The Annals of Probability
2003, Vol. 31, No. 1, 93–114

PREDICTION OF WEAKLY STATIONARY SEQUENCES ON POLYNOMIAL HYPERGROUPS

BY VOLKER HÖSEL AND RUPERT LASSER

GSF-National Research Center for Environment and Health

We investigate random sequences $(X_n)_{n\in\mathbb{N}_0}$ with spectral representation based on certain orthogonal polynomials, that is, random sequences that are weakly stationary with respect to polynomial hypergroups. We present various situations where one meets this kind of sequence. The main topic is on the one-step prediction. In particular, it is examined when the mean-squared error tends to zero. For many cases we present a complete solution for the problem of $(X_n)_{n\in\mathbb{N}_0}$ being asymptotically deterministic.

1. Introduction. The study of stochastic processes indexed by hypergroups was mainly motivated by the fact that averages of weakly stationary random sequences are in general not weakly stationary and thus classical theory does not apply. But, for a large class of averaging procedures one gets sequences of Karhunen type [cf. (10) below] and a covariance structure which allows stochastic analysis.

For example, certain statistical estimates $X_n = \sum_{k=-n}^{n} a_{n,k} Y_k$, $n \in \mathbb{N}_0$, of the constant mean M of a weakly stationary random sequence $(Y_n)_{n\in\mathbb{Z}}$ fulfill

$$E\big((X_n - M)\overline{(X_m - M)}\big) = \operatorname{cov}(X_n, X_m) = \sum_{k=|n-m|}^{n+m} g(n,m;k)\operatorname{cov}(X_k, X_0)$$

for all $n, m \in \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$. The coefficients $g(n,m;k)$ are linearization coefficients of orthogonal polynomial systems depending on the choice of the averaging coefficients $a_{n,k}$. More precisely [see Hösel and Lasser (1992)], if we select $a_{n,k} = a_{n,-k}$ for $k = -n, \dots, n$ with $\sum_{k=-n}^{n} a_{n,k} = 1$ such that the trigonometric polynomials

$$R_n(\cos t) = \sum_{k=-n}^{n} a_{n,k} e^{ikt}$$

form an orthogonal polynomial sequence $(R_n)_{n\in\mathbb{N}_0}$, and if the coefficients $g(n,m;k)$ in

$$R_n(x)R_m(x) = \sum_{k=|n-m|}^{n+m} g(n,m;k) R_k(x)$$

Received July 2001; revised December 2001.
AMS 2000 subject classifications. 60G10, 60G25, 43A62.
Key words and phrases. Prediction, random sequences, orthogonal polynomials, hypergroups.

are nonnegative, the sequence $(X_n)_{n\in\mathbb{N}_0}$ is weakly stationary with respect to the polynomial hypergroup induced by $(R_n)_{n\in\mathbb{N}_0}$.

The idea of studying random fields over hypergroups goes back to Lasser and Leitner (1989). Further contributions on this topic are contained in Leitner (1991), Lasser and Leitner (1990), Hösel and Lasser (1992), Bloom and Heyer [(1995), pages 546–552], Kakihara [(1997), pages 237–241], Rao (1989), Blower (1996), Heyer (1991, 2000) and Hösel (1998).

We start by recalling the precise definition of weak stationarity with respect to $(R_n)_{n\in\mathbb{N}_0}$. Then we present further situations where we meet stochastic sequences of such type. Our main purpose is to extend the results of Hösel and Lasser (1992) considerably, dealing with prediction problems. In this paper we present for many cases a complete solution of $(X_n)_{n\in\mathbb{N}_0}$ being asymptotically deterministic.

Note how our results could be applied in the scenario of classical weakly stationary processes $(Y_n)_{n\in\mathbb{Z}}$: the error of predicting Y_{n+1} from $Y_0, \dots, Y_n$ may not tend to zero with growing n. But, a variety of averages of this process are asymptotically deterministic (the prediction error tends to zero) and we even know the respective convergence orders.

2. Basic facts on polynomial hypergroups. Throughout this paper $(R_n)_{n\in\mathbb{N}_0}$ is a fixed polynomial sequence that induces a polynomial hypergroup on $\mathbb{N}_0$; that means that $(R_n(x))_{n\in\mathbb{N}_0}$ is a polynomial sequence with degree $(R_n) = n$, that is orthogonal with respect to a probability measure $\pi \in M^1(\mathbb{R})$. We assume that $R_n(1) = 1$ and that all the linearization coefficients $g(m, n; k)$, defined by

$$R_m(x)R_n(x) = \sum_{k=|n-m|}^{n+m} g(m, n; k)R_k(x) \tag{1}$$

are nonnegative. Then we denote the convex combination of point measures ε_k on $\mathbb{N}_0$ by

$$\varepsilon_m * \varepsilon_n := \sum_{k=|n-m|}^{n+m} g(m, n; k)\varepsilon_k$$

and call $\varepsilon_m * \varepsilon_n$ the convolution of $n \in \mathbb{N}_0$ and $m \in \mathbb{N}_0$. With $*$ as convolution, the identity map as involution and 0 as unit element, $\mathbb{N}_0$ becomes a commutative hypergroup that is called a polynomial hypergroup; see Bloom and Heyer (1995). Section 3.3 of Bloom and Heyer (1995) contains a long list of polynomial hypergroups. Every character on the polynomial hypergroup $\mathbb{N}_0$ is given by $\alpha_x : \mathbb{N}_0 \to \mathbb{R}$, $\alpha_x(n) = R_n(x)$, where $x \in D_s$ with

$$D_s = \bigl\{x \in \mathbb{R} : \{R_n(x) : n \in \mathbb{N}_0\} \text{ is bounded}\bigr\}. \tag{2}$$

The character space $\widehat{\mathbb{N}_0}$ is homeomorphic to D_s; see Bloom and Heyer (1995). In particular D_s is a compact subset of $\mathbb{R}$.

The orthogonal polynomials $R_n(x)$ are determined by their three-term recurrence relation

$$R_0(x) = 1, \qquad R_1(x) = \frac{1}{a_0}(x - b_0)$$

and for $n \in \mathbb{N}$,

$$R_1(x)R_n(x) = a_n R_{n+1}(x) + b_n R_n(x) + c_n R_{n-1}(x), \tag{3}$$

where $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, $(c_n)_{n\in\mathbb{N}}$ are three real-valued sequences with $a_n, c_n > 0$, $b_n \geq 0$ and $a_n + b_n + c_n = 1$. For $a_0, b_0 \in \mathbb{R}$ we assume $a_0 > 0$ and $a_0 + b_0 = 1$.

The linearization coefficients $g(m, n; k)$ can be directly calculated from the recurrence coefficients a_n, b_n, c_n; see Lasser (1983). The Haar measure h on the (discrete) polynomial hypergroup is given by $h(n) = g(n, n; 0)^{-1}$ and also determined by the weights

$$h(0) = 1, \quad h(1) = 1/c_1, \quad h(n) = \frac{\prod_{k=1}^{n-1} a_k}{\prod_{k=1}^{n} c_k}, \qquad n = 2, 3, \ldots . \tag{4}$$

The support of the orthogonalization measure π is contained in D_s. In fact we have

$$\operatorname{supp} \pi \subseteq D_s \subseteq [1 - 2a_0, 1] \tag{5}$$

and π is the Plancherel measure on D_s, that is,

$$\int_{D_s} R_n(x) R_m(x)\, d\pi(x) = \begin{cases} 0, & \text{if } n \neq m, \\ \dfrac{1}{h(n)}, & \text{if } n = m. \end{cases}$$

We essentially use a Bochner theorem for polynomial hypergroups; see Theorem 4.1.6 of Bloom and Heyer (1995) which states that for each bounded positive definite sequence $(d(n))_{n\in\mathbb{N}_0}$ there exists a unique $\mu \in M^+(D_s)$ such that

$$d(n) = \int_{D_s} R_n(x)\, d\mu(x). \tag{6}$$

Hereby positive definiteness stands for

$$\sum_{i,j=1}^{n} \lambda_i \overline{\lambda_j} \varepsilon_{m_i} * \varepsilon_{m_j}(d) \geq 0 \tag{7}$$

for all $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and $m_1, \ldots, m_n \in \mathbb{N}_0$.

