Journal of Computational and Applied Mathematics 233 (2010) 1355–1365

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Asymptotics of orthogonal polynomial's entropy

A.I. Aptekarev [a,*], J.S. Dehesa [b], A. Martinez-Finkelshtein [c]

[a] Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Miusskaya Sq. 4, 125047 Moscow, Russia
[b] Institute Carlos I of Theoretical and Computational Physics, Granada University, 18071 Granada, Spain
[c] Department of Statistics and Applied Mathematics, Almería University, 04120 Almería, Spain

ARTICLE INFO

Article history:
Received 26 March 2008

Dedicated to Professor Jesús S. Dehesa on the occasion of his 60th birthday.

Keywords:
Information entropy
Uncertainty in quantum mechanics
Orthogonal polynomials
Asymptotics
Szegő condition
Mutual energy
Logarithmic potential

ABSTRACT

This is a brief account on some results and methods of the asymptotic theory dealing with the entropy of orthogonal polynomials for large degree. This study is motivated primarily by quantum mechanics, where the wave functions and the densities of the states of solvable quantum-mechanical systems are expressed by means of orthogonal polynomials. Moreover, the uncertainty principle, lying in the ground of quantum mechanics, is best formulated by means of position and momentum entropies. In this sense, the behavior for large values of the degree is intimately connected with the information characteristics of high energy states. But the entropy functionals and their behavior have an independent interest for the theory of orthogonal polynomials. We describe some results obtained in the last 15 years, as well as sketch the ideas behind their proofs.

1. Introduction

Given a probability density ρ, the expression $-\ln(\rho)$ is known as the *surprise level function*, whose mean value (average lack of information or uncertainty) is identified with the *Shannon entropy* of ρ [1]. In particular, if $\{\rho_j\}_{j=1}^n$ is a discrete probability distribution ($\rho_j \geq 0, \sum \rho_j = 1$), the information entropy functional

$$s_n := -\sum_{j=1}^{n} \rho_j \ln(\rho_j) \tag{1}$$

measures, in common perception, the uncertainty associated with this probability distribution. Two extremal cases are

$$\text{Uniform distribution: } \{\rho_j\}_{j=1}^n = \left\{\frac{1}{n}, \dots, \frac{1}{n}\right\} \quad \Rightarrow \quad s_n = \ln(n),$$
$$\text{Dirac delta: } \{\rho_j\}_{j=1}^n = \{0, \dots, 0, 1, 0, \dots, 0\} \quad \Rightarrow \quad s_n = 0.$$

Jensen's inequality applied to (1) gives

$$0 \leq s_n \leq \ln(n),$$

showing that the uniform distribution, having the most uncertain outcome, has maximal entropy, while the deterministic event has the minimal one.

* Corresponding author.
E-mail addresses: aptekaa@keldysh.ru (A.I. Aptekarev), dehesa@ugr.es (J.S. Dehesa), andrei@ual.es (A. Martinez-Finkelshtein).

0377-0427/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2009.02.056

If ρ is a continuous probability distribution,

$$\rho(x) \geqslant 0, \quad x \in \mathbb{R}, \qquad \int_{\mathbb{R}} \rho(x)\mathrm{d}x = 1,$$

we can define by analogy the information entropy [1]

$$S_\rho = -\int \rho(x) \ln \rho(x)\mathrm{d}x, \tag{2}$$

known also as the *Boltzmann*, *Boltzmann–Shannon* or *differential* entropy, that characterizes the localization of the density of the distribution. The measures of information (1) and (2), although formally similar, have different properties. In particular, if ρ is a continuous probability distribution on $[a, b]$ and we introduce a discrete distribution

$$\rho_j = \int_{a+(j-1)h}^{a+jh} \rho(x)\mathrm{d}x, \quad j = 1, \dots, n, h = (b-a)/n,$$

then $S_\rho \approx s_\rho + \ln(h)$, showing in particular that S_ρ is unbounded. In this sense, it is more convenient to consider an entropy functional for two (probability) measures, μ and ν,

$$\mathcal{K}(\mu, \nu) = \begin{cases} -\displaystyle\int \ln\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)\mathrm{d}\nu, & \text{if } \nu \text{ is } \mu\text{-a.c.,} \\ -\infty, & \text{otherwise.} \end{cases} \tag{3}$$

This is the *Kullback–Leibler information*, also known as the *relative* or *mutual* entropy, which measures the "distance" between ν and μ. Obviously, if ν is μ-a.c., we can also rewrite it as

$$\mathcal{K}(\mu, \nu) = -\int \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \ln\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)\mathrm{d}\mu.$$

There is no a priori preferred notion of information measure in physical applications, but a relevant role played by the Boltzmann–Shannon entropy in quantum mechanics (and in particular, in the modern density functional theory [2]) is motivated in part by the entropic formulation of the uncertainty principle. Consider for instance a single particle system in D dimensions. For any quantum mechanical state the distribution density is $\rho(x) := |\Psi(x)|^2$, where $\Psi(x)$ is the corresponding wave function or physical solution of the associated Schrodinger equation. If γ is the distribution density in the momentum space, then the Heisenberg's uncertainty principle for the quantum mechanical system is a consequence of the following inequality (see [3,4]):

$$S_\rho + S_\gamma \geqslant D(1 + \ln \pi).$$

For the fundamental quantum mechanical systems (harmonic oscillator, hydrogen atom) the relevant component of the probability density of physical states are expressed by means of orthogonal polynomials (Gegenbauer, Laguerre, Hermite). This brought up the study of the entropy functionals for orthogonal polynomials (see [5] as well as the survey [6]).

