Stationary Processes, 1963-64
K. Ito

VI. PREDICTION AND MOVING AVERAGE REPRESENTATION

Linear Problem. (discrete time parameter). Let x_n $n \in Z$, be a complex-valued weakly stationary stochastic sequence with mean 0. We have the following decompositions:

$$r(n) = r_x(n) = \int_\Gamma e^{i2\pi\lambda n}\, dF(\lambda) \qquad \Gamma = R|Z$$

$$x_n = \int_\Gamma e^{i2\pi\lambda n}\, dM(\lambda), (M(\Lambda_1), M(\Lambda_2)) = F(\Lambda_1 \cap \Lambda_2)$$

x_n is called trivial if $\Gamma(0) = 0$, i.e., if $x_n = 0$.

1. Definitions.

$L_{mn}(x)$ = closed linear subspace of $L^2(\Omega, \mathcal{B}, P)$ spanned by x_k, $m \le k \le n$.

$$L_n(x) = L_{-\infty,n}(x)$$

$$L(x) = L_\infty(x) = L_{-\infty,\infty}(x) = \bigvee_n L_n(x)$$

$L_{-\infty}(x) = \wedge L_n(x)$ (= the space of remote past), $m < n \Longrightarrow L_m(x) \subset L_n(x)$

shift operator U : unitary operator U determined by $U x_n = x_{n+1}$, $n \in Z$,
$U^n L_m(x) = L_{m+n}(x)$

Definition 1. x_n is called (purely) non-deterministic if $L_{-\infty}(x) = 0$ and (purely) deterministic if $L_{-\infty}(x) = L(x)$.

Definition 2. $x_n^d = P_{L_{-\infty}(x)} \cdot x_n$, $n \in Z$, is a deterministic stationary sequence and is called the deterministic part of x_n.

Definition 3. $x_n^i = P_{L(x) \ominus L_{-\infty}(x)} \cdot x_n$ is a non-deterministic stationary sequence and is called the non-deterministic part of x_n.

Definition 4. $x_n = x_n^i + x_n^d$ is called the Wold decomposition of x_n

Corollary 1. $L(x^d) = L_{-\infty}(x)$, $L(x^i) = L(x) \ominus L_{-\infty}(x)$.

Definition 5. Two stationary sequences x_n and y_n which may be defined on different probability spaces are called (weakly) equivalent (in symbol $x_n \sim y_n$) if $r_x(n) \equiv r_y(n)$.

Corollary 2. $x_n \sim y_n$ iff $\exists$ isomorphism $V: L(x) \to L(y)$ with $Vx_n = y_n$, $n \in Z$.

Corollary 3. A stationary sequence equivalent to a deterministic one is also deterministic. Similarly for "non-deterministic".

Definition 6. An orthonormal sequence ξ_n $((\xi_n, 1) = 0, (\xi_n, \xi_m) = \delta_{nm})$ is called white light or white noise.

The Hinčin measure of ξ_n is the uniform distribution on Γ:

$$r_\xi(n) = \delta_{no} = \int_\Gamma e^{i2\pi\lambda n}\, d\lambda,$$

from which the adjective "white" comes.

Corollary 4. $\xi_n = \int_\Gamma e^{i2\pi\lambda n} M_\xi(d\lambda)$, $(M(\Lambda_1), M(\Lambda_2)) = \int_{\Lambda_1 \cap \Lambda_2} d\lambda$

Corollary 5. A white noise is non-deterministic.

Corollary 6. A sequence equivalent to a white noise is also a white noise, and any two white noises are equivalent.

Definition 7. A stationary sequence x_n is said to have linear regression if it satisfies a linear difference equation with constant coefficients:

$$x_n + a_1 x_n + \cdots + a_m x_{n-m} = 0.$$

Corollary 7.

x_n has linear regression

<==> the Hinčin measure increases only with a finite number of jumps.

2. Moving Average Representation.

Given a white noise ξ_n and a (non-random) two-sided sequence $a \equiv (a_n) \in \ell^2(Z)$, form

$$y_n = \sum_m a_{n-m} \xi_m$$

i.e.,
$$y = a * \xi$$

Then y_n is also a stationary sequence with

$$r_y(n) = \int_\Gamma e^{i2\pi\lambda n} \left| a(e^{-i2\pi\lambda}) \right|^2 d\lambda$$

where

$$a(\zeta) = \sum_{n \in Z} a_n \zeta^n$$

In fact

$$y_n = \sum_m a_{n-m} \int_\Gamma e^{i2\pi\lambda m} M_\xi(d\lambda) = \int_\Gamma e^{i2\pi\lambda n} a(e^{-i2\ \lambda}) M_\xi(d\lambda)$$

so that

$$r_y(n) = (y_n, y_o) = \int_\Gamma e^{i2\pi\lambda n} |a(e^{-i2\pi\lambda})|^2 d\lambda$$

Theorem 1. $L(a * \xi) \subset L(\xi)$ and $L(a * \xi) = L(\xi) \iff a(e^{-i2\pi\lambda}) \neq 0$ a.e.

Proof. Use the following Tauberian Theorem for Fourier series:
Let $a \in \ell^2(Z)$.

$$\text{c.}\ell\text{.m}\{\alpha * a,\ \alpha \in \ell^1(Z)\} = \ell^2(Z)$$

$$\iff a(e^{-i2\pi\lambda}) \neq 0, \text{ a.e.}$$

Definition 1. Let x_n be a stationary sequence. If we have

$$x \sim a * \xi \qquad a \in \ell^2(Z), \qquad \xi = \text{white noise},$$

$a * \xi$ is called a moving average representation of x.

Corollary 1. If x has a moving average representation $a * \xi$ with $a(e^{-i2\pi\lambda}) \neq 0$, a.e., then we can find a white noise $\eta_n \in L(x)$ such that

$$x = a * \eta$$

Proof. If $a(e^{-i2\pi\lambda}) \neq 0$, then

$$L(y) = L(\xi), \qquad \text{for } y = a * \xi$$

Since $x \sim y$, we have an isomorphism $V : L(x) \to L(y)$ such that $Vx_n = y_n$. Set $\eta_n = V^{-1}\xi_n$. Then $y = a * \xi$ goes over into $x = a * \eta$ by V^{-1}.

