ON GENERALIZED SUM RULES FOR JACOBI MATRICES

F. NAZAROV, F. PEHERSTORFER1, A. VOLBERG AND P. YUDITSKII2

Abstract. This work is in a stream (see e.g. [6], [8], [9], [5]) initiated by a paper of Killip and Simon [7], an earlier paper [3] also should be mentioned here. Using methods of Functional Analysis and the classical Szegő Theorem we prove sum rule identities in a very general form. Then, we apply the result to obtain new asymptotics for orthonormal polynomials.

1. Introduction

1.1. Finite dimensional perturbation of the Chebyshev matrix. Let \{e_n\}_{n \geq 0} be the standard basis in \(l^2(\mathbb{Z}_+).\) Let \(J\) be a Jacobi matrix defining a bounded self-adjoint operator on \(l^2(\mathbb{Z}_+):\)

\[ Je_n = p_n e_{n-1} + q_n e_n + p_{n+1} e_{n+1}, \quad n \geq 1, \]

and

\[ J e_0 = q_0 e_0 + p_1 e_1. \]

Under the condition \(p_n > 0,\) the vector \(e_0\) is cyclic for \(J.\) The function

\[ r(z) = \langle (J - z)^{-1} e_0, e_0 \rangle \]

is called the resolvent function. It has the representation

\[ r(z) = \int \frac{d\sigma(x)}{x - z}. \]

The measure \(\sigma, d\sigma \geq 0,\) is called the spectral measure of \(J.\)

Using a three term recurrence relation for orthonormal polynomials \(\{P_n(z)\}_{n \geq 0}\) with respect to \(\sigma\) one can restore the coefficient sequences of \(J\)

\[ zP_n(z) = p_n P_{n-1}(z) + q_n P_n(z) + p_{n+1} P_{n+1}(z), \quad n \geq 1, \]

and

\[ zP_0(z) = q_0 P_0(z) + p_1 P_1(z). \]

With a given \(J\) we associate a sequence \(J(n)\) defined by

\[ p(n)_k = \begin{cases} p_k, & k < n \\ 1, & k \geq n \end{cases}, \]

\[ q(n)_k = \begin{cases} q_k, & k < n \\ 0, & k \geq n \end{cases}. \]

Date: December 3, 2003.

The work was supported by the Austrian Science Fund FWF, project number: P16390-N04.

Keywords: orthogonal polynomials, asymptotics, Szegő condition.

2000 AMS Subject classification: primary 47B36, secondary 42C05.
$J(n)$ is a finite dimensional perturbation of the “free” (Chebyshev) matrix $J_0 = S_+ + S^{*}_+ + S_n e_n = e_{n+1}$.

Note that $r_0(z) = ((J_0 - z)^{-1}e_0,e_0) = -\zeta$, where $1/\zeta + \zeta = z$, $\zeta \in \mathbb{D}$, that is $\zeta = \frac{z - \sqrt{z^2 - 4}}{2}$. Further, in terms of orthonormal polynomials $r(n)(z) = (J(n) - z)^{-1}e_0,e_0) = \frac{p_n Q_n(z) - \zeta Q_{n-1}(z)}{p_n P_n(z) - \zeta P_{n-1}(z)}$,

where $Q_n$ are so called orthonormal polynomials of the second kind $Q_n(z) := \int P_n(x) - \zeta P_{n-1}(x) dx$.

They satisfy the same three term recurrence relation as $P_n$’s but with a different initial condition). What is important for us

(1) $\sigma'(n)_{a.c.}(x) = \frac{1}{\pi} \text{Im}(r(n)(x + i0)) = \frac{-\text{Im}\zeta(x + i0)}{\pi |p_n P_n(x) - \zeta(x + i0) P_{n-1}(x)|^2}$.

and

(2) $\sigma(J(n)) \cap \{\mathbb{R} \setminus [-2,2]\} = \{z \in \mathbb{C} \setminus [-2,2]: p_n P_n(z) - \zeta(z) P_{n-1}(z) = 0\}$.

The perturbation determinant of $J(n)$ with respect to $J_0$ is well defined and we can introduce a function

$$\Delta_n(\zeta) = \frac{1}{\prod_{j=1}^{n-1} p_j} \det(J(n) - z)(J_0 - z)^{-1}.$$ 

By definition

(3) $\log \Delta_n(z) = -t(n)_0 - \sum_{k \geq 1} \frac{t(n)_k}{k z^k}$

where

$$t(n)_0 = \sum_{j=1}^{n-1} \log p_j, \quad t(n)_k = \text{tr}(J(n)^k - J_0^k), \quad k \geq 1.$$ 

On the other hand one can find the determinant by a direct calculation and then

$$\Delta_n(z) = (p_n P_n(z) - \zeta P_{n-1}(z))\zeta^n,$$

as before $1/\zeta + \zeta = z$, $\zeta \in \mathbb{D}$.

Therefore $\Delta_n(z)$ has explicit representation in terms of coefficients of $J(n)$, on the other hand it has nice analytic properties: its zeros in $\mathbb{C} \setminus [-2,2]$ are simple and related to the eigenvalues of $J(n)$ in this region (see (2)); it has no poles; and by

(4) $$|\Delta_n(x + i0)|^2 = \frac{1}{2\pi} \frac{\sqrt{4 - x^2}}{\sigma'(n)_{a.c.}}.$$ 

That is, we can restore $\Delta_n(z)$ only in terms of these (partial) spectral data (see the next subsection).
1.2. The Killip–Simon functional via spectral data.

**Definition 1.1.** Let \( J \) be a Jacobi matrix with a spectrum on \([-2, 2] \cup X\), where the only possible accumulation points of \( X = \{x_k\} \) are \( \pm 2 \). Following to Killip and Simon, to a given nonnegative polynomial \( A \) we associate the functional that might diverge only to \(+\infty\)

\[
\Lambda_A(J) := \sum_X F(x_k) + \frac{1}{2\pi} \int_{-2}^{2} \log \left( \frac{\sqrt{4 - x^2}}{2\pi \sigma_{n,c.}} \right) A(x) \sqrt{4 - x^2} \, dx,
\]

where

\[
F(x) = \begin{cases} 
A(x) \sqrt{x^2 - 4} & \text{for } x > 2, \\
-\int_{-2}^{x} A(x) \sqrt{x^2 - 4} \, dx & \text{for } x < -2.
\end{cases}
\]

Let us point out that the Killip–Simon functional \( \Lambda_A(J) \) is defined in terms of the spectral data of \( J \) only. Let us demonstrate how to obtain for a finite dimensional perturbation \( J_0 \) a representation of \( \Lambda_A(J(n)) \) in terms of the recurrence coefficients.

