Journal of Approximation Theory **117**, 152–178 (2002)
doi:10.1006/jath.2002.3691

L^p-Norms and Information Entropies of Charlier Polynomials

Lars Larsson-Cohn

Department of Mathematics, Uppsala University, Box 480, 751 06 Uppsala, Sweden
E-mail: llc@math.uu.se

Communicated by W. Van Assche

Received September 05, 2001; revised January 06, 2002; accepted February 08, 2002

We derive asymptotics for the L^p-norms and information entropies of Charlier polynomials. The results differ to some extent from previously studied orthogonal polynomials, for example, the L^p-norms show a peculiar behaviour with two thresholds. Some complications arise because the measure involved is discrete.

Key Words: Charlier polynomials; L^p-norms; information entropy.

1. INTRODUCTION

There has been a substantial recent activity on the information entropy and the related L^p-norms of orthogonal polynomials. Most of the effort has been put into entropy studies (e.g. [1–4, 10, 11, 14, 16, 18, 19], and the references given therein), a project that was initiated in 1994 [17]. This interest has a quantum mechanical origin, in fact being motivated by an entropy version of Heisenberg's uncertainty principle [7]. However, there has also been applications of L^p-norms to operators and extremization on the Wiener space [24, 25]. For an updated survey, see [15].

Since all cases studied hitherto have involved continuous measures, it would be interesting to investigate the situation for a discrete one. The present paper deals with the non-classical Charlier polynomials, orthogonal with respect to a Poisson distribution. These were introduced by Charlier in 1906 [12] on treating a function expansion problem. They are important in probability theory, for example they appear in expansions of the Edgeworth type in convergence to a Poisson distribution [5]. As we shall see, the discreteness of the measure complicates matters. In return, the results that emerge are interesting and somewhat different from previously studied cases.

The paper is organized as follows: The results are formulated and discussed in Section 2 and proven in Sections 4 and 5. An intermediate

0021-9045/02 $35.00

section treats certain cases of the polynomial asymptotics. The main results are Theorems 2.1 and 2.7.

This paper is a shortened version of [26], where further information can be found.

1.1. Notation and Preliminaries

Let $a > 0$. The Charlier polynomials $C_n(x; a)$ may be defined by

$$C_n(x;a) = \sum_{k=0}^{n} \binom{n}{k} (x)_k (-a)^{n-k}, \tag{1.1}$$

where $(x)_k = x(x-1)\cdots(x-k+1)$ is the falling factorial power. They satisfy the recurrence formula

$$C_{n+1}(x;a) = (x-n-a)C_n(x;a) - anC_{n-1}(x;a) \tag{1.2}$$

and, most important, the orthogonality relation

$$\sum_{x=0}^{\infty} C_m(x;a)C_n(x;a)w(x) = a^n n! \delta_{mn}, \tag{1.3}$$

where w is the weight function of a Poisson distribution with parameter a:

$$w(x) = \frac{a^x e^{-a}}{x!}, \qquad x = 0, 1, 2, \ldots$$

As a general reference we mention [32]. We shall suppress the dependence on a, writing simply $C_n(x)$. We point out that one has the possibility of different normalizations. Our choice, giving monic polynomials, is common, but in the context of entropies, orthonormal polynomials are more appropriate. They are denoted by $\hat{C}_n = C_n/\sqrt{a^n n!}$. We also note that the cases $p \leqslant 1$ of Theorem 2.1 and (1.1) suggest $C_n/(-a)^n$ as a natural normalization, see also Remark 2.3.

Our computations rely heavily on recent results on strong asymptotics of the Charlier polynomials. A recurring theme is the complication due to the fact that we are dealing with sums rather than integrals. In the context of L^p-norms, $p \neq 2$, these sums can be handled with the Euler–Maclaurin summation, leading to integrals whose asymptotics can be established by a technique related to the saddle point method. (The classical saddle point method, which has been found useful for Hermite polynomials [25], is applicable only when $p = 1$.) The hardest part is to analyse the integrand close to its minimum, for which we use the so-called Lambert W function and heavy computations.

The situation is quite different when $p = 2$ or when entropies are studied. This is due to the fact that the oscillations in the important region are not resolved by sampling over the integers. We shall therefore rely on a combination of Fourier expansions, Poisson's summation formula and the method of stationary phase. It is interesting to note that, although the oscillations are not resolved, we still extract a factor $\pi^{-1}B(p+\frac{1}{2},\frac{1}{2})$, cf. (5.1), just like when applying the Fejér–Bernstein lemma to fast oscillations in an integral, the typical case for continuous measures [1,25]. The final step when computing the entropies is a well-known differentiation procedure, justified by Montel's theorem from complex analysis.

We shall mainly be concerned with asymptotics as $n \to \infty$. Therefore, we let $f \prec g$ have the strong meaning that $f = O(g/n^s)$ for any given $s > 0$, at least if some constants are properly chosen. Moreover, $\sim$ denotes asymptotic equivalence in the same sense, i.e. $f \sim g$ iff $f - g \prec g$, whereas $\asymp$ means equality within constant factors. We shall find it convenient to put $v = n/a$. Finally, c is a positive and finite constant, not necessarily the same on each occurrence, and indicator functions (characteristic functions) are denoted by **1**.

2. MAIN RESULTS

We turn to the formulation and discussion of our main results. The proofs will follow in Sections 4 and 5. We state the results for fixed $a > 0$, but they obviously extend to a in compact subsets of $(0, \infty)$ [21].

2.1. L^p-norms

All L^p-norms will be taken with respect to w unless otherwise stated, so that $\|f\|_p = \{\sum_{x=0}^{\infty} |f(x)|^p w(x)\}^{1/p}$. The following theorem describes the asymptotics of these norms of the Charlier polynomials. Recall that $\|C_n\|_2 = \sqrt{a^n n!}$ and $v = n/a$.

THEOREM 2.1. *The following holds as $n \to \infty$:*
(a) *If $2 < p < \infty$, then*

$$\|C_n\|_p = c(p)(n!)^{1-1/p} a^{n/p} n^{-(p-1)^2/2p^2} e^{\Sigma_1}, \tag{2.1}$$

where Σ_1 is an asymptotic series in n with leading term $av^{1-1/p}$, see (4.9) *and* (4.16).
(b) *If $1 < p < 2$, then*

$$\|C_n\|_p = c(p)(n!)^{1-1/p} a^{n/p} n^{-(p-1)/2p^2} e^{\Sigma_2}, \tag{2.2}$$

where Σ_2 *is an asymptotic series in n with leading term* $av^{1/p}$, *see* (4.17) *and* (4.18).

(c) *If* $p = 1$, *then*

$$\|C_n\|_1 = e^{-2a}(2a)^n(1 + O(n^{-1})).$$

(d) *If* $0 < p < 1$, *then*

$$\|C_n\|_p = a^n e^{\Sigma_3}, \tag{2.3}$$

where Σ_3 *is an asymptotic series in n with leading term* $\frac{a}{p}v^p$, *see* (4.19) *and* (4.22).

The constants $c(p)$ *are given by*

$$c(p) = \frac{1}{(2\pi a^{1/p})^{(p-1)/2p} p^{1/2p}}, \qquad p > 2,$$

$$c(p) = \frac{1}{(2\pi a^{p+1})^{(p-1)/2p} p^{1/2p}}, \qquad 1 < p < 2.$$

Remark 2.2. The presence of the Σ's makes these formulas a little untransparent. For concreteness, take $p = 3$ and $a = 1$. Then, $v = n$, $\delta = \varepsilon = \eta = \frac{1}{3}$ in (4.9), and

$$\|C_n\|_3 = \frac{(n!)^{2/3} n^{-2/9}}{(2\pi)^{1/3} 3^{1/6}} \exp\left(n^{2/3} - \frac{2}{3}n^{1/3} + \frac{4}{9} + O(n^{-1/3})\right).$$

To obtain such formulas one needs Σ_i to an absolute error of $o(1)$, for which our estimations suffice if $0 < p < \frac{2}{3}$, $\frac{4}{3} < p < \frac{8}{5}$, or if $\frac{8}{3} < p < 4$. By computing more terms in the asymptotic series, one can, in principle, do the same for any p, although this seems like a hard task if p is close to 1, 2 or ∞.

We remark that the results may be stated in a conciser, but less informative form. For example, it follows from (4.14) together with the subsequent argument that

$$\|C_n\|_p^p = g(\beta_0)e^{-F(\beta_0)}\sqrt{\frac{2\pi}{F''(\beta_0)}}\exp\left\{-\frac{p}{16a} - \frac{ap}{2}v^{\delta-\varepsilon}\right\}(1 + o(1))$$

if $p > \frac{7}{3}$. Here g and F are as in (4.2) and (4.3), and $\beta_0 > 1$ is the zero of F'. The problem is to compute $F(\beta_0)$ to sufficient accuracy.