Theorem 2. In order for x_n to have a moving average representation, it is necessary and sufficient that the Hinčin measure of x is absolutely continuous.

Proof. (i) Assume that $x \sim y = a * \xi$. Then

$$r_x(n) = r_y(n) = \int_\Gamma e^{i2\pi\lambda n}|a(e^{-i2\pi\lambda})|^2\, d\lambda$$

(ii) Assume that

$$(x_p\ x_2) = \int_\Gamma e^{i2\pi(p-2)\lambda} f(\lambda)\, d\lambda, \qquad f \in L^1(\Gamma)$$

Then $\sqrt{f(\lambda)} \in L^2(\Gamma)$. Consider the Fourier expansion of $\sqrt{f(\lambda)}$

$$\sqrt{f(\lambda)} = \sum_m a_m\, e^{-i2\pi\lambda m} \equiv a(e^{-i2\pi\lambda})$$

with $a = (a_m) \in \ell^2(Z)$. Let ξ_n be a white noise. Then

$$y = a * \xi \sim x$$

because

$$r_y(n) = \int e^{i2\pi\lambda n}|a(e^{-i2\pi\lambda})|^2\, d\lambda$$

$$= \int e^{i2\pi\lambda n} f(\lambda)\, d\lambda = r_x(n).$$

<u>Remark 1.</u> The representation is not unique. We can use the expansion of any function $\sqrt{f(\lambda)}\, e^{i\phi(\lambda)}$, $\phi(\lambda)$ being real and measurable.

<u>Remark 2.</u> $f(\lambda) \neq 0$ (a.e.) $\iff a(e^{-2\pi\lambda}) \neq 0$ (a.e.)

$$\iff L(a * \xi) = L(\xi)$$

$$\implies x = a * \eta, \quad \forall_n \; \eta_n \in L(x), \quad (\eta_n) = \text{white noise.}$$

<u>Definition 2.</u> A moving average representation $x \sim a * \xi$ is called <u>backward</u> if $a_n = 0$ for $n < 0$ and <u>forward</u> if $a_n = 0$ for $n > 0$.

<u>Corollary 2.</u> If $x \sim a * \xi$ is backward, then

$$L_n(a * \xi) \subset L_n(\xi)$$

<u>Theorem 3.</u> The following three conditions are equivalent for a non-trivial stationary sequence x_n.

(i) x_n has a backward moving average representation,

(ii) x_n is non-deterministic,

(iii) the Hinčin measure of x_n is absolutely continuous with the density $f(\lambda)$ satisfying

$$\int_\Gamma \log f(\lambda)\, d\lambda > -\infty .$$

<u>**Proof.**</u> (i) $\Longrightarrow$ (ii)

$$L_n(a * \xi) \subset L_n(\xi)$$

and so

$$L_{-\infty}(a * \xi) \subset L_{-\infty}(\xi) = 0$$

Therefore $a * \xi$ is non-deterministic and so is x.

(ii) $\rightarrow$ (i)

$$L_{-1}(x) \underset{\neq}{\subset} L_0(x)$$

$$\left[\begin{array}{l} \because L_{-1}(x) = L_0(x) \Longrightarrow L_{n-1}(x) = U^n L_{-1}(x) = U^n L_0(x) \\ \quad = L_n(x) \Longrightarrow L_\infty(x) = L_{-\infty}(x) = 0 \\ \quad \Longrightarrow x_n \text{ is trivial (contrary to the assumption)} \end{array}\right]$$

$$\therefore\ x_0 \notin L_{-1}(x),\ \xi_0' = x_0 - P_{L_{-1}(x)} \cdot x_0 \neq 0$$

Set $\xi_0 = \dfrac{\xi_0'}{\|\xi_0'\|}$ and $\xi_n = U^n \xi_0$. Then

$$L_0(x) = L_{-1}(x) \oplus \{\xi_0\},$$

and applying U^n to both sides, we get

$$L_n(x) = L_{n-1}(x) \oplus (\xi_n)$$

$$L_0(x) = L_{-\infty}(x) + \text{c.l.m.}\ \{\xi_0, \xi_{-1}, \xi_{-2}, \dots\}$$

$$= \text{c.l.m.}\ \{\xi_0, \xi_{-1}, \xi_{-2}, \dots\}$$

$$x_0 = \sum_m a_m \xi_{-m}$$

where

$$a_m = \begin{cases} (x_0, \xi_{-m}) & m \geq 0 \\ 0 & m < 0 \end{cases}$$

Then

$$x_n = U^n x_0 = \sum_m a_m U^n \xi_{-m} = \sum a_m \xi_{n-m}$$

$$= \sum a_{n-m} \xi_m$$

To prove (i) <==> (iii) we shall use the following known

<u>Theorem</u>: In order for a non-negative $f \in L^1(\Gamma)$ to be expressed as

$$f(\lambda) = \left| \sum_{n \geq 0} a_n e^{-i2\pi\lambda n} \right|^2$$

with $0 < \sum |a_n|^2 < \infty$, $a_n = 0$ for $n < 0$, it is necessary and sufficient that

$$\int_\Gamma \log f(\lambda)\, d\lambda > -\infty$$

(This theorem will be included in the results in the Section 4)

(i) → (iii) If $x \sim a * \xi$ (backward), then

$$r_x(n) = r_{a*\xi}(n) = \int_\Gamma e^{i2\pi\lambda n} \Big| \sum_{n \geq 0} a_n e^{-i2\pi\lambda n} \Big|^2 d\lambda$$

and so the Hinčin measure of x_n is absolutely continuous with the density

$$f(\lambda) = \Big| \sum_{n \geq 0} a_n e^{-i2\pi\lambda n} \Big|^2 .$$

Now use the above theorem.

