A Large Deviation Principle for Polynomial Hypergroups

1991 MSC Subject classification: 60F10, 60J10, 43A62

Abstract

Let $S_n,\ n \geqslant 1$, be a random walk on a polynomial hypergroup $(\mathbb{N}_0, *)$, i.e. a Markov chain on the nonnegative integers with stationary transition probabilities $P_{ij} = \delta_i * \mu(\{j\})$ where μ is a fixed probability measure on $\mathbb{N}_0$. Under certain conditions on this measure the principle of large deviations is shown for the distributions of S_n/n. This result comprises the large deviation principle for birth and death random walks associated with the polynomials generating the polynomial hypergroup.

1 Introduction

Random walks on polynomial hypergroups provide a unified frame for studying isotropic random walks on a variety of algebraic structures, comprising random walks on free groups, on homogeneous trees, on the dual spaces of the Gelfand pairs $(SO(n),\, SO(n-1))$ ($n \geqslant 3$) and on the dual of the compact group $SU(2)$ (see the survey [8]). Limit theorems for such random walks include laws of large numbers, central limit theorems and laws of the iterated logarithm (see [17] – [20]).

The purpose of this paper is to derive another limit theorem (Theorem 1), namely an analogue of Cramér's theorem concerning large deviations for sums of independent identical distributed random variables (see for instance [16], Section 3). A special case of the result presented here will be the large deviation principle for random walks on the dual of $SU(2)$ proved in [2]. A related result on random walks on the dual of an arbitrary compact semisimple Lie group is given in [5].

The paper is organized as follows: After recalling some basic facts concerning random walks on polynomial hypergroups and the abstract large deviation principle, the main result is stated in section 2. In section 3, a moment generating function for polynomial hypergroups is introduced which will be used in section 4, together with a technical result on bounds of the orthogonal polynomials outside the interval of orthogonality, to prove the main theorem. The explicit form of the rate function for birth and death random walks is calculated in section 5. Finally, we present some examples.

2 The main result

Polynomial hypergroups. This paragraph describes the class of polynomial hypergroups to be used later. Details on (polynomial) hypergroups in general and many examples may be found in [8],[10] and [13].

Let $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, $(c_n)_{n\in\mathbb{N}}$ be sequences of real numbers satisfying $a_n,\, c_n > 0, b_n \geqslant 0$ and $a_n + b_n + c_n = 1$ $(n \in \mathbb{N})$. Assume further that there exist $\alpha := \lim_{n\to\infty} a_n \in]0,\,1[$, $\beta := \lim_{n\to\infty} b_n$ and $\gamma := \lim_{n\to\infty} c_n \in]0,\,1[$, so $\alpha+\beta+\gamma = 1$. Note that polynomial hypergroups can be defined without this restriction (see [10]). Using Favard's theorem (see, for instance [3], Theorem 1.4.4) we define a sequence of orthogonal polynomials by

$$\begin{aligned} P_0(x) = 1, \qquad & P_1(x) = 2\sqrt{\alpha\gamma}x + \beta \\ P_1(x)P_n(x) \quad = \quad & a_n P_{n+1}(x) + b_n P_n(x) + c_n P_{n-1}(x). \end{aligned}$$

$x_0 := \frac{1-\beta}{2\sqrt{\alpha\gamma}} = \frac{\alpha+\gamma}{2\sqrt{\alpha\gamma}} \geqslant 1$ is a normalization constant giving $P_0(x_0) = 1$, and then $P_n(x_0) = 1$ for all $n \in \mathbb{N}$ by induction.

This normalization of the polynomials differs from the normalization used in [10] and is usually employed when dealing with probabilistic limit theorems on polynomial hypergroups (see [17]–[19]).

Suppose that all the linearization coefficients $g(m, n, k)$, defined by

$$P_n(x)P_m(x) = \sum_{k=|n-m|}^{n+m} g(m,n,k)P_k(x),$$

are nonnegative. If we define a convolution of point measures on $\mathbb{N}_0$ by

$$\delta_m * \delta_n = \sum_{k=|n-m|}^{n+m} g(m,n,k)\delta_k \tag{2.1}$$

$\mathbb{N}_0$ becomes a commutative hypergroup $(\mathbb{N}_0, *)$ with 0 as unit element and the identity as involution and is called a *polynomial hypergroup*.

On any polynomial hypergroup there exists a *Haar measure* m (i.e. a positive measure m satisfying $\delta_k * m = m$ for every $k \in \mathbb{N}_0$), which is uniquely determined by $m(\{0\}) = 1$. Note that the positivity of the linearization coefficients implies that $h(n) := m(\{n\}) \geqslant 1$ for every n and $\alpha \geqslant \gamma$.
By our assumptions the orthogonality measure π of the polynomials $P_n(x)$ is in Nevai's class $M(0, 1)$ (see [12]) and we have $\alpha > \gamma$ if and only if $x_0 \notin \operatorname{supp} \pi$.