Remark 2.3. For the sake of completeness, we briefly discuss the case $p = 0$. Let $\|f\|_0 = \exp\int \log|f|$, a geometric mean of $|f|$. This is a natural definition, since $\|f\|_0 = \lim_{p\to 0} \|f\|_p$ for any $f \in L^{0+} := \bigcup_{p>0} L^p$ on probability spaces [20]. (Note, however, that other definitions of L^0 and $\|\cdot\|_0$ exist in the literature [6, 24].) Now, it readily follows from (3.2) that

$$\log |C_n(x)| = x \log v + n \log a + O((1 + x^2)/n)$$

for integral $x \in [0, (1-\varepsilon)n]$, $0 < \varepsilon < 1$. Hence,

$$\sum_{x=0}^{(1-\varepsilon)n} \log |C_n(x)| w(x) = a \log v + n \log a + O(n^{-1}).$$

If we could estimate the corresponding sum with $x > (1-\varepsilon)n$ properly, we would thus have

$$\|C_n\|_0 \stackrel{?}{=} a^n v^a (1 + O(n^{-1})). \tag{2.4}$$

Since upper bounds are trivial, e.g. $|C_n(x)| \leqslant 2^n x! \max(1, a^n)$, the problem is to give lower bounds, i.e. to show that $C_n(x)$ is sufficiently far away from zero. By analysing the proof of Theorem 3.1, notably the fact that z stays away from 1, one sees that the regions occurring there cause no trouble. Hence, the question boils down to giving lower bounds on $|C_n(x)|$ for integers $|x - n| \leqslant m\sqrt{n}$, $m > 2\sqrt{a}$, which seems difficult due to the irregular oscillations in that region.

Note that formal differentiation of (2.3) gives the same result, but that such a procedure is not easy to justify. In any case, $\limsup_{n\to\infty} \|C_n\|_0 / (a^n v^a) \leqslant 1$. We also remark that numerical evidence supports (2.4).

Remark 2.4. It is interesting to compare these results to the previously investigated Jacobi and Hermite polynomials [1, 25]. For these polynomials there is (unless $\alpha, \beta \leqslant -\frac{1}{2}$ in the Jacobi case) a threshold value p_0 with the following property: all L^p-norms with $p < p_0$ grow at the same rate; on the threshold the growth is a little stronger, after which it increases quickly with p. For Hermite polynomials, $p_0 = 2$; for Jacobi polynomials, p_0 can, depending on the parameters α and β, take any value in $(2, \infty)$.

In the present case there are two threshold values: $p = 1$ and 2. However, $p = 2$ appears to be a "weak" threshold, cf. Remark 2.6. On the contrary, $p = 1$ has many of the characteristics of a typical threshold, including the rapid change of dominating region, cf. Remark 2.5. However, the L^p growth rate increases also for $p < 1$: $\|C_n\|_p = o(\|C_n\|_q)$ whenever $0 \leqslant p < q < \infty$. This seems in fact to be a phenomenon, not previously observed for orthogonal polynomials.

The striking behaviour with two thresholds resembles the multimodal oscillations of the Charlier polynomials, although we do not know if there is a deeper connection.

Remark 2.5. A related question is where the main contribution to the norms comes from. If $p \neq 2$, the proofs in Section 4 and [26] show that the bulk of the mass is contained in Gaussian peaks, the centre and width of which are given in Table I. For $p = 2$ the situation is different. Refining the argument of Section 5, it is not hard to show that the mass is smeared out over the interval $|x - n| \leqslant 2\sqrt{an} + cn^{1/6} \log^{2/3} n$; dominant in the sense of $\sim$ above. The same statement applies to the entropies. On the other hand, the L^0-mass seems to follow, without normalization, a shifted Poisson distribution, cf. Remark 2.3.

The case $p = 2$ is interesting from a general point of view. Orthogonal polynomials always have oscillating regions, and one would expect the L^2-mass to be concentrated to these, since this is where the orthogonality "takes place". The Charlier polynomials have a multiple-mode of oscillations, in effect being oscillating for $0 \leqslant x \leqslant n + 2\sqrt{an}$, but the important oscillations seem to be the central ones, close to $x = n$.

Remark 2.6. As mentioned, the threshold $p = 2$ is "weak" in many senses; the behaviour for $1 < p < 2$ and $p > 2$ show large, though not complete, similarities, cf. Remark 2.9. It is worth noting that parts of the similarities may be viewed as passing to the conjugate exponent $p' = p/(p-1)$; for example this is true for the values in Table I. Changing p into p' also takes Σ_1 into Σ_2, as far as we have computed them, but with the sign of some terms reversed. We do not know whether a duality argument might explain these symmetries.

2.2. *Information Entropies*

The Boltzmann–Shannon information entropy of a probability density $\rho(x)$ on $\mathbf{R}^d$ is defined as $S(\rho) = -\int \rho \log \rho \, dx$ [31]. In quantum mechanical

TABLE I
The Centre (Dominating Term) and Width of the Gaussian Peaks Contributing to the L^p-Norms, cf. Remark 2.5. The Values are Given in Units of $\beta = x/n$, Whereas $\nu = n/a$. For $p = 2$ and 0 the Mass Distribution is Non-Gaussian

p	Centre	Width
$(2, \infty)$	$1 + \nu^{-1/p}$	$n^{-(p+1)/2p}$
$(1, 2)$	$1 - \nu^{-(p-1)/p}$	$n^{-(2p-1)/2p}$
1	$1/2$	$n^{-1/2}$
$(0, 1)$	ν^{p-1}	$n^{-(2-p)/2}$

applications one typically has $\rho = |\Psi|^2$, Ψ being the wave function. For many systems, Ψ is given in terms of orthogonal polynomials P_n with respect to some measure μ. The Boltzmann–Shannon entropy is then closely related to functionals of the form

$$S_n(P) = \int P_n^2 \log P_n^2 \, d\mu,$$

which [17] has called the *entropy* of P_n, supposed to be orthonormalized. Note that $S_n(P) \geqslant 0$ if μ is a probability measure by Jensen's inequality. If the distribution is discrete, as in our case, all integrals should be replaced by sums. In particular,

$$S_n(\hat{C}) = \sum_{x=0}^{\infty} \hat{C}_n(x)^2 \log \hat{C}_n(x)^2 w(x). \tag{2.5}$$

Our result about the Charlier entropies is the following.

THEOREM 2.7. *Let $\hat{C}_n$ be the orthonormalized Charlier polynomials. Then, with notation* (2.5),

$$S_n(\hat{C}) = (n+a)\log \frac{n}{ae} + 3a + 1 - \frac{1}{2}\log 2\pi a + o(1) \tag{2.6}$$

as $n \to \infty$.

Remark 2.8. This $n \log n$ growth seems to be new. Earlier studied entropies grow like n (Freud, Laguerre) or are bounded (Jacobi and some other polynomials on compact intervals) [1]. This discrepancy vanishes partly if we instead consider the Boltzmann–Shannon entropy $B_n(p) := -\int p_n^2 w \log(p_n^2 w)$ or the corresponding sum. Namely, the Charlier polynomials satisfy $B_n(\hat{C}) = -N_n'(1) = \frac{1}{2}\log n + c + o(1)$, cf. (5.4). From [1,3] it is easily seen that $B_n = c_1 \log n + c_2 + o(1)$ for Freud and Laguerre polynomials as well.

We remark that $B_n(\hat{C}) = \log \sqrt{n} + O(1)$ which is reasonable, since the number of contributing integers is of the order $\sqrt{n}$, and the (unit) L^2-mass is fairly uniformly distributed among these, cf. Remark 2.5.

Remark 2.9. Theorem 2.7 implies that

$$\frac{d}{dp}\|C_n\|_p \bigg|_{p=2} = \frac{1}{4}\|C_n\|_2 S_n(\hat{C}) = \frac{1}{4}\|C_n\|_2 \left\{ n \log \frac{n}{ae} + O(\log n) \right\}, \tag{2.7}$$

which is interesting in the light of Theorem 2.1. Namely, the latter asserts that, for fixed $p > 2$, $\|C_n\|_p = (n!)^{1-1/p} a^{n/p} \exp\{n^{1-1/p}(c + o(1))\}$. If formal differentiation was allowed,

$$\frac{d}{dp}\|C_n\|_p = \|C_n\|_p\left\{\frac{\log n!}{p^2} - \frac{n\log a}{p^2} + O(n^{1-1/p}\log n)\right\},$$

which, for $p = 2$, is (2.7) with a larger error. A similar remark applies for $p<2$. This nicely illustrates the weakness of the threshold $p = 2$.

