Publ. Math. Debrecen
54 / 1-2 (1999), 103–129

Berry–Esséen-type inequalities for ultraspherical expansions

By MICHAEL VOIT (Tübingen)

Abstract. This paper contains several variants of Berry–Esséen-type inequalities for *ultraspherical* expansions of probability measures on $[0, \pi]$. Similar to the classical results on $\mathbb{R}$, proofs will be based in some cases on ultraspherical analogues of Fejér-kernels. The inequalities in this paper in particular lead to relations between the spherical-cap-distance of probability measures on unit spheres $S^d \subset \mathbb{R}^{d+1}$ and the norms of associated L^2-convolution operators. Moreover, the inequalities will be used to derive the order of convergence for some central limit theorems on $[0, \pi]$ and on S^d; the limit distributions there are analogues of Gaussian measures and the uniform distribution.

1. Introduction

The classical Berry–Esséen-inequality relates the $\|.\|_\infty$–distance of distribution functions on $\mathbb{R}$ with their Fourier transforms. There exist well-known extensions of this inequality to $\mathbb{R}^d$ for $d \geq 2$, the torus $T := \{z \in \mathbb{C} : |z| = 1\}$ and T^d; see [2], [7], [8], [13]. On the other hand, little seems to be done for similar estimations for other classical orthogonal expansions of distribution functions. For instance, one can ask for such estimates on $[0, \infty[$ in terms of Hankel transforms and their applications to radial probability measures on $\mathbb{R}^d$. Or, one can study expansions of measures on $[-1, 1]$ with respect to ultraspherical or Jacobi polynomials. This has applications to probability theory on spheres and projective spaces.

Mathematics Subject Classification: Primary 60E15, Secondary 60B15, 33C55, 26D05, 42C10, 43A62.
Key words and phrases: Berry–Esséen-inequality, ultraspherical polynomials, random walks on spheres, central limit theorem, uniform distribution on spheres, spherical cap distance.

Similarly, one can study Jacobi transforms of measures on $[0,\infty[$ with applications to radial distributions on hyperbolic spaces. The purpose of this paper is to encourage the study of Berry–Esséen-type estimates for such transforms and their applications. We however here restrict our attention to ultraspherical expansions.

We next describe the main results of this paper. Consider the ultraspherical polynomials $(R_n^{(\alpha)})_{n\geq 0}$ on $[-1,1]$ of index $\alpha \geq -1/2$ normalized by $R_n^{(\alpha)}(1)=1$ (this is the best normalization in probability theory and harmonic analysis). The ultraspherical coefficients of a probability measure μ on $[0,\pi]$ (of index α) then are defined by

$$\widehat{\mu}(n) := \int_0^\pi R_n^{(\alpha)}(\cos t)d\mu(t) \quad (n\geq 0).$$

The main result of this paper will be the following Berry–Esséen-type inequality for the distribution functions F_μ, F_ν of probability measures μ,ν on $[0,\pi]$ respectively:

1.1. Theorem. *For $\alpha \geq -1/2$ there is a universal constant $M_\alpha > 0$ as follows: If F_ν satisfies the Lipschitz condition*

$$|F_\nu(x) - F_\nu(x+h)| \leq mh \max_{z\in[x,x+h]} \sin^{2\alpha+1} z \quad \text{for } x,h\geq 0 \text{ with } x+h\leq \pi$$

with some constant $m>0$, then for all $N\geq 1$,

$$\|F_\mu - F_\nu\|_\infty \leq M_\alpha \cdot \left(\frac{m}{N} + \sum_{n=1}^N n^{\alpha-1/2}\cdot |\widehat{\mu}(n)-\widehat{\nu}(n)|\right).$$

Several variants of this result will be also discussed. We mention that the proof of Theorem 1.1 will be based on a smoothing procedure using ultraspherical analogues of Fejér kernels introduced by Lasser and Obermaier [11]. The second half of this paper will be devoted to applications of Theorem 1.1. The first two applications in Sections 5 and 6 concern the rate of convergence of two central limit theorems for certain random walks on $[0,\pi]$. In the first case, the limits are ultraspherical analogues of Gaussian distributions. In the second case, the limit is given by $d\omega_\alpha(t) := c_\alpha \cdot \sin^{2\alpha+1} t\, dt$ on $[0,\pi]$ with a suitable constant $c_\alpha > 0$. This measure is the orthogonality measure of the trigonometric ultraspherical polynomials $R_n^{(\alpha)}(\cos t)$ and forms the ultraspherical analogue of

the uniform distribution (which appears for $\alpha = -1/2$). The random walks mentioned above are homogeneous Markov chains on $[0,\pi]$ whose transition probabilities are defined in terms of the so-called ultraspherical convolution of measures on $[0,\pi]$. This convolution is discussed in [3], [18], [19]; for general random walks on hypergroups we refer to [4].

For $\alpha = d/2 - 1$ with $d \in \mathbb{N}$, $d \geq 2$, the ultraspherical random walks on $[0,\pi]$ mentioned above and their associated limit theorems admit an interpretation as projections of rotationally symmetric (or "isotropic") random walks on the unit sphere $S^d \subset \mathbb{R}^{d+1}$. In this way, the results of Sections 5 and 6 below lead to the following theorem:

1.2. Theorem. *Let $r \in [0,1/2]$, and let μ be a probability measure on $[0,\pi]$ with $\mu \neq \delta_0$. For $k \in \mathbb{N}$ define the contracted probability measure $\mu_{k,r} \in M^1([0,\pi])$ with*

$$\mu_{k,r}(A) := \mu(k^r A) \quad \text{for } A \subset [0,\pi] \text{ a Borel set.}$$

For each k, consider the isotropic random walk $(X_n^{k,r})_{n\geq 0}$ on S^d starting in some fixed North Pole x_0 of S^d such that in each step of time an angular jump with distribution $\mu_{k,r}$ appears, i.e., the random variables $\angle(X_n^{k,r}, X_{n+1}^{k,r})$ are independent and $\mu_{k,r}$-distributed. Then, after k steps of time, the random variable $\angle(X_k^{k,r}, x_0)$ (i.e., the angular distance of $X_k^{k,r}$ from the North pole) has some distribution $\mu_{k,r}^{(k)}$. The distribution functions of these distributions on $[0,\pi]$ have the following properties for $k \to \infty$:

(1) *If $r = 1/2$, then $\|F_{\mu_{k,r}^{(k)}} - F_\nu\|_\infty = O(1/k)$ for some specific ultraspherical Gaussian distribution ν on $[0,\pi]$ (ν is the angular part of some Gaussian measure on S^d centered at x_0).*

(2) *If $r \in [0,1/2[$, then $\|F_{\mu_{k,r}^{(k)}} - F_{\omega_{d/2-1}}\|_\infty = O(e^{-ck^{1-2r}})$ with some (known) constant $c > 0$ (notice that $\omega_{d/2-1}$ is the angular part of the uniform distribution on S^d).*

The final application of Theorem 1.1 in Section 7 concerns the spherical cap distance on S^d which is frequently used to measure how well points are distributed on S^d; see [5], [12], [15]. Theorem 1.1 relates this distance with the norm distance of convolution operators on L^2-spaces which can be usually handled much better. For the sphere S^2, we in particular recover results which are already contained in Lubotzky, Phillips, and

Sarnack [12]. These results were used there to study the order of convergence (with respect to the spherical cap distance) of a recursive algorithm of distributing points on S^2 uniformly.

This paper is organized as follows: In Section 2 we establish basic notations and facts on ultraspherical expansions; in Section 3 we collect some smoothing inequalities. The central Section 4 then contains several Berry–Esséen-type estimations. Sections 5–7 finally are devoted to the applications mentioned above.

Acknowledgement. Parts of this paper were written while the author was a visiting lecturer at the University of Virginia in Charlottesville. He would like to thank the Department of Mathematics there for its hospitality. Moreover, it is a pleasure to thank G. Pap and M. Rösler for some discussions, and K.-J. Förster for his hint to reference [9].

2. Ultraspherical expansions

The purpose of this section is to introduce some notations and facts about ultraspherical polynomials and expansions. Most results are more or less well-known.