Let μ be a positive unit Borel measure on $\mathbb{R}$ and let

$$p_n(x) = \kappa_n \prod_{j=1}^{n} \left(x - \zeta_j^{(n)}\right), \quad \kappa_n > 0, \quad n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}, \tag{4}$$

denote the corresponding sequence of *orthonormal* polynomials such that

$$\int p_n(x)p_m(x)\mathrm{d}\mu(x) = \delta_{mn}, \quad m, n \in \mathbb{N}_0.$$

Then we can define the sequence of probability measures ν_n, absolutely continuous with respect to μ, given by

$$\mathrm{d}\nu_n(x) = p_n^2(x)\mathrm{d}\mu(x), \quad n \in \mathbb{N}_0, \tag{5}$$

(note that $\nu_0 = \mu$). These measures are usually related with the quantum-mechanical probability distribution of physical states, and are standard objects of study in the analytic theory of orthogonal polynomials. As it was shown in [7], ν_n is associated with the behavior of the ratio p_{n+1}/p_n as $n \to \infty$.

The relative entropy

$$E_n := \mathcal{K}(\mu, \nu_n) = -\int p_n^2(x) \ln\left(p_n^2(x)\right)\mathrm{d}\mu(x) \tag{6}$$

is called the (continuous) information *entropy of orthogonal polynomials* $\{p_n\}$. Obviously, this is not the only way to define an entropy associated with orthogonal polynomials (4). For instance, for $j \in \{1, \dots, n\}$, let

$$\psi_i = \ell_n(\zeta_j^{(n)})\, p_{i-1}^2(\zeta_j^{(n)}), \quad i = 1, \dots, n,$$

where $\ell_n^{-1}(x) := \sum_{k=0}^{n-1} p_k^2(x)$ is the Christoffel function. Then $\{\psi_1^2, \ldots, \psi_n^2\}$ is a discrete probability distribution with the associated Shannon entropy

$$S_{n,j} := -\sum_{i=1}^{n} \psi_i^2 \ln\left(\psi_i^2\right).$$

An explicit formula for this entropy for Chebyshev polynomials of the first and second kind has been obtained recently in [8]. Result of several numerical experiments have been presented therein, suggesting that after an appropriate rescaling and normalization, entropies $S_{n,j}$ have a limit as $n \to \infty$.

Here we focus on the continuous entropy E_n, and review some ideas appeared along the 15 years history of research on orthogonal polynomials entropy. More precisely, we highlight some results and methods regarding the asymptotics of this functional for large n:

$$E_n \sim ? \quad \text{as } n \to \infty.$$

From the quantum-mechanical point of view this problem is related to the information characteristics of the highly excited states (i.e. Rydberg states). But the notion of entropy is intrinsically relevant for the theory of orthogonal polynomials. The Szegő constant is an entropic quantity, and properties of the entropy allow to prove the well known Szegő asymptotics, as shown in [9, Ch. 2]. As a consequence of the results exposed below, we may conclude that entropy and some other information theoretic measures capture in fact some fine features of the sequence $\{p_n\}$.

In the next Section 2 we discuss the basic asymptotic formula for entropy of polynomial sequences orthogonal on the interval of the real axis and its connection with Bernstein–Szegő asymptotics of orthogonal polynomials. Section 3 is devoted to the L^p-norms method of proof of the asymptotic results. Then in Section 4 we consider a logarithmic potential theory approach to the asymptotic of orthogonal polynomials entropy. Here the mutual energy of the zero counting measure and ν_n play an important role. Section 5 is devoted to the modern state of the asymptotics of orthogonal polynomials entropy in the Szegő class.

2. Bernstein–Szegő asymptotic formula and entropy of polynomials orthogonal on the interval

Let us describe first a "straightforward" approach to the asymptotics of E_n, namely replacing p_n by their asymptotic expression directly in (6). We assume in this section that μ is a measure on $[-1, 1]$, absolutely continuous with respect to the Lebesgue measure, and $d\mu(x) = w(x)dx$. A relevant character on the segment is the Chebyshev weight

$$\rho(x) = \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}},$$

along with the *equilibrium measure* η, absolutely continuous with respect to the Lebesgue measure on $[-1, 1]$, given by $d\eta(x) = \rho(x)dx$. Also, $w_0(x) = w(x)/\rho(x)$ is called the *trigonometric weight* corresponding to w, and

$$\gamma(x) := \frac{1}{2\pi} \int_{-1}^{1} \frac{\ln w_0(t) - \ln w_0(x)}{t - x} \sqrt{\frac{1-x^2}{1-t^2}} dt$$

is the *harmonic conjugate* of the function $\ln w_0(x)$. We may also define

$$R_n(x) := \sqrt{\frac{2}{w_0(x)}} \cos(n \arccos x + \gamma(x)). \tag{7}$$

S. N. Bernstein and G. Szegő showed that under certain assumptions, p_n and R_n are close on $[-1, 1]$ (see [10,11]). Namely, Bernstein's condition is

$$(B): \quad \begin{array}{l} 0 < \lambda < w_0(x) < L, \quad x \in [-1, 1] \\ |w_0(x+\delta) - w_0(x)| |\ln \delta|^{1+\varepsilon} < K, \quad (\varepsilon > 0, K > 0) \end{array} \tag{8}$$

while the (weaker) condition assumed by Szegő is

$$(S): \quad \int_{-1}^{1} \ln w(x) d\eta(x) > -\infty. \tag{9}$$

The following theorem holds:

Theorem 1 (*Bernstein–Szegő*). *If $w \in (S)$, then $\|p_n - R_n\|_{L^2_w[-1,1]} = o(1)$ as $n \to \infty$. Moreover, if $w \in (B)$, then $\|p_n - R_n\|_{C[-1,1]} = o(1)$ for $n \to \infty$.*

For the situation considered here the correct asymptotic formula for the entropy functional (6) was derived in the paper [12]. Formally substituting (7) in (6) we have

$$E_n \simeq -\int_{-1}^{1} \ln\left(\frac{2\cos^2(n\theta+\gamma(x))\rho(x)}{w(x)}\right) 2\cos^2(n\theta+\gamma(x))\mathrm{d}\eta(x),$$

where $x=\cos\theta$. Neglecting (for the moment) the phase shift $\gamma(x)$ we get

$$E_n \simeq -\int_{-1}^{1} \ln(2\cos^2(n\theta))\, 2\cos^2(n\theta)\mathrm{d}\eta(x) - \int_{-1}^{1} \ln\left(\frac{\rho(x)}{w(x)}\right) 2\cos^2(n\theta)\mathrm{d}\eta(x). \tag{10}$$

The first integral on the r.h.s. does not depend on the weight of orthogonality w; actually, it is the entropy of Chebyshev polynomials of the first kind. This functional has been explicitly evaluated in [13]:

$$E^{(0)} := -\int_{-1}^{1} \ln(2\cos^2(n\theta))\, 2\cos^2(n\theta)\mathrm{d}\eta(x) = \ln(2) - 1; \tag{11}$$

observe that it does not depend on n; it is a long standing conjecture of L. Golinskii that this feature is actually another characterization of Chebyshev polynomials of the first kind.

