}^{\nu-1} \mathbf{C}_n\,\mathbf{C}_n^*.$$

Proof. We have

$$[e^{-\nu i\theta}\mathbf{\Phi}(e^{i\theta})]_{0+} = \sum_{k=0}^{\infty} \mathbf{C}_{\nu+k}\,e^{ki\theta} = e^{-\nu i\theta}\{\mathbf{\Phi}(e^{i\theta}) - \sum_{k=0}^{\nu-1} \mathbf{C}_k\,e^{ki\theta}\}.$$

This is in $\mathbf{L}_\infty^{0+}$, since cf. 5.5 (a) each term inside { } is in $\mathbf{L}_\infty$. Also, $\mathbf{\Phi}^{-1} \in \mathbf{L}_\infty^{0+}$; hence

$$\mathbf{Y}_\nu(e^{i\theta}) = [e^{-\nu i\theta}\mathbf{\Phi}(e^{i\theta})]_{0+}\,\mathbf{\Phi}^{-1}(e^{i\theta}) \in \mathbf{L}_\infty^{0+}.$$

By the Convolution Rule I, 3.9 (d), the kth Fourier coefficient of $\mathbf{Y}_\nu$ is easily seen to be $\mathbf{E}_{\nu k}$ for $k \geqslant 0$ (and of course, zero for $k < 0$). Thus as $N \to \infty$

$$\sum_{k=0}^{N} \mathbf{E}_{\nu k}\,e^{ki\theta} \to \mathbf{Y}_\nu(e^{i\theta}) \quad \textit{in the } \mathbf{L}_2\textit{-norm}.$$

But in view of the Boundedness Condition, $\mathbf{L}_2$ and $\mathbf{L}_{2,F}$ are the same topological space. Hence as $N \to \infty$,

$$\sum_{k=0}^{N} \mathbf{E}_{\nu k}\,e^{ki\theta} \to \mathbf{Y}_\nu(e^{i\theta}) \quad \textit{in the } \mathbf{L}_{2,F}\textit{-norm}.$$

Since by Corollary 4.11, $\hat{\mathbf{f}}_\nu$ in $\mathfrak{M}_\infty$ corresponds to $\mathbf{Y}_\nu$ in $\mathbf{L}_{2,F}$, it follows (cf. 4.8) that

$$\sum_{k=0}^{N} \mathbf{E}_{\nu k}\,\mathbf{f}_{-k} \to \hat{\mathbf{f}}_\nu, \quad \text{as } N \to \infty.$$

(Q.E.D.)

We thus get the prediction $\hat{\mathbf{f}}_\nu$ as the sum of a series converging in-the-mean. The sequence $(\mathbf{E}_{\nu k})_{k=0}^{\infty}$, where $\mathbf{E}_{\nu k}$ is as in 5.7, is the matricial *weighting function* in the time-domain in the discrete parameter case. It involves the Fourier coefficients of $\mathbf{\Phi}$, $\mathbf{\Phi}^{-1}$ alone.

If our aim is to perform the computations digitally, the expressions for $\hat{\mathbf{f}}_\nu$ and $\mathbf{G}_\nu$ given in 5.7 could themselves be used. If, however, the prediction is to be done by an analogue computer, then we must shift from the time-domain to the frequency-domain. Now as shown in 4.11 and in the proof of 5.7, the function

5.8 $$\mathbf{Y}_\nu(e^{i\theta}) = [e^{-\nu i\theta}\mathbf{\Phi}(e^{i\theta})]_{0+}\cdot\mathbf{\Phi}^{-1}(e^{i\theta}) \tag{5.8}$$

is in $\mathbf{L}_2\,(=\mathbf{L}_{2,F})$ and corresponds to the prediction $\hat{\mathfrak{f}}_\nu \in \mathfrak{M}_\infty$. Hence $\mathbf{Y}_\nu$ is the matricial *frequency-response* or *transfer-function* of the electric filter for which the response will be the prediction $\hat{\mathfrak{f}}_\nu(\omega)$ when the past part $(\mathfrak{f}_n(\omega))_{-\infty}^{0}$ of the multiple time-series $(\mathfrak{f}_n(\omega))_{-\infty}^{\infty}$ is fed in as input.

We should mention here a lacuna in prediction theory, which is present even for the case $q=1$, viz. the absence of a spectral characterization of processes for which the expression for $\hat{\mathfrak{f}}_\nu$ given in 5.7 is valid in the usual $\mathfrak{L}_2$-sense. As 5.7 shows, the Boundedness Condition is certainly sufficient, but it seems the weaker conditions $\mathbf{F}'$, $\mathbf{F}'^{-1}\in\mathbf{L}_1$ should suffice. For $q=1$ Kolmogorov [5, Theorem 24] has shown that these conditions characterise processes, which are non-deterministic with regard to *both past and future*, i.e. for which $\mathfrak{f}_n \notin \mathfrak{S}(\mathfrak{f}_k)_{k\neq n}$. Another unsettled question pertains to the case when the series expansion in 5.7 is not valid when convergence is taken in the usual $\mathfrak{L}_2$-sense. Does this expansion become valid when convergence is interpreted in some more subtile summability sense? (Cf. Doob [4, p. 564, Sec. 2].). We shall not discuss these questions in this paper.

6. Determination of the generating function and the linear predictor

In this section we shall express the generating function of a S.P. $(\mathfrak{f}_n)_{-\infty}^{\infty}$ satisfying the Boundedness Condition 5.1 in terms of the spectral density $\mathbf{F}'$ by following an iterative procedure of the type discussed heuristically in Sec. 3. We shall then derive computable expressions for the linear predictor and the prediction error matrix.

By 1.5 (c) the Boundedness Condition 5.1 implies that

$$\left|\frac{2}{\lambda'+\lambda}\mathbf{F}'(e^{i\theta})-\mathbf{I}\right|_B \leqslant \frac{\lambda'-\lambda}{\lambda'+\lambda}<1.$$

Now let $\tilde{\mathbf{F}}' = 2\mathbf{F}'/(\lambda'+\lambda)$ and $\tilde{\mathfrak{f}}_n = \sqrt{2/(\lambda'+\lambda)}\,\mathfrak{f}_n$. Then $\tilde{\mathbf{F}}'$ will be the spectral density of the process $(\tilde{\mathfrak{f}}_n)_{-\infty}^{\infty}$. The generating functions of $(\tilde{\mathfrak{f}}_n)_{-\infty}^{\infty}$ and $(\mathfrak{f}_n)_{-\infty}^{\infty}$ will be connected by the relation $\tilde{\mathbf{\Phi}} = \sqrt{2/(\lambda'+\lambda)}\,\mathbf{\Phi}$. In determining the generating function of a process satisfying 5.1 there is therefore no loss of generality in making the following assumption.

6.1 ASSUMPTION. *$(\mathfrak{f}_n)_{-\infty}^{\infty}$ is a regular, full-rank process with spectral density $\mathbf{F}' = \mathbf{I}+\mathbf{M}$, where*

$$\mu = \operatorname*{ess.\,l.u.b.}_{0\leqslant\theta\leqslant 2\pi} |\mathbf{M}(e^{i\theta})|_B < 1.$$

The iterative scheme (3.10) now suggests the study of the operator $\mathfrak{P}$ defined as follows.

6.2 DEFINITION. *For all* $\mathbf{\Phi} \in \mathbf{L}_2$, $\mathfrak{P}(\mathbf{\Phi}) = (\mathbf{\Phi}\mathbf{M})_+$.

Since for $\mathbf{\Phi} \in \mathbf{L}_2$ and $\mathbf{M} \in \mathbf{L}_\infty$ we have $\mathbf{\Phi}\mathbf{M} \in \mathbf{L}_2$, this definition makes sense. Some easily established properties of $\mathfrak{P}$ are stated below.