2.1. Ultraspherical expansions. Consider the ultraspherical polynomials

$$R_n^{(\alpha)}(x) := {}_2F_1(-n, n+2\alpha+1; \alpha+1; (1-x)/2) \quad (x \in \mathbb{R},\, n \geq 0) \tag{2.1}$$

of index $\alpha \geq -1/2$ which are normalized by $R_n^{(\alpha)}(1) = 1$ and orthogonal on $[-1,1]$ with respect to the measure $(1-x^2)^\alpha dx$. The trigonometric polynomials $p_n^{(\alpha)}(t) := R_n^{(\alpha)}(\cos t)$ form an orthogonal basis of $L^2([0,\pi], \omega_\alpha)$ with respect to the probability measure

$$d\omega_\alpha(t) := c_\alpha \sin^{2\alpha+1} t\, dt \quad \text{with} \quad c_\alpha := \frac{\Gamma(2\alpha+2)}{\Gamma(\alpha+1)^2 2^{2\alpha+1}}. \tag{2.2}$$

Denote the Banach space of all Borel measures on $[0,\pi]$ by $M_b([0,\pi])$. Then, the ultraspherical expansion coefficients of $f \in L^1([0,\pi], \omega_\alpha)$ and $\mu \in M_b([0,\pi])$ are given by

$$\widehat{f}(n) := \int_0^\pi R_n^{(\alpha)}(\cos t) f(t) d\omega_\alpha(t) \quad \text{and} \quad \widehat{\mu}(n) := \int_0^\pi R_n^{(\alpha)}(\cos t) d\mu(t) \tag{2.3}$$

for $n \in \mathbb{Z}_+ := \{0, 1, 2, \dots\}$. We next define the ultraspherical Plancherel weights

$$h_n^{(\alpha)} := \left(\int_0^\pi R_n^{(\alpha)}(\cos t)^2 d\omega_\alpha(t) \right)^{-1} = \frac{(2n + 2\alpha + 1) \cdot (2\alpha + 1)_n}{(2\alpha + 1) \cdot n!} \quad (n \in \mathbb{Z}_+) \tag{2.4}$$

and introduce the associated weighted spaces $l^p(\mathbb{Z}_+, h)$. For $g \in l^1(\mathbb{Z}_+, h)$, its inverse ultraspherical transform $\check{g} \in C_b([0, \pi])$ is given by

$$\check{g}(x) := \sum_{k=0}^{\infty} h_k^{(\alpha)} g(k) R_k^{(\alpha)}(\cos x) \quad (x \in [0, \pi]). \tag{2.5}$$

The following facts are well-known (see [4] in setting of commutative hypergroups):

2.2. Facts.

(1) If $f \in L^1([0, \pi], \omega_\alpha)$ and $\mu \in M_b([0, \pi])$, and if $\|\mu\|$ denotes the total variation norm of μ, then $\|\widehat{f}\|_\infty \le \|f\|_1$ and $\|\widehat{\mu}\|_\infty \le \|\mu\|$.

(2) Inversion formula: If $f \in L^1([0, \pi], \omega_\alpha)$ with $\widehat{f} \in l^1(\mathbb{Z}_+, h)$, then $(\widehat{f})^\vee = f$ and $\|f\|_\infty = \|(\widehat{f})^\vee\|_\infty \le \|\widehat{f}\|_1$.

(3) Plancherel formula: If $f \in L^2([0, \pi], \omega_\alpha) \subset L^1([0, \pi], \omega_\alpha)$, then $\widehat{f} \in l^2(\mathbb{Z}_+, h)$ with $\|f\|_2 = \|\widehat{f}\|_2$.

We next give a collection of useful uniform estimations for ultraspherical polynomials.

2.3. Lemma. *Let $\alpha \ge -1/2$, $n \in \mathbb{Z}_+$, and $\theta \in [0, \pi]$. Then:*

(1) $|R_n^{(\alpha)}(\cos\theta)| \le 1$;

(2) $|\sin\theta \cdot R_n^{(\alpha+1)}(\cos\theta)| \le \dfrac{2(\alpha+1)}{\sqrt{(n+1)(n+2\alpha+2)}}$.

(3) $M_\alpha^1 := \sup\limits_{\varphi \in [0,\pi], n \ge 1} n^{\alpha+3/2} \left| \sin^{\alpha+3/2}\varphi \cdot R_{n-1}^{(\alpha+1)}(\cos\varphi) \right|$
$\le \left(\dfrac{\alpha+2}{2} \right)^{\alpha+1/2} \cdot \dfrac{(2\alpha+2)\Gamma(2\alpha+3)}{\Gamma(\alpha+3/2)}$.

(4) $M_\alpha^2 := \sup_{\varphi\in[0,\pi],n\geq 1} n^{\alpha+3/2}\left|\sin^{2\alpha+2}\varphi\cdot R_{n-1}^{(\alpha+1)}(\cos\varphi)\right|$
$\leq \left(\frac{\alpha+2}{2}\right)^{\alpha+1/2}\cdot\frac{\Gamma(2\alpha+3)}{\Gamma(\alpha+3/2)}.$

PROOF. For Part (1) we refer to Section 7.32 of SZEGÖ [16].

(2) We proceed as in Theorem 7.32.1 of [16]. In fact, the case $n=0$ is clear. For $n\geq 1$, consider

$$f_n(x) := R_n^{(\alpha)}(x)^2 + \frac{1-x^2}{n(n+2\alpha+1)}\Big(\frac{d}{dx}R_n^{(\alpha)}(x)\Big)^2.$$

The differential equation for Jacobi polynomials (Equation (4.2.1) of SZEGÖ [16]) yields that

$$f_n'(x) = \frac{2(2\alpha+1)x}{n(n+2\alpha+1)}\Big(\frac{d}{dx}R_n^{(\alpha)}(x)\Big)^2.$$

Hence, the nonnegative f_n attains its maximum on $[-1,1]$ at the boundary. As

$$\frac{d}{dx}R_n^{(\alpha)}(x) = \frac{n(n+2\alpha+1)}{2(\alpha+1)}R_{n-1}^{(\alpha+1)}(x) \tag{2.6}$$

(see (4.7.14) of SZEGÖ [16]) and $f_n(\pm 1)=1$, Part (2) follows.

(3) Corollary 1.8 of FÖRSTER [9] yields that for $\varphi\in[0,\pi]$, $n\geq 0$,

$$\begin{aligned}&\left|\sin^{\alpha+3/2}\varphi\cdot R_n^{(\alpha+1)}(\cos\varphi)\right|\\ &\leq (2\alpha+2)\frac{\Gamma(n/2+\alpha+3/2)}{\Gamma(\alpha+3/2)\Gamma(n/2+1)}\binom{n+2\alpha+2}{n}^{-1}.\end{aligned} \tag{2.7}$$

As $a^b\leq\Gamma(a+b)/\Gamma(a)\leq(a+b)^b$ for $a,b\geq 0$, Part (3) now follows readily.

(4) This follows from Eq. (19) of DURAND [6]. □

2.4. Ultraspherical convolution of measures on [0,π]. By Gegenbauer's product formula, the ultraspherical polynomials of index $\alpha>-1/2$ satisfy

$$\begin{aligned}&R_k^{(\alpha)}(\cos s)\cdot R_k^{(\alpha)}(\cos t)\\ &= c_{\alpha-1/2}\int_0^\pi R_k^{(\alpha)}(\cos s\cos t+\sin s\sin t\cos z)(\sin z)^{2\alpha}dz\end{aligned} \tag{2.8}$$

for $s, t \in [0, \pi]$ with $c_{\alpha-1/2}$ being given according to Eq. (2.2); see Eq. (2.23) of ASKEY [1]. For the limit case $\alpha = -1/2$, this formula degenerates into

$$\begin{aligned} &R_k^{(-1/2)}(\cos s) \cdot R_k^{(-1/2)}(\cos t) \\ &\quad = \frac{1}{2}\Big(R_k^{(-1/2)}(\cos|s-t|) + R_k^{(-1/2)}(\cos(\pi - |\pi - s - t|))\Big). \end{aligned} \tag{2.9}$$

For all $\alpha \geq -1/2$ and $s, t \in [0, \pi]$ we hence find unique probability measures, say $\delta_s *_\alpha \delta_t$ (or, for short, $\delta_s * \delta_t$) on $[0, \pi]$ with

$$R_k^{(\alpha)}(\cos s) \cdot R_k^{(\alpha)}(\cos t) = \int_0^\pi R_k^{(\alpha)}(\cos u) d(\delta_s * \delta_t)(u) \quad \text{for all } k \geq 0. \tag{2.10}$$

This ultraspherical convolution $\delta_s * \delta_t$ of point measures can be extended uniquely to a bilinear, weakly continuous, and probability preserving convolution $*$ on the Banach space $M_b([0, \pi])$ of all (complex) Borel measures on $[0, \pi]$. In particular, $(M_b([0, \pi]), *)$ is a commutative Banach algebra, and the convolution establishes a hypergroup structure on $[0, \pi]$. For details see [3], [4], [18], [19]. The product formulas (2.8) and (2.9) imply that the ultraspherical coefficients of the convolution product of $\mu, \nu \in M_b([0, \pi]$ satisfy

$$(\mu * \nu)^\wedge(n) = \widehat{\mu}(n) \cdot \widehat{\nu}(n) \quad \text{for all } n \geq 0. \tag{2.11}$$

3. Smoothing inequalities and Fejér kernels

The proof of the Berry–Esséen-type inequality 1.1 depends on a smoothing procedure using ultraspherical analogues of Fejér kernels. These kernels and their applications to smoothing will be discussed in this section. We always assume that $\alpha \geq -1/2$ holds.

3.1. Notations.

(1) The space of all (Borel) probability measures on $[0, \pi]$ is denoted by $M^1([0, \pi])$.

(2) The characteristic function of a set $A \subset \mathbb{R}$ is denoted by 1_A, and the support of a function f or a measure μ by $\operatorname{supp} f$ or $\operatorname{supp} \mu$ respectively.

