A Matrix View on the Root Distribution for Orthogonal Polynomials *

To my Italian friends in Pisa and around

E. Tyrtyshnikov †

Abstract

A new message is a simple matrix-based proof of the equal distribution property for the roots of polynomials orthogonal on an interval.

1 Introduction

Consider polynomials $p_n(x) = p_{0n} + p_{1n}x + \ldots + p_{nn}x^n$ (with all the coefficients real and the senior one positive) orthogonal in the following sense:

$$\int_{-1}^{1} p_m(x)p_n(x)d\sigma(x) = \delta_{mn}.$$

(As usual, δ_{mn} is 1 for $m = n$, and 0 otherwise.)

Here, $\sigma(x)$ is a monotonic function with infinitely many points of growth. For such a $\sigma(x)$, we always have a uniquely determined infinite sequence of orthogonal polynomials.

The elementary facts about the orthogonal polynomials are the following:

- They satisfy the three-term recurrences

$$xp_n(x) = \beta_{n-1}p_{n-1}(x) + \alpha_n p_n(x) + \beta_n p_{n+1}(x).$$

- Their roots are all inside $[-1, 1]$, and, moreover, the roots of $p_n(x)$ and $p_{n+1}(x)$ satisfy the so-called interlacing inequalities.

Any other fact from a good lot of them [1, 5] is not that elementary. Not in the least this applies to the celebrated equal distribution property of the roots for any $\sigma(x)$ from a pretty wide class.

---

*This work was perfomed when the author was a visiting professor at Dipartimento di Matematica, Università di Pisa, supported by the CNR, Italy. We also acknowledge the support, in part, by the Russian Fund of Basic Research and Volkswagen-Stiftung.

†Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina 8, Moscow 117333, Russia

To be precise, recall that two sequences of n-tuples of real numbers, $\{a_{in}\}_{i=1}^n$ and $\{b_{in}\}_{i=1}^n$, are said to be equally distributed if, for any $F(x)$ continuous with a bounded support,

$$\lim_{n\to\infty} \left(\frac{1}{n} \sum_{i=1}^{n} F(a_{in}) - \frac{1}{n} \sum_{i=1}^{n} F(b_{in}) \right) = 0.$$

With this definition, we have the following theorem.

THEOREM 1.1. *Assume that a_{in} and b_{in} are the roots of orthogonal polynomials associated with any two functions $\sigma_a(x)$ and $\sigma_b(x)$ providing that*

$$\int_{-1}^{1} \frac{\log \sigma'(x)}{\sqrt{1-x^2}} dx > -\infty \tag{1}$$

(for $\sigma = \sigma_a$ and σ_b). Then the roots a_{in} and b_{in} are equally distributed.

The proof is given in [5, 2]. However, it involves many nontrivial constructions far short of the matrix analysis framework. In this note, we propose a simple matrix-based proof.

2 Main result

Below we present a new proof of Theorem 1.1.

1. The roots of $p_n(x)$ coincide with the eigenvalues of the following Hermitian tridiagonal matrix:

$$A_n = \begin{bmatrix} \alpha_0 & \beta_1 & & \\ \beta_1 & \alpha_1 & \beta_2 & \\ \dots & \dots & \dots & \\ & \beta_{n-2} & \alpha_{n-2} & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}. \tag{2}$$

   (This can be found anywhere, for example, in [7].)

2. Given two sequences of Hermitian matrices A_n and B_n, assume that there exist Hermitian matrices E_n and R_n such that

$$A_n - B_n = E_n + R_n, \quad ||E_n||_F^2 = o(n), \quad \operatorname{rank} R_n = o(n). \tag{3}$$

   Then the eigenvalues of A_n and B_n are equally distributed.

   (This is the matrix indication for the equal distribution property proposed in [6].)

   It is more convenient, sometimes, to prove that, for any $\varepsilon > 0$, it is possible to split $A_n - B_n = E_{\varepsilon n} + R_{\varepsilon n}$ so that $||E_{\varepsilon n}||_F^2 \leq \varepsilon n$ and $\operatorname{rank} R_{\varepsilon n} \leq \varepsilon n$ for all n sufficiently large. This implies that there exist some E_n and R_n satisfying (3).