First, let us note that the function \( \log \Delta_n(z) \) is well defined in the upper half plane, in fact, in the domain \( \mathbb{C} \setminus \sigma(J(n)) \). Moreover, the boundary values of the real part \( \text{Re} \log \Delta_n(x + it) \), \( x \in [-2, 2] \), are given by (4). For \( x \geq 2 \) the imaginary part of \( \log \Delta_n(z) \) (that is the argument of \( \Delta_n(\zeta) \)) is of the form

\[
\frac{1}{\pi} \arg \Delta_n(x + i0) = \#\{y \in \sigma(J(n)) : y \geq x\}
\]

and similarly,

\[
\frac{1}{\pi} \arg \Delta_n(x + i0) = -\#\{y \in \sigma(J(n)) : y \leq x\}
\]

for \( x \leq -2 \). Therefore, multiplying \( \log \Delta_n(z) \) by \( A(z)\sqrt{z^2 - 4} \), where \( A(z) \) is the given nonnegative polynomial, we get a function with the following representation

\[
A(z)\sqrt{z^2 - 4} \log \Delta_n(z) = B_n(z) + \int_{\sigma(J(n))} \frac{d\lambda_n}{x - z},
\]

where \( B_n(z) \) is a (real) polynomial of degree one bigger than \( A \) and

\[
\lambda'_n(x) = \begin{cases} 
\frac{1}{2\pi} A(x) \sqrt{4 - x^2} \log \frac{1}{2\pi \sigma_{n,c.}}, & x \in [-2, 2] \\
A(x) \sqrt{x^2 - 4} \{y \in \sigma(J(n)) : y \geq x\}, & x \geq 2 \\
A(x) \sqrt{x^2 - 4} \{y \in \sigma(J(n)) : y \leq x\}, & x \leq -2
\end{cases}
\]

Thus the functional \( \Lambda_A(J(n)) = \int d\lambda_n \).

Let us mention that the polynomial \( B_n(z) \) is determined uniquely by (7) since

\[
\int_{\sigma(J(n))} \frac{d\lambda_n}{x - z} = -\int \frac{d\lambda_n}{z} - \ldots = O\left( \frac{1}{z} \right), \quad z \to \infty.
\]

Let us define

\[
\Phi(z) = \text{Const} + a_1 z + \cdots + a_{m+2} z^{m+2}
\]

by

\[
\Phi'(z) = z A(z) - \frac{1}{\pi} \int_{-2}^{2} \frac{A(x) - A(z)}{x - z} \sqrt{4 - x^2} \, dx.
\]
Note that

\[ (9) \quad A(z)\sqrt{z^2 - 4} = \frac{1}{\pi} \int_{-2}^{2} \frac{A(x)}{x - z} \sqrt{4 - x^2} \, dx + \Phi'(z). \]

Therefore, using (3), (7), (8) and (9) we get

\[ (10) \quad \int d\lambda_n = -at(n)_{0} + a_{1}t(n)_{1} + 2a_{2} \frac{t(n)_{2}}{2} + \cdots + (m + 2)a_{m+2} \frac{t(n)_{m+2}}{(m+2)} \]

\[ = -at(n)_{0} + \text{tr}\{\Phi(J(n)) - \Phi(J_{0})\}, \]

where we put

\[ a = \frac{1}{\pi} \int_{-2}^{2} A(x)\sqrt{4 - x^2} \, dx. \]

Note, if \( A(z) = 1 \), that is \( a = 2 \), \( \Phi(z) = \text{Const} + z^2/2 \), then we are in the Killip–Simon case \([7]\):

\[ \int d\lambda_n = \frac{t(n)_{2}}{2} - 2t(n)_{0} = -1 + \sum_{k=1}^{\infty} (p(n)_{k}^2 - 1 - \log p(n)_{k}^2) + \frac{1}{2} \sum_{k=0}^{\infty} q(n)_{k}. \]

For a more general example see Appendix.

1.3. The Killip–Simon functional via coefficient sequences. For a bounded operator \( G \) in \( l^2(\mathbb{Z}_+) \) we denote \( G^{(k)} := (S_{+}^k)^{k}GS_{+}^{k} \).

Lemma 1.2. For all \( k \geq 1 \) and \( n \geq l - 1 \)

\[ (J^{(k)})^{l}e_n = (J^{(l)})^{(k)}e_n. \]

Proof. Let us mention that the decomposition of the vector \( J^{l}e_{k+n} \) begins with the basic’s vector \( e_{k+n-l} \). Therefore the orthoprojector \( P_{k-1} \) onto the subspace spanned by \( \{e_0, \ldots, e_{k-1}\} \) annihilates this vector, \( P_{k-1}J^{l}e_{k+n} = 0 \). Thus, by induction,

\[ (J^{(k)})^{l+1}e_n = J^{(k)}(J^{(l)})^{l}e_n = J^{(k)}(J^{(l)})^{(k)}e_n = (S_{+}^k)^{k}J^{l}S_{+}^{k}(S_{+}^k)^{k}J^{l}e_n = (S_{+}^k)^{k}J^{l+1}e_{k+n} = (J^{(l+1)})^{(k)}e_n. \]

\( \square \)

For a bounded Jacobi matrix \( J \) (and a polynomial \( A \)) let us define a function of a finite number of variables

\[ h_{A} = h_{A}(J) := -a \log p_{m+2} + \langle \{\Phi(J) - \Phi(J_{0})\}e_{m+1}, e_{m+1} \rangle. \]

Note that due to the previous lemma

\[ h_{A} \circ \tau^{k} = -a \log p_{m+k+2} + \langle \{\Phi(J^{(k)}) - \Phi(J_{0})\}e_{m+1}, e_{m+1} \rangle \]

\[ = -a \log p_{m+k+2} + \langle \{\Phi(J) - \Phi(J_{0})\}e_{m+k+1}, e_{m+k+1} \rangle, \]

where \( \tau \) acts just as a shift of indexes. In this case the series

\[ \sum_{k \geq 0} h_{A} \circ \tau^{k} \]

may not converge, but the generic term is well define.
Definition 1.3. With a given Jacobi matrix $J$ and a polynomial $A$ of degree $m$ we associate the series

$$H_A(J) := \sum_{k=0}^{m} (-a \log p_{k+1} + \langle \{\Phi(J) - \Phi(J_0)\} e_k, e_k\rangle) + \sum_{k \geq 0} h_A \circ \tau^k. \tag{11}$$

Note that $H_A(J(n))$ is just a finite sum, in fact $h \circ \tau^k$ vanishes starting with a suitable $k$, moreover $H_A(J(n)) = \Lambda_A(J(n))$.

1.4. Results.

Theorem 1.4. Let $A$ be a nonnegative polynomial. The spectral measure $\sigma$ of a Jacobi matrix $J$ with a spectrum of the form $[-2, 2] \cup X$, where $\pm 2$ are the only possible accumulation points of the discrete set $X$, satisfies $\Lambda_A(J) < \infty$ if and only if series $\{1\}$ converges; moreover $H_A(J) = \Lambda_A(J)$.

In a sense our result is a kind of “existence theorem”. To balance the situation we derive from it the following application. (We conjectured this result in a note mentioned in §).

