C. R. Math. Rep. Acad. Sci. Canada Vol. **29** (2) 2007, pp. 52–60

THE DYADIC DISTRIBUTION AND ITS ORTHOGONAL POLYNOMIALS

PETER C. GIBSON

Presented by George Elliott, FRSC

ABSTRACT. An open inverse problem that generalizes the classical moment problem is to construct all probability distributions on the real line whose sequence of orthogonal polynomials includes a prescribed subsequence. We have recently solved this problem for a class of subsequences that arise naturally in the context of iterative quadrature schemes, thereby making it possible to construct previously unknown distributions whose orthogonal polynomials have exotic properties. The results are illustrated here by an example: we explicitly construct a distribution on the interval $[-1, 1]$, such that for every $k \geq 1$, its degree $2^k - 1$ orthogonal polynomial divides that of degree $2^{k+1} - 1$, and the zeros of these are equally spaced. Equal spacing of the zeros contrasts starkly with the generic asymptotic behaviour predicted by Szegö's classical theorem.

RÉSUMÉ. Un problème inverse qui reste ouvert et qui généralise le problème classique des moments est de construire toutes les lois de probabilité sur la droite réelle dont la suite des polynômes orthogonaux associée comprend une sous-suite prescrite. On a récemment résolu le problème pour une classe de sous-suites qui provient naturellement des schémas de quadrature iteratifs, ce qui rend possible la construction de lois de probabilités nouvelles dont les polynômes orthogonaux ont des propriétes exotiques. Les résultats sont illustrés ici par un exemple: on construit explicitement une loi sur l'intervalle $[-1, 1]$, tel que pour tout $k \geq 1$, son polynôme orthogonal de degré $2^k - 1$ divise celui de degré $2^{k+1} - 1$, et les zéros de ceux-ci sont également distribués. La distribution égale des zéros se differencie de la distribution asymptotique générique prédite par le théorème classique de Szegö.

1. Introduction. We present a new example of a probability distribution on the line, which we call the dyadic distribution, that serves to illustrate the exotic structure possible for a sequence of orthogonal polynomials when no restrictions are placed on its generating distribution. This fits into the framework of, and complements, ongoing research into the asymptotic distribution of zeros of orthogonal polynomials, such as that of Denisov and Simon [3] and Peherstorfer [6], among others.

The theoretical existence of the dyadic distribution, without explicit formulas, was recently established in [5], which we now discuss briefly, fixing notation as

Received by the editors on March 13, 2007.
AMS Subject Classification: 42C05, 44A60.
Keywords: orthogonal polynomials, location of zeros, moment problems.
© Royal Society of Canada 2007.

follows. Given a probability distribution $d\sigma$ on the real line with respect to which polynomials are integrable, let $p_0^\sigma, p_1^\sigma, p_2^\sigma, \ldots$ denote its associated sequence of monic, orthogonal polynomials. The question of whether one can construct a distribution to realize a prescribed sequence as its orthogonal polynomials is in fact a re-casting of the classical moment problem [1], since the sequence $\{p_n^\sigma\}$ carries precisely the same information about $d\sigma$ as does the moment sequence $\{\mu_n^\sigma\}$, defined by $\mu_n^\sigma = \int x^n \, d\sigma(x)$. One of two main results in [5] can be viewed as solving a variant of the classical moment problem, in which, rather than prescribing the full sequence of orthogonal polynomials, only a thin subsequence is prescribed, with the additional restriction that earlier terms in the subsequence divide all later terms, or, in other words, have all their zeros in common with later terms. A basic constraint imposed by standard interlacing results is that if $m < n$, then p_m^σ and p_n^σ share at most $n - m - 1$ zeros. Thus, for example, if p_m^σ divides p_n^σ then $n \geq 2m + 1$. The "fast iterative" distributions constructed in [5] have the property that they contain a subsequence $\{p_{\nu_k}^\sigma\}_{k=1}^\infty$ such that $p_{\nu_k}^\sigma$ divides $p_{\nu_{k+1}}^\sigma$ for every $k \geq 1$, and where the indices ν_k grow as *slowly as possible* in the sense that $\nu_{k+1} = 2\nu_k + 1$. (For a generic distribution $d\sigma$, two polynomials p_m^σ and p_n^σ can be expected to have no common zeros at all, except possibly 0 in the case of symmetric distributions. See [4].)