For any probability measure μ on $\mathbb{N}_0$ its *Fourier transform* $\hat{\mu}(x)$ is the continuous real-valued function

$$D_S \to \mathbb{R}, \quad \hat{\mu}(x) = \sum_{k=0}^{\infty} \mu(\{k\}) P_k(x),$$

where $D_S := \{x \in \mathbb{R} \mid |P_n(x)| \leq 1 \; \forall \, n \in \mathbb{N}\}$. $\hat{\mu}(x)$ is also uniquely defined (with the possible value $+\infty$) for any x with $P_n(x) \geqslant 0$ for every n.
Let $\mu, \nu \in M^1(\mathbb{N}_0)$ be probability measures. Then it is easily seen that $\widehat{\mu * \nu}(x) = \hat{\mu}(x) \cdot \hat{\nu}(x)$ if $x \in D_S$ or if $P_n(x) \geqslant 0$ for every n and $\hat{\mu}(x), \hat{\nu}(x)$ are both finite.
Random walks. We shall study Markov chains S_n with stationary transition probabilities on $\mathbb{N}_0$ which are homogeneous with respect to the convolution of the polynomial hypergroup in the following sense:
$S_0 = 0$ and there exists a probability measure μ on $\mathbb{N}_0$ such that

$$P(S_n = j \mid S_{n-1} = i) = \delta_i * \mu(\{j\}) \quad (n \in \mathbb{N}, i, j \in \mathbb{N}_0).$$

A Markov chain fulfilling these conditions is called *random walk* with law μ. For general properties of such random walks we refer to [7] and [8]. Note that in the case of $\mu = \delta_1$ this definition agrees with that given in [9]. It is immediate from the definition that the distribution of the variables S_n is given by the n-fold convolution product $\mu^{(n)}$.
The abstract large deviation principle. Consider a sequence F_n of probability measures on a polish space E converging weakly to a degenerate distribution at some point $x_0 \in E$ (in the main theorem of this paper F_n will be the distribution of S_n/n and E the interval $[0, \infty[$).
The abstract definition of the large deviation principle is as follows (see [16]):
Definition:
Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures on a polish space E and $\{a_n\}_{n\in\mathbb{N}}$ a divergent sequence of positive numbers. We say that $\{F_n\}$ satisfies the *large deviation principle* with constants $\{a_n\}$ and rate function $I : E \to [\,0, \infty\,[$, if the following conditions hold:

(i) I is lower semicontinuous and has compact level sets, i.e. for each $m \geqslant 0$ $\{x \mid I(x) \leq m\}$ is compact.

(ii) For each closed subset A of E

$$\limsup_{n\to\infty} \frac{1}{a_n} \log F_n(A) \leq - \inf_{x\in A} I(x).$$

(iii) For each open subset G of E

$$\liminf_{n\to\infty} \frac{1}{a_n} \log F_n(G) \geqslant - \inf_{x\in G} I(x).$$

Here and in the following log denotes the natural logarithm.

We need the following constant θ_0, related to the normalization constant x_0 by

$$\begin{aligned}\theta_0 &= \cosh^{-1} x_0 && (x_0 \geqslant 1)\\ &= \log(x_0 + \sqrt{x_0^2 - 1}) = \log(\tfrac{\alpha+\gamma+|\alpha-\gamma|}{2\sqrt{\alpha\gamma}})\\ &= \log\sqrt{\tfrac{\alpha}{\gamma}} \geqslant 0 && (\alpha \geqslant \gamma).\end{aligned}$$

Now we can state the main result of this paper:

Theorem 1. *Let* $(\mathbb{N}_0, *)$ *be a polynomial hypergroup as defined above and let* μ *be a probability measure on* $\mathbb{N}_0$ *with finite support. Denote by* F_n *the distribution of* S_n/n $(n \in \mathbb{N})$ *where* S_n *is a random walk with law* μ.
Then the sequence $\{F_n\}_{n\in\mathbb{N}}$ *of probability measures satisfies the principle of large deviations with constants {n} and the convex rate function*

$$I(x) = \begin{cases} +\infty & \text{if } x \notin [\,0,\,k\,] \\ \sup_{t\geqslant -\theta_0}\{tx - c(t)\} & \text{if } x \in [\,0,\,k\,] \end{cases},$$

where k is the right endpoint of the support of μ, θ_0 *is as above and*

$$c(t) = \log \hat{\mu}(\cosh(t + \theta_0)) \text{ for } t \geqslant -\theta_0 \ .$$

Furthermore, $E_*(X)$ *is the unique minimum point of* $I(x)$ *where* X *is a* $\mathbb{N}_0$*-valued random variable with law* μ *and* $E_*(X)$ *is defined as in section 3.*

Imposing stronger assumptions on the polynomial hypergroup we can prove a complete analogue of Cramér's Theorem for measures with possibly infinite support. The additional condition in the following Corollary is satisfied e.g. by Tchebichef polynomials of the first kind or by Griñspun polynomials (see section 6 or [10], Example 3.(g)(ii)) and their two parameter extension (see [11]).