6.3 LEMMA. (a) $\mathfrak{P}$ *is in the Banach-algebra* $\mathfrak{B}$ *of* 1.16; *more fully, it is a bounded linear operator on* $\mathbf{L}_2$ *into* $\mathbf{L}_2^+$ *and* $|\mathfrak{P}| \leqslant \mu < 1$, ($\mu$ *as in* 6.1).

(b) *If* $\mathfrak{J}$ *is the unit of* $\mathfrak{B}$, *then* $\mathfrak{J} + \mathfrak{P}$ *is invertible and*

$$(\mathfrak{J} + \mathfrak{P})^{-1} = \mathfrak{J} - \mathfrak{P} + \mathfrak{P}^2 - \mathfrak{P}^3 + \cdots,$$

where the last series converges absolutely in $\mathfrak{B}$, *being in fact dominated by the convergent geometric series* $\sum \mu^k$.

(c) $$\mathfrak{P}(\mathbf{I}) = \mathbf{M}_+, \quad \mathfrak{P}^2(\mathbf{I}) = (\mathbf{M}_+\mathbf{M})_+, \quad \mathfrak{P}^3(\mathbf{I}) = \{(\mathbf{M}_+\mathbf{M})_+\mathbf{M}\}_+,$$

and so on.

(d) $$\|\mathfrak{P}^k(\mathbf{I})\| \leqslant \mu^k \sqrt{q}.$$

In view of 6.3 (b) the following definition makes sense.

6.4. DEFINITION. *Let* $\mathbf{\Psi} = (\mathfrak{J} + \mathfrak{P})^{-1}(\mathbf{I}) = \sum_{k=0}^{\infty} (-1)^k \mathfrak{P}^k(\mathbf{I})$.

The last series is absolutely convergent in the $\mathbf{L}_2$-norm, since by 6.3 (d)

$$\sum_{k=0}^{\infty} \|\mathfrak{P}^k(\mathbf{I})\| \leqslant \sqrt{q} \sum_{k=0}^{\infty} \mu^k < \infty.$$

By 6.3 (c) and the fact that $\mathbf{L}_2^{0+}$ is closed we have

6.5 $$\mathbf{\Psi} = \mathbf{I} - \mathbf{M}_+ + (\mathbf{M}_+\mathbf{M})_+ - \{(\mathbf{M}_+\mathbf{M})_+\mathbf{M}\}_+ + \cdots \in \mathbf{L}_2^{0+}. \tag{6.5}$$

The function $\mathbf{\Psi}$ is thus derivable from the spectral density by an iterative method. We shall now show that the generating function $\mathbf{\Phi}$ of our S.P. and its innovation matrix $\mathbf{G}$ are easily obtainable from $\mathbf{\Psi}$.

6.6 THEOREM. (a) $\mathbf{\Psi} = \sqrt{\mathbf{G}}\,\mathbf{\Phi}^{-1}$, (b) $\mathbf{\Psi}\mathbf{F}'\mathbf{\Psi}^* = \mathbf{G}$.

Proof. Since the S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ satisfies the Assumption 6.1, it certainly satisfies the Boundedness Condition 5.1. Hence by Theorem 5.5 (a) $\mathbf{\Phi}^{-1} \in \mathbf{L}_\infty^{0+} \subseteq \mathbf{L}_2$. The function $\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}$ thus lies in the domain of $\mathfrak{P}$. We shall show that

$$(\mathfrak{J}+\mathfrak{P})\,(\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1})=\mathbf{I}. \tag{1}$$

Since by 6.3 (b) $\mathfrak{J}+\mathfrak{P}$ is one-one, it will follow from (1) and 6.4 that $\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}=\boldsymbol{\Psi}$, and therefore (cf. (2.5)) that

$$\boldsymbol{\Psi}\,\mathbf{F}'\,\boldsymbol{\Psi}^* = \sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}\cdot\boldsymbol{\Phi}\,\boldsymbol{\Phi}^*\cdot(\boldsymbol{\Phi}^{-1})^*\sqrt{\mathbf{G}}=\mathbf{G}.$$

To prove (1) we note that by (2.5) $\boldsymbol{\Phi}(0)=\sqrt{\mathbf{G}}$. Since $\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}\in\mathbf{L}_\infty^{0+}$, therefore $\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}(0)=\mathbf{I}$ and hence

$$\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}=\mathbf{I}+(\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1})_+. \tag{2}$$

Now since $\mathbf{F}'=\boldsymbol{\Phi}\,\boldsymbol{\Phi}^*$, therefore

$$\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}+(\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1})\,\mathbf{M}=\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}\,\mathbf{F}'=\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^*\in\mathbf{L}_2^{0-}.$$

Hence $$(\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1})_+ + \{(\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1})\,\mathbf{M}\}_+=\mathbf{0},$$

i.e. by (2) $$(\mathfrak{J}+\mathfrak{P})\,(\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1})=\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1}+\{(\sqrt{\mathbf{G}}\,\boldsymbol{\Phi}^{-1})\,\mathbf{M}\}_+=\mathbf{I}.$$

This establishes (1). (Q.E.D.)

The prediction error matrix $\mathbf{G}$, and thence the generating function $\boldsymbol{\Phi}=\boldsymbol{\Psi}^{-1}\sqrt{\mathbf{G}}$, are therefore expressible in terms of the spectral density $\mathbf{F}'$. Somewhat different expressions for $\mathbf{G}$ and $\boldsymbol{\Phi}$ can be obtained as follows. Since $\mathbf{M}^*=\mathbf{M}$, we have from (6.5) and 1.15 (b) (4),

6.7 $$\boldsymbol{\Psi}^*=\mathbf{I}-\mathbf{M}_-+(\mathbf{M}\,\mathbf{M}_-)_- - \{\mathbf{M}\,(\mathbf{M}\,\mathbf{M}_-)_-\}_- + \cdots \in L_2^{0-}. \tag{6.7}$$

Now let

6.8 $$\boldsymbol{\chi}=\mathbf{F}'\,\boldsymbol{\Psi}^*. \tag{6.8}$$

By Theorem 6.6, $\boldsymbol{\chi}=\boldsymbol{\Psi}^{-1}\,\mathbf{G}=\boldsymbol{\Phi}\sqrt{\mathbf{G}}\in\mathbf{L}_\infty^{0+}$. Since $\boldsymbol{\Phi}_0=\sqrt{\mathbf{G}}$, it follows that $\boldsymbol{\chi}_0=\mathbf{G}$; whence $\boldsymbol{\chi}=\boldsymbol{\Phi}\sqrt{\boldsymbol{\chi}_0}$. Thus

6.9 Corollary. *Let* $\boldsymbol{\chi}$ *be defined as in* (6.8) *and* (6.7). *Then*

$$\mathbf{G}=\boldsymbol{\chi}_0,\quad \boldsymbol{\Phi}=\boldsymbol{\chi}\,(\sqrt{\boldsymbol{\chi}_0^{-1}}).$$

We shall now express the linear predictor and prediction error matrix in terms of $\boldsymbol{\Psi}$.