The second integral on the r.h.s. of (10) has a limit for $n\to\infty$:

$$\int_{-1}^{1} \ln\left(\frac{\rho_0(x)}{w(x)}\right) 2\cos^2(n\theta)\mathrm{d}\eta(x) \quad\longrightarrow\quad -\int_{-1}^{1} \ln(w_0(x))\,\mathrm{d}\eta(x) = -\mathcal{K}(\mu,\eta).$$

The existence of the limit and its value is a more or less evident fact, and it rigorously follows from a useful lemma proved in [14]:

Lemma 2. *Let $g\in C(\mathbb{R})$, $g(\theta+\pi)=g(\theta)$, $f\in L^1([0,\pi])$ and $\gamma(\theta)$ be a measurable and almost everywhere finite on $[0,\pi]$ function. Then for $n\to\infty$,*

$$\int_0^{\pi} g(n\theta+\gamma(\theta))f(\theta)\mathrm{d}\theta \longrightarrow \frac{1}{\pi}\int_0^{\pi} g(\theta)\mathrm{d}\theta \int_0^{\pi} f(\theta)\mathrm{d}\theta.$$

Remark 3. For $f(\theta)\equiv 1$, $g(\theta)=\cos^p(\theta)$ and $\gamma\in C([0,\pi])$ this result was proved by Bernstein (see [10, p. 12]). For $\gamma(\theta)\equiv 0$ and $g\in L^\infty[0,\pi]$ the statement is the well known Fejer's Lemma which establishes convergence to zero of the Fourier coefficients.

We note that this lemma justifies our omission of the phase shift γ above. Thus, taking into account that under Bernstein condition (8) we can rigorously substitute asymptotics (7) in the entropy functional (6), we obtain our first result:

Theorem 4. *Let $\{p_n\}_{n=0}^{\infty}$ be a system of orthonormal polynomials (4) with respect to an absolutely continuous measure $\mathrm{d}\mu(x)=w(x)\mathrm{d}x$ on $[-1,1]$ satisfying the Bernstein condition (8). Then the entropy of orthogonal polynomials (6) converges when $n\to\infty$:*

$$\lim_n E_n = E_\infty := E^{(0)} - \mathcal{K}(\mu,\eta), \tag{12}$$

where $\mathcal{K}$ is the mutual entropy (3), η is the equilibrium measure of $[-1,1]$, and constant $E^{(0)}$ is defined in (11).

We see that the limit of the entropy of orthogonal polynomial sequences with respect to a general weight satisfying the Bernstein condition (8) is equal to the entropy of Chebyshev polynomials of the first kind reduced by the mutual (relative) entropy between the measure of orthogonality μ and the equilibrium (Chebyshev) measure η of the interval. Since the mutual entropy (3) is always non-negative (Jensen's inequality), Chebyshev polynomials exhibit the asymptotically maximal entropy, in agreement with our intuition that these polynomials are the "most uniform" ones. Another observation is that right-hand side in (12) exists under the Szegő condition (9), which is milder than what we assumed in Theorem 4, namely the Bernstein condition (8). It is still a challenging open problem to prove (or present a counterexample) that formula (12) for the entropy of orthogonal polynomials on the interval is valid in the whole Szegő class.

3. L^p norms method

If we want to prove (12) under milder conditions on the weight, we cannot simply substitute the asymptotics of p_n into the entropy functional. The logarithmic singularity present therein creates additional complications that we can overcome using the L^p norms method. The idea is very simple. Consider the functional

$$R_\rho^p := \frac{1}{1-p}\ln\int |\rho(x)|^p\mathrm{d}x,$$

called *Renyi entropy*. The identity

$$\lim_{p\to 1} R_\rho^p = S_\rho$$

reduces the problem to the estimates of the L^p norms. In consequence, if we consider the L^p_μ norm of our orthonormal polynomials p_n,

$$N_n(p) := \int |p_n(x)|^p \mathrm{d}\mu(x), \tag{13}$$

then

$$E_n = -\lim_{p\to 1} \frac{1}{p-1} \ln \int |p_n(x)|^{2p} \mathrm{d}\mu(x) = -N_n'(2p)\big|_{p=1}.$$

Suppose that the sequence $\{N_n(2p)\}$ is uniformly convergent on the interval $p \in [0, t]$; we denote its limit by $N(2p)$. Functions $N_n(2p)$ are holomorphic in the half-plane $\mathrm{Re}(p) > 0$, and uniformly bounded in $0 < \mathrm{Re}(p) < t$, being so on the interval $(0, t)$. Hence, it is a normal family, whose limit points coincide with function $N(2p)$, holomorphic in $0 < \mathrm{Re}(p) < t$. These considerations justify taking limits in the derivative:

$$\lim_n E_n = -\lim_n N_n'(2p)\big|_{p=1} = -N'(2p)|_{p=1} =: E_\infty. \tag{14}$$