Corollary 1. *Let* $(\mathbb{N}_0, *)$ *be a polynomial hypergroup as defined above and further assume that the Haar weights* $h(n) := m(\{n\})$ *are uniformly bounded.*
Let μ *be a probability measure on* $\mathbb{N}_0$ *with* $M_\mu(t) := \hat{\mu}(\cosh t) < \infty$ *for every* $t \in \mathbb{R}$ *and denote by* F_n *the distribution of* S_n/n $(n \in \mathbb{N})$ *where* S_n *is a random walk with law* μ.
Then the sequence $\{F_n\}_{n\in\mathbb{N}}$ *of probability measures satisfies the principle of large deviations with constants {n} and the convex rate function*

$$I(x) = \begin{cases} +\infty & \text{if } x < 0 \\ \sup_{t\geqslant 0}\{tx - \log M_\mu(t)\} & \text{if } x \geqslant 0 \end{cases}.$$

Furthermore, $x = 0$ *is the unique minimum point of* $I(x)$.

Remarks. (i) In the Theorem we have $\mathrm{domI} = \{x : I(x) < \infty\} \subset [\,0,\,k\,]$. This agrees with the fact that for S_n as in the Theorem we have $P(\,S_n/n \in [\,0,\,k\,]\,) = 1$ for every n.

(ii) The Theorem can be regarded as a result on the rate of convergence in the weak law of large numbers: S_n/n converges in probability exponentially fast to its limit (see [6], Theorem II.6.3).

(iii) The unique minimum point of $I(x)$ is strictly positive if and only if $x_0 \notin \operatorname{supp} \pi$. This seems to be the only property of the orthogonality measure π that matters for our result. Theorem 4 below shows this at least for the birth and death random walk.

(iv) A related result for Sturm-Liouville hypergroups on $[0, \infty[$ will be presented in a forthcoming paper.

3 The (modified) moment generating function

In this section the analogue of the moment generating function for polynomial hypergroups will be defined. For this, the following definition of (modified) moments on polynomial hypergroups introduced by Voit [17] is needed (see also [20] for a discussion of moment functions on arbitrary hypergroups).
Let the functions $\varphi_{n,\theta}$ and m_n $(\theta \in \mathbb{C}, n \in \mathbb{N})$ be defined by

$$\varphi_{n,\theta}(k) := \left(\frac{\partial}{\partial t}\right)^n P_k(\cosh t)|_{t=\theta} \quad \text{and} \quad m_n(k) := \varphi_{n,\theta_0}(k).$$

The functions m_n are called *moments*.
For any $\mathbb{N}_0$-valued random variable X with law μ

$$E_*(X) := E(m_1(X)) = \sum_{k=0}^{\infty} m_1(n)\mu(\{n\})$$

is called the *modified expectation* of X (E denotes the usual expectation of a random variable). It has the following properties ($n \in \mathbb{N}$):

(i) $m_1(n) \equiv 0$ if $\alpha = \gamma$ and $m_1(n) > 0$ if $\alpha > \gamma$.

(ii) $E_*(S_n) = nE_*(X)$ for each random walk with law μ, where X is a random variable with law μ.

Proofs of these facts may be found in [17] or [20].

The following Lemma will be used in section 4 to replace e^{nt} by $P_n(\cosh(t + \theta_0))$ for $t \geqslant -\theta_0$.

Lemma 1. *Let the orthogonal polynomials $P_n(x)$ and θ_0 be defined as above.*

(i) For $t \geqslant 0$

$$\gamma_n e^{nt} \leq P_n(\cosh(t+\theta_0)) \leq e^{nt}$$

where $0 < \gamma_n \leq 1/2$ is monotonically decreasing and $\lim_{n\to\infty} \frac{1}{n} \log \gamma_n = 0$.
If furthermore the Haar weights are uniformly bounded there exists a constant $0 < \gamma \leq 1/2$ with

$$\gamma e^{nt} \leq P_n(\cosh t) \leq e^{nt}.$$

(ii) Let $\alpha > \gamma$ i.e $\theta_0 > 0$. For $-\theta_0 \leq t < 0$

$$e^{nt} \leq P_n(\cosh(t+\theta_0)) \leq \delta_n e^{nt}$$

where $\delta_n \geqslant 2$ is monotonically increasing and $\lim_{n\to\infty} \frac{1}{n} \log \delta_n = 0$.

Proof. The assumptions imply the following property (T) (see [12]):

$$P_n(x) = \sum_{k=0}^{n} h(n,k) T_k(x)$$

for every $x \in \mathbb{R}$, where $h(n,k) \geqslant 0$, $\sum_{k=0}^{n} h(n,k) \cosh k\theta_0 = 1$ and $T_n(x)$ denote the Tchebichef polynomials of the first kind.

(i) For $t \geqslant 0$ we have

$$\frac{1}{2} e^{k(t+\theta_0)} \leq \cosh k(t+\theta_0) \leq \cosh k\theta_0 e^{kt}.$$

Thus

$$P_n(\cosh(t+\theta_0)) = \sum_{k=0}^{n} h(n,k) \cosh k(t+\theta_0) \leq \sum_{k=0}^{n} h(n,k) \cosh k\theta_0 \; e^{kt} \leq e^{nt}$$

and

$$P_n(\cosh(t+\theta_0)) \geqslant \frac{1}{2} \sum_{k=0}^{n} h(n,k) \; e^{k(t+\theta_0)} \geqslant \frac{1}{2} h(n,n) \left(\sqrt{\alpha/\gamma}\right)^n e^{nt}.$$