To clarify we note that

$$\varepsilon_l * \varepsilon_m(d) = \sum_{k=|m-l|}^{m+l} g(l, m; k)\, d(k).$$

DEFINITION 2.1. A sequence $(X_n)_{n\in\mathbb{N}_0}$ of square integrable (complex-valued) random variables on a probability space (Ω, P) is called *weakly stationary on the polynomial hypergroup* $\mathbb{N}_0$ *induced by* $(R_n)_{n\in\mathbb{N}_0}$ provided the function $d(m,n) := E(X_m\overline{X_n})$ is bounded and fulfills

$$d(m,n) = \sum_{k=|m-n|}^{m+n} g(m,n;k)d(k,0). \tag{8}$$

We will write $d(m)$ instead of $d(m,0)$. Notice that we do not make any assumptions on the mean values $E(X_n)$. Obviously $d(m,n)$ is the usual covariance only if $E(X_n) = 0$ for all $n \in \mathbb{N}_0$. Now it is easily shown that $(d(n))_{n\in\mathbb{N}_0}$ is a bounded positive definite sequence in the sense of (7). Hence there exists a unique bounded positive Borel measure μ on D_s such that for all $n \in \mathbb{N}_0$,

$$E(X_k\overline{X_0}) = d(k) = \int_{D_s} R_k(x)\, d\mu(x), \qquad k \in \mathbb{N}_0. \tag{9}$$

We call μ the spectral measure of $(X_n)_{n\in\mathbb{N}_0}$. Combining (1) and (8) we get from (9), for all $n,m \in \mathbb{N}_0$,

$$E(X_m\overline{X_n}) = d(m,n) = \int_{D_s} R_m(x)R_n(x)\, d\mu(x), \tag{10}$$

showing that $(X_n)_{n\in\mathbb{N}_0}$ belongs to the Karhunen class. Moreover, a Cramér representation theorem is valid. A straightforward modification of known arguments yields an orthogonal stochastic measure $Z : \mathcal{B} \to L^2(P)$, $\mathcal{B}$ being the Borel σ-algebra on D_s such that

$$X_k = \int_{D_s} R_k(x)\, dZ(x), \tag{11}$$

where $\|Z(A)\|_2^2 = \mu(A)$ for all $A \in \mathcal{B}$.

In fact, the following characterizations of weak stationarity hold.

THEOREM 2.1. *For* $(X_n)_{n\in\mathbb{N}_0} \subset L^2(\Omega, P)$ *the following statements are equivalent*:

(i) $(X_n)_{n\in\mathbb{N}_0}$ *is weakly stationary on the polynomial hypergroup* $\mathbb{N}_0$ *induced by* $(R_n)_{n\in\mathbb{N}_0}$.

(ii) $E(X_m\overline{X_n}) = \int_{D_s} R_m(x)R_n(x)\, d\mu(x)$ *for every* $n,m \in \mathbb{N}_0$, *where* μ *is a bounded positive Borel measure on* D_s.

(iii) $X_k = \int_{D_s} R_k(x)\, dZ(x)$ *for every* $k \in \mathbb{N}_0$, *where* Z *is an orthogonal stochastic measure on* D_s.

PROOF. (i) $\Rightarrow$ (ii) has already been shown. (ii) $\Rightarrow$ (iii) is based on the isometric isomorphism Φ between $L^2(D_s,\mu)$ and $H = \mathrm{span}\{X_n : n \in \mathbb{N}_0\}^- \subseteq L^2(\Omega, P)$ determined by $\Phi(R_n) = X_n$. The stochastic measure is defined by

$Z(A) = \Phi(\chi_A)$, A being a Borel subset of D_s. The construction parallels the classical case [see Shiryayev (1984), pages 395–403].

Finally, assuming (iii) we get for every $m, n \in \mathbb{N}_0$,

$$E(X_m\overline{X_n}) = \int_{D_s} R_m(x)R_n(x)\,d\mu(x),$$

where μ is defined by $\mu(A) = \|Z(A)\|_2^2$, $A \in \mathcal{B}$. The linearization of $R_m(x)R_n(x)$ gives

$$E(X_m\overline{X_n}) = \sum_{k=|n-m|}^{m+n} g(m,n;k)E(X_k\overline{X_0}). \qquad \square$$

3. Occurrence of weakly stationary sequences on polynomial hypergroups. In Lasser and Leitner (1989) and Hösel and Lasser (1992) we have shown that statistical estimates X_n of the constant mean of a weakly stationary random sequence form a weakly stationary sequence on a polynomial hypergroup $\mathbb{N}_0$. We shall now describe further situations where one meets such random sequences.

3.1. *Real and imaginary parts.* Let $(Y_n)_{n\in\mathbb{Z}}$ be a weakly stationary complex-valued process with symmetry, that is $Y_{-n} = \overline{Y_n}$. The random sequence $(U_n)_{n\in\mathbb{N}_0}$ of the real parts $U_n = \operatorname{Re} Y_n = \frac{1}{2}(Y_n + Y_{-n})$, $n \in \mathbb{N}_0$, is no longer weakly stationary in the usual sense. However, one can easily check that

$$E(U_mU_n) = \tfrac{1}{2}E(U_{n+m}U_0) + \tfrac{1}{2}E(U_{|n-m|}U_0).$$

That means $(U_n)_{n\in\mathbb{N}_0}$ is a weakly stationary random sequence on the polynomial hypergroup $\mathbb{N}_0$ induced by the Chebyshev's polynomials $T_n(x)$ of the first kind. $T_n(x)$ are orthogonal on $[-1,1]$ with respect to $d\pi(x) = \pi^{-1}(1-x^2)^{-1/2}\,dx$. They belong to the class of Jacobi polynomials $R_n^{(\alpha,\beta)}(x)$; see Bloom and Heyer [(1995), 3.3.1], $\alpha = -\frac{1}{2}$, $\beta = -\frac{1}{2}$.

Denote the imaginary part of Y_n by

$$V_n = \operatorname{Im} Y_n = \frac{1}{2i}(Y_n - Y_{-n}).$$

Since $V_0 = 0$ we have

$$E(U_nU_0) + iE(V_nU_0) = E(Y_n\overline{Y_0}) = \int_{-\pi}^{\pi} \cos(nt)\,d\mu(t) + i\int_{-\pi}^{\pi} \sin(nt)\,d\mu(t),$$

where $\mu \in M^+(]-\pi,\pi])$ is the spectral measure of $(Y_n)_{n\in\mathbb{Z}}$.

The random sequence $(U_n)_{n\in\mathbb{N}_0}$ is weakly stationary with respect to $T_n(x) = \cos(nt)$, where $x = \cos t$ for $x \in [-1,1]$, $t \in [0,\pi]$. Hence we have a unique spectral representation

$$E(U_nU_0) = \int_0^{\pi} \cos(nt)\,d\nu(t),$$

$\nu \in M^+([0,\pi])$. Now it is clear that $\nu = \mu|[0,\pi] + \mu_1|]0,\pi[$, where μ_1 is the image measure of μ under the mapping $t \to -t$. Further we observe that

$$\begin{aligned} E(U_m U_n) + E(V_m V_n) &= \operatorname{Re} E(Y_m \overline{Y_n}) \\ &= \int_{-\pi}^{\pi} \cos((m-n)t)\, d\mu(t) = \int_0^{\pi} \cos((m-n)t)\, d\nu(t) \\ &= \int_0^{\pi} \cos(mt)\cos(nt)\, d\nu(t) + \int_0^{\pi} \sin(mt)\sin(nt)\, d\nu(t). \end{aligned}$$

Since $E(U_m U_n) = \int_0^{\pi} \cos(mt)\cos(nt)\, d\nu(t)$, it follows that

$$E(V_m V_n) = \int_0^{\pi} \sin(mt)\sin(nt)\, d\nu(t).$$

Define, for $n \in \mathbb{N}_0$,

$$X_n := \frac{1}{n+1} V_{n+1} = \frac{1}{n+1} \operatorname{Im} Y_{n+1}.$$

Then

$$E(X_m X_n) = \int_0^{\pi} \frac{\sin((m+1)t)}{(m+1)\sin t} \frac{\sin((n+1)t)}{(n+1)\sin t} (\sin t)^2\, d\nu(t).$$

By Theorem 2.1 we see that $(X_n)_{n\in\mathbb{N}_0}$ is a weakly stationary random sequence on the polynomial hypergroup $\mathbb{N}_0$ induced by the Chebyshev's polynomials $U_n(x)$ of the second kind. Notice that $U_n(x) = R_n^{(1/2,1/2)}(x) = \frac{\sin((n+1)t)}{(n+1)\sin t}$, $x = \cos t$.