Beyond interlacing, there also exist metric constraints governing the zeros of orthogonal polynomials. For example, consider the sequence of polynomials whose zeros are evenly spaced in the open interval $(-1, 1)$:

$$p_n(x) = \prod_{k=1}^{n} \Big(x - \frac{2k - n - 1}{n + 1} \Big). \tag{1}$$

The zeros of these polynomials interlace, but they are not the orthogonal polynomials of any distribution. (In fact, there are distributions $d\sigma$ such that $p_n^\sigma = p_n$ for $n = 1, 2, 3$, but the equality inevitably breaks down at $n = 4$.) The well-known asymptotic result of Szegö [7, Chapter 12.7] asserts that the orthogonal polynomials of a wide class of distributions on $[-1, 1]$ have zeros that tend to be distributed like the those of the Chebyshev polynomials, eventually clustering at the endpoints.

It is with reference to the above discussion that the dyadic distribution is an extreme contrast to the generic situation: an infinite subsequence of its orthogonal polynomials is such that the earlier terms in the subsequence divide all later terms, and each term in the subsequence has equally spaced zeros.

We cite [7] and [2] for the basic theory of orthogonal polynomials, and in particular their relations to discrete distributions and Jacobi matrices. Before describing the dyadic distribution, we recall some particular facts from the standard theory that will be needed in later arguments.

1.1. Preliminary facts. One basic technical fact that we need is the following. Given a distribution $d\sigma$ on $\mathbb{R}$ with respect to which polynomials are integrable and whose support contains at least $m + 1$ points, there is a unique distribution of the form

$$d\sigma_m(x) = \sum_{i=1}^{m} c_i \delta(x - \xi_i) \tag{2}$$

that has the property that for every polynomial p of degree at most $2m-1$,

$$\int_{-\infty}^{\infty} p \, d\sigma_m = \int_{-\infty}^{\infty} p \, d\sigma. \tag{3}$$

This is of course the m-point Gauss rule for $d\sigma$. We reserve the notation $d\sigma_m$ for this particular distribution determined by $d\sigma$. $d\sigma_m$ is supported precisely at the zeros of p_m^σ, the degree m orthogonal polynomial generated by $d\sigma$, and is uniquely determined by the pair of polynomials $p_{m-1}^\sigma, p_m^\sigma$. The Christoffel numbers c_i in (2) are given in terms of $p_{m-1}^\sigma, p_m^\sigma$ by the formula

$$c_i = \frac{C_1}{p_{m-1}^\sigma(\xi_i)(p_m^\sigma)'(\xi_i)}, \tag{4}$$

where $C_1 > 0$ is the appropriate normalizing constant. $d\sigma_m$ is also uniquely determined by the pair of polynomials $p_m^\sigma, p_{m+1}^\sigma$. In terms of these the Christoffel numbers c_i in (2) are given by the formula

$$c_i = \frac{C_2}{(p_m^\sigma)'(\xi_i)p_{m+1}^\sigma(\xi_i)}, \tag{5}$$

where $C_2 < 0$ is again the appropriate normalization constant. (See [7, p. 48].)