Denoting the leading coefficient of $P_n(x)$ by σ_n it follows immediately that

$$h(n,n) = \frac{\sigma_n}{2^{n-1}} = \frac{2(\sqrt{\alpha\gamma})^n}{\prod_{k=1}^{n-1} a_k}$$

and thus

$$\frac{1}{2} h(n,n) \left(\sqrt{\alpha/\gamma}\right)^n = \frac{\alpha^n}{\prod_{k=1}^{n-1} a_k}.$$

If this sequence is not monotonically decreasing, let $\sigma_k := \sup_{n\geqslant k} a_n$ and

$$\gamma_n := \frac{\alpha^n}{\prod_{k=1}^{n-1}\sigma_k},$$

which has the desired properties since $\sigma_k \to \alpha$ for $k \to \infty$.
Now suppose that $h(n) \leq M$. Then it is easily seen that $\alpha = \gamma$ and $\theta_0 = 0$. Denoting by $p_n(x)$ the corresponding orthonormal polynomials we have for every $t \geqslant 0$

$$\frac{p_n(\cosh t)}{e^{nt}} = \sqrt{h(n)}\frac{P_n(\cosh t)}{e^{nt}} \leq \sqrt{M}.$$

Thus the orthogonality measure corresponding to the polynomials $P_n(x)$ satisfies Szegö's condition on $[-1, 1]$ (see [15],p. 247), and

$$h(n,n) = \frac{\sigma_n}{2^{n-1}} = \frac{\lambda_n}{2^{n-1}\sqrt{h(n)}} \geqslant \frac{\lambda_n}{2^{n-1}\sqrt{M}},$$

where λ_n denotes the leading coefficient of $p_n(x)$. Since $\lambda_n/2^n$ tends to a positive limit ([15],Theorem 3.5) there exists a constant $\gamma > 0$ with $h(n,n)/2 \geqslant \gamma$.

(ii) For $-\theta_0 \leq t < 0$ we have $\cosh k(t+\theta_0) \geqslant \cosh k\theta_0 e^{kt}$ and thus $P_n(\cosh(t+\theta_0)) \geqslant e^{nt}$. Now define orthogonal polynomials $Q_n(x)$ by

$$Q_n(x) := \frac{P_n(x)}{P_n(1)}.$$

The coefficients $\tilde{a}_n, \tilde{b}_n, \tilde{c}_n$ of the 3-term-recurrence-relation of $Q_n(x)$ satisfy $\tilde{a}_n \to \tilde{\alpha}, \tilde{b}_n \to \tilde{\beta}$ and $\tilde{c}_n \to \tilde{\gamma}$ with $\tilde{\alpha} = \tilde{\gamma}$ (see [18], 2.12). Using part (i) we get

$$P_n(\cosh(t+\theta_0)) = \frac{Q_n(\cosh(t+\theta_0))}{Q_n(\cosh\theta_0)} \leq \tilde{\gamma}_n^{-1}\ e^{-n\theta_0}\ e^{n(t+\theta_0)} \leq \tilde{\gamma}_n^{-1}\ e^{nt}.$$

The proof is completed by setting $\delta_n = \tilde{\gamma}_n^{-1}$. □

Remark. In general, it is impossible to replace the lower bound γ_n in part (i) by a strictly positive constant whenever the Haar weights are unbounded. A counterexample is provided by the Tchebichef polynomials of the second kind. These have the explicit representation

$$P_n(\cos t) = P_n^{(1/2,1/2)}(\cos t) = \frac{\sin(n+1)t}{(n+1)\sin t}$$

and consequently

$$P_n(\cosh t) = \frac{\sinh(n+1)t}{(n+1)\sinh t} = e^{nt}\frac{1-e^{-2(n+1)t}}{(n+1)(1-e^{-2t})}.$$

For $\mu \in M^1(\mathbb{N}_0)$ define its *(usual) moment generating function* $\Lambda_\mu(t)$ by

$$\Lambda_\mu(t) = \int_{\mathbb{N}} e^{tx}\, d\mu(x).$$

Theorem 2. *Let $(\mathbb{N}_0, *)$ be a polynomial hypergroup as defined in section 2 and let μ be a probability measure on $\mathbb{N}_0$ for which there exists $t_0 > \theta_0$ with $\Lambda_\mu(t_0) < \infty$. Then the function $M_\mu(t) : [-\theta_0,\, t_0 - \theta_0] \to]\,0, \infty[$ defined by*

$$t \mapsto \hat{\mu}(\cosh(t+\theta_0)) = \sum_{k=0}^{\infty} \mu(\{k\}) P_k(\cosh(t+\theta_0))$$

is of class C^∞ and

$$E(m_n(X)) = \mu_n = \left(\frac{\partial}{\partial t}\right)^n \hat{\mu}(\cosh(t+\theta_0))|_{t=0} = \left(\frac{\partial}{\partial t}\right)^n M_\mu(t)|_{t=0}$$

for every $n \in \mathbb{N}$.

