LIMIT THEOREMS FOR HYPERGROUPS ON $\mathbb{R}^+$

BY

HANS MARTIN ZEUNER

TECHNICAL REPORT NO. 274
MAY 1987

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT DMS86-00235

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
LIMIT THEOREMS FOR HYPERGROUPS ON IR+

BY

HANSMARTIN ZEUNER

TECHNICAL REPORT NO. 274
MAY 1987

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT DMS86-00235

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Limit theorems for hypergroups on $\mathbb{R}_+$

by

Hansmartin Zeuner

Mathematisches Institut
der Universität Tübingen
Auf der Morgenstelle 10
D-7400 Tübingen
Federal Republic of Germany
currently
Stanford University
Department of Statistics
Sequoia Hall
Stanford, California 94305
USA
1. Introduction

Being aware of the difficulties in formulating and proving a law of large numbers for random walks on an arbitrary locally compact group, it seems hopeless to try the same attempt on the even more general structure of a hypergroup. However, since these difficulties arise in part from the complicated geometric structure of many of the groups considered, one might expect that it is possible to obtain limit theorems on hypergroups which are of particularly simple geometry. This has successfully been done by Eymard, Roynette, Gallardo, and Ries (see Eymard, Roynette [7], Gallardo, Ries [14], and Gallardo [13]) in the case of the hypergroups on \( \mathbb{N} \) related to the Gegenbauer polynomials. Whereas on the real line (with the usual topology) there is exactly one structure as a topological group, there is an abundant collection of hypergroups on the half line \( \mathbb{R}_+ \) (see Chébli [5], Zeuner [28]). In the case of the Chébli-Trimèche hypergroups, where enough analytical tools are developed, the law of large numbers and a central limit theorem will be proved in this article.

If \( X_1, X_2, \ldots \) are i.i.d. random variables with values in a group, the corresponding random walk is the sequence \( S_1, S_2, \ldots \) defined by \( S_n = X_n X_{n-1} \cdots X_1 \). Since the operation on a hypergroup is only defined in terms of the convolution of measures, the random walk \( (S_n : n \geq 1) \) can only be defined by its distribution and not as a function of \( (X_n : n \geq 1) \). The notion of concretization and a randomized multiplication is introduced in 3.3 in order to obtain an explicit construction of \( (S_n : n \geq 1) \) in terms of \( (X_n : n \geq 1) \) for every 2nd countable locally compact hypergroup.

As in the classical case, the moments of a random variable are introduced, both to formulate the conditions under which a particular limit theorem holds, and to calculate the actual value of the limit. This has to be done by a modified definition to fit with the hypergroup operation. The first and second moments are in very close connection with the notion of the dispersion of a probability measure used in Faraut [8] and Trimèche [23]. As to be expected by the results of Guivarc'h [15], two different situations occur depending on the parameter \( \varrho \) which determines the growth of the hypergroup. If \( \varrho > 0 \) the hypergroup is of exponential growth and the expectation of every nonzero random variable is strictly positive. This result includes the symmetric spaces of rank one of non-compact type and corresponds to Guivarc'h [15], corollaire on page 77. If \( \varrho = 0 \) the expectation of every random variable is 0 and so is the limit of \( \frac{1}{n} S_n \) in probability. This result should be compared with Guivarc'h [15], théorème 3 on page 72.

In both cases a strong law of large numbers will be proved. Apart from the different and less general situation in this article the main difference with the results in Guivarc'h [15] is the fact that for hypergroups on \( \mathbb{R}_+ \) the law of large numbers can be formulated without the use of a gauge
function and the a.s. limit of $\frac{1}{n}S_n$ can be calculated explicitly.

Two different kinds of central limit theorems on $(\mathbb{R}_+, \ast)$ are considered in this article. The central limit theorem which states that the sum of many small random variables has approximately a Gaussian distribution has been proved by Trimèche [23] and is included (in a special case) into this article to point out the difference to another kind of central limit theorem which describes the asymptotic behaviour of the distribution of $S_n$ as $n \to \infty$. In the case of $(\mathbb{R}, +)$ there is no distinction between these two formulations of the central limit theorem since $a \cdot (X + Y) = aX + aY$, but this relation does not hold for the randomized addition on $\mathbb{R}_+$. It will be shown for every sequence $(X_n : n \geq 1)$ of i.i.d. random variables with finite variance that $S_n$ is asymptotically normal if $\theta > 0$ and that $\frac{1}{\sqrt{n}}S_n$ converges to a Rayleigh distribution under certain conditions on the hypergroup if $\theta = 0$.

The proofs are generalizations of the methods used in the classical case of $(\mathbb{R}, +)$. The main difference comes from the fact that the asymptotical properties of the characters and related functions which are used in the proofs are very well known for $(\mathbb{R}, +)$ and $\chi_\lambda(x) = e^{i\lambda x}$, but have to be shown in every particular situation in the case of a hypergroup.
2. Preliminaries

2.1. Let $K$ be a hypergroup in the sense of Jewett [19]; this means that $K$ is a locally compact space with an associative convolution $(x, y) \mapsto \varepsilon_x * \varepsilon_y \in \mathcal{M}^1(K)$ (the space of probability measures on $K$) such that there exist a neutral element $e \in K$ and an inversion $x \mapsto x'$ satisfying certain conditions (see Heyer [18], Jewett [19], or Spector [22] for details). In the cases considered in this article (except in the third paragraph) $K$ will be Hermitian (i.e. $x' = x$ for all $x \in K$); in particular this implies the commutativity of $K$.

The dual $\hat{K}$ of the Hermitian hypergroup $K$ is the space of all real valued multiplicative functions $\varphi$ on $K$ with $\varphi(e) = \| \varphi \|_\infty = 1$ (see Jewett [19], 6.3). For every probability measure $P$ on $K$ the Fourier transform $FP$ is the continuous real-valued function $\varphi \mapsto FP(\varphi) := \int \varphi \, dP$ on $\hat{K}$. It is a well known fact that the uniqueness theorem and the continuity theorem for the Fourier transform are valid for many commutative hypergroups (see Bloom, Heyer [3], Heyer [18], and Jewett [19]).

2.2. In the sequel we consider the class of Chebli-Trimèche hypergroups on $K := \mathbb{R}_+$: For every function $A$ on $\mathbb{R}_+$ (which turns out to be the Lebesgue density of a Haar measure of $K$) satisfying $A(0) = 0$, $A$ strictly increasing and unbounded, $A'(x) = \frac{\partial}{\partial x} + B(x)$ on a neighbourhood of 0 (where $\alpha > 0$ and $B$ is an odd $C^\infty$-function on $\mathbb{R}$), there exists a unique hypergroup structure on $\mathbb{R}_+$ such that

$$\frac{\partial}{\partial x} (A(x) A(y) \int f \, d\varepsilon_x * \varepsilon_y) = \frac{\partial}{\partial y} (A(x) A(y) \int f \, d\varepsilon_x * \varepsilon_y)$$

for every even $C^\infty$-function $f$ on $\mathbb{R}$ and $x, y \in \mathbb{R}_+$.

The neutral element of this hypergroup is 0, the inversion is the identity map and the multiplicative functions are precisely the solutions $\varphi_\lambda (\lambda \in \mathbb{C})$ of the differential equation

$$\varphi''_\lambda + \frac{A'}{A} \varphi'_\lambda + (\varphi^2 + \lambda^2) \varphi_\lambda = 0, \quad \varphi_\lambda(0) = 1, \varphi'_\lambda(0) = 0.$$

The dual of this hypergroup is $\hat{K} = \{ \varphi_\lambda : \lambda \in \mathbb{R}_+ \cup i[0, \varrho] \}$ where $\varrho \geq 0$ is defined as

$$\varrho := \frac{1}{2} \lim_{x \to \infty} \frac{A'(x)}{A(x)}.$$ In the following we will always identify $\hat{K}$ with the set of parameters $\mathbb{R}_+ \cup [0, \varrho]$.

The proof of the preceding results can be found in Chébli [5] and Zeuner [28].