(3) The distribution function F_μ of a probability measure $\mu \in M^1([0,\pi])$ is defined by

$$F_\mu(\theta) := \mu([0,\theta]) \quad \text{for } \theta \in \mathbb{R}.$$

We mention explicitly that this notation will be also used for $\theta \notin [0,\pi]$.

(4) We say that a distribution function F_μ satisfies the Lipschitz condition $L^\alpha(m)$ with some constant $m > 0$ if

$$|F_\mu(x) - F_\mu(x+h)| \le mh \max_{z\in[x,x+h]} \sin^{2\alpha+1} z$$

$$\text{for } x, h \ge 0 \quad \text{with } x + h \le \pi.$$

3.2. Remark. The Lipschitz condition 3.1(4) is an indispensible tool for the smoothing inequalities below. In our applications of Berry–Esséen-type inequalities below, this Lipschitz condition is always satisfied. In fact, if $\mu \in M^1([0,\pi])$ has a continuous ω_α-density f, then the Lipschitz condition holds with $m = c_\alpha \|f\|_\infty$. In particular, for $\mu = \omega_\alpha$ one has $m = c_\alpha$.

We next turn to some estimations for the ultraspherical convolution on $M^1([0,\pi])$.

3.3. Lemma. *Let $\mu, \nu, \varrho \in M^1([0,\pi])$, $t, x \in [0,\pi]$, and $T \ge 0$. Then,*

(1) $\operatorname{supp}\varrho \subset [0,T]$ *implies* $\mu * \varrho([0,t-T]) \le \mu([0,t]) \le \mu * \varrho([0,t+T])$, *and*

(2) $|\int_0^\pi (\delta_z * \delta_x)(1_{[0,t]}) d(\mu-\nu)(z)| \le 2\|F_\mu - F_\nu\|_\infty$.

Proof.

(1) As $\operatorname{supp}(\delta_x * \delta_y) \subset [|x-y|, \min(x+y, 2\pi-x-y)]$ for $x, y \in [0,\pi]$, the function

$$\varrho * 1_{[0,t+T]}(x) := \int_0^T (\delta_x * \delta_y)(1_{[0,t+T]}) d\varrho(y) \quad (x \in [0,\pi])$$

satisfies $0 \le \varrho * 1_{[0,t+T]} \le 1$ on $[0,\pi]$ with $\varrho * 1_{[0,t+T]}(x) = 1$ for $x \in [0,t]$. Hence,

$$\mu * \varrho([0,t+T]) = \int_0^\pi \int_0^T (\delta_x * \delta_y)(1_{[0,t+T]}) d\varrho(y) d\mu(x) \ge \mu([0,t]).$$

The second inequality can be checked in the same way.

(2) Assume first that $\alpha > -1/2$. Section 2.4 and Fubini's Theorem yield that

$$\Big|\int_0^\pi (\delta_z * \delta_x)(1_{[0,t]})d(\mu-\nu)(z)\Big| \tag{3.1}$$

$$\le c_{\alpha-1/2}\int_0^\pi \Big|\int_0^\pi 1_{[\cos t,1]}(\cos z\cos x+\sin z\sin x\cos w)d(\mu-\nu)(z)\Big| \cdot \sin^{2\alpha} w\,dw$$

As the function $z \mapsto \cos z\cos x+\sin z\sin x\cos w$ has at most one local extremum in $]0,\pi[$ for all $x, w \in [0,\pi]$, it follows that

$$\Big|\int_0^\pi 1_{[\cos t,1]}(\cos z\cos x+\sin z\sin x\cos w)d(\mu-\nu)(z)\Big| \le 2\|F_\mu - F_\nu\|_\infty$$

for $x, z \in [0,\pi]$. Part (2) now follows from (3.1) and the observation that $\int_0^\pi \sin^{2\alpha} w\,dw = c_{\alpha-1/2}^{-1}$. The limit case $\alpha = -1/2$ can be checked similarly. □

Lemma 3.3 leads to the following smoothing inequality which is motivated by [12].

3.4. Lemma. *Let $T > 0$, and $\mu, \nu, \varrho \in M^1([0,\pi])$ with $\operatorname{supp}\varrho \subset [0,T]$ such that F_ν has the Lipschitz property $L^\alpha(m)$ for some $m > 0$. Then, for all $\theta \in [0,\pi]$,*

$$|F_\mu(\theta) - F_\nu(\theta)| \le 2mT \cdot \max_{z\in[\theta-2T,\theta+2T]\cap[0,\pi]} \sin^{2\alpha+1} z + \max_{i\in\{\pm1\}} \big|F_{\mu*\varrho}(\theta+iT) - F_{\nu*\varrho}(\theta+iT)\big|.$$

Proof. For abbreviation, we put $R := \max_{z\in[\theta-2T,\theta+2T]\cap[0,\pi]} \sin^{2\alpha+1} z$. The Lipschitz condition $L^\alpha(m)$ and Lemma 3.3(1) imply that for $0 \le \theta \le \pi - 2T$,

$$\nu([0,\theta]) \ge \nu([0,\theta+2T]) - 2mTR \ge \nu*\varrho([0,\theta+T]) - 2mTR.$$

It is also clear from the Lipschitz condition $L^\alpha(m)$ that for $\theta \in [\pi-2T,\pi]$,

$$\nu([0,\theta]) \ge 1 - 2mTR \ge \nu*\varrho([0,\theta+T]) - 2mTR.$$

In summary, applying Lemma 3.3(1) also to μ, we see that for all $\theta \in [0,\pi]$,

$$(\mu-\nu)([0,\theta]) \le (\mu*\varrho-\nu*\varrho)([0,\theta+T]) + 2mTR.$$

In the same way, $(\mu-\nu)([0,\theta]) \ge (\mu*\varrho-\nu*\varrho)([0,\theta-T]) - 2mTR$. This completes the proof. □

We next turn to a smoothing inequality which involves ultraspherical Fejér kernels.

3.5. Ultraspherical Fejér kernels. Let $(R_n^{(\alpha,\beta)})_{n\ge 0}$ be the Jacobi polynomials with indices $\alpha,\beta \ge -1/2$ and normalization $R_n^{(\alpha,\beta)}(1)=1$. Following Lasser and Obermaier [11], we introduce the ultraspherical Fejér kernels as follows: Using Eq. (4.5.3) of [16] together with the normalization of the Jacobi polynomials above, we first observe that

$$\sum_{k=0}^{n} h_k^{(\alpha)} R_k^{(\alpha)} = \left(\sum_{k=0}^{n} h_k^{(\alpha)}\right) R_n^{(\alpha+1,\alpha)} = d_n^{(\alpha)} R_n^{(\alpha+1,\alpha)} \tag{3.2}$$

with

$$d_n^{(\alpha)} = \frac{\Gamma(n+2\alpha+2)\cdot(n+\alpha+1)}{\Gamma(2\alpha+2)\cdot(\alpha+1)\cdot n!} = \sum_{k=0}^{n} h_k^{(\alpha)}. \tag{3.3}$$

Hence,

$$\begin{aligned}\int_0^\pi R_n^{(\alpha+1,\alpha)}(\cos t)^2 d\omega_\alpha(t) &= \frac{1}{{d_n^{(\alpha)}}^2}\sum_{k=0}^{n}(h_k^{(\alpha)})^2 \int_0^\pi R_k^{(\alpha)}(\cos t)^2 d\omega_\alpha(t)\\ &= \frac{1}{{d_n^{(\alpha)}}^2}\sum_{k=0}^{n} h_k^{(\alpha)} = 1/d_n^{(\alpha)}. \end{aligned}\tag{3.4}$$

Therefore, for $n \ge 0$, the ultraspherical Fejér kernels

$$F_n^{(\alpha)}(t) := d_n^{(\alpha)}(R_n^{(\alpha+1,\alpha)}(\cos t))^2 \ge 0 \quad (n\ge 0,\, t\in[0,\pi]) \tag{3.5}$$

are the densities of the probability measures $\varrho_n^{(\alpha)} := F_n^{(\alpha)}\omega_\alpha$ on $[0,\pi]$ with $\widehat{\varrho_n^{(\alpha)}}(k) = 0$ for $k > 2n$ (for the latter we refer to [11]). It is also known that the probability measures $\varrho_n^{(\alpha)}$ tend weakly to the point measure δ_0 for $n\to\infty$ (this follows from results in [11] together with Lévy's continuity theorem for ultraspherical expansions; cf. Section 4.2 of [4]). We here need the following quantitative result for the smoothing inequality 3.7 below:

3.6. Lemma. *There is a constant $M_\alpha > 0$ such that for all $\epsilon \in]0, \pi/2]$ and $n \geq 1$,*

$$\varrho_n^{(\alpha)}([\epsilon, \pi]) \leq M_\alpha/\epsilon n.$$

Proof. In the following, $M_1, M_2, \ldots$ denote constants depending on α only. Eq. (7.32.6) of Szegö [16] and our normalization of the Jacobi polynomials imply that

$$|R_n^{(\alpha+1,\alpha)}(\cos\theta)| \cdot \theta^{\alpha+3/2} \leq M_1/n^{\alpha+3/2} \quad \text{for} \quad \theta \in [0, \pi/2],\ n \geq 1.$$