Proof. First note that $\Lambda_\mu(t) < \infty$ for every $t \leq t_0$. Thus $M_\mu(t)$ is well defined and finite by Lemma 1. For $t \geqslant -\theta_0$ property (T) yields

$$\begin{aligned}
\varphi_n(k, t+\theta_0) &= \sum_{l=0}^{k} h(k,l)\, l^n\, 1/2(\, e^{l(t+\theta_0)} + (-1)^n e^{-l(t+\theta_0)}) \\
&\leq \sum_{l=0}^{k} h(k,l)\, l^n\, e^{l(t+\theta_0)} \leq k^n\, e^{k(t+\theta_0)}.
\end{aligned}$$

Since the moment generating function is finite on $(0,\ t_0)$ it is analytic on $(0,\ t_0)$ with

$$\left(\frac{\partial}{\partial t}\right)^n \Lambda_\mu(t) = \int_0^\infty x^n\, e^{tx}\, d\mu(x)$$

for $t \in (0,\ t_0)$. This means that $k \mapsto k^n\, e^{kt} \in L^1(\mu)$ for every $t \in (0,\ t_0)$ and $n \geqslant 1$. By Lebesgue's dominated convergence theorem, the proof is complete. □

Remarks. (i) The Theorem shows that on polynomial hypergroups the function $M_\mu(t) = \hat{\mu}(\cosh(t+\theta_0))$ $(t \geqslant -\theta_0)$ is the natural analogue of the usual moment generating function. Therefore, $M_\mu(t)$ will be called *(modified) moment generating function* of the measure μ. Since $f(t) = \hat{\mu}(\cos(t+i\theta_0))$ is the Fourier transform of the measure μ, we have $M_\mu(t) = f(it)$ $(t \geqslant -\theta_0)$ as in the classical case.

(ii) The conclusions of this Theorem are comparable to the differentiability properties of the Fourier transform of the measure μ (see Theorem 1 and Theorem 2 in [17]).

4 Proof of the large deviation principle

The following theorem of Ellis will be used:

Theorem 3 ([6], Theorem II.6.1)**.** *Let W_n be an arbitrary sequence of random variables with values in $\mathbb{R}$ and (a_n) a divergent sequence of positive numbers. Define for $t \in \mathbb{R}$:*

$$c_n(t) := \frac{1}{a_n} \log E(\exp(tW_n)).$$

Assume that

(a) Each $c_n(t)$ is finite for every $t \in \mathbb{R}$.

(b) $c(t) := \lim_{n\to\infty} c_n(t)$ exists, is finite and differentiable for every $t \in \mathbb{R}$.

Then the distributions F_n of W_n/a_n satisfy the large deviation principle with constants (a_n) and the convex rate function

$$I(x) = \sup_{t\in\mathbb{R}}\{tx - c(t)\}.$$

Thus, for proving Theorem 1 it suffices to check the conditions of Theorem 3. This is partly accomplished by the following Lemma.

Lemma 2. *Let μ be a probability measure on $\mathbb{N}_0$ with $\Lambda_\mu(t) < \infty$ for every $t \in \mathbb{R}$ and let S_n be a random walk with law μ. Define the functions $c_n(t)$ for S_n with $a_n = n$ as in Theorem 3. Then*

(i) Each $c_n(t)$ is finite for every $t \in \mathbb{R}$.

(ii) $c(t) = \lim_{n\to\infty} c_n(t)$ exists for every $t \in \mathbb{R}$ and is finite.

Proof. We have for $m, n \in \mathbb{N}$ (see Equation (2.1)):

$$\exp(t(m * n)) \le \exp(t(m+n)) \quad t \geqslant 0, \tag{4.1}$$

$$\exp(t(m * n)) \geqslant \exp(t(m+n)) \quad t \le 0, \tag{4.2}$$

and

$$\int_0^\infty e^{tx}\, d\mu^{(n+m)}(x) = \int_0^\infty \exp(t(x*y))\, d\mu^{(m)}(x) d\mu^{(n)}(y), \tag{4.3}$$

where $\exp(t(x*y)) := \int e^{tz}\, d(\delta_x * \delta_y)(z)$.

(i) Thus using (4.1)–(4.3) and Jensen's inequality we have for $t < 0$

$$0 \geqslant c_n(t) \geqslant 1/n \log(\Lambda_\mu(t))^n = \log \Lambda_\mu(t) \geqslant t \int_0^\infty x\, d\mu(x) > -\infty$$

and analogously for $t \geqslant 0$

$$0 \le c_n(t) \le \log \Lambda_\mu(t) < +\infty.$$

(ii) By (4.1)–(4.3) the function $f(n) = \log \Lambda_{\mu^{(n)}}(t)$ is subadditive for each fixed $t \geqslant 0$ and the function $g(n) = -\log \Lambda_{\mu^{(n)}}(t)$ is subadditive for each fixed $t \leq 0$. The conclusion follows from Lemma 3.1.3 on subadditive functions in [4]. □

Proof of Theorem 1:
By assumption there exists a $k \in \mathbb{N}_0$ with $\mu = \sum_{j=0}^{k} \mu_j \, \delta_j$ and it is obvious that $\Lambda_\mu(t) < \infty$ for every $t \in \mathbb{R}$. By Lemma 2 it remains to prove the differentiability of $c(t)$ for every $t \in \mathbb{R}$ and the form of $I(x)$.