2.3. The most important technical tool used in this article is the Laplace representation for the multiplicative functions $\varphi_\lambda (\lambda \in \mathbb{C})$ proved in Chébli [5], Proposition 1-IV: For every $x \in \mathbb{R}_+$ there exists a probability measure $\nu_x$ on $[-x, x]$ such that

$$\varphi_\lambda(x) = \int e^{-(s+i\lambda)} \nu_x(dt) \quad \text{for } x \in \mathbb{R}_+, \lambda \in \mathbb{C}.$$
Furthermore the measure $\tau_x$ with the density $t \mapsto e^{-\xi t}$ with respect to $\nu_x$ is a symmetric sub-probability measure on $\mathbb{R}$ which depends continuously on $x$ (in the weak topology on $M^b(\mathbb{R})$). Therefore $\tau$ may be considered as a sub-Markovian kernel from $\mathbb{R}_+$ into $\mathbb{R}$ and it follows from the Laplace representation for $\varphi_\lambda$ that for every $P \in M^1(\mathbb{R}_+)$ we have

$$\mathcal{F}P(\lambda) = \hat{\tau} \hat{P}(\lambda) \quad \text{for all } \lambda \in \mathbb{R}_+$$

where $\tau P(A) = \int \tau_x(A) P(dx)$ for every Borel measurable subset $A$ of $\mathbb{R}$, and $\hat{}$ denotes the usual Fourier transform on the real line (which should be well distinguished from the Fourier transform $\mathcal{F}$ of $\mathbb{R}_+$ considered as a hypergroup).
3. Concretizations of hypergroups

The main problem which makes it difficult to state probabilistic results on a hypergroup \( K \) is the fact that the definition does not allow us to define the "product" of two independent random variables \( X \) and \( Y \) with values in \( K \) as a \( K \)-valued random variable \( X \cdot Y \) directly. It is clear, however, that the distribution of this product — if it exists — should be \( P_X \ast P_Y \). It is the purpose of this paragraph to construct such a random variable, unifying the different approaches which have been made in concrete examples.

3.1. Definition: Let \((K, \ast)\) be a hypergroup (not necessarily commutative), \( \mu \) a probability measure on a compact set \( M \) and \( \Phi : K \times K \times M \to K \) be Borel-measurable. \((M, \mu, \Phi)\) is called a concretization of \((K, \ast)\) if

\[
\mu\{\Phi(x, y, \cdot) \in A\} = (\varepsilon_x \ast \varepsilon_y)(A) \quad \text{for } x, y \in K, A \in \mathcal{B}(K).
\]

Here \( \mathcal{B}(K) \) denotes the Borel \( \sigma \)-field of \( K \).

3.2. Examples:

3.2.1. Let \( G \) be a locally compact group and \( \ast \) the convolution defined by the group operation. If we define \( \Phi(x, y, 1) := xy \) for \( x, y \in G \) then \( \{(1), \varepsilon_1, \Phi\} \) is a concretization of \((G, \ast)\).

3.2.2. More generally let \( H \) be a compact subgroup of \( G \) and \((G//H, \ast)\) the double coset hypergroup (see Jewett [19]). Then \((H, \omega_H, \Phi)\) is a concretization of \((G//H, \ast)\) if we define \( \Phi(x, y, h) := H \varphi(x)h \varphi(y)H \) where \( \varphi : G//H \to G \) is measurable and satisfies \( x = H \varphi(x)H \) for every \( x \in G//H \) (if \( G \) is locally compact, metrizable, and separable the existence of \( \varphi \) follows from Bondar [4]).

3.2.3. Let \( K := \mathbb{R}_+^\ast, \varepsilon_x \ast \varepsilon_y := \frac{1}{2}\varepsilon_{|x|y} + \frac{1}{2}\varepsilon_{x+y}, M := \{-1, 1\}, \mu := \frac{1}{2}\varepsilon_{-1} + \frac{1}{2}\varepsilon_{1}, \) and \( \Phi(x, y, \lambda) := |x + \lambda y| \). Then \((M, \mu, \Phi)\) is a concretization of \((\mathbb{R}_+^\ast, \ast)\).

3.2.4. Let \( \alpha > -\frac{1}{2} \) and \((\mathbb{R}_+^\ast, \ast)\) be the hypergroup defined in Kingman [21] by \( \varepsilon_x \ast \varepsilon_y := c_\alpha \int_{-1}^{\lambda} \frac{d \lambda}{\sqrt{x^2 + y^2 - 2\lambda xy}} (1 - \lambda^2)^{\alpha - \frac{1}{2}}, (x, y \in \mathbb{R}_+) \) with \( c_\alpha := \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha + \frac{1}{2}), \sqrt{\pi}}. \) A concretization of this hypergroup is given by \( M := [-1, 1], \mu := g \cdot \lambda_{[-1, 1]} \) (where \( g(\eta) := c_\alpha \cdot (1 - \eta^2)^{\alpha - \frac{1}{2}} \) and \( \lambda_{[-1, 1]} \) denotes the Lebesgue measure on \([-1, 1]) \) and \( \Phi(x, y, \lambda) := \sqrt{x^2 + y^2 - 2\lambda xy}. \)

3.2.5. Let \( \alpha > -\frac{1}{2} \) and \([0, \pi], \ast\) the hypergroup defined in Bingham [2] by

\[
\varepsilon_x \ast \varepsilon_y := c_\alpha \int_{-1}^{\infty} \varepsilon_{\arccos(x \cos y + \lambda \sin x \sin y)} (1 - \lambda^2)^{\alpha - \frac{1}{2}} d \lambda \quad \text{for } x, y \in [0, \pi].
\]

A concretization of this hypergroup is given by \( M := [-1, 1], \mu := g \cdot \lambda_{[-1, 1]} \) as in 3.2.4 and \( \Phi(x, y, \lambda) := \arccos(x \cos y + \lambda \sin x \sin y). \)
3.2.6. If we choose $M$ and $\mu$ as in 3.2.4 and $\Phi(x, y, \lambda) := \text{arch}(chxchy + \lambda shzshy)$ for $x, y \in \mathbb{R}_+, \lambda \in [-1, 1]$, we obtain a concretization of the hyperbolic hypergroup (see Karpelevich, Tutubalin, Shur [20], Tutubalin [25], and Zeuner [27]).

3.2.7. Let $\alpha > \beta > -\frac{1}{2}$ and consider the hypergroup operation on $\mathbb{R}_+$ defined in Flensted-Jensen, Koornwinder [11]. In this case a concretization is given by

$$M := [0, 1] \times [0, \pi], \mu := g \cdot \lambda^2_M$$

(with $g(r, \theta) := c_{\alpha, \beta}(1 - r^2)^{\alpha - \beta - 1}r^{2\beta + 1}(\sin \theta)^{2\beta}$, and

$$\Phi(x, y, (r, \theta)) := \text{arch}\left[\frac{1}{2}(1 + chx)(1 + chy) + \frac{r^2}{2}(1 - chx)(1 - chy) + r \cos \theta \cdot shx \cdot shy\right].$$

3.2.8. Consider the convolution

$$\epsilon_x \ast \epsilon_y := \sum_{s=0}^{x+y} \frac{|x - y| + 2s + 1}{(x + 1)(y + 1)} \epsilon_{|x-y|+2s}(x, y \in \mathbb{N})$$

on the nonnegative integers $\mathbb{N}$ (compare Eymard, Roynette [7] and Gallardo, Ries [14]). In [14] a concretization of this hypergroup is given by $M := S^2$ (the sphere in $\mathbb{R}^3$), $\mu$ the uniform distribution on $S^2$ and

$$\Phi(x, y, D) := |x - y| + 2\left[\frac{1}{2}([((x + 1)e_0 + (y + 1)D \parallel |x - y|)]\right]$$

(where $e_0$ is a fixed unit vector in $\mathbb{R}^3$ and $\parallel \parallel$ the Euclidean norm).

3.3. Definition: In the sequel let $(M, \mu, \Phi)$ be a concretization of the hypergroup $(K, \ast)$ and $(\Omega, \mathcal{F}, P)$ be a probability space. If $X$ and $Y$ are $K$-valued random variables and if $\Lambda$ is an $M$-valued random variable, independent from $(X, Y)$ and satisfying $P_\Lambda = \mu$ we define

$$X^\Lambda Y := \Phi(X, Y, \Lambda).$$

This is a $K$-valued random variable.

More generally let $(X_n : n \geq 1)$ be a sequence of $K$-valued random variables and $(\Lambda_n : n \geq 1)$ be a sequence of $M$-valued random variables with $P_{\Lambda_n} = \mu$ for $n \geq 1$ and such that $X_1, \Lambda_1, X_2, \Lambda_2, \ldots$ are independent. Then we define $\Lambda \prod_{j=1}^n$ recursively by

$$\Lambda_0 \prod_{j=1}^0 X_j := e \text{ and } \Lambda_1 \prod_{j=1}^n X_j := \Phi(X_n, \Lambda \prod_{j=1}^{n-1} X_j, \Lambda_n)$$

for $n \geq 1$. 
It is clear that \((\Lambda^n \prod_{j=1}^n X_j : n \in \mathbb{N})\) is a (non homogeneous) Markov chain, the transition kernel being

\[
P\{\Lambda^n \prod_{j=1}^n X_j = A \mid \Lambda^{n-1} \prod_{j=1}^{n-1} X_j = x\} = (P_{X^n} \ast \varepsilon_x)(A) \quad \text{P.a.s.}
\]

If the hypergroup is commutative we will write \(X^\Lambda Y\) instead of \(X^\Lambda Y\) and \(\sum_{j=1}^n \Lambda_j\) instead of \(\prod_{j=1}^n \Lambda_j\).