3.2. *Stationary increments.* A random sequence $(Y_n)_{n\in\mathbb{Z}}$ is called a sequence with stationary increments if all $E(Y_{n+k} - Y_k)$ depend only on $n \in \mathbb{Z}$ and $E((Y_{n_1+k} - Y_k)(\overline{Y}_{n_2+k} - \overline{Y_k}))$ depend only on $n_1, n_2 \in \mathbb{Z}$; see Yaglom [(1987), Section 23]. It is readily seen [cf. Yaglom (1987), equation (4.227)] that

$$Y_{n+k} - Y_k = \int_{-\pi}^{\pi} e^{ikt} \frac{e^{int} - 1}{e^{it} - 1}\, dZ(t),$$

where Z is an orthogonal stochastic measure on $]-\pi,\pi]$. Hence we get for arbitrary $n \in \mathbb{N}_0$,

$$Y_{n+1} - Y_{-n} = \int_0^{\pi} \frac{\sin((2n+1)(t/2))}{\sin(t/2)}\, d\tilde{Z}(t),$$

where $\tilde{Z}$ is the orthogonal stochastic measure on $[0,\pi]$ defined by $\tilde{Z} = Z|[0,\pi] + Z_1$, $Z_1(]a,b]) = Z([-b,-a[)$ for $]a,b] \subseteq]0,\pi]$ and $Z_1(\{0\}) = 0$. Putting

$$X_n := \frac{1}{2n+1}(Y_{n+1} - Y_{-n}),$$

we have a weakly stationary sequence $(X_n)_{n\in\mathbb{N}_0}$ on the polynomial hypergroup $\mathbb{N}_0$ induced by the Jacobi polynomials $R_n^{(1/2,-1/2)}(x) = \frac{\sin((2n+1)(t/2))}{(2n+1)\sin(t/2)}$, $x = \cos t$.

3.3. *Coefficients of random orthogonal expansions for density estimation.* Suppose that the distribution of a random variable X is absolutely continuous with respect to a positive Borel measure π on the interval $[-1,1]$; that is, $P(X \in A) = \int_A f(x)\,d\pi(x)$, where $f \in L^1([-1,1],\pi)$, $f \geq 0$. Consider the sequence $(p_n)_{n\in\mathbb{N}_0}$ of polynomials that are orthonormal with respect to π, and let as before $R_n(x) = p_n(x)/p_n(1)$. Further, assume that $(R_n)_{n\in\mathbb{N}_0}$ induces a polynomial hypergroup on $\mathbb{N}_0$. Given independent random variables $X_1, X_2, \dots, X_N$ equally distributed as X, the unknown density function $f(x)$ can be estimated by the random orthogonal expansion

$$f_N(\omega; x) := \sum_{k=0}^{q(N)} a_{N,k} c_{N,k}(\omega) p_k(x).$$

Here $q(N)$ is the truncation point and $a_{N,k}$ are numerical coefficients to be chosen in an appropriate manner; see Lasser, Obermaier and Strasser (1993) and Devroye and Györfi (1985). The random coefficients are given by

$$c_{N,k}(\omega) := \frac{1}{N} \sum_{j=1}^{N} p_k\big(X_j(\omega)\big).$$

Define $C_k := c_{N,k}/p_k(1)$ for each $k \in \mathbb{N}_0$. The random sequence $(C_k)_{k\in\mathbb{N}_0}$ is weakly stationary on the polynomial hypergroup $\mathbb{N}_0$ induced by $(R_k)_{k\in\mathbb{N}_o}$. In fact, we have

$$C_n = \frac{1}{N} \sum_{j=1}^{N} R_n(X_j)$$

and hence

$$\begin{aligned}
E(C_m C_n) &= \frac{1}{N^2} \sum_{i,j=1}^{N} E\big(R_m(X_i) R_n(X_j)\big) \\
&= \frac{1}{N^2} \sum_{i,j=1}^{N} \int_{-1}^{1} R_m(x) R_n(x) f(x)\,d\pi(x) \\
&= \int_{-1}^{1} R_m(x) R_n(x) f(x)\,d\pi(x) \\
&= \sum_{k=|n-m|}^{n+m} g(m,n;k) \int_{-1}^{1} R_k(x) R_0(x) f(x)\,d\pi(x) \\
&= \sum_{k=|n-m|}^{n+m} g(n,m;k) E(C_k C_0).
\end{aligned}$$

Estimating the density function f by random orthogonal expansion is hence strongly related to estimating the spectral measure $f\pi$ of the random sequence $(C_k)_{k\in\mathbb{N}_0}$, where $(C_k)_{k\in\mathbb{N}_0}$ is weakly stationary with respect to $R_k = p_k/p_k(1)$.

3.4. *Stationary radial stochastic processes on homogeneous trees.* We denote by T a homogeneous tree of degree $q \geq 1$ with metric d. Let G be the isometry group of T, $t_0 \in T$ an arbitrary but fixed knot of T and let H be the stabilizer of t_0 in G. We identify T with the coset space G/H and call a mapping on T radial if it depends only on $|t| = d(t, t_0)$. A square integrable stochastic process $(X_t)_{t \in T}$ is called stationary, if there exists a function $\phi : \mathbb{N}_0 \to \mathbb{R}$ such that

$$E(X_s \overline{X_t}) = \phi(d(s,t))$$

for all s and t in T [cf. Arnaud (1994)]. It is known [Arnaud (1994)] that the following spectral representation is true:

$$E(X_s \overline{X_t}) = \int_{-1}^{1} R_{d(s,t)}(x)\, d\mu(x),$$

where $(R_n(x))_{n \in \mathbb{N}_0}$ are the orthogonal polynomials (also called Cartier–Dunau polynomials) corresponding to homogeneous trees of degree q. The polynomials $R_n(x)$ are determined by

$$R_0(x) = 1, \qquad R_1(x) = x$$

and

$$R_1(x) R_n(x) = \frac{q}{q+1} R_{n+1}(x) + \frac{1}{q+1} R_{n-1}(x), \qquad n \in \mathbb{N}.$$

They induce a polynomial hypergroup on $\mathbb{N}_0$ [see Lasser (1983)].

We assume, moreover, that the stationary stochastic process $(X_t)_{t \in T}$ is radial; that is, $X_t = X_s$ if $|t| = |s|$. Putting $X_n := X_t$ whenever $|t| = n$ we get a well-defined random sequence. We have

$$\int_H R_{d(h(s),t)}(x)\, d\beta(h) = R_{|s|}(x) R_{|t|}(x) = \sum_{k=||s|-|t||}^{|s|+|t|} g(|s|, |t|, k) R_k(x),$$

where β is the Haar measure on the compact stabilizer H, $h(s)$ the action of $h \in H$ on $T = G/H$, and the explicit form of the linearization coefficients $g(|s|, |t|, k)$ for $m, n \in \mathbb{N}, m \leq n$ given by

$$\begin{aligned}
g(m, n, n-m) &= \frac{1}{(q+1)q^{m-1}}, \\
g(m, n, n+m) &= \frac{q}{q+1}, \\
g(m, n, n+m-2k) &= \frac{q-1}{(q+1)q^k} \qquad && \text{for } k \in \{1, 2, \ldots, m-1\}, \\
g(m, n, n+m-k) &= 0 && \text{for } k \in \{1, 3, \ldots, 2m-1\}
\end{aligned}$$

[cf., e.g., Lasser (1983), Voit (1990), Cowling, Meda and Setti (1998)].

From the above spectral representation we can derive

$$E(X_m\overline{X_n}) = \int_{-1}^{1} R_m(x)R_n(x)\,d\mu(x).$$

By Theorem 1.1 we obtain that $(X_n)_{n\in\mathbb{N}_0}$ is weakly stationary on the polynomial hypergroup $\mathbb{N}_0$ induced by the Cartier–Dunau polynomials $(R_n)_{n\in\mathbb{N}_0}$.

4. One-step prediction. Having presented situations where we meet random sequences that are weakly stationary on the polynomial hypergroup $n \in \mathbb{N}_0$ we now study the prediction problem. In the sequel fix an orthogonal polynomial sequence $(R_n)_{n\in\mathbb{N}_0}$ which induces a polynomial hypergroup on $\mathbb{N}_0$. Let $(X_n)_{n\in\mathbb{N}_0}$ be a weakly stationary random sequence on $\mathbb{N}_0$ induced by $(R_n)_{n\in\mathbb{N}_0}$. The problem in one-step prediction can be formulated as follows. Given $n \in \mathbb{N}_0$ denote the linear space generated by $X_0,\dots,X_n$ by $H_n := \operatorname{span}\{X_0,\dots,X_n\} \subseteq \operatorname{span}\{X_k : k \in \mathbb{N}_0\}^- \subseteq L^2(P)$. We want to characterize the prediction $\hat{X}_{n+1} \in H_n$ of X_{n+1} with the minimum property

$$\|\hat{X}_{n+1} - X_{n+1}\|_2 = \min\{\|Y - X_{n+1}\|_2 : Y \in H_n\}. \tag{12}$$

It is well known that $\hat{X}_{n+1} = P_{H_n}X_{n+1}$, where P_{H_n} is the orthogonal projection from $L^2(P)$ to H_n. The problem is to determine the coefficient $b_{n,k}$, $k = 0,\dots,n$, in the representation

$$\hat{X}_{n+1} = \sum_{k=0}^{n} b_{n,k}X_k \tag{13}$$

and to decide whether the prediction error

$$\delta_n := \|\hat{X}_{n+1} - X_{n+1}\|_2 \tag{14}$$

converges to zero as n tends to infinity. We call $(X_n)_{n\in\mathbb{N}_0}$ *asymptotically R_n-deterministic* if $\delta_n \to 0$ as $n \to \infty$.