(i) First let $t \geqslant 0$. By means of Lemma 1 we obtain for $n \in \mathbb{N}$ and $0 \leq j \leq nk$:

$$\gamma_{nk} \, e^{jt} \leq P_j(\cosh(t+\theta_0)) \leq e^{jt}.$$

This yields

$$\Lambda_{\mu^{(n)}}(t) \geqslant \sum_{j=0}^{nk} P_j(\cosh(t+\theta_0)) \, \mu^{(n)}(\{j\}) = (\hat{\mu}(\cosh(t+\theta_0)))^n \geqslant \gamma_{nk} \Lambda_{\mu^{(n)}}(t)$$

and thus

$$0 \leq c_n(t) - \log M_\mu(t) \leq -1/n \log \gamma_{nk}.$$

By Lemma 1 this means $c(t) = \log M_\mu(t)$ for $t \geqslant 0$.
Next let $-\theta_0 \leq t < 0$. As above, we obtain

$$0 \leq \log M_\mu(t) - c_n(t) \leq 1/n \log \delta_{nk}$$

and $c(t) = \log M_\mu(t)$ for $-\theta_0 \leq t < 0$.
Finally let $t < -\theta_0$. We have $c(t) \leq 0$ for $t \leq 0$ and furthermore $c_+^{'}(-\theta_0) = 0$ since

$$c_+^{'}(-\theta_0) = \frac{\sum_{j=0}^{k} \mu_j \varphi_{1,0}(j)}{\hat{\mu}(1)}$$

and

$$\varphi_{1,0}(j) = \begin{cases} \frac{\partial}{\partial t} P_j(\cosh t)|_{t=0} & \text{if } \alpha = \gamma \\ P_j(1) \frac{\partial}{\partial t} Q_j(\cosh t)|_{t=0} & \text{if } \alpha > \gamma, \end{cases}$$

where $Q_j(x)$ is defined as in the proof of Lemma 1(ii). Now [17], Lemma 3.2.3 shows $\varphi_{1,0}(j) = 0$ for every j. Thus for $t \leq -\theta_0$, $c(t)$ is a finite convex function bounded to the left of the minimum point $t = -\theta_0$. Hence it is constant and $c(t) = c(-\theta_0)$ for every $t \leq -\theta_0$.

(ii) Next we prove the form of $I(x)$. For $t \geqslant 0$ we have $c(t) \leq \log \Lambda_\mu(t) \leq kt$ (see Lemma 2). But then

$$I(x) \geqslant \sup_{t \geqslant 0} \{tx - kt\} = +\infty$$

for $x > k$. As $c(t)$ is constant for $t \leq -\theta_0$ we get for $x < 0$

$$I(x) \geqslant \sup_{t \leq -\theta_0} \{tx - c(t)\} = \sup_{t \leq -\theta_0} \{tx\} - c(-\theta_0) = +\infty.$$

Finally let $0 \leq x \leq k$. Then $tx - c(t) \leq -\theta_0 x - c(-\theta_0)$ for every $t \leq -\theta_0$. This means

$$I(x) = \sup_{t \in \mathbb{R}} \{tx - c(t)\} = \sup_{t \geqslant -\theta_0} \{tx - c(t)\}.$$

$E_*(X)$ is the unique minimum point of $I(x)$ by Theorem 2 and [6], Theorem II.6.3. □

Proof of Corollary 1:
Under the assumption of the Corollary we have $\alpha = \gamma$ and by Lemma 1

$$\gamma e^{nt} \leq P_n(\cosh(t)) \leq e^{nt}$$

for $t \geqslant 0$ $(n \in \mathbb{N})$ where $0 < \gamma \leq 1/2$. Thus $\Lambda_\mu(t) \leq 1/\gamma \ \hat{\mu}(\cosh t) < \infty$. As in the proof of Theorem 1 we obtain

$$c(t) = \begin{cases} 0 & \text{if } t < 0 \\ \log M_\mu(t) & \text{if } t \geqslant 0 \end{cases}.$$

This yields the conclusion as in the proof of Theorem 1. □

5 The rate function for birth and death random walks

In the special case that $\mu = \delta_1$ the random walk is called a *birth and death random walk* (see [9]). Its rate function will now be calculated. This theorem may be compared to [4], Exercise 1.2.11.

Theorem 4. *Let $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}, (c_n)_{n\in\mathbb{N}}$ be sequences of real numbers satisfying the assumptions of section 2. Denote by S_n the random walk with the transition probabilities*

$$P(S_{n+1} = j|\ S_n = i) = \begin{cases} a_n & j = i + 1 \\ b_n & j = i \\ c_n & j = i - 1 \\ 1 & i = 0,\ j = 1 \\ 0 & |j - i| > 1. \end{cases}$$

Then the distributions F_n of S_n/n satisfy the principle of large deviations with constants $\{n\}$ and the rate function

$$I(x)=\begin{cases} +\infty & x\notin[0,\,1]\\ \log\frac{1}{2\sqrt{\alpha\gamma}+\beta} & x=0\\ x\log\dfrac{\left(\beta x+\sqrt{4\alpha\gamma+(\beta^2-4\alpha\gamma)x^2}\right)^2}{2\alpha\left(4\alpha\gamma(1-x)+\beta(\beta x+\sqrt{4\alpha\gamma+(\beta^2-4\alpha\gamma)x^2})\right)} & 0<x<1\\ +(1-x)\log\dfrac{(1-x)(\beta x+\sqrt{4\alpha\gamma+(\beta^2-4\alpha\gamma)x^2})}{4\alpha\gamma(1-x)+\beta(\beta x+\sqrt{4\alpha\gamma+(\beta^2-4\alpha\gamma)x^2})} & \\ \log\frac{1}{\alpha} & x=1. \end{cases}$$

