The Annals of Probability
1997, Vol. 25, No. 1, 457–477

RATE OF CONVERGENCE TO GAUSSIAN MEASURES ON *n*-SPHERES AND JACOBI HYPERGROUPS

By Michael Voit

Universität Tübingen

In this paper we prove central limit theorems of the following kind: let $S^d \subset \mathbb{R}^{d+1}$ be the unit sphere of dimension $d \geq 2$ with uniform distribution ω_d. For each $k \in \mathbb{N}$, consider the isotropic random walk $(X_n^k)_{n\geq 0}$ on S^d starting at the north pole with jumps of fixed sizes $\angle(X_n^k, X_{n-1}^k) = \pi/\sqrt{k}$ for all $n \geq 1$. Then there is some $k_0(d)$ such that for all $k \geq k_0(d)$, the distributions ϱ_k of X_k^k have continuous, bounded ω_d-densities f_k. Moreover, there is a (known) Gaussian measure ν on S^d with ω_d-density h such that $\|f_k - h\|_\infty = O(1/k)$ and $\|\varrho_k - \nu\| = O(1/k)$ for $k \to \infty$, where $O(1/k)$ is sharp. We shall derive this rate of convergence in the central limit theorem more generally for a quite general class of isotropic random walks on compact symmetric spaces of rank one as well as for random walks on $[0, \pi]$ whose transition probabilities are related to product linearization formulas of Jacobi polynomials.

Introduction. The purpose of this paper is to derive a Berry–Esséen-type central limit theorem for isotropic random walks on unit spheres $S^d \subset \mathbb{R}^{d+1}$ for $d \geq 2$ and, more generally, on compact symmetric spaces of rank 1. To stress this analogy, let us first recapitulate the classical setting (see [1], [8] and [9] for details).

Let $(Y_l)_{l\geq 1}$ be a sequence of i.i.d. $\mathbb{R}$-valued, centered random variables having third moments with common distribution $\mu \in M^1(\mathbb{R})$. Then, for $k \in \mathbb{N}$, the normalized variables $Y_l/\sqrt{k}$ are μ_k distributed with $\mu_k(A) := \mu(\sqrt{k}A)$ ($A \subset \mathbb{R}$ a Borel set), and the random variables $X_n^{(k)} := (1/\sqrt{k})\sum_{l=1}^n Y_l$ are $\mu_k^{(n)}$-distributed (with respect to usual convolution powers) and form a random walk $(X_n^{(k)})_{n\geq 0}$ on $\mathbb{R}$. The theorem of Berry and Esséen now states that the distribution functions of $\mu_k^{(k)}$ tend uniformly to a normal distribution function with $O(1/\sqrt{k})$.

Now consider the unit sphere $S^d \subset \mathbb{R}^{d+1}$ of dimension $d \geq 2$ with fixed North Pole $x_0 \in S^d$ and uniform distribution ω_d; regard the group $SO(d)$ of rotations as the subgroup of $SO(d+1)$ stabilizing x_0. A stationary random walk $(X_n)_{n\geq 0}$ on S^d is called isotropic (with respect to x_0) if it starts in x_0 at time 0 and if its transition probabilities are $SO(d+1)$-invariant. This yields that the distributions of these chains are $SO(d)$-invariant, and that they are determined completely by the distribution $\mu \in M^1([0, \pi])$ of the angles

$$\angle(X_n, X_{n-1}) \subset [0, \pi], \qquad n \in \mathbb{N}.$$

Received April 1995; revised January 1996.

AMS 1991 *subject classifications*. Primary 60J15; secondary 60F05, 60B10, 33C25, 42C10, 43A62.

Key words and phrases. Random walks on n-spheres, central limit theorem, Gaussian measures, compact symmetric spaces of rank one, total variation distance, Jacobi polynomials.

between two successive jumps. In this way, if some probability measure $\mu \in M^1([0,\pi])$ is given, we form the probability measure $\mu_k \in M^1([0,\pi/\sqrt{k}])$ with $\mu_k(A) := \mu(\sqrt{k}A)$ ($A \subset [0,\pi]$ a Borel set) and consider the associated random walk $(X_n^{(k)})_{n\geq 0}$ on S^d starting in x_0. We shall prove that under a mild restriction on μ ("μ must not be concentrated in 0 too much") there is some $k_0(d)$ such that for all $k \geq k_0(d)$, the distributions ϱ_k of $X_k^{(k)}$ have continuous, bounded ω_d-densities f_k. Moreover, there is a (known) Gaussian measure ν on S^d with ω_d-density h such that

$$\|f_k - h\|_\infty = O(1/k) \quad \text{and} \quad \|\varrho_k - \nu\| = O(1/k) \quad \text{for } k \to \infty,$$

where $O(1/k)$ is sharp. For Gaussian measures on spheres see, for instance, [12] and [28]. Our convergence result is stronger than the classical theorem of Berry and Esséen in two ways: the rate of convergence is $O(1/k)$ instead of $O(1/\sqrt{k})$, and we have convergence with respect to the total variation norm which is stronger than uniform convergence of the distribution functions.

The convergence result above will be derived in a more general setting. First of all, it holds for all compact symmetric spaces of rank 1 (i.e., projective spaces are also included) while the second generalization is motivated by its proof; it is as follows: Using the projection $\tilde{d}\colon S^d \to [0,\pi]$, $x \mapsto \angle(x, x_0)$, we see that for all $k \in \mathbb{N}$ the sequences $(\tilde{d}(X_n^{(k)}))_{n\geq 0}$ form Markov chains on $[0,\pi]$ starting in 0 where their transition probabilities are related to an abstract convolution structure on $[0,\pi]$ induced by product linearization formulas of the spherical functions, that is, in our setting, certain Jacobi polynomials (see Helgason [13]). It is now a matter of fact that Jacobi polynomials form such a convolution structure on $[0,\pi]$ for all indices $\alpha \geq \beta > -1$ with $\beta \geq -1/2$ or $\alpha + \beta \geq 0$, that this convolution structure is a hypergroup and that it makes sense to study random walks on $[0,\pi]$ whose transition probabilities are closely related to this hypergroup structure (see [10, 11], [15], [3], [14]). In this way, it is natural to establish a limit theorem for random walks on arbitrary Jacobi hypergroups on $[0,\pi]$ from which the limit results for compact symmetric spaces of rank 1 then easily follow.

This paper is organized as follows: in Section 1 we first recapitulate some facts on isotropic random walks and Gaussian measures on compact symmetric spaces of rank 1 and on Jacobi hypergroups on $[0,\pi]$. The main result will be stated in Theorem 1.6 in the latter setting. After a brief discussion, we shall transfer it to a central limit theorem on unit spheres. Section 2 deals with a main ingredient of the proof of Theorem 1.6, namely inequalities relating norms of functions and measures on $[0,\pi]$ with norms of their Jacobi–Fourier transforms. We state and prove these inequalities in the general setting of compact commutative hypergroups and then specialize it to the case of Jacobi hypergroups on $[0,\pi]$. An inequality of this type was used first by Diaconis and Shashahani [7] for compact (or, more specifically, for finite) groups. It recently turned out that inequalities of this kind have many applications in probability theory, mainly for bounds of convergence to uniformity; see, for instance, [6], [18], [19], [20], [25]. Finally, the proof of Theorem 1.6 will be completed in Section 3 by using the whole machinery of (asymptotic) properties of Jacobi polynomials.

1. Isotropic random walks on compact symmetric spaces of rank one and on Jacobi hypergroups.

1.1. *Compact two-point homogeneous spaces.* Let (X, d) be a (nontrivial) compact connected two-point homogeneous space; that is, X admits a compact group G of isometries such that for all $x, y, u, v \in X$ with $d(x, y) = d(u, v)$ there exists some $g \in G$ with $g(x) = u$ and $g(y) = v$. By the classification of Wang [27], these spaces are exactly the compact symmetric spaces of rank 1, and are equal either to the sphere $S^n \subset \mathbb{R}^{n+1}$ $(n \geq 1)$, to the projective space $\mathbb{P}^n(\mathbb{K})$ $(n \geq 2$, and $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or the field of quaternions), or to the projective plane $\mathbb{P}^2(\mathbb{O})$ over the octonions.