From Hilbert space theory we know that $\hat{X}_{n+1}$ can be characterized by the property

$$E\big((X_{n+1} - \hat{X}_{n+1})\overline{Y}\big) = \langle X_{n+1} - \hat{X}_{n+1}, Y\rangle = 0 \tag{15}$$

for all $Y \in H_n$. From (15) we get for $b = (b_{n,0},\dots,b_{n,n})^T$ the linear equation

$$\Phi^T b = \varphi, \tag{16}$$

where $\varphi = (E(X_{n+1}\overline{X_0}), E(X_{n+1}\overline{X_1}),\dots,E(X_{n+1}\overline{X_n}))^T$ and Φ is the $(n+1)\times(n+1)$-matrix $\Phi = (E(X_i\overline{X_j}))_{0\le i,j\le n}$. Matrices structured like Φ can be seen as generalizations of Toeplitz-matrices with regard to the orthogonal polynomial sequence $(R_n)_{n\in\mathbb{N}_0}$. The dependency of the i,j-entry on $i-j$ in the Toeplitz case is substituted by the condition

$$\Phi_{i,j} = \varepsilon_i * \varepsilon_j(d) = \sum_{k=|i-j|}^{i+j} g(i,j;k)\,d(k) \qquad \text{with } d(k) = E(X_n\overline{X_0}).$$

Methods of fast inversion of matrices of this type are not studied in the numerical literature, as far as the authors know. A method resembling the Durbin–Levinson algorithm [cf. Brockwell and Davies (1991)] will be investigated at another place. Here we stress the theory of prediction.

In Hösel and Lasser (1992) we showed that

$$\delta_n = \frac{\sigma_{n+1}}{\rho_{n+1}}, \tag{17}$$

where $\sigma_n = \sigma_n(\pi)$ is the leading coefficient of $R_n(x) = \sigma_n x^n + \cdots$ and $\rho_n = \rho_n(\mu)$ is the positive leading coefficient of the polynomials $q_n(x) = \rho_n x^n + \cdots$, that are orthonormal with respect to the spectral measure μ. If the orthogonalization measure π satisfies the Kolmogorov–Szegö property, we can give a rather complete characterization when $\delta_n \to 0$ holds.

4.1. *Kolmogorov–Szegö class.* To each polynomial sequence $(p_n)_{n\in\mathbb{N}_0}$ orthonormal with respect to a probability measure $\nu \in M^1([-1,1])$ one can associate a unique polynomial sequence $(\psi_n)_{n\in\mathbb{N}_0}$ on $[-\pi,\pi]$ orthonormal with respect to a measure α given by

$$d\alpha(t) = |\sin t|\, d\nu(\cos t), \qquad t \in [-\pi,\pi]. \tag{18}$$

Denote the positive leading coefficients of $\psi_n(t)$ by $\rho_n(\alpha)$ and those of $p_n(t)$ by $\rho_n(\nu)$. We apply some important results on orthogonal polynomials on the unit circle. The original references are Geronimus (1960) and Szegö (1975); see also Lubinsky (1987).

If the Radon–Nikodym derivative α' of α fulfils the Kolmogorov–Szegö property, that is, $\ln(\alpha') \in L^1([-\pi,\pi])$ or equivalently

$$\int_{-\pi}^{\pi} \ln(\alpha'(t))\, dt > -\infty, \tag{19}$$

then $\rho_n(\alpha)$ converges monotonically increasingly towards the geometric mean

$$\rho(\alpha) := \exp\left(-\frac{1}{4\pi}\int_{-\pi}^{\pi} \ln(\alpha'(t))\, dt\right) \tag{20}$$

[see Lubinsky (1987), Theorem 3.4]. If the Kolmogorov–Szegö property is not valid we have $\lim_{n\to\infty} \rho_n(\alpha) = \infty$. Transferring this result to $[-1,1]$ we have [see Geronimus (1960), Theorem 9.2] the following.

PROPOSITION 4.1. *Let $(p_n)_{n\in\mathbb{N}_0}$ be an orthonormal polynomial sequence with respect to a measure $\nu \in M^1([-1,1])$, $(|\operatorname{supp}\nu| = \infty)$. Then alternatively we have*:

(i) *If $\int_{-1}^{1} \ln(\nu'(x))/\sqrt{1-x^2}\, dx > -\infty$, then there exist positive constants C_1, C_2 such that for all $n \in \mathbb{N}_0$,*

$$C_1 \le \frac{\rho_n(\nu)}{2^n} \le C_2.$$

(ii) *If $\int_{-1}^{1} \ln(\nu'(x))/\sqrt{1-x^2}\, dx = -\infty$, then $\lim_{n\to\infty} \frac{\rho_n(\nu)}{2^n} = \infty$.*

PROOF. Consider the orthonormal polynomial sequence $(\psi_n)_{n\in\mathbb{N}_0}$ on $[-\pi,\pi]$ with leading coefficients $\rho_n(\alpha)$, where α is the measure as in (18). The coefficients $\rho_n(\alpha)$ and $\rho_n(\nu)$ satisfy the inequalities [see Geronimus (1960), equation (9.9)]

$$\frac{\rho_{2n-1}(\alpha)}{2\sqrt{\pi}} \le \frac{\rho_n(\nu)}{2^n} \le \frac{\rho_{2n}(\alpha)}{\sqrt{\pi}}. \tag{21}$$

Moreover, we have

$$\int_{-\pi}^{\pi} \ln(\alpha'(t))\,dt = 2\int_{-1}^{1} \frac{\ln(\nu'(x))}{\sqrt{1-x^2}}\,dx,$$

and the assertions (i) and (ii) follow by the convergence of $\rho_n(\alpha)$ towards $\rho(\alpha)$ or towards infinity, respectively. □

For the orthonormal version $p_n(x)$ of $R_n(x)$ we have $p_n(x) = \sqrt{h(n)}R_n(x)$. Hence $\sigma_n(\pi) = \frac{\rho_n(\pi)}{\sqrt{h(n)}}$ and by (17) we obtain immediately the theorem.

THEOREM 4.2. *Let $(X_n)_{n\in\mathbb{N}_0}$ be a weakly stationary random sequence on the polynomial hypergroup $\mathbb{N}_0$ induced by $(R_n)_{n\in\mathbb{N}_0}$. Suppose that the orthogonalization measure π of $(R_n)_{n\in\mathbb{N}_0}$ has support contained in $[-1,1]$ and fulfils the Kolmogorov–Szegö property on $[-1,1]$, that is,*

$$\int_{-1}^{1} \frac{\ln(\pi'(x))}{\sqrt{1-x^2}}\,dx > -\infty. \tag{22}$$

Then $(X_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic if and only if

$$\lim_{n\to\infty} \frac{\sqrt{h(n)}\rho_n(\mu)}{2^n} = \infty. \tag{23}$$

In particular we have:

(i) *If $\lim_{n\to\infty} h(n) = \infty$, then $(X_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic without any assumption on the spectral measure μ. For the prediction errors we have*

$$\delta_n = O\left(\frac{1}{\sqrt{h(n+1)}}\right) \qquad \textit{as } n\to\infty.$$

(ii) *If $\{h(n) : n\in\mathbb{N}_0\}$ is bounded, then $(X_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic if and only if μ does not fulfil the Kolmogorov–Szegö property on $[-1,1]$.*

(iii) *If $\{h(n) : n\in\mathbb{N}_0\}$ is unbounded, then $(X_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic provided μ does not fulfil the Kolmogorov–Szegö property on $[-1,1]$.*

PROOF. We know that $\delta_n = \frac{\sigma_{n+1}(\pi)}{\rho_{n+1}(\mu)} = \frac{\rho_{n+1}(\pi)}{\sqrt{h(n+1)}\rho_{n+1}(\mu)}$.