Remark 2.10. Results on entropies always have bearing on logarithmic potential theory. Namely, the logarithmic potential of a Borel measure μ on $\mathbf{C}$ is defined as $V(z;\mu) = -\int \log|z-x|\,d\mu(x)$. If we take $\mu = \sum_{x=0}^{\infty} w(x)\delta_x$ as the Poisson measure and put $d\nu_n(x) = \hat{C}_n(x)^2\,d\mu(x)$, then $S_n(\hat{C}) = 2\log k_n - 2\sum_{j=1}^{n} V(\zeta_{n,j};\nu_n)$, where $\zeta_{n,j}$ are the zeros and $k_n = (a^n n!)^{-1/2}$ is the leading coefficient of $\hat{C}_n$ [15, Sect. 3]. It follows from Theorem 2.7 that

$$-\sum_{j=1}^{n} V(\zeta_{n,j};\nu_n) = n\log\frac{n}{e} + \frac{2a+1}{4}\log\frac{ne^2}{a} + o(1)$$

as $n\to\infty$. Note that $\zeta_{n,j}$ are the local minima of $V(\cdot;\nu_n)$ [19].

3. ASYMPTOTICS OF THE CHARLIER POLYNOMIALS

Unlike most classical polynomials, the Charlier polynomials do not satisfy a second-order linear differential equation, rendering the task of establishing sharp asymptotics more difficult. The first approach, due to Maejima and Van Assche [27, 33] was probabilistic and valid for $x<0$. Goh [22] used integral representations, and his results were improved by Rui and Wong [30], still covering only $\varepsilon n \leqslant x \leqslant Mn$.

A completely different method was used in an ingenious paper by Dunster [21], who, via a hypergeometric representation, derived a differential equation for the Charlier polynomials with the roles of the parameter a and the variable x reversed. This enabled him to use the theory of asymptotics for differential equations [8, 28] to prove complete and uniform asymptotics for all real x, even uniformly in a, subject to certain restrictions.

We shall localize and extend Dunster's results to suit our needs. Note that (b) is a sharpened version of Goh's Theorem 1 [22]. However, (a) does not resemble his Theorem 7, due to the fact that the zeros of C_n lie close to the integers, making the leading term vanish there.

For the reader's convenience, we collect some notation used in the theorem. Thus, $x = n\beta$ and $\rho = n|1-\beta|$. In addition, $z = (n+\frac{1}{2})\sqrt{\zeta}/\rho$, where ζ is given by (3.6), and

$$\psi(z) = \operatorname{arcsech} z - \sqrt{1-z^2} - \log\frac{2}{ez} - \frac{1}{4}z^2 = \frac{z^4}{32} + O(z^6) \tag{3.1}$$

as $z \to 0$.

THEOREM 3.1. *Let $M > 1$ and $m > 2\sqrt{a}$ be fixed constants. Moreover, let ρ, ψ, and z be as in the proof given below, see* (3.8)–(3.9), *or the last paragraph above. Then the following hold as $n \to \infty$:*

(a) *If $0 \leqslant \beta \leqslant 1 - mn^{-1/2}$ and $n\beta$ is an integer, then*

$$C_n(n\beta) = \frac{(-1)^{n(1-\beta)} a^n e^{-a\beta/(1-\beta)}}{(1-\beta)^{n(1-\beta)+1/2}} \left(\frac{n}{ae}\right)^{n\beta} e^{-\rho\psi(z)} \left(1 + O\left(\frac{1}{n(1-\beta)^2}\right)\right). \tag{3.2}$$

(b) *If $1 + mn^{-1/2} \leqslant \beta \leqslant M$, then*

$$C_n(n\beta) = \frac{n!\beta^{n\beta+1/2} e^{-a/(\beta-1)}}{\sqrt{2\pi n}(\beta-1)^{n(\beta-1)+1/2}} e^{-\rho\psi(z)} \left(1 + O\left(\frac{1}{n(\beta-1)^2}\right)\right). \tag{3.3}$$

If β is bounded away from 1, then $\rho\psi(z) = O(n^{-1})$, and the factor $e^{-\rho\psi(z)}$ can be ignored.

Proof. The proof of (b) is similar to the case $\beta \geqslant \varepsilon$ of (a), and so we only prove (a). Fixing a small number ε, we divide this into the cases $\beta \leqslant \varepsilon$ and $\beta \geqslant \varepsilon$.

Let us start with the former, assuming without loss of generality that $x \geqslant 10$, say. (Otherwise, $\theta = O(n^{-1})$ in (3.4), and the result is immediate.) By (1.1),

$$C_n(x) = \sum_{k=0}^{x} \binom{n}{k} (x)_k (-a)^{n-k} =: \sum_{k=0}^{x} T_k,$$

since x is an integer. Introducing

$$q_k = \frac{T_{x-k}}{T_{x-k+1}} = -\frac{a(x-k+1)}{k(n-x+k)} = -\frac{\theta}{k} + O(n^{-1}), \tag{3.4}$$

where

$$\theta = \frac{ax}{n-x},$$

we have $C_n(x) = T_x(1 + q_1 + q_1q_2 + \cdots + q_1 \cdots q_x)$. Taking ε small, we may assume that $\theta \leqslant \frac{1}{2}$. Now, by (3.4),

$$q_1 \cdots q_k = \frac{(-\theta)^k}{k!} + \eta_k$$

with $|\eta_k| \leqslant (\theta + O(n^{-1}))^k - \theta^k$. Hence,

$$C_n(x) = T_x\left(\sum_{k=0}^{x} \frac{(-\theta)^k}{k!} + \eta\right)$$

with

$$|\eta| \leqslant \sum_{k=0}^{\infty}((\theta + O(n^{-1}))^k - \theta^k) = O(n^{-1}).$$

Since $\sum_{k>x} \theta^k/k! = O(n^{-1})$ as well, we conclude that $C_n(x) = T_x e^{-\theta}(1 + O(n^{-1}))$, and the result follows from Stirling's formula.

For the case $\varepsilon \leqslant \beta \leqslant 1 - mn^{-1/2}$ we shall use Dunster's Subcase IIa [21]. Combined with [21, Sect. 4] and estimate (3.11) from [8] this gives, for integers $x = n\beta$ in the interval under consideration,

$$\begin{aligned} C_n(n\beta) = (-1)^{n(1-\beta)} n! e^{a/2} \beta^{n\beta/2+1/4} \left(\frac{ae}{n}\right)^{n(1-\beta)/2} \\ \times J_{n(1-\beta)}((n + \tfrac{1}{2})\sqrt{\zeta})(1 + O(n^{-1})), \end{aligned} \tag{3.5}$$

where J is a Bessel function of the first kind and

$$\zeta = \zeta(t) = c_1 t + c_2 t^2 + O(t^3) \tag{3.6}$$

(uniformly in β) is an analytic function of $t = a/(n + \frac{1}{2}) \to 0$. The first two coefficients of this Taylor expansion are given by

$$c_1 = \frac{4}{e} A =: \frac{4}{e}\left(\frac{n\beta + \frac{1}{2}}{n + \frac{1}{2}}\right)^{-(n\beta+1/2)/n(1-\beta)},$$

$$c_2 = \frac{c_1}{(1-\beta)^2}\left(1 + \frac{1}{2n}\right)^2 \left(\frac{c_1}{2} - \frac{1 + \beta + n^{-1}}{1 + \frac{1}{2}n^{-1}}\right).$$

We turn our interest to the Bessel function. By Eq. (7.16) of [28, Chap. 10],

$$J_\rho(\rho z) = \frac{e^{-\rho\gamma}}{\sqrt{2\pi\rho}}(1 + O(\rho^{-1} + z^2)) \tag{3.7}$$

uniformly in $0 < z \leqslant z_0 < 1$ as $\rho \to \infty$, where

$$\gamma = \operatorname{arcsech} z - \sqrt{1 - z^2} = : \log\frac{2}{ez} + \frac{1}{4}z^2 + \psi(z). \tag{3.8}$$