Consider the compact stabilizer subgroup H of some point $x_0 \in X$. Then the homogeneous space G/H (equipped with the quotient topology) can be identified with X, and the double coset space $G//H := \{HgH\colon g \in G\}$ with the space $X^H := \{H(x)\colon x \in X\}$ of all H-orbits in X where both spaces again carry the quotient topology. As G acts in a two-point homogeneous way, X^H can also be identified with the compact interval $D := \{d(x, x_0)\colon x \in X\} \subset [0, \infty[$ via $H(x) \simeq d(H(x), x_0)$. Without loss of generality we assume that $D = [0, \pi]$. In this way,

$$G//H \simeq X^H \simeq [0, \pi]. \tag{1.1}$$

Now consider the Banach-$*$-algebra $M_b(G)$ of all (signed) Borel measures on G with the usual convolution and involution of measures. If $\omega_H \in M_b(G)$ is the Haar measure of H normalized by $\omega_H(H) = 1$, then $M_b(G||H) := \{\mu \in M_b(G)\colon \omega_H * \mu * \omega_H = \mu\}$ is the Banach-$*$-subalgebra of $M_b(G)$ consisting of all H-biinvariant measures; the canonical projection $\pi\colon G \to G//H$, $g \mapsto HgH$, induces an isometric isomorphism $\tilde{\pi}$ between the Banach spaces $M_b(G||H)$ and $M_b(G//H)$. Now transfer convolution and involution to $M_b(G//H)$ such that $\tilde{\pi}$ becomes an isomorphism of Banach-$*$-algebras. In the two-point homogeneous setting above, these algebras are commutative, and the identification (1.1) leads to an isomorphic commutative Banach-$*$-algebra structure on the Banach space $M_b([0, \pi])$. It is well known (Helgason [13] or Section 3.5 of Bloom and Heyer [3]) that these convolution structures on $[0, \pi]$ are related to product formulas for Jacobi polynomials $(P_n^{(\alpha, \beta)})_{n \in \mathbb{N}_0}$ where α, β depend on X as follows.

If $X = S^n$ $(n \geq 1)$, then $\alpha = \beta = (n - 2)/2$.
If $X = \mathbb{P}^n(\mathbb{R})$ $(n \geq 2)$, then $\alpha = (n - 2)/2$, $\beta = -1/2$.
If $X = \mathbb{P}^n(\mathbb{C})$ $(n \geq 2)$, then $\alpha = n - 1$, $\beta = 0$.
If $X = \mathbb{P}^n(\mathbb{H})$ $(n \geq 2)$, then $\alpha = 2n - 1$, $\beta = 1$.
If $X = \mathbb{P}^2(\mathbb{O})$, then $\alpha = 7$, $\beta = 3$.

1.2. *Isotropic random walks on compact two-point homogeneous spaces.* Let X be a two-point homogeneous space as above. A stationary Markov chain $(S_n)_{n\geq 0}$ on $X$is called an isotropic random walk if it starts at some point $x_0 \in X$ (at time 0), and if its transition probabilities are G invariant, that is,

$$\mathsf{P}\big(S_n \in A \,|\, S_{n-1} = x\big) = \mathsf{P}\big(S_n \in g(A) \,|\, S_{n-1} = g(x)\big) \tag{1.2}$$

for $n \in \mathbb{N}$, $x \in X$, $g \in G$ and Borel sets $A \subset X$. In order to study the distributions of $(S_n)_{n\geq 0}$, consider the stabilizer H of the starting point $x_0 \in X$ and use the identifications $X \simeq G/H$ and $G//H \simeq X^H \simeq [0, \pi]$. Then the mapping $\tilde{d}$: $X \to [0, \pi]$, $x \mapsto d(x, x_0)$, corresponds to the canonical projection $G/H \to G//H$ and the projection $(\tilde{d}(S_n))_{n\in\mathbb{N}_0}$ becomes a Markov chain on $[0, \pi]$ by (1.2). Moreover, the transition probabilities of this Markov chain are related to the convolution structure on $[0, \pi]$ by

$$P\big(\tilde{d}(S_n) \in A \,|\, \tilde{d}(S_{n-1}) = z\big) = (\delta_z * \mu)(A) \quad \text{for } n \in \mathbb{N},\ z \in [0, \pi] \tag{1.3}$$

and for Borel sets $A \subset [0, \pi]$ where $\mu \in M^1([0, \pi])$ is just the distribution of the jump distances from S_{n-1} to S_n. Integration with respect to the kernel $(z, A) \mapsto \delta_z * \mu(A)$ and induction yield that the distribution of $\tilde{d}(S_n)$ is given by the n-fold convolution power $\mu^{(n)}$ for $n \geq 0$ where $\mu^{(n)}$ contains all information of the distribution of S_n. In summary, isotropic random walks on X starting in x_0 are described completely by μ.

1.3. *Random walks on Jacobi hypergroups.* For $\alpha, \beta > -1$, we define the Jacobi polynomials by

$$P_n^{(\alpha,\beta)}(x) := {}_2F_1(-n, n+\alpha+\beta+1; \alpha+1; (1-x)/2), \qquad x \in \mathbb{R},\ n \geq 0,$$

which are normalized by $P_n^{(\alpha,\beta)}(1) = 1$ and orthogonal on $[-1, 1]$ with respect to the weight $(1-x)^\alpha(1+x)^\beta$. It was shown by Gasper [10, 11] that exactly for indices α, β in the region

$$P := \{(\alpha, \beta) \in \mathbb{R}^2\colon \alpha \geq \beta > -1 \text{ and } (\beta \geq -1/2 \text{ or } \alpha + \beta \geq 0)\} \tag{1.4}$$

there exist (unique) probability measures $\delta_s * \delta_t$ on $[0, \pi]$ for all $s, t \in [0, \pi]$ such that

$$P_n^{(\alpha,\beta)}(\cos s) P_n^{(\alpha,\beta)}(\cos t) = \int_{[0,\pi]} P_n^{(\alpha,\beta)}(\cos u)\, d(\delta_s * \delta_t)(u), \qquad n \in \mathbb{N}_0. \tag{1.5}$$

For $\alpha = \beta \geq -1/2$, (1.5) is just Gegenbauer's product formula. The convolution $\delta_s * \delta_t$ determined by (1.5) can be extended uniquely to a bilinear, commutative, associative and weakly continuous convolution $*$ on the Banach space $M_b([0, \pi])$ of all (signed) Borel measures on $[0, \pi]$. Moreover, $*$ establishes a hypergroup structure on $[0, \pi]$, and for certain α, β this convolution is exactly the convolution derived in Section 1.1. For details, see [3], [13] and [15]. We do not need further details of this hypergroup here and omit an explicit representation of its convolution.

Now choose a probability measure $\mu \in M^1([0,\pi])$ and introduce the Markov chain $(X_n)_{n\geq 0}$ on $[0,\pi]$ associated with μ and the Jacobi-type hypergroup above as follows: the chain starts at 0 at time 0, and the transition probabilities satisfy

$$P\big(X_n \in A \mid X_{n-1} = z\big) = (\delta_z * \mu)(A) \quad \text{for } n \in \mathbb{N},\ z \in [0,\pi] \tag{1.6}$$

and $A \subset [0,\pi]$ a Borel set. As in Section 1.2, we obtain that X_n is $\mu^{(n)}$ distributed for each $n \in \mathbb{N}$ with respect to Jacobi convolution powers of indices $(\alpha,\beta) \in P$.

Before we are able to state the main result of this paper, we need two further definitions. The first one concerns Gaussian distributions on Jacobi hypergroups on $[0,\pi]$.

1.4. *Gaussian measures.* For all $(\alpha,\beta) \in P$, we define the functions

$$q(n) := q^{(\alpha,\beta)}(n) := -\left(\frac{d}{d\theta}\right)^2 P_n^{(\alpha,\beta)}(\cos\theta)\Big|_{\theta=0} = \frac{n(n+\alpha+\beta+1)}{2(\alpha+1)} \tag{1.7}$$

and

$$h_n := h_n^{(\alpha,\beta)} := \frac{(2n+\alpha+\beta+1)(\alpha+\beta+1)_n(\alpha+1)_n}{(\alpha+\beta+1)n!(\beta+1)_n} \tag{1.8}$$

for $n \in \mathbb{N}_0$. Then for all $\sigma^2 > 0$, the heat kernel

$$h_{\sigma^2}^{(\alpha,\beta)}(\theta,\varphi) := \sum_{n=0}^{\infty} h_n^{(\alpha,\beta)} \exp(-\sigma^2 q(n)/2) P_n^{(\alpha,\beta)}(\cos\theta) P_n^{(\alpha,\beta)}(\cos\varphi), \qquad \theta,\varphi \in [0,\pi] \tag{1.9}$$

is a positive continuous function on $[0,\pi]\times[0,\pi]$. The continuity follows immediately from $|P_n^{(\alpha,\beta)}(x)| \leq 1$ for $n \in \mathbb{N}_0$, $x \in [-1,1]$; the positivity is shown, for instance, in Bochner [4]. We mention that this positivity also follows from Theorem 1.6; see Remark 1.8. The probability measure

$$\begin{aligned} d\nu_{\sigma^2}(\theta) \\ := h_{\sigma^2}^{(\alpha,\beta)}(\theta,0)\frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)2^{\alpha+\beta+1}}(\sin\theta)^{2\alpha+1}(1+\cos\theta)^{\beta-\alpha}\,d\theta \end{aligned} \tag{1.10}$$

on $[0,\pi]$ is called the Gaussian measure with “variance” σ^2. The coefficients $h_n^{(\alpha,\beta)}$ can be regarded as weights of the Plancherel measure (cf. Section 2.3). Moreover, the measure

$$d\omega^{(\alpha,\beta)}(\theta) := \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)2^{\alpha+\beta+1}}(\sin\theta)^{2\alpha+1}(1+\cos\theta)^{\beta-\alpha}\,d\theta, \qquad \theta \in [0,\pi]$$

appearing in (1.10) is the normalized Haar measure of our hypergroup structure.