By the Kolmogorov–Szegö property for π and Proposition 4.1 there exist positive constants C_1, C_2 such that

$$C_1 \leq \frac{\rho_{n+1}(\pi)}{2^{n+1}} \leq C_2 \qquad \text{for } n \in \mathbb{N}_0$$

and hence

$$C_1 \leq \sqrt{h(n+1)}\frac{\rho_{n+1}(\mu)}{2^{n+1}}\delta_n \leq C_2 \qquad \text{for } n \in \mathbb{N}_0.$$

Now applying Proposition 4.1 again, every statement follows. □

From the preceding theorem we know that $(X_n)_{n\in\mathbb{N}_0}$ shows a prediction behavior similar to classical weakly stationary processes provided $\{h(n) : n \in \mathbb{N}_0\}$ is bounded and the orthogonalization measure π fulfils the Kolmogorov–Szegö condition. We now prove for the case $D_s = [-1, 1]$ that if $(h(n))_{n\in\mathbb{N}_0}$ does not converge to infinity, the Kolmogorov–Szegö property of π is valid.

PROPOSITION 4.3. *Let $(R_n)_{n\in\mathbb{N}_0}$ induce a polynomial hypergroup on $\mathbb{N}_0$, and suppose that $D_s = [-1, 1]$. If π does not fulfil the Kolmogorov–Szegö property, then* $\lim_{n\to\infty} h(n) = \infty$.

PROOF. Consider the representation of $R_n(x)$ by Chebyshev's polynomials of the first kind,

$$R_n(x) = \sum_{k=0}^{n} a_{n,k} T_k(x), \tag{24}$$

where $T_k(x) = \cos(k \arccos x) = 2^{k-1}x^k + \cdots$. Since $|R_n(x)| \leq 1$ for every $x \in D_s$, $n \in \mathbb{N}_0$, we obtain, by applying the orthogonalization measure $d\nu(x) = \frac{1}{\pi}\frac{dx}{\sqrt{1-x^2}}$ of $(T_k)_{k\in\mathbb{N}_0}$,

$$1 \geq \int_{-1}^{1} R_n^2(x)\, d\nu(x) \geq a_{n,n}^2 \int_{-1}^{1} T_n^2(x)\, d\nu(x) = \frac{a_{n,n}^2}{2}.$$

Comparing the leading coefficients in (24), we get

$$a_{n,n} = \frac{\sigma_n(\pi)}{2^{n-1}} = \frac{\rho_n(\pi)}{\sqrt{h(n)}2^{n-1}},$$

and hence $2h(n) \geq \left(\rho_n(\pi)/2^{n-1}\right)^2$.

By Proposition 4.1 we have $\lim_{n\to\infty} \rho_n(\pi)/2^n = \infty$ and hence $\lim_{n\to\infty} h(n) = \infty$. □

Many examples of orthogonal polynomial sequences $(R_n)_{n\in\mathbb{N}_0}$ inducing a polynomial hypergroup can be found in Bloom and Heyer (1995); see also Lasser (1983, 1994). We consider a few more examples to determine the asymptotic behavior of the prediction error.

1. We consider $R_n^{(\alpha,\beta)}$, the Jacobi polynomials with $\alpha \geq \beta > -1$ and $\alpha + \beta + 1 \geq 0$. The orthogonalization measure is $d\pi(x) = c_{\alpha\beta}(1-x)^\alpha(1+x)^\beta\,dx$. The Haar weights (4) are
$$h(n) = \frac{(2n+\alpha+\beta+1)(\alpha+\beta+1)_n(\alpha+1)_n}{(\alpha+\beta+1)n!(\beta+1)_n}.$$
Evidently π satisfies the Kolmogorov–Szegö property. If $\alpha \neq -\frac{1}{2}$ we have $\lim_{n\to\infty} h(n) = \infty$. Hence every weakly stationary random sequence on the polynomial hypergroup $\mathbb{N}_0$ induced by $(R_n^{(\alpha,\beta)})_{n\in\mathbb{N}_0}$ is asymptotically $R_n^{(\alpha,\beta)}$-deterministic.
   For the prediction error we have
$$\delta_n = O(n^{-\alpha-1/2}) \qquad \text{as } n \to \infty.$$
If $\alpha = -\frac{1}{2}$, then $R_n^{(-1/2,-1/2)}(x) = T_n(x)$ are the Chebyshev's polynomials of the first kind. The Haar weights are $h(0) = 1$, $h(n) = 2$ for $n \in \mathbb{N}$, and Theorem 4.2(ii) can be utilized.
2. Next is $R_n^{(\nu)}(x;\alpha)$, the associated ultraspherical polynomials with $\alpha > -\frac{1}{2}$, $\nu \geq 0$. These polynomials are studied in detail in Lasser [(1994), (3)]. The Haar weights (4) are
$$h(n) = \frac{(2n+2\alpha+2\nu+1)}{4\alpha^2(2\alpha+2\nu+1)(\nu+1)_n(2\alpha+\nu+1)_n}\big((2\alpha+\nu)_{n+1} - (\nu)_{n+1}\big)^2$$
and the orthogonalization measure π on $[-1,1]$,
$$d\pi(x) = c_{\alpha\nu}g(x)(1-x^2)^\alpha\,dx$$
with
$$g(\cos t) = \big|{}_2F_1\big(\tfrac{1}{2}-\alpha, \nu; \nu+\alpha+\tfrac{1}{2}; e^{2it}\big)\big|^{-2}, \qquad x = \cos t.$$
To show that π fulfils the Kolmogorov–Szegö property it is convenient to look at the leading coefficients. From Lasser [(1994), equation (3.10)] we have
$$\rho_n(\pi) = 2^n\left(\frac{(\nu+\alpha+\frac{3}{2})_n(\nu+\alpha+\frac{1}{2})_n}{(\nu+1)_n(\nu+2\alpha+1)_n}\right)^{1/2}$$
and the asymptotic properties of the Gamma function yield
$$\lim_{n\to\infty}\frac{\rho_n(\pi)}{2^n} = \left(\frac{\Gamma(\nu+1)\Gamma(\nu+2\alpha+1)}{\Gamma(\nu+\alpha+\frac{3}{2})\Gamma(\nu+\alpha+\frac{1}{2})}\right)^{1/2}.$$
Proposition 4.1 yields the Kolmogorov–Szegö property of π. Using once more the asymptotic properties of the Gamma function we see that $h(n) = O(n^{2\alpha+1})$ and hence we have for the prediction error
$$\delta_n = O(n^{-\alpha-1/2}) \qquad \text{as } n \to \infty.$$
Note that we enlarge the domain of α compared to Hösel and Lasser [(1992), Corollary 2].

3. Finally, we have $R_n(x; \nu, \kappa)$, the Bernstein–Szegö polynomials with $\nu, \kappa \geq 0$, $\kappa - 1 < \nu < 1$. The polynomials under consideration are orthogonal with respect to the measure on $[-1, 1]$,

$$d\pi(x) = c_{\nu\kappa} \frac{dx}{g(x)\sqrt{1-x^2}},$$

where $g(x) = |\nu e^{2it} + \kappa e^{it} + 1|^2$, $x = \cos t$, is a polynomial with $g(x) > 0$ for all $x \in [-1, 1]$. By Szegö (1975) these polynomials can be represented explicitly by Chebyshev's polynomials of the first kind:

$$\begin{aligned} R_n(x; \nu, \kappa) &= \frac{1}{\nu+\kappa+1}\big(T_n(x) + \kappa T_{n-1}(x) + \nu T_{n-2}(x)\big), \qquad n \geq 2, \\ R_1(x; \nu, \kappa) &= \frac{1}{\nu+\kappa+1}\big((\nu+1)T_1(x) + \kappa T_0(x)\big), \\ R_0(x; \nu, \kappa) &= 1. \end{aligned}$$

An easy calculation shows

$$\begin{aligned} R_1(x; \nu, \kappa) R_n(x; \nu, \kappa) = {} & \frac{\nu+1}{2(\nu+\kappa+1)} R_{n+1}(x; \nu, \kappa) \\ & + \frac{\kappa}{\nu+\kappa+1} R_n(x; \nu, \kappa) \\ & + \frac{\nu+1}{2(\nu+\kappa+1)} R_{n-1}(x; \nu, \kappa) \end{aligned} \tag{25}$$

for $n \geq 3$. It is straightforward to calculate the linearization coefficients and to check directly that $(R_n(x; \nu, \kappa))_{n \in \mathbb{N}_0}$ induces a polynomial hypergroup on $\mathbb{N}_0$ provided $\nu, \kappa \geq 0$ and $\kappa - 1 < \nu < 1$. Since the recurrence coefficients in (25) are constant for $n \geq 3$ the Haar weights are bounded, and Theorem 4.2(ii) can be applied.