Thus, $\psi(z) = z^4/32 + O(z^6)$ as $z \to 0$. Moreover, the assumption $m > 2\sqrt{a}$ implies that z is bounded away from one, so that γ stays away from zero. We have $\rho = n|1 - \beta|$ and

$$z = \frac{(n + \frac{1}{2})\sqrt{\zeta}}{n|1 - \beta|} = \frac{\sqrt{c_1 a}}{\sqrt{n}|1 - \beta|}\Big(1 + \frac{\mu}{n} + O(n^{-2})\Big) \tag{3.9}$$

with

$$\mu = \frac{ac_2}{2c_1} + \frac{1}{4}.$$

(The absolute values make these expressions valid in the case (b) too.) Hence,

$$\gamma = \frac{1}{2}\log n + \log\frac{1 - \beta}{\sqrt{aeA}} + \left(\frac{aA}{e(1 - \beta)^2} - \mu\right)\frac{1}{n} + \psi(z) + O\left(\frac{1}{n^2(1 - \beta)^2}\right).$$

Inserting this into (3.7) gives

$$J_{n(1-\beta)}((n + \tfrac{1}{2})\sqrt{\zeta}) = \frac{n^{-n(1-\beta)/2}}{\sqrt{2\pi n(1 - \beta)}}\left(\frac{\sqrt{aeA}}{1 - \beta}\right)^{n(1-\beta)} e^{-\rho\psi(z)}$$
$$\times \exp\left\{(1 - \beta)\left(\mu - \frac{aA}{e(1 - \beta)^2}\right)\right\}\left(1 + O\left(\frac{1}{n(1 - \beta)^2}\right)\right),$$

which, combined with (3.5) and a little algebra, gives the desired result. ∎

4. L^p-NORMS: PROOF OF THEOREM 2.1

4.1. *The Case* $p > 2$

We start with the case $p > 2$. This is, along with the case $1 < p < 2$, the hardest one, and we shall discuss it in some detail, treating the other cases more briefly. Let m, M and $\beta = x/n$ be as in Theorem 3.1, and write

$$C_n = C_n^{(1)} + C_n^{(2)} + C_n^{(3)}$$

with

$$C_n^{(1)}(x) = C_n(x)\mathbf{1}\{x < n + m\sqrt{n}\},$$

$$C_n^{(2)}(x) = C_n(x)\mathbf{1}\{n + m\sqrt{n} \leqslant x \leqslant Mn\}.$$

We deal with $C_n^{(2)}$ first, this being the main term. Thus

$$\beta \in I := [\tilde{a}, \tilde{b}]$$

with

$$\tilde{a} = 1 + mn^{-1/2}, \qquad \tilde{b} = M.$$

By (3.3) and Stirling's formula, we then have

$$|C_n^{(2)}(x)|^p w(x) = n^{-1} g(\beta) e^{-F(n;\beta)} e^{-np(\beta-1)\psi(z)} \left(1 + O\left(\frac{1}{n(1-\beta)^2}\right)\right) \tag{4.1}$$

with $\psi(z)$ as in Section 3 (cf. (3.1)),

$$g(\beta) = \frac{n(n!)^p e^{-a}}{(2\pi n)^{(p+1)/2}} \beta^{(p-1)/2} e^{-p\beta/16a} \tag{4.2}$$

and

$$\begin{aligned} F(n;\beta) = {} & \frac{ap}{\beta - 1} + \frac{p}{2}\log(\beta - 1) - \frac{p\beta}{16a} + n\beta \log\frac{n}{ae} \\ & + n(p(\beta - 1)\log(\beta - 1) - (p-1)\beta\log\beta). \end{aligned} \tag{4.3}$$

To reduce the notational burden, we suppress the dependence of F on n and write $F(n;\beta) =: F(\beta)$. Note, however, that this dependence is central for the coming asymptotics.

The main part of the computation of $\|C_n^{(2)}\|_p$ is the estimation of the sum

$$S := n^{-1} \sum_{\beta \in n^{-1}\mathbf{Z} \cap I} g(\beta) e^{-F(\beta)}.$$

We shall approximate S by an integral using the Euler–Maclaurin summation formula [9]. Adapted to the present range of β values, the latter reads as

$$S = \int_{\tilde{a}}^{\tilde{b}} h(\beta)\,d\beta + \frac{1}{(2k+1)!n^{2k+1}} \int_{\tilde{a}}^{\tilde{b}} h^{(2k+1)}(\beta)\tilde{B}_{2k+1}(n\beta)\,d\beta$$
$$+ \left\{ \frac{h(\tilde{a}) + h(\tilde{b})}{2n} + \sum_{j=1}^{k} \frac{B_{2j}}{(2j)!n^{2j}} [h^{(2j-1)}]_{\tilde{a}}^{\tilde{b}} \right\} =: S_1 + S_2 + S_3, \tag{4.4}$$

where $h = ge^{-F}$, B_{2j} are the Bernoulli numbers, $\tilde{B}_{2k+1}$ denotes the 1-periodic extension of the $(2k+1)$th Bernoulli polynomial from the interval $[0, 1]$, and $[f]_a^b = f(b) - f(a)$. We have made the harmless assumption that $n\tilde{a}$ and $n\tilde{b}$ are integers.

We treat the main term S_1 first, which requires a study of the function F. Differentiating, we find

$$F'(\beta) = -ap\left(\frac{1}{\beta - 1} - \frac{1}{4a}\right)^2 + n\log\frac{n}{a} + n(p\log(\beta - 1) - (p-1)\log\beta),$$

$$F''(\beta) = \frac{2ap}{(\beta-1)^3} - \frac{p}{2(\beta-1)^2} + n\left(\frac{p}{\beta-1} - \frac{p-1}{\beta}\right). \tag{4.5}$$

Now, it is easy to see that $F''(\beta) > 0$ on I for large n. Moreover, $F'(\tilde{a}) = -(p/2 - 1)n\log n + O(n) \to -\infty$ and $F'(\tilde{b}) = n\log n + O(n) \to +\infty$ as $n \to \infty$. Thus, at least for large n, F' has a unique zero, say $\beta_0 \in I$, corresponding to a strict, global minimum of F.

As usual in such contexts, the main contribution to the integral comes from a small neighbourhood of β_0. We shall, therefore, calculate $\beta_0 = \beta_0(n)$ and, most important, $F(\beta_0)$, to some accuracy. First, note that $\lim_{n\to\infty} F'(\beta) = +\infty$ for any fixed $\beta \in I$. Thus, for large n, $1 < \beta_0 < \beta$, and so $\beta_0 \to 1$. Putting

$$y = \frac{1}{\beta - 1} - \frac{1}{4a},$$

the equation $F'(\beta) = 0$ can be written as

$$y^2 + \frac{n}{a}\log y = \frac{n}{2a}U \tag{4.6}$$

with

$$U = \frac{2}{p}\log\frac{n}{a} - \frac{2(p-1)}{p}\log\beta - 2\log\left(1 + \frac{1}{4ay}\right) = \frac{2}{p}\log\frac{n}{a} + o(1). \tag{4.7}$$

Equation (4.6) has the implicit solution

$$y = e^{U/2} G\left(\frac{2ae^U}{n}\right), \tag{4.8}$$

where $G(x) = \sqrt{W(x)/x} = 1 - \frac{1}{2}x + \frac{5}{8}x^2 - \frac{49}{48}x^3 + \cdots$ and W, known as the Lambert W function [13], satisfies $W(x)e^{W(x)} = x$.

The above can be used to calculate y (and thus β) iteratively. We shall find it convenient to introduce the following notation:

$$\delta = 1/p \in (0, \tfrac{1}{2}), \qquad \varepsilon = 1 - 2\delta \in (0,1), \qquad \eta = \min(\delta, \varepsilon). \tag{4.9}$$

Moreover, $v = n/a$ as always. Thus, $U = 2\delta \log v + o(1)$. Inserting this into (4.8) gives $y = v^\delta + o(v^\delta)$, so that $U = 2\delta \log v + O(v^{-\delta})$. Starting from this, one then iterates, letting (4.7) and (4.8) feed each other. Each iteration reduces the error by a factor v^η. The rather tedious calculations can be found in [26]. The result, to the fourth order, is

$$\begin{aligned} y_0 = {} & v^\delta - v^{\delta-\varepsilon} - \left(\frac{p-1}{p} + \frac{1}{4a}\right) + \frac{5}{2}v^{\delta-2\varepsilon} + 2\left(\frac{p-1}{p} + \frac{1}{4a}\right)v^{-\varepsilon} \\ & + \frac{p-1}{2p^2}v^{-\delta} - \frac{49}{6}v^{\delta-3\varepsilon} - 8\left(\frac{p-1}{p} + \frac{1}{4a}\right)v^{-2\varepsilon} \\ & - \left(\frac{(p-1)(2ap+p-a)}{2ap^2} + \frac{1}{16a^2}\right)v^{-\delta-\varepsilon} \\ & + \frac{(p-1)(p-2)}{3p^3}v^{-2\delta} + O(v^{\delta-4\eta}), \end{aligned}$$

where y_0 corresponds to β_0. This, in turn, means that $\beta_0 = 1 + v^{-\delta} + O(v^{-\delta-\eta})$.