Since by (2.5) and Theorem 5.5 (a), $\boldsymbol{\Phi}$, $\boldsymbol{\Phi}^{-1}\in\mathbf{L}_\infty^{0+}$, it follows from Theorem 6.6 that $\boldsymbol{\Psi}$, $\boldsymbol{\Psi}^{-1}\in\mathbf{L}_\infty^{0+}$. Let

$$\boldsymbol{\Psi}(e^{i\theta})=\sum_0^\infty \mathbf{A}_k\,e^{ki\theta},\qquad \boldsymbol{\Psi}^{-1}(e^{i\theta})=\sum_0^\infty \mathbf{B}_k\,e^{ki\theta}.$$

From (6.5) we find that $\mathbf{A}_0 = \mathbf{I}$ and for $m > 0$

6.10
$$\mathbf{A}_m = -\mathbf{\Gamma}'_m + \sum_n \mathbf{\Gamma}'_n \mathbf{\Gamma}'_{m-n} - \sum_n \sum_p \mathbf{\Gamma}'_p \mathbf{\Gamma}'_{n-p} \mathbf{\Gamma}'_{m-n} + \cdots, \tag{6.10}$$

where $\mathbf{\Gamma}'_k$ is the kth Fourier coefficient of $\mathbf{M}$ and all subscripts run from 1 to ∞. The coefficients $\mathbf{A}_k$ are thus determinable.(1) The coefficients $\mathbf{B}_k$ can be found from the recurrence relations

6.11
$$\left.\begin{array}{c} \mathbf{A}_0 \mathbf{B}_0 = \mathbf{I} = \mathbf{B}_0 \mathbf{A}_0 \\ \mathbf{A}_0 \mathbf{B}_1 + \mathbf{A}_1 \mathbf{B}_0 = \mathbf{0} = \mathbf{B}_0 \mathbf{A}_1 + \mathbf{B}_1 \mathbf{A}_0 \\ \mathbf{A}_0 \mathbf{B}_2 + \mathbf{A}_1 \mathbf{B}_1 + \mathbf{A}_2 \mathbf{B}_0 = \mathbf{0} = \mathbf{B}_0 \mathbf{A}_2 + \mathbf{B}_1 \mathbf{A}_1 + \mathbf{B}_2 \mathbf{A}_0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{array}\right\} \tag{6.11}$$

Since $\mathbf{A}_0 = \mathbf{I}$ matrix inversion will not be encountered in finding the $\mathbf{B}_k$.

Now let $\mathbf{C}_k$, $\mathbf{D}_k$ be the kth Fourier coefficients of $\mathbf{\Phi}$, $\mathbf{\Phi}^{-1} \in \mathbf{L}_\infty^{0+}$. Then by Theorem 6.6 $\mathbf{A}_k = \sqrt{\mathbf{G}}\, \mathbf{D}_k$, $\mathbf{B}_k = \mathbf{C}_k \sqrt{\mathbf{G}^{-1}}$. Hence by Theorem 5.7

$$\mathbf{E}_{\nu k} = \sum_{n=0}^{k} \mathbf{C}_{\nu+n} \mathbf{D}_{k-n} = \sum_{n=0}^{k} \mathbf{B}_{\nu+n} \mathbf{A}_{k-n}$$

$$\mathbf{G}_\nu = \sum_{n=0}^{\nu-1} \mathbf{C}_n \mathbf{C}_n^* = \sum_{n=0}^{\nu-1} \mathbf{B}_n \mathbf{G} \mathbf{B}_n^*.$$

$\mathbf{G}$ may be evaluated from the formula $\mathbf{G} = \mathbf{\Psi} \mathbf{F}' \mathbf{\Psi}^*$. Since $\mathbf{G}$ is constant we may take instead the average:

6.12
$$\mathbf{G} = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{\Psi}(e^{i\theta}) \mathbf{F}'(e^{i\theta}) \mathbf{\Psi}^*(e^{i\theta})\, d\theta. \tag{6.12}$$

We may sum up these results as follows.

6.13 THEOREM. *Let the q-ple stationary S.P. satisfy Assumption* 6.1. *Then for* $\nu > 0$,

$$\hat{\mathbf{f}}_\nu = (\mathbf{f}_\nu \mid \mathfrak{M}_0) = \sum_{k=0}^{\infty} \mathbf{E}_{\nu k} \mathbf{f}_{-k}, \qquad \mathbf{E}_{\nu k} = \sum_{n=0}^{k} \mathbf{B}_{\nu+n} \mathbf{A}_{k-n}$$

$$\mathbf{G}_\nu = (\mathbf{f}_\nu - \hat{\mathbf{f}}_\nu,\ \mathbf{f}_\nu - \hat{\mathbf{f}}_\nu) = \sum_{k=0}^{\nu-1} \mathbf{B}_n \mathbf{G} \mathbf{B}_n^*,$$

where $\mathbf{A}_k$, $\mathbf{B}_k$ (*the Fourier coefficients of the function* $\mathbf{\Psi}$ *of* (6.5) *and* $\mathbf{\Psi}^{-1}$) *and* $\mathbf{G}$ *are given by* (6.10), (6.11) *and* (6.12).

We thus have an explicit method of computing the weighting factors $\mathbf{E}_{\nu k}$ in the time-domain. It easily follows from (5.8) and Theorem 6.6 that the corresponding transfer-function in the frequency-domain can be expressed in the form

(1) Since $\mathbf{F}' = \mathbf{I} + \mathbf{M}$, we obviously have $\mathbf{\Gamma}'_0 = (\mathbf{f}_0, \mathbf{f}_0) - \mathbf{I}$ and $\mathbf{\Gamma}'_n = (\mathbf{f}_n, \mathbf{f}_0)$, $n \neq 0$.

6.14 $$\mathbf{Y}_\nu(e^{i\theta}) = [e^{-\nu i\theta}\,\mathbf{\Psi}^{-1}(e^{i\theta})]_{0+}\,\mathbf{\Psi}(e^{i\theta}). \tag{6.14}$$

Since for purposes of prediction the Assumption 6.1 is no stronger than the Boundedness Condition 5.1, we have solved the Prediction Problem 2.2 and the corresponding problem in the frequency-domain for processes satisfying the latter condition.

7. Estimation of the spectral density

In this section we shall consider the computation of the spectral density function of a q-ple, regular, full-rank S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ from its correlation matrices $\mathbf{\Gamma}_n$, which in turn are to be derived from time series observations in the past. We shall show that on account of the errors inherent in all observation and estimation such an empirically determined spectral density will satisfy the Boundedness Condition 5.1. This condition will thus be fulfilled in many practical cases of prediction.

Suppose that the correlation matrices $\mathbf{\Gamma}_n$ have been obtained and we wish to estimate the spectral density $\mathbf{F}$. In practice we will know the values of only a finite number of $\mathbf{\Gamma}_n$. A natural approach would therefore be to take the Cesaro partial sums of the Fourier series of $\mathbf{F}$:

7.1 $$\mathbf{F}_N(e^{i\theta}) = \sum_{n=-N+1}^{N-1}\left(1-\frac{|n|}{N}\right)\mathbf{\Gamma}_n\,e^{ni\theta} \tag{7.1}$$

as estimates of $\mathbf{F}$. This has the merit that as $N\to\infty$, $\mathbf{F}_N(e^{i\theta})\to\mathbf{F}(e^{i\theta})$, a.e., as follows from a trivial matricial extension of the Fejer-Lebesgue Theorem. Also by a simple rearrangement of terms we get

$$\begin{aligned}\mathbf{F}_N(e^{i\theta}) &= \frac{1}{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\mathbf{\Gamma}_{j-k}\,e^{i(j-k)\theta}\\ &= \frac{1}{N}\left(\sum_{j=1}^{N}e^{ji\theta}\,\mathbf{f}_j,\ \sum_{k=1}^{N}e^{ki\theta}\,\mathbf{f}_k\right),\end{aligned}$$

so that

7.2 $$\mathbf{F}_N(e^{i\theta}) \textit{ is non-negative hermitian for } 0\leqslant\theta\leqslant 2\pi. \tag{7.2}$$

Since every Laurent polynomial in $e^{i\theta}$ with matrix coefficients is bounded, we see that every such estimate $\mathbf{F}_N$ will be bounded above.

When we take into account not only the evaluation of the spectral density from the $2N-1$ correlation matrices $\mathbf{\Gamma}_n$ with $|n|<N$, but also the derivation of these matrices from empirical data on time series, we find that each estimate $\mathbf{F}_N$ is also bounded away from zero, in the sense that its eigenvalues are bounded away from 0. To see this we shall first discuss a way of estimating the matrix coefficient $(1-|n|/N)\,\mathbf{\Gamma}_n$ of $e^{ni\theta}$ in the Cesaro partial sum (7.1). Since $\mathbf{\Gamma}_{-n}=\mathbf{\Gamma}_n^*$, it suffices to take $0\leqslant n<N$.

