TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 00, Number 0, Pages 000–000
S 0002-9947(XX)0000-0

VERBLUNSKY COEFFICIENTS AND NEHARI SEQUENCES

YUKIO KASAHARA AND NICHOLAS H. BINGHAM

Abstract. We are concerned with a rather unfamiliar condition in the theory of the orthogonal polynomials on the unit circle. In general, the Szegö function is determined by its modulus, while the condition in question is that it is also determined by its argument, or in terms of the function theory, that the square of the Szegö function is rigid. In prediction theory, this is known as a spectral characterization of complete nondeterminacy for stationary processes, studied by Bloomfield, Jewel and Hayashi (1983) going back to a small but important result in the work of Levinson and McKean (1964). It is also related with the cerebrated result of Adamyan, Arov and Krein (1968) for the Nehari problem, and there is a one-one correspondence between the Verblunsky coefficients and the negatively indexed Fourier coefficients of the phase factor of the Szegö function, which we call a Nehari sequence. We presents some fundamental results on the correspondence, including extensions of the strong Szegö and Baxter's theorems.

1. Introduction

Let μ be a probability measure on the unit circle $\mathbb{T} = \{|z| = 1\}$ parameterized by $z = e^{i\theta}$. It is called *trivial* if it is supported on a finite set. In the other case, so *nontrivial*, μ generates a system of orthogonal polynomials on the unit circle (OPUC) obtained by the Gram–Schmidt method applied to $1, z, z^2, \dots$ in the Hilbert space $L^2(\mu)$. Let $\Phi_0, \Phi_1, \Phi_2, \dots$ be the monic OPUC, and put $a_n = \Phi_n(0)$ for $n = 1, 2, \dots$. These constants appear in the Szegö recurrence formula [Sz1]

$$\Phi_{n+1} = z\Phi_n + a_{n+1} z^n \overline{\Phi}_n, \tag{1.1}$$

and form a sequence $a = (a_1, a_2, \dots)$ on the open unit disc $\mathbb{D} = \{|z| < 1\}$. Every sequence on $\mathbb{D}$ arises in this way from a unique nontrivial probability measure on $\mathbb{T}$. The last fact was first proved by Verblunsky [V1, V2] in a slightly different context, and a_n take his name, the *Verblunsky coefficients* of μ (also called Schur, Szegö or Geronimus coefficients or parameters). It is a central question in the spectral theory of OPUC how the properties of a correspond to the properties of μ and vice-versa. The strong Szegö theorem [Sz2] and Baxter's theorem [B] to be discussed later are cerebrated results in this direction. For OPUC, see Simon [Si1, Si2, Si3].

Let m be the normalized Lebesgue measure on $\mathbb{T}$. For $1 \le p \le \infty$, denote by L^p the Lebesgue space on $\mathbb{T}$ with respect to m, and by H^p the allied Hardy space on $\mathbb{D}$. As usual, a function in H^p will be identified with its boundary value function in L^p. For a textbook account of Hardy functions, see Hoffman [Ho]. A nonzero function g

2000 *Mathematics Subject Classification.* Primary 42C05; Secondary 42A10, 42A70.
Key words and phrases. Orthgonal polynomials on the unit circle, Verblunky coefficients, Nehari problem, rigid functions.

in H^1 is called *rigid* (or *strongly outer*) if it is determined by its argument, or more specifically, by the unimodular function $g/|g|$ up to a positive constant factor. A rigid function is outer, but the converse is not true. For rigidity and related topics, see Sarason [S3] and Poltoratski–Sarason [PS]. Rigidity arises as key ingredients in many problems, some of which will be discussed soon below and later, while it does not seem to have received considerable attention in OPUC. One of the chief aims of this paper is to elaborate its role in OPUC.

Write $d\mu = wdm + d\mu_s$, where μ_s is the singular part. When $\log w \in L^1$, there is a unique outer function h in H^2 such that $h(0) > 0$ and $w = |h|^2$ a.e. on $\mathbb{T}$. It is called the *Szegö function*, and plays a significant role in the spectral theory of OPUC. Note that h^2 is an outer function in H^1 with $h^2/|h^2| = h/\bar{h}$. In their investigation for the dependence structure of continuous-time stationary Gaussian processes in terms of weighted L^2-space on the line, Levinson–McKean [LM] found a spectral density relevant to rigidity. For its analogue on the circle (known as a characterization of *complete nondeterminacy* for discrete-time stationary processes, studied by Bloomfield *et al.* [BJH]), let us call w a *Levinson–McKean weight* (for short, LM-weight) if

$$\mu_s = 0, \quad \log w \in L^1, \quad \text{and} \quad h^2 \text{ is rigid.} \tag{1.2}$$

See also Dym–McKean [DM] for background. Under the condition (1.2), the space $L^2(\mu)$ has a certain structure that permits us to handle OPUC Φ_n in a particular manner (using von Neumann's alternating projections theorem) of the kind treated by Seghier [Se], Inoue [In1, In2, In3], Inoue–Kasahara [IK1, IK2], and Bingham *et al.* [BIK] in the context of prediction theory. In [LM], the key phrase in (1.2) is actually described as "h is determined by its *phase factor* $h/\bar{h}$", and this means that μ is determined by it, or equivalently, by the two-sided sequence of its Fourier coefficients. In fact, rigidity enables us to specify μ only from the negative half of the sequence, as explained now.

Rigidity also appears in the so-called *Nehari problem*, in which, given a sequence $\gamma = (\gamma_1, \gamma_2, \ldots)$ of complex numbers, one seeks functions ϕ in the unit ball of L^∞ such that $\gamma_n = \int_{\mathbb{T}} e^{in\theta}\phi\, dm$ for all $n = 1, 2, \ldots$. The name came from Nehari [Ne] proving that a solution exists if and only if the Hankel matrix made up from γ acts as a contraction on the space ℓ^2. The problem has more than one solution if and only if γ consists of the negatively indexed Fourier coefficients for the phase factor of some function in H^2. In this indeterminate case, γ will be called a *Nehari sequence*. Adamyan–Arov–Krein [AAK] described the solution set of an indeterminate Nehari problem using an outer function h in H^2 with the following properties: h has unit norm, h^2 is rigid, and $h/\bar{h}$ solves the problem, namely, for $n = 1, 2, \ldots$,

$$\gamma_n = \int_{\mathbb{T}} e^{in\theta}(h/\bar{h})dm. \tag{1.3}$$

Moreover, h can be chosen so $h(0) > 0$. Then h is uniquely determined by γ, and this relation sets up a one-one correspondence between the set of Nehari sequences and the set of LM-weights. For textbook treatments of the Nehari problem, see Peller [P] (operator-theoretic approach originating with [AAK]) and Garnett [G] (function-theoretic approach).

To sum up the above discussion involved with (1.1)–(1.3), a probability measure μ with LM-weight has two kinds of parameters, its Verblunsky coefficients a and its Nehari sequence γ. Thus, besides the 'central question' mentioned earlier, it is also of interest to study its analogue for a and γ. The main purpose of this paper is to present some fundamental results for these questions. The outline is as follows. The next section is devoted to a review of background information on Hardy functions and the concept of rigidity. In Section 3, our research starts with a discussion on the structure of $L^2(\mu)$, stemming from a discovery of Levison–McKean, and related matters in OPUC. Section 4 proceeds to the Adamyan–Arov–Krein theory for establishing some basic relations between a and γ. The final section provides the strong Szegö and Baxter's theorems in terms of a and γ.

2. Rigidity

In this preparatory section, we recall some basic matters in the theory of Hardy functions and the concept of rigidity.