This can be used to compute $F(\beta_0)$, a task that is simplified a little by taking into account that $F'(\beta_0) = 0$. Another page of straightforward but tiresome calculations [26] results in

$$-F(\beta_0) = -(n \log n - n) + n \log a + \frac{1}{2}\log\frac{n}{a} + a + \frac{p}{16a} + p\tilde{\Sigma}_1,$$

where $\tilde{\Sigma}_1$ is an asymptotic series in v, starting with

$$\tilde{\Sigma}_1 = av^{\delta+\varepsilon} - av^{\delta} + \frac{a(p-1)}{2p}v^{\varepsilon} + \frac{a}{2}v^{\delta-\varepsilon} + \frac{a(p-2)}{p} + \frac{a(p-1)(2p-3)}{6p^2}v^{\varepsilon-\delta} + O(v^{1-4\eta}). \tag{4.10}$$

This looks complicated, but we stress that several cancellations take place during the computations, suggesting that there is a simpler way to arrive at the result. For example, the terms $v^{2\delta}$, $v^{2\delta-\varepsilon}$, and $v^{2\delta-2\varepsilon}$ all cancel, making it plausible that the error in (4.10) is actually $O(v^{\delta+\varepsilon-3\eta})$. As for the leading term, note that $\delta+\varepsilon = 1-1/p \in (\frac{1}{2}, 1)$. Using Stirling once again, one finds

$$e^{-F(\beta_0)} = \frac{a^n}{n!}\sqrt{\frac{2\pi n^2}{a}}e^{a+p/16a+p\tilde{\Sigma}_1}(1+O(n^{-1})). \tag{4.11}$$

Moreover,

$$F''(\beta_0) = apv^{1+\delta}(1+O(v^{-\eta})). \tag{4.12}$$

We turn to the local approximation of F near β_0. Let

$$F_j = F^{(j)}(\beta_0), \qquad j \geqslant 0,$$

be the derivatives of F at β_0, and similarly for g. Put

$$\omega = \sqrt{\frac{\lambda \log n}{F_2}},$$

where λ is a large constant, and consider the disc

$$J = \{\beta; |\beta-\beta_0| \leqslant \omega\}$$

in the complex plane. We shall see that the bulk of the contribution comes from $J \cap \mathbf{R}$, which we, by abuse of notation, call J as well.

First, note that if ϕ is an analytic function, $\phi(0) = \phi'(0) = 0$, $\phi''(0) \neq 0$ and $\sup|\phi^{(3)}| \leqslant A$ in a suitable neighbourhood of the origin, then $\phi(z) \neq 0$ for $0 < |z| < 3|\phi''(0)|/A$. In the region $|\beta-\beta_0| < \frac{1}{2}v^{-\delta}$ (say) we have $|F^{(3)}(\beta)| \leqslant c\, n^{1+2\delta}$. Since $F_2 \asymp n^{1+\delta}$ this means that $F(\beta)-F_0 \neq 0$ for $0 < |\beta-\beta_0| \leqslant cn^{-\delta}$. In particular,

$$f(\beta) = \sqrt{F(\beta)-F_0}$$

is analytic on J, and we choose it to be increasing on the real line. We note in passing that $|F^{(j)}| \leqslant cn^{1+(j-1)\delta}$ on J, $j \geqslant 2$, as follows from (4.5).

Now, $F(\beta_0 \pm \omega) - F_0 \geqslant cF_2\omega^2 = c\lambda \log n$, so that $e^{-(F-F_0)} \leqslant n^{-c\lambda}$ on $I \backslash J$. Since anything that occurs in front of this exponential in (4.4), i.e. products of derivatives of g and F, are bounded by fixed (depending on k only) powers of n, we see that by choosing λ large enough, everything outside J is negligible in the sense of $\prec$ in Section 1.1. In particular, $S_3 \prec S_1$. Moreover,

$$S_1 \sim e^{-F_0} \int_J g(\beta) e^{-f(\beta)^2}\, d\beta.$$

On J we introduce $u = f(\beta)$ as a new variable, ranging over the interval

$$J_u := f(J) \supset \{u;\ |u| \leqslant c\sqrt{\lambda \log n}\}.$$

Hence,

$$\int_J g(\beta) e^{-f(\beta)^2}\, d\beta = \int_{J_u} \frac{g(\beta)}{f'(\beta)} e^{-u^2}\, du \tag{4.13}$$

and we must investigate f' more carefully. Expanding F in a Taylor series around β_0 and differentiating formally (which can be justified e.g. by the Cauchy integral formula), one readily finds

$$f'(\beta) = \sqrt{\frac{F_2}{2}} \left(1 + \frac{F_3}{3F_2}(\beta - \beta_0) + O(n^{2\delta}(\beta - \beta_0)^2)\right)$$

on J. Moreover, $\beta - \beta_0 = u\sqrt{2/F_2} + O(u^2/n)$. Hence, the right-hand side of (4.13) equals

$$g_0 \sqrt{\frac{2}{F_2}} \int_{J_u} \left(1 + u\left(\frac{g_1}{g_0} - \frac{F_3}{3F_2}\right)\sqrt{\frac{2}{F_2}} + O\left(\frac{u^2}{n^{\delta+\varepsilon}}\right)\right) e^{-u^2}\, du$$

$$= g_0 \sqrt{\frac{2\pi}{F_2}} (1 + O(n^{-\delta-\varepsilon}))$$

and so

$$S_1 = g_0 e^{-F_0} \sqrt{\frac{2\pi}{F_2}}((1 + O(n^{-\delta-\varepsilon})). \tag{4.14}$$

It remains to take care of S_2. Since any derivative of g is bounded by a constant times g itself, we have, for any $s > 0$,

$$|S_2| \leqslant \frac{c_k}{n^{2k+1}} \int_J \left(\sum_\alpha |f_\alpha(\beta)| \right) g(\beta) e^{-F(\beta)}\, d\beta + O(S_1/n^s),$$

summing over finitely many f_α, each of which is a product of derivatives of F (no undifferentiated functions) of orders summing up to at most $2k+1$, and c_k is a constant depending on k only. Moreover, $|F'(\beta)| \leqslant cF_2|\beta - \beta_0| \leqslant cn^{(1+\delta)/2}\sqrt{\log n}$ on J. Hence, we can estimate each f_α by a constant times $(F')^\tau F^{(\sigma_1)} \cdots F^{(\sigma_m)}$, with $\sigma_i \geqslant 2$ and $\tau + \sum_i \sigma_i \leqslant 2k+1$. Thus,

$$|f_\alpha| \leqslant cn^{\tau(1+\delta)/2 + \sum (1+(\sigma_i - 1)\delta)} \log^{\tau/2} n \leqslant cn^{(m+\tau/2)(1-\delta)+\delta(2k+1)} \log^{\tau/2} n.$$

But $2k+1 \geqslant \tau + \sum \sigma_i \geqslant \tau + 2m$, so this is bounded by

$$n^{(1+\delta)(2k+1)/2} \log^{\tau/2} n \leqslant cn^{3(2k+1)/4}.$$

Since we may take k as large as we please, we see that $S_2 \prec S_1$. Recalling (4.2), (4.11), (4.12), and (4.14), we have shown that

$$S \sim S_1 = \left(\frac{n!}{\sqrt{2\pi}} \right)^{p-1} a^n n^{-(p-1)^2/2p} \frac{e^{p\tilde{\Sigma}_1}}{a^{(p-1)/2p}\sqrt{p}} \tag{4.15}$$

(the errors are absorbed into $\tilde{\Sigma}_1$).

Now, the interval I is chosen so that the O-term in (4.1) is bounded on I and is $O(n^{-\varepsilon})$ on J. Hence, the only problem in passing from S to $\|C_n^{(2)}\|_p^p$ lies in the factor $e^{-np(\beta-1)\psi(z)}$. This will result only in a small correction of $\tilde{\Sigma}_1$; we sketch the reason for this, omitting the details.