(iii) → (i) If $\int_\Gamma \log f(\lambda)\, d\lambda > -\infty$, then the above theorem shows that $f(\lambda)$ can be expressed as

$$f(\lambda) = \Big|\sum a_n e^{-i2\pi\lambda n}\Big|^2 \qquad (a_n = 0 \text{ for } n < 0),$$

and so $x \sim a * \xi$ (ξ = white noise) because

$$r_{a*\xi}(n) = \int e^{i2\pi\lambda n} f(\lambda)\, d\lambda = r_x(n).$$

3. Generalized Poisson Formula.

Theorem of Fatou. Let $u(z)$ be a harmonic function and suppose that

(1) $$\sup_{0 \leq r < 1} \int_{-\pi}^{\pi} |u(re^{i\theta})| d\theta < \infty.$$

Then u can be expressed by generalized Poisson formula,

(2) $$u(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(\theta - \varphi)\, \mu(d\varphi)$$

$$P_r(\mu) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2} \qquad \text{(Poisson kernel)}$$

$$\mu = w^* - \lim_{r \uparrow 1} \mu_r, \qquad \mu_r(d\theta) = u(re^{i\theta})\, d\theta$$

The boundary value $u(e^{i\theta})$ of $u(z)$ exists a.e., and equals the density of the absolutely continuous part of μ_θ a.e. (To be more precise, where $\mu'(\theta) = \lim_{\epsilon \downarrow 0} \mu(\theta - \epsilon, \mu + \epsilon)/2\epsilon$ exists, the non-tangential limit of $u(z)$ as $z \to e^{i\theta}$ exists and equals $\mu'(\theta)$.)

Note. 1. In the general Poisson formula the a.e. existing boundary value $u(e^{i\theta})$ determines only the absolutely continuous part of $\mu(d\theta)$ and so does not always determine the behavior of $u(z)$ in $|z| < 1$. To determine $u(z)$, $|z| < 1$, completely, we should know, besides the boundary value $u(e^{i\theta})$, the singular part $s(d\theta)$ of $\mu(d\theta)$, i.e., the weak* limit of $s_r(d\theta) = [u(re^{i\theta}) - u(e^{i\theta})]\, d\theta$ as $r \uparrow 1$.

Note 2 If u is expressed as $\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta-\varphi)\,\mu(d\varphi)$ with bounded variation signed measure μ of [illegible], then u satisfies (1)

Function Class H_1 . (Hardy class of order 1)

Definition 1. A function analytic in $|z| < 1$ is said to belong to H_1, iff

$$(3) \qquad \sup_{0 \le r < 1} \int_{-\pi}^{\pi} |g(re^{i\theta})| d\theta < \infty$$

Applying Faton's theorem to the real and imaginary parts, we have

Generalized Poisson formula for H_1-functions. Any $g \in H_1$ has boundary values $g(e^{i\theta})$ a.e. and

$$(4) \qquad g(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} Pr(\theta - \varphi) \, [g(e^{i\varphi}) \, d\varphi + s(d\varphi)]$$

where s is a complex-valued singular measure of bounded variation defined by

$$(5) \qquad s = w^* - \lim_{r \uparrow 1} s_r, \qquad s_r(d\theta) = (g(re^{i\theta}) - g(e^{i\theta})) \, d\theta$$

<u>Fundtion Class</u> H_2 <u>(Hardy class of order 2)</u>.

<u>Definition 2</u>. A function analytic in $|z| < 1$ is said to belong to H_1 iff

(6) $$\sup_{0 \leq r < 1} \int_{-\pi}^{\pi} |g(re^{i\theta})|^2 \, d\theta < \infty$$

Consider the power series expansion of $g(z)$:

(7) $$g(z) = \sum_{n \geq 0} a_n z^n$$

Then (6) is equivalent to

(8) $$\sum_{n \geq 0} |a_n|^2 < \infty$$

Since $g \in H_2 \subset H_1$, $g(e^{i\theta}) = \text{a.e.} \lim_{r \uparrow 1} g(re^{i\theta})$ exists. It follows from (7) that

(9) $$g(e^{i\theta}) = \underset{n \to \infty}{\text{l.i.m.}} \sum_{n=0}^{N} a_n e^{in\theta} = \underset{r \uparrow 1}{\text{l.i.m.}}\, g(re^{i\theta})$$

from which we have

(10) $$\frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{i\theta}) e^{-in\theta} \, d\theta = \begin{cases} a_n & n \geq 0 \\ 0 & n < 0 \end{cases}$$

$$(11) \qquad g(re^{i\theta}) = \sum_{n \geq 0} a_n r^n e^{in\theta}$$

$$= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} \int_{-\pi}^{\pi} g(e^{i\varphi}) e^{-in\varphi} d\varphi$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{i\varphi}) \sum_{n=-\infty}^{\infty} r^{|n|} e^{in(\theta-\varphi)} d\varphi$$

and so

$$(12) \qquad g(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(\theta-\varphi) g(e^{i\varphi}) d\varphi$$

This is Poisson formula for H_2 functions.

Relation Between H_2 and L^2_+.

Let L^2_+ be the subspace of $L^2(-\pi, \pi)$ generated by $e^{in\theta}$, $n \geq 0$. Then

$$T : g(z) \equiv \sum_{n \geq 0} a_n z^n \to (Tg)(e^{i\theta}) = \sum_{n \geq 0} a_n e^{in\theta} ,$$

$$\left(\sum_{n \geq 0} |a_n|^2 < \infty \right)$$

determines a one-to-one mapping from H_2 onto L^2_+ and we have

(i) $(Tg)(e^{i\theta})$ is the a.e. boundary value function of g and so can be written as $g(e^{i\theta})$.

(ii) $g_r(e^{i\theta}) = g(re^{i\theta}) \in L^2_+$ and $\|g_r\| \uparrow \|Tg\|$ as $r \uparrow 1$. (Notice that $\|g_r\| = \sum_n |a_n|^2 r^{2n}$).

(iii) <u>Poisson Formula</u>.

$$g(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (Tg)(e^{i\varphi}) \, P_r(\theta - \varphi) \, d\varphi, \qquad z = re^{i\theta} \qquad (0 \le r < 1)$$

(iv)
$$|g(z)| \le \|Tg\| \frac{1 + |z|}{1 - |z|}$$

(v)
$$g(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Re[(Tg)(e^{i\varphi})] \cdot \frac{e^{i\varphi} + z}{e^{i\varphi} - z} \, d\varphi + i\beta$$

We have already proved (i), (ii) and (iii).

To prove (iv),

$$|g(re^{i\theta})| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |Tg(e^{i\varphi})| \, \frac{1 + r}{1 - r} \, d\varphi$$

$$\le \|Tg\| \frac{1 + r}{1 - r} \qquad \text{(by the Schwarz inequality)}$$

To prove (v), let $h(z)$ be the integral in the right side.