These facts can be turned around as follows. Let $S = \{s_1, \dots, s_m\}$ and $T = \{t_1, \dots, t_{m+1}\}$ be any two sets of numbers that satisfy

$$t_1 < s_1 < t_2 < s_2 < \cdots < t_m < s_m < t_{m+1},$$

and let p and q be the monic polynomials having (simple) zeros S and T respectively. Then there exists a unique distribution $d\gamma$ supported on T with the property that $\operatorname{supp} d\gamma_m = S$. The uniquely determined distribution $d\gamma$ and its m-point Gauss rule $d\gamma_m$ are given by the respective formulas

$$d\gamma(x) = C \sum_{i=1}^{m+1} \frac{\delta(x - t_i)}{p(t_i)q'(t_i)}, \text{ where } C = \Big(\sum_{i=1}^{m+1} \frac{1}{p(t_i)q'(t_i)} \Big)^{-1}, \tag{6}$$

$$d\gamma_m(x) = C' \sum_{i=1}^{m} \frac{\delta(x - s_i)}{p'(s_i)q(s_i)}, \text{ where } C' = \Big(\sum_{i=1}^{m} \frac{1}{p'(s_i)q(s_i)} \Big)^{-1}. \tag{7}$$

These formulas will be used in the proof of our main theorem.

2. The dyadic distribution.

2.1. Construction. Let $\mathbb{D}$ denote the set of dyadic rational numbers in the open interval $(-1,1)$. A dyadic rational of the form $\lambda = r/2^j$, where r is odd and $j \geq 1$, is said to have order j; the number 0 is deemed to be the unique dyadic rational of order 0. $\mathbb{D}_j$ denotes the set of dyadic rationals in $(-1,1)$ of order at most j. The numbers

$$\psi(\lambda, n) = \frac{\binom{2^{n+1}-2}{2^n(1+\lambda)-1}}{\binom{2^{n+1}-2}{2^n-1}}, \text{ where } n \geq 1 \text{ and } \lambda \in \mathbb{D}_n, \tag{8}$$

play a central role in our considerations. Note that $\psi(-\lambda, n) = \psi(\lambda, n)$, that $\psi(\lambda, n)$ can be expressed as a ratio of rising factorials,

$$\psi(\lambda, n) = \frac{\big(2^n(1-|\lambda|)\big)\big(2^n(1-|\lambda|)+1\big)\cdots(2^n-1)}{(2^n)(2^n+1)\cdots\big(2^n(1+|\lambda|)-1\big)}, \tag{9}$$

and that for fixed λ of order $j \geq 1$ the sequence $\{\psi(\lambda, n)\}_{n=j}^{\infty}$ is strictly decreasing.

We define a non-negative function w on $\mathbb{R}$ in terms of the $\psi(\lambda, n)$ as follows. If $\lambda \in \mathbb{D}$ has order j, where $j \geq 1$, then set

$$w(\lambda) = 2^{-j}\psi(\lambda, j) - \sum_{n=j+1}^{\infty} 2^{-n}\psi(\lambda, n), \tag{10}$$

and otherwise set $w(\lambda) = 0$. If $\lambda \in \mathbb{D}$ has positive order then the expression (10) defining $w(\lambda)$ is strictly positive. Furthermore, we show in Section 2.2 (Corollary 3) that

$$\sum_{\lambda \in \mathbb{D}} w(\lambda) = 1.$$

This paves the way for our main definition.

Definition 1. Let $d\alpha$ denote the probability distribution defined by

$$d\alpha(x) = \sum_{\lambda \in \mathbb{D}} w(\lambda)\delta(x - \lambda),$$

where δ denotes the Dirac delta function. We shall refer to $d\alpha$ as the *dyadic distribution* on $[-1,1]$.

The accumulation function of $d\alpha$ is

$$\alpha(x) = \sum_{\lambda \in \mathbb{D}} w(\lambda)H(x - \lambda), \tag{11}$$

where H denotes the Heaviside function:

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0. \end{cases}$$

Our main result is the proof in Section 2.4 that an infinite subsequence of the orthogonal polynomials generated by $d\alpha$ have zeros that are equally spaced in $(-1, 1)$. As a first step we establish some basic properties of $d\alpha$.