As $d_n^{(\alpha)} = O(n^{2\alpha+2})$, it follows that

$$\varrho_n^{(\alpha)}([\epsilon, \pi/2]) \leq \frac{M_2}{n} \int_\epsilon^{\pi/2} \frac{\sin^{2\alpha+1}\theta}{\theta^{2\alpha+3}} d\theta \leq \frac{M_3}{n\epsilon}. \tag{3.6}$$

Moreover, (7.32.6) and (4.1.3) of Szegö [16] imply that for $\theta \in [\pi/2, \pi]$ and $n \geq 1$,

$$|R_n^{(\alpha+1,\alpha)}(\cos\theta)| \cdot (\pi - \theta)^{\alpha+1/2} \leq M_4/n^{\alpha+3/2}.$$

Hence,

$$\varrho_n^{(\alpha)}([\pi/2, \pi]) \leq \frac{M_5}{n} \int_{\pi/2}^{\pi} \frac{\sin^{2\alpha+1}\theta}{(\pi-\theta)^{2\alpha+1}} d\theta \leq \frac{M_6}{n}. \tag{3.7}$$

The lemma is now a consequence of (3.6) and (3.7). □

3.7. Proposition. *For all $\alpha \geq -1/2$ there is a constant $R_\alpha > 0$ such that the Fejér measures $\varrho_n^{(\alpha)} \in M^1([0, \pi])$ $(n \geq 1)$ have the following property: If $\mu, \nu \in M^1([0, \pi])$ with F_ν satisfying the Lipschitz condition $L^\alpha(m)$ for $m > 0$, then*

$$\|F_\mu - F_\nu\|_\infty \leq 2\|F_{\mu * \varrho_n^{(\alpha)}} - F_{\nu * \varrho_n^{(\alpha)}}\|_\infty + R_\alpha m/n.$$

Proof. Fix $n \in \mathbb{N}$, and put $A := \|F_\mu - F_\nu\|_\infty$ and $A_n := \|F_{\mu * \varrho_n^{(\alpha)}} - F_{\nu * \varrho_n^{(\alpha)}}\|_\infty$. For each $\epsilon > 0$ we find $x_0 \in [0, \pi]$ with $|F_\mu(x_0) - F_\nu(x_0)| \geq A - \epsilon$. Assume now that $F_\mu(x_0) - F_\nu(x_0) \geq A - \epsilon$ holds (the case with the converse sign can be handled in the same way). Define $h := A/(4m)$ and $t := x_0 + h$.

Then for $x \in [0,h]$, we have $t - x = x_0 + h - x \geq x_0 \geq 0$. Hence, with Lemma 3.3(1),

$$\begin{aligned}\int_0^\pi (\delta_z * \delta_x)(1_{[0,t]})d(\mu-\nu)(z) &\geq F_\mu(t-x) - F_\nu(t+x)\\ &\geq F_\mu(x_0) - F_\nu(x_0) - (x+h)m\\ &\geq A - \epsilon - 2hm \geq A/2 - \epsilon.\end{aligned} \tag{3.8}$$

It follows from Lemma 3.6 and Lemma 3.3(2) that

$$\begin{aligned}((\mu-\nu)*\varrho_n^{(\alpha)})([0,t]) &= \int_0^h \int_0^\pi (\delta_z * \delta_x)(1_{[0,t]})d(\mu-\nu)(z)d\varrho_n^{(\alpha)}(x)\\ &\quad + \int_h^\pi \int_0^\pi (\delta_z * \delta_x)(1_{[0,t]})d(\mu-\nu)(z)d\varrho_n^{(\alpha)}(x)\\ &\geq (A/2-\epsilon) - \frac{M_\alpha}{hn} \cdot \max_{x\in[0,\pi]} \Big| \int_0^\pi (\delta_z * \delta_x)(1_{[0,t]})d(\mu-\nu)(z)\Big|\\ &\geq (A/2-\epsilon) - \frac{2M_\alpha A}{hn} = (A/2-\epsilon) - \frac{8M_\alpha m}{n}\end{aligned}$$

With $\epsilon \to 0$, it follows that $A/2 \leq A_n + 8M_\alpha m/n$ as claimed. □

4. Berry–Esséen-type inequalities

In this section we collect several versions of Berry–Esséen-type inequalities. The following theorem lists several versions without smoothing. Part (4) there seems to be the most natural and useful extension of the classical setting. As usual, we assume $\alpha \geq -1/2$.

4.1. Theorem. *Let $\theta \in [0,\pi]$ and $\mu,\nu \in M^1([0,\pi])$ such that the distribution function F_ν of ν satisfies the Lipschitz condition $L^\alpha(m)$ for some $m > 0$. Then, with the constants $c_\alpha, M_\alpha^1, M_\alpha^2 > 0$ of Section 2, the following estimations hold:*

(1) $|F_\mu(\theta) - F_\nu(\theta)| \leq \sin^{2\alpha+2}\theta \cdot \dfrac{c_\alpha}{2(\alpha+1)} \displaystyle\sum_{n=1}^\infty h_n^{(\alpha)} \cdot |\widehat{\mu}(n) - \widehat{\nu}(n)|;$

(2) $|F_\mu(\theta) - F_\nu(\theta)| \leq \sin^{2\alpha+1}\theta \cdot c_\alpha \displaystyle\sum_{n=1}^\infty \frac{h_n^{(\alpha)}}{\sqrt{n(n+2\alpha+1)}} \cdot |\widehat{\mu}(n) - \widehat{\nu}(n)|;$

(3) $|F_\mu(\theta) - F_\nu(\theta)| \le \sin^{\alpha+1/2}\theta \cdot \dfrac{c_\alpha \cdot M_\alpha^1}{2(\alpha+1)} \displaystyle\sum_{n=1}^{\infty} \frac{h_n^{(\alpha)}}{n^{\alpha+3/2}} \cdot |\widehat{\mu}(n) - \widehat{\nu}(n)|;$

(4) $|F_\mu(\theta) - F_\nu(\theta)| \le \dfrac{c_\alpha \cdot M_\alpha^2}{2(\alpha+1)} \displaystyle\sum_{n=1}^{\infty} \frac{h_n^{(\alpha)}}{n^{\alpha+3/2}} \cdot |\widehat{\mu}(n) - \widehat{\nu}(n)|.$

PROOF. Assume first that $\widehat{\mu}, \widehat{\nu} \in l^1(\mathbb{Z}_+, h)$ holds (notice that this is no essential restriction in Part (1)). In this case, the inversion formula in Section 2.2 ensures that

$$f_{\mu-\nu}(\varphi) := \sum_{n=1}^{\infty} h_n^{(\alpha)} R_n^{(\alpha)}(\cos\varphi) \cdot (\widehat{\mu}(n) - \widehat{\nu}(n))$$

is continuous on $[0,\pi]$ with $\mu - \nu = f_{\mu-\nu}\omega_\alpha$. As

$$\frac{d}{d\theta}\Big(\sin^{2\alpha+2}\theta \cdot R_{n-1}^{(\alpha+1)}(\cos\theta)\Big) = 2(\alpha+1)\sin^{2\alpha+1}\theta \cdot R_n^{(\alpha)}(\cos\theta) \tag{4.1}$$

(see the Rodrigues formula (4.7.12) of [16]), it follows that

$$\begin{aligned} F_\mu(\theta) - F_\nu(\theta) &= \mu([0,\theta]) - \nu([0,\theta]) = c_\alpha \int_0^\theta f_{\mu-\nu}(\varphi)\sin^{2\alpha+1}\varphi d\varphi \\ &= c_\alpha \int_0^\theta \sum_{n=1}^{\infty} h_n^{(\alpha)} R_n^{(\alpha)}(\cos\varphi) \cdot (\widehat{\mu}(n) - \widehat{\nu}(n))\sin^{2\alpha+1}\varphi d\varphi \\ &= c_\alpha \sum_{n=1}^{\infty} \Big(\int_0^\theta R_n^{(\alpha)}(\cos\varphi)\sin^{2\alpha+1}\varphi d\varphi\Big) h_n^{(\alpha)} \cdot (\widehat{\mu}(n) - \widehat{\nu}(n)) \\ &= \frac{c_\alpha}{2(\alpha+1)} \sum_{n=1}^{\infty} \sin^{2\alpha+2}\theta \cdot R_{n-1}^{(\alpha+1)}(\cos\theta) \cdot h_n^{(\alpha)} \cdot (\widehat{\mu}(n) - \widehat{\nu}(n)). \end{aligned} \tag{4.2}$$