Then

$$\Re h(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Re[(g)(e^{i\varphi})]\, P_r(\theta - \varphi)\, d\varphi, \qquad z = re^{i\theta}$$

$$= \Re g(z)$$

which implies that $h(z) - g(z)$ is a pure imaginary constant, because $h(z)$ and $g(z)$ are analytic.

4. Factorisation theorem for functions in H_2.

Using the classical Poisson formula we can easily prove

Theorem 1 (Elementary factorisation theorem). Let $g(z)$ be a function analytic on the unit closed disk ~~(with no zero points on the boundary)~~ If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the zero points* of $g(z)$ in $|z| < 1$, then

$$(1) \quad g(z) = \alpha \cdot \prod_{k=1}^{n} \frac{z - \alpha_k}{1 - \bar{\alpha}_k z} \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\varphi} + z}{e^{i\varphi} - z} \log\left|g(e^{i\varphi})\right| d\varphi \right\},$$

where α is a constant of modulus 1, and

$$(2) \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|g(e^{i\varphi})\right| d\varphi = \log|a| + \sum_{\substack{k=1 \\ \alpha_k \neq 0}}^{n} \log \frac{1}{|\alpha_k|} ,$$

where a is the first non-vanishing coefficient of the power series expansion of $g(z)$.

The purpose of this section is to extend this theorem to functions in H_2.

Lemma 1. Let $g(z)$ be analytic in $|z| < 1$ and assume ~~(that $g(z) \neq 0$ on $|z|=r$)~~ and that $g(0) \neq 0$. Then if $\alpha_1, \ldots, \alpha_n$ are the zero points of g in $|z| < r$,

$$(3) \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|g(re^{i\varphi})\right| d\varphi = \log|g(0)| + \sum_{\substack{k=1 \\ \alpha_k \neq 0}}^{n} \log\frac{r}{|\alpha_k|} \geq \log|g(0)|.$$

---

* Unless otherwise stated, we repeat every multiple root by its multiplicity.

Lemma 2. If $g \in H_2$ and if $g(0) \neq 0$, then

(4) $$\int_{-\pi}^{\pi} \left|\log|g(re^{i\varphi})|\right| d\varphi \leq \int_{-\pi}^{\pi} |g(e^{i\varphi})|^2 \, d\varphi \; - \; 2\pi \log|g(0)| \qquad \text{for } 0 \leq r \leq 1 .$$

Lemma 3. If $g \in H_2$, then

(5) $$|g(re^{i\theta})|^2 \quad \leq \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(e^{i\varphi})|^2 P_r(\theta-\varphi)\, d\varphi$$

(5') $$\int_{-\pi}^{\pi} |g(\rho e^{i\varphi})|^2 P_{r/\rho}(\theta-\varphi)\, d\varphi \leq \int_{-\pi}^{\pi} |g(e^{i\varphi})|^2 P_r(\theta-\varphi)\, d\varphi \quad \text{for } 0 \leq r \leq \rho \leq 1$$

(6) $$\log|g(re^{i\theta})| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \log|g(e^{i\varphi})| P_r(\theta - \varphi)\, d\varphi .$$

Lemma 4. If $g \in H_2$ and if $\alpha_1, \alpha_2, \ldots$ be the zero points of g in $|z| < 1$, then

(7) $$\sum_{n \geq 0} (1 - |\alpha_n|) < \infty .$$

Lemma 5. If $|\alpha_n| < 1$ and if $\sum_n (1 - |\alpha_n|) < \infty$,

then the infinite product

(8) $$\prod_n \frac{-\bar{\alpha}_n}{|\alpha_n|} \frac{z - \alpha_n}{1 - \bar{\alpha}_n z} \qquad \left(\text{convention: } \frac{\bar{\alpha}_n}{|\alpha_n|} = 1 \text{ if } \alpha_n = 0\right)$$

defines an analytic function $B(z) \in H_2$ with the zero points $\{\alpha_n\}$ and satifies

(9) $|B(z)| \leq 1$ in $|z| < 1$,

(10) $|B(e^{i\varphi})| = 1$ a.e.

<u>Definition</u> 1. B(z) in Lemma 5 is called the <u>Blaschke product</u> with zero points $\{\alpha_n\}$.

Given $g \in H_2$, let $\{\alpha_n\}$ be its zero points in $|z| < 1$. Then we can define the Blaschke product with the roots $\{\alpha_n\}$ by Lemma 3 and Lemma 4. Then we have

Lemma 6. $h = g/B \in H_2$ and has no zero points in $|z|<1$ and $|h(e^{i\varphi})| = |g(e^{i\varphi})|$ a.e.

If $h \in H_2$ and if h has no roots in $|z|<1$, then

Lemma 7.

$$h(z) = \alpha \cdot \exp\left\{ \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{i\varphi}+z}{e^{i\varphi}-z} \left[\log|h(e^{i\varphi})|\,d\varphi - \sigma(d\varphi)\right]\right\}, \tag{11}$$

where α is a constant of modulus 1 and σ is a bounded non-negative singular measure defined by $d\sigma = \underset{r\uparrow 1}{w^*\text{-lim}} \left(\log|h(e^{i\varphi})| - \log|h(re^{i\varphi})|\right) d\varphi$.

Definition 2. A function $f(z)$ is called an <u>outer function</u> with the (real) <u>generating density</u> $\omega(e^{i\varphi})$, if it is expressed as

$$f(z) = \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{i\varphi}+z}{e^{i\varphi}-z}\,\omega(e^{i\varphi})\,d\varphi\right\}, \quad \omega(e^{i\varphi}) \in L^1(-\pi, \pi). \tag{12}$$

Notice that an outer function does <u>not always</u> belong to H_2.

Definition 3. A function $f(z)$ is called a <u>singular function</u> with <u>generating singular measure</u> $d\sigma$, if it is expressed as

$$f(z) = \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{i\varphi}+z}{e^{i\varphi}-z}(-\sigma)(d\varphi)\right\}, \tag{13}$$

where σ is a bounded non-negative singular measure on $(-\pi, \pi)$. Notice that a singular function belongs to H_2.