The past values $\mathbf{x}_{-k} = \mathbf{f}_{-k}(\omega)$, $k > 0$, of particular time series of the S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ can be found from observation. From the record of these observations for the ith and jth components:

$$x_{-k}^{(i)} = f_{-k}^{(i)}(\omega), \qquad x_{-k}^{(j)} = f_{-k}^{(j)}(\omega), \qquad k \geq 0,\ 1 \leq i, j \leq q \tag{1}$$

we can compute the one-sided time-average

$$\gamma_n^{(N)}(i, j, \omega) = \frac{1}{N} \sum_{k=n}^{N-1} f_{n-k}^{(i)}(\omega)\, \overline{f_{-k}^{(j)}(\omega)}, \qquad 0 \leq n < N, \tag{2}$$

in which the number of terms is $N - n$, and therefore depends on the lead n. (Since $0 \leq n < N$, each sum will have at least one term.) The reason for this choise of the number of terms is to make the expected value of $\gamma_n^{(N)}(i, j, \omega)$ equal to the (i, j)th entry of the desired matrix $(1 - n/N)\,\mathbf{\Gamma}_n$:

$$\int_\Omega \gamma_n^{(N)}(i, j, \omega)\, dP(\omega) = \frac{1}{N} \sum_{k=n}^{N-1} (f_{n-k}^{(i)}, f_{-k}^{(j)}) = \left(1 - \frac{n}{N}\right) \gamma_n^{ij},$$

where $\gamma_n^{ij} = (f_n^{(i)}, f_0^{(j)})$ is the (i, j)th entry of $\mathbf{\Gamma}_n$. Putting $\mathbf{\Gamma}_n^{(N)}(\omega) = \gamma[{}_n^{(N)}(i, j, \omega)]$, it follows that

$$\overline{\mathbf{\Gamma}}_n^{(N)} = \int_\Omega \mathbf{\Gamma}_n^{(N)}(\omega)\, dP(\omega) = \left(1 - \frac{n}{N}\right) \mathbf{\Gamma}_n. \tag{3}$$

We must now evaluate the expected value $\overline{\mathbf{\Gamma}}_n^{(N)}$ in (3). In the Wiener-Kolmogorov prediction theory the shift operator U of the S.P. $(\mathbf{f}_n)_{-\infty}^{\infty}$ is generated by a measure-preserving transformation T on the probability space Ω onto itself (cf. Doob [4, p. 461–464]). In this case

$$f_{n+k}^{(i)}(\omega) = (U^k f_n^{(i)})(\omega) = f_n^{(i)}(T^k \omega). \tag{4}$$

In the light of the theorems of von Neumann [8] and Oxtoby and Ulam [7], which assert roughly that every measure-preserving transformation can be resolved into ergodic components, and that nearly all continuous measure-preserving transformations are ergodic, we may take the transformation T in (4) to be ergodic. Since every function $f_k^{(i)}$ is in $\mathfrak{L}_2$, it follows from (2) that the entries of the random matrix-valued function $\mathbf{\Gamma}_n^{(N)}$ are in L_1 on Ω. Hence by a trivial matricial extension of Birkhoff's Ergodic Theorem [1], for almost all ω

$$\lim_{\nu \to \infty} \frac{1}{\nu + 1} \{\mathbf{\Gamma}_n^{(N)}(\omega) + \mathbf{\Gamma}_n^{(N)}(T^{-1}\omega) + \cdots + \mathbf{\Gamma}_n^{(N)}(T^{-\nu}\omega)\} = \overline{\mathbf{\Gamma}}_n^{(N)}. \tag{5}$$

Now from (2) and (4) we find that for $0 \leq \mu \leq \nu$,

$$\begin{aligned}\gamma_n^{(N)}(i, j, T^{-\mu}\omega) &= \frac{1}{N}\sum_{k=n}^{N-1} f_{n-k}^{(i)}(T^{-\mu}\omega)\cdot\overline{f_{-k}^{(j)}(T^{-\mu}\omega)}\\ &= \frac{1}{N}\sum_{k=n}^{N-1} f_{n-k-\mu}^{(i)}(\omega)\cdot\overline{f_{-k-\mu}^{(j)}(\omega)}\\ &= \sum_{\lambda=n+\mu}^{N+1-\mu} f_{n-\lambda}^{(i)}(\omega)\cdot\overline{f_{-\lambda}^{(j)}(\omega)}.\end{aligned}$$

This average thus extends only into the past, and is therefore computable from the observed values (1) of the component time series of the process. Consequently, the average with $\nu+1$ terms on the L.H.S. of (5) is also computable from such observed values. For sufficiently large ν we may take this average as an approximation [1] to the expected value $\overline{\mathbf{\Gamma}}_n^{(N)}$, i.e. by (3) to the desired matrix $(1-n/N)\mathbf{\Gamma}_n$. We thus have a method of approximating to these matrices by using data collected from time series observations.

Now in measuring the values of a time series, random instrumental errors will be inevitable, so that the result of measurement will be

$$\mathbf{g}_n(\omega) = \mathbf{f}_n(\omega) + \boldsymbol{\varphi}_n(\omega), \tag{6}$$

$\boldsymbol{\varphi}_n(\omega)$ being the "noise" or disturbance caused by measurement. We may assume that the conditions of measurement are kept constant, and that the errors are mutually independent. It will then follow that $(\boldsymbol{\varphi}_n)_{-\infty}^{\infty}$ is a q-ple *white noise process*, i.e.

$$(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_n) = \gamma_{mn}\mathbf{\Lambda}, \tag{7}$$

where $\mathbf{\Lambda}$ is a fixed matrix. To obtain $\mathbf{\Lambda}$ we note that a measurement of the vector $\mathbf{x}_n = \mathbf{f}_m(\omega)$ consists actually of q simple measurements, one for each component. The measurement of the ith component will involve an error $\phi_n^{(i)}(\omega)$ of absolute value λ_i, say. These errors being independent, we will have

$$(\phi_n^{(i)}, \phi_n^{(j)}) = \int_\Omega \phi_n^{(i)}(\omega)\,\overline{\phi_n^{(j)}(\omega)}\,dP(\omega) = \delta_{ij}\lambda_i^2,$$

which shows that

$$\mathbf{\Lambda} \textit{ is a diagonal matrix with positive entries.} \tag{8}$$

Finally, since the errors of measurement will be independent of the size of the measured quantity, we will have

$$(\mathbf{f}_m, \boldsymbol{\varphi}_n) = \mathbf{0}, \qquad -\infty < m, n < \infty. \tag{9}$$

---

(1) In this paper we will not discuss the difficult question of the mode of approximation nor the question as to how large ν must be taken in order to secure a given degree of approximation.

From (6), (7) and (9), we get

$$\mathbf{\Gamma}'_{m-n} = (\mathbf{g}_m, \mathbf{g}_n) = (\mathbf{f}_m, \mathbf{f}_n) + \delta_{mn}\mathbf{\Lambda} = \mathbf{\Gamma}_{m-n} + \delta_{mn}\mathbf{\Lambda}, \tag{10}$$

which shows that the actually observed process $(\mathbf{g}_n)_{-\infty}^{\infty}$ is stationary. Since in any practical case the correlation matrices computed from time series data will be the $\mathbf{\Gamma}'_n = (\mathbf{g}_n, \mathbf{g}_0)$ and not the $\mathbf{\Gamma}_n = (\mathbf{f}_n, \mathbf{f}_0)$, the empirically derived spectral density will not be $\mathbf{F}_N$ as given in (7.1) but rather (cf. (10))

$$\begin{aligned}\mathbf{G}_N(e^{i\theta}) &= \sum_{n=-N+1}^{N-1}\left(1 - \frac{|n|}{N}\right)\mathbf{\Gamma}'_n e^{ni\theta} \\ &= \sum_{n=-N+1}^{N-1}\left(1 - \frac{|n|}{N}\right)\mathbf{\Gamma}_n e^{ni\theta} + \mathbf{\Lambda} \\ &= \mathbf{F}_N(e^{i\theta}) + \mathbf{\Lambda}.\end{aligned}$$

By (7.2) $\mathbf{G}_N(e^{i\theta}) \succ \mathbf{\Lambda}$, and hence by (8) the eigenvalues of $\mathbf{G}_N$ are bounded away from zero.