### 3.4. Proposition

Let \((X,Y,\Lambda)\) be independent and \(P_\Lambda = \mu\). Then \(P_{X^\Lambda Y} = P_X \ast P_Y\).

**Proof:**

\[
P_{X^\Lambda Y}(A) = P\{\Phi(X,Y,\Lambda) \in A\}
= \int \int P\{\Phi(x,y,\Lambda) \in A\} P_X(dx)P_Y(dy)
= \int \int \mu\{\Phi(x,y,\cdot) \in A\} P_X(dx)P_Y(dy)
= \int \int \varepsilon_x \ast \varepsilon_y(A) P_X(dx)P_Y(dy)
= (P_X \ast P_Y)(A) \quad \text{for } A \in B(K).
\]

### 3.5. Remark

It is clear that a concretization is not uniquely determined by the hypergroup and hence the same is true for \(X^\Lambda Y\). However, the following proposition shows that the joint distribution of \(X, Y,\) and \(X^\Lambda Y\) does not depend on the choice of the concretization.

### 3.6. Proposition

Let \(X,Y,X',\) and \(Y'\) be \(K\)-valued random variables with \(P(X,Y) = P(X',Y')\). Furthermore let \((M,\mu,\Phi)\) and \((M',\mu',\Phi')\) be concretizations of \((K,\ast)\) and \(\Lambda,\Lambda'\) be \(M\)-valued resp. \(M'\)-valued random variables such that \(\Lambda\) is independent of \((X,Y)\) and \(\Lambda'\) is independent of \((X',Y')\) with \(P_\Lambda = \mu\) and \(P_{\Lambda'} = \mu'\). Then \(P_{(X,Y,X^\Lambda Y)} = P_{(X',Y',X'^\Lambda Y')}\).

**Proof:** For every \(A,B,C \in B(K)\) we have \(P_{(X,Y,X^\Lambda Y)}(A \times B \times C)\)

\[
= \int_{A \times B} P\{\Phi(x,y,\Lambda) \in C\} P_{(X,Y)}(d(x,y))
= \int_{A \times B} \varepsilon_x \ast \varepsilon_y(C) P_{(X',Y')} (d(x,y))
= \int_{A \times B} P\{\Phi'(x,y,\Lambda') \in C\} P_{(X',Y')} (d(x,y))
= P_{(X',Y',X'^\Lambda Y')} (A \times B \times C).
\]

### 3.7. Corollary

If \((X_n,\Lambda_n : n \in \mathbb{N})\) are independent with \(P_{\Lambda_n} = \mu\) then the distribution of \((X_n,\Lambda^n \prod_{j=1}^n X_j : n \geq 1)\) does not depend on the concretization \((M,\mu,\Phi)\) or on the choice of \((\Lambda_n : n \geq 1)\).
3.8. Proposition: Let \((K, \ast)\) be a (locally compact) hypergroup with countable base of topology. Then there exists a mapping \(\Phi : K \times K \times [0, 1] \rightarrow K\) such that \(([0, 1], \lambda_{[0,1]}, \Phi)\) is a concretization of \((K, \ast)\).

Proof: We will only treat the case that \(K\) is not countable (if \(K\) is at most countable we may construct \(\Phi\) in the same way as below without worrying about measurability). It follows from the assumptions that there exists a bimeasurable bijection \(\psi : K \rightarrow [0, 1]\). The induced mapping \(\mu \mapsto \psi(\mu)\) from \(\mathcal{M}^1(K)\) onto \(\mathcal{M}^1([0, 1])\) is Borel measurable and hence so is the mapping \(p : [0, 1]^2 \rightarrow \mathcal{M}^1([0, 1])\) defined by

\[ p(x, y) := \psi(e^{-\psi^{-1}(x)} \ast e^{-\psi^{-1}(y)}). \]

It is a well known fact that for every \(\mu \in \mathcal{M}^1[0, 1]\) there is a unique left continuous increasing function \(\varphi_\mu : [0, 1] \rightarrow [0, 1]\) such that \(\varphi_\mu(\lambda_{[0,1]}) = \mu\), namely

\[ \varphi_\mu(\eta) = 0 \vee \sup \{ z \in [0, 1] : \mu([0, z]) < \lambda \}. \]

We will prove that the mapping \(\Phi_0 : [0, 1]^3 \rightarrow [0, 1]\) defined by \(\Phi_0(x, y, \lambda) := \varphi_{p(x, y)}(\lambda)\) for all \(x, y, \lambda \in [0, 1]\) is Borel measurable. It follows from the left sided continuity of \(\varphi_\mu\) that it suffices to show that for every \(\lambda \in [0, 1]\) the mapping

\[ \mu \mapsto \varphi_\mu(\lambda) = 0 \vee \sup \{ z \in [0, 1] : \mu([0, z]) < \lambda \} \]

is Borel measurable. This, however, is a consequence of

\[ \{ \mu \in \mathcal{M}^1([0, 1]) : \varphi_\mu(\lambda) \leq \varepsilon \} = \{ \mu \in \mathcal{M}^1([0, 1]) : \int_{[0,\varepsilon]} d\mu \geq \lambda \}. \]

The mapping \(\Phi : K \times K \times [0, 1] \rightarrow K\) can therefore be defined by

\[ \Phi(h, k, \lambda) := \psi^{-1}(\Phi_0(\psi(h), \psi(k), \lambda)). \]

3.9. We are now considering the special cases \(K = \mathbb{R}_+\) and \(K = [0, 1]\) (see Achour, Trimèche [1], Chébli [5], and Zeuner [28]). It follows from [28] that we may suppose without loss of generality that

\[ \min \text{ supp } \varepsilon_x \ast \varepsilon_y = |x - y| \quad \text{for } x, y \in K \]

and

\[ \max \text{ supp } \varepsilon_x \ast \varepsilon_y = x + y \quad \text{if } x, y \in K \quad \text{(and } x + y \leq 1 \text{ in the case } K = [0, 1]). \]
Since in these cases there is no need for a Borel isomorphism \( \psi \) in the proof of 3.8, we get the following additional properties of \( \Phi \):

\[
\Phi(x, y, 0) = |x - y| \quad \text{for } x, y \in K
\]

and

\[
\Phi(x, y, 1) = x + y \quad \text{for } x, y \in K \quad (\text{and } x + y \leq 1 \text{ in the case } K = [0, 1]).
\]

Furthermore \( \Phi(x, 0, \lambda) = \Phi(0, x, \lambda) = x \). Every hypergroup on \( \mathbb{R}_+ \) or \([0, 1]\) is commutative (Zeuner [28], corollary 2.4) and hence

\[
\Phi(x, y, \lambda) = \Phi(y, x, \lambda) \quad \text{for } x, y \in K, \ \lambda \in [0, 1].
\]

It is easy to prove that for every \( \lambda \in [0, 1] \) the mapping \( \Phi(\cdot, \cdot, \lambda) : K \times K \to K \) is lower semicontinuous. Under the additional assumptions that \( \varepsilon_x \ast \varepsilon_y \) is diffuse for \( x, y > 0 \) and \( \supp \varepsilon_x \ast \varepsilon_y = [|x - y|, x + y] \) (which happens to be true if \( (\mathbb{R}_+, \ast) \) is a Chébli-Trimèche hypergroup as shown by Trimèche [24], §8), \( \Phi(\cdot, \cdot, \lambda) \) is continuous for every \( \lambda \in [0, 1] \).
4. Moments

From now on let \((K, \ast)\) be a Chébli-Trimèche hypergroup on \(\mathbb{R}_+\) (see 2.2). It is proved in Chébli [5] that \(\varphi_\lambda(x)\) is an analytic function of \(\lambda\). The derivations of \(\varphi_\lambda(x)\) with respect to \(\lambda\) will be the most important tool to define moments for each probability measure on \(\mathbb{R}_+\) in a way which is consistent with the convolution structure.

4.1. Definition: For every \(\lambda \in \mathbb{C}, x \in \mathbb{R}_+\) and \(n \geq 0\) let \(\varphi_{n,\lambda}(x) := \left(\frac{\partial}{\partial \mu}\right)^n \varphi_{\lambda+\mu}(x)\big|_{\mu=0}\) and \(m_n(x) := \varphi_{n,\lambda}\).