Definition 4. A function f analytic in $|z|<1$ is called an <u>inner function</u> if $|f(z)| \le 1$ in $|z|<1$ (from which it follows that $f \in H_2$) and if $|f(e^{i\varphi})| = 1$ a.e.

Corollary 1. Both Blaschke products and singular functions are inner.

Corollary 2. If $f \in H_2$ and if $|f(e^{i\varphi})| = 1$ a.e. then f is inner (use Lemma 3)

<u>Theorem</u> 2 (<u>Factorization theorem for functions in H_2</u>).

Any function $g \in H_2$ can be factorised as

(14) $$g = \alpha \cdot B \cdot g_s \cdot s$$

- α : a constant of modulus 1
- B : Blaschke product
- g_s : an outer function in H_2
- s : a singular function

Such factorization is unique; if there is any such factorization of g in the form (14), then

B is the Blaschke product with the same zero points $\{\alpha_k\}$ as g,

g_s is the outer function with generating density $\log|g(e^{i\varphi})|$,

and

s is the singular function with generating singular measure $\sigma(d\varphi)$ defined by

(15) $$d\sigma(\varphi) = \underset{r \uparrow 1}{w^*\text{-lim}} \left[\log|g(e^{i\varphi})| - \log|g(re^{i\varphi})|\right] d\varphi .$$

We have also

(16) $$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|g(e^{i\varphi})|\, d\varphi = \log|a| + \sum_{\alpha_k \neq 0} \log\frac{1}{|\alpha_k|} + \int_{-\pi}^{\pi} d\sigma(\varphi),$$

where a is the first non-vanishing coefficient of the power series expansion of g.

<u>Corollary</u> 2. $g \in H_2$ is a Blaschke product iff g is an inner function with $\log|g(re^{i\varphi})|$ uniformly (in $0 \le r < 1$) integrable on $[-\pi, \pi)$.

<u>Corollary</u> 3. $g \in H_2$ is an outer function iff $\log|g(re^{i\varphi})|$ is uniformly (in $0 \le r < 1$) integrable on $[-\pi, \pi)$ and $g(z)$ has no zero points in $|z|<1$.

<u>Corollary</u> 4. $g \in H_2$ is a singular function iff g is an inner function which has no zero points in $|z| < 1$.

5. <u>Determining all backward moving average representations of any non-deterministic stationary sequence</u>.

In this section we shall determine all backward representations <u>of the form</u>

(1) $$x = a * \xi, \qquad L(x) = L(\xi), \qquad U\xi_n = \xi_{n+1}$$

for any (non-trivial) non-deterministic stationary sequence x_n.

If x_n has any backward representation, then x_n must be non-deterministic and x_n has a unique backward representation of the form (1) with the same coefficients. This will justify that we consider only non-deterministic stationary sequences and only backward representations of the form (1).

Since x_n is non-deterministic, its Hinčin measure $dF(\lambda)$ is absolutely continuous with density function $f(\lambda)$ satisfying

(2) $$\int_\Gamma \log f(\lambda)\, d\lambda > -\infty$$

or equivalently (because of the integrability of f)

(2') $$\int_\Gamma |\log f(\lambda)|\, d\lambda < \infty$$

In Section 2 we saw that $x = a * \xi$ is a backward representation, iff

$$(3) \quad \begin{cases} a_n = 0 & n < 0 \\ \Sigma\, |a_n|^2 < \infty & \\ |\Sigma\, a_n\, e^{-i2\pi n\lambda}|^2 = f(\lambda) & \text{a.e.} \end{cases}$$

Using $\{a_n\}$, we shall introduce an analytic function

$$(4) \qquad a(z) = \sum_n a_n z^n \equiv \sum_{n \geq 0} a_n z^n$$

Equation (3) can be written in terms of $a(z)$ as

$$(3') \quad \begin{cases} a \in H_2 \\ |a\,(e^{-i2\pi\lambda})| = |\lim_{r \uparrow 1} a\,(r\, e^{-i2\pi\lambda})| = \sqrt{f(\lambda)} \qquad \text{a.e.} \end{cases}$$

By the factorization theorem in Section 4, a can be expressed as

$$(5) \qquad a(z) = \alpha\, B(z) \cdot s(z) \cdot g_s(z)$$

where B is the Blaschke product, s is a singular function and g_s is an outer function with generating density $\log \sqrt{f(\lambda)}$ $(\equiv \log|a(e^{-i2\pi\lambda})|)$

Write g_i for $\alpha \cdot B(z) \cdot s(z)$. Then $g_i(z)$ is an inner function which will be called the <u>inner part</u> of a, while g_s will be called the (standard) <u>outer part</u> of a.

<u>Theorem 1</u>. If x is a white noise and if $x = a * \xi$ is a backward representation, then $a(z)$ is an inner function and ξ can be expressed by x as

$$(6) \qquad \xi = a^* * x, \qquad a^*_n = \bar{a}_{-n}$$

i.e.,

$$(7) \qquad \xi_n = \sum_{k \geq 0} \bar{a}_k x_{n+k} \qquad \text{(forward representation)}$$

<u>Proof</u>. Let $\xi_n = \int_\Gamma e^{+i2\pi\lambda n} dM(\lambda)$ be the Kolmogorov-Cramer representation of ξ_n. Since ξ_n is a white noise,

$$(M(\Lambda_1), M(\Lambda_2)) = \int_{\Lambda_1 \cap \Lambda_2} d\lambda$$

Then

$$x_n = \int_\Gamma e^{i2\pi\lambda n} a(e^{-i2\pi\lambda}) \, dM(\lambda)$$

Since x_n is also a white noise, we have

$$|a(e^{-i2\pi\lambda})|^2 \, d\lambda = d\lambda$$

i.e., $|a(e^{-i2\pi\lambda})| = 1$ a.e. But $a(z) \in H_2$, and so $a(z)$ is inner. Now observe

$$\sum_{k \geq 0} \bar{a}_k \, x_{n+k} = \int_\Gamma e^{i2\pi\lambda n} \sum_k \bar{a}_k \, e^{+i2\pi\lambda k} \, dM(\lambda)$$

$$= \int_\Gamma e^{i2\pi\lambda n} \, \overline{a(e^{-i2\pi\lambda})} \, a(e^{-i2\pi\lambda}) \, dM(\lambda)$$

$$= \int_\Gamma e^{i2\pi\lambda n} |a(e^{-i2\pi\lambda})|^2 \, dM(\lambda)$$

$$= \int_\Gamma e^{i2\pi\lambda n} \, dM(\lambda) = \xi_n ,$$

which proves (7).