Theorem 1.5. Let $A(x)$ be a nonnegative polynomial of degree $m$ with all zeros on $[-2, 2]$. Let a measure $\sigma$ supported on $[-2, 2] \cup X$ satisfy the condition $\int d\lambda < \infty$, where

$$\lambda'(x) = \lambda'(x; \sigma) = \begin{cases} \frac{1}{2\pi} A(x) \sqrt{4 - x^2} \log \left( \frac{1}{2\pi} \frac{1-x}{\sigma_{\mu}(x)} \right), & x \in [-2, 2] \\ A(x) \sqrt{x^2 - 4} \{y \in X : y \geq x\}, & x \geq 2 \\ A(x) \sqrt{x^2 - 4} \{y \in X : y \leq x\}, & x \leq -2 \end{cases}. \tag{12}$$

Then the sequence of orthonormal polynomials $P_n(z) = P_n(z; \sigma)$, normalized by

$$\zeta^{n+1} \sqrt{z^2 - 4} P_n(z) \exp \left( -\frac{\tilde{B}_n(z)}{A(z) \sqrt{z^2 - 4}} \right) = 1 + O \left( \frac{1}{z^{m+2}} \right),$$

the polynomial $\tilde{B}_n(z)$ (of degree $m + 1$) is determined uniquely by the condition

$$\log \{\zeta^{n+1} \sqrt{z^2 - 4} P_n(z)\} - \frac{\tilde{B}_n(z)}{A(z) \sqrt{z^2 - 4}} = O \left( \frac{1}{z^{m+2}} \right),$$

converges uniformly on compact subsets of the domain $\mathbb{C} \setminus [-2, 2]$ to the holomorphic function

$$D(z) := \exp \left( \frac{1}{A(z) \sqrt{z^2 - 4}} \int \frac{d\lambda}{x - z} \right). \tag{13}$$

Note that as well as in the Szegő case the limit function $D(z)$ can be expressed only in terms of $\sigma'_{a,c}$. X.

2. Semicontinuity of Szegő Type Functional

For a measure $\mu$ on the unit circle $\mathbb{T}$ we denote by $Sz(\mu)$ the functional

$$Sz(\mu) = \int_{\mathbb{T}} \log \frac{d\mu_{a,c}}{dm} dm.$$

Recall the main property of this functional

$$Sz(\mu) = \inf \{ \log \int_{\mathbb{T}} |1 - f|^2 \, dp(t) : f \text{ is a polynomial, } f(0) = 0 \}.$$
Lemma 2.1. Let $\mu_k$ converge weakly to $\mu$. Then

$$\limsup Sz(\mu_k) \leq Sz(\mu).$$

Proof. Since for every $\epsilon$ there exists a polynomial $g$, $g(0) = 0$, such that

$$\log \int_T |1 - g|^2 d\mu(t) \leq Sz(\mu) + \epsilon,$$

starting from a suitable $k$ we have

$$\log \int_T |1 - g|^2 d\mu_k(t) \leq Sz(\mu) + 2\epsilon.$$

But for every $k$

$$Sz(\mu_k) = \inf \left\{ \log \int_T |1 - f|^2 d\mu_k(t) : f \text{ is a polynomial, } f(0) = 0 \right\} \leq \log \int_T |1 - g|^2 d\mu_k(t).$$

Thus (14) is proved. □

Lemma 2.2. Let $\rho$ be a normalized nonnegative weight, i.e., $\rho \geq 0$, $\int_T \rho dm = 1$, such that $\rho \log \rho \in L^1$. Assume that $\mu_k$ converges weakly to $\mu$. Then

$$\liminf \int_T \log dm d(\mu_k) a.e. \rho dm \geq \int_T \log dm d\mu a.e. \rho dm.$$

Proof. Define a map $\psi : T \to T$ by $\psi(e^{i\theta}) = \exp(i \int_0^\theta \rho(e^{i\theta}) d\theta)$ and denote by $\phi$ the inverse map, $\psi \circ \phi = \text{id} : T \to T$. Let us apply Lemma 2.2 to the sequence $\tilde{\mu}_n := \mu_n \circ \phi$ that converges weakly to $\tilde{\mu} := \mu \circ \phi$.

$$\liminf \int_T \log dm d(\tilde{\mu}_k) a.e. dm \geq \int_T \log dm d\tilde{\mu} a.e. dm.$$

Making the inverse change of variable in each integral we have

$$\liminf \int_T \log dm d(\mu_k) a.e. dm \geq \int_T \log dm d\mu a.e. dm.$$

Since $\rho \log \rho \in L^1$ we get (15). □

Corollary 2.3.

$$\liminf_{n \to \infty} \Lambda_A(J(n)) \geq \Lambda_A(J).$$

Proof. Outside of $[-2, 2]$ we apply the Fatou Lemma, e.g. [53], p. 17, and on $[-2, 2]$ we apply Lemma 2.2. □

3. Lemma on positiveness and its consequences

For a given interval $I$, $0 \in I$, let $h \in C(I^l)$ be such that $h(0, ..., 0) = 0$. Then

$$H(\xi) = \sum_{i=0}^\infty h(x_{i+1}, x_{i+2}, ..., x_{i+l})$$

is well defined on

$$I_0^\infty = \{ \xi : \xi = (x_0, x_1, ..., x_n, 0, 0...) \}. $$
Lemma 3.1. Assume that $H$ is bounded from below, $H(x) \geq C$ for all $x \in I_0^\infty$. Then there exists a function $g$ of the form
\[
g(x_1, \ldots, x_l) = h(x_1, \ldots, x_l) + \gamma(x_2, \ldots, x_l) - \gamma(x_1, \ldots, x_{l-1}), \quad \gamma \in C(I^{l-1}),
\]
such that $g \geq 0$.

First we prove a sublemma.

Lemma 3.2. The set $G$, consisting of functions of the form
\[
G = \{ g(x_1, \ldots, x_l) + \gamma(x_1, \ldots, x_{l-1}) - \gamma(x_2, \ldots, x_l) \},
\]
where $g \in C(I^l), g \geq 0, g(0) = 0, \gamma \in C(I^{l-1}),$ is closed in $C(I^l)$.

Proof. We give a proof in the case of two variables (the general case can be considered in a similar way).

Let
\[
(16) \quad h(x, y) = \lim \{ g_n(x, y) + \gamma_n(x) - \gamma_n(y) \},
\]
Assuming the normalization $\gamma_n(0) = 0$ we get a uniform bound for $\gamma_n$,
\[
-1 - h(0, x) \leq \gamma_n(x) \leq h(x, 0) + 1.
\]
Therefore there exists a subsequence that converges weakly, say, in $L^2$. Then, using the Mazur Theorem, see e.g.\ref{13}, p. 120, and convexity of $G$ we can find a sequence $\gamma_n^{(1)}(x)$ and corresponding sequence of $g_n^{(1)}(x, y) \geq 0$ such that $\gamma_n^{(1)} \to \gamma_1, g_n^{(1)} \to g_1$ in $L^2$ strongly and we still have \ref{13}.

Thus, there exists a representation
\[
(17) \quad h(x, y) = g_1(x, y) + \gamma_1(x) - \gamma_1(y)
\]
that holds almost everywhere, and the function $\gamma_1(x)$, in fact, because of uniform boundness, belongs to $L^\infty$.

Starting with this place we will show that there exists a representation for $h(x, y)$ of the form \ref{17} but with continuous functions $\gamma$ and $g \geq 0$. First, let us construct a function $\gamma_2$ which is defined for all $x \in I$ and such that $\gamma_2(x) - \gamma_2(y) \leq h(x, y)$ holds everywhere.