3. For the Chebyshev polynomials, we have

$$\alpha_n = 0, \quad \beta_n = \frac{1}{2} \quad \forall\, n.$$

Their roots are the eigenvalues of

$$B_n = \begin{bmatrix} 0 & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ \dots & \dots & \dots & & \\ & \frac{1}{2} & 0 & \frac{1}{2} \\ & & \frac{1}{2} & 0 \end{bmatrix}. \tag{4}$$

4. If the relation (1) is fulfilled, then [5]

$$\lim_{n\to\infty} \alpha_n = 0, \quad \lim_{n\to\infty} \beta_n = \frac{1}{2}. \tag{5}$$

5. Using (5), it is an easy matter to show that, for the matrices defined by (2) and (4), we have exactly (3). Let $\varepsilon > 0$, and take up $N = N(\varepsilon)$ such that $|\alpha_n| \le \varepsilon, \quad |\beta_n - \frac{1}{2}| \le \varepsilon$ for all $n \ge N$. Set

$$E_{\varepsilon n} \equiv \begin{bmatrix} 0 & 0 \\ 0 & \{A_n - B_n\}_{i,j=N+1}^{n} \end{bmatrix}, \quad R_{\varepsilon n} \equiv A_n - B_n - E_{\varepsilon n}.$$

Then $A_n - B_n = E_{\varepsilon n} + R_{\varepsilon n}$ and, obviously,

$$||E_{\varepsilon n}||_F^2 \le \varepsilon\, n, \quad \operatorname{rank} R_{\varepsilon n} \le 2N.$$

Since ε is arbitrary, this implies (3).

We conclude immediately that the roots of any orthogonal polynomials are distributed equally with the roots of Chebyshev polynomials. Thus, the proof of Theorem 1.1 is completed.

We could stop at this. However, since the proof of the limiting relations (5) in [5] is somewhat "distributed" through several chapters, it might be useful to sketch a straightforward way to them. We will do this in the next section.

3 Why 0 and $\frac{1}{2}$?

We answer the question through a sequence of 7 articles as follows.

1. Consider the orthogonal polynomials $\phi_n(z) = \phi_{0n} + \phi_{1n}z + \ldots + \phi_{nn}z^n$ on a circle (let $\phi_{nn} > 0$):

$$(\phi_m, \phi_n) \equiv \frac{1}{2\pi} \int_0^{2\pi} \phi_m(e^{\mathrm{i}t})\, \overline{\phi_n(e^{\mathrm{i}t})}\, d\mu(t) = \delta_{mn}. \tag{6}$$

Denote by M_n the matrix of moments:

$$M_n = [m_{kl}] = [(e^{\mathrm{i}kt}, e^{\mathrm{i}lt})]_{kl=0}^{n}.$$

Then, from (6) it readily follows that

$$\begin{bmatrix} \phi_{00} & & & \\ \phi_{01} & \phi_{11} & & \\ \dots & \dots & \dots & \\ \phi_{0n} & \phi_{1n} & \dots & \phi_{nn} \end{bmatrix} M_n \begin{bmatrix} \overline{\phi}_{00} & \overline{\phi}_{01} & \dots & \overline{\phi}_{0n} \\ & \overline{\phi}_{11} & \dots & \overline{\phi}_{1n} \\ & \dots & \dots & \dots \\ & & & \overline{\phi}_{nn} \end{bmatrix} = I.$$

Therefore, the columns

$$y_n = \phi_{nn} \begin{bmatrix} \overline{\phi}_{0n} \\ \dots \\ \overline{\phi}_{nn} \end{bmatrix}, \qquad x_n = J_n\, \overline{y}_n = \phi_{nn} \begin{bmatrix} \phi_{nn} \\ \dots \\ \phi_{0n} \end{bmatrix}$$

are the last and the first ones in M_n^{-1}. Here, J_n puts the entries into the reverse order. (The assertion for y_n stems directly from the above matrix equation. Concerning x_n, we need to recognize in M_n a Toepitz matrix, that implies $M_n^T = J_n M_n J_n$ and, hence, $(M_n^{-1})^T = J_n (M_n^{-1})^T J_n$. Take into account also that $M_n = M_n^*$.)

2. In the field of matrix computations, it is well-known that, for the inverses to nonsingular Toeplitz matrices embedded to some larger matrix, x_{n+1} and y_{n+1} are some linear combinations of specifically augmented x_n and y_n (this is behind Levinson's algorithm; see [3, 4]):

$$x_{n+1} = c_n \begin{bmatrix} x_n \\ 0 \end{bmatrix} + d_n \begin{bmatrix} 0 \\ y_n \end{bmatrix}. \tag{7}$$

(The equation for y_{n+1} can be obtained from this by applying J_{n+1} to both sides.)