If $x = a * \xi$ is a backward representation, then we have

(8) $$L_n(x) \subset L_n(\xi)$$

and

(9) $$L(x) = L(\xi)$$

Equation (8) is clear because $x = a * \xi$ is backward. To prove (9), notice that

$$\left|\sum_{n \geq 0} a_n e^{-i2\pi\lambda n}\right|^2 = f(\lambda) \neq \text{a.e.}$$

by virtue of (2), so that $(b * a : b \in \ell^1(Z))$ is dense in $\ell^2(Z)$ by the Tauberian theorem.

Keeping (8) and (9) in mind, we shall introduce

Definition 1. A backward representation $x = a * \xi$ is called canonical if

(10) $$L_n(x) = L_n(\xi) \quad \text{for every } n \quad (\leftrightarrows L_0(x) = L_0(\xi))$$

There exists at least one canonical representation of any given non-deterministic stationary sequence x_n. Consider the innovation

$$\xi_n^s = \frac{x_n - P_{L_{n-1}(x)} \cdot x_n}{\|x_n - P_{L_{n+1}(x)} \cdot x_n\|}$$

and set

$$a_n^s = (x_o, \xi_{-n}^s).$$

Then it is easy to see that $x = a^s * \xi^s$ is a canonical backward representation.

Definition 2. $x = a^s * \xi^s$ is called a standard backward representation.

Theorem 2.

$$\text{(11)} \qquad a^s(Z) = \exp\left\{\int_\Gamma \frac{e^{-i2\pi\lambda} + Z}{e^{-i2\pi\lambda} + Z} \log \sqrt{f(\lambda)}\, d\lambda\right\}$$

Proof. Let

$$\text{(12)} \qquad a^s(z) = a^i(z)\, a^o(z)$$

be the factorization of a^s into its inner and outer parts. Since $|a^s(e^{-i2\pi\lambda})|^2 = f(\lambda)$ a.e., $a^o(z$ must equal the right side of (11).

To prove (11), it is enough to prove $a^i(z) \equiv 1$. Set

(13) $$\xi = a^i * \xi^s$$

Using the Kolmogorov-Cramer representation of ξ^s, ξ is also a white noise and

(14) $$x = a^o * \xi$$

By Theorem 1 we can derive from (13)

(15) $$\xi_n^s = \sum_{k \geq 0} \overline{a_k^i}\, \xi_{n+k}$$

Since ξ^s is the innovation of x, we have

$$\xi_n^s \in L_n(x) \subset L_n(\xi) \quad \text{(by (14))}$$

i.e., $\xi_n^s \perp \xi_{n+k}$ $(k > 0)$, which, combined with (15), implies

(16) $$\overline{a_k^i} = 0 \qquad (k > 0) \quad \text{i.e., } a^i(z) = a_o^i$$

It is clear that $|a_o^i| = 1$ ($\because$ $\xi_o = a_o^i\, \xi^s$). Thus $a^s(z) = a_o^i \cdot a^o(z)$. But it is clear that

$$a^o(0) > 0, \qquad a^s(0) > 0$$

and so $a^i_o = 1$.

<u>Theorem 3</u>. If $x = a * \xi$ is any backward representation, and if $a(z) = a^i(z)\, a^o(z)$ is the factorization of $a(z)$ into the inner and outer parts of $a(z)$ ($a^o(z) \equiv a^s(z)$ by Theorem 2), then

$$(17) \qquad \xi_n = \sum_{k \geq 0} \overline{a^i_k}\, \xi^s_{n+k}$$

<u>Proof</u>. Since a^i is inner, $\eta = a^i * \xi$ is a white noise and $x = a^s * \eta$, so that $\eta = \xi^s$, i.e., $\xi^s = a^i * \xi$, which implies (17) by Theorem 1.

<u>Corollary 2</u>. $x = a * \xi$ is a canonical backward representation, iff $a(z) = \alpha\, a^s(z)$, α being a constant of modulus 1.

<u>Remark</u>. Let $x = a * \xi$ be any backward representation. Then $L(x) = L(\xi)$ and $L_n(x) \subset L_n(\xi)$. Therefore $\xi_n \in L(\xi) \ominus L_{n-1}(\xi) \subset L(x) \ominus L_{n-1}(x)$ $= L(\xi^s_n, \xi^s_{n+1}, \dots)$. Therefore $\xi_n = \sum_{k \geq 0} b_k\, \xi^s_{n+k}$. ($b_k$ does not depend on n because of (1).) Theorem 3 shows that we can express b_k as $\overline{a^i_k}$ using the inner part $a^i(z)$ of $a(z)$.

In the canonical backward representation $x = a * \xi$ we do not need the future information of x_n to construct ξ_n.

6. Prediction. Let x_n be any stationary sequence with the Hinčin measure dF and the Kolmogorov-Cramer orthogonal random measure $dM(\lambda)$. Let $R^1 = A \cup S$, $A \cap S = \emptyset$ be the decomposition such that $F_a(\Lambda) = F(\Lambda \cap A)$ is absolutely continuous and $F_s(\Lambda) = F(\Lambda \cap S)$ is singular. It is clear that $F'(\lambda) = F_a'(\lambda)$ $(= f(\lambda))$, $F_s'(\lambda) = 0$ a.e. Set

(1) $$M_a(\Lambda) = M(\Lambda \cap A), \qquad M_s(\Lambda) = M(\Lambda \cap S)$$

and

(2) $$x_n^a = \int e^{-i2\pi\lambda n}\, dM_a(\lambda), \qquad x_n^s = \int e^{-i2\pi\lambda n}\, dM_s(\lambda)$$

Then

(3) $$\begin{aligned} x_n &= x_n^a + x_n^s \\ L(x^a) &\perp L(x^s) \\ L(x) &= L(x^a) \oplus L(x^s) \end{aligned}$$

Recall that x_n can be decomposed as

$$x_n = x_n^i + x_n^d, \qquad x_n^i : \text{non-deterministic}, \; x_n^d : \text{deterministic}$$

(4) $L(x^i) \perp L(x^d)$

$$L(x) = L(x^i) \oplus L(x^d)$$

Let us examine the relation between two decompositions (3) and (4).