Proof. The first part follows from Theorem 1. It remains to prove the form of the rate function for $0\le x\le 1$.
In the case of a birth and death random walk we have

$$c^{'}(t)=\begin{cases}\frac{\sinh(t+\theta_0)}{\cosh(t+\theta_0)+c/2} & t\geqslant-\theta_0\\ 0 & t\le-\theta_0\end{cases}.$$

where $c:=\frac{\beta}{\sqrt{\alpha\gamma}}$. Thus for $0<x<1$ the equation $c^{'}(t)=x$ has the unique solution

$$t(x)=\log\frac{\sqrt{\gamma}(cx+\sqrt{4(1-x^2)+c^2x^2})}{2\sqrt{\alpha}(1-x)}.$$

By [6],Theorem VI.5.3. we have $I(x)=xt(x)-\log(2\sqrt{\alpha\gamma}\cosh(t(x)+\theta_0)+\beta)$ and a straightforward calculation yields the form of $I(x)$ for $0<x<1$.
Finally $I(0)$ and $I(1)$ are easily obtained from the formulas $I(0)=\lim_{x\downarrow 0}I(x)$ and $I(1)=\lim_{x\uparrow 1}I(x)$ ([6], Theorem VI.3.2.). □

Remark. Note that in the case of a birth and death random walk the rate function depends only on the values of α,β and γ.

6 Examples

Jacobi polynomials. Let $a,b\in\mathbb{R}$ with $a\geqslant b>-1$ and $a+b+1\geqslant 0$. Define

$$a_n=\frac{2(n+a+b+1)(n+a+1)(a+b+2)}{(2n+a+b+2)(2n+a+b+1)2(a+1)}$$
$$b_n=\frac{a-b}{2(a+1)}\left(1-\frac{(a+b+2)(a+b)}{(2n+a+b+2)(2n+a+b)}\right)$$
$$c_n=\frac{2n(n+b)(a+b+2)}{(2n+a+b+1)(2n+a+b)2(a+1)}$$

These sequences induce polynomial hypergroups and the corresponding orthogonal polynomials are the *Jacobi polynomials* $P_n^{(a,b)}(x)$ (cf. [10], 3.(a)). We have

$$\alpha = \frac{a+b+2}{4(a+1)} = \gamma \qquad \text{and} \qquad \beta = \frac{a-b}{2(a+1)}$$

and consequently $x_0 = 1$ and $\theta_0 = 0$.

Random walks on the dual spaces of the Gelfand pairs $(SO(n), SO(n-1))$ $(n \geqslant 3)$**.** Given a Gelfand pair (G, K) the set S of all bounded spherical functions may be regarded as the dual space of (G, K). If G is compact, all bounded spherical functions are positive definite and we can define a probability preserving convolution $*$ on the set of bounded complex measures on S (see for example [14], section I.9). A random walk on S is then a stationary Markov chain X_n with transition probabilities

$$P(X_{n+1} \in A | X_n = x) = \delta_x * \mu(A) \quad (n \in \mathbb{N}, x \in S, A \subset S).$$

For $(G, K) = (SO(n), SO(n-1))$ S may be identified with $\{P_n^{(a,a)}(x) \,|\, n \in \mathbb{N}_0\ ,\ x \in [-1, 1]\}$ where $a = \frac{1}{2}(n-2)$ (see [1]) and the convolution agrees with the one introduced in section 2.

Grinspun polynomials $T_n(x; a)$ (cf. [10], 3(g)ii). These polynomials provide an example of a polynomial hypergroup with bounded Haar measure. Fix $a \geqslant 2$. Then the defining sequences $(a_n)_{n\in\mathbb{N}_0}$ and $(c_n)_{n\in\mathbb{N}_0}$ are given by

$$a_1 = \frac{a-1}{a} \quad c_1 = \frac{1}{a} \quad \text{and } a_n = \frac{1}{2} = c_n \quad n \geqslant 2.$$

Isotropic random walks on trees. Let T_{a+1} be a homogeneous tree of order $a+1$, that is an infinite connected graph, in which every node belongs to $a+1$ edges. Let Γ be the vertex set of T_{a+1} and let G be the automorphism group of Γ. An isotropic stationary random walk on T_{a+1} is a Markov chain $(\tilde{S}_n)_{n\in\mathbb{N}_0}$ on Γ that is starting at a vertex v_0 and whose transition probabilities

$$P(u, v) = P(\tilde{S}_{n+1} = v | \tilde{S}_n = u) \qquad (n \in \mathbb{N}_0, u, v \in \Gamma)$$

are preserved by all elements of G. This is equivalent to the condition

$$P(u, v) = \frac{1}{|\{w \in \Gamma : d(w, v) = d(u, v)\}|} \mu(\{d(u, v)\}) \tag{6.1}$$

where d is the usual metric on T_{a+1} and μ is a probability measure on $\mathbb{N}_0$.
If $\pi : \Gamma \to \mathbb{N}_0$ is defined by $\pi(w) = d(w, v_0)$, then $S_n := \pi(\tilde{S}_n)$ is a Markov chain on $\mathbb{N}_0$ and we have

$$P(S_{n+1} = l \,|\, S_n = k) = \frac{\sum_{w\in\Gamma, d(w,v_0)=k} P(\tilde{S}_n = w) P(d(\tilde{S}_{n+1}, v_0) = l \,|\, \tilde{S}_n = w)}{\sum_{w\in\Gamma, d(w,v_0)=k} P(\tilde{S}_n = w)}.$$