Some elementary properties of the functions \(\varphi_{n,\lambda}\) and \(m_n\) will be proved first.

4.2. Let \(L\) be the differential operator on \(\mathbb{R}_+\) defined by \(Lf(x) = -f''(x) - \frac{A'(x)}{A(x)} f'(x)\) for \(x > 0\) and \(f \in C^2(\mathbb{R}_+)\) with \(f'(0) = 0\). By differentiating the differential equation

\[ L\varphi_\lambda = (\varphi^2 + \lambda^2)\varphi_\lambda, \quad \varphi_\lambda(0) = 1, \varphi_\lambda'(0) = 0 \]

with respect to \(\lambda\) we obtain

\[ L\varphi_{n,\lambda} = (\varphi^2 + \lambda^2)\varphi_{n,\lambda} + 2in\lambda \varphi_{n-1,\lambda} - n(n-1)\varphi_{n-2,\lambda}, \quad \varphi_{n,\lambda}(0) = \varphi_{n,\lambda}'(0) = 0 \]

and especially

\[ Lm_n = -2n\varphi m_{n-1} - n(n-1)m_{n-2}, \quad m_n(0) = m_n'(0) = 0 \text{ for } n \geq 1 \]

(with \(m_0(x) = 1\) for every \(x \in \mathbb{R}_+\)).

4.3. It follows from the Laplace representation (2.3) that

\[ \varphi_{n,\lambda}(x) = \int_{-\infty}^{\infty} t^n e^{-it \lambda} \nu_x(dt) = \int_{-\infty}^{\infty} t^n e^{-it \lambda} \tau_x(dt) \]

and

\[ m_n(x) = \int_{-\infty}^{\infty} t^n \nu_x(dt) \text{ for } x \in \mathbb{R}_+, \lambda \in \mathbb{C}, n \geq 1. \]

4.4. If \(\lambda \in \mathbb{R}_+\) \(\varphi_{n,\lambda}\) is real valued for every \(n \geq 1\) since \(\varphi_\lambda\) is real valued. For \(\lambda \in \mathbb{R}_+\) \(\varphi_{n,\lambda}\) is real valued if \(n\) is even and \(i\varphi_{n,\lambda}\) is real if \(n\) is odd. This follows since \(\varphi_\lambda(x)\) is an analytic function of \(\lambda\) and \(\varphi_\lambda\) is real for \(\lambda \in \mathbb{R}_+\).

It follows from \(m_n(x) = \int_0^{\infty} t^n (e^{\delta t} + (-1)^n e^{-\delta t}) \tau_x(dt)\) that \(m_n \geq 0\) for every \(n \geq 1\).

4.5. We now have to study the two cases \(\varrho = 0\) and \(\varrho > 0\) separately. We begin with the case \(\varrho = 0\). It is clear that \(m_n = 0\) if \(n\) is odd.

4.6. Lemma: Let \(\varrho = 0, \lambda \in \mathbb{R}_+, \text{ and } n \in \mathbb{N}\). Then

a) \(m_{2k} \leq 1 + m_{2n}\) for every \(k < n\),
b) \(|\varphi_{2n,\lambda}| \leq m_{2n}\), and

c) \(|\varphi_{2n-1,\lambda}| \leq 1 + m_{2n}\).

**Proof:**

a) \(m_{2k}(x) = \int t^{2k} \nu_x(dt) \leq \int (1 + t^{2n}) \nu_x(dt) = 1 + m_{2n}(x)\)

b) \(|\varphi_{2n,\lambda}(x)| \leq \int |t^{2n} e^{-it\lambda}| \nu_x(dt) = \int t^{2n} \nu_x(dt) = m_{2n}(x)\)

c) \(|\varphi_{2n-1,\lambda}(x)| \leq \int |t^{2n-1} e^{-it\lambda}| \nu_x(dt) = \int |t|t^{2n-1} \nu_x(dt) \leq \int (1 + t^{2n}) \nu_x(dt) = 1 + m_{2n}(x)\).

4.7. **Theorem:** Let \(P\) be a probability measure on \(\mathbb{R}_+\) and \(n \geq 1\). Then the following conditions are equivalent:

(i) \(\int m_{2n} dP\) is finite,

(ii) \(FP(k)(\lambda) = i^k \int \varphi_{k,\lambda} dP\) for all \(k \leq 2n, \lambda \in \mathbb{R}_+\). In particular \(FP^{(2k)}(0) = \int m_{2k} dP\).

**Proof:** "i \(\implies\) ii": By 4.6 a) and induction \(FP^{(2n-2)}\) exists. From 4.6 c) and b) we obtain

\(|\frac{\varphi_{2n-2,\lambda} - \varphi_{2n-2,\mu}}{\lambda - \mu}| \leq m_{2n} + 1\)

and

\(|\frac{\varphi_{2n-1,\lambda} - \varphi_{2n-1,\mu}}{\lambda - \mu}| \leq m_{2n}\)

and it follows from the dominated convergence theorem that \(FP^{(2n-1)}(\lambda)\) and \(FP^{(2n)}(\lambda)\) exist and equal \(t^{2n-1} \cdot \int \varphi_{2n-1,\lambda} dP\) and \(t^{2n} \cdot \int \varphi_{2n,\lambda} dP\) respectively. In particular

\(FP^{(2n-1)}(0) = \int \varphi_{2n-1,0} dP = \int m_{2n-1} dP = 0\).

"ii \(\implies\) i": Because of \(FP^{(2n-1)}(0) = 0\) the 2n-th derivative \(FP^{(2n)}(0)\) equals

\(2 \lim_{h \to 0} \frac{1}{h^2} (FP^{(2n-2)}(h) - FP^{(2n-2)}(0))\). Since \(\varphi_{2n-2,\lambda} \leq m_{2n-2}\) by 4.6 b) we may apply Fa-tou's lemma to obtain

\[0 \leq \int m_{2n} dP\]

\[= \int (\frac{\partial}{\partial h})^2 \varphi_{2n-2,h} |_{h=0} dP\]

\[= 2 \int \lim_{h \to 0} \frac{1}{h^2} (m_{2n-2} - \varphi_{2n-2,h}) dP\]

\[\leq 2 \liminf_{h \to 0} \frac{1}{h^2} \int m_{2n-2} - \varphi_{2n-2,h} dP\]

\[= 2 \liminf_{h \to 0} \frac{1}{h^2} (FP^{(2n-2)}(0) - FP^{(2n-2)}(h))\]

\[= -2FP^{(2n)}(0) < \infty\).

4.8. **Remark:** The condition \(FP^{(2n-1)}(0) = 0\) — which does not occur in the usual formulation of this theorem on \(\mathbb{R}\) — cannot be dismissed. For example in the case of Kingman's
hypergroups \((A(x) = x^{\alpha}, \alpha > 0, \text{see } 3.2.4)\) the (Cauchy type) distribution \(P\) with density
\[ x \mapsto \frac{2^{\frac{n}{2}+1}}{\sqrt{\pi} \Gamma(\frac{n}{2}+\frac{1}{2})} \frac{x^n}{(1+x^2)^{\frac{n}{2}+\frac{1}{2}}} \]
has the Fourier transform \(\lambda \mapsto e^{-\lambda}\) (see Erdélyi [6], 8.6.(4)) which is
infinitely often differentiable on \(\mathbb{R}_+\) but \(\int m_2 \, dP = \infty\).

Let us now turn to the case \(\varrho > 0\).

4.9. Lemma: Let \(\varrho > 0\) and \(n \geq 0\). Then

\( a) \ \varphi_{n,i\lambda} > 0 \) if \(\lambda \in [0, \varrho]\),

\( b) \ \varphi_{n,i\lambda} \leq \varphi_{n,i\mu} \) if \(0 \leq \lambda \leq \mu \leq \varrho\), and

\( c) \ \frac{m_n - \varphi_{n,i\lambda}}{\varrho - \lambda} \leq \frac{m_n - \varphi_{n,i\mu}}{\varrho - \mu} \) if \(0 \leq \lambda \leq \mu < \varrho\).

Proof: \( a) \ \varphi_{n,i\lambda}(x) = \int t^n e^{-t(e^{-\lambda})} \nu_x(dt) > 0\).

\[ \varphi_{n,i\lambda}(x) = \int_{-x}^{-x} t^n e^{-t(e^{-\lambda})} \nu_x(dt) = \int_0^x t^n (e^{t\lambda} + (-1)^n e^{-t\lambda}) \tau_x(dt) \]
\[ \leq \int_0^x t^n \cdot (e^{t\mu} + (-1)^n e^{-t\mu}) \tau_x(dt) \]
\[ = \varphi_{n,i\mu}(x) \]

since sinh and cosh are increasing functions.