For $1 \leq p \leq \infty$, let L^p be the Lebesgue space with respect to the normalized Lebesgue measure m on the unit circle $\mathbb{T}$, and write $\|\cdot\|_p$ for the usual L^p-norm. The Hardy space H^p is a Banach space of analytic functions $f(z)$ on the unit disc $\mathbb{D}$ such that $\sup_{r<1}\|f_r\|_p < \infty$, the supremum being the norm. Here, $f_r(e^{i\theta}) = f(re^{i\theta})$ and its L^p-norm is nondecreasing in $r < 1$, so that the supremum coincides with the limit as $r \to 1$. An H^p-function $f(z)$ has its boundary-value function $f = \lim_{r\to 1} f_r$ in L^p, from which the original function can be recovered by the Poisson integral. In what follows, these two functions will not be distinguished, and H^p will be regarded as a closed subspace of L^p, so that $H^p = \{f \in L^p \mid \int_{\mathbb{T}} e^{in\theta} f dm = 0 \text{ for } n = 1, 2, \ldots\}$. Note that $H^p \supset H^q$ for $1 \leq p < q \leq \infty$. A nonzero function $f \in H^1$ satisfies $\log|f| \in L^1$, and so it has positive modulus, namely, $|f| > 0$ a.e.. A function $j \in H^\infty$ is called *inner* if $|j| = 1$ a.e.. By the maximum principle, $|j(z)| < 1$ on $\mathbb{D}$ unless j is constant. A nonzero function $g \in H^1$ is called *outer* if

$$g(z) = c \cdot \exp\left(\int_{\mathbb{T}} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log|g(e^{i\theta})| dm\right) \qquad (z \in \mathbb{D}), \tag{2.1}$$

in which $c = g(0)/|g(0)|$, or equivalently, if it has the following extremal property: if $f \in H^1$ satisfies $|f| \leq |g|$ a.e., then $|f(z)| \leq |g(z)|$ on $\mathbb{D}$. Every nonzero function $f \in H^1$ is expressed as $f = jg$ with inner j and outer g satisfying $|f| = |g|$ a.e., and this inner-outer factorization is unique up to constant factors of modulus one.

In a Banach space, a point of the unit ball B is called an *exposed point* of B if there exists a linear functional on the space that attains its norm at the point and only there, and it is necessarily an *extreme point* of B, that is, a point in B which is not a proper convex combination of two distinct points in B. As is well-known, de Leeuw–Rudin [dLR] showed that an H^1-function is an extreme point of the unit ball of H^1 if and only if it is an outer function with unit norm. They also discussed the exposed points of the ball. A linear functional L on H^1 that has g in the unit ball of H^1 attaining $L(g) = \|L\|$ can be expresses as $L(f) = \int_T \phi f dm$ $(f \in H^1)$ with $\phi = \|L\| \cdot |g|/g$, and then $L(f) = \|L\|$ implies $f/|f| = g/|g|$ a.e. as well as $\|f\|_1 = 1$. Hence, the exposed point in question is an H^1-function with unit norm for which no other functions in the unit sphere of H^1 have the same argument. This gives rise to the concept of rigidity to be introduced now.

A nonzero function $g \in H^1$ is called *rigid* (Sarason [S2]) if it is determined by its argument up to a positive constant factor, or more precisely, if $f/|f| = g/|g|$ a.e. implies $f = cg$ for some $c > 0$, subject to $f \in H^1$. Such g is also called *strongly outer* (Nakazi [N1]). It is a set routine to compare two function $f = jo$

and $g = (1+j)^2 o$ with inner j and outer o. Note that g is outer because $1+j$ is so (see [Ho, p.76]). In view of $(1+j)^2 = j|1+j|^2$, they enjoy $f/|f| = g/|g|$, but $g/f = |1+j|^2$ is non-constant unless j is constant. This shows on one hand that non-outer functions cannot be rigid, and on the other hand that there are nonrigid outer functions. Unfortunately, no one has found a structural characterization of rigid functions, and there is no good answer to the question which outer functions are rigid. Each of the following two conditions is sufficient for a function in H^1 to be rigid. One is that its reciprocal is also in H^1, and the other is that its real part is nonnegative (see Yabuta [Y]). See Sarason [S3] and Poltoratski–Sarason [PS] for further information on rigidity.

3. Levinson–McKean weights

In this section, we discuss the structure of $L^2(\mu)$ with LM-weight and relevant matters, including a useful expansion formula for OPUC.

Let μ be a probability measure on $\mathbb{T}$. Denote by $H^2(\mu)$ and $H^2_-(\mu)$ the closed subspaces of $L^2(\mu)$ spanned by $1, z, z^2, \ldots$ and $\bar{z}, \bar{z}^2, \ldots$, respectively, and by H^2 and H^2_- those spaces for the normalized Lebesgue measure m. This kind of set-up is often used for studying a discrete-time, wide-sense, stationary stochastic process with spectral measure μ, and $H^2(\mu)$ and $H^2_-(\mu)$ are interpreted as its *past* and *future*. The process is called *completely nondeterministic* if $H^2_-(\mu) \cap H^2(\mu) = \{0\}$, according to Sarason [S1]. Recall $d\mu = w dm + d\mu_s$, where μ_s is the singular part. The Szegö alternative asserts that

$$L^2(\mu_s) = \cap_n [z^n H^2(\mu)] \subsetneq H^2(\mu) \subsetneq L^2(\mu)$$

if $\log w \in L^1$, otherwise $\cap_n [z^n H^2(\mu)] = H^2(\mu) = L^2(\mu)$ (cf. Dym–McKean [DM]). This shows that $\mu_s = 0$ and $\log w \in L^1$ are necessary for complete nondeterminacy. If $\log w \in L^1$, there is a unique outer function $h \in H^2$ satisfying $w = |h|^2$ a.e. and $h(0) > 0$. It is called the *Szegö function*, and given by

$$h(z) = \exp\left(\int_{\mathbb{T}} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log \sqrt{w}\, dm \right) \qquad (z \in \mathbb{D}).$$

For $n = 1, 2, \ldots$, let P_n be the space of polynomials with degree less than n, and put $P_0 = \{0\}$. Let us now recall the definition (1.2) of an LM-weight. It originated with Levinson–KcKean [LM, p. 105], in which they found (1.2) to be a spectral characterization of the identity

$$z^n H^2_-(\mu) \cap H^2(\mu) = P_n \tag{3.1}$$

for $n = 1$ (actually, this is an analogue of their result in the weighted L^2-space on the real line, but it can be proved in just the same way as they did). It seems that their discovery had not received considerable attention until Bloomfield *et al.* [BJH, Proposition 5] characterized complete nondeterminacy in terms of the exposed points in the unit ball of H^1. It is easy to see that the occurrence of (3.1) does not depend on $n = 0, 1, 2, \ldots$, as long as μ is nontrivial. Indeed, one has

$$z^{n-1} H^2_-(\mu) \cap H^2(\mu) = [z^n H^2_-(\mu) \cap H^2(\mu)] \cap \bar{z}\, [z^n H^2_-(\mu) \cap H^2(\mu)],$$

while if $\log w \in L^1$, it follows from $z^{n+1} H^2_-(\mu) = z^n H^2_-(\mu) + z^n \mathbb{C}$ that

$$z^{n+1} H^2_-(\mu) \cap H^2(\mu) = [z^n H^2_-(\mu) \cap H^2(\mu)] + z^n \mathbb{C}.$$

So, by induction, (3.1) holds for every $n = 0, 1, 2, \ldots$ if it holds for some n (cf. Inoue–Kasahara [IK2, Theorem 2.3]), except for the trivial case that μ is supported on a finite set, which makes $P_n = L^2(\mu)$ for all $n \geq \sharp\,\mathrm{supp}(\mu)$.

The following is a consequence of the above discussion on the structure of $L^2(\mu)$ with LM-weight (cf. [IK2, Remark 2]).