Gaussian measures on $[0,\pi]$ associated with Jacobi polynomials have been studied by many mathematicians from different points of view; see [16], [17],

[20]. Last, we mention that for the cases considered in Section 1.1, the Gaussian measures above are in fact projections of Gaussian measures on the underlying symmetric space. For spheres, see also [12] and [28].

1.5. *Condition* (G). We say that a probability measure μ on $[0, \pi]$ has growth property (G) at point 0 if there exist constants $c, p > 0$ such that

$$\mu([0, \varepsilon]) \le c \cdot \varepsilon^p \quad \text{for all } \varepsilon \in [0, \pi]. \tag{G}$$

THEOREM 1.6. *Let μ be a probability measure on $[0, \pi]$ with property* (G) *and variance*

$$\sigma^2 := \int_0^\pi x^2 \, d\mu(x).$$

For $k \in \mathbb{N}$, define the probability measure μ_k on $[0, \pi/k]$ by

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

Then, for all Jacobi hypergroups on $[0, \pi]$ with indices $(\alpha, \beta) \in P$ with $\alpha > -1/2$, there exists some $k_0 = k_0(\alpha, \beta, \mu)$ such that for each $k \ge k_0$, the distribution $\mu_k^{(k)}$ has a continuous, bounded $\omega^{(\alpha,\beta)}$-density f_k, and the densities f_k tend uniformly on $[0, \pi]$ to the Gaussian density $h_{\sigma^2}^{(\alpha,\beta)}(\,\cdot\,, 0)$, with

$$\|f_k - h_{\sigma^2}^{(\alpha,\beta)}(\,\cdot\,, 0)\|_\infty = O(1/k) \qquad \text{for } k \to \infty. \tag{1.11}$$

In particular, the convolution powers $\mu_k^{(k)}$ tend to the Gaussian measure ν_{σ^2} with respect to the total variation norm with

$$\|\mu_k^{(k)} - \nu_{\sigma^2}\| = O(1/k) \qquad \text{for } k \to \infty. \tag{1.12}$$

Finally, $O(1/k)$ in (1.11) and (1.12) is sharp for all measures $\mu \ne \delta_0$ on $[0, \pi]$.

REMARK 1.7. Comparing Theorem 1.6 with the theorem of Berry–Esséen, we have a better order of convergence, and in particular we have convergence with respect to the total variation norm, which is stronger than uniform convergence of the distribution functions. We suppose that the reason for this is as follows: the proof of the Berry–Esséen theorem needs a smoothing procedure (see [9] for details) which is in particular indispensible for discrete measures. This smoothing is not necessary in our setting, because even for discrete measures μ [with property (G)] the measures $\mu_k^{(k)}$ become absolutely continuous with respect to Haar measure for k sufficiently large (see also the proof of Theorem 1.6 in Section 3).

The case $\alpha = \beta = -1/2$, which corresponds to the torus $\{z \in \mathbb{C}\colon |z| = 1\}$ and which is not covered by Theorem 1.6, is in between the classical setting on $\mathbb{R}$ and the setting of Theorem 1.6. We hope to analyze this interesting case in a forthcoming paper.

REMARK 1.8. If μ is any probability measure on $[0,\pi]$ with $\mu \neq \delta_0$, then the convolution powers $\mu_k^{(k)}$ tend weakly to ν_{σ^2}. This weak result was proved, for instance, in [24], and may be seen as a result in the long history of triangular central limit theorems. Closely related central limit theorems can be found in [5], [22], [30] and [3].

Finally, we mention that the convergence of the probability measures $\mu_k^{(k)}$ to ν_{σ^2} ensures in particular that the heat kernel $h_{\sigma^2}(\theta, 0)$ is always nonnegative. The nonnegativity of the general heat kernel then follows from

$$h_{\sigma^2}(\theta,\varphi) = h_{\sigma^2}(\theta,0) * \delta_\varphi, \qquad \theta,\varphi \in [0,\pi]$$

with respect to the Jacobi convolution on $[0,\pi]$.

REMARK 1.9. If a probability measure μ on $[0,\pi]$ has property (G), then in particular μ has no point mass in 0. Moreover, Conditon (G) is equivalent to

$$\int_0^\pi t^{-q}\, d\mu(t) < \infty \quad \text{for some } q > 0. \tag{1.13}$$

In fact, if (1.13) holds, then, by the Cauchy–Schwarz inequality,

$$\mu([0,\varepsilon]) = \int_0^\varepsilon t^{q/2} t^{-q/2}\, d\mu(t) \le \left(\int_0^\varepsilon t^q\, d\mu(t) \int_0^\pi t^{-q}\, d\mu(t)\right)^{1/2} = O(\varepsilon^{q/2})$$

as claimed. Conversely, if property (G) holds, then for $0 < q < p$ we obtain

$$\begin{aligned}\int_0^1 t^{-q}\, d\mu(t) &\le \mu([0,1])2^q + \sum_{n=2}^\infty \mu([0,1/n])((n+1)^q - n^q)\\ &\le \mu([0,1])2^q + \sum_{n=2}^\infty n^{-p} n^{q-1} < \infty\end{aligned}$$

as desired.

At the end of this section we transfer Theorem 1.6 to a limit result for isotropic random walks on the unit sphere $S^d \subset \mathbb{R}^{d+1}$, $d \ge 2$, with uniform distribution ω_d. For this we use the north pole $x_0 = (1,0,\dots,0)$ and regard $SO(d)$ as stabilizer of x_0 in $SO(d+1)$. In this setting, Theorem 1.6 is as follows.

THEOREM 1.10. *Let μ be a probability measure on* $[0,\pi]$ *with property* (G) *and variance $\sigma^2 := \int_0^\pi x^2\, d\mu(x)$. For each $k \in \mathbb{N}$, let $(X_n^k)_{n\ge 0}$ be the isotropic random walk on S^d starting in x_0 such that the normalized jump distances*

$$\sqrt{k}\cdot \angle(X_n^k, X_{n-1}^k) \text{ are } \mu\text{-distributed for all } n.$$

Then there is some $k_0 = k_0(d,\mu)$ such that for each $k \ge k_0$, the distributions ϱ_k of X_k^k have continuous, bounded ω_d-densities f_k, and the densities f_k tend *uniformly on S^d to the Gaussian density*

$$h_{\sigma^2}(\theta) = \sum_{n=0}^{\infty} \frac{2n+d-1}{d-1} \binom{n+d-1}{d} \times \exp\left(\frac{-n(n+d-1)\sigma^2}{2d}\right) P_n^{((d-2)/2,\,(d-2)/2)}(\cos\theta) \tag{1.14}$$

(with respect to polar coordinates $x_1 = \cos\theta$, $x_2 = \sin\theta\cos\varphi, \ldots$ on S^d) with

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

In particular, the distributions ϱ_k tend to the Gaussian measure $\nu_{\sigma^2} := h_{\sigma^2}\omega_d$ on S^d with $\|\varrho_k - \nu_{\sigma^2}\| = O(1/k)$ for $k \to \infty$.

It is clear that Theorem 1.10 can be stated for projective spaces in the same way; we here omit details. Finally, we remark that Bingham [2] studied nonisotropic random walks on $S^2 \subset \mathbb{R}^3$, which are also closely related to Jacobi convolutions. It is clear that Theorem 1.10 can also be stated for these random walks.