\( c) \ \frac{m_n(x) - \varphi_{n,i\lambda}(x)}{\varrho - \lambda} = \int_0^x t^n \frac{(e^{t\varrho} + (-1)^n e^{-t\varrho}) - (e^{t\lambda} + (-1)^n e^{-t\lambda})}{\varrho - \lambda} \tau_x(dt) \]
\[ \leq \int_0^x t^n \frac{(e^{t\varrho} + (-1)^n e^{-t\varrho}) - (e^{t\mu} + (-1)^n e^{-t\mu})}{\varrho - \mu} \tau_x(dt) \]
\[ = \frac{m_n(x) - \varphi_{n,i\mu}(x)}{\varrho - \mu} \]

since sinh and cosh are both convex functions.

4.10. Lemma: Let \(\varrho > 0\). Then \(m_n(x) \leq (\frac{\varrho}{\varrho})^n + \varrho m_{n+1}(x)\) for all \(x \geq 0, n \in \mathbb{N}\).

Proof: If \(n\) is odd we conclude from \(\sinh y \leq y \cosh y\) for all \(y \geq 0\) that

\[ m_n(x) = \int_0^x t^n \cdot \sinh(t \varrho) \tau_x(dt) \]
\[ \leq \varrho \int_0^x t^{n+1} \cosh(t \varrho) \tau_x(dt) \]
\[ = \varrho m_{n+1}(x) \] for all \(x \geq 0\).

If \(n\) is even we use the inequality \(\cosh y \leq \sinh y + 1_{[0,1]}(y)\) to obtain \(y^n \cosh y \leq y^{n+1} \sinh y + 2^n\) which implies

13
\[ m_n(x) = \int_0^x t^n \chi(t\rho) \tau_x(dt) \]
\[ \leq \int_0^x (q t^{n+1} \chi(t\rho) + \left( \frac{2}{\rho} \right)^n) \tau_x(dt) \]
\[ \leq q m_{n+1}(x) + \left( \frac{2}{\rho} \right)^n \text{ for all } x \geq 0. \]

4.11. Theorem: Let \( \rho > 0, n \geq 1, \) and \( P \) be a probability measure on \( \mathbb{R}_+ \). Then the following conditions are equivalent:

(i) \( \int m_n dP \) is finite,

(ii) \( \lambda \mapsto \mathcal{F}P(i\lambda) \) is \( n \) times differentiable on \([0, \rho] \).

In both cases \( (\frac{\partial}{\partial \lambda})^k \mathcal{F}P(i\lambda) = \int \varphi_{k,i\lambda} dP \) for all \( k \leq n, \lambda \in [0, \rho] \).

Proof: “\( i \implies ii \)” It follows from 4.10 and induction that \( f: \lambda \mapsto \int \varphi_{i\lambda} dP \) is \((n-1)\) times differentiable on \([0, \rho] \) and \( f^{(n-1)}(\lambda) = \int \varphi_{n-1,i\lambda} dP \). The mean value theorem and 4.9 b) imply that \(|\varphi_{n-1,i\lambda}(x) - \varphi_{n-1,i\mu}(x)| \leq |\lambda - \mu| \cdot m_n(x)\). By Lebesgue’s theorem we obtain that \( f^{(n-1)} \) is differentiable and \( (\frac{\partial}{\partial \lambda})^n \mathcal{F}P(i\lambda) = f^{(n)}(\lambda) = \int \varphi_{n,i\lambda} dP \).

“\( ii \implies i \)” By induction the first \( n-1 \) moments \( \int m_k dP \) \((k \leq n-1)\) are finite and \( f^{(n-1)}(\lambda) = \int \varphi_{n-1,i\lambda} dP \). It follows from 4.10 c) that the difference quotients \( \frac{m_{n-1+i\lambda} - \varphi_{n-1,i\lambda}}{\rho - \lambda} \)
(\( \lambda < \rho \)) approach \( m_n \) increasingly as \( \lambda \nearrow \rho \) and hence by the theorem of monotone convergence \( \int m_n dP = \frac{\partial}{\partial \lambda} \int \varphi_{n-1,i\lambda} dP|_{\lambda=\rho} \) exists and equals \( f^{(n)}(\rho) \) which is finite.

4.12. Remark: Let \( \rho > 0 \) and \( \int m_n dP \) be finite. Then it follows from 4.3 and 4.9 b) that \( |\varphi_{n,\lambda}(x)| \leq m_n(x) \) for all \( x \geq 0 \) and \( \lambda \in \mathbb{C} \) such that \( |\Re\lambda| \leq \rho \). Therefore the function \( \lambda \mapsto \int \varphi_{\lambda} dP \) is \( n \) times differentiable in this strip and in particular \( \eta \mapsto \int \varphi_{\eta+i\rho} dP \) is \( n \) times differentiable on \( \mathbb{R} \). This fact will be used later.

4.13. Remark: Let \( \rho \geq 0 \). Then a sufficient condition for \( E(m_n(X)) \) being finite (where \( X \) is a \( \mathbb{R}_+ \)-valued random variable) is \( E(X^n) < \infty \). This follows from the inequality \( m_n(x) \leq x^n \) for all \( x \geq 0 \) which is a consequence of the fact that the measure \( \mu_x \) in 4.3 is supported by \([-x,x]\).

4.14. Theorem: Let \( \rho \geq 0, X \) and \( Y \) be independent \( \mathbb{R}_+ \)-valued random variables such that \( E(m_n(X)) \) and \( E(m_n(Y)) \) are finite.

a) Then \( E(m_n(X + Y)) \) is finite.

b) If \( \rho > 0 \) we obtain

\[
E(m_n(X + Y)) = \sum_{k=0}^n \binom{n}{k} E(m_k(X)) E(m_{n-k}(Y))
\]
c) If \( \varrho = 0 \) and \( n = 2\ell \) we obtain

\[
E(m_{2\ell}(X^{\text{A}}Y)) = \sum_{j=0}^{\ell} \binom{2\ell}{2j} E(m_{2j}(X)) E(m_{2\ell-2j}(Y)).
\]

**Proof:** a) follows from the fact that the product of two \( n \) times differentiable functions is again \( n \) times differentiable and theorems 4.7 and 4.11. b) and c) are a consequence of Leibniz's rule.
5. Expectation

5.1. In this paragraph the special properties of the function $m_1$ will be considered. Since $m_1 = 0$ if $\varrho = 0$ we will assume $\varrho > 0$ throughout the whole paragraph. The function $m_1$ has already been defined in Faraut [8] and Trimèche [23] under the name "forme quadratique généralisée". It will be used to define a modified expectation for every $\mathbb{R}_+$-valued random variable consistent with the hypergroup structure (see 5.6).

5.2. Examples If $A$ is of the form $A(x) = (sh x)^{\alpha}$ for some $\alpha > 0$, the function $m_1$ can be written down in closed form for some values of $\alpha$. According to Faraut [8], $m_1(x) = 2 \ln \cosh \frac{x}{2}$ if $\alpha = 1$ and $m_1(x) = x \coth x - 1$ if $\alpha = 2$. If $\alpha = 3/2$ one calculates $m_1(x) = 2 \ln \cosh \frac{x}{2} + \frac{1}{2}(\tanh \frac{x}{2})^2 (x \geq 0)$. If $A(x) = (ch x)^2$ then $m_1(x) = x \tanh x$.

5.3. Remark: By integrating the differential equation for $m_1$ (see 4.2), one obtains

$$m_1(x) = 2\varrho \int_0^x \frac{1}{A(y)} \int_0^y A(z)dz \; dy \quad \text{for} \; x \geq 0.$$ 

This formula has been used in Gallardo [12].

5.4. Definition: Let $(\mathbb{R}_+, \ast)$ be a Chébli-Trimèche hypergroup with the corresponding function $m_1$. Then for every $\mathbb{R}_+$-valued random variable $X$, $E_\ast(X) := E(m_1(X))$ is called the $\ast$-expectation of $X$.

5.5. Remark Although $E_\ast(X) = 0$ holds for every random variable $X$ in the case $\varrho = 0$, the $\ast$-expectation does not lose its entire sense as can be seen from theorem 8.4. If $\varrho > 0$ the $\ast$-expectation is also called "dispersion" (see Faraut [8], Gallardo [12], and Trimèche [23]). We reserve the expression "variance" (which is also used for $E_\ast(X)$ in the literature) to the corresponding number related with $m_2$. Theorems 7.6, 7.7, and 10.2 are the motivation for this terminology.