Since $P(\tilde{S}_n = w)$ and $P(d(\tilde{S}_{n+1}, v_0) = l \,|\, \tilde{S}_n = w)$ depend only on $d(w, v_0)$ and not on w itself, we obtain

$$(6.2) \qquad P(S_{n+1} = l \,|\, S_n = k) = P(d(\tilde{S}_{n+1}, v_0) = l \,|\, \tilde{S}_n = w)$$

for any $w \in \Gamma$ satisfying $d(w, v_0) = k$.
Denoting the stabilizer of v_0 by K we see that the state space of $(S_n)_{n \in \mathbb{N}_0}$ is actually the set of K-orbits of T_{a+1}. Denoting the canonical convolution on $G//K$ (see [14], section I.5 and I.7) by $*$, equations (6.1) and (6.2) imply

$$P(S_{n+1} = l \,|\, S_n = k) = \delta_k * \mu(\{l\}).$$

This convolution agrees with the convolution of the polynomial hypergroup generated by the sequences

$$a_n = \frac{a}{a+1} \qquad c_n = \frac{1}{a+1}$$

(cf. [18], section 5). The corresponding orthogonal polynomials are the *Cartier polynomials.*

ACKNOWLEDGEMENT:
It is a pleasure to thank M. Voit for suggesting this problem and for several helpful discussions. I would also like to thank the referee whose suggestions led to an improvement of the presentation of this paper.

References

[1] R. Askey, N. H. Bingham: Gaussian processes on compact symmetric spaces. Z. Wahrsch. verw. Geb. **37**, 127–143 (1976)

[2] W. Cegla, J.T. Lewis, G.A. Raggio: The free energy of quantum spin systems and large deviations. Comm. Math. Phys. **118**, 337–354 (1988)

[3] T.S. Chihara: An Introduction to Orthogonal Polynomials. New York: Gordon&Breach 1978

[4] J.D. Deuschel,D.W. Stroock: Large Deviations. New York: Academic Press 1989

[5] N.D. Duffield: A large deviation principle for the reduction of product representations. Proc. Amer. Math. Soc. **109**, 503–515 (1990)

[6] R.S. Ellis: Entropy, Large Deviations and Statistical Mechanics. Berlin Heidelberg New York: Springer 1985

[7] L. Gallardo, O. Gebuhrer: Marches aléatoires et hypergroupes. Expo. Math. **5**, 41–73 (1987)

[8] H. Heyer: Probability theory on hypergroups: A survey. In H. Heyer(ed.): Probability Measures on Groups VII. (Lect. Notes Math., vol. 1064, pp. 481–550) Berlin Heidelberg New York: Springer 1984

[9] S. Karlin, J. McGregor: Random walks. Ill. J. Math. **3**, 66–81 (1959)

[10] R. Lasser: Orthogonal polynomials and hypergroups. Rend. Math. **2**, 185–209 (1983)

[11] R. Lasser: Lacunarity with respect to orthogonal polynomials. Acta Sci. Math. **47**, 391–403 (1984)

[12] R. Lasser, M. Rösler: A note on property (T) of orthogonal polynomials. Arch. Math. **60**, 459–463 (1993)

[13] R. Lasser: Orthogonal polynomials and hypergroups II - the symmetric case. Trans. Amer. Math. Soc. **341**, 749 – 770 (1994)

[14] G. Letac: Problèmes classiques de probabilité sur un couple de Gelfand. In D. Duguè et al.(eds.): Analytical Methods in Probability Theory. (Lect. Notes Math., vol. 861, pp. 93–127) Berlin Heidelberg New York: Springer 1981

[15] D.S. Lubinsky: A survey of general orthogonal polynomials for weights on finite and infinite intervals. Acta Appl. Math. **10**, 237–296 (1987)

[16] S.R.S. Varadhan: Large Deviations and Applications. Philadelphia: SIAM 1984

[17] M. Voit: Laws of large numbers for polynomial hypergroups and some applications. J. Theoret. Probab. **3**, 245–266 (1990)

[18] M. Voit: Central limit theorems for random walks on $\mathbb{N}_0$ that are associated with orthogonal polynomials. J. Multiv. Anal. **34**, 290–322 (1990)

[19] M. Voit: A law of the iterated logarithm for a class of polynomial hypergroups. Mh. Math. **109**, 311–326 (1990)

[20] Hm. Zeuner: Moment functions and laws of large numbers on hypergroups. Math. Z. **211**, 369–407 (1992)