It is not hard to see that one needs to only consider the first term $z^4/32$ in the Taylor series of $\psi(z)$, leading to an extra term $ap/2v(\beta-1)^3$ in F. This changes β_0 into $\tilde{\beta}_0$, say, but still $\tilde{\beta}_0 = 1 - v^{-\delta} + O(v^{-\delta-\eta})$. Now, within a region $|\beta - \tilde{\beta}_0| < \frac{1}{2}v^{-\delta}$, $n(\beta-1)z^4$ varies only within constant factors. It follows that essentially all the mass still lies in J. But there,

$$-np(\beta-1)\psi(z) = -\frac{ap}{2}v^{\delta-\varepsilon} - \frac{3ap}{v(\beta_0-1)^4\sqrt{F_2}}u + O(v^{-2\varepsilon}u^2).$$

The first term above is constant and adds into $p\tilde{\Sigma}_1$. The remaining ones result in a relative error $O(v^{\delta-\varepsilon-2\eta})$, which may be absorbed into $\tilde{\Sigma}_1$. Thus, (4.15) holds with S replaced by $\|C_n^{(2)}\|_p^p$ provided that one changes $\tilde{\Sigma}_1$

into $\Sigma_1 = \tilde{\Sigma}_1 - \frac{a}{2}v^{\delta-\varepsilon}$:

$$\begin{aligned}\Sigma_1 &= av^{\delta+\varepsilon} - av^{\delta} + \frac{a(p-1)}{2p}v^{\varepsilon} \\ &\quad + \frac{a(p-2)}{p} + \frac{a(p-1)(2p-3)}{6p^2}v^{\varepsilon-\delta} + O(v^{1-4\eta}).\end{aligned} \tag{4.16}$$

Taking pth roots in (4.15) with these corrections, we get (2.1) with $C_n^{(2)}$ in the place of C_n.

We finally estimate $C_n^{(1)}$ and $C_n^{(3)}$. Let us start with the latter, i.e. $x > M$. Provided that $M \geqslant 1 + a$, the modulus of the terms in (1.1) is increasing, so that $|C_n^{(3)}(x)| \leqslant (n+1)x^n$ and

$$|C_n^{(3)}(x)|^p w(x) \leqslant c^x x^{np-x} \leqslant \left(\frac{c}{x}\right)^{x/2}.$$

Summing over $x \geqslant Mn$ gives $\|C_n^{(3)}\|_p^p = o(1) \prec \|C_n^{(2)}\|_p^p$.

As for $C_n^{(1)}$, split the interval $[0, n + m\sqrt{n})$ into the three parts $[0, n - m\sqrt{n})$, $[n - m\sqrt{n}, n)$ and $[n, n + m\sqrt{n})$, denoting the corresponding polynomials $C_n^{(1,1)}$ through $C_n^{(1,3)}$. The first part is easily estimated by means of (3.2). For the second one, note that if $x = n - O(\sqrt{n})$, then the modulus of the summand in (1.1) is maximized for $k = k_0 = n - O(\sqrt{n})$. It follows that

$$|C_n^{(1,2)}(x)| \leqslant (n+1)\binom{n}{k_0}(x)_{k_0} a^{n-k_0} \leqslant n^{c\sqrt{n}} n!.$$

This gives $\|C_n^{(1,2)}\|_p^p \leqslant n^{c\sqrt{n}} a^n (n!)^{p-1} \prec \|C_n^{(2)}\|_p^p$, since $\Sigma_1 \asymp n^{1-1/p}$ and $1 - 1/p > \frac{1}{2}$. Estimating the third part similarly completes the proof of (2.1).

4.2. *The Case* $1 < p < 2$

As mentioned, this case is very similar to the one just discussed, and so we refer to [26] for the proof. For the statement of Theorem 2.1 we only mention that with

$$\delta = 1 - 1/p \in (0, \tfrac{1}{2}), \qquad \varepsilon = 1 - 2\delta \in (0,1), \quad \eta = \min(\delta, \varepsilon) \tag{4.17}$$

we have

$$\Sigma_2 = av^{\delta+\varepsilon} - av^{\delta} - \frac{a}{2p}v^{\varepsilon} + \frac{a(p-2)}{p} + \frac{a(3-p)}{6p^2}v^{\varepsilon-\delta} + O(v^{1-4\eta}). \tag{4.18}$$

4.3. The Case $p = 1$

The threshold value $p = 1$ turns out to be the simplest case, due to the fact that the $n \log n$-term in F vanishes. Namely, let m be as in Theorem 3.1, put $C_n = C_n^{(1)} + C_n^{(2)}$, the main term being

$$C_n^{(1)}(x) = C_n(x)\mathbf{1}\{x \in I\},$$

where

$$I = [0, 1 - mn^{-1/2}].$$

Aiming for $\|C_n^{(1)}\|_1$, we investigate the sum

$$S := n^{-1} \sum_{\beta \in n^{-1}\mathbf{Z} \cap I} g(\beta) e^{-nF(\beta)},$$

where

$$g(\beta) = a^n \sqrt{\frac{n}{2\pi}} \frac{e^{-a/(1-\beta)}}{\sqrt{\beta(1-\beta)}}$$

and

$$F(\beta) = \beta \log \beta + (1 - \beta) \log(1 - \beta);$$

cf. (3.2). This notation has the same significance, and hence *not* the same wording, as that of Section 4.1. Now, F is minimized at $\beta = \frac{1}{2}$. Euler–Maclaurin shows that $S \sim \int_J g(\beta) e^{-nF(\beta)}\, d\beta$, with

$$J = \{\beta;\ |\beta - \tfrac{1}{2}| \leqslant \sqrt{\lambda \log n / n}\},$$

λ a large constant. By the classical saddle point method [9],

$$S = g(\tfrac{1}{2}) e^{-F(1/2)} \sqrt{\frac{2\pi}{nF''(1/2)}} (1 + O(n^{-1})) = e^{-2a}(2a)^n (1 + O(n^{-1})).$$

Since the factor $e^{-\rho\psi(z)}$ in (3.2) is now insignificant, we may replace S by $\|C_n^{(1)}\|_1$.

As for the remainder, split $C_n^{(2)}$ into two parts. The first one, corresponding to $n - m\sqrt{n} < x < n$ is treated as in the end of Section 4.1. Following [25] we estimate the second one, say $C_n^{(2,2)}$, by Lyapounov's inequality, which, for a function f on a finite measure space with total mass A and $0 < p \leqslant q < \infty$, reads $\|f\|_p \leqslant A^{1/p - 1/q} \|f\|_q$. We take $q = 2$ and note that $C_n^{(2,2)}$ lives on the half-line $[n, \infty)$, having w-mass $A \leqslant ca^n/n!$.

Thus,

$$\|C_n^{(2,2)}\|_1 \leqslant A^{1-1/2}\|C_n\|_2 \leqslant ca^n \prec \|C_n^{(1)}\|_1$$

and the proof is complete.

4.4. The Case $0<p<1$

We finally discuss the case $0<p<1$. This is again similar to, but much simpler than $p>2$. Take $\gamma \in (0,1)$, and write $C_n = C_n^{(1)} + C_n^{(2)}$ with $C_n^{(1)}$ living on

$$\beta \in I := [n^{-1}, \gamma].$$

Adjusting the summand by a relative error of $O(\beta + 1/n\beta)$, $\|C_n^{(1)}\|_p^p$ goes into

$$S := \frac{a^{np}e^{-a}}{\sqrt{2\pi n}} \sum_{\beta \in n^{-1}\mathbf{Z} \cap I} e^{-F(\beta)}$$

with

$$F(\beta) = n(1-p)\beta \log \frac{n}{ae} + \frac{1}{2}\log \beta + n(\beta \log \beta + p(1-\beta)\log(1-\beta)),$$

cf. (3.2). As usual, F has a minimum at $\beta_0 \in I$, but $\beta_0 \to 0$ this time. Substituting $y = 1/\beta$ and putting, besides $v = n/a$,

$$\delta = 1-p, \qquad \varepsilon = 1-\delta = p, \qquad \eta = \min(\delta, \varepsilon), \tag{4.19}$$

the equation $F'(\beta) = 0$ takes the form

$$\log y - \frac{y}{2n} = U$$

with $U = \delta \log v - p \log(1-\beta)$. This equation has the implicit solution

$$y = 2nT\left(\frac{e^U}{2n}\right), \tag{4.20}$$

where $T(x) = -W(-x) = x + x^2 + \frac{3}{2}x^3 + \cdots$ and W is again the Lambert W function, cf. Section 4.1. (T is known as the tree function, being the generating function for the number of trees on n vertices [13].) Moreover,

$$\frac{e^U}{2n} = \frac{v^{-\varepsilon}}{2a}(1-y^{-1})^{-p}. \tag{4.21}$$

Proceeding as in Section 4.1, letting (4.20) and (4.21) feed each other, one readily computes the first few terms in the asymptotic expansion for y_0:

$$y_0 = v^\delta + p + \frac{1}{2a} v^{\delta-\varepsilon} + O(v^{\delta-2\eta}).$$

In particular, $\beta_0 = v^{-\delta} + O(v^{-\delta-\eta})$. This gives, recalling that $F'(\beta_0) = 0$,

$$-F(\beta_0) = n(1-p)\beta_0 - np\log(1-\beta_0) - \frac{1}{2}\log\beta_0 + \frac{1}{2} = \frac{\delta}{2}\log v + a + p\Sigma_3$$

with

$$\Sigma_3 = \frac{a}{p} v^\varepsilon - \frac{a}{2} v^{\varepsilon-\delta} - \frac{a}{p} + O(v^{\varepsilon-2\eta}) \tag{4.22}$$

and $F''(\beta_0) = av^{1+\delta}(1 + O(v^{-\eta}))$, see [26] for the computational details.