In [14] a detailed study of the L^p-norms asymptotics for the most important classes of orthogonal polynomials with respect to an absolutely continuous measure has been carried out. Here we review the main results of [14]. For an absolutely continuous measure $\mathrm{d}\mu(x) = w(x)\mathrm{d}x$ on the interval $[-1, 1]$ define

$$N(p) := \frac{2^{p/2}}{\pi^2} \frac{\Gamma(1/2)\Gamma(p/2+1/2)}{\Gamma(p/2+1)} \int_0^\pi w_0^{1-p/2}(\cos\theta)\mathrm{d}\theta, \tag{15}$$

where $\Gamma(x)$ is the Euler gamma function. Then

Theorem 5. (1) *Let $w \in (B)$, then uniformly for $p \in (0, \infty)$,*

$$N_n(p) = N(p) + o(1), \quad n \to \infty. \tag{16}$$

(2) *Let $w \in (S)$, then the asymptotic formula (15)–(16) is valid uniformly for $p \in [0, 2]$.*

The fact that the asymptotic formula (16) for general Szegő weights is restricted to the interval $[0, 2]$ is not accidental. The following result holds:

Theorem 6. *For any $\varepsilon > 0$ there exists an absolutely continuous measure $\mathrm{d}\mu(x) = w(x)\mathrm{d}x$, $w \in (S)$, such that the $L_\mu^{2+\varepsilon}$ norm of the orthonormal polynomials with respect to μ is unbounded.*

This theorem is a corollary of the asymptotic behavior of the norms of Jacobi polynomials, orthogonal with respect to the weight

$$w_{\alpha,\beta}(x) = (1-x)^\alpha(1+x)^\beta, \quad \alpha, \beta > -1.$$

Theorem 7. *Let $w_{\alpha,\beta}$ be a Jacobi weight, and $\alpha \geqslant \beta > -1$. Then*

(1) *If $-1 < \alpha \leqslant -1/2$, then the asymptotic formula (16) holds uniformly for $p \in [0, \infty)$.*
(2) *If $\alpha > -1/2$ and $p_0 = 2 + 2/(2\alpha+1)$, then*
  (a) *the asymptotic formula (16) is valid for $p \in [0, p_0)$;*
  (b) *$N_n(p_0) = c_{\alpha,\beta} \ln n(1 + o(1))$, $n \to \infty$, where*

$$c_{\alpha,\beta} = \pi^{-(p/2+1)} \frac{\Gamma(p/2+1/2)\Gamma(1/2)}{\Gamma(p/2+1)} \begin{cases} 2^{\frac{\alpha-\beta+1}{2\alpha+1}}, & \alpha > \beta, \\ 2^{1+\frac{1}{2\alpha+1}}, & \alpha = \beta. \end{cases}$$

  (c) *if $p > p_0$ then for every $n \geqslant n_0$ there exist constants $c_1 = c_1(\alpha, \beta, p)$ and $c_2 = c_2(\alpha, \beta, p)$ satisfying*

$$c_1 \leqslant \frac{N_n(p)}{n^{(2\alpha+1)(p/2-1)-1}} \leqslant c_2.$$

Remark 8. (1) Observe the different rate of growth of the L^p norms of Jacobi polynomials for different values of p.
(2) In the case of Jacobi weight $w_{\alpha,\beta}$ it is possible to compute $N(p)$ in (15) explicitly (see [14]):

$$N(p) = \frac{2^{(\alpha+\beta)(1-p/2)+1}}{\pi^{p/2+1}} \frac{\Gamma(p/2+1/2)\Gamma(1/2)}{\Gamma(p/2+1)} \frac{\Gamma(\rho(2-p)+1/2)\Gamma(\sigma(2-p)+1/2)}{\Gamma((\rho+\sigma)(2-p)+1)},$$

where $\rho = \frac{\alpha}{2} + \frac{1}{4}$, $\sigma = \frac{\beta}{2} + \frac{1}{4}$.

As a consequence of these results, using (14), we get the asymptotic formulas for the entropy of orthogonal polynomials. Taking derivatives in (15) we obtain

$$-E_\infty = \ln 2 + \psi(3/2) - \psi(2) - \frac{1}{\pi}\int_0^\pi \ln w_0(\theta)\mathrm{d}\theta,$$

where $\psi(z)$ is the digamma function. Making a change of variables in the integral in the r.h.s. and using the identity $\psi(3/2) - \psi(2) = 1 - \ln 2$ we obtain again the result (12) stated in Theorem 4, but now as a consequence of Theorem 5. Moreover, from Theorem 7 we have for the Jacobi weights (which are from the Szegő class but do not satisfy Bernstein's condition) that (12) holds also for Jacobi polynomials with all admissible values of the parameters. In this case, using the remark above, the relative entropy (3) can be evaluated explicitly in terms of the beta function B, showing that for the weight $w_{\alpha,\beta}$,

$$E_\infty = \ln 2\pi - 1 - 2(\alpha+\beta+1)\ln 2 - \ln B(\alpha+1, \beta+1). \tag{17}$$

The method of L^p norms works also in the case of an unbounded support of the orthogonality measure. For an absolutely continuous measure $\mathrm{d}\mu(x) = w(x)\mathrm{d}x$ on $\mathbb{R}$ it is more convenient however to study the asymptotic behavior of the *weighted* L^p norms,

$$\widehat{N}_n(p) := \int_{\mathbb{R}} \left|p_n(x)\sqrt{w_\lambda(x)}\right|^{2p} \mathrm{d}x. \tag{18}$$

In [14] the case of the Freud weights

$$w_\lambda(x) = \exp\left\{-|x|^\lambda\right\}, \quad x \in \mathbb{R},\ \lambda > 1, \tag{19}$$

was studied, yielding the following theorem:

Theorem 9. *Let w_λ be a Freud weight, then as $n \to \infty$,*

$$\widehat{N}_n(p) = \left(\frac{2}{\pi}\right)^p \frac{\Gamma(p+1/2)}{\Gamma(p+1)} \frac{\Gamma(1-p/2)}{\Gamma(3/2-p/2)} x_n^{1-p}(1+o(1)) \tag{20}$$

uniformly on $p \in [0, 4/3)$, where

$$x_n = \left(\frac{2n+1}{2\beta}\right)^{1/\lambda}, \qquad \beta = \frac{\Gamma(\lambda/2+1/2)}{\Gamma(1/2)\Gamma(\lambda/2)}. \tag{21}$$