2. The Fourier transform on compact commutative hypergroups and some inequalities.

The main ingredients of the proof of Theorem 1.6 are inequalities of the Diaconis–Shashahani type. As we think that these (more or less obvious) inequalities have many applications, we derive them here in a quite general setting, namely for compact commutative hypergroups. We first recapitulate some general facts about Fourier transform on compact commutative hypergroups for the convenience of the reader; for further details we refer to Section 2.2 of [3] and [14]. After having established the claimed inequality in this general setting, we return to Jacobi hypergroups on $[0, \pi]$ at the end of this section.

2.1. *Fourier transforms on compact commutative hypergroups.* A compact commutative hypergroup K is a compact Hausdorff space K together with an abstract convolution $*$ on the Banach space $M_b(K)$ of all (signed) Borel measures on K such that $(M_b(K), *)$ becomes a commutative Banach algebra. Moreover, for some further technical conditions, the following facts are required.

1. If $M^1(K)$ is the space of all probability measures on K, then $\mu * \nu \in M^1(K)$ for all $\mu, \nu \in M^1(K)$.
2. There exists an identity $e \in K$ with $\delta_e * \mu = \mu$ for all $\mu \in M^1(K)$, and there is a continuous involution $\cdot^-\colon K \to K$ with $\delta_{\overline{x}} * \delta_{\overline{y}} = (\delta_y * \delta_x)^-$ and with

$$e \in \operatorname{supp}(\delta_x * \delta_y) \iff x = \overline{y} \quad \text{for all } x, y \in K.$$

The most familiar examples are given by the usual convolutions of measures on compact commutative groups. Moreover, the above conditions are satified for the Jacobi convolutions on $[0, \pi]$ with $e = 0$ and the identity mapping as involution.

By Theorem 7.2A of [14], each compact commutative hypergroup K admits a unique normalized Haar measure ω, that is, $\omega \in M^1(K)$ satisfies $\omega * \mu = \omega$ for all $\mu \in M^1(K)$. The dual space $\widehat{K}$ of K is given by

$$\widehat{K} := \Big\{\alpha\colon K \to \mathbb{C} \text{ continuous, } \alpha \not\equiv 0, \\ \int_K \alpha(z)\, d(\delta_x * \delta_{\bar{y}})(z) = \alpha(x)\overline{\alpha(y)} \text{ for } x, y \in K\Big\},$$

where $\widehat{K}$ carries the discrete topology. The Fourier transform of $\mu \in M_b(K)$ and a function $f \in L^1(K, \omega)$ is given by

$$\widehat{\mu}(\alpha) := \int_K \overline{\alpha(x)}\, d\mu(x) \quad \text{and} \\ \widehat{f}(\alpha) := \widehat{f\omega}(\alpha) := \int_K \overline{\alpha(x)} f(x)\, d\omega(x), \qquad \alpha \in \widehat{K}. \tag{2.1}$$

As $|\alpha(x)| \le 1$ for all $\alpha \in \widehat{K}$ and $x \in K$ (see Section 2.2 of [3]), we obtain that

$$|\widehat{\mu}(\alpha)| \le \|\mu\| \quad \text{and} \quad |\widehat{f}(\alpha)| \le \|f\|_1 \tag{2.2}$$

for all $\alpha \in \widehat{K}$, $\mu \in M_b(K)$ and $f \in L^1(K, \omega)$.

The Haar measure ω admits a dual Plancherel measure π on $\widehat{K}$ defined by

$$\pi(\{\alpha\}) := \left(\int_K |\alpha|^2\, d\omega\right)^{-1}, \qquad \alpha \in \widehat{K}. \tag{2.3}$$

This measure leads to a Plancherel formula; that is, for $f \in L^2(K, \omega)$ we have

$$\int_K |f|^2\, d\omega = \int_{\widehat{K}} |\widehat{f}|^2\, d\pi = \sum_{\alpha \in \widehat{K}} \pi(\{\alpha\}) |\widehat{f}(\alpha)|^2. \tag{2.4}$$

The inverse Fourier transform of a function $g \in L^1(\widehat{K}, \pi)$ is given by

$$\check{g}(x) := \sum_{\alpha \in \widehat{K}} g(\alpha)\pi(\{\alpha\})\alpha(x), \qquad x \in K, \tag{2.5}$$

where $\check{g}$ is a continuous function on K with

$$\|\check{g}\|_\infty \le \|g\|_1. \tag{2.6}$$

Finally, we recapitulate the following inversion theorem: if for some $\mu \in M_b(K)$ its Fourier transform $\widehat{\mu}$ is contained in $L^1(\widehat{K}, \pi)$, then μ has a continuous ω density with $\mu = (\widehat{f})^\vee \omega$; see Theorem 2.2.36 in [3].

We are now in a position to derive the following inequalities.

LEMMA 2.2. *Let μ be a measure on a compact commutative hypergroup K.*

(i) *If μ has a continuous ω-density f, then*

$$\|\mu\|^2 \le \int_K |f|^2\, d\omega = \|\widehat{f}\|_2^2 = \|\widehat{\mu}\|_2^2 = \sum_{\alpha \in \widehat{K}} \pi(\{\alpha\}) |\widehat{\mu}(\alpha)|^2, \tag{2.7}$$

where the right-hand side may be equal to infinity.

(ii) *If* $\widehat{\mu} \in L^1(\widehat{K}, \pi)$*, then* μ *has a continuous* ω*-density* f *with*

$$\|\mu\| \le \|f\|_\infty \le \|\widehat{f}\|_1 = \|\widehat{\mu}\|_1. \tag{2.8}$$

Both results are obvious from (2.4) and (2.6), respectively, where in the second case the inversion formula has to be employed. We here note that the "Diaconis–Shashahani-type inequality" in (2.7) is stronger than (2.8) if one is interested in estimates of $\|\mu\|$ only. However, $\|f\|_\infty$ can be estimated only by using (2.8).

A similar, slightly weaker version of part (i) of Lemma 2.2 was recently shown by Ross and Xu [19]; they applied their result to rates of convergence of Metropolis Markov chains. A discrete version of 2.2 was also applied in [25] to the Ehrenfest urn.

We next turn to the concrete examples of Section 1.

2.3. *Examples.* For all $\alpha \ge \beta > -1$ with $\beta \ge -1/2$ or $\alpha + \beta \ge 0$, the product linearization (1.5) of Jacobi polynomials induces a compact commutative hypergroup structure on $[0, \pi]$. Its normalized Haar measure is given by

$$\frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)2^{\alpha+\beta+1}}(\sin\theta)^{2\alpha+1}(1+\cos\theta)^{\beta-\alpha}\,d\theta, \qquad \theta \in [0, \pi] \tag{2.9}$$

and the dual space by $\widehat{K} = \{P_n;\ n \in \mathbb{N}_0\}$ (in fact, "$\supset$" follows from the linearization (1.5) while the converse inclusion needs more care; cf. [15]). Moreover, as

$$\begin{aligned}\left(\int_0^\pi |P_n^{(\alpha,\beta)}(\cos\theta)|^2\,d\omega(\theta)\right)^{-1}&\\ = \frac{(2n+\alpha+\beta+1)(\alpha+\beta+1)_n(\alpha+1)_n}{(\alpha+\beta+1)n!(\beta+1)_n} &= h_n^{(\alpha,\beta)}\end{aligned} \tag{2.10}$$

(see, for instance, equation (4.3.3) in [21]), the Plancherel measure on $\widehat{K}$ is given by the constants $h_n^{(\alpha,\beta)}$. In view of the Fourier transform (2.1), Lemma 2.2 is as follows: if $\mu \in M_b([0, \pi])$ has a ω-density f, then

$$\|\mu\|^2 \le \|\widehat{\mu}\|_2^2 = \sum_{n=0}^\infty h_n^{(\alpha,\beta)}\left|\int_0^\pi P_n^{(\alpha,\beta)}(\cos\theta)\,d\mu(\theta)\right|^2 \tag{2.11}$$

and

$$\|\mu\| \le \|f\|_\infty \le \|\widehat{f}\|_1 = \sum_{n=0}^\infty h_n^{(\alpha,\beta)}\left|\int_0^\pi P_n^{(\alpha,\beta)}(\cos\theta)\,d\mu(\theta)\right|. \tag{2.12}$$

In order to apply (2.11) or (2.12) to the proof of Theorem 1.6, we set $\mu := \mu_k^{(k)} - \nu_{\sigma^2}$ (cf. Section 1.4 and Theorem 1.6). Then

$$\begin{aligned}\widehat{\mu_k^{(k)}}(n) = (\widehat{\mu}_k(n))^k &= \left(\int_0^\pi P_n^{(\alpha,\beta)}(\cos(\theta))\,d\mu_k(\theta)\right)^k\\ &= \left(\int_0^\pi P_n^{(\alpha,\beta)}\big(\cos\big(t/\sqrt{k}\big)\big)\,d\mu(t)\right)^k,\end{aligned} \tag{2.13}$$

and, by (2.10) and the orthogonality of Jacobi polynomials,

$$\widehat{\nu}_{\sigma^2}(n) = \exp\left(-\frac{q(n)\sigma^2}{2}\right) \tag{2.14}$$

for all $n \geq 0$. Hence, (2.12) leads to

$$\begin{aligned} \|\mu_k^{(k)} - \nu_{\sigma^2}\| &\leq \|f_k - h_{\sigma^2}(\,\cdot\,,0)\|_\infty \\ &\leq \sum_{n=0}^{\infty} h_n \left| \left(\int_0^\pi P_n^{(\alpha,\beta)}\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t) \right)^k \right. \\ &\qquad\qquad \left. - \exp\left(-\frac{q(n)\sigma^2}{2}\right) \right| \end{aligned} \tag{2.15}$$

whenever the right-hand side of (2.15) is finite and where f_k is the ω-density of $\mu_k^{(k)}$. Equation (2.11) leads to a similar inequality.