The argument then goes as in Section 4.1, and results in $S = a^{np} e^{p\Sigma_3}$, which is (2.3) with S in the place of $||C_n||_p^p$. As for the rest, we only mention that $C_n^{(2)}$ is most easily estimated by Lyapounov's inequality as in Section 4.3, but with $q = 1$, using the just proven L^1 result.

5. ENTROPIES: PROOF OF THEOREM 2.7

We turn to the information entropy

$$S_n(\hat{C}) = \sum_{x=0}^{\infty} \hat{C}_n(x)^2 \log \hat{C}_n(x)^2 w(x)$$

of the ortho*normal* Charlier polynomials $\hat{C}_n = C_n/\sqrt{a^n n!}$. Note that $S_n(\hat{C}) = 4\frac{d}{dp}||\hat{C}_n||_p$ evaluated at the threshold $p = 2$, but that this derivative cannot be calculated directly from Theorem 2.1, cf. Remark 2.9. Instead, we adopt a technique introduced by Aptekarev *et al.* [1]. For p close to 1, define

$$N_n(p) = \sum_{x=0}^{\infty} (\hat{C}_n(x)^2 w(x))^p.$$

Then

$$\begin{aligned} N_n'(1) &= \sum \hat{C}_n(x)^2 \log \hat{C}_n(x)^2 w(x) + \sum \hat{C}_n(x)^2 w(x) \log w(x) \\ &=: S_n(\hat{C}) + T_n(\hat{C}). \end{aligned}$$

It turns out that $T_n(\hat{C})$ is fairly simple to compute. Let us therefore start with $N_n'(1)$. We shall see that most of the contribution comes from the central region $|x-n|<2\sqrt{an}$, whence Theorem 4 of Goh [22] suits our needs. With the notation

$$x = n + a + \xi\sqrt{n}$$

the latter asserts that

$$\hat{C}_n(x) = \frac{\sqrt{n!}e^{\xi^2/4+a/2}}{a^{n/2}(an)^{1/4}\sqrt{\pi\sin\theta}}\left(\frac{n}{a}\right)^{(x-n)/2}(\cos\varphi + O(n^{-1/4}))$$

uniformly on

$$I_\varepsilon := \{\xi; |\xi| \leqslant 2\sqrt{a} - \varepsilon\}$$

for any $\varepsilon > 0$. Here

$$\cos\theta = \frac{\xi}{2\sqrt{a}}$$

and

$$\varphi = 2\sqrt{an}(\theta\cos\theta - \sin\theta) + a(\theta - \sin\theta\cos\theta) + \frac{\pi}{4}.$$

By abuse of notation, we sometimes consider I_ε as a set of corresponding values of x. We shall also write o_ε for little order as $\varepsilon \to 0$ (rather than $n \to \infty$). Now, restrict x to some I_ε. A straightforward computation shows that

$$\hat{C}_n(x)^2 w(x) = \frac{1}{\pi\sqrt{an}\sin\theta}(\cos^2\varphi + O(n^{-1/4})).$$

As an approximation of $N_n(p)$ we consider the sum

$$N_n(p;\varepsilon) = \left(\frac{1}{\pi\sqrt{an}}\right)^p \sum_{I_\varepsilon} \frac{|\cos\varphi|^{2p}}{|\sin\theta|^p},$$

where the sum is taken over those $\xi \in I_\varepsilon$ such that $x = n + a + \xi\sqrt{n}$ is an integer.

Expand $|\cos\varphi|^{2p}$ in a Fourier series $\sum_{m\in\mathbf{Z}} b_m e^{im\varphi}$. After differentiation, it follows from a theorem of Zygmund [34, VI. (3.6)] that $\sum_{m\in\mathbf{Z}} |m|^{1/2}|b_m|$ is bounded for p in a (complex) neighbourhood of 1. Moreover,

$$b_0 = \frac{1}{2\pi}\int_0^{2\pi} |\cos\varphi|^{2p}\,d\varphi = \pi^{-1}B(p+\tfrac{1}{2},\tfrac{1}{2}), \tag{5.1}$$

where B is the Beta function. Hence, $N_n(p;\varepsilon) = \sum_{m\in\mathbf{Z}} U_m$ with

$$U_m = b_m \left(\frac{1}{\pi\sqrt{an}}\right)^p \sum_{I_\varepsilon} \frac{e^{im\varphi}}{|\sin\theta|^p}.$$

The terms with $m \neq 0$ are small due to cancellations as we shall see. Summing over $\xi \in I_\varepsilon$ and using Euler–Maclaurin, the main term is found to be

$$U_0 = M_n(p)(1 + o_\varepsilon(1) + O(n^{-1/2}))$$

with

$$M_n(p) = \frac{2\sqrt{an}}{\pi}\left(\frac{1}{\pi\sqrt{an}}\right)^p B(p+\tfrac{1}{2},\tfrac{1}{2})B(1-\tfrac{p}{2},\tfrac{1}{2}).$$

We turn to the estimation of the remainder terms. Let us, for $m \neq 0$, write

$$U_m = b_m \left(\frac{1}{\pi\sqrt{an}}\right)^p \sum_{I_\varepsilon} f(\xi)e^{im\sqrt{n}g(\xi)}, \tag{5.2}$$

where

$$f(\xi) = \frac{1}{|\sin\theta|^p},$$

$$g(\xi) = 2\sqrt{a}(\theta\cos\theta - \sin\theta) + O(n^{-1/2}).$$

Put f and g equal to zero outside I_ε. If we for notational simplicity assume a to be integral, the sum in (5.2) can be written as

$$\sum_{\xi\in n^{-1/2}\mathbf{Z}} f(\xi)e^{im\sqrt{n}g(\xi)} = \sum_{k\in\mathbf{Z}} h(k) = \sum_{k\in\mathbf{Z}} \hat{h}(k),$$

where we have used Poisson's summation formula with

$$h(x) = f(n^{-1/2}x)e^{im\sqrt{n}g(n^{-1/2}x)}$$

and

$$\hat{h}(k) := \int_{\mathbf{R}} h(x)e^{-2\pi ikx}\,dx = \sqrt{n}\int_{\mathbf{R}} f(y)e^{i\sqrt{n}(mg(y)-2\pi ky)}\,dy.$$

Since $g'(y) = \theta + O(n^{-1/2})$, the phase above can have a stationary point only if $|k| \leqslant c|m|$, in which case the method of stationary phase [23] gives

$$|\hat{h}(k)| \leqslant \frac{c\sqrt{n}}{(|m|\sqrt{n})^{1/2}} = \frac{cn^{1/4}}{\sqrt{|m|}}.$$

For large k we can do better. Namely, if $|k| > c|m|$, then, by Hörmander's bound [23, Theorem 7.7.1],

$$(k\sqrt{n})^2 \left| \int_{\mathbf{R}} f(y) \exp\left(ik\sqrt{n}\left(\frac{m}{k}g(y) - 2\pi y\right)\right) dy \right| \leqslant c,$$

so that $|\hat{h}(k)| \leqslant c/(k^2\sqrt{n})$. Summing over k, this yields

$$|U_m| \leqslant c|b_m| \left(\frac{1}{\pi\sqrt{an}}\right)^p \left(|m|\frac{n^{1/4}}{\sqrt{|m|}} + \frac{1}{|m|\sqrt{n}}\right) \leqslant c|m|^{1/2}|b_m|n^{-1/4}U_0.$$

Hence, $\sum_{m \neq 0} |U_m| \leqslant cn^{-1/4}U_0$ and $N_n(p;\varepsilon) = U_0(1 + O(n^{-1/4}))$. Putting everything together,

$$\begin{aligned}\sum_{x \in I_\varepsilon} (\hat{C}_n(x)^2 w(x))^p &= N_n(p;\varepsilon)(1 + O(n^{-1/4})) \\ &= M_n(p)(1 + o_\varepsilon(1) + O(n^{-1/4})). \qquad (5.3)\end{aligned}$$

By an estimation similar to the ones in Section 4 one finds that the x outside I_ε contribute no more than $o_\varepsilon(1)$ times this. Thus, (5.3) holds with the leftmost term replaced by $N_n(p)$. It follows that

$$N_n(p) = M_n(p)(1 + o(1))$$

as $n \to \infty$.