If we differentiate the weighted norm $\widehat{N}_n(p)$ with respect to p at $p = 1$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}p}\int_{-\infty}^{+\infty} \left|p_n(x)\sqrt{w_\lambda(x)}\right|^{2p} \mathrm{d}x\bigg|_{p=1} = \int_{-\infty}^{+\infty} \ln\left|p_n^2(x)w_\lambda(x)\right| p_n^2(x)w_\lambda(x)\mathrm{d}x.$$

Hence,

$$E_n = -\frac{\mathrm{d}}{\mathrm{d}p}\int_{-\infty}^{+\infty} \left|p_n\sqrt{w_\lambda}\right|^{2p} \mathrm{d}x\bigg|_{p=1} - \int_{-\infty}^{+\infty} p_n^2 w_\lambda \ln w_\lambda \mathrm{d}x = -I_1 + I_2.$$

Integrating the second integral by parts and taking into account the orthonormality of p_n's it is not difficult to show that

$$I_2 = -\frac{2n+1}{\lambda}.$$

Function $G_n(p) = x_n^{p-1}\widehat{N}_n(p)$, with x_n defined in (21), is holomorphic in the band $0 < \mathrm{Re}(p) < 2$, and

$$G_n'(p) = \widehat{N}_n'(p)x_n^{p-1} + \widehat{N}_n(p)x_n^{p-1}\ln x_n.$$

Differentiating equation (20) at $p = 1$ (a procedure that can be formally justified in the same way as for a bounded interval of orthogonality) we obtain

$$\lim_n G_n'(1) = I_2 + \lim_n \ln x_n = G'(1) = \frac{\mathrm{d}}{\mathrm{d}p}\left[\left(\frac{2}{\pi}\right)^p \frac{\Gamma(p+1/2)}{\Gamma(p+1)} \frac{\Gamma(1-p/2)}{\Gamma(3/2-p/2)}\right]\bigg|_{p=1}.$$

Using the well-known identities for the Euler's gamma function and its logarithmic derivative it is easy to establish that $G'(1) = 1 - \ln\pi$. In this way, taking into account (21), we conclude that

$$I_1 = -\frac{1}{\lambda}\ln\frac{n}{\beta} + 1 - \ln\pi + o(1), \quad n \to \infty.$$

Summarizing:

Theorem 10. *Let w_λ be a Freud weight (19). Then for $n \to \infty$ we have*

$$E_n = -\frac{2n+1}{\lambda} + \frac{1}{\lambda} \ln 2n - \frac{1}{\lambda} \ln \frac{\sqrt{\pi}}{2} \frac{\Gamma(\lambda/2)}{\Gamma(\lambda/2+1/2)} - 1 + \ln \pi + o(1).$$

A particular case of this result is the asymptotic formula for the entropy of Hermite polynomials, orthonormal on $\mathbb{R}$ with respect to the weight $w(x) = w_2(x) = e^{-x^2}$:

$$E_n = -n + \ln \sqrt{2n} - 3/2 + \ln \pi + o(1), \quad n \to \infty.$$

A similar asymptotic formula for the entropy of Laguerre polynomials, orthonormal on $\mathbb{R}_+$ with respect to the weight $w^{(\alpha)}(x) = x^\alpha e^{-x}$, $\alpha > -1$, was previously obtained in [15]:

$$E_n = -2n + (\alpha + 1) \ln n - \alpha - 2 + \ln 2\pi + o(1), \quad n \to \infty.$$

We conclude this section with some remarks. The study of the L^p norms of orthogonal polynomials is of independent interest in the theory of general orthogonal and extremal polynomials. On one hand, this problem is connected with the classical research of S.N. Bernstein on the asymptotics of the L^p extremal polynomials in [10], that received further development in recent papers [16,17]. On the other hand, it is a generalization of the widely known problem of Steklov on the estimation of the L^∞ norms of polynomials orthonormal with respect to a positive weight (see [18]). Indeed, for $p = 1$ the norms are bounded (they are just equal to 1), but for $p = \infty$ (as it has been shown in [19]) they may grow to infinity. What happens with the boundedness of the L^p-norms of the orthonormal polynomials for the intermediate values of $1 < p < \infty$?

4. Logarithmic potential theory approach

The entropic functional (6) can be restated in terms of the logarithmic potential theory, which gives an additional interesting perspective of the problem. If μ and ν are Borel (generally speaking, real signed) measures on $\mathbb{C}$, we denote by

$$V(z; \mu) = -\int \ln |z - t| d\mu(t)$$

the logarithmic potential of μ, and by

$$I[\nu, \mu] = \int V(z; \nu) d\mu(z) = -\iint \ln |z - t| d\nu(t) d\mu(z)$$

the mutual energy of μ and ν. Observe that $I(\mu) = I[\mu, \mu]$ is the standard logarithmic energy of μ.

It is well-known that the asymptotic behavior of the n-th root of the orthonormal polynomials (4) is closely related with the sequence of discrete probability measures, called the normalized *zero counting* measures for p_n,

$$\lambda_n := \frac{1}{n} \sum_{j=1}^{n} \delta_{\zeta_j^{(n)}},$$

via the formula

$$\log |p_n(z)| = \ln \kappa_n - nV(z; \lambda_n). \tag{22}$$

Recall that another important sequence of probability measures ν_n associated to $\{p_n\}$ was introduced in (5). A remarkable identity connecting the entropy (6) with these two sequences of measures is [20,21]