To complete the proof of Theorem 1.6, we have to estimate the right-hand side of (2.15); this will be done in Section 3.

3. Proof of the central limit theorem. We first give an outline of the proof of Theorem 1.6 for the convenience of the reader. The proofs of several technical steps of the proof will be postponed; they are contained in a sequence of lemmas following the main body of the proof. Note that the superscripts (α, β) are often omitted from now on.

3.1. *Outline of the proof.* First of all, (2.15) states that

$$\begin{aligned} \|\mu_k^{(k)} - \nu_{\sigma^2}\| &\leq \|f_k - h_{\sigma^2}(\,\cdot\,,0)\|_\infty \\ &\leq \sum_{n=0}^{\infty} h_n \left| \left(\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t) \right)^k \right. \\ &\qquad\qquad \left. - \exp\left(-\frac{q(n)\sigma^2}{2}\right) \right| \\ &=: R_k. \end{aligned} \tag{3.1}$$

Now let $A > 0$ be a constant which will be fixed later. Then

$$\begin{aligned} R_k \leq{}& \sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} h_n \left| \left(\int_0^\pi P_n(\cos(t/\sqrt{k}))\, d\mu(t) \right)^k - \left(1 - \frac{q(n)\sigma^2}{2k}\right)^k \right| \\ &+ \sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} h_n \left| \left(1 - \frac{q(n)\sigma^2}{2k}\right)^k - \exp\left(-\frac{q(n)\sigma^2}{2}\right) \right| \\ &+ \sum_{n\geq\lfloor A\sqrt{k}\rfloor} h_n \exp\left(-\frac{q(n)\sigma^2}{2}\right) \\ &+ \sum_{n\geq\lfloor A\sqrt{k}\rfloor} h_n \left| \int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t) \right|^k. \end{aligned} \tag{3.2}$$

We next have to estimate the four sums on the right-hand side of (3.2). In order to handle the first one, we shall prove in Lemma 3.3 below that

$$\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right)d\mu(t) = 1 - \frac{q(n)\sigma^2}{2k} + \frac{n^4 H}{k^2}, \tag{3.3}$$

where $H = H(n, k, \mu)$ is bounded by some constant independent of k and n. This fact will be used in the proof of Lemma 3.4 below to show that, if the constant $A > 0$ is sufficiently small, we find some constant C_1 with

$$\sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} h_n \left|\left(\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right)d\mu(t)\right)^k - \left(1 - \frac{q(n)\sigma^2}{2k}\right)^k\right| \le \frac{C_1}{k}. \tag{3.4}$$

We next turn to the second term. As

$$q(n) = \frac{n(n+\alpha+\beta+1)}{2(\alpha+1)}, \qquad n \in \mathbb{N}_0, \tag{3.5}$$

we may choose the constant $A > 0$ sufficiently small such that

$$\frac{q(n)\sigma^2}{2k} \le 1 \quad \text{for all } n \le A\sqrt{k}. \tag{3.6}$$

If we now use

$$h_n = \frac{(2n+\alpha+\beta+1)(\alpha+\beta+1)_n(\alpha+1)_n}{(\alpha+\beta+1)n!(\beta+1)_n} = O(n^{2\alpha+1}) \tag{3.7}$$

as well as the well-known inequality

$$0 \le e^{-z} - (1 - z/t)^t \le e^{-z}z^2/t \quad \text{for } t \ge 1,\ z \in \mathbb{R},\ |z| \le 1, \tag{3.8}$$

then we find constants $C_2, C_3 > 0$ such that

$$\begin{aligned} &\sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} h_n \left|\left(1 - \frac{q(n)\sigma^2}{2k}\right)^k - \exp\left(-\frac{q(n)\sigma^2}{2}\right)\right| \\ &\qquad \le C_2 \sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} n^{2\alpha+1} n^4 \exp\left(-\frac{q(n)\sigma^2}{2}\right)\frac{1}{k} \le C_3/k. \end{aligned} \tag{3.9}$$

Moreover, (3.7) and (3.8) also lead to

$$\sum_{n \ge \lfloor A\sqrt{k}\rfloor} h_n \exp(-q(n)\sigma^2/2) \le \sum_{n \ge \lfloor A\sqrt{k}\rfloor} n^{2\alpha+1}\exp(-C_4 n^2) = O(C_5^k) \tag{3.10}$$

for suitable constants $C_4 > 0$ and $0 < C_5 < 1$. Finally, it will be shown in Lemma 3.8 that for $\alpha > -1/2$,

$$\sum_{n \ge \lfloor A\sqrt{k}\rfloor} h_n \left|\int_0^\pi P_n\big(\cos(t/\sqrt{k})\big)\,d\mu(t)\right|^k \le C_6 \cdot C_7^k \tag{3.11}$$

for suitable constants $C_6 > 0$ and $0 < C_7 < 1$ and for k sufficiently large. In summary, (3.2), (3.4), (3.9), (3.10) and (3.11) yield that $R_k \le C/k$ for some constant C and for k sufficiently large. Therefore, in view of the inequality (3.1), Theorem 1.6 follows.

In order to complete the proof of Theorem 1.6, we still have to establish (3.3), (3.4) and (3.11). The proof of (3.3) will be based on the following Taylor formula for Jacobi polynomials.

LEMMA 3.2. *For all $\alpha \geq \beta > -1$ there exists a constant $C > 0$ such that, for all $n \in \mathbb{N}_0$ and $x \in [-1, 1]$,*

$$\left| P_n^{(\alpha,\beta)}(x) - \left(1 - (1-x)\frac{(n+\alpha+\beta+1)n}{2(\alpha+1)}\right)\right| \leq Cn^4(1-x)^2.$$

PROOF. Taylor's formula ensures that for each $x \in [-1, 1]$ there exists some $z \in [x, 1]$ with

$$\left|P_n^{(\alpha,\beta)}(x) - \left(1 - (1-x)P_n^{(\alpha,\beta)'}(1)\right)\right| \leq \tfrac{1}{2}(1-x)^2 P_n^{(\alpha,\beta)''}(z).$$

In view of our normalization $P_n^{(\alpha,\beta)}(1) = 1$, (4.21.7) and (4.1.1) of [21] lead to

$$P_n^{(\alpha,\beta)'}(1) = \frac{(n+\alpha+\beta+1)n}{2(\alpha+1)} = q(n).$$

Also, (4.21.7) and (4.1.1) of Szegö [21] yield

$$P_n^{(\alpha,\beta)''}(z) = \frac{n(n-1)(n+\alpha+\beta+1)(n+\alpha+\beta+2)}{4(\alpha+1)(\alpha+2)} P_{n-2}^{(\alpha+2,\beta+2)}(z).$$

As $|P_{n-2}^{(\alpha+2,\beta+2)}(z)| \leq 1$, the proof of the lemma is complete. □

Lemma 3.2 has the following consequence.

LEMMA 3.3. *Assume that $\alpha \geq \beta > -1$. Let μ be a probability measure on $[0, \pi]$ with $\sigma^2 = \int_0^\pi t^2\, d\mu(t)$. Then, for all $n, k \in \mathbb{N}$,*

$$\int_0^\pi P_n^{(\alpha,\beta)}\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t) = 1 - \frac{q(n)\sigma^2}{2k} + \frac{n^4 H}{k^2},$$

where $H = H(n, k, \mu)$ is bounded by some constant independent of k and n.