2.2. Basic properties. It will be useful to work with the following approximations to the function w. For each $k \geq 0$, set $w^k(\lambda) = 0$ if $\lambda \notin \mathbb{D}_k$, and for $\lambda \in \mathbb{D}_k$ set

(12)
$$w^k(\lambda) = \begin{cases} 2^{-k}\psi(\lambda, k) & \text{if } \lambda \text{ has order } k, \\ 2^{-j}\psi(\lambda, j) - \displaystyle\sum_{n=j+1}^{k} 2^{-n}\psi(\lambda, n) & \text{if } \lambda \text{ has order } j \text{ and } 1 \leq j < k, \\ 2^{-k} & \text{if } \lambda = 0. \end{cases}$$

Note that w^k is strictly positive on $\mathbb{D}_k$. Moreover, with some careful bookkeeping one can check that

$$\sum_{\lambda \in \mathbb{D}_k} w^k(\lambda) = 1.$$

Let $d\alpha^k$ denote the associated distribution supported on $\mathbb{D}_k$, defined as

$$d\alpha^k(x) = \sum_{\lambda \in \mathbb{D}_k} w^k(\lambda)\delta(x - \lambda), \tag{13}$$

and let α^k be the accumulation function

$$\alpha^k(x) = \sum_{\lambda \in \mathbb{D}_k} w^k(\lambda)H(x - \lambda). \tag{14}$$

We omit the proof of the following estimate.

Lemma 2. $\sum_{\lambda \in \mathbb{D}} |w(\lambda) - w^k(\lambda)| = O(2^{-k/2})$.

The lemma leads immediately to two useful facts.

Corollary 3. $\sum_{\lambda \in \mathbb{D}} w(\lambda) = \lim_{k\to\infty} \sum_{\lambda \in \mathbb{D}} w^k(\lambda) = 1$.

Corollary 4. *$\alpha^k \to \alpha$ uniformly.*

We are now in a position to move on to the structure of the orthogonal polynomials generated by $d\alpha$.

2.3. The orthogonal polynomials.

THEOREM 1. *For each $k \geq 1$, the degree $2^k - 1$ monic, orthogonal polynomial of the dyadic distribution is*

$$p^{\alpha}_{2^k-1}(x) = \prod_{\lambda \in \mathbb{D}_{k-1}} (x - \lambda).$$

PROOF. For each $k \geq 1$, set

$$s_k = \frac{2^{2^{k+1}-k-3}}{\binom{2^{k+1}-2}{2^k-1}}, \tag{15}$$

$$u^k(\lambda) = \begin{cases} 2^{-2^{k+1}+3}\binom{2^{k+1}-2}{2^k(1+\lambda)-1} & \lambda \in \mathbb{D}_{k-1}, \\ 0 & \text{otherwise}, \end{cases} \tag{16}$$

$$v^k(\lambda) = \begin{cases} 2^{-2^{k+1}+3}\binom{2^{k+1}-2}{2^k(1+\lambda)-1} & \lambda \in \mathbb{D}_k \setminus \mathbb{D}_{k-1}, \\ 0 & \text{otherwise}. \end{cases} \tag{17}$$

It follows directly from these definitions that each of u^k and v^k is non-negative, $\sum_\lambda u^k(\lambda) = \sum_\lambda v^k(\lambda) = 1$, and

$$w^k(\lambda) = w^{k-1}(\lambda) - s_k u^k(\lambda) + s_k v^k(\lambda). \tag{18}$$

Setting

$$d\varphi^k(x) = \sum_\lambda u^k(\lambda)\delta(x-\lambda) \quad \text{and} \quad d\theta^k(x) = \sum_\lambda v^k(\lambda)\delta(x-\lambda), \tag{19}$$

the equation (18) is equivalent to

$$\begin{aligned} d\alpha^k &= d\alpha^{k-1} - s_k d\varphi^k + s_k d\theta^k \\ &= d\alpha^{k-1} + s_k(d\theta^k - d\varphi^k). \end{aligned} \tag{20}$$