4.2. *The general case.* There are important orthogonal polynomial sequences inducing a polynomial hypergroup but not belonging to the Kolmogorov–Szegö class [e.g., the Cartier–Dunau polynomials of Section 3.4]. Often it is not possible or at least very difficult to decide the membership to the Kolmogorov–Szegö class. Throughout this subsection we suppose on π only that the corresponding orthogonal polynomials induce a polynomial hypergroup on $\mathbb{N}_0$. $(X_n)_{n \in \mathbb{N}_0}$ will be a weakly stationary random sequence on this polynomial hypergroup with spectral measure μ.

A standard procedure in the Hilbert space $L^2(D_s, \mu)$ yields

$$\delta_n^2 = \frac{\Delta(X_0, \ldots, X_{n+1})}{\Delta(X_0, \ldots, X_n)}, \tag{26}$$

where

$$\Delta(X_n, \dots, X_{n+m}) = \det \begin{pmatrix} \langle X_n, X_n \rangle & \cdots & \langle X_n, X_{n+m} \rangle \\ \vdots & & \vdots \\ \langle X_{n+m}, X_n \rangle & \cdots & \langle X_{n+m}, X_{n+m} \rangle \end{pmatrix}.$$

Notice that $\Delta(X_0, \dots, X_n) > 0$ for all $n \in \mathbb{N}_0$ if and only if $\operatorname{supp}\mu$ is infinite. We assume further on that $|\operatorname{supp}\mu| = \infty$.

The following sequence of upper bounds holds for δ_n:

$$\begin{aligned} \delta_n^2 &= \frac{\Delta(X_0, \dots, X_{n+1})}{\Delta(X_0, \dots, X_n)} \le \frac{\Delta(X_1, \dots, X_{n+1})}{\Delta(X_1, \dots, X_n)} \le \cdots \\ &\le \frac{\Delta(X_n, X_{n+1})}{\Delta(X_n)} \le \Delta(X_{n+1}); \end{aligned} \tag{27}$$

see Mitrinović [(1970), page 46]. Because of Proposition 4.3 we concentrate on the case $h(n) \to \infty$. The following lemma is needed.

LEMMA 4.4. *Let $(R_n)_{n\in\mathbb{N}_0}$ induce a polynomial hypergroup on $\mathbb{N}_0$, and assume that $h(n) \to \infty$. Then for every $k, l \in \mathbb{N}_0$ holds $g(n, n + l; k) \to 0$ as $n \to \infty$.*

PROOF. For $n \ge k$ we have, with the Cauchy–Schwarz inequality,

$$\begin{aligned} g(n, n+l; k) &= h(k) \int_{D_s} R_n(x) R_{n+l}(x) R_k(x)\, d\pi(x) \\ &\le h(k) g(n, n; 0)^{1/2} \left(\int_{D_s} \big(R_{n+l}(x) R_k(x)\big)^2 \, d\pi(x) \right)^{1/2}. \end{aligned}$$

Since $|R_{n+l}(x) R_k(x)| \le 1$ for $x \in D_s$ we see that $g(n, n + l; k) \to 0$ as $n \to \infty$. $\square$

THEOREM 4.5. *Let $(X_n)_{n\in\mathbb{N}_0}$ be a weakly stationary random sequence on the polynomial hypergroup $\mathbb{N}_0$ induced by $(R_n)_{n\in\mathbb{N}_0}$. Assume that $h(n) \to \infty$ as $n \to \infty$. If $d(n) = E(X_n \overline{X_0})$ tends to zero, the random sequence $(X_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic.*

PROOF. We have $E(X_n\overline{X_n}) = \sum_{k=0}^{2n} g(n, n; k)\, d(k)$. Since $\sum_{k=0}^{2n} g(n, n; k) = 1$ and $g(n, n; k) \to 0$ as $n \to \infty$ by Lemma 4.4, Toeplitz's Lemma [see Knopp (1922), page 377] yields $\Delta(X_n) = E(X_n\overline{X_n}) \to 0$, and by (27) we have $\delta_n \to 0$. $\square$

REMARK 1. Theorem 4.5 improves Theorem 2 of Hösel and Lasser (1992), where we had to assume that the recurrence coefficients a_n, b_n and c_n are convergent.

REMARK 2. If the spectral measure μ is absolutely continuous with respect to the measure π, we can deduce that $d(n) \to 0$ [see Bloom and Heyer (1995), Theorem 2.2.32(vi)].

We apply (27) further on to deal also with spectral measures μ that contain discrete or singular parts.

THEOREM 4.6. *Let $(X_n)_{n\in\mathbb{N}_0}$ be a weakly stationary random sequence on the polynomial hypergroup $\mathbb{N}_0$ induced by $(R_n)_{n\in\mathbb{N}_0}$, where $h(n) \to \infty$ as $n \to \infty$. Assume that the spectral measure has the form*

$$\mu = f\pi + \mu_0 + \mu_1,$$

where $f \in L^1(D_s, \pi)$, $\operatorname{supp}\mu_0 \subseteq D_{s,0} := \{x \in D_s : R_n(x) \to 0 \text{ for } n \to \infty\}$ and $\operatorname{supp}\mu_1 \subseteq D_s \backslash D_{s,0}$ is a finite set or empty. If $m = |\operatorname{supp}\mu_1| > 0$ and there is some $n_0 \in \mathbb{N}_0$ such that

$$\inf_{n\geq n_0} \det \begin{pmatrix} \langle R_n, R_n\rangle_{\mu_1} & \cdots & \langle R_n, R_{n+m-1}\rangle_{\mu_1} \\ \vdots & & \vdots \\ \langle R_{n+m-1}, R_n\rangle_{\mu_1} & \cdots & \langle R_{n+m-1}, R_{n+m-1}\rangle_{\mu_1} \end{pmatrix} > 0, \tag{28}$$

then the random sequence $(X_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic.

PROOF. If $\mu_1 = 0$ we have $\Delta(X_n) = E(X_n\overline{X_n}) = \int_{D_s} R_n^2(x) f(x)\, d\pi(x) + \int_{D_s} R_n^2(x)\, d\mu_0(x)$. By Theorem 4.5 (and Remark 2 above) and the theorem of dominated convergence we have $\Delta(X_n) \to 0$ and the statement follows.

If $m = |\operatorname{supp}\mu_1| > 0$, we consider the Gramian determinants $\Delta(X_n, \ldots, X_{n+m-1})$ and $\Delta(X_n, \ldots, X_{n+m})$. Expanding the determinants we obtain

$$\Delta(X_n, \ldots, X_{n+k}) = \det \begin{pmatrix} \langle R_n, R_n\rangle_{\mu_1} & \cdots & \langle R_n, R_{n+k}\rangle_{\mu_1} \\ \vdots & & \vdots \\ \langle R_{n+k}, R_n\rangle_{\mu_1} & \cdots & \langle R_{n+k}, R_{n+k}\rangle_{\mu_1} \end{pmatrix} + \sum_{\sigma\in I_k} \varphi_\sigma,$$

where $|I_k|$ is finite and independent of n, whereas the φ_σ's are products of $\langle R_{n+i_1}, R_{n+j_1}\rangle_{f\pi+\mu_0}$ and $\langle R_{n+i_2}, R_{n+j_2}\rangle_{\mu_1}$ containing at least one factor $\langle R_{n+i}, R_{n+j}\rangle_{f\pi+\mu_0}$. Since all $\langle R_{n+i}, R_{n+j}\rangle_{f\pi+\mu_0}$ tend to zero as $n \to \infty$, we conclude from the assumption that $1/\Delta(X_n, \ldots, X_{n+m-1})$ is bounded for $n \geq n_0$. Since $|\operatorname{supp}\mu_1| = m$, the $(m+1) \times (m+1)$-determinant,

$$\det \begin{pmatrix} \langle R_n, R_n\rangle_{\mu_1} & \cdots & \langle R_n, R_{n+m}\rangle_{\mu_1} \\ \vdots & & \vdots \\ \langle R_{n+m}, R_n\rangle_{\mu_1} & \cdots & \langle R_{n+m}, R_{n+m}\rangle_{\mu_1} \end{pmatrix}$$

has to be zero. Hence $\Delta(X_n, \ldots, X_{n+m})$ tends to zero with $n \to \infty$. By (27) we get $\delta_n \to 0$. $\square$

A typical situation in which to use Theorem 4.6 is when π is even and $\operatorname{supp}\mu_1 \subseteq D_s \backslash D_{s,0} = \{-1, 1\}$. Since $R_n(1) = 1$ and $R_n(-1) = (-1)^n$ we obtain, for $\mu_1 = \alpha\varepsilon_1 + \beta\varepsilon_{-1}$,

$$\det\begin{pmatrix} \langle R_n, R_n\rangle_{\mu_1} & \langle R_n, R_{n+1}\rangle_{\mu_1} \\ \langle R_{n+1}, R_n\rangle_{\mu_1} & \langle R_{n+1}, R_{n+1}\rangle_{\mu_1} \end{pmatrix} = 4\alpha\beta.$$

The part of the spectral measure contained in $D_{s,0} = \{x \in D_s : R_n(x) \to 0\}$ is easy to deal with, as we have already seen in the proof of Theorem 4.6. Hence we want to derive results on the size of $D_{s,0}$.