The theorem under our additional assumption is now a consequence of Lemma 2.3. We next turn to the general case in (2)–(4); we first prove Part (2). For $n \ge 1$, choose $g_N \in L^1([0,\pi], \omega_\alpha)$ with $g_N \ge 0$, $\operatorname{supp} g_N \subset [0, 1/(2N)]$, and $\|g_N\|_1 = 1$. Then

$$g_N * g_N(\theta) := \int_0^\pi g_N(\varphi) \cdot (\delta_\varphi * \delta_\theta)(g_N) d\omega_\alpha(\varphi)$$

defines a function $g_N * g_N \in L^1([0,\pi],\omega_\alpha)$ with $g_N * g_N \geq 0$, $\mathrm{supp}(g_N * g_N) \subset [0, 1/N]$, $\|g_N * g_N\|_1 = 1$, and $(g_N * g_N)^\wedge = \widehat{g}_N^2 \in l^1(\mathbb{Z}_+, h)$. Hence, if $\varrho_N := (g_N * g_N)\omega_\alpha \in M^1([0,\pi])$, we may apply (2)–(5) to $\varrho_N * \mu$ and $\varrho_N * \nu$. In particular, Part (2) in the special case above and Lemma 3.4 imply that for $\theta \in [0,\pi]$ and $N \in \mathbb{N}$,

$$(\mu-\nu)([0,\theta]) \leq 2m/N + \max_{i\in\{\pm 1\}} \big(|F_{\varrho_N*\mu}(\theta + i/N) - F_{\varrho_N*\nu}(\theta + i/N)|\big)$$

$$\leq 2m/N + \max_{i\in\{\pm 1\}} (\sin^{2\alpha+1}(\theta + i/N)) \tag{4.3}$$

$$\times\, c_\alpha \sum_{n=1}^{\infty} \frac{h_n^{(\alpha)}}{\sqrt{n(n+2\alpha+1)}} \cdot |\widehat{g}_N(n)|^2 |\widehat{\mu}(n) - \widehat{\nu}(n)|$$

We may assume

$$R := \sum_{n=1}^{\infty} \frac{h_n^{(\alpha)}}{\sqrt{n(n+2\alpha+1)}} |\widehat{\mu}(n) - \widehat{\nu}(n)| < \infty.$$

As $|\widehat{g}_N(n)| \leq 1$ and $\widehat{g}_N(n) \to 1$ for $N \to \infty$ by the construction of g_N, the dominated convergence theorem ensures that the right hand side of (4.3) tends to $\sin^{2\alpha+1}(\theta) \cdot c_\alpha \cdot R$ which implies Part (2). Parts (3) and (4) follow in the same way. □

4.2. Remarks.

(1) Up to the precise constant, Theorem 4.1(1) can be also derived as follows: If $\|\widehat{\mu} - \widehat{\nu}\|_1 < \infty$ holds, then $\mu - \nu$ admits an ω_α-density $f_{\mu-\nu}$ with $\|f_{\mu-\nu}\|_\infty \leq \|\widehat{\mu} - \widehat{\nu}\|_1$. It follows for $\theta \in [0, \pi/2]$ that for some constant R_α,

$$|F_\mu(\theta) - F_\nu(\theta)| \leq \int_0^\theta |f_{\mu-\nu}| d\omega_\alpha \leq \|f_{\mu-\nu}\|_\infty \cdot c_\alpha \int_0^\theta \sin^{2\alpha+1}\varphi d\varphi$$
$$\leq R_\alpha \|f_{\mu-\nu}\|_\infty \sin^{2\alpha+2}\theta.$$

The same conclusion works also for $\theta \in [\pi/2, \pi]$ by taking $\int_\theta^\pi$ instead of $\int_0^\theta$.

(2) For the classical case $\alpha = -1/2$ (corresponding to symmetric probability measures on the torus), parts (2)–(4) of Theorem 4.1 are equivalent up to the precise constant.

Theorem 4.1(4) and the smoothing with the Fejér kernel lead to the following result:

4.3. Theorem. *Let $\mu, \nu \in M^1([0,\pi])$ such that F_ν has the Lipschitz property $L^\alpha(m)$ for some $m > 0$. Then there is a universal constant $M_\alpha > 0$ such that*

$$\|F_\mu - F_\nu\|_\infty \le M_\alpha \cdot \Big(\frac{m}{N} + \sum_{n=1}^{N} \frac{h_n^{(\alpha)}}{n^{\alpha+3/2}} \cdot |\widehat{\mu}(n) - \widehat{\nu}(n)|\Big) \quad \text{for all } \ N \ge 1.$$

Proof. Assume without loss of generality $N \ge 2$. Put $k := \lfloor N/2 \rfloor$. By Proposition 3.7, the Fejér measures $\varrho_k^{(\alpha)}$ satisfy

$$\|F_\mu - F_\nu\|_\infty \le 2\|F_{\mu * \varrho_k^{(\alpha)}} - F_{\nu * \varrho_k^{(\alpha)}}\|_\infty + R_\alpha m/k.$$

Now apply Theorem 4.1(4) to $\mu * \varrho_k^{(\alpha)}$ and $\nu * \varrho_k^{(\alpha)}$ (notice that $\nu * \varrho_k^{(\alpha)}$ satisfies the Lipschitz condition $L^\alpha(m)$ for some $m > 0$). As $\widehat{\varrho}_k^{(\alpha)}(n) = 0$ for $n > 2k$, and as

$$|(\mu * \varrho_k^{(\alpha)})^\wedge(n) - (\nu * \varrho_k^{(\alpha)})^\wedge(n)| = |\widehat{\mu}(n) - \widehat{\nu}(n)| \cdot \widehat{\varrho}_k^{(\alpha)}(n) \le |\widehat{\mu}(n) - \widehat{\nu}(n)|,$$

the theorem follows readily. □

4.4. Remark. Theorem 4.3 can be slightly improved by not estimating the Fejér weights $\widehat{\varrho}_N^{(\alpha)}(n)$ by 1. However it seems to be difficult to compute $\widehat{\varrho}_N^{(\alpha)}(n)$ sufficiently explicitly in order to obtain a considerable improvement of Theorem 4.3.

We next turn to a application of Theorem 4.1(3) and the smoothing in Lemma 3.4:

4.5. Theorem. *Let $\mu, \nu \in M^1([0,\pi])$ such that the distribution function F_ν of ν satisfies the Lipschitz condition $L^\alpha(m)$ for some $m > 0$. Then there is a universal constant $M_\alpha > 0$ such that for all $\theta, T \in [0,\pi]$,*

$$|F_\mu(\theta) - F_\nu(\theta)| \le 2mT \cdot w(\theta,T)^{2\alpha+1} + M_\alpha w(\theta,T)^{\alpha+1/2}$$
$$\times\Big(\sum_{n=1}^{\lfloor 1/T \rfloor} n^{\alpha-1/2}|\widehat{\mu}(n) - \widehat{\nu}(n)| + T^{-(\alpha+3/2)} \sum_{n=\lfloor 1/T \rfloor}^{\infty} \frac{|\widehat{\mu}(n) - \widehat{\nu}(n)|}{n^2}\Big)$$

with $w(\theta,T) := \max\limits_{z\in[\theta-T,\theta+T]\cap[0,\pi]} \sin z$.

PROOF. Consider the probability measure $\varrho_T^\alpha := c_T \cdot \omega_\alpha|_{[0,T]}$ with the constant

$$c_T := \Big(c_\alpha \int_0^T \sin^{2\alpha+1} t\, dt\Big)^{-1} = O(T^{-(2\alpha+2)}).$$

Eq. (4.1) and Lemma 2.3(3) show that for $n \geq 0$, $T > 0$ and suitable $M_1, M_2, M_3 > 0$,

$$\begin{aligned}|\widehat{\varrho}_T^\alpha(n)| &= c_T\Big|\int_0^T \sin^{2\alpha+1} t R_n^{(\alpha)}(\cos t)dt\Big| \\ &\leq \frac{M_1}{T^{2\alpha+2}} \cdot \big|\sin^{2\alpha+2} T \cdot R_{n-1}^{(\alpha+1)}(\cos T)\big| \\ &\leq \frac{M_2}{T^{2\alpha+2}} \cdot \frac{\sin^{\alpha+1/2} T}{n^{\alpha+3/2}} \leq \frac{M_3}{T^{\alpha+3/2} n^{\alpha+3/2}}.\end{aligned}$$

Therefore, for $n \geq 1$,

$$|(\mu * \varrho_T^\alpha)^\wedge(n) - (\nu * \varrho_T^\alpha)^\wedge(n)| \leq M_3 |\widehat{\mu}(n) - \widehat{\nu}(n)| \cdot \min\Big(1, (Tn)^{-(\alpha+3/2)}\Big).$$

A combination of Theorem 4.1(3) and Lemma 3.4 now leads to the claim. □

4.6. Corollary. *Let* $(\mu_N)_{N\geq 1}$ *be a sequence in* $M^1([0,\pi])$, *and let* $\nu \in M^1([0,\pi])$ *such that* F_ν *satisfies the Lipschitz condition* $L^\alpha(m)$ *for some* $m > 0$. *If*

$$\sup_{n\in\mathbb{N}} |\widehat{\mu}_N(n) - \widehat{\nu}(n)| = O(N^{-A}) \quad \text{for } N \to \infty$$

holds for some $A > 0$, *then for all constants* $B > 0$ *and* $r \in [0, A/(2\alpha+2)]$,

$$\sup_{\theta\in[0,BN^{-r}]} |F_{\mu_N}(\theta) - F_\nu(\theta)| = O\Big(\frac{1}{N^{[A+r(2\alpha+2)(\alpha+1/2)]/(\alpha+3/2)}}\Big)$$

$$\text{for } N \to \infty.$$

In particular, for $r = 0$ *and* $r = A/(2\alpha+2)$ *respectively:*

(1) $\|F_{\mu_N} - F_\nu\|_\infty = O(N^{-A/(\alpha+3/2)})$ *for* $N \to \infty$;