Thus the empirically derived spectral density will not only be bounded above but also bounded away from zero, in the sense that its eigenvalues will be bounded away from zero. Denoting this spectral density by $\mathbf{F}$ instead of $\mathbf{F}_N$ or $\mathbf{G}_N$, we may by 1.5 (a) restate its boundedness property in the form 5.1, and sum up the preceeding discussion as in the next theorem.

7.3 THEOREM. *On account of errors of observation and estimation, any estimate of the spectral density of a regular full rank* S.P., *derived from its correlation matrices, which are obtained by averaging time series data, will satisfy the Boundedness Condition* 5.1.

There are physical processes in which periodicities, though imperfect, are so marked that it is untenable to postulate regularity, and it becomes convenient to admit non-absolutely continuous spectral distributions. In such cases the foregoing considerations will not of course apply. Cases are also conceivable in which we may be able to hypothesize the values of the correlation matrices $\mathbf{\Gamma}_n$ from a theoretical study of the process without recourse to sampling. If the hypothetical $\mathbf{\Gamma}_n$ do not die down with sufficient rapidity as $n \to \pm\infty$, the foregoing remarks would again be inapplicable.

8. A general factorization algorithm

In this section we shall show how the iterative method developed in Sec. 6 to get the generating function can be generalized to solve the following problem.

8.1. Factorization Problem. *Given a* $q \times q$ (1) *matrix-valued function* $\mathbf{F}$ *on* C *such that* $\mathbf{F} \in \mathbf{L}_1$ *and* $\log |\Delta \mathbf{F}| \in L_1$, *to find functions* $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$ *on* C *with the properties*

$$\mathbf{F}(e^{i\theta}) = \mathbf{\Phi}_1(e^{i\theta})\, \mathbf{\Phi}_2(e^{i\theta}), \quad \text{a.e.} \tag{1}$$

$$\mathbf{\Phi}_1 \in \mathbf{L}_2^{0+}, \qquad \mathbf{\Phi}_2 \in \mathbf{L}_2^{0-} \tag{2}$$

$$|\Delta\{\mathbf{\Phi}_1(0)\}|^2 = \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log|\Delta\{\mathbf{F}(e^{i\theta})\}|\, d\theta\right]. \tag{3}$$

We shall solve this proplem under the

8.2 ASSUMPTION. $\mathbf{F}(e^{i\theta}) = \mathbf{I} + \mathbf{M}(e^{i\theta})$, $\quad |\mathbf{M}(e^{i\theta})|_B \leqslant \mu < 1$, a.e.

Our method will work in different settings, e.g. when C is replaced by an annulus A, and $\mathbf{F}$ by a matrix-valued function holomorphic on A, and with suitable restrictions even for functions whose values are operators on an infinite-dimensional space, cf. [6]; but we shall confine ourselves here to the version given in 8.1. Such factorization problems have no bearing on prediction theory except when $\mathbf{F}$ is hermitian-valued, but are important in other branches of analysis, cf. [6, Sec. 1].

8.3 DEFINITION. *With any function* $\mathbf{M} \in \mathbf{L}_\infty$ *we associate two operators on* $\mathbf{L}_2$ *defined by*

$$\mathfrak{P}_+(\mathbf{\Phi}) = (\mathbf{\Phi}\mathbf{M})_+, \quad \mathfrak{P}_-(\mathbf{\Phi}) = (\mathbf{M}\mathbf{\Phi})_-, \quad \mathbf{\Phi} \in \mathbf{L}_2.$$

Some easily established properties of these operators are stated in the next lemma.

8.4 LEMMA. (a) $\mathfrak{P}_+$, $\mathfrak{P}_-$ *are in the Banach algebra* $\mathfrak{B}$ *of* 1.16; *more fully, they are bounded linear operators on* $\mathbf{L}_2$ *into* $\mathbf{L}_2^+$, $\mathbf{L}_2^-$, *respectively, and*

$$|\mathfrak{P}_+|,\ |\mathfrak{P}_-| \leqslant \mu, \quad \text{where } \mu = \operatorname*{ess.\,l.u.b.}_{0 \leqslant \theta \leqslant 2\pi} |\mathbf{M}(e^{i\theta})|_B.$$

(b) *If* $\mathfrak{J}$ *is the unit of* $\mathfrak{B}$, *then* $\mathfrak{J} + \mathfrak{P}_+$, $\mathfrak{J} - \mathfrak{P}_-$ *are invertible and*

$$(\mathfrak{J} + \mathfrak{P}_\pm)^{-1} = \mathfrak{J} - \mathfrak{P}_\pm + \mathfrak{P}_\pm^2 - \cdots,$$

where the last series is absolutely convergent in $\mathfrak{B}$, *being in fact dominated by the convergent geometric series* $\Sigma\, \mu^k$.

(c) $$\mathfrak{P}_+^{n+1}(\mathbf{\Phi}) = \{\mathfrak{P}_+^n(\mathbf{\Phi}) \cdot \mathbf{M}\}_+, \quad \mathfrak{P}_-^{n+1}(\mathbf{\Phi}) = \{\mathbf{M} \cdot \mathfrak{P}_-^n(\mathbf{\Phi})\}_-.$$

(d) $$\mathfrak{P}_+(\mathbf{I}) = \mathbf{M}_+, \ \mathfrak{P}_+^2(\mathbf{I}) = (\mathbf{M}_+\mathbf{M})_+, \ \mathfrak{P}_+^3(\mathbf{I}) = \{(\mathbf{M}_+\mathbf{M})_+\mathbf{M}\}_+, \ \ldots$$

$$\mathfrak{P}_-(\mathbf{I}) = \mathbf{M}_-, \ \mathfrak{P}_-^2(\mathbf{I}) = (\mathbf{M}\mathbf{M}_-)_-, \ \mathfrak{P}_-^3(\mathbf{I}) = \{\mathbf{M}(\mathbf{M}\mathbf{M}_-)_-\}_-, \ \ldots$$

(1) Not necessarily hermitian.

(e) $$\|\mathfrak{P}_+^n(\mathbf{I})\|, \quad \|\mathfrak{P}_-^n(\mathbf{I})\| \leqslant \mu^n \sqrt{q}.$$

The following definition therefore makes sense.

8.5 Definition. (a) $\boldsymbol{\Psi}_{0+} = (\mathfrak{J} + \mathfrak{P}_+)^{-1}(\mathbf{I}), \quad \boldsymbol{\Psi}_{0-} = (\mathfrak{J} + \mathfrak{P}_-)^{-1}(\mathbf{I}).$

(b) $\mathbf{G} = \boldsymbol{\Psi}_{0+}(\mathbf{I} + \mathbf{M})\boldsymbol{\Psi}_{0-}.$

We shall now prove the crucial result that the function $\mathbf{G}$ is constant-valued, the constant being an invertible matrix. This will be done by considering the Fourier series of $\mathbf{G}$. We shall first show that $\mathbf{G} \in \mathbf{L}_1$, and therefore has such a series.