The appearance of (7) suggests that we have some relations between $\phi_n(z)$ and $\phi_n^*(z)$, their reversal polynomials, defined as follows:

$$\phi_n^*(z) \;=\; \overline{\phi}_{nn} + \ldots + \overline{\phi}_{1n} z^{n-1} + \overline{\phi}_{0n} z^n \;=\; z^n\, \overline{\phi}_n(z^{-1}).$$

The familiar two-term recurrences come up directly from (7):

$$\phi_{n+1}(z) = \gamma_n z \phi_n(z) + \delta_n \phi_n^*(z), \quad \phi_{n+1}^*(z) = \overline{\delta}_n z \phi_n(z) + \gamma_n \phi_n^*(z), \tag{8}$$

where, as is easy to check,

$$\gamma_n = \frac{\phi_{n+1\, n+1}}{\phi_{nn}}, \quad \delta_n = \frac{\phi_{0\, n+1}}{\phi_{nn}}.$$

3. If we would have wanted to make a step towards computations, we need to determine the unknown coefficients in (7). To this end, multiply (7) by M_{n+1}, and then, since M_{n+1} is Toeplitz,

$$1 = c_n + d_n G_n, \qquad 0 = c_n \overline{G}_n + d_n,$$

where, as is readily seen,

$$G_n = m_{-1} y_{0n} + \ldots + m_{-n} y_{nn} = \overline{m_n x_{0n} + \ldots + m_1 x_{nn}}.$$

It follows that

$$c_n = \frac{1}{1 - |G_n|^2}, \quad d_n = -\overline{G}_n c_n \quad \Rightarrow \quad c_n^2 - |d_n|^2 = c_n. \tag{9}$$

Consequently,

$$\phi_{n+1\, n+1}^2 - \phi_{nn}^2 = |\phi_{0 n+1}|^2 \quad \Rightarrow \quad \gamma_n^2 - |\delta_n|^2 = 1. \tag{10}$$

It implies that the senior coefficient ϕ_{nn} increases monotonically.

4. We want to have ϕ_{nn} converge to some finite limit. In line with the above, it is sufficient to know that ϕ_{nn} is upper-bounded. For this, we make use of a short yet elegant chain of inequalities (due to Szegö) as follows:

$$\begin{aligned}\phi_{nn}^{-2} &= \frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\phi_n^*(e^{\mathrm{i}t})}{\phi_{nn}}\right|^2 d\mu(t) \geq \frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\phi_n^*(e^{\mathrm{i}t})}{\phi_{nn}}\right|^2 \mu'(t)\, d\mu(t)\\ &\geq \exp\left\{\frac{1}{2\pi}\int_0^{2\pi}\log\left(\left|\frac{\phi_n^*(e^{\mathrm{i}t})}{\phi_{nn}}\right|^2 \mu'(t)\right) d\mu(t)\right\}\\ &= \exp\left\{\frac{1}{2\pi}\int_0^{2\pi}\log\mu'(t)\, dt\right\}.\end{aligned}$$

(Some explanation, for the last passage, will appear in a minute.)

It is clear now that if we assume that

$$\frac{1}{2\pi}\int_0^{2\pi}\log\mu'(t)dt \;>\; -\infty, \tag{11}$$

then, for some finite number Φ,

$$\phi_{nn}\to\Phi. \quad\Rightarrow\quad \phi_{0n}\to 0,\quad \gamma_n\to 1,\quad \delta_n\to 0. \tag{12}$$

Moreover, allowing for (10), we have

$$\sum_{n=0}^{\infty}|\delta_n|^2 \;<\; +\infty, \qquad \sum_{n=0}^{\infty}|\phi_{0n}|^2 \;<\; +\infty. \tag{13}$$

Note that, in the above chain, we use that

$$\exp\left\{\frac{1}{2\pi}\int_0^{2\pi}\log\left(\left|\frac{\phi_n^*(e^{\mathrm{i}t})}{\phi_{nn}}\right|^2\right) dt\right\} \;=\; 1,$$

for $\phi_n^*(0)/\phi_{nn} = 1$ and the function under the integral is analytic. The latter emanates from the fact that all zeroes of $\phi_n^*(z)$ lie strictly outside the unit circle.