Set $\gamma_2(x_0) = \limsup_{\delta \to 0} \frac{1}{\delta} \int_{x_0 - \delta}^{x_0 + \delta} \gamma_1$. Note that $\gamma_2(x) = \gamma_1(x)$ (a.e.). To show that $\gamma_2(x) - \gamma_2(y) \leq h(x, y)$ for all $(x, y) \in I^2$ we average the inequality with $\gamma_1$ over rectangles $[x_0 - \delta \leq x < x_0 + \delta, y_0 - \delta \leq y \leq y_0 + \delta]$ and take the upper limit when $\delta \to 0$. Since
\[
\limsup(a + b) \geq \limsup a + \liminf b
\]
we get the inequality we need. Next, we construct an upper semicontinuous function $\gamma_3(x_0) = \limsup_{x \to x_0} \gamma_2(x)$.

Let $\Gamma$ be the set of upper semicontinuous functions defined on $I$ with normalization $\gamma(0) = 0$ and such that $\gamma(x) - \gamma(y) \leq h(x, y)$. The previous construction shows that $\Gamma \neq \emptyset$. Now, the key point is to consider the function
\[
\gamma_4(x) := \sup \{ \gamma(x) : \gamma \in \Gamma \}.
\]
It belongs to $\Gamma$ since $\sup \{ \beta_1(x), \beta_2(x) \} \in \Gamma$ if only $\beta_1(x) \in \Gamma, \beta_2(x) \in \Gamma$.

We claim that $\gamma_4(x)$ is lower semicontinuous. Assume, on the contrary, that it is not. This means that there exist $\delta > 0$, a point $x_0 \in I$ and a sequence $\{x_n\}$, $\lim x_n = x_0$, such that $\gamma_4(x_n) \leq -h(x_0, x) - \delta$. Let us mention that $x_0 \neq 0$ since
\[
-h(0, x) \leq \gamma(x) \leq h(x, 0),
\]
and hence \( \lim_{x \to 0} \gamma(x) = 0 = \gamma(0) \) for all \( \gamma \in \Gamma \).

The function \( h(x, y) \) is continuous therefore we can choose such \( N \) that
\[
|h(x_N, y) - h(x_0, y)| \leq \delta/2
\]
for all \( y \in I \).

Let
\[
\gamma_5(x) = \begin{cases} 
\gamma_4(x), & x \neq x_N \\
\gamma_4(x_N) + \delta/2, & x = x_N.
\end{cases}
\]

Let us check that \( \gamma_5 \in \Gamma \). It is upper semicontinuous, \( \gamma_5(0) = 0 \). Further, for \( y \neq x_N \) we have
\[
\gamma_5(x_N) - \gamma_5(y) = \gamma_4(x_N) + \delta/2 - \gamma_4(y) \\
\leq \gamma_4(x_0) - \delta/2 - \gamma_4(y) \\
\leq h(x_0, y) - \delta/2 \leq h(x_N, y).
\]

Moreover the inequality \( \gamma_5(x) - \gamma_5(y) \leq h(x, y) \) holds on the line \( y = x_N \) and for all other values of \( x \) and \( y \).

On the other hand it could not be in the class, since
\[
\gamma_5(x_N) > \sup\{\gamma(x_N), \gamma \in \Gamma\}.
\]

Therefore we arrive to a contradiction. Thus \( \gamma_4(x) \) is simultaneously upper and lower semicontinuous, that is \( \gamma_4(x) \) is a continuous function. The lemma is proved.

\[\square\]

**Proof of Lemma 3.1.** If not then \( h \) does not belong to the closed convex set \( G \). Therefore there exists a measure \( \mu \in C(I^l)^*, d\mu \geq 0 \), such that
\[
(18) \quad \int h(x) \, d\mu(x) < 0
\]
and
\[
\int (\gamma(x_2, ..., x_l) - \gamma(x_1, ..., x_{l-1})) \, d\mu(x) = 0.
\]

In other words
\[
(19) \quad \int_{z \in I} d\mu(y, z) = \int_{z \in I} d\mu(z, y)
\]
for all \( y \in I^{l-1} \).

Without lost of generality we may assume that \( \mu \) is absolutely continuous, moreover \( d\mu = w(x_1, ..., x_l) \, dx_1 \ldots dx_l \), \( w \neq 0 \) a.e. Note that condition (19) is now of the form
\[
(20) \quad \int_{z \in I} w(y, z) \, dz = \int_{z \in I} w(z, y) \, dz, \quad y \in I^{l-1}.
\]

We want to get a contradiction between (18) and \( H \geq C \) by extending the functional related to \( w \) on functions on \( I_0^\infty \).

We can normalize \( w \) by the condition \( \int h(x) = 1 \). Let us think on \( w \) as on the probability
\[
w(y) dy = P\{x_i \in (y_i, y_i + dy_i), \ i = 1, ..., l\},
\]
and we want
\[
(21) \quad P\{x_i + k \in (y_i, y_i + dy_i), \ i = 1, ..., l\} = w(y) \, dy, \ \text{for all } k,
\]
that is the probability should be shift invariant. Actually we will define on $I_N$ step by step for increasing $N$ probabilistic measures

$$
\rho(x_1, \ldots, x_N)dx_1 \ldots dx_N
$$

using a conditional probability.

For $N \geq l$ inductively define

$$
\rho(x_1, \ldots, x_N, x_{N+1})dx_1 \ldots dx_N dx_{N+1} := \frac{w(x_{N+2-l}, \ldots, x_N, x_{N+1})dx_{N+1}}{\int w(x_{N+2-l}, \ldots, x_N, v)dv}.
$$

Now we have to check that (21) holds true.

If $k \neq N + 1 - l$ then (21) holds by the induction conjecture since

$$
\int I \rho(x_1, \ldots, x_N, x_{N+1})dx_{N+1} = \rho(x_1, \ldots, x_N).
$$

In case $k = N + 1 - l$ we have

$$
\int I \rho(x_1, \ldots, x_{N+1-l}, y_1, \ldots, y_l)dx_1 \ldots dx_{N+1-l} = \int \left( \int_{x \in I_N} \rho(x, x_{N+1-l}, y_1, \ldots, y_l)dx \right) dx_{N-l} \frac{w(y_1, \ldots, y_l)}{\int w(y_1, \ldots, y_{l-1}, v)dv} = \int w(x_{N+1-l}, y_1, \ldots, y_{l-1})dx_{N-l} \frac{w(y_1, \ldots, y_l)}{\int w(y_1, \ldots, y_{l-1}, v)dv}.
$$

Making use of (20) we get

$$
\int I \rho(x_1, \ldots, x_{N+1-l}, y_1, \ldots, y_l)dx_1 \ldots dx_{N+1-l} = w(y_1, \ldots, y_l)
$$

that is (21) is proved.

Now we are in a position to finish Lemma’s proof. For $x$’s of the form $x = (x, 0, \ldots)$, $x \in I_N$, we can integrate $H$ against $\rho$:

$$
\int_{x \in I_N} H(x) \rho(x) \geq C.
$$

On the other hand using the definition of $H$ and the key property of $\rho$ we get

$$
C \leq \int_{x \in I_N} H(x) \rho(x) \leq (l - 1) ||b|| + (N - l + 1) \int_I b(y)w(y)dy.
$$

Since $N$ is arbitrary large, (18) contradicts to (22). □

**Corollary 3.3.** For a nonnegative polynomial $A$ there exist continuous functions $g_A$ and $\gamma_A$ such that

$$
h_A = g_A + \gamma_A \circ \tau - \gamma_A
$$

and $g_A \geq 0$.