The key observation is that, setting $S = \mathbb{D}_{k-1}$ and $T = \mathbb{D}_k \setminus \mathbb{D}_{k-1}$, the distributions $d\gamma$ and $d\gamma_m$ defined according to the formulas (6) and (7) are equal to $d\theta$ and $d\varphi$ respectively. (The fact that points in S and T are equally spaced makes this a straightforward calculation.) Thus $d\varphi = d\theta_m$, where $m = 2^k - 1$ is the size of $\mathbb{D}_{k-1}$. This means that for every polynomial p of degree at most $2m - 1$,

$$\int_{-\infty}^{\infty} p\, d\theta - \int_{-\infty}^{\infty} p\, d\varphi = 0,$$

which implies in turn by equation (20) that

$$\int_{-\infty}^{\infty} p\, d\alpha^k = \int_{-\infty}^{\infty} p\, d\alpha^{k-1}.$$

Iterating the above reasoning yields more generally that for every $n \geq 0$ and every polynomial p of degree at most $2m - 1$ $(= 2^{k+1} - 3)$,

$$\int_{-\infty}^{\infty} p \, d\alpha^{k+n} = \int_{-\infty}^{\infty} p \, d\alpha^{k-1}. \tag{21}$$

By Corollary 4, letting $n \to \infty$ in (21) yields that

$$\int_{-\infty}^{\infty} p \, d\alpha = \int_{-\infty}^{\infty} p \, d\alpha^{k-1},$$

from which it follows that

$$d\alpha_{2^k - 1} = d\alpha^{k-1}. \tag{22}$$

Therefore the zeros of $p^{\alpha}_{2^k-1}$ coincide with the support of $d\alpha^{k-1}$, which is $\mathbb{D}_{k-1}$ by construction. ■

Of course the points of $\mathbb{D}_{k-1}$ are equally spaced:

$$\mathbb{D}_{k-1} = \left\{ \frac{n - 2^{k-1}}{2^{k-1}} \;\middle|\; 1 \leq n \leq 2^k - 1 \right\}.$$

And, since $\mathbb{D}_k \subset \mathbb{D}_{k+n}$ for any n, it follows immediately from the theorem that $p^{\alpha}_{2^k-1}$ divides $p^{\alpha}_{2^{k+n}-1}$.

2.4. Structure of the Jacobi matrix. We recall that any sequence of orthogonal polynomials on $\mathbb{R}$ obeys a 3-term recurrence of the form

$$p^{\sigma}_{n+1}(x) = (x - a_{n+1}) p^{\sigma}_n(x) - b_n^2 p^{\sigma}_{n-1}(x) \quad n \geq 1, \tag{23}$$

with starting values $p^{\sigma}_0 = 1$, $a_1 = \mu^{\sigma}_1$ (the first moment of $d\sigma$) and $p^{\sigma}_1(x) = x - a_1$. The values a_n and b_n are the diagonal and next-to-diagonal entries of the Jacobi matrix of $d\sigma$, defined as

$$J_{\sigma} = \begin{pmatrix} a_1 & b_1 & 0 & 0 & \cdots \\ b_1 & a_2 & b_2 & 0 & \cdots \\ 0 & b_2 & a_3 & b_3 & \cdots \\ 0 & 0 & b_3 & a_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{24}$$

where, according to the standard theory, $a_n \in \mathbb{R}$ and $b_n > 0$, for each $n \geq 1$.

Since the dyadic distribution is symmetric about 0, its Jacobi matrix J_{α} has zero diagonal. Thus the dyadic Jacobi matrix is represented by the sequence of positive terms $b_1, b_2, \ldots$. A plot of several hundred terms of the sequence reveals some intriguing structure. This is most clearly brought out if, instead of b_n, one plots $4(b_n)^2 - 1$ against $\log_2 n$, as in Figure 1.