$$E_n = -2 \ln \kappa_n + 2n I[\lambda_n, \nu_n]. \tag{23}$$

The behavior of leading coefficients κ_n (or equivalently, of the L^2 norms of the monic orthogonal polynomials) is well understood for a wide class of measures. Hence, this identity shifts the attention to the determination of the mutual energy of two basic sequences of measures of the theory of orthogonal polynomials, that are in a certain sense complementary. The zero-counting measure λ_n is concentrated at the zeros of p_n, and under very general assumptions weakly converges (as $n \to \infty$) to the equilibrium measure of the support of the orthogonality measure. By (22), the potential of this limiting measure describes the n-th root asymptotics of the orthogonal polynomials. Measure ν_n, on the contrary, has zero density at the zeros of p_n. This measure plays an important role in the investigation of the ratio asymptotics of orthogonal polynomials and of the limiting behavior of the corresponding recurrence coefficients. Actually, convergence of this measure to the equilibrium measure of an interval is equivalent to the existence of the limits of the recurrent coefficients. A class of orthogonality measures $\mathcal{M}$ for which the corresponding recurrence coefficients have limits is called the *Nevai class*. Thus, if μ is supported on the interval $[-1, 1]$, then $\mu \in \mathcal{M}$ if and only if ν_n weakly converge to η. A celebrated theorem of Rakhmanov [7] states that *Erdös condition* ($\mu'(x) > 0$, $x \in [-1, 1]$ a.e.) is sufficient for μ to belong to the class $\mathcal{M}$. A straightforward application of the representation (23) is the weak asymptotics of the entropy E_n:

Proposition 11. *If $\mu \in \mathcal{M}$, then $E_n = o(n)$ as $n \to \infty$.*

Indeed, this follows from the fact that for $\mu \in \mathcal{M}$ both measures ν_n, λ_n weakly converge to the equilibrium measure η, and by lower semi-continuity of the energy,

$$\liminf_n I[\lambda_n, \nu_n] = \gamma,$$

the Robin constant of the support of μ ($\gamma = \ln(2)$ if the support of μ is $[-1, 1]$). On the other hand, $\mu \in \mathcal{M}$ is also sufficient for $\ln \kappa_n^{1/n}$ to converge to the same Robin constant. Finally, by Jensen's inequality, $E_n \leqslant 0$, which yields the assertion above.

Remark 12. Proposition 11 does not hold for measures with unbounded support. For instance, for the Freud weights (19),

$$\lim_n \frac{E_n}{n} = -\frac{2}{\lambda}.$$

There are other equivalent representations of the mutual energy $I(\lambda_n, \nu_n)$, and by (23), for the entropy E_n, that are (at least, potentially) useful. For instance, using Fubini's theorem we can readily express this mutual energy in terms of the logarithmic potential of the measure ν_n (see (4)):

$$I[\lambda_n, \nu_n] = \frac{1}{n} \sum_{j=1}^{n} V(\zeta_j^{(n)}; \nu_n). \tag{24}$$

It is a remarkable fact that the potential in the r.h.s. is actually evaluated at its local minima:

Proposition 13. *For the logarithmic potential of the measure ν_n we have*

$$\frac{\mathrm{d}}{\mathrm{d}x} V(x, \nu_n)\Big|_{x=\zeta_j^{(n)}} = 0, \qquad \frac{\mathrm{d}^2}{\mathrm{d}x^2} V(x, \nu_n)\Big|_{x=\zeta_j^{(n)}} > 0.$$

Indeed, the first equality above follows from the identity

$$\int \frac{p_n^2(x)}{t-x} \mathrm{d}\mu(x) = p_n(t) \int \frac{p_n(x)}{t-x} \mathrm{d}\mu(x),$$

while the inequality holds because

$$\int \frac{p_n^2(x)}{(t-x)^2} \mathrm{d}\mu(x) > 0.$$

When μ is supported on $[-1, 1]$, another curious representation for the mutual energy in terms of the generalized moments of ν_n and λ_n was found in [21]. Let $T_k(x) = \cos(k \arccos x)$ denote, as usual, the Chebyshev polynomials of the first kind, and let

$$c_{k,n} = \int T_k(x) \mathrm{d}\lambda_n(x), \qquad m_{k,n} = \int T_k(x) \mathrm{d}\nu_n(x), \quad k, n \geq 0. \tag{25}$$

Obviously, $|c_{k,n}| \leq 1$ and $|m_{k,n}| \leq 1$ for all values of k and n.

Theorem 14. *With the assumptions and notation explained above,*

$$I[\lambda_n, \nu_n] = \ln 2 + 2 \sum_{k=1}^{\infty} \frac{c_{k,n} m_{k,n}}{k}, \tag{26}$$

where the series on the right-hand side is convergent.

Moreover, if we denote

$$M_n := \sup_{x \in [-1,1]} \int_{-1}^{1} \left| \frac{p_n^2(x) - p_n^2(t)}{x-t} \right| \mathrm{d}\mu(x) < +\infty, \tag{27}$$

then for $N \in \mathbb{N}$ we have

$$\left| I[\lambda_n, \nu_n] - \ln 2 - 2 \sum_{k=1}^{N} \frac{c_{k,n} m_{k,n}}{k} \right| \leq \frac{4M_n}{N+1}. \tag{28}$$

Formula (26) was used successfully in [21] as the cornerstone of an efficient algorithm for the numerical evaluation of the entropy on a finite interval. It would be interesting to explore its asymptotic implications, as well as to extend it to the case of the unbounded support.

5. Asymptotics in the Szegő class and beyond

As we mentioned, establishing the asymptotics (12) in the whole Szegő class is still an open problem. The most recent result in this sense is a characterization of this behavior in terms of the growth of p_n's that has been obtained in [22]. Assume again that $d\mu(x) = w(x)dx$ is an absolutely continuous measure on the interval $[-1, 1]$. For $M > 0$ let us denote

$$\Delta_n(M) := \{x \in [-1, 1] : p_n^2(x)w_0(x) \geq M\}, \tag{29}$$

where $w_0(x) = w(x)/\rho(x)$ is the trigonometric weight introduced in Section 2.