5.6. Proposition: Let $X$ and $Y$ be independent random variables. Then $E_\ast(X \oplus Y) = E_\ast(X) + E_\ast(Y)$.

The proof follows from 4.14. □

5.7. Lemma: If $\varrho > 0$ then $\lim_{x \to -\infty} \frac{m_1(x)}{x} = 1$.

Proof: Suppose that $m_1''$ takes negative values. Since $m_1''(0) = \frac{2\varrho}{2\varrho + 1} > 0$ there exists $x_0 > 0$ with $m_1'(x_0) > 0$, $m_1''(x_0) < 0$ and $m_1'''(x_0) < 0$. This would imply that $m_1' \cdot \frac{A'}{A}$ and $m_1''$ are strictly decreasing in a neighbourhood of $x_0$. But this is impossible since $m_1'' + m_1' \frac{A'}{A} = 2\varrho$ by 4.2. From this contradiction we conclude that $m_1'' \geq 0$ and $m_1'$ is increasing. Suppose now that
\[ \beta := \lim_{x \to \infty} m'_1(x) < 1. \] This implies \( m''_1 = 2 \rho - \frac{A'}{A} \cdot m'_1 > 2 \rho - \frac{A'}{A} \beta. \) When \( \frac{A'}{A}(x) \) is close enough to \( 2 \rho \) the last number becomes strictly positive and hence \( m''_1(x) \) is bounded away from 0 for large enough \( x. \) This is a contradiction with \( \sup_{x \geq 0} m'_1(x) = \beta < 1. \) On the other hand, from \( 2 \rho m'_1 \leq m''_1 + \frac{A'}{A} m'_1 = 2 \rho \) we obtain \( m'_1 \leq 1 \) and hence \( m'_1(x) \not\to 1 \) as \( x \to \infty. \) This implies \( m'_1(x) \not\to 1. \)

5.8. **Corollary:** Let \( \rho > 0 \) and \( X \) be a \( \mathbb{R}_+ \)-valued random variable. Then \( E_*(X) \) is finite if and only if \( E(X) \) is finite.

5.9. **Proposition:** Let \( \rho > 0 \) and \( X \) be a \( \mathbb{R}_+ \)-valued random variable with \( 0 \leq E_*(X) \leq +\infty. \) Then \( \frac{\partial}{\partial \lambda} E(\varphi_{\lambda}(X))|_{\lambda=\rho} = E_*(X). \)

**Proof:** If \( E_*(X) \) is finite this is theorem 4.11. If \( E_*(X) = \infty \) this follows from 4.9 and the theorem of monotone convergence. \( \blacksquare \)
6. Variance

We will now explore the properties of the function $m_2$ to obtain a modification of the notion of variance in a similar way as for the expectation in the last paragraph.

6.1. Examples: If $A(x) = x^\alpha$ ($\alpha \geq 0$) we obtain $m_2(x) = \frac{1}{\alpha + 1} x^2$. If $A(x) = (\text{sh} x)^2$, $m_2(x) = x^2 + 2 - 2x \text{ coth} x$ (Tutubalin [25], p. 191) and in the case $A(x) = (\text{ch} x)^2$ we have $m_2(x) = x^2$.

6.2. Lemma: $m_1(x)^2 \leq m_2(x) \leq x^2$ for every $x \geq 0$.

Proof: The first inequality follows from 4.3 and Jensen’s inequality; the second has already been proved in 4.13.

6.3. Corollary: Let $\varrho > 0$. Then $\lim_{x \to \infty} \frac{1}{x^2} m_2(x) = 1$.

6.4. Lemma: If $\varrho = 0$ then $m_2$ is a convex function and $\lim_{x \to \infty} \frac{m_2(x)}{x} = +\infty$.

Proof: The convexity of $m_2$ follows in the same way as the convexity of $m_1$ in the first part of the proof of 5.7. The assumption that $m_2'$ is bounded leads to a contradiction since it implies $\lim_{x \to -\infty} m_2''(x) = 2 - \lim_{x \to -\infty} m_2'(x) \frac{A'(x)}{A(x)} = 2$. Hence $m_2'$ is unbounded. But from $m_2'(x) \not\to \infty$ as $x \to \infty$ follows $\lim_{x \to \infty} \frac{m_2(x)}{x} = +\infty$ by the mean value theorem.

6.5. Lemma. Suppose that $\varrho = 0$ and $\{x \frac{A'(x)}{A(x)} : x > 0\}$ is bounded. Then there exists $\gamma > 0$ such that $m_2(x) \geq \gamma x^2$ for $x \geq 0$.

Proof: It follows from the differential equation $m_2'' + \frac{A'}{A} m_2' = 2$ that the function $\psi$ defined on $\mathbb{R}_+$ by

$$
\psi(x) = \begin{cases} 
\frac{m_2'(x)}{x} & \text{for } x > 0 \\
m_2''(0) = \frac{2}{\alpha + 1} & \text{for } x = 0
\end{cases}
$$

satisfies the differential equation

$$
x \psi'(x) + \left(x \frac{A'(x)}{A(x)} + 1\right) \psi(x) = 2, \quad \psi(0) = \frac{2}{\alpha + 1}.
$$

Let $b$ be an upper bound for $x \frac{A'(x)}{A(x)}$ ($x > 0$). Then for every $x$ such that $\psi(x) < \frac{2}{b+1}$ we obtain $\psi'(x) > 0$ and so $\psi$ is certainly bounded from below by $2\gamma$ where $\gamma := \min(\frac{1}{b+1}, \frac{1}{\alpha + 1})$. But this implies $m_2'(x) \geq 2\gamma x$ and hence the result.

6.6. Definition: In order to define the $*$-variance for every Chébli-Trimèche hypergroup $(\mathbb{R}_+, *)$ with corresponding functions $m_1$ and $m_2$ we introduce the function $v : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$, $v(x, \xi) := m_2(x) - 2\xi m_1(x) + \xi^2$ ($\lambda, \xi \geq 0$) which is non negative by 6.2. For every $\mathbb{R}_+$-valued
random variable $X$ such that $E(m_2(X)) < \infty$ the function $\xi \mapsto E(v(X, \xi))$ on $\mathbb{R}$ takes its minimum at $\xi = E_*(X)$, this value being $V_*(X) := E(v(X, E_*(X))) = E(m_2(X)) - E(m_1(X))^2 \geq 0$. If $E(m_2(X)) = \infty$ we define $V_*(X) = \infty$. $V_*(X)$ is called the $*$-variance of $X$.

6.7. **Remark:** In the case $\rho = 0$ $V_*(X)$ equals $E(m_2(X))$; this number is called the “dispersion” of $X$ in Trimèche [23].

6.8. **Remark:** $V_*(X)$ is strictly positive unless $X = 0$ $P$-a.s. At the first look it is surprising that the $*$-variance of a constant $X \neq 0$ does not equal zero. But it reflects the fact that $X \# Y$ is not a constant even if $X > 0$ and $Y > 0$ are constant.

6.9. **Proposition:** If $\rho > 0$ or if $\{x - \frac{A'(x)}{A(x)} : x > 0\}$ is bounded, $V_*(X)$ exists if and only if $E(X^2) < \infty$.

**Proof:** This is a consequence of 6.3 in the first case, and 6.2 and 6.5 in the second.

6.10. **Proposition:** Let $X$ and $Y$ be independent $\mathbb{R}_+$-valued random variables. Then $V_*(X \# Y) = V_*(X) + V_*(Y)$. Both sides of this equation may be infinite.

**Proof:** If $E(m_2(X))$ or $E(m_2(Y))$ are infinite it follows from 4.14 that $E(m_2(X \# Y))$ and hence $V_*(X \# Y)$ equals $+\infty$. Let us therefore suppose that $V_*(X) < \infty$ and $V_*(Y) < \infty$. It follows from 4.14 that

$$
V_*(X \# Y) = E(m_2(X \# Y)) - E(m_1(X \# Y))^2
= E(m_2(X)) + 2E(m_1(X))E(m_1(Y)) + E(m_2(Y)) - E(m_1(X))^2 - 2E(m_1(X))E(m_1(Y)) - E(m_1(Y))^2
= V_*(X) + V_*(Y).
$$
7. Laws of large numbers in the case of exponential growth

Recall that \((M, \mu, \Phi)\) denotes a fixed concretization of a Chébli-Trimèche hypergroup \((\mathbb{R}_+, +)\).