PROOF. As μ is a probability measure, Lemma 3.2 and $|1 - \cos x| \leq x^2/2$ imply that

$$\int_0^\pi P_n^{(\alpha,\beta)}\left(\cos(t/\sqrt{k})\right) d\mu(t) = 1 - q(n)\int_0^\pi \left(1 - \cos(t/\sqrt{k})\right) d\mu(t) + R_1$$

with

$$|R_1| \leq \frac{n(n-1)(n+\alpha+\beta+1)(n+\alpha+\beta+2)}{8(\alpha+1)(\alpha+2)4k^2} \int_0^\pi t^4\, d\mu(t) = \frac{n^4}{k^2} O(1). \tag{3.12}$$

Moreover, since $|1-\cos x - x^2/2| \le x^4/24$ and $q(n) = O(n^2)$,

$$q(n)\left|\int_0^\pi \left(1-\cos\left(\frac{t}{\sqrt{k}}\right) - \frac{t^2}{(2k)}\right) d\mu(t)\right| = \frac{n^2}{k^2} O(1). \tag{3.13}$$

Lemma 3.3 now follows from (3.12) and (3.13). □

Lemma 3.3 has the following consequence.

LEMMA 3.4. *Assume that $\alpha, \beta, \mu, \sigma^2$ are given as in Lemma 3.3. Then there exists a constant $A > 0$ such that*

$$\sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} h_n \left| \left(\int_0^\pi P_n^{(\alpha,\beta)}\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t)\right)^k - \left(1 - \frac{q(n)\sigma^2}{2k}\right)^k \right| = O\left(\frac{1}{k}\right).$$

PROOF. Let H be given as in Lemma 3.3. Choose $A > 0$ sufficiently small such that

$$\frac{n^4|H|}{k^2} \le \frac{1}{2}\frac{q(n)\sigma^2}{2k} \quad \textit{for all } n \le A\sqrt{k}. \tag{3.14}$$

Applying the mean value theorem to the function $x \mapsto (1-x)^k$ together with the inequality $(1 - z/k)^k \le e^{-z}$ [$z, k > 0$; cf. (3.8)], we obtain from (3.14) that

$$\begin{aligned} &\left| \left(1 - \frac{q(n)\sigma^2}{2k} + \frac{n^4 H}{k^2}\right)^k - \left(1 - \frac{q(n)\sigma^2}{2k}\right)^k \right| \\ &\quad \le \frac{n^4|H|}{k^2} k \max\left[\left(1 - \frac{q(n)\sigma^2}{2k} + \frac{n^4 H}{k^2}\right), \left(1 - \frac{q(n)\sigma^2}{2k}\right)\right]^{k-1} \\ &\quad \le \frac{n^4|H|}{k}\left(1 - \frac{q(n)\sigma^2}{4k}\right)^{k-1} \le \frac{2n^4|H|}{k} \exp\left(\frac{-q(n)\sigma^2}{4}\right) \end{aligned} \tag{3.15}$$

for $n \le A\sqrt{k}$. As $h_n = O(n^{2\alpha+1})$ by (3.7), it follows that

$$\sum_{n=0}^{\lfloor A\sqrt{k}\rfloor} h_n \left| \left(1 - \frac{q(n)\sigma^2}{2k} + \frac{n^4 H}{k^2}\right)^k - \left(1 - \frac{q(n)\sigma^2}{2k}\right)^k \right| = O\left(\frac{1}{k}\right). \tag{3.16}$$

Thus, Lemma 3.4 is a consequence of Lemma 3.3.

We still have to establish (3.11). Its proof depends on estimations of the growth of $|P_n^{(\alpha,\beta)}(\cos\theta)|$ for large n and small θ. For this, we use Hilb's formula and compare the Jacobi polynomials with the (spherical) Bessel functions Λ_α which are given by

$$\Lambda_\alpha(x) := \Gamma(\alpha+1)\left(\frac{x}{2}\right)^{-\alpha} J_\alpha(x) := \Gamma(\alpha+1)\sum_{n=0}^{\infty} \frac{(-1)^n (x/2)^{2n}}{n!\,\Gamma(n+\alpha+1)}, \qquad x \in \mathbb{C}, \alpha > -1$$

and normalized by $\Lambda_\alpha(0) = 1$. With this normalization, Hilb's formula is as follows.

LEMMA 3.5. *For $\alpha \ge \beta > -1$ with $\beta \ge -1/2$ or $\alpha+\beta \ge 0$, let the remainder $R_n^{(\alpha,\beta)}(\theta)$ be given by*

$$R_n^{(\alpha,\beta)}(\theta) := P_n^{(\alpha,\beta)}(\cos\theta) - \Lambda_\alpha\left(\theta\left(n + \frac{\alpha+\beta+1}{2}\right)\right).$$

Then for each $c > 0$ there is a constant $M > 0$ with

$$R_n^{(\alpha,\beta)}(\theta) \le \begin{cases} Mn^{-2}, & \text{if } 0 \le \theta \le c/n, \\ Mn^{-2}(\theta n)^{1/2-\alpha}, & \text{if } c/n \le \theta \le \pi/2. \end{cases}$$

PROOF. Using our normalization $P_n^{(\alpha,\beta)}(1) = \Lambda_\alpha(0) = 1$, we may rewrite Hilb's formula for Jacobi polynomials (see Theorem 8.21.12 in [21]) as follows:

$$\left(\frac{\sin\theta}{\theta}\right)^{\alpha+1/2}\left(\cos\left(\frac{\theta}{2}\right)\right)^{\beta-\alpha} P_n^{(\alpha,\beta)}(\cos\theta) - \Lambda_\alpha\left(\theta\left(n + \frac{\alpha+\beta+1}{2}\right)\right)$$
$$= \begin{cases} \theta^2 O(1), & \text{if } 0 \le \theta \le c/n, \\ \theta^{1/2-\alpha} n^{-3/2-\alpha} O(1), & \text{if } c/n \le \theta \le \pi/2. \end{cases}$$

(c some constant). As $|P_n^{(\alpha,\beta)}(\cos\theta)| \le 1$ for $\theta \in \mathbb{R}$, Lemma 3.5 follows immediately where in particular in the first case the condition $\theta \le c/n$ has to be used.

We also notice here the following obvious consequence of the asymptotic formula for Bessel functions (see [29], page 368).

LEMMA 3.6. $\Lambda_\alpha(x) = O(1/x^{\alpha+1/2})$ *for $x \to \infty$.*

Assume from now on that $\alpha \ge \beta > -1$ and the probability measure μ is given as in Theorem 1.6. Then, in particular, $\alpha > -1/2$ holds, and μ has the following technical property.

LEMMA 3.7. *If a probability measure μ in $[0,\pi]$ satisfies Condition* (G)*, then there exist constants $M, r > 0$ such that for each $\varepsilon > 0$ there exists $\delta > 0$ with*

$$\mu([0,\delta]) < \varepsilon \quad \text{and} \quad \int_\delta^\pi \frac{1}{t^{\alpha+1/2}}\, d\mu(t) \le M\varepsilon^{-r}.$$

PROOF. By Condition (G), we find $c, p > 0$ with $\mu([0,\delta]) < c\delta^p$ for all $\delta \ge 0$. Hence, if $\varepsilon > 0$ is given, and if $\delta := (\varepsilon/c)^{1/p}$, then $\mu([0,\delta]) < \varepsilon$ and

$$\int_\delta^\pi t^{-(\alpha+1/2)}\, d\mu(t) \le \delta^{-(\alpha+1/2)} = c^{(\alpha+1/2)/p}\varepsilon^{-(\alpha+1/2)/p}$$

as claimed. □

Using these preparations, we now proceed with the proof of (3.11).