Theorem 1. If $\int \log f(\lambda)\, d\lambda = -\infty$, then x_n^a, x_n^s and x_n are all deterministic and

$$x_n^i = 0, \qquad x_n^d = x_n = x_n^a + x_n^s$$

If $\int \log f(\lambda)\, d\lambda > -\infty$, then

$$x_n^i = x_n^a, \qquad x_n^d = x_n^s$$

Definition 1. $x_{n,m} \equiv P_{L_m} \cdot x_n$ is called the predictor of x_n with the information up to time m, and $e^2_{n,m} \equiv \|x_n - P_{L_m(x)} \cdot x_n\|^2$ is called the mean square error of this predictor $e^2_{n,m}$ depends only on $n-m$ and so we can write e^2_{n-m} for $e^2_{n,m}$.

If $\int f(\lambda)\, d\lambda = -\infty$, then x_n is deterministic and so $x_{n,m} = x_n$ for every m. Therefore there is no problem of prediction in this case.

__Theorem 2__. If $\int f(\lambda)\, d\lambda > -\infty$, then

$$e_m^2 = \sum_{k=0}^{m-1} |a_k|^2 ,$$

where a_k, $k = 0, 1, 2, \ldots$ are determined by

$$\sum_{k \geq 0} a_k z^k = \exp \left\{ \frac{1}{2} \int_{\Gamma} \frac{e^{-i2\pi\lambda} + z}{e^{-i2\pi\lambda} - z} \log f(\lambda)\, d\lambda \right\}$$

The predictor $x_{n,m}$ is obtained as follows:

$$\xi_n = \frac{x_n - P_{L_{n-1}(x)} \cdot x_n}{\| x_n - P_{L_{n-1}(x)} \cdot x_n \|} \quad (\text{innovation})$$

$$a_k = (x_0, \xi_{-k}) = (x_n, \xi_{n-k})$$

$$x_{n,m} = \sum_{k \geq n-m} a_k \xi_{n-k} + x_n^s$$

x_n^s = singular part of x_n (see theorem 1)

7. Concrete expression of innovation and predictor.

Let x_n be non-deterministic. Then the Hinčin measure is absolutely continuous with the density $f(\lambda)$ satisfying

$$(1) \qquad \int \log f(\lambda)\, d\lambda > -\infty$$

We shall introduce the following functions analytic in $|z| < 1$

$$(2) \qquad a(z) \equiv \Sigma\, a_k\, z^k = \exp\left\{ \frac{1}{2} \int_\Gamma \frac{e^{-i2\pi\lambda} + z}{e^{-i2\pi\lambda} - z} \log f(\lambda)\, d\lambda \right\}$$

$$(3) \qquad b(z) \equiv \Sigma\, b_k\, z^k = a(z)^{-1}$$

$$(4) \qquad A_\ell(z) = \sum_{k \geq \ell} a_k\, z^k$$

$$(5) \qquad C_\ell(z) = \sum_{k \geq \ell} c_{\ell,k}\, z^k = A_\ell(z)\, b(z)$$

It is clear that $a(z)$ is a standard outer function and belongs to H_2. Therefore $\Sigma\, |a_n|^2 < \infty$, so that $A_\ell(z) \in H_2$. b does not necessarily belong to H_2, but the a.e. boundary value function $b(e^{-i2\pi\lambda})$ exists and equals $a(e^{-i2\pi\lambda})^{-1}$ a.e. Similarly for $C_\ell(z)$.

Now observe the Kolmogorov-Cramer representation for x_n and ξ_n

(6) $$x_n = \int e^{i2\pi\lambda n}\, dM_x(\lambda)$$

(7) $$\xi_n = \int e^{i2\pi\lambda n}\, dM_\xi(\lambda)$$

Then it is clear that

(8) $$dM_x(\lambda) = a(e^{-i2\pi\lambda})\, dM_\xi(\lambda)$$

(9) $$dM_\xi(\lambda) = b(e^{-i2\pi\lambda})\, dM_x(\lambda)$$

Thus we have

(10) $$\xi_n = \int e^{i2\pi\lambda n}\, b(e^{-i2\pi\lambda})\, dM(\lambda)$$

(11) $$\begin{aligned} x_{n,n-\ell} &= \int e^{i2\pi\lambda n}\, A_\ell\,(e^{-i2\pi\lambda})\, dM_\xi(\lambda) \\ &= \int e^{i2\pi\lambda n}\, A_\ell\,(e^{-i2\pi\lambda})\, b(e^{-i2\pi\lambda})\, dM(\lambda) \\ &= \int e^{i2\pi\lambda n}\, C_\ell\,(e^{-i2\pi\lambda})\, dM(\lambda) \end{aligned}$$

from which we have <u>formal expansions</u>

(10') $$\xi_n = \sum_{k \geq 0} b_k\, x_{n-k}$$

(11') $$x_{n,n-\ell} = \sum_{k \geq \ell} c_{\ell,k}\, x_{n-k}$$

Theorem 1. If $f(\lambda)$ is essentially bounded (i.e., $\exists M < \infty$ such that $f(\lambda) < M$ a.e.), and if $\int f(\lambda)^{-1} d\lambda < \infty$, then the formal expansions (10'), (11') converge and the equalities are true.