We restrict ourselves from now on to the case that π is even, that is, $b_n = 0$ for all $n \in \mathbb{N}_0$. At first we consider the case of $a_n \to a$ as $n \to \infty$, where $\frac{1}{2} < a < 1$. The Cartier–Dunau polynomials $R_n(x)$ with $q \geq 2$ belong to this class.

We know then that $\operatorname{supp}\pi = [-\gamma, \gamma]$ is a proper subset of $D_s = [-1, 1]$ with $\gamma = 2\sqrt{a(1-a)}$. Note that $\gamma = 1$ exactly when $a = \frac{1}{2}$. In fact, using theorems of Blumenthal and Poincaré this is proved in Lasser [(1994), Theorem 2.2]. In addition we can obtain the following result. Compare also Voit [(1991), Theorem 8.2(4)].

PROPOSITION 4.7. *Assume that* $(R_n)_{n\in\mathbb{N}_0}$ *is even and induces a polynomial hypergroup on* $\mathbb{N}_0$. *Furthermore, let* $a_n \to a$ *with* $\frac{1}{2} < a < 1$. *Then* $R_n(x) \to 0$ *as* $n \to \infty$ *for each* $x \in]-1, 1[$.

PROOF. Fix $\alpha_0 \in]\gamma, 1[$. By the separating property of zeros of $R_n(x)$ [see Szegö (1975)], we get $R_n(\alpha_0) > 0$. Now $Q_n(x) = R_n(\alpha_0 x)/R_n(\alpha_0)$ defines a polynomial hypergroup on $\mathbb{N}_0$. The recurrence relation of the $Q_n(x)$ is given by

$$x\,Q_n(x) = \tilde{a}_n Q_{n+1}(x) + \tilde{c}_n Q_{n-1}(x),$$

where

$$\tilde{a}_n = a_n \frac{R_{n+1}(\alpha_0)}{R_n(\alpha_0)\alpha_0}, \qquad \tilde{c}_n = c_n \frac{R_{n-1}(\alpha_0)}{R_n(\alpha_0)\alpha_0}.$$

By the Poincaré theorem [cf. the proof of Theorem 2.2 in Lasser (1994)], we have

$$\lim_{n\to\infty} \frac{R_n(\alpha_0)}{R_{n-1}(\alpha_0)} < 1.$$

Since $\tilde{a}_n$ is convergent, the dual space of the polynomial hypergroup induced by $(Q_n)_{n\in\mathbb{N}_0}$ is $[-1, 1]$. In particular,

$$|R_n(\alpha_0 x)/R_n(\alpha_0)| \leq 1$$

for all $x \in [-1, 1]$, $n \in \mathbb{N}_0$. That means $\lim_{n\to\infty} |R_n(\alpha_0 x)| \leq \lim_{n\to\infty} R_n(\alpha_0) = 0$ for all $x \in [-1, 1]$. Since $\alpha_0 \in]\gamma, 1[$ can be chosen arbitrarily we have $\lim_{n\to\infty} R_n(x) = 0$ for all $x \in]-1, 1[$. $\square$

COROLLARY 4.8. *Let $(X_n)_{n\in\mathbb{N}_0}$ be a weakly stationary random sequence on the polynomial hypergroup $\mathbb{N}_0$ induced by $(R_n)_{n\in\mathbb{N}_0}$. Suppose that π is even and $a_n \to a$ with $\frac{1}{2} < a < 1$. Then $(X_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic.*

PROOF. By Proposition 4.3 we get $\lim_{n\to\infty} h(n) = \infty$ and the statement follows from Proposition 4.7. □

If $a_n \to \frac{1}{2}$, the problem is much more involved. A sufficient condition for $R_n(x) \to 0$ for every $x \in]-1, 1[$ can be derived by making use of the Turan determinant. Given the orthogonal polynomials $(R_n)_{n\in\mathbb{N}_0}$ fulfilling (3), the Turan determinant is defined by

$$\theta_n(x) = h(n)\left(R_n^2(x) - \frac{a_n}{a_{n-1}} R_{n-1}(x) R_{n+1}(x)\right). \tag{29}$$

PROPOSITION 4.9. *Assume that $(R_n)_{n\in\mathbb{N}_0}$ is even and induces a polynomial hypergroup on $\mathbb{N}_0$. Further let $a_n \to \frac{1}{2}$ as $n \to \infty$. The following inequalities for $\theta_n(x)$ are valid:*

(i) *$\theta_n(x)/h(n) \le C_1(R_{n-1}^2(x) + R_n^2(x) + R_{n+1}^2(x))$ for all $x \in \mathbb{R}$, $n \in \mathbb{N}$, where $C_1 > 0$ is a constant independent of x and n.*

(ii) *$|\theta_n(x) - \theta_{n-1}(x)|/h(n) \le C_2|a_{n-1}c_n - a_{n-2}c_{n-1}|(R_{n-1}^2(x) + R_n^2(x))$ for all $x \in [-1, 1]$, $n \in \mathbb{N}$, where $C_2 > 0$ is a constant independent of x and n.*

(iii) *Given some $\delta \in]0, 1[$ there is $N \in \mathbb{N}$ such that $\theta_n(x)/h(n) \ge C_3 (R_{n-1}^2(x) + R_n^2(x))$ for all $x \in [-1+\delta, 1-\delta]$, $n \ge N$, where $C_3 > 0$ is a constant independent of x and n.*

PROOF. (i) Since $\lim_{n\to\infty}(a_n/a_{n-1}) = 1$ and $2|R_{n-1}(x)R_{n+1}(x)| \le R_{n-1}^2(x) + R_{n+1}^2(x)$, the inequality of (i) follows immediately.

(ii) By using the recurrence relation (3) we get

$$\theta_n(x) = h(n)R_n^2(x) + h(n-1)R_{n-1}^2(x) - \frac{x}{a_{n-1}} h(n) R_{n-1}(x) R_n(x) \tag{30}$$

and

$$\begin{aligned}\theta_n(x) &= h(n)R_n^2(x) + h(n+1)\frac{a_n c_{n+1}}{a_{n-1}c_n} R_{n+1}^2(x) \\ &\quad - h(n)\frac{a_n x}{a_{n-1}c_n} R_n(x) R_{n+1}(x).\end{aligned} \tag{31}$$

Employing (30) to θ_n and (31) to θ_{n-1} gives

$$\begin{aligned}\theta_n(x)-\theta_{n-1}(x) &= h(n)\Big(1-\frac{a_{n-1}c_n}{a_{n-2}c_{n-1}}\Big)R_n^2(x)\\ &\qquad + \Big(h(n-1)\frac{a_{n-1}}{a_{n-2}c_{n-1}}-h(n)\frac{1}{a_{n-1}}\Big)xR_{n-1}(x)R_n(x)\\ &= h(n)\Big(\frac{a_{n-2}c_{n-1}-a_{n-1}c_n}{a_{n-2}c_{n-1}}R_n^2(x)\\ &\qquad + \frac{a_{n-1}c_n-c_{n-1}a_{n-2}}{c_{n-1}a_{n-2}a_{n-1}}xR_{n-1}(x)R_n(x)\Big).\end{aligned}$$

Since $a_n\to\frac{1}{2}$, $c_n\to\frac{1}{2}$ we obtain for $|x|\le 1$,

$$|\theta_n(x)-\theta_{n-1}(x)|/h(n)\le C_2|a_{n-1}c_n-a_{n-2}c_{n-1}|\big(R_{n-1}^2(x)+R_n^2(x)\big).$$

(iii) Applying (30) we have

$$\theta_n(x)=h(n)\Big(R_n(x)-\frac{x}{2a_{n-1}}R_{n-1}(x)\Big)^2+\Big(1-\frac{x^2}{4c_na_{n-1}}\Big)h(n-1)R_{n-1}^2(x)$$

and

$$\theta_n(x)=h(n-1)\Big(R_{n-1}(x)-\frac{x}{2c_n}R_n(x)\Big)^2+\Big(1-\frac{x^2}{4c_na_{n-1}}\Big)h(n)R_n^2(x).$$