7.1. Proposition: Let \(X, Y, \text{ and } \Lambda \) be independent \(\mathbb{R}_+, \mathbb{R}_+, \text{ and } M\)-valued random variables such that \(P_{\Lambda} = \mu\).

a) If \(E_*(X)\) and \(E_*(Y)\) are finite, then \(E(m_1(X) X Y) | X = m_1(X) + E_*(Y)\) \(P\)-almost surely.

b) If \(V_*(X)\) and \(V_*(Y)\) are finite, then

\[
E(v(X X Y), E_*(X X Y)) | X = v(X, E_*(X)) + V_*(Y) \quad P\text{- a.s.}
\]

c) If \(|\Im \lambda| \leq \varrho, then

\[
E(\varphi_*(X X Y) | X) = \varphi_*(X) \cdot E(\varphi_*(Y)) \quad P\text{- a.s.}
\]

Proof: a) Let \(A \in \mathcal{B}(\mathbb{R}_+)\). Then by 5.6 it follows

\[
E(1_{\{X \in A\}} \cdot m_1(X X Y)) = E(m_1(1_{\{X \in A\}} X X Y)) - E(1_{\{X \notin A\}} m_1(Y))
\]
\[
= E(m_1(1_{\{X \in A\}} X) + E_*(Y) - P\{X \notin A\} E_*(Y)
\]
\[
= E(1_{\{X \in A\}} \cdot [m_1(X) + E_*(Y)]).
\]

b) For every \(A \in \mathcal{B}(\mathbb{R}_+)\) we conclude from 4.14 that

\[
E(1_{\{X \in A\}} \cdot m_2(X X Y)) = E(m_2(1_{\{X \in A\}} X X Y)) - E(1_{\{X \notin A\}} m_2(Y))
\]
\[
= E(m_2(1_{\{X \in A\}} X) + 2E(m_1(1_{\{X \in A\}} X)) E_*(Y)
\]
\[
+ E(m_2(Y)) - P\{X \notin A\} E(m_2(Y))
\]
\[
= E(1_{\{X \in A\}} \cdot [m_2(X) + 2m_1(X) E_*(Y) + E(m_2(Y))]).
\]

Therefore \(E(m_2(X X Y) | X) = m_2(X) + 2m_1(X) E_*(Y) + E(m_2(Y))\) \(P\)-almost surely and hence

\[
E(v(X X Y, E_*(X X Y)) | X) = E(m_2(X X Y) | X) - 2E_*(X X Y) E(m_1(X X Y) | X) + E_*(X X Y)^2
\]
\[
= m_2(X) + 2m_1(X) E_*(Y) + E(m_2(Y))
\]
\[
- 2(E_*(X) + E_*(Y))(m_1(X) + E_*(Y)) + (E_*(X) + E_*(Y))^2
\]
\[
= m_2(X) - 2m_1(X) E_*(Y) + E_*(X)^2 + E(m_2(Y)) - E_*(Y)^2
\]
\[
= v(X, E_*(X)) + V_*(Y) \quad P\text{- almost surely.}
\]
c) Since \( \varphi_{\lambda} \) is a bounded multiplicative function we obtain for every \( A \in B(\mathbb{R}_+) \)

\[
E(1_{\{X \in A\}} \cdot \varphi_{\lambda}(X^A Y)) = E(\varphi_{\lambda}((1_{\{X \in A\}} X)^A Y)) - E(1_{\{X \in A\}} \varphi_{\lambda}(Y)) \\
= E(\varphi_{\lambda}(1_{\{X \in A\}} X))E(\varphi_{\lambda}(Y)) - P\{X \notin A\}E(\varphi_{\lambda}(Y)) \\
= (E(1_{\{X \in A\}} \varphi_{\lambda}(X)) + P\{X \notin A\})E(\varphi_{\lambda}(Y)) - P\{X \notin A\}E(\varphi_{\lambda}(Y)) \\
= E(1_{\{X \in A\}} \cdot \varphi_{\lambda}(X)E(\varphi_{\lambda}(Y))).
\]

7.2. Notation: For the rest of this article we suppose that \( X_1, X_2, \ldots, \Lambda_1, \Lambda_2, \ldots \) are independent and \( \mathbb{R}_+ \)-resp. \( M \)-valued random variables such that \( P_{\Lambda_{n}} = \mu \) for every \( n \geq 1 \). It follows from 3.3 that the process \( \{S_n : n \geq 0\} \) where \( S_n := \sum_{j=1}^{n} X_j \) is a (non homogeneous) Markov chain.

7.3. Corollary: a) If \( E_{\ast}(X_n) < \infty \) resp. \( V_{\ast}(X_n) < \infty \) for \( n \geq 1 \) then \( (m_1(S_n) : n \in \mathbb{N}) \) resp. \( (v(S_n, E_{\ast}(S_n)) : n \in \mathbb{N}) \) are submartingales with respect to the canonical filtration.

b) If \( \lambda \in \mathbb{R}_+ \cup [0, \varrho] \) then \( (\varphi_{\lambda}(S_n) : n \in \mathbb{N}) \) is a supermartingale.

Proof: From 7.1 a) we obtain for every \( n \geq 1 \)

\[
E(m_1(S_n)|S_{n-1}) = E(m_1(X_n^A S_{n-1})|S_{n-1}) = m_1(S_{n-1}) + E_{\ast}(X_n) \geq m_1(S_{n-1}) \quad P\text{-almost surely.}
\]

The other assertions can be proved in the same way.

Note that this corollary holds for any hypergroup \( K \) and \( \mathbb{R}_+ \)-valued functions \( m_1 \) and \( m_2 \) on \( K \) such that 4.14 and \( m_2 \geq m_1^2 \) hold.

For the rest of this paragraph we suppose \( \varrho > 0 \). In view of \( A(x) \geq A(1) \cdot e^{2\varrho(x-1)} \) for \( x \geq 1 \) this implies that \( (\mathbb{R}_+, \ast) \) is of exponential growth.

7.4. Theorem: Let \( \{X_n : n \geq 1\} \) be an independent series of \( \mathbb{R}_+ \)-valued random variables such that \( \sum_{n=1}^{\infty} \frac{1}{n^2} V_{\ast}(X_n) < \infty \). Then

\[
\frac{1}{n}(S_n - m_1^{-1}(E_{\ast}(S_n))) \to 0 \quad P\text{-almost surely.}
\]

Proof: Let \( \varepsilon > 0 \). For every \( r \in \mathbb{N} \) we define \( s_r := E_{\ast}(S_r) \) and

\[
A_r := \{|m_1(S_n) - s_r| \geq \varepsilon \cdot n \text{ for some } n \in [2^r, 2^{r+1}[\cap \mathbb{N}].
\]

The probability of this event is

\[
P(A_r) \leq P(|m_1(S_n) - s_r| \geq \varepsilon \cdot 2^r \text{ for some } n < 2^{r+1}) \\
= P(m_2(S_n) - 2s_r m_1(S_n) + s_r^2 \geq \varepsilon^2 \cdot 2^{2r} + m_2(S_n) - m_1(S_n)^2 \text{ for some } n < 2^{r+1}) \\
\leq P(v(S_n, s_n) \geq \varepsilon^2 \cdot 2^{2r} \text{ for some } n < 2^{r+1})
\]

21
by 6.2. Applying 7.3 a) and Kolmogorov's inequality for positive submartingales (see Feller [9], p. 241) we obtain

\[ P(A_r) \leq \frac{1}{e^2 \cdot 2^{2r}} \cdot E(v(S_{2^r+1-1, s_{2^r+1-1}})) = \frac{1}{e^2 \cdot 2^{2r}} \sum_{j=1}^{2^r+1-1} V_*(X_j). \]

This implies

\[
\varepsilon^2 \sum_{r \geq 0} P(A_r) \leq 2^{2r} \sum_{r \geq 0} \sum_{j=1}^{2^r+1-1} V_*(X_j) \\
= \sum_{j \geq 1} V_*(X_j) \sum_{2^r+1 \geq j} 4^{-r} \\
\leq \frac{16}{3} \sum_{j \geq 1} \frac{V_*(X_j)}{j^2} < \infty.
\]

Hence by the Borel-Cantelli Lemma

\[ P\{\frac{1}{n}|m_1(S_n) - s_n| \geq \varepsilon \text{ infinitely often} \} = P(\limsup \limits_{r \to \infty} A_r) = 0 \]

for every \( \varepsilon > 0 \) and thus

\[ P\{\frac{1}{n}(m_1(S_n) - s_n) \to 0 \} = 1. \]

Since \( (m_1^{-1})'(t) \searrow 1 \) as \( t \to \infty \) there is a number \( a > 0 \) such that \( |m_1^{-1}(x) - m_1^{-1}(y)| \leq 2|x - y| + a \) for all \( x, y \in \mathbb{R}_+ \). Therefore the last equation implies \( P\{\lim \limits_{n \to \infty} \frac{1}{n}(S_n - m_1^{-1}(s_n)) = 0 \} = 1. \)

7.5. Remarks

7.5.1. If in the situation of the preceding theorem we assume additionally that \( \frac{1}{n}E_*(S_n) \) is bounded we obtain

\[ \lim_{n \to \infty} \frac{1}{n}(S_n - E_*(S_n)) = 0 \quad \text{P-a.s..} \]

This is a consequence of

\[ \frac{1}{n}(m_1^{-1}(E_*(S_n)) - E_*(S_n)) = \frac{1}{n}E_*(S_n) \cdot \left( \frac{m_1^{-1}(E_*(S_n))}{E_*(S_n)} - 1 \right) \to 0 \]

(compare 5.7).