(2) *For each constant* $B > 0$,

$$\sup_{\theta\in[0,BN^{-A/(2\alpha+2)}]} |F_{\mu_N}(\theta) - F_\nu(\theta)| = O(N^{-A}) \quad \textit{for } N \to \infty.$$

PROOF. This follows readily from Theorem 4.5. In fact, if $r \in [0, A/(2\alpha+2)]$, then

$$s := [A - (\alpha + 1/2)r]/(\alpha + 3/2) \geq r$$

holds. Now take $T = T(N) = N^{-s}$ in Theorem 4.5 and observe that $w(\theta, T(N)) = O(N^{-r})$ holds uniformly for $\theta \in [0, BN^{-r}]$. This leads to the claim. □

5. Application to a central limit theorem

In this section we apply Theorem 4.3 to the rate of convergence in a central limit theorem concerning ultraspherical convolutions. For an interpretation of this convolution on $M^1([0,\pi])$ of index α in terms of radial random walks on the unit spheres $S^n \subset \mathbb{R}^n$ with $\alpha = n/2 - 1$ we refer to [3], [19] and references there. We need some preparations.

5.1. Gaussian measures. Let $\alpha \geq -1/2$ and define the function

$$q(n) := q^{(\alpha)}(n) := -\left(\frac{d}{d\theta}\right)^2 R_n^{(\alpha)}(\cos\theta)\Big|_{\theta=0} = \frac{n(n+2\alpha+1)}{2(\alpha+1)} \quad (n \geq 0). \tag{5.1}$$

Then for $\sigma^2 > 0$, the heat kernel

$$h_{\sigma^2}^{(\alpha)}(\theta) := \sum_{n=0}^{\infty} h_n^{(\alpha)} e^{-\sigma^2 q(n)/2} R_n^{(\alpha)}(\cos\theta) \quad (\theta, \varphi \in [0,\pi]) \tag{5.2}$$

is a positive continuous function on $[0,\pi]$. The probability measure

$$d\nu_{\sigma^2}^{\alpha}(\theta) := h_{\sigma^2}^{(\alpha)}(\theta) \cdot d\omega_\alpha(\theta) \tag{5.3}$$

on $[0,\pi]$ is called the Gaussian measure with "variance" σ^2.

5.2. Central limit theorem. *For $\mu \in M^1([0,\pi])$ define its "variance"*

$$\sigma^2 := \int_0^\pi x^2 d\mu(x) \geq 0 \tag{5.4}$$

which is consistent with the notation in 5.1 *for Gaussian measures. For $k \in \mathbb{N}$, consider the contracted probability measure $\mu_k \in M^1([0,\pi])$ with*

$$\mu_k(A) := \mu(\sqrt{k}A) \quad \text{for } A \subset [0,\pi] \quad \text{a Borel set.} \tag{5.5}$$

It is well-known (see Voit [17]) that for $\alpha \geq -1/2$ the ultraspherical convolution powers

$$\mu_k^{(k)} := \mu_k * \mu_k * \ldots * \mu_k \in M^1([0,\pi])$$

tend weakly to the Gaussian measure $\nu_{\sigma^2}^\alpha$ for $k \to \infty$. Moreover, under some additional conditions, the following strong convergence result is known; see [19]:

5.3. Theorem. *Let $\alpha > -1/2$ and $\mu \in M^1([0,\pi])$. Assume there exist constants $c, p > 0$ such that $\mu([0,\epsilon]) \leq c \cdot \epsilon^p$ for all $\epsilon \in [0,\pi]$ (which means that "μ is not concentrated at 0 too much"). Then there exists $k_0 = k_0(\alpha,\mu)$ such that for each $k \geq k_0$, the measure $\mu_k^{(k)}$ has a continuous, bounded $\omega^{(\alpha)}$-density f_k. Moreover,*

$$\|f_k - h_{\sigma^2}^{(\alpha)}(\,.\,,0)\|_\infty = O(1/k) \quad \text{and} \quad \|\mu_k^{(k)} - \nu_{\sigma^2}^\alpha\| = O(1/k) \quad \text{for } k \to \infty.$$

We now show that the results of Section 4 lead to the weaker convergence result $\|F_{\mu_k^{(k)}} - F_{\nu_{\sigma^2}^\alpha}\|_\infty = O(1/k)$ where here the non-concentration condition is not needed:

5.4. Theorem. *For all $\alpha > -1/2$ and $\mu \in M^1([0,\pi])$ with $\mu \neq \delta_0$,*

$$\|F_{\mu_k^{(k)}} - F_{\nu_{\sigma^2}^\alpha}\|_\infty = O(1/k) \quad \text{for } k \to \infty.$$

The proof of Theorem 5.4 relies on the following asymptotic estimations (see also [19]):

5.5. Lemma. *The following results hold for $k \to \infty$:*

(1) *There exists a constant $A = A(\mu, \alpha) > 0$ such that*

$$\sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} h_n^{(\alpha)} \Big| (\mu_k^{(k)})^\wedge(n) - (\nu_{\sigma^2}^{\alpha})^\wedge(n) \Big| = O(1/k).$$

(2) *For each $A > 0$ there exists a constant $C_1 = C_1(\mu, \alpha) \in]0,1[$ such that*

$$\sum_{n=\lfloor A\sqrt{k}\rfloor}^{\infty} h_n^{(\alpha)} |(\nu_{\sigma^2}^{\alpha})^\wedge(n)| = O(C_1^k).$$

(3) *For all constants $0 < A < B$, there is a constant $C_2 = C_2(\mu, \alpha) \in]0,1[$ with*

$$\sum_{n=\lfloor A\sqrt{k}\rfloor}^{\lfloor B\sqrt{k}\rfloor} h_n^{(\alpha)} |(\mu_k^{(k)})^\wedge(n)| = O(C_2^k).$$

(4) *There exist constants $B = B(\mu, \alpha) > A$ and $C_3 = C_3(\mu, \alpha) \in]0,1[$ with*

$$\sum_{n=\lfloor B\sqrt{k}\rfloor}^{k} h_n^{(\alpha)} |(\mu_k^{(k)})^\wedge(n)| = O(C_3^k).$$

PROOF. For Part (1) we refer to Eq.(3.9) and Lemma 3.4 in VOIT [19]; for Part (2) see Eq. (3.10) in [19]. Moreover, Part (3) is shown on p. 474 of [19]. Notice that for these results the non-concentration condition of Theorem 5.3 is not needed there.

It remains to prove Part (4): As $\mu \neq \delta_0$, there exists $a > 0$ with $c := \mu([0,a]) < 1$. Using $|P_n^{(\alpha)}(x)| \leq 1$ for $n \geq 0$, $x \in [-1,1]$ and $|P_n^{(\alpha)}(\cos t)| = O((tn)^{-(\alpha+1/2)})$ for $t \in [0, \pi/2]$, $n \geq 0$ (see Section 7.32 of SZEGÖ [16]), we find a constant $S > 0$ such that for all $n \geq 1$ and k sufficiently large,

$$\begin{aligned} |(\mu_k^{(k)})^\wedge(n)| &= \Big| \Big(\int_0^a + \int_a^\pi \Big) P_n^{(\alpha)}(\cos(t/\sqrt{k})) d\mu(t) \Big|^k \\ &\leq \left(c + (1-c) S \cdot \frac{\sqrt{k}^{\alpha+1/2}}{n^{\alpha+1/2}} \right)^k . \end{aligned} \tag{5.6}$$

Now choose $B > A$ with $c + (1-c)S/B =: T < 1$. Then for all $n \geq B\sqrt{k}$, we obtain $|(\mu_k^{(k)})^\wedge(n)| \leq T^k$, and thus, for some $C_3 \in]0,1[$,

$$\sum_{n=\lfloor B\sqrt{k}\rfloor}^{k} h_n^{(\alpha)} |(\mu_k^{(k)})^\wedge(n)| = O\Big(\sum_{n=\lfloor B\sqrt{k}\rfloor}^{k} n^{2\alpha+1} T^k \Big) \leq O(C_3^k). \tag{5.7}$$

□

PROOF of Theorem 5.4. Lemma 5.5 implies that for $k \to \infty$,

$$\sum_{n=0}^{k} h_n^{(\alpha)} \Big| (\mu_k^{(k)})^\wedge(n) - (\nu_{\sigma^2}^{\alpha})^\wedge(n) \Big| = O(1/k).$$

Theorem 5.4 is now an immediate consequence of Theorem 4.3 with $N := k$. □

5.6. Remarks.

(1) Theorem 5.4 is not valid for $\alpha = -1/2$. In this case one has the weaker sharp order

$$\| F_{\mu_k^{(k)}} - F_{\nu_{\sigma^2}^{\alpha}} \|_\infty = O(1/\sqrt{k}) \quad \text{for } k \to \infty.$$

In fact, for any $t \in]0,\pi[$, the convolution product $\delta_t^{(k)}$ has a mass of order at least $O(1/\sqrt{k})$ at 0 (this follows easily from the corresponding well-known result for centered Binomial distributions on $\mathbb{R}$).