**Proof.** Note that $H_A(J(n))$ are uniformly bounded from below. □

**Corollary 3.4.** Let $J$ be such that $p_n \to 1$ and $q_n \to 0$. Then

$$
H_A(J) := \sum_{k=0}^{m} (-a \log p_{k+1} + (\Phi(J) - \Phi(J_0))e_k, e_k) - \gamma_A + \sum_{k \geq 0} g_A \circ \tau^k.
$$
That is the series with positive terms $\sum_{k \geq 0} g_A \circ \tau^k$ converges if and only if the series $\sum_{k \geq 0} h_A \circ \tau^k$ converges.

**Proof.** We use representation (23) and continuity of $\gamma_A$. \hfill $\square$

4. Proof of the Main Theorem

Assume that for a given $J$ its spectral measure $\sigma$ is such that $\Lambda_A(J) < \infty$, see definition (5). Note that due to Denisov–Rakhmanov Theorem [4]

(25) $p_n(\sigma) \to 1, \quad q_n(\sigma) \to 0$

and we can use (24) as a definition of $H_A(J)$.

With the measure $\sigma$ let us associate a measure $\sigma_\epsilon$ that we get by using the following two regularizations. First, we add to its absolutely continuous part the component $\epsilon dx$, that is $(\sigma'_\epsilon)_{a.c.} = \sigma'_{a.c.} + \epsilon$. Second, we leave just a finite number of the spectral points outside of $[-2, 2]$, say, that one that belongs to $\mathbb{R} \setminus [-2-\epsilon, 2+\epsilon]$.

It is important that

(26) $p_n(\sigma_\epsilon) \to p_n(\sigma), \quad q_n(\sigma_\epsilon) \to q_n(\sigma)$

for a fixed $n$ as $\epsilon \to 0$. The measure $\sigma_\epsilon$ satisfies the conditions of Szegő’s Theorem, and therefore $\zeta^n P_n(z, \sigma_\epsilon) \to \Delta(z; \sigma_\epsilon)$ uniformly on compact subsets of $\mathbb{C} \setminus [-2, 2]$. Here $\Delta(z; \sigma_\epsilon)$ is defined by

$$\Delta(z; \sigma_\epsilon) = \exp \left\{ \sqrt{z^2 - 4} \int \frac{1}{x - z} \frac{d\lambda(x; \sigma_\epsilon)}{x^2 - 4} \right\}. $$

In other words

$$\log \Delta(z; J(n; \sigma_\epsilon)) \to \log \Delta(z; \sigma_\epsilon), \quad n \to \infty,$$

uniformly on $\mathbb{C} \setminus \text{supp}(\sigma_\epsilon)$.

Finally, since (all) coefficients in decomposition (3) of $\log \Delta(z; J(n; \sigma_\epsilon))$ at infinity converge to the corresponding coefficients of $\log \Delta(z; \sigma_\epsilon)$ we get

$$H_A(J_\epsilon(n)) \to \Lambda_A(J_\epsilon), \quad n \to \infty.$$ 

Evidently $\Lambda_A(J_\epsilon) \leq \Lambda_A(J)$. Therefore for every $\delta$ there exists $n_0$ such that

$$H_A(J_\epsilon(n)) \leq \Lambda_A(J) + \delta$$

for all $n \geq n_0$. Since in the case under consideration $H_A$ is (basically) a series with positive terms, we get that every partial sum is bounded

$$H^N_A(J_\epsilon(n)) \leq \Lambda_A(J) + \delta.$$ 

Note that the left–hand side does not depend on $n$ if $n$ is big enough. Thus

$$H^N_A(J_\epsilon) \leq \Lambda_A(J).$$

Now, for a fixed $N$ let us pass to the limit as $\epsilon \to 0$. Due to (20) and continuity of $g_A$, for all $N$

$$H^N_A(J) \leq \Lambda_A(J).$$
But this means that
\[ \limsup H_A(J(n)) = \limsup \Lambda_A(J(n)) \leq \Lambda_A(J). \]

Using Corollary 2.3 we get
\[ H_A(J) = \lim H_A(J(n)) = \lim \Lambda_A(J(n)) = \Lambda_A(J). \]

Finally, starting with the condition that series \((11)\) converges we conclude that
\[ \limsup H_A(J(n)) = \limsup \Lambda_A(J(n)) < \infty. \]
Therefore, due to Corollary 2.3, we have \(\Lambda_A(J) < \infty\) and this completes the proof.

5. Asymptotic of orthonormal polynomials

Proof of Theorem 1.5. First let us mention that simultaneously with the convergence
\[ \Lambda(J(n)) = \int d\lambda_n \rightarrow \Lambda(J) = \int d\lambda, \]
we proved

\[ (27) \quad \lim_{n \rightarrow \infty} \int P(x) d\lambda_n(x) = \int P(x) d\lambda(x) \]

for every \(P(x) = Q^2(x)\) and hence \((27)\) holds for all polynomials. Since the variations of \(\lambda_n\)'s are uniformly bounded and since there is a finite interval \([\alpha_1, \alpha_2]\) containing the support of each measure \(\lambda_n\) in the family, \(\lambda_n\) converges weakly to \(\lambda\).

We will estimate the difference
\[ \left| \int \frac{d\lambda_n}{x-z} - \int \frac{d\lambda}{x-z} \right| \]
on a system of contours of the form
\[ \tau = \{z = x + iy : a \leq x \leq b, \ y = \pm c; \ |y| \leq c, \ x = a, b \} \]
that shrink to the interval \([-2, 2]\).

Integrating by parts, on a horizontal line we have
\[
\left| \int \frac{(\lambda - \lambda_n) dx}{(x-z)^2} \right| \leq \int_{\alpha_1}^{\alpha_2} \frac{|\lambda - \lambda_n| dx}{c^2} + |\lambda(\alpha_2) - \lambda_n(\alpha_2)| \int_{\alpha_2}^{\infty} \frac{dx}{|x-z|^2} \\
\leq \int_{\alpha_1}^{\alpha_2} \frac{|\lambda - \lambda_n| dx}{c^2} + |\lambda(\alpha_2) - \lambda_n(\alpha_2)|. 
\]

Since the \(\lambda_n(x)\) are uniformly bounded and \(\lim_{n \rightarrow \infty} \lambda_n(x) = \lambda(x)\) for all \(x\), the above estimate shows that for every \(\epsilon > 0\) there exists \(n_0\) such that
\[ \left| \int \frac{d\lambda_n}{x-z} - \int \frac{d\lambda}{x-z} \right| \leq \epsilon, \ n \geq n_0, \]
when \(z\) runs on a horizontal line of the contour \(\tau\).