The proof for this is more or less straightforward: it is based chiefly on (8), and uses, in particular, the imposing Cristoffel–Darboux formula: [1]

$$\sum_{k=0}^{n}\phi_k(x)\,\overline{\phi_n(y)} = \frac{\phi_{n+1}^*(x)\,\overline{\phi_{n+1}^*(y)} - \phi_{n+1}(x)\,\overline{\phi_{n+1}(y)}}{1-x\overline{y}}. \tag{14}$$

Let $x = y$ and $\phi_{n+1}^*(x) = 0$. If $1-|x|^2 > 0$, then it contradicts (14). If $|x| = 1$, then $\phi_{n+1}(x) = \phi_{n+1}^*(x) = 0 \quad\Rightarrow\quad \phi_n(x) = \phi_n^*(x) = 0$, and so on. Eventually, we conclude that x is

[1] This can be readily derived from (8) by induction, and relies, above all, on the relation $\gamma_n^2 - |\delta_n|^2 = 1$, which we have derived above from Levinson's formulas.

the root for $\phi_0(x)$ and $\phi_0^*(x)$, which is impossible.

5. We need to prove, to boot, that

$$\phi_{1n} \to 0 \tag{15}$$

and

$$\phi_{n\,n+1} - \phi_{n-1\,n} \to 0. \tag{16}$$

On equating the coefficients at z and z^n in the first relation from (8),

$$\phi_{1\,n+1} = \gamma_n \phi_{0n} + \delta_n \overline{\phi}_{n-1\,n}, \quad \phi_{n\,n+1} = \gamma_n \phi_{n-1\,n} + \delta_n \overline{\phi}_{0n}.$$

Since $\gamma_n \to 1$, $\delta_n \to 0$, and $\phi_{0n} \to 0$, we obtain the said limits as soon as we show that $\phi_{n-1\,n}$ is bounded. Using the same relation from (8) recursively, we have

$$|\phi_{n-1\,n}| \ \le\ c_{n-1}\,|\delta_{n-1}|\,|\phi_{0\,n-1}| \ +\ \ldots\ +\ c_0\,|\delta_0|\,|\phi_{00}|,$$

where

$$c_j \ =\ \gamma_{n-1}\cdots\gamma_j \ \le\ \frac{\phi_{n-1\,n-1}}{\phi_{00}} \ \le\ \frac{\Phi}{\phi_{00}}.$$

The boundedness of $|\phi_{n-1\,n}|$ follows by applying the Cauchy inequality and recollecting (13).

6. For any $\sigma(x)$, the orthogonal polynomials $p_n(x)$ on $[-1,1]$ can be juxtaposed to some orthogonal polynomials $\phi_n(z)$ on the unit circle [1, 5]. The latter ones emerge if we take

$$\mu(t) = \begin{cases} -\sigma(\cos t),\ 0 \le t \le \pi, \\ \sigma(\cos t),\ \pi \le t \le 2\pi. \end{cases}$$

Since $\mu(t)$ is an odd function with respect to π, the matrix of moments M_n (see Article 1) is now a real symmetric positive definite matrix. Hence, from the Cholesky decomposition of M_n^{-1}, the coefficients of the polynomials orthogonal on the circle for $\mu(t)$ are real. Note also that, for any functions $f(x)$ and $g(x)$,

$$\int_{-1}^{1} p(x)g(x)\,d\sigma(x) \ =\ \frac{1}{2}\int_{0}^{2\pi} p(\cos x)g(\cos x)\,d\mu(t). \tag{17}$$

It is not difficult to verify that, for $z = e^{\mathrm{i}t}$,

$$\begin{aligned} \frac{\phi_{2n}(z) + \phi_{2n}^*(z)}{z^n} &= \sum_{k=0}^{2n} (\phi_{k\,2n} + \overline{\phi}_{2n-k,2n})\, e^{\mathrm{i}kt - \mathrm{i}nt} \\ &= 2\,\phi_{n\,2n} + 2\sum_{k=1}^{n} (\phi_{n-k\,2n} + \phi_{n+k,2n})\,\cos kt \end{aligned}$$

is a polynomial in

$$x = \frac{1}{2}(z + z^{-1}) = \cos t.$$

Any two of them, for different n, are orthogonal with respect to $\sigma(x)$, because

$$\int_0^{2\pi} \sum_{k=0}^{2n} (\phi_{k\,2n} + \phi_{2n-k\,2n})\, e^{i(k-n)}\, e^{imt}\, d\mu(t) =$$

$$\int_0^{2\pi} \phi_{2n}(e^{it})\, e^{i(m-n)}\, d\mu(t) + \overline{\int_0^{2\pi} \phi_{2n}(e^{it})\, e^{-i(m+n)}\, d\mu(t)} = 0$$

for $-n+1 \le m \le n-1$ (the first integral is equal to 0 for $-n+1 \le m \le n$ while the second one is for $-n \le m \le n-1$). More precisely [1, 5],

$$p_n(x) = \frac{\phi_{2n}(z) + \phi^*_{2n}(z)}{z^n}\, \frac{1}{c(1+\nu_n)}, \tag{18}$$

where

$$c = \sqrt{2\pi}, \quad 1 + \nu_n = \sqrt{1 + \frac{\phi_{\ 0\,2n}}{\phi_{\ 2n\,2n}}}.$$