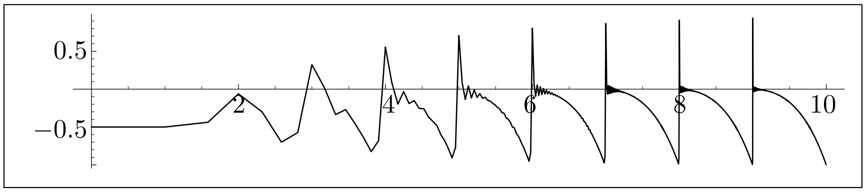

Figure 1: A plot of $4(b_n)^2 - 1$ versus $\log_2 n$ $(1 \leq n \leq 1022)$.

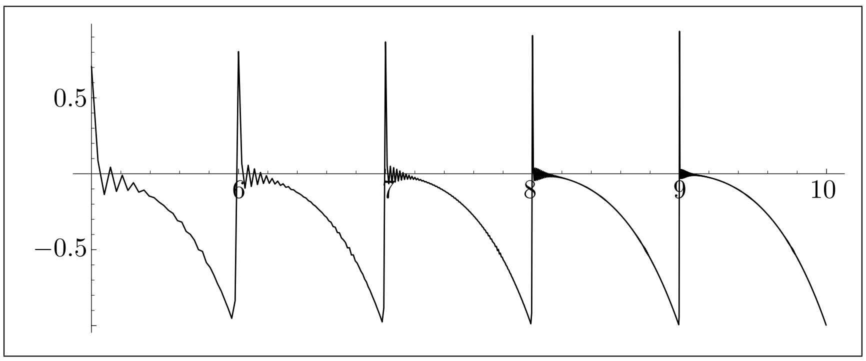

Figure 2: A closer view.

It looks like an electrocardiogram! More precisely, it seems clear that the sequence $\{b_n\}$ is asymptotically $\log_2$ periodic. The zoomed image in Figure 2 of the latter 5 cycles shows more clearly the progressive stabilization of the picture.

Away from the positive peaks at powers of 2 (which in Figures 1 and 2 occur at integer values of $\log_2 n$), it seems reasonable to conjecture that b_n has the asymptotic structure

$$b_n \sim \frac{1}{2}\sqrt{1 - (\log_2 n - [\log_2 n])^{\beta}}, \tag{25}$$

where $[\log_2 n]$ denotes the integer part of $\log_2 n$, and where the constant β lies between 2 and 3.

3. Conclusion. In summary, the dyadic distribution we have presented captures in concrete form some of the strange possibilities realizable within the context of orthogonal polynomials generated by completely general probability distributions on $\mathbb{R}$. The apparent asymptotic $\log_2$ periodicity of the dyadic distribution's Jacobi matrix constitutes an intriguing fact that calls for deeper explanation.

References

1. N. I. Akhiezer, *The Classical Moment Problem.* Hafner, New York, 1965.
2. C. de Boor and G. H. Golub, *The numerically stable reconstruction of a Jacobi matrix from spectral data.* Linear Algebra Appl. **21** (1978), 245–260.
3. S. A. Denisov and B. Simon, *Zeros of orthogonal polynomials on the real line.* J. Approx. Theory **121** (2003), 357–364.
4. P. C. Gibson, *Common zeros of two polynomials in an orthogonal sequence.* J. Approx. Theory **105** (2000), 129–132.
5. ———, *Fast iterative Gaussian quadrature.* Preprint, March 2007.
6. F. Peherstorfer, *On the zeros of orthogonal polynomials: The elliptic case.* Constr. Approx. **20** (2004), 377–397.
7. G. Szegö, *Orthogonal Polynomials.* Amer. Math. Soc., New York, 1959.

Department of Mathematics & Statistics
York University
Toronto, ON
M3J 1P3
email: pcgibson@yorku.ca