In particular it follows that

$$\theta_n(x)\ge\Big(1-\frac{x^2}{4c_na_{n-1}}\Big)h(n-1)R_{n-1}^2(x)$$

and

$$\theta_n(x)\ge\Big(1-\frac{x^2}{4c_na_{n-1}}\Big)h(n)R_n^2(x).$$

Since $x\in[-1+\delta,1-\delta]$ there is a constant C_3 and $N\in\mathbb{N}$ such that

$$\theta_n(x)/h(n)\ge C_3\big(R_{n-1}^2(x)+R_n^2(x)\big)\qquad\text{for all } n\ge N.\qquad\square$$

THEOREM 4.10. *Assume that $(R_n)_{n\in\mathbb{N}_0}$ is even and induces a polynomial hypergroup on $\mathbb{N}_0$. Let $a_n\to\frac{1}{2}$ and $h(n)\to\infty$ as $n\to\infty$. Further, suppose that*

$$\sum_{n=1}^{\infty}|a_nc_{n+1}-a_{n-1}c_n|<\infty. \tag{32}$$

Then every weakly stationary random sequence $(X_n)_{n\in\mathbb{N}_0}$ on the polynomial hypergroup $\mathbb{N}_0$ induced by $(R_n)_{n\in\mathbb{N}_0}$ is asymptotically R_n-deterministic.

PROOF. It suffices to prove that $R_n(x) \to 0$ as $n \to \infty$ for every $x \in]-1, 1[$. Proposition 4.9(ii) and (iii) imply

$$\begin{aligned}|\theta_n(x) - \theta_{n-1}(x)| &\le h(n)C_2|a_{n-1}c_n - a_{n-2}c_{n-1}|\big(R_{n-1}^2(x) + R_n^2(x)\big)\\ &\le \frac{C_2}{C_3}|a_{n-1}c_n - a_{n-2}c_{n-1}|\theta_n(x) = \varepsilon_n\theta_n(x)\end{aligned}$$

for all $n \ge N$, where $\varepsilon_n = \frac{C_2}{C_3}|a_{n-1}c_n - a_{n-2}c_{n-1}|$. Hence

$$\frac{1}{1+\varepsilon_n}\theta_{n-1}(x) \le \theta_n(x) \le \frac{1}{1-\varepsilon_n}\theta_{n-1}(x)$$

for every $n \ge N$. Since $\sum_{n=2}^{\infty} \varepsilon_n$ is convergent, $\theta_n(x)$ is convergent, too. Applying once more Proposition 4.9(iii) and $h(n) \to \infty$, we get $R_n(x) \to 0$. □

We conclude this paper with some further examples.

4. The Cartier–Dunau polynomials that are essential for the investigation of stationary radial stochastic processes on homogeneous trees have the property that $a_n = q/(q+1)$. Hence for $q \ge 2$ we can use Corollary 4.8 and get that $\delta_n \to 0$ for the corresponding random sequences $(X_n)_{n\in\mathbb{N}_0}$.
5. $R_n(x; \beta|q)$ the (continuous) q-ultraspherical polynomials with $-1 < \beta < 1$ and $0 < q < 1$. Their hypergroup structure is studied in Lasser (1983) [cf. also Bressoud (1981)]. The recurrence coefficients a_n, c_n are not given in explicit form, but we know the asymptotic behavior; see Bloom and Heyer [(1995), page 168]. In fact,

$$\alpha_n = \frac{1}{2} + \frac{1}{n} + o\left(\frac{1}{n}\right)$$

and hence $a_n \to \frac{1}{2}$ as $n \to \infty$. It now follows with elementary computations from (4) that $h_n \to \infty$ with growing n.

To check property (32) of Theorem 4.10 we investigate the monotonicity of the sequence $(\alpha_n)_{n\in\mathbb{N}}$. A direct calculation shows that $(\alpha_n)_{n\in\mathbb{N}}$ is increasing if $\beta \le q$ and decreasing if $q \le \beta$. In both cases we have $\sum_{n=1}^{\infty} |\alpha_{n+1} - \alpha_n| = |\frac{1}{4} - \alpha_1|$. Now Theorem 4.10 implies that every corresponding random sequence is asymptotically $R_n(\cdot; \beta|q)$-deterministic.

REFERENCES

ARNAUD, J. P. (1994). Stationary processes indexed by a homogeneous tree. *Ann. Probab.* **22** 195–218.

BLOOM, W. R. and HEYER, H. (1995). *Harmonic Analysis of Probability Measures on Hypergroups*. de Gruyter, Berlin.

BLOWER, G. (1996). Stationary processes for translation operators. *Proc. London Math. Soc.* **72** 697–720.

BRESSOUD, D. M. (1981). Linearization and related formulas for q-ultraspherical polynomials. *SIAM J. Math. Anal.* **12** 161–168.

BROCKWELL, P. J. and DAVIES, R. A. (1991). *Time Series*: *Theory and Methods*. Springer, New York.

CHIHARA, T. S. (1978). *An Introduction to Orthogonal Polynomials*. Gordon and Breach, New York.

COWLING, M., MEDA, ST. and SETTI, A. G. (1998). An overview of harmonic analysis on the group of isometries of a homogeneous tree. *Expo. Math.* **16** 385–424.

DEVROYE, L. and GYÖRFI, L. (1985). *Nonparametric Density Estimation*: *The L^1-View*. Wiley, New York.

GERONIMUS, Y. L. (1960). *Polynomials Orthogonal on a Circle and Interval*. Pergamon, Oxford.

HEYER, H. (1991). Stationary random fields over hypergroups. In *Gaussian Random Fields* 197–213. World Scientific, Singapore.

HEYER, H. (2000). The covariance distribution of a generalized random field over a commutative hypergroup. *Contemp. Math.* **261** 73–82.

HÖSEL, V. (1998). On the estimation of covariance functions on P_n-weakly stationary processes. *Stochastic Anal. Appl.* **16** 607–629.

HÖSEL, V. and LASSER, R. (1992). One-step prediction for P_n-weakly stationary processes. *Monatsh. Math.* **113** 199–212.

KAKIHARA, Y. (1997). *Multidimensional Second Order Stochastic Processes*. World Scientific, Singapore.

KNOPP, K. (1922). *Theorie und Anwendung der unendlichen Reihen*. Springer, Berlin.

LASSER, R. (1983). Orthogonal polynomials and hypergroups. *Rend. Mat.* **3** 185–209.

LASSER, R. (1994). Orthogonal polynomials and hypergroups II—The symmetric case. *Trans. Amer. Math. Soc.* **341** 749–770.

LASSER, R. and LEITNER, M. (1989). Stochastic processes indexed by hypergroups I. *J. Theoret. Probab.* **2** 301–311.

LASSER, R. and LEITNER, M. (1990). On the estimation of the mean of weakly stationary and polynomial weakly stationary sequences. *J. Multivariate Anal.* **35** 31–47.

LASSER, R., OBERMAIER, J. and STRASSER, W. (1993). On the consistency of weighted orthogonal series density estimators with respect to L^1-norm. *Nonparametric Statist.* **3** 71–80.

LEITNER, M. (1991). Stochastic processes indexed by hypergroups II. *J. Theoret. Probab.* **4** 321–332.

LUBINSKY, D. S. (1987). A survey of general orthogonal polynomials for weights on finite and infinite intervals. *Acta Appl. Math.* **10** 237–296.

MITRINOVIĆ, D. S. (1970). *Analytic Inequalities*. Springer, Berlin.

RAO, M. M. (1989). Bimeasures and harmonizable processes. *Probability Measures on Groups IX*. *Lecture Notes in Math.* **1379** 254–298. Springer, Berlin.

SHIRYAYEV, A. N. (1984). *Probability*. Springer, New York.

SZEGÖ, G. (1975). *Orthogonal Polynomials*. Amer. Math. Soc., Providence, RI.

VOIT, M. (1990). Central limit theorems for random walks on $\mathbb{N}_0$ that are associated with orthogonal polynomials. *J. Multivariate Anal.* **34** 290–322.

VOIT, M. (1991). Factorization of probability measures on symmetric hypergroups. *J. Austral. Math. Soc. Ser. A* **50** 417–467.

YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions* I. Springer, New York.

GSF-NATIONAL RESEARCH CENTER
FOR ENVIRONMENT AND HEALTH
INSTITUTE OF BIOMATHEMATICS
AND BIOMETRY
INGOLSTÄDTER LANDSTRASSE 1
85764 NEUHERBERG
GERMANY
E-MAIL: lasser@gsf.de
hoesel@t-online.de