Theorem 3.1. *Let μ be a probability measure on $\mathbb{T}$. If μ satisfies* (1.2)*, then* (3.1) *holds for every $n = 0, 1, 2, \ldots$. Conversely, if μ is nontrivial and* (3.1) *holds for some $n = 0, 1, 2, \ldots$, then μ satisfies* (1.2)*.*

Remark 3.2. To the best of our knowledge, the full statement of Theorem 3.1 has never appeared before, while the essence of the theorem was clarified in the 1980s. For instance, the point of the matter is implicitly given in [BJH, Propositions 9] as a spectral characterization of k-step complete nondeterminacy. Some relevant results may also be found in Hayashi [Ha, Theorem 1], Nakazi [N2, Theorem 9] and Younis [Yo, Theorem 4].

Let us turn to the monic OPUC Φ_n, which are the projections of z^n onto $P_n^{\perp}$ in $L^2(\mu)$. If $\mu_s = 0$ and $\log w \in L^1$, the mapping $f \to hf$ from $L^2(\mu)$ to L^2 is a Hilbert space isomorphism, and Beurling's theorem (see [Ho, p.101]) implies

$$h[z^n H^2_-(\mu) \cap H^2(\mu)] = z^n (h/\bar{h}) H^2_- \cap H^2.$$

This is quite useful in studying the OPUC Φ_n for an LM-weight. Indeed, by (3.1), the image $h\Phi_n$ is obtained by projecting $z^n h$ onto the orthogonal complement of $z^n(h/\bar{h})H^2_- \cap H^2 = hP_n$ in L^2, and this kind of projection can be described in terms of the Toeplitz and Hankel operators, as we now explain.

Let $\phi \in L^\infty$, and write M_ϕ for the multiplication operator on L^2 defined by $M_\phi f = \phi f$. The Toeplitz operator T_ϕ on H^2 and the Hankel operator H_ϕ from H^2 to H^2_- with symbol ϕ are defined by

$$T_\phi = PM_\phi|_{H^2}, \qquad H_\phi = P_- M_\phi|_{H^2},$$

where P and P_- are the orthogonal projection operators of L^2 onto H^2 and H^2_-, respectively. Clearly, $T_\phi + H_\phi = M_\phi|_{H^2}$. Their adjoint operators are given by

$$T^*_\phi = PM_{\bar{\phi}}|_{H^2} = T_{\bar{\phi}}, \qquad H^*_\phi = PM_{\bar{\phi}}|_{H^2_-}.$$

The projection just stated can be treated in the following approximation scheme.

Lemma 3.3. *Let ϕ be a unimodular function in L^∞, and let Q_ϕ be the orthogonal projection operator of H^2 onto $\phi H^2_- \cap H^2$. Then, for $f \in H^2$, it holds that*

$$f - Q_\phi f = \sum_{k=0}^{\infty} T_\phi [H^*_\phi H_\phi]^k T^*_\phi f,$$

the sum converging strongly in H^2.

Proof. Let $|\phi| = 1$. Then $M_\phi P_- M_{\bar{\phi}}$ is the orthogonal projection operator of L^2 onto ϕH^2_-. Since $H^*_{\bar{\phi}} H_{\bar{\phi}} = PM_\phi P_- M_{\bar{\phi}}|_{H^2}$, it follows from von Neumann's alternating projections theorem (see Halmos [H, Problem 122]) that $(H^*_{\bar{\phi}} H_{\bar{\phi}})^n$ converges to Q_ϕ as $n \to \infty$ in the strong operator topology. Recall that

$$I - H^*_{\bar{\phi}} H_{\bar{\phi}} = T_\phi T^*_\phi, \qquad H^*_{\bar{\phi}} H_{\bar{\phi}} T_\phi = T_\phi H^*_\phi H_\phi$$

(see Peller [P, (3.1.5) and (4.4.1)]). Repeated use of these relations gives

$$I - (H_{\bar\phi}^* H_{\bar\phi})^n = \sum_{k=0}^{n-1} T_\phi [H_\phi^* H_\phi]^k T_\phi^*,$$

whence the lemma follows. □

Remark 3.4. If $I - H_\phi^* H_\phi$ is invertible, the above result may be rewritten as Seghier's formula [Se, Proposition 3], namely,

$$f - Q_\phi f = \phi(I - H_\phi^* H_\phi)^{-1} T_\phi^* f - H_\phi (I - H_\phi^* H_\phi)^{-1} T_\phi^* f.$$

For convenience, write T_n and H_n for the Toeplitz and Hankel operators with symbol $\phi = z^n(h/\bar h)$. By Theorem 3.1, when μ has an LM-weight, its monic OPUC admit the following expansion.

Proposition 3.5. *If μ satisfies* (1.2)*, then, for every $n = 0, 1, 2, \ldots$,*

$$h\Phi_n = h(0) \sum_{k=0}^{\infty} T_n [H_n^* H_n]^k 1,$$

the sum converging strongly in H^2.

Proof. Apply Lemma 3.3 to $f = z^n h$, noting $T_n^*(z^n h) = P\bar h = h(0)$. □

The above is a consequence from Theorem 3.1 that if μ has an LM-weight,

(3.2) $$\Phi_n \text{ is perpendicular to } z^n H_-^2(\mu) \cap H^2(\mu).$$

To see the converse implication, recall Hayashi's criterion for rigidity [Ha, p.695]:

$$h^2 \textit{ is rigid if and only if } h \textit{ is perpendicular to } (h/\bar h) H_-^2 \cap H^2,$$

provided that h is a nonzero function in H^2. This can be extended to the following characterization of LM-weights in terms of OPUC.

Theorem 3.6. *Let μ be a probability measure on $\mathbb{T}$. If μ satisfies* (1.2)*, then* (3.2) *holds for every $n = 0, 1, 2, \ldots$. Conversely, if μ is nontrivial and* (3.2) *holds for some $n = 0, 1, 2, \ldots$, then μ satisfies* (1.2).

Proof. The converse half is to be checked. Since μ is nontrivial, Φ_n never vanishes, and (3.2) has a proper meaning. If it holds for some $n = 1, 2, \ldots$, by symmetry, $z^n \overline{\Phi}_n$ is perpendicular to $z[z^n H_-^2(\mu) \cap H^2(\mu)]$, so that Φ_{n-1} is perpendicular to $z^{n-1} H_-^2(\mu) \cap H^2(\mu)$ in view of the inverse Szegö recursion

$$z\Phi_{n-1} = \rho_n^{-2} [\Phi_n - a_n z^n \overline{\Phi}_n],$$

where $\rho_n = \sqrt{1 - |a_n|^2}$, and thus, by induction, (3.2) holds for $n = 0$. Therefore, in any case, the assumption implies that 1 is perpendicular to $H_-^2(\mu) \cap H^2(\mu)$. The Szegö alternative now allows us to have $\mu = |h|^2 m$, showing that h is perpendicular to $(h/\bar h) H_-^2 \cap H^2$. Consequently, by Hayashi's criterion, μ has an LM-weight. □

Remark 3.7. It is also interesting to see Hayashi's criterion through our approximation scheme. Take a nonzero function h in H^2 and project it onto $(h/\bar h) H_-^2 \cap H^2$. Then, since $T_{h/\bar h}^* h = \overline{h(0)}$, it follows from Lemma 3.3 that

$$\|h - Q_{h/\bar h} h\|_2^2 = (h - Q_{h/\bar h} h, h)_2 = |h(0)|^2 \sum_{k=0}^{\infty} ([H_{h/\bar h}^* H_{h/\bar h}]^k 1, 1)_2,$$

which, by Hayashi's criterion, coincides with $\|h\|_2^2$ if and only if h^2 is rigid. This should be compared with [Ha, Theorem 8], which states that the equality holds in $|h(0)|/\inf_{f\in H^2}\|T_{h/\bar{h}}(1+zf)\|_2 \le \|h\|_2$ if and only if h^2 is rigid.

4. Verblunsky coefficients and Nehari sequences

In this section, after recalling the result of Adamyan–Arov–Krein, we study how a Nehari sequence corresponds, via an LM-weight, to its Verblunsky coefficients.