Arguing much the same, it is not hard to show that $|N_n(p)| \leqslant c|M_n(p)|$ uniformly on compact subsets of the strip $\{p; 0 \leqslant \text{Re } p < 2\}$. But by Montel's classical theorem [29, Theorem 14.6], any uniformly bounded sequence of analytic functions that converges pointwise must in fact converge uniformly on compact sets. In particular, $N_n/M_n \to 1$ uniformly on a neighbourhood of $p = 1$, and so, by differentiation,

$$N_n'(1) = M_n(1)\left(\frac{N_n(1)M_n'(1)}{M_n(1)^2} + o(1)\right) = \log\frac{e}{2\pi\sqrt{an}} + o(1). \qquad (5.4)$$

We also need $T_n(\hat{C})$, which, as often in such contexts, is much simpler to compute. (See, however [16], which is devoted to this quantity in more difficult cases.) For Freud polynomials, T_n can even be calculated exactly [1].

Although this seems not to be the case here, we can easily find good approximations. Namely, since we need to only consider $x = n + O(\sqrt{n})$,

$$\log w(x) = -x\log\frac{n}{ae} - (x-n)\left(1+\frac{1}{2n}\right) - \frac{(x-n)^2}{2n} + \frac{(x-n)^3}{6n^2} - a - \frac{1}{2}\log 2\pi n + O(n^{-1}). \tag{5.5}$$

On the other hand, using the recurrence and orthogonality relations (1.2) and (1.3) it is straightforward to compute the following "central moments":

$$\sum_{x=0}^{\infty}(x-n)\hat{C}_n(x)^2 w(x) = a,$$

$$\sum_{x=0}^{\infty}(x-n)^2\hat{C}_n(x)^2 w(x) = 2an + a^2 + a,$$

$$\sum_{x=0}^{\infty}(x-n)^3\hat{C}_n(x)^2 w(x) = 6a^2 n + a^3 + 3a^2 + a.$$

This gives, together with (5.5),

$$-T_n(\hat{C}) = (n+a)\log\frac{n}{ae} + 3a + \frac{1}{2}\log 2\pi n + O(n^{-1}) \tag{5.6}$$

and (2.6) follows by adding (5.4) and (5.6).

ACKNOWLEDGMENTS

The author wishes to thank his advisor Svante Janson for valuable discussions, and Dunster for kindly sending him a preliminary version of the paper [21]. Thanks also to the referees for their careful reading of the manuscript and for a number of improving suggestions.

REFERENCES

1. A.I. Aptekarev, V.S. Buyarov, and J.S. Dehesa, Asymptotic behaviour of the L^p-norms and the entropy for general orthogonal polynomials (Russian), *Mat. Sb.* **185** (1994), 3–30 (Translation in *Russian Acad. Sci. Sb. Math.* **82** (1995), 373–395).

2. A.I. Aptekarev, V.S. Buyarov, J.S. Dehesa, and W. Van Assche, Asymptotics of entropy integrals for orthogonal polynomials (Russian), *Dokl. Akad. Nauk.* **346** (1996), 439–441.
3. A.I. Aptekarev, V.S. Buyarov, J.S. Dehesa, and R.J. Yáñez, Strong asymptotics of Laguerre polynomials and information entropies of two-dimensional harmonic oscillator and one-dimensional Coulomb potentials, *J. Math. Phys.* **39** (1998), 3050–3060.
4. A.I. Aptekarev, J.S. Dehesa, and R.J. Yáñez, Spatial entropy of central potentials and strong asymptotics of orthogonal polynomials, *J. Math. Phys.* **35** (1994), 4423–4428.
5. A.D. Barbour, L. Holst, and S. Janson, "Poisson Approximation," Oxford Univ. Press, Oxford, 1992.
6. J. Bergh and J. Löfström, "Interpolation Spaces," Springer, New York, 1976.
7. I. Białynicki-Birula and J. Mycielski, Uncertainty relations for information entropy in wave mechanics, *Comm. Math. Phys.* **44** (1975), 129–132.
8. W.G.C. Boyd and T.M. Dunster, Uniform asymptotic solutions of a class of second-order linear differential equations having a turning point and a regular singularity, with an application to Legendre functions, *SIAM J. Math. Anal.* **17** (1986), 422–450.
9. N.G. de Bruijn, "Asymptotic Methods in Analysis," North-Holland, Amsterdam, 1958.
10. V.S. Buyarov, J.S. Dehesa, A. Martínez-Finkelshtein, and E.B. Saff, Asymptotics of the information entropy for Jacobi and Laguerre polynomials with varying weights, *J. Approx. Theory* **99** (1999), 153–166.
11. V.S. Buyarov, P. López-Artés, A. Martínez-Finkelshtein, and W. Van Assche, Information entropy of Gegenbauer polynomials, *J. Phys. A* **33** (2000), 6549–6560.
12. C.V.L. Charlier, Über die Darstellung willkürlicher Funktionen, *Ark. Mat. Astron. och Fysik* **2** (1906), 35pp.
13. R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth, On the Lambert *W* function, *Adv. Comput. Math.* **5** (1996), 329–359.
14. J.S. Dehesa, R. González-Férez, W. Van Assche, and R.J. Yáñez, Entropic integrals of hyperspherical harmonics and spatial entropy of *D*-dimensional central potentials, *J. Math. Phys.* **40** (1999), 5675–5686.
15. J.S. Dehesa, A. Martínez-Finkelshtein, and J. Sánchez-Ruiz, Quantum information entropies and orthogonal polynomials, *J. Comput. Appl. Math.* **133** (2001), 23–46.
16. J.S. Dehesa and J. Sánchez-Ruiz, Entropic integrals of orthogonal hypergeometric polynomials with general supports, *J. Comput. Appl. Math.* **118** (2000), 311–322.
17. J.S. Dehesa, W. Van Assche, and R.J. Yáñez, Position and momentum information entropies of the *D*-dimensional harmonic oscillator and hydrogen atom, *Phys. Rev. A* **50** (1994), 3065–3079.
18. J.S. Dehesa, W. Van Assche, and R.J. Yáñez, Entropy of orthogonal polynomials with Freud weights and information entropies of the harmonic oscillator potential, *J. Math. Phys.* **36** (1995), 4106–4118.
19. J.S. Dehesa, W. Van Assche, and 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.
20. R.A. DeVore and G.G. Lorentz, "Constructive Approximation," Springer, New York, 1993.
21. T.M. Dunster, Uniform asymptotic expansions for Charlier polynomials, *J. Approx. Theory* **112** (2001), 93–133.
22. W.M. Goh, Plancherel–Rotach asymptotics for the Charlier polynomials, *Constr. Approx.* **14** (1998), 151–168.
23. L. Hörmander, "The Analysis of Linear Partial Differential Operators I," Springer, New York, 1990.
24. S. Janson, "Gaussian Hilbert Spaces," Cambridge Univ. Press, Cambridge, UK, 1997.

25. L. Larsson-Cohn, L^p-norms of Hermite polynomials and an extremal problem on Wiener chaos, *Ark. Mat.* **40** (2002). To appear.
26. L. Larsson-Cohn, L^p-norms and information entropies of Charlier polynomials, *U.U.D.M. Report* 2001:22, Uppsala, 2001.
27. M. Maejima and W. Van Assche, Probabilistic proofs of asymptotic formulas for some classical polynomials, *Math. Proc. Cambridge Philos. Soc.* **97** (1985), 499–510.
28. F.W.J. Olver, "Asymptotics and Special Functions," Academic Press, New York, 1974.
29. W. Rudin, "Real and Complex Analysis," McGraw-Hill, New York, 1987.
30. B. Rui and R. Wong, Uniform asymptotic expansions of Charlier polynomials, *Methods Appl. Anal.* **1** (1994), 294–313.
31. C.E. Shannon, A mathematical theory of communication, *Bell System Technol. J.* **27** (1948), 379–423, 623–656.
32. G. Szegő, "Orthogonal Polynomials," American Mathematical Society, New York, 1939.
33. W. Van Assche, "Asymptotics for Orthogonal Polynomials," Lecture Notes in Mathematics, Vol. 1265, Springer, New York, 1980.
34. A. Zygmund, "Trigonometric Series," Vol. I, Cambridge Univ. Press, Cambridge, UK, 1959.