Next, let us consider, say, the right vertical line on \(\tau\). Assume that \(b\) is between of two consequent points \(x_{k+1} < x_k\) of the set \(X\). We can even specify \(b = (x_{k+1} + x_k)/2\). The point is that starting with a suitable \(n\) the interval \([b - \delta/2, b + \delta/2]\) is in a gap of the support of \(\lambda - \lambda_n\). Here \(\delta := (x_k - x_{k+1})/2\). Put \(\hat{\lambda}(x) = \lambda(x) - \lambda(b)\)
and \( \tilde{\lambda}_n(x) = \lambda_n(x) - \lambda_n(b) \). Doing basically the same as on a horizontal line, we get

\[
\left| \int_{b+\delta/2}^{\infty} \frac{\tilde{\lambda} - \tilde{\lambda}_n}{(x - z)^2} \, dx \right| \leq \int_{b+\delta/2}^{\infty} \frac{|\tilde{\lambda} - \tilde{\lambda}_n|}{(\delta/2)^2} \, dx + |\tilde{\lambda}(\alpha_2) - \tilde{\lambda}_n(\alpha_2)| \int_{\alpha_2}^{\infty} \frac{dx}{|x - z|^2},
\]

and the same estimation for \( \int_{-\infty}^{b-\delta/2} \).

In other words the estimation

(28) \[
\left| A(z) \sqrt{z^2 - 4 \log \Delta_n(z)} - B_n(z) - \int_{x - z}^{\infty} \frac{d\lambda}{x - \lambda} \right| \leq \epsilon
\]

holds on the rectangle \( \tau \) if \( n \geq n_0 \).

Introduce the holomorphic function \( D(z) \) by (18), \( z \in \mathbb{C} \setminus [-2, 2] \), and consider the difference

\[
\left| \Delta_n(z) e^{\frac{B_n(z)}{A(z) \sqrt{z^2 - 4}}} - D(z) \right| = |D(z)| \left| \frac{A(z) \sqrt{z^2 - 4} \log \Delta_n(z) - B_n(z) - \int_{x - z}^{\infty} \frac{d\lambda}{x - \lambda}}{A(z) \sqrt{z^2 - 4}} \right| - 1
\]

on the contour \( \tau \). Due to (28) the difference is uniformly small on the contour and therefore also in the exterior of the rectangle.

Thus we have

(29) \[
\zeta^n (p_n P_n(z) - \zeta \tilde{P}_{n-1}(z)) \exp \left( -\frac{B_n(z)}{A(z) \sqrt{z^2 - 4}} \right) \to D(z)
\]

uniformly in the domain \( \mathbb{C} \setminus [-2, 2] \). Let us derive from this an asymptotic for the orthonormal polynomials properly.

First of all due to (24) we have (21)

\[
\frac{P_{n-1}(z)}{p_n P_n(z)} \to \zeta
\]

uniformly in \( \mathbb{C} \setminus [-2, 2] \). Therefore from (24) we get

(30) \[
\zeta^n P_n(z) \exp \left( -\frac{B_n(z)}{A(z) \sqrt{z^2 - 4}} \right) \to \frac{D(z)}{1 - \zeta^2}.
\]

Next we will adjust a bit the polynomials \( B_n \) in (30).

Let \( \tilde{J}(n) \) be \( n \times n \) matrix with coefficients \( p_k, q_k \), respectively \( \tilde{J}_0(n) \) is \( n \) by \( n \) matrix that we obtain cutting the Chebyshev matrix \( J_0 \). Recall that

\[
P_n(z) = \frac{1}{p_1 \ldots p_n} \det(z - \tilde{J}(n))
\]

in particular

\[
\det(z - \tilde{J}_0(n)) = \frac{\zeta^{-n-1} - \zeta^{n+1}}{\zeta^{-1} - \zeta}.
\]

That is

\[
\frac{1}{p_1 \ldots p_n} \det(z - \tilde{J}_0(n)) = (\zeta^{-1} - \zeta) \frac{\zeta^{n+1} P_n(z)}{1 - \zeta^{2n+2}}.
\]
and hence
\[
\log(\zeta^{n+1} \sqrt{\pi^2 - 4P_n(z)}) = -\log(p_1...p_n) - \frac{\text{tr}(\bar{J}(n) - \bar{J}_0(n))}{z} - \frac{\text{tr}(\bar{J}^2(n) - \bar{J}_0^2(n))}{2z^2} - \ldots.
\]

Thus we can substitute \(B_n(z)\) by the polynomial \(\tilde{B}_n(z)\), which is uniquely defined by
\[
\log(\zeta^{n+1} \sqrt{\pi^2 - 4P_n(z)}) - \tilde{B}_n(z)A(z)\sqrt{\pi^2 - 4} = O\left(\frac{1}{z^{m+2}}\right).
\]

by condition (25) for any fixed \(k\)
\[
\text{tr}(\bar{J}^k(n) - \bar{J}_0^k(n)) \rightarrow 0, \quad n \rightarrow \infty.
\]

\[\square\]

6. Appendix: Laptev–Naboko–Safronov Example

It is more convenient (uniform) to use two sided Jacobi matrices acting in \(l^2(\mathbb{Z})\).

In particular, then the function \(H_{\mathcal{A}}(J)\) is positive.

6.1. Positive definite Hankel minus Toeplitz. Recall that the Chebyshev polynomials of the second kind \(U_l(z)\) form an orthogonal system with respect to the weight \(\sqrt{\pi^2 - 4}\),

\[
\frac{1}{\pi} \int_{-2}^{2} U_l(x)U_k(x)\sqrt{\pi^2 - 4 - x^2} \, dx = 2\delta_{k,l},
\]

where

\[
U_l(z) := \frac{\zeta^{-l} - \zeta^l}{\zeta^{-1} - \zeta}, \quad z = \zeta^{-1} + \zeta.
\]

Note also that the following map transforms the polynomials of the second kind into the Chebyshev polynomials of the first kind

\[
zU_l(z) = \frac{1}{\pi} \int_{-2}^{2} \frac{U_l(x) - U_l(z)}{x - z} \sqrt{\pi^2 - 4 - x^2} \, dx = T_l(z).
\]

**Lemma 6.1.** For \(m \neq n\)

\[
H_{U_mU_n}(J) = \text{tr} \left\{ \frac{T_{m+n}}{m+n} \frac{T_{m-n}}{|m-n|} \right\}_j,
\]

and

\[
H_{U_m^2}(J) = \text{tr} \left\{ \frac{T_{2n}}{2n} - \sum_i \log p_i^2 \right\}_j = \text{tr} \left\{ \frac{T_{2n}^2}{2n} - \sum_i \log p_i^2 \right\}_j.
\]

**Proof.** We have

\[
\Phi'(z) = zU_m(z)U_n(z) - \frac{1}{\pi} \int U_m(x)U_n(x)\sqrt{\pi^2 - 4 - x^2} \, dx
\]

\[
- \frac{1}{\pi} \int U_m(x)U_n(z)\sqrt{\pi^2 - 4 - x^2} \, dx.
\]
Using (31), (32), (33) we have for $m > n$
\[
\Phi'(z) = zU_m(z)U_n(z) - \frac{1}{\pi} \int \frac{U_m(x) - U_m(z)}{x - z} \sqrt{4 - x^2} dx U_n(z) = T_m(z)U_n(z) - U_{m+n}(z) - U_{m-n}(z).
\]
Since $T'_k = kU_k$, $k \geq 1$, we get
\[
\Phi(z) = \frac{T_{m+n}(z)}{m+n} - \frac{T_{m-n}(z)}{m-n} + \text{const}.
\]
By orthogonality also
\[
a = \frac{1}{\pi} \int_{-2}^{2} U_m(x)U_n(x) \sqrt{4 - x^2} dx = 0.
\]
Thus (34) is proved. A proof of (35) requires just a minor modification. □

**Proposition 6.2.** Let $J$ be a finite dimensional perturbation of $J_0$. Define
\[
a_k(J) = \begin{cases} 
\text{tr}\left\{ \frac{T_k}{k} \right\}_{J_0}, & k \geq 1 \\
\sum_i \log p_i^k & k = 0
\end{cases}
\]
Then the matrix $\{a_{k+l}(J) - a_{k-l}(J)\}_{k,l \geq 1}$ is positive.