Theorem 15. *Assume that the weight w belongs to the Szegő class (S). Then, for all $M > 2$, as $n \to \infty$,*

$$E_n(w) = E^{(0)} - \mathcal{K}(\mu, \eta) - \int_{\Delta_n(M)} p_n^2(x) \ln^+(p_n^2(x))\, w(x)dx + o(1), \tag{30}$$

where $\ln^+(x) = \max\{\ln(x), 0\}$, $x > 0$.

As a simple consequence of the above formula, we conclude that if the weight $w \in (S)$, then

$$\limsup_{n\to\infty} E_n(w) \leq E^{(0)} - \mathcal{K}(\mu, \eta),$$

and (12) holds if and only if there exists a constant $M > 2$, such that

$$\lim_n \int_{\Delta_n(M)} p_n^2(x) \ln^+(p_n^2(x))\, w(x)dx = 0; \tag{31}$$

in this case (31) is valid for for all $M > 2$.

Some sufficient conditions for (31) were specified in [22]. For instance, if there exists $\varepsilon > 0$ such that

$$\sup_n \int_{-1}^{1} \left(\ln^+(p_n^2(x))\right)^{1+\varepsilon} p_n^2(x)w(x)dx < \infty, \tag{32}$$

then we have (12). In particular, these results supersede and generalize those obtained earlier in [14].

Moreover, if we know that $\nu_n \to \eta$ as $n \to \infty$ in the weak-* topology (for instance, if w satisfies the Erdös condition, $w(x) > 0$ a.e. on $[-1, 1]$), then it follows from the weak upper semicontinuity of the mutual entropy [23, Corollary 5.3] that $\limsup E_n \leq \mathcal{K}(\mu, \eta)$. In particular, it shows that if such a weight is not in the Szegő class, then

$$\lim_{n\to\infty} E_n = -\infty.$$

But how fast the entropy E_n diverges to $-\infty$ if $w \notin (S)$?

There is another motivation to extend these results beyond the Szegő class. Recall that the Erdös condition on the weight is sufficient to assure that both λ_n and ν_n tend (as $n \to \infty$) to the equilibrium distribution η on $[-1, 1]$. In particular, from the convexity properties of the mutual energy it follows that

$$\lim_{n\to\infty} I[\lambda_n, \nu_n] = I(\eta) = \ln(2).$$

What is more surprising is that in the Szegő class, the next term of the asymptotic expansion of $I[\lambda_n, \nu_n]$ also exhibits a "universal" behavior, in the sense that it does not depend on the choice of the weight w. Namely, from Theorem 15 it follows

Corollary 16. *Assume that w is a weight in the Szegő class (S) such that (12) holds. Then the mutual energy $I(\lambda_n, \nu_n)$ has the following asymptotic expansion:*

$$I[\lambda_n, \nu_n] = \ln(2) - \frac{1}{2n} + o\left(\frac{1}{n}\right), \quad n \to \infty. \tag{33}$$

It is easy to conjecture that the 1/2 coefficient of n^{-1} is the logarithmic capacity of $[-1, 1]$. Can we drop the assumption $w \in (S)$ in Corollary 16?

The Pollaczek polynomials constitute the first and the best known example of a family of orthogonal polynomials on $[-1, 1]$ with respect to a weight *not satisfying* the Szegő condition. The *symmetric* Pollaczek polynomials, $p_n^\lambda(x; a)$, depend on two real parameters, $\lambda > 0$, $a \geq 0$, and may be defined by the recurrence relation

$$xp_n^\lambda(x; a) = a_{n+1}\, p_{n+1}^\lambda(x; a) + a_n\, p_{n-1}^\lambda(x; a), \qquad p_{-1}^\lambda(x; a) = 0, \qquad p_0^\lambda(x; a) = 1,$$

with the coefficients

$$a_n = \frac{1}{2}\sqrt{\frac{n(n + 2\lambda - 1)}{(n + \lambda + a)(n + \lambda + a - 1)}}.$$

Pollaczek polynomials $p_n^\lambda(x; 0)$ (that is, for $a = 0$) reduce to orthonormal Gegenbauer polynomials with parameter λ; for $a > 0$ the orthogonality weight $w = w_\lambda(\cdot; a) \notin (S)$.

Theorem 17 ([24]). *For the symmetric Pollaczek weight* $w = w_\lambda(\cdot; a)$, *with* $a \geq 0$ *and* $\lambda \geq 1$,

$$E_n = -2a\ln(n) + \tau(\lambda; a) + o(1), \quad n \to \infty,$$

where

$$\tau(\lambda; a) := 2a - 1 + \ln\left(\frac{\Gamma(\lambda + a)\Gamma(\lambda + a + 1)}{\Gamma(2\lambda)}\right). \tag{34}$$

Moreover,

$$I[\lambda_n, \nu_n] = \ln(2) - \frac{1-2a}{2n} + o\left(\frac{1}{n}\right), \quad n \to \infty. \tag{35}$$

Remark 18. The value $\tau(\lambda; 0) = -1 + \log(\Gamma(\lambda)\Gamma(\lambda+1)/\Gamma(2\lambda))$ matches E_∞ in (17) for orthonormal Gegenbauer polynomials ($\alpha = \beta = \lambda - 1/2$), found in [14].

Comparing (33) and (35) we see that we cannot omit the conjecture $w \in (S)$ from Corollary 16. Observe however that the second term of asymptotics is still independent of the main parameter λ, and coincides with (33) for $a = 0$. An interesting open problem is to compute the n^{-1} term in the asymptotic expansion (33) for a general class of weights.

Finally, one more closing remark is appropriate. The limit in (12) contains two terms. The "universal" term, $E^{(0)}$, depends only on the equilibrium measure, and according to [22], is present in the asymptotic expressions beyond the Szegő class. The other term, given by the relative entropy $\mathcal{K}(\mu, \eta)$, encodes the features of the orthogonality measure and its class; it blows up as soon as we drop the (S) condition. Comparing (12) with the identity (23) we may conclude that peculiarities of the asymptotic behavior of the leading coefficient κ_n are mirrored by the asymptotics of the entropy E_n.