Indeed, using (17), we have

$$\frac{1}{2} \int_0^{2\pi} |\phi_{2n}(e) + \overline{\phi}_{2n}\, e^{i\,2nt}|\, d\mu(t) = \int_0^{2\pi} |\phi_{2n}(e)|^2\, d\mu(t) + \mathrm{Re} \int_0^{2\pi} \phi^2_{2n}(e)\, e^{-i\,2nt}\, d\mu(t),$$

where the first integral is equal to 2π while the second is

$$\mathrm{Re} \int_0^{2\pi} \phi_{2n}(e)\, \phi_{0\,2n}\, e^{-i\,2nt}\, d\mu(t) = 2\pi\, \frac{\phi_{0\,2n}}{\phi_{2n\,2n}}.$$

What is important for us is only that c is a constant and $\nu_n \to 0$. As is easy to see, if $\sigma(x)$ is subject to (1), then $\mu(t)$ satisfies (11). That provides us with the limiting relations for the coefficients of $\phi_{2n}(z)$. With that much, it is obvious that $\nu_n \to 0$.

From (18), we also have

$$c(1+\nu_n)\, p_n(x) = 2\, \phi_{n\,2n} + 2 \sum_{k=1}^{n} (\phi_{n+k\,2n} + \phi_{n-k\,2n})\, \cos kt. \tag{19}$$

7. Write the three-term recurrences for $p_n(x)$ in the form

$$p_{n+1}(x) = \frac{1}{\beta_n}\, x\, p_n(x) - \frac{\alpha_n}{\beta_n}\, p_n(x) - \frac{\beta_{n-1}}{\beta_n}\, p_{n+1}(x). \tag{20}$$

On the strength of (19), the ratio of the senior coefficients for p_{n+1} and p_n is equal to

$$\frac{1}{\beta_n} = 2\, \frac{1+\nu_n}{1+\nu_{n+1}}\, \frac{\phi_{2(n+1)\,2(n+1)} + \phi_{0\,2(n+1)}}{\phi_{2n\,2n} + \phi_{0\,2n}}.$$

The first coefficient (equal to 2) comes from the expansion of $\cos nt$ in the powers of $\cos t = x$. From (12) we arrive immediately at $\beta_n^{-1} \to 2$.

Further, on the comparison of the coefficients at x^n in both sides of (20),

$$2\,\frac{\phi_{2(n+1)-1\;2(n+1)}+\phi_{1\;2(n+1)}}{1+\nu_{n+1}}=\frac{1}{\beta_n}\,\frac{\phi_{2n-1\;2n}+\phi_{1\;2n}}{1+\nu_n}-\frac{\alpha_n}{\beta_n}\,\frac{\phi_{2n\;2n}+\phi_{0\;2n}}{1+\nu_n}.$$

Here, we use that the expansion of $\cos nt$ in the powers of $\cos t$ contains $\cos^k t$ only for $k=n,\,n-2,\,\ldots$, that is, for all k of the same parity as n. From the above, it is clear now that $\alpha_n \to 0$.

I would like to thank Dario Bini, Stefano Serra Capizzano, Paolo Tilli, and Nickolai Zamarashkin for their discussion and encouragement.

References

[1] Ya. L. Geronimus, *Polynomials Orthogonal on a Circle and Interval*, International Series of Monographs on Pure and Applied Mathematics, Vol. 18. Pergamon Press, Oxford, 1960.

[2] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkely and Los Angeles, 1958.

[3] G. Heing and K. Rost, *Algebraic Methods for Toeplitz-Like Matrices and Operators*, Akademie-Verlag, Berlin, 1984.

[4] N. Levinson, The Wiener RMS (Root-Mean-Square) error criterion in filter design and prediction, *J. Math. and Phys.* 25: 261–278 (1947).

[5] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, Volume 23. Providence, Rhod Island, 1939.

[6] E. E. Tyrtyshnikov. A unifying approach to some old and new theorems on distribution and clustering, *Linear Algebra Appl.* (1996).

[7] E. E. Tyrtyshnikov. *A Brief Introduction to Numerical Analysis*, Birkhauser, Boston, 1997.