(2) Obviously, Theorem 5.4 implies Theorem 1.2(1). We also mention that for $\alpha \geq 0$, the families $(\mu_k^{(n)})_{n\geq 0} \subset M^1([0,\pi])$ of convolution powers admit interpretations as distributions of angular parts of certain random walks on the sphere S^2 hving a "compass"; for details see BINGHAM [3].

(3) By PAP and VOIT [14], there exists a short Edgeworth expansion associated with Theorem 5.3 under the assumptions there. A short look into [14] and Lemma 5.5 yield that a short Edgeworth expansion also exists in the setting of Theorem 5.4.

6. Applications to the convergence to the uniform distribution

In this section we prove an analogue of Theorem 5.4 where here, after a different normalization, the limit distribution is the measure ω_α.

6.1. A limit theorem with the uniform distribution as limit. Let $\alpha \geq -1/2$, $r \in [0, 1/2[$ and $\mu \in M^1([0,\pi])$ with $\mu \neq \delta_0$. For $k \in \mathbb{N}$, consider the concentrated probability measure $\mu_{k,r} \in M^1([0,\pi])$ with

$$\mu_{k,r}(A) := \mu(k^r A) \quad \text{for } A \subset [0,\pi] \quad \text{a Borel set.}$$

By Voit [17], the ultraspherical convolution powers $\mu_{k,r}^{(k)}$ tend weakly to the uniform distribution ω_α for $k \to \infty$. This generalizes the well-known fact for $r = 0$ that $\mu^{(k)}$ tends weakly to ω_α. We now use Theorem 4.3 to derive the following rate of convergence:

6.2. Theorem. *Let $\alpha > -1/2$, $r \in]0, 1/2[$ and $\mu \in M^1([0,\pi])$ with $\mu \neq \delta_0$. Define $\sigma^2 := \int_0^\pi x^2 d\mu(x) > 0$. Then for each $\epsilon > 0$,*

$$\|F_{\mu_{k,r}^{(k)}} - F_{\omega_\alpha}\|_\infty = O\Big(\exp\Big(-\Big(\frac{\sigma^2}{4(\alpha+1)} - \epsilon\Big) \cdot k^{1-2r}\Big)\Big) \quad \text{for } k \to \infty.$$

The proof of this result will be based on the following estimations:

6.3. Lemma. *The following results hold for $k \to \infty$ in the setting of Theorem 6.2:*

(1) *There exists a constant $A = A(\mu, \alpha, r) > 0$ with*

$$\sum_{n=1}^{\lfloor Ak^r \rfloor} h_n^{(\alpha)} \big|(\mu_{k,r}^{(k)})^\wedge(n)\big| = O\Big(\exp\Big(-\Big(\frac{\sigma^2}{4(\alpha+1)} - \epsilon\Big) \cdot k^{1-2r}\Big)\Big).$$

(2) *For all constants $0 < A < B$, there is a constant $E \in]0,1[$ with*

$$\sum_{n=\lfloor Ak^r \rfloor}^{\lfloor Bk^r \rfloor} h_n^{(\alpha)} |(\mu_{k,r}^{(k)})^\wedge(n)| = O(E^k).$$

(3) *There exist constants $B = B(\mu, \alpha, r) > A$ and $D = D(\mu, \alpha, r) > 1$ with*

$$\sum_{n=\lfloor Bk^r \rfloor}^{\lfloor D^k \rfloor} h_n^{(\alpha)} |(\mu_{k,r}^{(k)})^\wedge(n)| = O(D^{-k}).$$

PROOF. In the following, let $C_1, C_2, \dots$ be constants depending on α, r, μ only.

(1) Lemma 3.2 of [19] ensures that for $n \geq 0$ and $x \in [-1,1]$,

$$\Big| R_n^{(\alpha)}(x) - \Big(1 - (1-x)\frac{(n+2\alpha+1)n}{2(\alpha+1)}\Big)\Big| \leq C_1 n^4 (1-x^2)$$

It follows readily (cf. the proof of Lemma 3.3 in [19]) that

$$|\widehat{\mu}_{k,r}(n)| = \Big|\int_0^\pi R_n^{(\alpha)}(\cos(t/k^r))d\mu_{k,r}(t)\Big| \leq \Big|1 - \frac{q(n)\sigma^2}{2k^{2r}} + \frac{C_2 n^4}{k^{4r}}\Big|$$
$$\leq \Big|1 - \frac{n^2\sigma^2}{4(\alpha+1)k^{2r}} + \frac{C_2 n^4}{k^{4r}}\Big|$$

for $n \leq C_3 k^r$ and C_2, C_3 suitable where $q(n)$ is defined in (5.1). Now choose $A > 0$ with $A \leq C_3$ and $C_2 A \leq \epsilon/2$. Then, for $1 \leq n \leq Ak^r$,

$$|\widehat{\mu}_{k,r}(n)|^k \leq \Big|1 - \Big(\frac{\sigma^2}{4(\alpha+1)} - \frac{\epsilon}{2}\Big)\cdot\frac{n^2}{k^{2r}}\Big|^k \leq \exp\Big(-\Big(\frac{\sigma^2}{4(\alpha+1)} - \frac{\epsilon}{2}\Big)k^{1-2r}\Big).$$

As $h_n^{(\alpha)} = O(n^{2\alpha+1})$, Part (1) follows readily.

(2) This can be shown in the same way as the corresponding result for $r = 1/2$ on p. 474 of [19]. We omit the lenghty details here and mention only that the proof uses Hilb's formula for ultraspherical polynomials (see Theorem 8.21.12 of SZEGÖ [16]).

(3) Similar as in the proof of (5.6) above, we have

$$|(\mu_{k,r}^{(k)})^\wedge(n)| \leq \Big(c + (1-c)C_4\frac{k^{r(\alpha+1/2)}}{n^{\alpha+1/2}}\Big)^k$$

for all $n \geq 1$ and k sufficiently large with $c \in]0,1[$. The proof can now be completed in the same way as in Lemma 5.5(4) for $r = 1/2$. □

PROOF of Theorem 6.2. Lemma 6.3 implies that for $k \to \infty$ and suitable $D > 1$,

$$\sum_{n=1}^{\lfloor D^k \rfloor} h_n^{(\alpha)} |(\mu_{k,r}^{(k)})^\wedge(n)| = O\Big(\exp\Big(-\Big(\frac{\sigma^2}{4(\alpha+1)} - \epsilon\Big)\cdot k^{1-2r}\Big)\Big).$$

Theorem 6.2 is now a consequence of Theorem 4.3 by choosing $N := \lfloor D^k \rfloor$ there. □

6.4. Remarks.

(1) For $r = 0$ and $\alpha > -1/2$, the methods above lead to a constant $C \in]0,1[$ with

$$\|F_{\mu_{k,r}^{(k)}} - F_{\omega_\alpha}\|_\infty = O(C^k) \quad \text{for } k \to \infty.$$

(2) Similar to Theorem 5.3, Theorem 6.2 and the first part of this remark are not valid for $\alpha = -1/2$. In fact, in this case, one usually has the order $O(1/\sqrt{k})$ only.

(3) In Voit [18], the order of convergence of the heat kernels $h_t^{(d)}$ on the spheres S^d to 1 is investigated when the time t and the dimension d tend to ∞ in certain coupled ways. This is done in [18] with respect to the $\|\,.\,\|_\infty$- and the $\|\,.\,\|_1$-norm. It might be interesting to explore whether the preceding results lead to similar convergence results with respect to the uniform convergence of distribution functions.

7. The spherical cap distance and distributing points on spheres

In this section we translate the Berry–Esséen-inequalities of Section 4 to the setting of probability measures on unit spheres. This leads to important applications.