8.6 Lemma. (a) $$\boldsymbol{\Psi}_{0+} = \mathbf{I} - \mathfrak{P}_+(\mathbf{I}) + \mathfrak{P}_+^2(\mathbf{I}) - \cdots \in \mathbf{L}_2^{0+},$$
$$\boldsymbol{\Psi}_{0-} = \mathbf{I} - \mathfrak{P}_-(\mathbf{I}) + \mathfrak{P}_-^2(\mathbf{I}) - \cdots \in \mathbf{L}_2^{0-},$$

the infinite series being absolutely convergent in the norm of $\mathbf{L}_2$, *cf.* (1.7).

(b) $\mathbf{G} \in \mathbf{L}_1$.

Proof. (a) The series expansions obviously follow from the last definition, and the expansions given in 8.4. Since the ranges of $\mathfrak{P}_+$, $\mathfrak{P}_-$ are included in $\mathbf{L}_2^+$, $\mathbf{L}_2^-$, and these are (closed) subspaces of $\mathbf{L}_2$, it follows from the expansions that $\boldsymbol{\Psi}_{0+} \in \mathbf{L}_2^{0+}$ and $\boldsymbol{\Psi}_{0-} \in \mathbf{L}_2^{0-}$. Also these series converge absolutely in the $\mathbf{L}_2$-norm, since by 8.4 (e)

$$\sum_0^\infty \|\mathfrak{P}_\pm^n(\mathbf{I})\| \leqslant \sqrt{q} \cdot \sum_0^\infty \mu^n < \infty.$$

(b) follows from (a), since $\mathbf{I} + \mathbf{M} \in \mathbf{L}_\infty$. (Q.E.D.)

8.7 Theorem. (a) $$(\mathbf{I} + \mathbf{M})\boldsymbol{\Psi}_{0-} = \mathbf{I} + (\mathbf{M}\boldsymbol{\Psi}_{0-})_{0+} \in \mathbf{L}_2^{0+},$$
$$\boldsymbol{\Psi}_{0+}(\mathbf{I} + \mathbf{M}) = \mathbf{I} + (\boldsymbol{\Psi}_{0+}\mathbf{M})_{0-} \in \mathbf{L}_2^{0-}.$$

(b) $\mathbf{G} = \text{const.} = \mathbf{I} + (\mathbf{M}\boldsymbol{\Psi}_{0-})_0 = \mathbf{I} + (\boldsymbol{\Psi}_{0+}\mathbf{M})_0.$

(c) $\mathbf{I} + \mathbf{M}$, $\boldsymbol{\Psi}_{0+}$, $\boldsymbol{\Psi}_{0-}$ *are invertible* a.e. *on* C, *and* $(\mathbf{I} + \mathbf{M})^{-1} \in \mathbf{L}_\infty$.

(d) $\mathbf{G}$ *is invertible.*

Proof. (a) Since $\mathbf{I} + \mathbf{M} \in \mathbf{L}_\infty$ and $\boldsymbol{\Psi}_{0-} \in \mathbf{L}_2$, therefore $(\mathbf{I} + \mathbf{M})\boldsymbol{\Psi}_{0-} \in \mathbf{L}_2$. Also by 1.15 (b) (1) and 8.3

$$\begin{aligned}(\mathbf{I} + \mathbf{M})\boldsymbol{\Psi}_{0-} &= \boldsymbol{\Psi}_{0-} + (\mathbf{M}\boldsymbol{\Psi}_{0-})_- + (\mathbf{M}\boldsymbol{\Psi}_{0-})_{0+} \\ &= (\mathfrak{J} + \mathfrak{P}_-)(\boldsymbol{\Psi}_{0-}) + (\mathbf{M}\boldsymbol{\Psi}_{0-})_{0+}.\end{aligned}$$

By 8.5 (a), the first term on the right is $\mathbf{I}$. This gives the first relation in (a). The second is proved similarly.

(b) By 8.5 (b) $\mathbf{G}=\mathbf{\Psi}_{0+}(\mathbf{I}+\mathbf{M})\mathbf{\Psi}_{0-}$. Now $\mathbf{\Psi}_{0+}\in\mathbf{L}_2^{0+}$, and as just shown $(\mathbf{I}+\mathbf{M})\mathbf{\Psi}_{0-}\in\mathbf{L}_2^{0+}$; hence by the Convolution Rule [I, 3.9 (d)] $\mathbf{G}\in\mathbf{L}_1^{0+}$. But we also know that $\mathbf{\Psi}_{0+}(\mathbf{I}+\mathbf{M})$, $\mathbf{\Psi}_{0-}\in\mathbf{L}_2^{0-}$; hence $\mathbf{G}\in\mathbf{L}_1^{0-}$. It follows that all the Fourier coefficients of $\mathbf{G}$, except the 0^{th} vanish, i.e. $\mathbf{G}=\text{const}$.

Next, since the range of $\mathfrak{P}_+$ is included in, $\mathbf{L}_2^+$, therefore by 8.6 (a) $(\mathbf{\Psi}_{0+})_0=\mathbf{I}$. We may accordingly write $\mathbf{\Psi}_{0+}=\mathbf{I}+\mathbf{\Psi}_+$, where $\mathbf{\Psi}_+\in\mathbf{L}_2^+$. This fact together with the first equality in (a) entail that

$$\begin{aligned}\mathbf{G}&=\mathbf{\Psi}_{0+}(\mathbf{I}+\mathbf{M})\mathbf{\Psi}_{0-}=\mathbf{\Psi}_{0+}\{\mathbf{I}+(\mathbf{M}\mathbf{\Psi}_{0-})_{0+}\}.\\&=\mathbf{I}+\mathbf{\Psi}_++(\mathbf{I}+\mathbf{\Psi}_+)(\mathbf{M}\mathbf{\Psi}_{0-})_{0+}\\&=\mathbf{I}+\mathbf{\Psi}_++(\mathbf{M}\mathbf{\Psi}_{0-})_{0+}+\mathbf{\Psi}_+(\mathbf{M}\mathbf{\Psi}_{0-})_{0+}.\end{aligned}$$

Since $\mathbf{G}$ is a constant, it follows that

$$\mathbf{G}=\mathbf{G}_0=\mathbf{I}+(\mathbf{M}\mathbf{\Psi}_{0-})_0.$$

The other expression for $\mathbf{G}$ is proved similarly.

(c) By Assumption 8.2, $\mathbf{I}+\mathbf{M}$ is in the Banach-algebra $\mathbf{L}_\infty$ at a distance μ less than 1 from $\mathbf{I}$. Hence it is invertible, and the function $(\mathbf{I}+\mathbf{M})^{-1}\in\mathbf{L}_\infty$. Next, since $\mathbf{\Psi}_{0+}\in\mathbf{L}_2^{0+}$, it follows from I, 3.13 (a) that $\Delta\mathbf{\Psi}_{0+}\in H_{2/q}$ on D_+. Also, $\mathbf{\Psi}_{0+}(0)=\mathbf{I}$, and therefore $\Delta\{\mathbf{\Psi}_{0+}(0)\}\neq 0$. Hence by the Riesz–Nevanlinna Theorem [I, 2.7], its radial limit can vanish almost nowhere on C. Hence $\mathbf{\Psi}_{0+}$ is invertible a.e. on C. By an inversion $z'=1/z$ of D_- onto D_+, we can show that the same is the case with the function $\mathbf{\Psi}_{0-}$.

(d) By 8.5 (b), $\mathbf{G}=\mathbf{\Psi}_{0+}(e^{i\theta})\{\mathbf{I}+\mathbf{M}(e^{i\theta})\}\mathbf{\Psi}_{0-}(e^{i\theta})$, a.e. Taking a θ for which all the three factors on the right are invertible, we see that $\mathbf{G}$ is invertible. (Q.E.D.)