By Herglotz's theorem, the association

$$\frac{1+zf(z)}{1-zf(z)} = \int_{\mathbb{T}} \frac{e^{i\theta}+z}{e^{i\theta}-z}d\mu \qquad (z\in\mathbb{D}) \tag{4.1}$$

sets up a one-one correspondence $f \leftrightarrow \mu$ between the unit ball of H^∞ and the set of probability measures on $\mathbb{T}$. Here, f is called the *Schur function* of μ, and the latter is trivial if and only if f is a finite Blaschke product. Also, since

$$w = \frac{1-|f|^2}{|1-zf|^2} \quad \text{a.e.}, \tag{4.2}$$

where w is the density in $d\mu = wdm + d\mu_s$, the Szegö condition $\log w \in L^1$ is fulfilled if and only if $\log(1-|f|) \in L^1$, which means that f is not an extreme point of the unit ball of H^∞ (see [Ho, p.138]). If f is not a finite Blaschke product, first putting $f_0 = f$ and then iterating

$$\alpha_n = f_n(0), \qquad f_{n+1}(z) = \frac{1}{z}\frac{f_n(z)-\alpha_n}{1-\bar{\alpha}_n f_n(z)}, \tag{4.3}$$

one has a sequence of functions $f_0, f_1, f_2, \ldots$, called the *Schur iterates* of f, in the unit ball of H^∞. The numbers α_n are called the *Schur parameters* of f. In fact, they are essentially the same as the Verblunsky coefficients a_n for μ, namely,

$$a_n = -\bar{\alpha}_{n-1}.$$

This fact is *Geronimus' theorem.* See Simon [Si1] for background.

Let $\gamma = (\gamma_1, \gamma_2, \ldots)$ be a sequence of complex numbers. In the Nehari problem, one seeks functions ϕ in the unit ball of L^∞ that have γ as its negatively indexed Fourier coefficients, that is, $\gamma_n = \int_{\mathbb{T}} e^{in\theta}\phi dm$ for all $n = 1, 2, \ldots$. In this paper, as mentioned earlier, γ is called a Nehari sequence if the problem has more than one solution. Adamyan–Arov–Krein [AAK] showed that such γ has an outer function h with unit norm that yields a bijection $u \to \phi$ from the unit ball of H^∞ onto the solution set, defined by the formula

$$\phi = (h/\bar{h}) - \frac{h^2(1-zf)(1-u)}{1-zfu}, \tag{4.4}$$

where f is the Schur function of $\mu = |h|^2 m$. In particular, h^2 is rigid, and this is essential for describing the full solutions; as long as $h \in H^2$ satisfies (1.3) and has unit norm, the formula (4.4) provides many solutions as u runs through the unit ball of H^∞. The above formula may be rewritten as

$$\phi = \frac{h(1-zf)(u-\overline{zf})}{\bar{h}(1-\overline{zf})(1-zfu)}.$$

So, $|\phi| = 1$ a.e. if and only if u is inner, in view of

$$|1-zfu|^2 - |u-\overline{zf}|^2 = (1-|f|^2)(1-|u|^2).$$

For $\lambda \in \mathbb{T}$, the solution ϕ corresponding to $u = \lambda$ is called *canonical*. It arises as the phase factor $\phi = h_\lambda/\bar{h}_\lambda$ of the function

$$h_\lambda = \sqrt{\lambda}\, \frac{h(1 - zf)}{1 - z(\lambda f)},$$

in which $\sqrt{\lambda}$ stands for one of the square-roots of λ. This h_λ has unit norm, and its square is rigid. In particular, it plays the same role as h in describing the full solution set. In this connection, the measures $\mu_\lambda = |h_\lambda|^2 m$, associated with $f_\lambda = \lambda f$, constitute the family of the Alexandrov–Clark measures of $\mu = |h|^2 m$ (see Poltoratski–Sarason [PS] and the references cited there).

Now, in the above result of [AAK], one may assume $h(0) > 0$ (otherwise take h_λ with $\sqrt{\lambda} = |h(0)|/h(0)$ instead of h). Then γ is associated, via the Szegö function h satisfying (1.3), with a probability measure $\mu = |h|^2 m$ with LM-weight. This defines a one-one correspondence between the set of Nehari sequences and the set of LM-weights. Moreover, by Verblunsky's theorem, a Nehari sequence γ is also associated with a unique Verblunsky coefficients a in such a way that

$$a = (a_1, a_2, \dots) \quad \leftrightarrow \quad \mu \quad \leftrightarrow \quad \gamma = (\gamma_1, \gamma_2, \dots). \tag{4.5}$$

In view of this, the question naturally arises as to how a Nehari sequence corresponds, via an LM-weight, to its Verblunsky coefficients.

The main aim here is to establish the following commutativity between the correspondence $a \leftrightarrow \gamma$ and the unilateral shift $(x_1, x_2, \dots) \to (x_2, x_3, \dots)$, and the 'inverse' relation. (These are relevant to the 'coefficient stripping' in [Si1].)

Theorem 4.1. *For $a = (a_1, a_2, \dots)$ corresponding to $\gamma = (\gamma_1, \gamma_2, \dots)$ via μ satisfying* (1.2) *as in* (4.5)*, the following hold:*

(i) *$\tilde{a} = (a_2, a_3, \dots)$ corresponds to $\tilde{\gamma} = (\gamma_2, \gamma_3, \dots)$, via a certain probability measure $\tilde{\mu}$ satisfying* (1.2)*.*

(ii) *For any $\zeta \in \mathbb{D}$, $a_\zeta = (\zeta, a_1, a_2, \dots)$ corresponds, via a certain probability measure μ_ζ satisfying* (1.2)*, to $\gamma_\zeta = (\omega_\zeta, \gamma_1, \gamma_2, \dots)$ with*

$$\omega_\zeta = \int_{\mathbb{T}} (h/\bar{h})\, dm - h(0)^2(1 - \zeta), \tag{4.6}$$

where h is the Szegö function of μ.

Before passing to the proof, let us discuss a pair of immediate by-products of the above theorem. In OPUC, one expects that the substantial properties of μ are determined by the asymptotic behavior of a. The following indicates that this also applies to LM-weights. (Compare [Si1, Theorem 3.4.4].)

Corollary 4.2. *Let μ and ν be two nontrivial probability measures on $\mathbb{T}$ with Verblunsky coefficients a and b, respectively. Suppose for some K and $k > -K$,*

$$a_n = b_{n+k} \qquad (n \geq K).$$

Then either both μ and ν satisfy (1.2)*, or neither does.*

This can be rephrased in terms of the Schur functions, as follows.

Corollary 4.3. *Let $f_0, f_1, f_2, \dots$ be the Schur iterates arising from a nontrivial probability measure. Then either all of them correspond, as the Schur functions, to probability measures obeying* (1.2)*, or none of them do.*

Theorem 4.1 (i) readily follows from the next expansion formula for the Verblunsky coefficients for an LM-weight. Recall that T_n and H_n stand for the Toeplitz and Hankel operators with symbol $z^n(h/\bar{h})$.

Proposition 4.4. *Let μ be a probability measure on $\mathbb{T}$. If it satisfies* (1.2)*, then*

$$a_n = \sum_{k=0}^{\infty} (T_n [H_n^* H_n]^k 1, 1)_2$$

holds for every $n = 1, 2, \ldots$.