Acknowledgements

AIA was supported by grants RFBR-08-01-00179 and NSh-4466.2006.1. JSD was partially supported by Ministerio de Educación y Ciencia (MEC) of Spain, grant FIS 2005-00973, and by Junta de Andalucía (JA), grant FQM-207. AMF acknowledges support from MEC under grant MTM2005-09648-C02-01, and from JA, grant FQM-229. Additionally, JSD and AMF were partially supported by the excellence grants FQM-481 and P06-FQM-01738 from JA.

Last but not least, AIA and AMF wish to take advantage of the recently celebrated 60th anniversary of the third coathor, JSD, to express our gratitude for his permanent support and friendship.

References

[1] C.E. Shannon, W. Weaver, The Mathematical Theory of Communication, AMS Colloquium Publications, University of Illinois Press, Urbana, 1949.
[2] R.M. Dreizler, E.K.U. Gross, Density Functional Theory: An Approach to the Quantum Mechanics, Springer Verlag, Heidelberg, 1990.
[3] K.I. Babenko, An inequality in the theory of Fourier integrals, Izv. Akad. Nauk SSSR Ser. Mat. 25 (1961) 531–542; Engl. Transl. Amer. Math. Soc., Transl., II. Ser. 44 (1965) 115–128.
[4] I. Bialynicki-Birula, J. Mycielsky, Uncertainty relations for information entropy in wave mechanics, Comm. Math. Phys. 44 (1975) 129–132.
[5] J.C. Angulo, J.S. Dehesa, Tight rigorous bounds to atomic information entropies, J. Chem. Phys. 97 (1992) 6485–6495.
[6] J.S. Dehesa, A. Martínez-Finkelshtein, J. Sánchez-Ruiz, Quantum information entropies and orthogonal polynomials, in: Proceedings of the Fifth International Symposium on Orthogonal Polynomials, Special Functions and their Applications (Patras, 1999), J. Comput. Appl. Math. 133 (2001) 23–46.
[7] E.A. Rakhmanov, On the asymptotics of the ratio of orthogonal polynomials, Math. USSR Sb. 32 (1977) 199–213.
[8] A.I. Aptekarev, J.S. Dehesa, A. Martínez-Finkelshtein, R. Yáñez, Discrete entropies of orthogonal polynomials, Constr. Approx., Online First, in press (doi:10.1007/s00365-008-9028-9). Also preprint arXiv:0710.2134.
[9] B. Simon, Orthogonal Polynomials on the Unit Circle I and II, in: AMS Colloquium Publications, vol. 54, American Mathematical Society, Providence, RI, 2005.
[10] S.N. Bernšteĭn, Sobranie sočineniĭ. Tom II. Konstruktivnaya teoriya funkciĭ [1931–1953], Izdat. Akad. Nauk SSSR, Moscow, 1954.
[11] G. Szegő, Orthogonal Polynomials, 4th ed., in: Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, RI, 1975.
[12] A.I. Aptekarev, J.S. Dehesa, R.J. Yáñez, Spatial entropy of central potentials and strong asymptotics of orthogonal polynomials, J. Math. Phys. 35 (9) (1994) 4423–4428.
[13] J.S. Dehesa, W. Van Assche, R.J. Yáñez, Information entropy of classical orthogonal polynomials and their application to the harmonic oscillator and Coulomb potentials, Methods Appl. Anal. 4 (1997) 91–110.
[14] A.I. Aptekarev, V.S. Buyarov, J.S. Dehesa, Asymptotic behavior of the L^p-norms and the entropy for general orthogonal polynomials, Russian Acad. Sci. Sb. Math. 82 (2) (1995) 373–395.
[15] A.I. Aptekarev, V. Buyarov, J.S. Dehesa, R. Yañez, Strong asymptotics for Laguerre polynomials and information entropy of 2D harmonic oscilator and Coulomb potentials, J. Math. Phys. 39 (1994) 1–11.
[16] V.A. Kaliaguine, On asymptotics of L_p extremal polynomials on a complex curve ($0 < p < \infty$), J. Approx. Theory 74 (2) (1993) 226–236.
[17] D.S. Lubinsky, E.B. Saff, Strong asymptotics for extremal polynomials associated with weights on **R**, in: Lecture Notes in Mathematics, vol. 1305, Springer-Verlag, Berlin, 1988.
[18] P.K. Suetin, V.A. Steklov's Problem in the Theory of Orthogonal Polynomials, in: Mathematical Analysis, vol. 15, Akad. Nauk SSSR Vsesojuz. Inst. Naučn. i Tehn. Informacii, Moscow, 1977, pp. 5–82. (in Russian).
[19] E.A. Rahmanov, Steklov's conjecture in the theory of orthogonal polynomials, Mat. Sb. (N.S.) 108(150) (4) (1979) 581–608; 640. English translation: Math. USSR Sb. 36 (4) (1980) 549–575.
[20] A.I. Aptekarev, V.S. Buyarov, J.S. Dehesa, W. Van Assche, Asymptotics for entropy integrals of orthogonal polynomials, Russian Acad. Sci. Dokl. Math. 53 (1996) 47–49.

[21] V. Buyarov, J.S. Dehesa, A. Martínez-Finkelshtein, J. Sánchez-Lara, Computation of the entropy of polynomials orthogonal on an interval, SIAM J. Sci. Comput. 26 (2) (2004) 488–509.
[22] B. Beckermann, A. Martínez-Finkelshtein, E.A. Rakhmanov, F. Wielonsky, Asymptotic upper bounds for the entropy of orthogonal polynomials in the Szegő class, J. Math. Phys. 45 (11) (2004) 4239–4254.
[23] R. Killip, B. Simon, Sum rules for Jacobi matrices and their applications to spectral theory, Ann. of Math. (2) 158 (1) (2003) 253–321.
[24] A. Martínez-Finkelshtein, J.F. Sánchez-Lara, Shannon entropy of symmetric Pollaczek polynomials, J. Approx. Theory 145 (1) (2007) 55–80.