LEMMA 3.8. *For each $A > 0$ there exist constants $C, D > 0$ and $0 < E < 1$ with*

$$\sum_{n\geq\lfloor A\sqrt{k}\rfloor} h_n \left|\int_0^\pi P_n\big(\cos(t/\sqrt{k})\big)\, d\mu(t)\right|^k \leq C\cdot E^k \quad \text{for all } k\geq D.$$

PROOF. During the proof of this lemma, $C_1, C_2, \ldots$ denote constants independent of k and n. According to Lemma 3.7 there exist constants $M, r > 0$ such that for

$$\varepsilon_{n,k} := \big(\sqrt{k}/n\big)^{(\alpha+1/2)/(1+r)} > 0, \tag{3.17}$$

we find $\delta_{n,k} > 0$ with

$$\mu([0,\delta_{n,k}]) < \varepsilon_{n,k} \quad \text{and} \quad \int_{\delta_{n,k}}^\pi \frac{1}{t^{\alpha+1/2}}\, d\mu(t) \leq M\varepsilon_{n,k}^{-r}. \tag{3.18}$$

As $|P_n^{(\alpha,\beta)}(x)| \leq 1$ for $x\in[-1,1]$ and $n\in\mathbb{N}_0$, we see that

$$\int_0^{\delta_{n,k}} \big|P_n\big(\cos(t/\sqrt{k})\big)\big|\, d\mu(t) \leq \varepsilon_{n,k}. \tag{3.19}$$

Hence, using the remainder $R_n^{(\alpha,\beta)}$ of Lemma 3.5, we get

$$\begin{aligned}
&\left|\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t)\right| \\
&\qquad \leq \varepsilon_{n,k} + \int_{\delta_{n,k}}^\pi \left|\Lambda_\alpha\left(\frac{t}{\sqrt{k}}\left(n+\frac{\alpha+\beta+1}{2}\right)\right)\right| d\mu(t) \\
&\qquad\quad + \int_{\delta_{n,k}}^\pi \left|R_n^{(\alpha,\beta)}\left(\frac{t}{\sqrt{k}}\right)\right| d\mu(t).
\end{aligned}$$

If $\alpha \geq 1/2$, then Hilb's formula 3.5 yields that $|R_n^{(\alpha,\beta)}(t/\sqrt{k})| \leq M/n^2$ and hence

$$\int_{\delta_{n,k}}^\pi \big|R_n^{(\alpha,\beta)}\big(t/\sqrt{k}\big)\big|\, d\mu(t) \leq M/n^2 \quad \text{for all } n, t, k$$

and a suitable constant M. Moreover, if $-1/2 < \alpha < 1/2$, then Lemma 3.5 yields that

$$\int_{\delta_{n,k}}^\pi \big|R_n^{(\alpha,\beta)}\big(t/\sqrt{k}\big)\big|\, d\mu(t) \leq M/n^2\left(1+\int_{\delta_{n,k}}^\pi \big(tn/\sqrt{k}\big)^{1/2-\alpha}\, d\mu(t)\right)$$

for a suitable constant M. Summing up, for $\alpha > -1/2$ we obtain that

$$\begin{aligned}
&\left|\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t)\right| \\
&\quad \le \varepsilon_{n,k} + \int_{\delta_{n,k}}^\pi \left|\Lambda_\alpha\left(\frac{t}{\sqrt{k}}\left(n + \frac{(\alpha+\beta+1)}{2}\right)\right)\right| d\mu(t) \\
&\qquad + \frac{C_1}{n^2}\left(1 + \left(\frac{\sqrt{k}}{n}\right)^{\alpha-1/2}\right).
\end{aligned}$$

If we apply Lemma 3.6 together with (3.17) and (3.18), then it follows that

$$\begin{aligned}
&\left|\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t)\right| \\
&\quad \le \varepsilon_{n,k} + C_2 \int_{\delta_{n,k}}^\pi \frac{1}{t^{\alpha+1/2}}\, d\mu(t) \left(\frac{\sqrt{k}}{n + (\alpha+\beta+1)/2}\right)^{\alpha+1/2} \\
&\qquad + \frac{C_1}{n^2}\left(1 + \left(\frac{\sqrt{k}}{n}\right)^{\alpha-1/2}\right) \\
&\quad \le C_3 \varepsilon_{n,k} + \frac{C_1}{n^2}\left(1 + \left(\frac{\sqrt{k}}{n}\right)^{\alpha-1/2}\right)
\end{aligned} \tag{3.20}$$

for $n \ge 1$. We next take some $B \ge A$ which will be fixed later. We use (3.20), the elementary inequality

$$(a+b)^k \le 2^{k-1}(a^k + b^k), \qquad a, b \ge 0,\ k \in \mathbb{N}_0,$$

as well as $h_n = O(n^{2\alpha+1})$ and conclude that

$$\begin{aligned}
&\sum_{n \ge \lfloor B\sqrt{k} \rfloor} h_n \left|\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t)\right|^k \\
&\quad \le C_4 \sum_{n \ge \lfloor B\sqrt{k} \rfloor} n^{2\alpha+1}\left[C_3\varepsilon_{n,k} + \frac{C_1}{n^2} + C_1 \frac{\sqrt{k}^{\alpha-1/2}}{n^{\alpha+3/2}}\right]^k \\
&\quad \le C_5 \sum_{n \ge \lfloor B\sqrt{k} \rfloor} n^{2\alpha+1}\left[C_6^k \frac{\sqrt{k}^{k(\alpha+1/2)/(1+r)}}{n^{k(\alpha+1/2)/(1+r)}} + \frac{C_7^k}{n^{2k}} + \frac{C_8^k \sqrt{k}^{k(\alpha-1/2)}}{n^{k(\alpha+3/2)}}\right] \\
&\quad \le C_9\Big[C_6^k \sqrt{k}^{k(\alpha+1/2)/(1+r)} \big(B\sqrt{k}\big)^{2\alpha+2-k(\alpha+1/2)/(1+r)} \\
&\qquad + C_7^k \big(B\sqrt{k}\big)^{2\alpha+2-2k} + C_8^k \sqrt{k}^{k(\alpha-1/2)} \big(B\sqrt{k}\big)^{2\alpha+2-k(\alpha+3/2)}\Big]
\end{aligned} \tag{3.21}$$

with constants C_i independent of k, n and B. Now choose k and B sufficiently large in view of the constants C_6, C_7, C_8, α and r such that the right-hand side of (3.21) has order $O(1/2^k)$ for $k \to \infty$.

We next turn to the sum for $\lfloor B\sqrt{k}\rfloor \geq n \geq \lfloor A\sqrt{k}\rfloor$. If we apply $h_n = O(n^{2\alpha+1})$ and the first case of Hilb's formula 3.5, then we obtain that

$$\begin{aligned}
&\sum_{n=\lfloor A\sqrt{k}\rfloor}^{\lfloor B\sqrt{k}\rfloor} h_n \left|\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t)\right|^k \\
&\qquad \leq C_{10}(B-A)\sqrt{k}\big(B\sqrt{k}\big)^{2\alpha+1} \\
&\qquad\quad \times \max_{\lfloor B\sqrt{k}\rfloor \geq n \geq \lfloor A\sqrt{k}\rfloor} \left|\int_0^\pi P_n\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right) d\mu(t)\right|^k \qquad (3.22)\\
&\qquad \leq C_{10}\big(B\sqrt{k}\big)^{2\alpha+2} \\
&\qquad\quad \times \max_{(x-(\alpha+\beta+1)/2)/\sqrt{k}\in[A,\,B]} \left(\left|\int_0^\pi \Lambda_\alpha\left(\frac{tx}{\sqrt{k}}\right) d\mu(t)\right| + \frac{C_{11}}{k}\right)^k
\end{aligned}$$

We next observe that $\mu \neq \delta_0$ and that

$$|\Lambda_\alpha(z)| < 1 \quad \text{for all } z > 0,\ \alpha > -1/2 \tag{3.23}$$

with $\Lambda_\alpha(0) = 1$. Therefore, the Hankel transform $\varphi(z) := \int_0^\pi \Lambda_\alpha(zt)\, d\mu(t)$ of μ satisfies $|\varphi(z)| < 1$ for all $z > 0$. As φ is continuous, it follows that there is a constant $C_{12} < 1$ with

$$\max_{(x-(\alpha+\beta+1)/2)/\sqrt{k}\in[A,\,B]} \left|\int_0^\pi \Lambda_\alpha\left(\frac{tx}{\sqrt{k}}\right) d\mu(t)\right| < C_{12} < 1.$$

Hence, if k is sufficiently large, then (3.22) implies that that we find constants C_{13} and $0 < E < 1$ with

$$\sum_{n=\lfloor A\sqrt{k}\rfloor}^{\lfloor B\sqrt{k}\rfloor} h_n \left|\int_0^\pi P_n\big(\cos\big(t/\sqrt{k}\big)\big)\, d\mu(t)\right|^k \leq C_{13}E^k.$$

This completes the proof of the lemma. □

The proof of the upper bounds of Theorem 1.6 is now complete.

3.9. *Remarks.*

1. In the proof of Theorem 1.6, property (G) of the measure μ was needed only for the final step of the proof in Lemma 3.8. In fact, it seems to be possible that

$$\sum_{n\geq\lfloor A\sqrt{k}\rfloor} h_n \left|\int_0^\pi P_n\big(\cos\big(t/\sqrt{k}\big)\big)\, d\mu(t)\right|^k$$

does not tend to 0 for $k \to \infty$ (or even diverges for all $k \in \mathbb{N}$) whenever the measure μ is concentrated "too much" in point 0. We however do not know whether conditions considerably weaker than (G) are sufficient to ensure that the statement of Lemma 3.8 remains valid.