7.1. The spherical cap distance. The spherical cap distance of probability measures μ, ν on the unit sphere $S^d \subset \mathbb{R}^{n+1}$ for $d \geq 2$ is given by

$$D_c(\mu,\nu) := \sup_J |\mu(J) - \nu(J)| \tag{7.1}$$

where J runs over all spherical caps $D(x,s) := \{y \in S^d : \angle(x,y) \leq s\}$ with $x \in S^d$, $s \in [0,\pi]$. This spherical cap distance is related to the uniform distance of distribution functions on $[0,\pi]$ as follows: For $x \in S^d$ consider the associated projection

$$p_x : S^d \to [0,\pi], \quad y \mapsto \angle(x,y); \tag{7.2}$$

now extend p_x to a projection of measures on S^d which is again denoted by p_x:

$$p_x : M^1(S^d) \longmapsto M^1([0,\pi]) \quad \text{with } p_x(\mu)(A) := \mu(p_x^{-1}(A)) \quad \text{for } A \subset [0,\pi].$$

Then, we have

$$D_c(\mu,\nu) = \sup_{x\in S^d} \|F_{p_x(\mu)} - F_{p_x(\nu)}\|_\infty. \tag{7.3}$$

There exists also a connection between L^2-convolution operators of measures on S^d and the ultraspherical coefficients (of index $\alpha = d/2-1$) of their projections:

7.2. The norm of L^2-convolution operators. For $d \geq 2$, the group $SO(d+1)$ acts on S^d in the natural way. The stabilizer $H_x \subset SO(d+1)$ of a "North pole" $x \in S^d$ is isomorphic with $SO(d)$. Identify

$$S^d \simeq SO(d+1)/H_x = \{gH_x : g \in SO(d+1)\}$$

and $[0,\pi] \simeq SO(d+1)//H_x = \{H_x g H_x : g \in SO(d+1)\}$ in the natural way such that the canonical projection $p_x : SO(d+1)/H_x \to SO(d+1)//H_x$ corresponds to the projection defined in Eq. (7.2). Moreover, let q_x be the canonical projection from $SO(d+1)$ to S^d.

For any Borel measure $\mu \in M_b(SO(d+1))$, consider the convolution operator

$$T_\mu f(z) := \int_{SO(d+1)} f(g(z))d\mu(g) \quad \text{on} \quad L^2(S^d, h_d)$$

where h_d is the uniform distribution on S^d. If one identifies the space of all H_x-invariant $L^2(S^d,h_d)$-functions with $L^2([0,\pi],\omega_{d/2-1})$, then the operators T_μ on $L^2(S^d,h_d)$ correspond to convolution operators $\tilde{T}_{p_x\circ q_x(\mu)}$ on $L^2([0,\pi],\omega_{d/2-1})$ given by

$$\tilde{T}_\varrho f(t) := \int_0^\pi (\delta_s *_{d/2-1} \delta_t)(f) d\varrho(s)$$

$$\text{for } \varrho \in M_b([0,\pi]),\ f \in L^2([0,\pi],\omega_{d/2-1}).$$

Hence, for all $x \in S^d$ and $\mu \in M_b(SO(d+1))$,

$$\|\tilde{T}_{p_x\circ q_x(\mu)}\|_{L^2([0,\pi],\omega_{d/2-1})} \leq \|T_\mu\|_{L^2(S^d,h_d)}. \tag{7.4}$$

Moreover, spectral theory for the commutative Banach-$*$-algebra $L^1([0,\pi],\omega_{d/2-1})$ (see, for instance, 2.2.4(v) of [4]) yields that

$$\|\tilde{T}_{p_x\circ q_x(\varrho)}\|_{L^2([0,\pi],\omega_{d/2-1})} = \sup_{n\geq 0} |\widehat{\varrho}(n)| \quad \text{for all } \varrho \in M_b([0,\pi]) \tag{7.5}$$

In summary, for all probability measures $\mu, \nu \in M^1(SO(d+1))$,

$$\sup_{x\in S^d, n\geq 0} |(p_x \circ q_x(\mu))^\wedge(n) - (p_x \circ q_x(\nu))^\wedge(n)| \leq \|T_{\mu-\nu}\|_{L^2(S^d,h_d)}. \tag{7.6}$$

Now one can combine (7.3) and (7.6) with the inequalities of Section 4 to obtain corresponding results for spherical cap distances. For instance, Corollary 4.6 leads to:

7.3. Corollary. *Let $(\mu_N)_{N\geq 1}$ be a sequence in $M^1(SO(d+1))$, and let $\nu \in M^1(SO(d+1))$ such that for all $x \in S^d$, the distribution function $F_{p_x\circ q_x(\nu)}$ satisfies the Lipschitz condition $L^{d/2-1}(m)$ for some $m > 0$. If*

$$\|T_{\mu_N} - T_\nu\| = O(N^{-A}) \quad \text{for } N \to \infty$$

and some $A > 0$, then for all constants $B > 0$ and $r \in [0, A/(2\alpha+2)]$,

$$\sup_{\theta\in[0,BN^{-r}]} |\mu_N(D(x,s)) - \nu(D(x,s))| = O\Big(\frac{1}{N^{[2A+rd(d-1)]/(d+1)}}\Big)$$

$$\text{for } N \to \infty.$$

In particular, for $r = 0$ and $r = A/d$ respectively:

(1) $D_c(\mu_N, \nu) = O(N^{-2A/(d+1)})$;

(2) *For all $B > 0$,* $\sup_{x\in S^d, s\in[0,BN^{-A/d}]} |\mu_N(D(x,s)) - \nu(D(x,s))| = O(N^{-A})$.

7.4. Remarks.

(1) For $d = 2$, $A = 1/2$, $r = 0$, and ν the uniform distribution on $SO(d+1)$, Corollary 7.3(1) is already implicitly shown in the proof of Theorem 2.5 of [12].

(2) If $r = A/d$ is maximal, and ν is again the uniform distribution on $SO(d+1)$, then Corollary 7.3 is not satisfying, as then the error term in 7.3(2) has the same order as

$$\sup_{x\in S^d, s\in[0,BN^{-A/d}]} |\nu(D(x,s))| = O(N^{-A}).$$

However, for intermediate constants r one obtains reasonable results: For instance, for $d = 2$, $A = 1/2$, and $r = 1/10$, one obtains

$$\sup_{x\in S^d,\, s\in[0,BN^{-1/10}]} |\mu_N(D(x,s)) - \nu(D(x,s))| = O(N^{-4/10}). \tag{7.7}$$

Local inequalities like (7.7) may be seen as supplements to [12] on the order of convergence of some recursive algorithm on distributing N points on S^2 uniformly.

References

[1] R. Askey, Orthogonal Polynomials and Special Functions, *SIAM*, 1975.

[2] A. C. Berry, The accuracy of the Gaussian approximation to the sum of independent variables, *Trans. Amer. Math. Soc.* **49** (1941), 122–136.

[3] N. H. Bingham, Random walks on spheres, *Z. Wahrscheinlichkeitsth. verw. Geb.* **22** (1972), 169–192.

[4] W. R. Bloom and H. Heyer, Harmonic Analysis of Probability Measures on Hypergroups, *De Gruyter*, 1995.

[5] M. Drmota and R. Tichy, Sequences, Discrepancies and Applications, vol. 1651, *Springer Lecture Notes in Math.*, 1997.

[6] L. Durand, Nicholson-type integrals for products of Gegenbauer functions and related topics, in: Theory and Applications of Special Functions (R. Askey, ed.), *Academic Press*, 1975, 353–374.

[7] C. G. Esséen, Fourier analysis of distribution functions, *Acta Math.* **77** (1945), 1–125.

[8] W. Feller, An Introduction to Probability Theory and its Applications, (2nd edn), vol. II, *New York, Wiley*, 1971.

[9] K.-J. Förster, Inequalities for ultraspherical polynomials and application to quadrature, *J. Comput. Appl. Math.* **49** (1993), 59–70.

[10] G. Gasper, Positivity and convolution structure for Jacobi series, *Ann. Math.* **93** (1971), 112–118.

[11] R. Lasser and J. Obermaier, On Fejér means with respect to orthogonal polynomials: a hypergroup-theoretic approach, in: Progress in approximation theory, *Academic Press*, *Boston, MA*, 1991, 551–565.

[12] A. Lubotzky, R. Phillips and P. Sarnack, Hecke operators and distributing points on the sphere I, *Comm. Pure Appl. Math.* **39** (1986), S149–S186.

[13] H. Niederreiter and W. Philipp, Berry–Esséen bounds and a theorem of Erdös and Turan on uniform distribution mod 1, *Duke Math. J.* **40** (1973), 633–649.

[14] G. Pap and M. Voit, Edgeworth expansion on n-spheres and Jacobi hypergroups, *Bull. Austral. Math. Soc.* **58** (1998), 393–401.

[15] W. M. Schmidt, Irregularities of distribution IV, *Invent. Math.* **7** (1969), 55–82.

[16] G. SZEGÖ, Orthogonal Polynomials, *Am. Math. Soc. Coll. Publ.* **23.**; *Providence, R. I.: Amer. Math. Soc.* (1959).

[17] M. VOIT, Central limit theorems for Jacobi hypergroups, in: Applications of hypergroups and related measure algebras (Seattle, 1993); *Contemp. Math.* **183**, 395–411 (1995).

[18] M. VOIT, Asymptotic behavior of heat kernels on spheres of large dimensions, *J. Multivariate Anal.* **59** (1996), 230–248.

[19] M. VOIT, Rate of convergence to Gaussian measures on n-spheres and Jacobi hypergroups, *Ann. Probab.* **25** (1997), 457–477.

MICHAEL VOIT
MATHEMATISCHES INSTITUT
UNIVERSITÄT TÜBINGEN
AUF DER MORGENSTELLE 10, 72076 TÜBINGEN
GERMANY

(Received September 6, 1997)