**Proof.** Put $A = |B|^2$ with $B = \sum_i U_i c_i$. Since $H_A(J) \geq 0$, due to Lemma 6.1 we get
\[
\sum_{k \geq 1, l \geq 1} \{a_{k+l}(J) - a_{k-l}(J)\} c_k c_l \geq 0.
\]
□

Note that continuous positive kernels of this kind are a classical object, see e.g. [1].

6.2. Laptev–Naboko–Safronov example: $A = U_l^2$. This case was considered in [2].

**Proposition 6.3.** Let $A(z) = U_l^2(z)$. Then $\Lambda_A(J) < \infty$ if and only if $T_l(J) - T_l(J_0)$ is Hilbert–Schmidt.

**Proof.** Due to Lemma 6.1
\[
H_A(J) = \text{tr} \frac{T_l^2(J) - T_l^2(J_0)}{2l} - 2 \sum \log p_i.
\]
Note that a row in the matrix $T_l(J)$ is of the form
\[
\langle e_i | T_l(J) = [\ldots, 0, (t_i)_{i-l}^i, (q_i)_{i-l}^i, \ldots],
\]
where $(t_i)_i = p_{i+1}p_{i+2}\ldots p_{i+l}$ and $(q_i)_i$ is a row–vector of dimension $2l - 1$. Therefore
\[
H_A(J) = \frac{1}{l} \left\{ \sum \frac{(q_i)_i}{2} (q_i)_i^* + \sum ((t_i)_l^2 - 1 - \log(t_i)_l^2) \right\}
\]
and the condition $H_A(J) < \infty$ is equivalent to $T_l(J) - T_l(J_0)$ is a Hilbert–Schmidt operator. □

It is possible to reformulate the above condition in terms of the coefficient sequences of $J$. 

Theorem 6.4. Let $A(z) = U_n^2(z)$. Then $\Lambda_A(J) < \infty$ if and only if
\begin{equation}
\{ \sum_{k=1}^{n} u_{j+k} \} \in l^2, \quad \{ \sum_{k=1}^{n} q_{j+k} \} \in l^2, \quad \{ u_j^2 \} \in l^2, \quad \{ q_j^2 \} \in l^2,
\end{equation}
where $u_j = p_j^2 - 1$.

A proof is split into several lemmas.

Lemma 6.5. Let $J = S^{-1}P + Q + PS$ and
\[ T_n(J) = \{ ... + \Lambda_0(n) + \Lambda_1(n) S + ... + \Lambda_n(n) S^n \}, \]
where $Q, P, \Lambda_k(n)$ are diagonal matrices. Then
\begin{align}
\Lambda_n(n) &= P \cdots P(-1) \cdots P(-n+1) \\
\Lambda_{n-1}(n) &= P \cdots P(-n+2) \{ Q + Q(-1) + ... + Q(-n+1) \}
\end{align}
and
\begin{equation}
\Lambda_{n-2}(n) = P \cdots P(-n+3) \{ ([P(1)]^2 - I + P^2 - I + ... + P(-n+3))^2 - I \}
+ Q[Q + Q(-1) + ... + Q(-n+2)] + Q(-1)[Q(-1) + ... + Q(-n+2)] + ...
+ Q(-n+2)Q(-n+2).
\end{equation}

Proof. All three formulas can be proved by induction using
\[ T_n(J) = JT_{n-1}(J) - T_{n-2}. \]

Let us prove (40). We have
\[ \Lambda_{n-2}(n) = S^{-1}P\Lambda_{n-1}(n-1)S + Q\Lambda_{n-2}(n-1) + PS\Lambda_{n-3}(n-1)S^{-1} - \Lambda_{n-2}(n-2). \]

Substituting (38) and (39) we get
\[ \Lambda_{n-2}(n) = S^{-1}PP\cdots P(-1) \cdots P(-n+2)S \\
+ QP \cdots P(-n+3) \{ Q + Q(-1) + ... + Q(-n+2) \} \\
+ PS\Lambda_{n-3}(n-1)S^{-1} - PP \cdots P(-1) \cdots P(-n+3) \\
= P \cdots P(-1) \cdots P(-n+3) \{ ([P(1)]^2 - I + Q + Q(-1) + ... + Q(-n+2)] \\
+ P\Lambda_{n-3}(n-1) \}. \]

Iterating the last relation we obtain (40). \qed

Lemma 6.6. If $T_n(J) - T_n(J_0)$ is Hilbert–Schmidt then relations (37) are fulfilled.

Proof. Since $\Lambda_n(n) - I$, $\Lambda_{n-1}(n)$ and $\Lambda_{n-2}(n)$ are Hilbert–Schmidt operators, using Lemma (55) we have
\begin{align}
\{ p_{1+i} \cdots p_{n+i} - 1 \} \in l^2 \\
\{ p_{1+i} \cdots p_{n-1+i} (q_i + \cdots + q_{n-1+i}) \} \in l^2
\end{align}
Having in mind (41) we simplify (42) and (43)

\[ \{q_i + \ldots + q_{n-1+i}\} \in l^2 \]

and

\[
\left\{ \sum_{k=1}^{i+n-1} (p_k^2 - 1) + \frac{1}{2} \sum_{k=i}^{i+n-2} q_k^2 + \frac{1}{2} \left( \sum_{k=i}^{i+n-2} q_k \right)^2 \right\} \in l^2.
\]

Now we wish to separate “p” and “q” conditions in (44). It is evident that \(a + b \in l^2\) implies \(a \in l^2\) and \(b \in l^2\) if only \(a_i \geq 0\) and \(b_i \geq 0\). Note that (44) implies \((p_1, \ldots, p_{n+i})^2/n - 1 \in l^2\). Thus using this condition and the inequality

\[
p_{i+n+i}^2 + \ldots + p_{n+i}^2/n - n \geq (p_1, \ldots, p_{n+i})^2/n - 1
\]

we get from (44) \(\{q_i^2\} \in l^2\) and \(\{\sum_{k=1}^{n} (p_{i+k}^2 - 1)\} \in l^2\).

Finally we note that

\[
(p_1 - 1)^2 + \ldots + (p_n - 1)^2 = (p_1^2 - 1) + \ldots + (p_n^2 - 1) - 2(p_1 - 1) + \ldots + (p_n - 1).
\]

Since

\[
2n(p_1 \ldots p_n)^{1/n} - 1 \leq 2(p_1 - 1) + \ldots + (p_n - 1)
\]

we have \(\{\sum_{k=1}^{n} (p_{i+k} - 1)\} \in l^2\) and therefore \((p_i - 1)^2 \in l^2\).

\[ \square \]

The following lemma can be shown by induction.