2. In order to make Theorem 1.6 more explicit, it would be interesting to have explicit bounds C in

$$\|\mu_k^{(k)} - \nu_{\sigma^2}\| \le C/k \quad \text{for all } k \in \mathbb{N},$$

where C may depend on α, β and μ. It seems to be possible to obtain reasonable explicit bounds in all steps of the proof above except possibly for the final step in the proof of Lemma 3.8, which would need a more careful discussion. However, in any case, the proof of Lemma 3.4 above shows the constant becomes very large for small variances σ^2.
3. Methods similar to the proof of Theorem 1.6 were applied in [25] to derive asymptotic results for Markov chains related to the Ehrenfest urn model. We think that the methods above can be also used to study the exact asymptotic behaviour of isotropic random walks and Gaussian semigroups on the spheres S^n for $n \to \infty$. Similar problems were recently discussed in [16], [20] and [23].

At the end of this paper, we still show that $O(1/k)$ in Theorem 1.6 is in fact sharp.

LEMMA 3.10. *For each probability measure $\mu \neq \delta_0$ on $[0,\pi]$ there exists a constant $c = c(\alpha,\beta,\mu) > 0$ such that*

$$\|\mu_k^{(k)} - \nu_{\sigma^2}\| \ge c/k \quad \textit{for all } k \in \mathbb{N}.$$

PROOF. We first use Taylor's expansion of $P_n^{(\alpha,\beta)}(\cos\theta)$ with $P_n^{(\alpha,\beta)'}(1)$ and

$$r(n) := P_n^{(\alpha,\beta)''}(1) = \frac{n(n-1)(n+\alpha+\beta+1)(n+\alpha+\beta+2)}{4(\alpha+1)(\alpha+2)}$$

as in the proof of Lemma 3.2. Thus,

$$\begin{aligned}
&P_n^{(\alpha,\beta)}\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right)\\
&\quad = 1 - \left(\frac{t^2}{2k} - \frac{t^4}{24k^2} + O\left(\frac{1}{k^3}\right)\right)q(n) + \left(\frac{t^2}{2k} + O\left(\frac{1}{k^2}\right)\right)^2 \frac{r(n)}{2} + O\left(\frac{1}{k^3}\right)\\
&\quad = 1 - \frac{t^2 q(n)}{2k} + \frac{t^4}{k^2}\left(\frac{q(n)}{24} + \frac{r(n)}{8}\right) + O(1/k^3)
\end{aligned}$$

for $k \to \infty$. Hence, with $m_4 := \int_0^\pi t^4\,d\mu(t)$,

$$\int_0^\pi P_n^{(\alpha,\beta)}\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right)d\mu(t) = 1 - \frac{\sigma^2 q(n)}{2k} + \frac{m_4}{k^2}\left(\frac{q(n)}{24} + \frac{r(n)}{8}\right) + O(1/k^3).$$

Therefore, the power series

$$(1 - at + bt^2)^{1/t} = e^{-a}\left(1 + (b - a^2/2)t + O(t^2)\right) \quad \text{for } t \to 0$$

leads to

$$\left(\int_0^\pi P_n^{(\alpha,\beta)}\left(\cos\left(\frac{t}{\sqrt{k}}\right)\right)d\mu(t)\right)^k = \exp\left(\frac{-\sigma^2 q(n)}{2}\right)\left(1+\frac{R(n)}{k}+O(1/k)\right) \tag{3.24}$$

for $k \to \infty$ with

$$R(n) := m_4\left(\frac{q(n)}{24}+\frac{r(n)}{8}\right) - \frac{\sigma^4 q(n)^2}{8}. \tag{3.25}$$

Using the definitions of $q(n)$ and $r(n)$ as well as $\alpha \geq \beta \geq -1/2$, we readily obtain that for all possible values of $\sigma^2 > 0$ and $m_4 > 0$ there exists some $n \geq 1$ with $R(n) \neq 0$. Hence, (2.2) yields

$$\|\mu_k^{(k)} - \nu_{\sigma^2}\| \geq \sup_{n\in\mathbb{N}} |\widehat{\mu}_k^{(k)}(n) - \widehat{\nu}_{\sigma^2}(n)| = O(1/k)$$

as claimed. □

REFERENCES

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

[2] Bingham, N. H. (1972). Random walks on spheres. *Z. Wahrsch. Verw. Gebiete* 22 169–192.

[3] Bloom, W. R. and Heyer, H. (1994). *Harmonic Analysis of Probability Measures on Hypergroups.* De Gruyter, Berlin.

[4] Bochner, S. (1979). Positivity of the heat kernel for ultraspherical polynomials and similar functions. *Arch. Rational Mech. Anal.* 70 211–217.

[5] Clerc, J. L. and Roynette, B. (1979). Un theoreme central-limite. In *Analyse Harmonique sur les groupes de Lie. Lecture Notes in Math.* 739 122–131. Springer, Berlin.

[6] Diaconis, P. (1988). *Group Representations in Probability and Statistics.* IMS, Hayward, CA.

[7] Diaconis, P. and Shashahani, M. (1986). Products of random matrices as they arise in the study of random walks on groups. *Contemp. Math.* 50 183–195.

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

[9] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* 2, 2nd ed. Wiley, New York.

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

[11] Gasper, G. (1972). Banach algebras for Jacobi series and positivity of a kernel. *Ann. Math.* 95 261–280.

[12] Hartman, P. and Watson, G. S. (1974). "Normal" distribution functions on spheres and the modified Bessel functions. *Ann. Probab.* 2 593–607.

[13] Helgason, S. (1962). *Differential Geometry and Symmetric Spaces.* Academic Press, New York.

[14] Jewett, R. I. (1975). Spaces with an abstract convolution of measures. *Adv. in Math.* 18 1–101.

[15] Lasser, R. (1983). Orthogonal polynomials and hypergroups. *Rend. Math. Appl.* 3 185–209.

[16] Lasser, R. (1987). Convolution semigroups on hypergroups. *Pacific J. Math.* 127 353–371.

[17] Mueller, C. and Weissler, F. B. (1982). Hypercontractivity for the heat semigroup for ultraspherical polynomials and on the n-sphere. *J. Funct. Anal.* 48 252–283.

[18] ROSENTHAL, J. (1994). Random rotations: characters and random walks on $SO(N)$. *Ann. Probab.* 22 398–423.

[19] ROSS, K. A. and XU, D. (1994). Hypergroup deformations and Markov chains. *J. Theoret. Probab.* 7 813–830.

[20] SALOFF-COSTE, L. (1994). Precise estimates on the rate of which certain diffisions tend to equilibrium. *Math. Z.* 217 641–677.

[21] SZEGÖ, G. (1959). *Orthogonal Polynomials.* Amer. Math. Soc., Providence, RI.

[22] TRIMECHE, K. (1978). Probabilités indéfiniment divisibles et théoreme de la limite centrale pour une convolution généralisée sur la demi-droite. *C. R. Acad. Sci. Paris. Ser. A* 286 63–66.

[23] VOIT, M. (1995). A central limit theorem for isotropic random walks on n-spheres for $n \to \infty$. *J. Math. Anal. Appl.* 189 215–224.

[24] VOIT, M. (1995). Central limit theorems for Jacobi hypergroups. In *Applications of Hypergroups and Related Measure Algebras. Contemp. Math.* 183 395–411.

[25] VOIT, M. (1996). Asymptotic distributions for the Ehrenfest urn and related random walks. *J. Appl. Probab.* 33 340–356.

[26] VOIT, M. (1996). Limit theorems for compact two-point homogeneous spaces of large dimensions. *J. Theoret. Probab.* 9 353–370.

[27] WANG, H. (1952). Two-point homogeneous spaces. *Ann. Math.* 55 177–191.

[28] WATSON, G. S. (1983). *Statistics on Spheres.* Wiley, New York.

[29] WHITTAKER, E. T. and WATSON, G. N. (1935). *A Course of Modern Analysis.* Cambridge Univ. Press.

[30] ZEUNER, HM. (1989). The central limit theorem for Chebli–Trimeche hypergroups. *J. Theoret. Probab.* 2 51–63.

[31] ZEUNER, H. (1989). One-dimensional hypergroups. *Adv. Math.* 76 1–18.

MATHEMATISCHES INSTITUT
UNIVERSITÄT TÜBINGEN
AUF DER MORGENSTELLE 10
72076 TÜBINGEN
GERMANY
E-MAIL: voit@mailserv.2dv.uni-teubingen.de