In view of 8.7 (c) we may invert the equation 8.5 (b) to get

$$\mathbf{I}+\mathbf{M}=\mathbf{\Psi}_{0+}^{-0}\mathbf{G}\mathbf{\Psi}_{0-}^{-1},\quad \text{a.e.}$$

We shall now show that inverses $\mathbf{\Psi}_{0+}^{-1}$, $\mathbf{\Psi}_{0-}^{-1}$ are themselves in $\mathbf{L}_2^{0+}$, $\mathbf{L}_2^{0-}$, respectively, so that we have a factorization of the desired kind.

8.8 LEMMA. (a) $\mathbf{\Psi}_{0+}^{-1}=\{\mathbf{I}+(\mathbf{M}\mathbf{\Psi}_{0-})_{0+}\}\mathbf{G}^{-1}\in\mathbf{L}_2^{0+}$

(b) $\mathbf{\Psi}_{0-}^{-1}=\{\mathbf{I}+(\mathbf{\Psi}_{0+}\mathbf{M})_{0-}\}\mathbf{G}^{-1}\in\mathbf{L}_2^{0-}$.

Proof. (a) By 8.7 (a) and 8.5 (b)

$$\mathbf{\Psi}_{0+}\{\mathbf{I}+(\mathbf{M}\mathbf{\Psi}_{0-})_{0+}\}\mathbf{G}^{-1}=\mathbf{\Psi}_{0+}\cdot(\mathbf{I}+\mathbf{M})\mathbf{\Psi}_{0-}\cdot\mathbf{G}^{-1}=\mathbf{G}\mathbf{G}^{-1}=\mathbf{I}.$$

Since our matrices are finite dimensional, we conclude that

$$\mathbf{\Psi}_{0+}^{-1}=\{\mathbf{I}+(\mathbf{M}\mathbf{\Psi}_{0-})_{0+}\}\mathbf{G}^{-1}.$$

From the constancy of $\mathbf{G}^{-1}$, it at once follows that the function on the right is in $\mathbf{L}_2^{0+}$.

(b) is proved similarly. (Q.E.D.)

We shall next prove the following relations.

8.9 Lemma. (a) $\int_0^{2\pi} \log |\Delta \{\mathbf{\Psi}_{0+}(e^{i\theta})\}| \, d\theta = 0.$

(b) $\int_0^{2\pi} \log |\Delta \{\mathbf{\Psi}_{0-}(e^{i\theta})\}| \, d\theta = 0.$

(c) $\int_0^{2\pi} \log |\Delta \{\mathbf{I} + \mathbf{M}(e^{i\theta})\}| \, d\theta = \log |\Delta(\mathbf{G})|.$

Proof. (a) As noted in the proof of 8.7 (c) $\Delta \mathbf{\Psi}_{0+} \in H_{2/q}$ on D_+ and $\Delta\{\mathbf{\Psi}_{0+}(0)\} = 1$. Hence by I, 2.6 (c) (2)

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

Since by 8.8 $\mathbf{\Psi}_{0+}^{-1} \in \mathbf{L}_2^{0+}$, we get in exactly the same way

$$0 = \log |\Delta \{\mathbf{\Psi}_{0+}^{-1}(0)\}| \leqslant \frac{1}{2\pi} \int_0^{2\pi} \log |\Delta \{\mathbf{\Psi}_{0+}^{-1}(e^{i\theta})\}| \, d\theta,$$

i.e.

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

From (1) and (2) we get the desired equality.

(b) can be established in the same way after an inversion $z' = 1/z$ of D_- onto D_+.

(c) From 8.5 (b)

$$|\Delta(\mathbf{G})| = |\Delta \mathbf{\Psi}_{0+}| \cdot |\Delta(\mathbf{I} + \mathbf{M})| \cdot |\Delta \mathbf{\Psi}_{0-}|, \quad \text{a.e.}$$

Taking logarithms and integrating over $[0, 2\pi]$, it follows from the constancy of $\mathbf{G}$ and the equalities (a) and (b) that

$$2\pi \cdot \log |\Delta(\mathbf{G})| = \int_0^{2\pi} \log \Delta |\{\mathbf{I} + \mathbf{M}(e^{i\theta})\}| \, d\theta.$$

(Q.E.D.)

To sum up, we have proved the following theorem.

8.10 Theorem. *If* (i) *the function* $\mathbf{M} \in \mathbf{L}_\infty$ *on* C, *and*

$$\mu = \operatorname*{ess.\,l.u.b.}_{0 \leqslant \theta \leqslant 2\pi} |\mathbf{M}(e^{i\theta})|_B < 1.$$

(ii) $$\Psi_{0+} = \mathbf{I} - \mathbf{M}_+ + (\mathbf{M}_+\mathbf{M})_+ - \{(\mathbf{M}_+\mathbf{M})_+\mathbf{M}\}_+ + \cdots$$
$$\Psi_{0-} = \mathbf{I} - \mathbf{M}_- + (\mathbf{M}\mathbf{M}_-)_- - \{\mathbf{M}(\mathbf{M}\mathbf{M}_-)_-\}_- + \cdots,$$

so that, $\mathbf{G}$ *being defined as in* 8.5 (b), *we have*

$$\Psi_{0+}^{-1} = \{\mathbf{I} + (\mathbf{M}\Psi_{0-})_{0+}\}\,\mathbf{G}^{-1}, \quad \Psi_{0-}^{-1} = \{\mathbf{I} + (\Psi_{0+}\mathbf{M})_{0-}\}\,\mathbf{G}^{-1}$$

then (a) $\Psi_{0+},\ \Psi_{0+}^{-1} \in \mathbf{L}_2^{0+}, \quad \Psi_{0-},\ \Psi_{0-}^{-1} \in \mathbf{L}_2^{0-}$,

(b) $\mathbf{I} + \mathbf{M}(e^{i\theta}) = \Psi_{+0}^{-1}(e^{i\theta})\,\mathbf{G}\,\Psi_{0-}^{-1}(e^{i\theta})$ a.e.,

(c) $$|\Delta(\mathbf{G})| = \exp\left[\frac{1}{2\pi}\int_0^{2\pi} \log|\Delta\{\mathbf{I} + \mathbf{M}(e^{i\theta})\}|\,d\theta\right].$$

Now let $\sqrt{\mathbf{G}}$ be any square root of $\mathbf{G}$. Then letting $\Phi_1 = \Psi_{0+}^{-1}\sqrt{\mathbf{G}}$, $\Phi_2 = \sqrt{\mathbf{G}}\,\Psi_{0-}^{-1}$ we get a solution of Factorization Problem 8.1 under the Assumption 8.2. We shall now show that when the values of $\mathbf{M}$ are hermitian this solution reduces to the one obtained in Sec. 6. Let $\mathbf{M} = \mathbf{M}^*$ on C. Then by 1.15 (b) (4),

$$\{\mathfrak{P}_+(\Phi)\}^* = \{(\Phi\mathbf{M})_+\}^* = \{(\Phi\mathbf{M})^*\}_- = (\mathbf{M}\Phi^*)_- = \mathfrak{P}_-(\Phi^*).$$

By induction it readily follows that

$$\{\mathfrak{P}_+^n(\mathbf{I})\}^* = \mathfrak{P}_-^n(\mathbf{I}),$$

whence by 8.6 (a) $\Psi_{0+}^* = \Psi_{0-}$. Hence $\mathbf{G} = \Psi_{0+}(\mathbf{I} + \mathbf{M})\Psi_{0+}^*$ is non-negative hermitian, in fact positive definite since it is invertible. Letting $\Phi_{0+} = \Psi_{0+}^{-1}\sqrt{\mathbf{G}}$, where $\sqrt{\mathbf{G}}$ is now the unique positive definite square-root of $\mathbf{G}$, the equality 8.10 (b) becomes

$$\mathbf{I} + \mathbf{M}(e^{i\theta}) = \Phi_{0+}(e^{i\theta}) \cdot \Phi_{0+}^*(e^{i\theta}), \quad \text{a.e.}$$