**Lemma 6.7.** Let \(J = J_0 + dJ\) then

\[
dT_J(J_0) e_0 = \sum_{k=0}^{l-1} S^{1-l} S^k [dJ + dJ^{(1-l)}] S^k e_0 =
\]

\[
\begin{bmatrix}
0 \\
p_{l+1} + \ldots + dp_0 \\
dq_{l+1} + \ldots + dq_0 \\
2dp_{l+2} + \ldots + 2dp_1 \\
dq_{l+2} + \ldots + dq_1 \\
2dp_{l+3} + \ldots + 2dp_2 \\
\vdots \\
dp_1 + \ldots + dp_l \\
0
\end{bmatrix}
\]

**Proof of the Theorem 6.4.** We only have to show that conditions (47) imply \(T(J) - T(J_0)\) is Hilbert–Schmidt. Note that each entry is a polynomial of \(q_j, u_i\) with \(u_i = p_i - 1\). Moreover, the linear term is described in Lemma 6.7. Note also that the sequences \(\{u_i q_{i+k}^j\}_i\), \(\{u_i q_{i+k}^j\}_i\), \(\{q_{i+k}^j\}_i\) belong to \(l^2\) for \(k + l \geq 2\). Thus, having in mind the structure of the matrix \(T(J) - T(J_0)\), we get that each diagonal forms an \(l^2\)-sequence, as was to be proved.

\[ \square \]
6.3. Simon’s conjecture. Since $H_A(J_0) = 0$ and $H_A(J) \geq 0$ the decomposition of $H_A$ about $J_0$ begins with a quadratic form, more exactly:

**Lemma 6.8.** Let $J = J_0 + dJ$ then the decomposition of $H_A$ about $J_0$ begins with

$$H_A(J) = \frac{1}{2} \langle dj | A(J_0) | dj \rangle + \ldots$$

where $\langle dj \rangle = \{ \ldots , 2dp_0 , dq_0 , 2dp_1 , dq_1 , \ldots \}$.

**Proof.** We start with the formula

$$dH_A(J) = \text{tr} \{ A(J) \text{Re}(Z^{-1} - Z) \, dJ \},$$

where $Z$ is the lower triangle solution of the equation $Z^{-1} + Z = J$. Note that the decomposition of $Z^{-1} - Z$ about $J_0$ is of the form

$$Z^{-1} - Z = S^{-1} - S + dJ - 2dZ + \ldots .$$

Using $dJ = -Z^{-1} dZZ^{-1} + dZ$ we get

$$-dZ |_{Z=S} = [ZdJZ + Z(-dZ)Z] |_{Z=S} = SdJS + S^2dJS^2 + \ldots .$$

Therefore the leading term in the decomposition of $\text{Re}(Z^{-1} - Z)$ is the Hankel operator

$$\Gamma = \ldots + S^{-1}dJS^{-1} + dJ + SdJS + \ldots ,$$

and

$$H_A(J) = \frac{1}{2} \text{tr} \{ A(J_0) \Gamma \, dJ \} + \ldots .$$

Let us mention that $\Gamma e_0 = dj$, thus we can rewrite this Hankel operator into the form

$$\Gamma = \sum S^k |dj \rangle \langle e_0 | S^k .$$

Since $A(J_0)$ and $S$ commute and $\Gamma S = S^{-1} \Gamma$ we get

$$\text{tr} \{ A(J_0) \Gamma \, dJ \} = \text{tr} \{ A(J_0) \Gamma (S^{-1} dP + dQ + dPS) \} = \text{tr} \{ A(J_0) \Gamma (2S^{-1} dP + dQ) \} .$$

Substituting $\Gamma$ we obtain

$$\text{tr} \{ A(J_0) \Gamma \, dJ \} = \text{tr} \{ A(J_0) (\sum S^k |dj \rangle \langle e_0 | S^k) (2S^{-1} dP + dQ) \}$$

$$= \text{tr} \{ A(J_0) |dj \rangle \langle e_0 | \sum (2S^{k-1} dPS^k + S^k dQS^k) \} .$$

But $\langle e_0 | \sum (2S^{k-1} dPS^k + S^k dQS^k) = \langle dj \rangle$ and this completes the proof. \qed

We believe that related to this quadratic form condition

$$\langle A(J_0) \, dJ , dJ \rangle < \infty ,$$

should play an important role in a counterpart of Simon’s conjecture formulated for the unit circle in several talks, for example [12]. Specifically, in Laptev–Naboko–Safronov case, where

$$A(J_0) = (I + S^2 + \ldots + S^{2l-2})^* (I + S^2 + \ldots + S^{2l-2}) ,$$
condition (17) means
\[
\{ dq_{i+1} + dq_{i+2} + \ldots + dq_{i+l} \} \in l^2(\mathbb{Z}),
\]
\[
\{ 2dp_{i+1} + 2dp_{i+2} + \ldots + 2dp_{i+l} \} \in l^2(\mathbb{Z}),
\]
compare (37).

REFERENCES

1. N. I. Akhiezer, The classical moment problem and some related questions in analysis. Translated by N. Kemmer, Hafner Publishing Co., New York 1965.
2. D. Damanik, R. Killip and B. Simon, Necessary and sufficient conditions in the spectral theory of Jacobi matrices and Schrödinger operators, [arXiv:math.SP/0309206]
3. P. Deift and R. Killip, On the absolutely continuous spectrum of one-dimensional Schrödinger operators with square summable potentials, Comm. Math. Phys. 203 (1999), no. 2, 341–347.
4. S. Denisov, On Rakhmanov’s theorem for Jacobi matrices, Proc. Amer. Math. Soc. 132 (2004), 847–852.
5. S. Denisov and S. Kupin, Non-Szego asymptotics for orthogonal polynomials on the unit circle, in preparation.
6. D. Hundertmark and B. Simon, Lieb-Thirring inequalities for Jacobi matrices, J. Approx. Theory 118 (2002), no. 1, 106–130.
7. R. Killip and B. Simon, Sum rules for Jacobi matrices and their applications to spectral theory, Ann. of Math. (2) 158 2003, 253–321.
8. S. Kupin, On sum rules of special form for Jacobi matrices, C. R. Math. Acad. Sci. Paris 336 (2003), no. 7, 611–614.
9. A. Laptev, S. Naboko and O. Safronov, On new relations between spectral properties of Jacobi matrices and their coefficients, Comm. Math. Phys. 241 (2003), 91–110.
10. F. Peherstorfer and P. Yuditskii, Asymptotics of orthonormal polynomials in the presence of a denumerable set of mass points, Proc. Amer. Math. Soc. 129 (2001), no. 11, 3213–3220.
11. E. Rakhmanov, The asymptotic behavior of the ratio of orthogonal polynomials, II. (Russian) Mat. Sb. (N.S.) 118(160) (1982), no. 1, 104–117, 143.
12. B. Simon, Sum Rules and the Spectral Theory of Orthogonal Polynomials, OPSFA, Copenhagen, August, 2003.
13. K. Yosida, Functional analysis. Third Edition. Die Grundlehren der mathematischen Wissenschaften, Band 123. Springer-Verlag, New York-Heidelberg, 1971.