Proof. It follows from (2) and (3) that

$$b(z) = \exp\{-\frac{1}{2}\int \cdots \log f(\lambda)\, d\lambda\}$$

$$= \exp\{\frac{1}{2}\int \cdots \log f(\lambda)^{-1}\, d\lambda\}$$

and $\int f(\lambda)^{-1} d\lambda < \infty$ implies that $b \in H_2$, so that $\Sigma |b_p|^2 < \infty$ and
$b(e^{-i2\pi\lambda}) = \sum_{k \geq 0} b_k e^{-i2\pi\lambda k}$ (convergence in $L^2(\Gamma\, d\lambda)$)

$$\|\xi_n - \sum_{k=0}^{p} b_k x_{n-k}\|^2$$

$$= \| \int e^{i2\pi\lambda n} \sum_{k > p} b_k e^{-i2\pi\lambda k}\, dM(\lambda)\|^2$$

$$= \int \Big|\sum_{k \geq p} \cdots\Big|^2 f(\lambda)\, d\lambda$$

$$\leq M \int \Big|\sum_{k > p} \cdots\Big|^2 d\lambda = M \sum_{k > p} |b_k|^2 \to 0 \qquad (p \to \infty)$$

Since we have

$$|b(e^{-i2\pi\lambda})|^2 = f(\lambda)^{-1}$$

and

$$C_\ell(z) = 1 - b(z) \sum_{k < \ell} a_k z^k ,$$

we have

$$\int |C_\ell(re^{-i2\pi\lambda})|^2 \, d\lambda \le 2 \int \left[1 + f(\lambda)^{-1} \left(\sum_{k < \ell} |a_k| \right)^2 \right] d\lambda$$

and therefore $C_\ell(z) \in H_2$. By the same argument as for ξ_n we have

$$\| x_{n,n-\ell} - \sum_{k=\ell}^{p} c_{\ell,k} x_{n-k} \|^2$$

$$\le M \sum_{k > p} |c_{\ell,k}|^2 \to 0 \qquad (p \to \infty)$$

In some cases the formal expansion does not converge but its Césaro sum converges and equals the right value, for example

<u>Theorem 2</u>. If $f(\lambda) = |1 - e^{-i2\pi\lambda}|^2 = 2(1 - \cos 2\pi\lambda)$, then

$$a(z) = 1 - z \qquad \text{(standard outer function)}$$

$$x_n = \xi_n - \xi_{n-1} \qquad \text{(standard representation)}$$

and

$$b(z) = \frac{1}{1-z} = 1 + z + z^2 + \cdots$$

The formal expansion for ξ_n i.e., $x_n + x_{n-1} + x_{n-2} + \cdots$ does not converge but we have, for its Césaro sum,

$$\lim_{q \to \infty} \frac{1}{q} \sum_{p=0}^{q-1} (x_n + x_{n-1} + \cdots + x_{n-p}) = \xi_n$$

(Notice that $x_{n,n-1} = -\xi_{n-1} = -\lim_{q \to \infty} \frac{1}{q} \sum_{p=0}^{q-1} (x_{n-1} + \cdots + x_{n-p-1})$ and $x_{n,n-\ell} = 0$ for $\ell > 1$.)

<u>Proof</u>.

$$\frac{1}{q} \sum_{p=0}^{q-1} (x_n + x_{n-1} + \cdots + x_{n-p}) = \xi_n - \frac{1}{q} (\xi_{n-1} + \xi_{n-2} + \cdots + x_{n-q})$$

and the norm of the second term is $q^{-1/2}$.

8. <u>Linear difference equation.</u>

Let us consider a linear difference equation

(1) $$a_0 x_n + a_1 x_{n-1} + \cdots + a_k x_{n-k} = \xi_n$$

where ξ_n is a given white noise.

<u>Theorem 1</u>. If there exists a stationary sequence x_n satisfying (1), then

(2) $a(z) \equiv \sum_{j=0}^{k} a_j z^j$ has no roots on $|z| = 1$

Conversely if (2) holds, then there exists a unique stationary sequence x_n satisfying (1).

Proof. Let us consider the Kolmogorov-Cramer representation of ξ_n

(3) $$\xi_n = \int e^{i2\pi\lambda n}\, dM_\xi(\lambda)$$

and assume that

(4) $$x_n = \int e^{i2\pi\lambda n}\, dM_x(\lambda)$$

satisfies (1). Then

(5) $$a(e^{-i2\pi\lambda})\, dM_x(\lambda) = dM_\xi(\lambda)$$

and so

(6) $$|a(e^{-i2\pi\lambda})|^2\, dF(\lambda) = d\lambda\ , \qquad dF(\lambda) = \text{the Hinčin measure of } x_n$$

Let N be the set of zero points of $a(e^{-i2\pi\lambda})$ (which is clearly a finite set) and let dG be the restriction of dF over N^c. Then

$$dF \geq dG = |a(e^{-i2\pi\lambda})|^{-2}\, d\lambda.$$

If $N \ni \lambda_0$, then $a(e^{-i2\pi\lambda}) \sim$ non-vanishing constant $\times (\lambda - \lambda_0)$ and so

$$\int dF \geq \int |a(e^{-i2\pi\lambda})|^{-2} d\lambda = \infty,$$

which is a contradiction. Thus N must be empty. This proves the first half of our theorem.

If (2) holds, then the above argument shows that

$$(7) \qquad dF = |a(e^{-i2\pi\lambda})|^{-2} d\lambda$$

and

$$(8) \qquad b(e^{-i2\pi\lambda}) \equiv a(e^{-i2\pi\lambda})^{-1} \in L^2(\Gamma, d\lambda)$$

so that x_n must be expressed as

$$(9) \qquad x_n = \int e^{i2\pi\lambda n} dM_x(\lambda) = \int e^{i2\pi\lambda n} b(e^{-i2\pi\lambda}) dM_\xi(\lambda)$$

and it is clear that x_n, thus defined, solves (1).

<u>Theorem</u> 2. Assume that (2) holds. Then $b(e^{-i2\pi\lambda}) \in L^2(\Gamma, d\lambda)$ and it can be expanded in Fourier series

(10) $$b(e^{-i2\pi\lambda}) = \sum_{j=-\infty}^{\infty} b_j e^{-i2\pi\lambda j}$$

and the solution x_n of (1) is given by

(11) $$x_n = \sum_j b_j \xi_{n-j}$$

If $a(z)$ has no roots in $|z| < 1$ (and so in $|z| \le 1$), then $b_j = 0$ $(j < 0)$ in (11) and (11) gives a canonical backward representation of x_n.

If $a(z)$ has roots in $|z| < 1$, then (11) is not backward.

Proof. The first and third parts are clear. To prove the second part, observe that

$$b(z) = a(z)^{-1}$$

$$= a_0^{-1}(1 - \alpha_1 z)^{-1} (1 - \alpha_2 z)^{-1} \cdots (1-\alpha_k z)^{-1}, \quad (|\alpha_j| < 1)$$

is an outer function in H_2.