Inverting, we get

$$\{\mathbf{I} + \mathbf{M}(e^{i\theta})\}^{-1} = \Phi_{0+}^{*-1}(e^{i\theta}) \cdot \Phi_{0+}^{-1}(e^{i\theta}), \quad \text{a.e.}$$

Since $\mathbf{I} + \mathbf{M} \in \mathbf{L}_\infty$ and by 8.7 (c), $(\mathbf{I} + \mathbf{M})^{-1} \in \mathbf{L}_\infty$, it follows that Φ_{0+}, $\Phi_{0+}^{-1} \in \mathbf{L}_\infty$. Since $\Phi_{0+} = \Psi_{0+}^{-1}\sqrt{\mathbf{G}}$, $\Phi_{0+}^{-1} = (\sqrt{\mathbf{G}})^{-1}\Psi_{0+}$, we conclude from 8.10 (a) that Φ_{0+}, $\Phi_{0+}^{-1} \in \mathbf{L}_\infty^{0+}$. We also note that $\Phi_{0+}(0) = \Psi_{0+}^{-1}(0)\sqrt{\mathbf{G}} = \sqrt{\mathbf{G}}$. Writing Ψ, Φ instead of Ψ_{0+}, Φ_{0+} we get the following theorem.

8.11 THEOREM. *If* (i) *the function* $\mathbf{M} \in \mathbf{L}_\infty$ *and has hermitian values, and*

$$\mu = \operatorname*{ess.\,l.u.b.}_{0 \leqslant \theta \leqslant 2\pi} |\mathbf{M}(e^{i\theta})|_B < 1,$$

(ii) $$\Psi = \mathbf{I} - \mathbf{M}_+ + (\mathbf{M}_+\mathbf{M})_+ - \{(\mathbf{M}_+\mathbf{M})_+\mathbf{M}\}_+ + \cdots,$$

(iii) [1] $$\mathbf{G} = \mathbf{\Psi}(e^{i\theta})\{\mathbf{I} + \mathbf{M}(e^{i\theta})\}\mathbf{\Psi}^*(e^{i\theta}) = \mathbf{I}_+(\mathbf{\Psi M})_0, \quad \text{a.e.}$$

(iv) $$\mathbf{\Phi} = \mathbf{\Psi}^{-1}\sqrt{\mathbf{G}},$$

then (a) $\mathbf{\Phi}, \mathbf{\Phi}^{-1} \in \mathbf{L}_\infty^{0+}$, *and* $\mathbf{\Phi}_+(0)$ *is positive definite.*

(b) $\mathbf{I} + \mathbf{M}(e^{i\theta}) = \mathbf{\Phi}(e^{i\theta}) \cdot \mathbf{\Phi}^*(e^{i\theta})$ a.e.

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

The "existence part" of this theorem is, of course, subsumed in our **I**, 7.13, but whereas the proof of the latter is indirect and non-constructive, we now have explicit expressions for the factors. We shall now show that two solutions $\mathbf{\Phi}$ of the Factorization Problem 8.1 with $\mathbf{F}$ non-negative hermitian-valued, which are such that $\mathbf{\Phi}, \mathbf{\Phi}^{-1} \in \mathbf{L}_2^{0+}$ can differ only by a constant unitary factor.

8.12 UNIQUENESS THEOREM. *If*

(i) $$\mathbf{\Phi}, \mathbf{\Phi}^{-1}, \mathbf{\Psi}, \mathbf{\Psi}^{-1} \in \mathbf{L}_2^{0+} \text{ on } C.$$

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

then there exists a unitary matrix U_0 *such that*

$$\mathbf{\Phi}(e^{i\theta}) = \mathbf{\Psi}(e^{i\theta})\mathbf{U}_0 \quad \text{a.e.} \tag{1}$$

Further $\mathbf{U}_0 = \mathbf{I}$ *if either* $\mathbf{\Phi}_+(0)$, $\mathbf{\Psi}_+(0)$ *are equal, or they are positive definite.*

Proof. By (i) the functions

$$\mathbf{U} = \mathbf{\Psi}^{-1}\mathbf{\Phi}, \quad \mathbf{U}^{-1} = \mathbf{\Phi}^{-1}\mathbf{\Psi} \tag{2}$$

are defined a.e. on C. Next, by (ii)

$$\mathbf{U} = \mathbf{\Psi}^{-1}\mathbf{\Phi} = \mathbf{\Psi}^*(\mathbf{\Phi}^*)^{-1} = (\mathbf{\Phi}^{-1}\mathbf{\Psi})^* = (\mathbf{U}^{-1})^*.$$

Hence $\mathbf{U}, \mathbf{U}^{-1}$ *are unitary-valued on* C *and so are in* $\mathbf{L}_\infty$. (3)

Since by (i) $\mathbf{\Psi}^{-1}$, $\mathbf{\Phi}$ and $\mathbf{\Phi}^{-1}$, $\mathbf{\Psi}$ are in $\mathbf{L}_2^{0+}$, therefore by the Convolution Rule [**I**, 3.9 (d)], $\mathbf{U}, \mathbf{U}^{-1}$ are in $\mathbf{L}_1^{0+}$. From (3) we conclude that $\mathbf{U}, \mathbf{U}^{-1} \in \mathbf{L}_\infty^{0+}$. It follows that if $\mathbf{U}(e^{i\theta}) = \sum_0^\infty \mathbf{A}_n e^{ni\theta}$, then

$$\mathbf{U}^{-1}(e^{i\theta}) = \mathbf{U}^*(e^{i\theta}) = \sum_0^\infty \mathbf{A}_n^* e^{-ni\theta} \in \mathbf{L}_\infty^{0-}.$$

---

[1] Cf. 8.5 (b) and 8.7 (b).

Since $\mathbf{U}^{-1} \in \mathbf{L}_\infty^{0+}$, we conclude that $\mathbf{A}_n = \mathbf{0}$, for $n > 0$, and $\mathbf{A}_0^* = \mathbf{A}_0^{-1}$, i.e. $\mathbf{A}_0$ is a unitary matrix $\mathbf{U}_0$. Thus the function $\mathbf{U}$ is constant-valued. The desired result (1) now follows from (2).

It is clear from (i) and (1) that

$$\mathbf{\Phi}_+(z) = \mathbf{\Psi}_+(z) \cdot \mathbf{U}_0, \quad z \in D_+.$$
$$\mathbf{\Phi}_+^{-1}(z) = \mathbf{U}_0^* \mathbf{\Psi}_+^{-1}(z), \quad z \in D_+.$$

Taking $z = 0$ and noting that $\Psi_+(0)$ is invertible, we get the remaining results. (Q.E.D.)

References

[1]. G. D. Birkhoff, Proof of the ergodic theorem. *Proc. Nat. Acad. Sci. U.S.A.*, 17 (1931), 656–660.

[2]. H. W. Bode & C. E. Shannon, Linear least square smoothing and prediction theory. *Proc. Inst. Radio Engineers*, 38 (1950), 417–425.

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

[4]. J. L. Doob, *Stochastic Processes*. New York, (1953).

[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]. P. Masani, The Laurent factorization of operator-valued functions. *Proc. London Math. Soc.*, 6 (1956), 59–69.

[7]. J. C. Oxtoby & S. M. Ulam, Measure-preserving homeomorphims and metrical transitivity. *Ann. of Math.*, 42 (1941), 874–920.

[8]. J. von Neummann, Zur Operatorenmethode in der klassischen Mechanik. *Ann. of Math.*, 33 (1932), 587–642.

[9]. ——, *Functional Operators*, vol. II. Princeton (1950).

[10]. N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series.* New York (1950).

[11]. ——, On the factorization of matrices. *Comment. Math. Helv.*, 29 (1955), 97–111.

[12]. N. Wiener & P. Masani, The prediction theory of multivariate stochastic processes, Part I. *Acta Math.*, 98 (1957), 111–150.