Proof. Since $a_n = \Phi_n(0)$, we have $h\Phi_n \equiv a_n h \pmod{zH^2}$. By Proposition 3.5,

$$a_n h(0) = (a_n h, 1)_2 = (h\Phi_n, 1)_2 = h(0) \sum_{k=0}^{\infty} (T_n [H_n^* H_n]^k 1, 1)_2.$$

Thus, the assertion follows. □

Proof of Theorem 4.1 (i). Let a, μ and γ be as in (4.5). Then $\tilde{\gamma} = (\gamma_2, \gamma_3, \ldots)$ is a Nehari sequence having a solution $\phi = z(h/\bar{h})$, and it corresponds to some probability measure $\tilde{\mu} = |\tilde{h}|^2 m$ satisfying (1.2). Write $\tilde{T}_n$ and $\tilde{H}_n$ for the Toeplitz and Hankel operators with symbol $z^n(\tilde{h}/\bar{\tilde{h}})$. Since the phase factor of $\tilde{h}$ solves the Nehari problem for $\tilde{\gamma}$, it follows that

$$\tilde{H}_n = H_{n+1}, \qquad \tilde{T}_n^* 1 = T_{n+1}^* 1.$$

So, by Proposition 4.4, the Verblunsky coefficients $\tilde{a}_n$ of $\tilde{\mu}$ satisfy

$$\tilde{a}_n = \sum_{k=0}^{\infty} (\tilde{T}_n [\tilde{H}_n^* \tilde{H}_n]^k 1, 1)_2 = \sum_{k=0}^{\infty} (T_{n+1} [H_{n+1}^* H_{n+1}]^k 1, 1)_2 = a_{n+1},$$

which means $\tilde{a} = (a_2, a_3, \ldots)$. □

Remark 4.5. By Proposition 4.4, the map $\gamma \to a$ from the set of Nehari sequences to the set of their Verblunsky coefficients may be visualized as

$$a_n = \sum_{k=0}^{\infty} {}^t\boldsymbol{\gamma}_n (\Gamma_n^* \Gamma_n)^k \boldsymbol{e},$$

where

$$\boldsymbol{\gamma}_n = \begin{pmatrix} \gamma_n \\ \gamma_{n+1} \\ \gamma_{n+2} \\ \vdots \end{pmatrix}, \quad \Gamma_n = \begin{pmatrix} \gamma_{n+1} & \gamma_{n+2} & \gamma_{n+3} & \cdots \\ \gamma_{n+2} & \gamma_{n+3} & \gamma_{n+4} & \cdots \\ \gamma_{n+3} & \gamma_{n+4} & \gamma_{n+5} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \boldsymbol{e} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

(From this, it is easily seen that $\lambda\gamma$, which is a Nehari sequence, corresponds to λa for any $\lambda \in \mathbb{T}$.) In this connection, when μ has an LM-weight, one may write

$$\|\Phi_n\|_\mu^2 = h(0)^2 \sum_{k=0}^{\infty} {}^t\boldsymbol{e}\, (\Gamma_n^* \Gamma_n)^k \boldsymbol{e}.$$

See Inoue [In1, In2, In3], Inoue–Kasahara [IK1, IK2] and Bingham *et al.* [BIK] for relevant results with application in prediction theory.

Theorem 4.1 (ii) can be derived from (i) with the aid of Geronimus' theorem and the next parameterization for the one-step extensions of a Nehari sequence $\gamma = (\gamma_1, \gamma_2, \dots)$. To state it, put

$$D(\gamma) = \{\omega_\zeta \mid \zeta \in \mathbb{D}\},$$

where ω_ζ is the constant displayed in (4.6). This is an open disc of radius $h(0)^2$ centered at $\int_{\mathbb{T}} (h/\bar{h}) dm - h(0)^2$. Also, it follows that $D(\gamma) \subset \mathbb{D}$.

Lemma 4.6. *Let $\gamma = (\gamma_1, \gamma_2, \dots)$ be a Nehari sequence. Then its one-step extension $(\omega, \gamma_1, \gamma_2, \dots)$ remains a Nehari sequence if and only if $\omega \in D(\gamma)$.*

Proof. For any $\zeta \in \mathbb{D}$, the extension $(\omega_\zeta, \gamma_1, \gamma_2, \dots)$ is a Nehari sequence, for

$$\phi = \bar{z}\left[(h/\bar{h}) - \frac{h^2(1-zf)(1-u)}{1-zfu}\right]$$

solves its Nehari problem as long as $u(0) = \zeta$. Every extension is of such form. Indeed, if $(\omega, \gamma_1, \gamma_2, \dots)$ is a Nehari sequence, the problem has a solution of the form $k/\bar{k}$ with $k \in H^2$. This means that $\gamma_n = \int_{\mathbb{T}} e^{in\theta} z(k/\bar{k}) dm$ for $n = 1, 2, \dots$, namely, $z(k/\bar{k})$ is a solution of the Nehari problem for the original sequence γ. So there is an inner function u such that

$$z(k/\bar{k}) = h/\bar{h} - \frac{h^2(1-zf)(1-u)}{1-zfu},$$

and the additional entry ω is evaluated as

$$\omega = \int_{\mathbb{T}} e^{i\theta}(k/\bar{k}) dm = \int_{\mathbb{T}} (h/\bar{h}) dm - h(0)^2(1-u(0)).$$

Here $u(0) \in \mathbb{D}$ because otherwise $u = \lambda$ ($\lambda = u(0) \in \mathbb{T}$) leads to $z(k/\bar{k}) = h_\lambda/\bar{h}_\lambda$, contradicting the fact that h_λ^2 is rigid. So ω lies in the disc $D(\gamma)$. $\square$

Proof of Theorem 4.1 (ii). Let a, μ and γ be as in (4.5). Take $\zeta \in \mathbb{D}$. Then, by Lemma 4.6, $\gamma_\zeta = (\omega_\zeta, \gamma_1, \gamma_2, \dots)$ is a Nehari sequence, and it corresponds to the Verblunsky coefficients, say, a_ζ of a probability measure $\mu_\zeta = |h_\zeta|^2 m$ obeying (1.2). By Theorem 4.1 (i), a_ζ is specified up its first entry, denoted by ζ_0, as

$$a_\zeta = (\zeta_0, a_1, a_2, \dots).$$

It remains to check $\zeta_0 = \zeta$. Let f and f_ζ be the Schur functions of $\mu = |h|^2 m$ and $\mu_\zeta = |h_\zeta|^2 m$, respectively. Geronimus' theorem plays a crucial role here. Namely, it implies both that $f_\zeta(0) = -\bar{\zeta}_0$ and that f_ζ has f as its first Schur iterate. The upshot from these two facts is that

$$f_\zeta(z) = \frac{zf(z) - \bar{\zeta}_0}{1 - \zeta_0 z f(z)},$$

and a direct computation involving (4.2) gives

$$h_\zeta(z) = \frac{\sqrt{1-|\zeta_0|^2}}{1+\bar{\zeta}_0 z} \cdot \frac{h(z)(1-zf(z))}{1-zf(z)b_{\zeta_0}(z)} \quad \text{with} \quad b_{\zeta_0}(z) = \frac{z+\zeta_0}{1+\bar{\zeta}_0 z},$$

whence

$$z(h_\zeta/\bar{h}_\zeta) = (h/\bar{h}) - \frac{h^2(1-zf)(1-b_{\zeta_0})}{1-zfb_{\zeta_0}}.$$

Since $b_{\zeta_0}(0) = \zeta_0$, the proof is finished by

$$\omega_\zeta = \int_{\mathbb{T}} e^{i\theta}(h_\zeta/\bar{h}_\zeta)dm = \int_{\mathbb{T}} (h/\bar{h})dm - h(0)^2(1-\zeta_0),$$

showing $\zeta_0 = \zeta$, as required. □

Since rigidity is a key concept of this paper, it would be worthwhile to single out the following outcome from the above proof, together with a direct proof.

Proposition 4.7. *Let h be a Szegö function. Suppose that h has unit norm and h^2 is rigid. Also, let f be the Schur function of $\mu = |h|^2 m$. Then for any $\zeta \in \mathbb{D}$, the Szegö function h_ζ, defined by*

$$h_\zeta(z) = \frac{\sqrt{1-|\zeta|^2}}{1+\bar{\zeta}z} \cdot \frac{h(z)(1-zf(z))}{1-zf(z)b_\zeta(z)} \quad \textit{with} \quad b_\zeta(z) = \frac{z+\zeta}{1+\bar{\zeta}z},$$

inherits its predecessor's property, viz, h_ζ has unit norm and h_ζ^2 is rigid.

A direct proof of Proposition 4.7. Thanks to Sarason [S2, Theorem 2], the two properties of h_ζ in question can be deduced from those of the function

$$g_\zeta(z) = \frac{h(z)(1-zf(z))}{1-zf(z)b_\zeta(z)}.$$

That is to say, since b_ζ is inner, Sarason's theorem implies that g_ζ has unit norm and g_ζ^2 is rigid. Now, (4.2) gives $|h(1-zf)|^2 = 1-|f|^2$, so that

$$|g_\zeta|^2 = \frac{1-|fb_\zeta|^2}{|1-zfb_\zeta|^2}.$$

Hence, it follows from $\|g_\zeta\|_2 = 1$ that

$$\int_{\mathbb{T}} \frac{e^{i\theta}+z}{e^{i\theta}-z}\,|g_\zeta|^2 dm = \frac{1+zf(z)b_\zeta(z)}{1-zf(z)b_\zeta(z)},$$

and since $b_\zeta(-\zeta) = 0$, the value of this function at $z = -\zeta$ is one. Therefore,

$$\|h_\zeta\|_2^2 = \int_T \frac{1-|\zeta|^2}{|1+\bar{\zeta}e^{i\theta}|^2}\,|g_\zeta|^2 dm = \mathrm{Re}\left[\int_{\mathbb{T}} \frac{e^{i\theta}-\zeta}{e^{i\theta}+\zeta}\,|g_\zeta|^2 dm\right] = 1.$$

To check rigidity, take an H^2-function k satisfying $k/\bar{k} = h_\zeta/\bar{h}_\zeta$, which implies that $(1+\bar{\zeta}z)k/\overline{(1+\bar{\zeta}z)k} = g_\zeta/\bar{g}_\zeta$. Since g_ζ^2 is rigid, $(1+\bar{\zeta}z)k$ is a real multiple of g_ζ, so that k is a real multiple of h_ζ. Thus, h_ζ^2 is rigid. □

Remark 4.8. The above fact provides another proof of Theorem 4.1 as follows. Let f be the Schur function that generates $a = (a_1, a_2, \ldots)$, and also let $\zeta \in \mathbb{D}$. By Geronimus' theorem, the Verblunsky coefficients $a_\zeta = (\zeta, a_1, a_2, \ldots)$ arise from

$$f_\zeta(z) = \frac{zf(z)-\bar{\zeta}}{1-\zeta zf(z)}.$$

Then Proposition 4.7 implies that the probability measure μ_ζ corresponding to a_ζ has an LM-weight $d\mu_\zeta = |h_\zeta|^2 dm$, and it is seen from the formula

$$z(h_\zeta/\bar{h}_\zeta) = (h/\bar{h}) - \frac{h^2(1-zf)(1-b_\zeta)}{1-zfb_\zeta}$$

that the Nehari sequence of μ_ζ is just as in Theorem 4.1 (ii), from which (i) readily follows. The detail is omitted.

5. Strong Szegö and Baxter's theorems

In this final section, we extend the Strong Szegö and Baxter's theorems to the correspondence between the Verblunsky coefficients and the Nehari sequences.

Let N^+ be the Smirnov class on $\mathbb{D}$, the linear space of all quotients u/v with $u, v \in H^\infty$, v being outer. A nonzero function $f \in N^+$ satisfies $\log|f| \in L^1$, and (2.1) defines the outer functions in N^+ as in the case of H^p. So, a nonzero function $f \in N^+$ admits the inner-outer factorization $f = jg$ with inner j and outer g. The Hardy space H^p is expressed as $H^p = L^p \cap N^+$. (This is also valid for $0 < p < 1$; in this case, H^p is defined in the same way as in Section 2, but it is not a Banach space.) These basic matters on the class N^+ may be found in Duren [D].

Under the restriction $\log w \in L^1$ for $d\mu = wdm + d\mu_s$, the correspondence (4.1) between Schur functions f and probability measures μ can be modified as follows.

Proposition 5.1. *The formula*

$$\frac{\sigma(z) + z\tau(z)}{\sigma(z) - z\tau(z)} = \int_{\mathbb{T}} \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu \qquad (z \in \mathbb{D}) \tag{5.1}$$

defines a one-one correspondence $\tau \leftrightarrow \mu$ between the Smirnov class N^+ and the set of probability measures with the Szegö condition $\log w \in L^1$, provided that σ is the unique outer function in N^+ determined by τ via

$$\sigma(0) > 0, \qquad |\sigma|^2 - |\tau|^2 = 1 \quad a.e..$$

Further, the Szegö function h of μ satisfies

$$h^{-1} = \sigma - z\tau \quad a.e., \qquad h = (h/\bar{h})\sigma + \overline{z\tau} \quad a.e..$$

Proof. Pick $\tau \in N^+$ and consider the outer function $\sigma \in N^+$ given by

$$\sigma(z) = \exp\left[\int_{\mathbb{T}} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log\sqrt{1 + |\tau|^2}\, dm\right] \qquad (z \in \mathbb{D}).$$

Then the quotient $f = \tau/\sigma$ lies in the unit ball of H^∞, and enjoys $\log(1 - |f|) \in L^1$ because of $1 - |f|^2 = |\sigma|^{-2}$. So, f is the Schur function of a probability measure μ with $\log w \in L^1$, and their relation (4.1) gives (5.1). The Szegö function h satisfies $h = 1/(\sigma - z\tau)$ because $w = 1/|\sigma - z\tau|^2$ follows from (4.2), and thus,

$$\bar{h}^{-1}\sigma + h^{-1}\overline{z\tau} = |\sigma|^2 - |\tau|^2 = 1,$$

showing $h = (h/\bar{h})\sigma + \overline{z\tau}$. The other half can be checked in a similar discussion on the pair $\sigma = 1/h(1 - zf)$ and $\tau = f/h(1 - zf)$ arising from f and h of μ. □

Remark 5.2. For $0 < p \le \infty$, one has $\sigma \in H^p$ if and only if $\tau \in H^p$. (Indeed, $|\tau|^p \le |\sigma|^p \le 2^{p/2}(|\tau|^p + 1)$ for $0 < p < \infty$.) Since $|h|^2 = \sigma h + \overline{z\tau h}$, μ is absolutely continuous if $\tau h \in H^1$ (or equivalently, if $\sigma h \in H^1$). In particular, μ satisfies (1.2) if $\tau \in H^2$, which implies $w^{-1} \in L^1$. Actually, τ belongs to H^2 if and only if $\mu_s = 0$, $\log w \in L^1$ and $h \in T_{h/\bar{h}} H^2$.

Let μ be a probability measure with $\log w \in L^1$, and let $\sigma, \tau \in N^+$ be as in (5.1). Define σ_n and τ_n for $n = 0, 1, \ldots$ by putting $\sigma_0 = \sigma$ and $\tau_0 = \tau$, and iterating

$$\sigma_{n+1} = \sigma_n + a_{n+1}\tau_n, \qquad z\tau_{n+1} = \tau_n + \bar{a}_{n+1}\sigma_n, \tag{5.2}$$

where a_n are the Verblunsky coefficients of μ. This may be regarded as a bridge between the Szegö recurrence and the Schur algorithm, in the following sense.

Proposition 5.3. *Let μ be a probability measure with $\log w \in L^1$, and let f be its Schur function. Then the monic OPUC Φ_n of μ satisfy*

$$h\Phi_n = z^n(h/\bar{h})\sigma_n + \overline{z\tau_n} \qquad a.e., \tag{5.3}$$

and the Schur iterates f_n of f enjoy

$$f_n(z) = \tau_n(z)/\sigma_n(z) \qquad (z \in \mathbb{D}). \tag{5.4}$$

In particular, both σ_n and τ_n belong to N^+, and σ_n are outer.

Proof. The above iteration gives

$$\begin{aligned} &z^{n+1}\bar{h}^{-1}\sigma_{n+1} + h^{-1}\overline{z\tau_{n+1}} \\ =&z[z^n\bar{h}^{-1}\sigma_n + h^{-1}\overline{z\tau_n}] + a_{n+1}z^n\overline{[z^n\bar{h}^{-1}\sigma_n + h^{-1}\overline{z\tau_n}]}. \end{aligned}$$

Since Proposition 5.1 says that $\bar{h}^{-1}\sigma + \bar{h}^{-1}\overline{z\tau} = 1$, this is nothing but the Szegö recurrence formula (1.1). That is to say, (5.3) holds for every $n = 0, 1, 2, \dots$. By Geronimus' theorem, namely, $a_{n+1} = -\bar{\alpha}_n$, the iteration also yields

$$\tau_{n+1}/\sigma_{n+1} = \frac{1}{z}\frac{(\tau_n/\sigma_n) - \alpha_n}{1 - \bar{\alpha}_n(\tau_n/\sigma_n)}, \qquad \sigma_{n+1} = \sigma_n[1 - \bar{\alpha}_n(\tau_n/\sigma_n)].$$

The former takes the same form as Schur's iteration (4.3), while the latter implies that σ_{n+1} is outer if σ_n is outer and if τ_n/σ_n is a Schur function. Note that τ_{n+1} belongs to N^+ if $f_n = \tau_n/\sigma_n$, because $\tau_n(0) + \bar{a}_{n+1}\sigma_n(0) = 0$ guarantees that $\tau_n + \bar{a}_{n+1}\sigma_n$ is divisible by z in N^+. Since $\sigma, \tau \in N^+$, σ is outer, and $f = \tau/\sigma$ as seen in the proof of Proposition 5.1, the remaining part is verified by induction. □

Remark 5.4. The above iteration gives $|\sigma_n|^2 - |\tau_n|^2 = \prod_{j=1}^n(1 - |a_j|^2)$. Naturally, one may also consider the 'normalized' pairs with $|s_n|^2 - |t_n|^2 = 1$, obeying

$$s_{n+1} = \rho_{n+1}^{-2}(s_n + a_{n+1}\tau_n), \qquad zt_{n+1} = \rho_{n+1}^{-2}(t_n + \bar{a}_{n+1}s_n),$$

where $\rho_n = \sqrt{1 - |a_n|^2}$. These in turn take care of the orthonormal polynomials φ_n of μ as $h\varphi_n = z^n(h/\bar{h})s_n + \overline{zt_n}$. Since $f_n = t_n/s_n$, Proposition 5.1 shows that $h_n = 1/(s_n - zt_n)$ is the Szegö function of a probability measure μ_n with f_n.

The following provides a trick in the proofs of the main results of this section.

Lemma 5.5. *Let X be a nontrivial linear subspace of N^+ such that*

$$\{zg \mid g \in X\} = \{g \in X \mid g(0) = 0\}.$$

If $\sigma, \tau \in X$, then $\sigma_n, \tau_n \in X$ for every $n = 0, 1, 2, \dots$, and conversely, if $\sigma_n, \tau_n \in X$ for some $n = 0, 1, 2, \dots$, then $\sigma, \tau \in X$.

Proof. Use Proposition 5.3, the iteration (5.2) and its inverse version

$$\sigma_{n-1} = \rho_n^{-2}[\sigma_n - a_n z\tau_n], \qquad \tau_{n-1} = \rho_n^{-2}[z\tau_n - \bar{a}_n\sigma_n],$$

where $\rho_n = \sqrt{1 - |a_n|^2}$. □

Remark 5.6. In the space L^2, the orthogonal complement of $z^n(h/\bar{h})H^2_- \cap H^2$ is the closure of $z^n(h/\bar{h})H^2 + H^2_-$. The case $X = H^2$ is of special interest in connection with the expansion formula in Proposition 3.5, modified as

$$h\Phi_n = h(0)\sum_{k=0}^{\infty}(z^n(h/\bar{h})[H_n^*H_n]^k 1 - H_n[H_n^*H_n]^k 1).$$

It is not difficult to see that $\sum_{k=1}^{\infty}[H_n^*H_n]^k 1$ converges in H^2 if and only if $\sigma, \tau \in H^2$. In this case, the above modified formula reduces to (5.3) with

$$\sigma_n = h(0)\sum_{k=0}^{\infty}[H_n^*H_n]^k 1, \qquad \tau_n = -h(0)\sum_{k=0}^{\infty}[H_n^*H_n]^k H_n^*\bar{z},$$

which solve $(I - H_n^*H_n)\sigma_n = h(0)$ and $(I - H_n^*H_n)\tau_n = -h(0)H_n^*\bar{z}$.

Here is the strong Szegö theorem [Sz2] (more precisely, Ibragimov's version [I], supplemented with a theorem of Golinskii–Ibragimov [GI]; for textbook treatment, see [Si2, Chapter 6]) in terms of a and γ.

Theorem 5.7. *For a nontrivial probability measure μ, the following are equivalent:*

$$\text{(i)} \quad \sum_{j=1}^{\infty} j|a_j|^2 < \infty; \qquad \text{(ii)} \quad \mu \textit{ satisfies } (1.2) \textit{ and } \sum_{j=1}^{\infty} j|\gamma_j|^2 < \infty.$$

Proof. Write $L_j(w) = \int_{\mathbb{T}} e^{-ij\theta}\log w\, dm$. First, assume (i). Then, by the Golinskii–Ibragimov theorem, $\mu_s = 0$, $\log w \in L^1$ and $\sum_{j=-\infty}^{\infty}|j||L_j(w)|^2 < \infty$, which implies $\sum_{j=1}^{\infty}|j||\gamma_j|^2 < \infty$ (see [Si2, Proposition 6.2.6]). Therefore, (ii) follows. Conversely, assume (ii). Then it can be shown that $\|H_n^*H_n\| \le \sum_{j=n+1}^{\infty} j|\gamma_j|^2$ in the operator norm on H^2, by handling

$$(H_n^*H_n g)(z) = \sum_{j=0}^{\infty}\left\{\sum_{k=0}^{\infty}\left(\sum_{l=n+1}^{\infty}\gamma_{j+l}\bar{\gamma}_{k+l}\right)g_k\right\}z^j,$$

where $g = \sum_{j=0}^{\infty} g_j z^j \in H^2$. Hence, one may take n large enough so $\|H_n^*H_n\| < 1$, making $\sum_{k=0}^{\infty}[H_n^*H_n]^k 1$ converge absolutely in H^2. Since this gives $\sigma_n, \tau_n \in H^2$, Proposition 5.1 and Lemma 5.5 find that $h^{-1} = \sigma - z\tau \in H^2$. Therefore, $w^{-1} \in L^1$. The results of Ibragimov–Solev [IS, Theorems 1 and 2] now implies $w^{-1} = |p|^2 v$, where p is a polynomials with roots on $\mathbb{T}$ and v is a nonnegative function such that $\sum_{j=-\infty}^{\infty}|j||L_j(v)|^2 < \infty$, as well as $v \in L^1$ (see [Si2, Proposition 6.1.5]). However, since $w \in L^1$, p is constant, so that $\sum_{j=-\infty}^{\infty}|j||L_j(w)|^2 < \infty$. Thus, by Ibragimov's theorem, (i) holds. □

The next is a similar extension for Baxter's theorem [B] (see [Si2, Chapter 5] for textbook treatment). Let ν be a *Beurling weight*, which is placed on $\mathbb{Z}$ in such a way that $\nu_j \ge 1$, $\nu_{-j} = \nu_j$ and $\nu_{j+k} \le \nu_j\nu_k$ for all $j, k \in \mathbb{Z}$. It is called a *strong Beurling weight* if $\inf_{j\in\mathbb{Z}}(j^{-1}\log\nu_j) = 0$. For example, $\nu_j = (|j|+1)^\alpha$ $(j \in \mathbb{Z})$ is a strong Beurling weight for $\alpha \ge 0$.

Theorem 5.8. *Let ν be a strong Beurling weight. Then, for a nontrivial probability measure μ, the following are equivalent:*

$$\text{(i)} \quad \sum_{j=1}^{\infty} \nu_j|a_j| < \infty; \qquad \text{(ii)} \quad \mu \textit{ satisfies } (1.2) \textit{ and } \sum_{j=1}^{\infty} \nu_j|\gamma_j| < \infty.$$

Proof. Write $\mathfrak{A}$ for the Beurling algebra of L^∞-functions $g = \sum_{j=-\infty}^{\infty} g_j e^{ij\theta}$ with $\|g\|_\nu = \sum_{j=-\infty}^{\infty}\nu_j|g_j| < \infty$. Assume (i). Then $\mu_s = 0$, $\log w \in L^1$ and $h, h^{-1} \in \mathfrak{A}$ (see [Si2, Theorem 5.2.2]). So, μ obeys (1.2) and $h/\bar{h} \in \mathfrak{A}$, whence (ii). Conversely, assume (ii). Then $H_n^*H_n$ satisfies $\|H_n^*H_n\| \le (\sum_{j=n+1}^{\infty}\nu_j|\gamma_j|)^2$ in $\mathfrak{A}$-norm. Hence, in the same way as the above proof, one has $\sigma, \tau \in \mathfrak{A}$. Thus, by Proposition 5.1, $h^{-1} \in \mathfrak{A}$ and $h \in L^\infty$. These imply (i) (see [Si2, Theorem 5.2.2]). □

References

[AAK] V. M. Adamjan,D. Z. Arov and M. G. Krein, Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and I. Schur, Funkcional. Anal. i Priložen. **2** (1968), 1–17; *Functional Anal. Appl.* **2** (1968), 269–281.

[B] G. Baxter, A convergence equivalence related to polynomials orthogonal on the unit circle, *Trans. Amer. Math. Soc.* **99** (1961), 471–487.

[BIK] N. H. Bingham, A. Inoue and Y. Kasahara, An explicit representation of Verblunsky coefficients, *Statist. Probab. Lett.*, in press.

[BJH] P. Bloomfield, N. P. Jewell and E. Hayashi, Characterizations of completely nondeterministic stochastic processes, *Pacific J. Math.* **107** (1983), 307–317.

[D] P. L. Duren, *Theory of H^p spaces*, Academic Press, New York, 1970.

[DM] H. Dym and H. P. McKean, *Gaussian processes, function theory, and the inverse spectral problem*, Academic Press, New York, 1976.

[G] J. B. Garnett *Bounded analytic functions.* Academic Press, New York, 1981.

[GI] B. L. Golinskii and I. A. Ibragimov, A limit theorem of G. Szegö, *Izv. Akad. Nauk SSSR Ser. Mat* 35 (1971), 408–427, [Russian]

[H] P. R. Halmos, *A Hilbert space problem book.* Springer-Verlag, New York, 1982.

[Ha] E. Hayashi, The solution sets of extremal problems in H^1, *Proc. Amer. Math. Soc.* **93** (1985), 690–696.

[Ho] K. Hoffman, *Banach spaces of analytic functions*, Prentice-Hall, 1962

[I] I. A. Ibragimov, A theorem of Gabor Szegö, *Mat. Zametki* **3** (1968), 693–702 [Russian].

[IS] I. A. Ibragimov and V. N. Solev, A condition for the regularity of a Gaussian stationary process, *Dokl. Akad. Nauk SSSR* bf 185 (1969), 509–512 [Russian]. English transl. in *Soviet Math. Dokl.* **10** (1969), 371–375.

[In1] A. Inoue, Asymptotics for the partial autocorrelation function of a stationary process, *J. Anal. Math.* **81** (2000), 65–109.

[In2] A. Inoue, Asymptotic behavior for partial autocorrelation functions of fractional ARIMA processes, *Ann. Appl. Probab.* **12** (2002), 1471–1491.

[In3] A. Inoue, AR and MA representation of partial autocorrelation functions, with applications, *Probab. Theory Related Fields* **140** (2008), 523-551.

[IK1] A. Inoue and Y. Kasahara, Partial autocorrelation functions of the fractional ARIMA processes with negative degree of differencing, *J. Multivariate Anal.* **89** (2004), 135–147.

[IK2] A. Inoue and Y. Kasahara, Explicit representation of finite predictor coefficients and its applications, *Ann. Statist.* **34** (2006), no. 2, 973–993.

[LM] N. Levinson and H. P. McKean, Weighted trigonometrical approximation on R^1 with application to the germ field of a stationary Gaussian noise. *Acta Math.* **112** (1964), 99–143.

[dLR] K. de Leeuw and W. Rudin, Extreme points and extremum problems in H_1, *Pacific J. Math.* **8** (1958), 467–485.

[N1] T. Nakazi, Exposed points and extremal problems in H^1, *J. Funct. Anal.* **53** (1983), 224-230.

[N2] T. Nakazi, Kernels of Toeplitz operators, *J. Math. Soc. Japan* **38** (1986), 607–616.

[Ne] Z. Nehari, On bounded bilinear forms, *Ann. Math.* **65** (1957), 153-162.

[P] V. V. Peller, *Hankel operators and their applications.* Springer-Verlag, New York, 2003.

[PS] A. Poltoratski and D. Sarason Aleksandrov-Clark measures, *Recent advances in operator-related function theory*, 1–14, *Contemp. Math.* **393**, Amer. Math. Soc., Providence, RI, 2006.

[S1] D. Sarason, *Function theory on the unit circle,* Notes for lectures given at a Conference at Virginia Polytechnic Institute and State University, 1978.

[S2] D. Sarason, Exposed points in H^1, I, *Oper. Theory Adv. Appl.* **41** (1989), 485–496.

[S3] D. Sarason, *Sub-Hardy Hilbert spaces in the unit disk,* Wiley, 1994.

[Se] A. Seghier, Prédiction d'un processus stationnaire du second ordre de covariance connue sur un intervalle fini, *Illinois J. Math.* **22** (1978), no. 3, 389–401.

[Si1] B. Simon, *Orthogonal polynomials on the unit circle, Part 1. Classical theory,* American Mathematical Society, Providence, RI, 2005.

[Si2] B. Simon, *Orthogonal polynomials on the unit circle, Part 2. Spectral theory,* American Mathematical Society, Providence, RI, 2005.

[Si3] B. Simon, *Szegö's theorem and its descendants. Spectral theory for L^2 perturbations of orthogonal polynomials,* Princeton University Press, Princeton, NJ, 2011.

[Sz1] G. Szegö, *Orthogonal Polynomials,* American Mathematical Society, New York, 1939

[Sz2] G. Szegö, On certain Hermitian forms associated with the Fourier series of a positive function. *Comm. Sém. Math. Univ. Lund* **1952** (1952), Tome Supplementaire, 228–238.

[V1] S. Verblunsky, On positive harmonic functions: A contribution to the algebra of Fourier series, *Proc. London Math. Soc.* **38** (1935), 125–157.

[V2] S. Verblunsky, On positive harmonic functions (second paper), *Proc. London Math. Soc.* **40** (1936), 290–320.

[Y] K. Yabuta, Some uniqueness theorems for $H^p(U^n)$ functions. *Tôhoku Math. J.* **24** (1972), 353–357.

[Yo] R. Younis, Hankel operators and extremal problems in H^1, *Integral Equations Operator Theory* **9** (1986), 893–904.

Department of Mathematics, Hokkaido University, Sapporo 060-0810, Japan
E-mail address: `y-kasa@math.sci.hokudai.ac.jp`

Department of Mathematics, Imperial College London, London SW7 2AZ
E-mail address: `n.bingham@ic.ac.uk`