7.5.2. If in the situation of the preceding theorem \( \eta := \lim \limits_{n \to \infty} \frac{1}{n}E_*(S_n) \) exists, then

\[ \lim_{n \to \infty} \frac{1}{n}S_n = \eta \quad \text{P-a.s..} \]

7.5.3. Under additional assumptions on the function \( A \) we can obtain \( m_1(x) = x + o(x) \) for \( x \to \infty \). Then the conclusion of theorem 7.4 may be written as \( \lim \limits_{n \to \infty} \frac{1}{n}(S_n - E_*(S_n)) = 0 \) P-a.s.
7.6. Corollary: Let \((X_n : n \geq 1)\) be an i.i.d. sequence of integrable random variables.

Then

\[
\frac{1}{n} S_n \to E_*(X_1) \quad P\text{-a.s.}
\]

Proof: Let \(a > 0\) be arbitrary, consider the truncated variables \(X_n^a := 1_{\{X_n < na\}} \cdot X_n\), and define \(S_0^a := 0\), \(S_n^a := S_{n-1}^a + \frac{1}{n} X_n^a\), \(s_n^a := E_*(S_n^a)\) for \(n \geq 1\). Using 6.2 we obtain

\[
\sum_{n=1}^{\infty} \frac{1}{n^2} V_*(X_n^a) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{j=0}^{\infty} E(1_{\{aj \leq X_n < a(j+1)\}} \cdot v(X_n^a, E_*(X_n^a)))
\]

\[
\leq \sum_{n=1}^{\infty} \frac{1}{n^2} \left\{ \sum_{j=0}^{n-1} P\{aj \leq X_n < a(j+1)\} \cdot (E_*(X_n)^2 + a^2(j+1)^2) + \sum_{j=n}^{\infty} P\{aj \leq X_n < a(j+1)\} \cdot E_*(X_n)^2 \right\}
\]

\[
\leq \sum_{n=1}^{\infty} \frac{1}{n^2} \left\{ \sum_{j=0}^{n-1} P\{aj \leq X_1 < a(j+1)\} \cdot a^2(j+1)^2 + \sum_{j=0}^{\infty} P\{aj \leq X_1 < a(j+1)\} \cdot E_*(X_1)^2 \right\}
\]

\[
= \sum_{j=0}^{\infty} P\{aj \leq X_1 < a(j+1)\} \cdot \left\{ a^2(j+1)^2 \cdot \sum_{n=j+1}^{\infty} \frac{1}{n^2} + E_*(X_1)^2 \cdot \sum_{n=1}^{\infty} \frac{1}{n^2} \right\}
\]

\[
\leq \sum_{j=0}^{\infty} 2a^2(j+1)P\{aj \leq X_1 < a(j+1)\} + \frac{\pi^2}{6} \cdot E_*(X_1)^2 < \infty.
\]

On the other hand it follows from \(\lim_{n \to \infty} E_*(X_n^a) = E_*(X_1)\) that

\[
\lim_{n \to \infty} \frac{1}{n} s_n^a = E_*(X_1).
\]

Hence 5.5.2 implies \(\lim_{n \to \infty} \frac{1}{n} S_n^a = E_*(X_1)\) P-a.s. for every \(a > 0\).

The probability of \(\Omega_a := \{X_n < na \text{ for all } n \geq 1\}\) is

\[
P(\Omega_a) = 1 - P\{X_n \geq na \text{ for some } n \geq 1\}
\]

\[
\geq 1 - \sum_{n=1}^{\infty} P\{X_n \geq na\}
\]

\[
\geq 1 - \frac{1}{a} E(X_1).
\]

The construction of \(S_n^a\) implies that \(S_n = S_n^a\) on \(\Omega_a\) and hence \(\lim_{n \to \infty} \frac{1}{n} S_n = E_*(X_1)\) P-a.s. on \(\Omega_a\).

Since \(P(\Omega_a) \to 1\) as \(a \to \infty\) the corollary is proved.
7.7. Theorem: Let \((X_n : n \geq 1)\) be an i.i.d. sequence of random variables with 
\(E^*(X_1) = +\infty\). Then \(\frac{1}{n} S_n \to \infty \quad P\text{-a.s.}\).

Proof: Let \(a\) be an arbitrary positive number. We will show that \(P\{\frac{1}{n} S_n < a \text{ i.o.}\} = 0\). This can be done by proving
\[
\sum_{n \geq 1} P\{\frac{1}{n} S_n < a\} < \infty
\]
and using the Borel-Cantelli lemma.

Consider the functions \(\lambda \mapsto E(\varphi_{i(e-\lambda)}(X_1))\) and \(\lambda \mapsto e^{-a\lambda}\). Since the derivations at 0 of these functions are \(-\infty\) (by 5.9) and \(-a\) there exists \(\lambda \in [0, q]\) such that
\(0 < e^{a\lambda} \cdot E(\varphi_{i(e-\lambda)}(X_1)) < 1\).

Therefore from \(\varphi_{i(e-\lambda)}(x) = \int_{-x}^{x} e^{-t\lambda} \nu_x(dt) \geq e^{-\lambda x}\) it follows
\[
P\{\frac{1}{n} S_n < a\} = P\{\varphi_{i(e-\lambda)}(S_n) > \varphi_{i(e-\lambda)}(an)\}
\leq P\{\varphi_{i(e-\lambda)}(S_n) > e^{-\lambda an}\}
\leq e^{\lambda an} \cdot E(\varphi_{i(e-\lambda)}(S_n))
= (e^{\lambda a} \cdot E(\varphi_{i(e-\lambda)}(X_1)))^n
\]

and finally
\[
\sum_{n \geq 1} P\{\frac{1}{n} S_n < a\} \leq \sum_{n \geq 1} (e^{\lambda a} E(\varphi_{i(e-\lambda)}(X_1)))^n < \infty. \quad \blacksquare
\]

7.8. Remark: In 7.2 we have only considered the case of a random walk starting at the neutral element 0 of \((\mathbb{R}_+, *)\). However, 7.6 and 7.7 (and clearly 7.4) remain valid if the starting point is arbitrarily distributed. A short look at the proofs of 7.6 and 7.7 shows that it suffices to suppose that \((X_2, X_3, \ldots)\) are identically distributed \((X_1\) even does not need to be integrable): \(X_1 = S_1\) can then be considered as the starting point of the random walk \((S_n : n \geq 1)\).
8. Laws of large numbers in the case of exponential boundedness

8.1. In this paragraph we suppose that \( \rho = 0 \). This implies \( E_*(X) = 0 \) for every random variable and therefore we expect the law of large numbers of a particularly simple form. For example if (in the terminology of 7.2) the variances \( V_*(X_j) = E(m_2(X_j)) \) are bounded by some constant \( b > 0 \) we obtain for every \( \varepsilon > 0 \)

\[
P\{\frac{1}{n} S_n \geq \varepsilon\} = P\{m_2(S_n) \geq m_2(n\varepsilon)\} \leq \frac{V_*(S_n)}{m_2(n\varepsilon)} \leq \frac{nb}{m_2(n\varepsilon)} \to 0
\]

by 6.4 and hence

\[
\frac{1}{n} S_n \to 0 = E_*(X_j) \quad \text{in probability.}
\]

8.2. However, the proof of a strong law becomes more difficult and requires some restrictions concerning the function \( m_2 \). For the rest of this paragraph we have to suppose that for every \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that \( m_2(\varepsilon x) \geq \delta m_2(x) \) for every \( x \geq 0 \).

8.3. Examples:

8.3.1. If \( \{x \frac{A'(x)}{A(x)} : x > 0\} \) is bounded, then 8.2 holds. This follows from 6.5. This criterion is useful if \( \frac{A'}{A} \) decreases fast. The opposite case is considered in the following example.

8.3.2. Suppose that there is a \( c > 0 \) with \( \frac{A'(x)}{A(x)} \geq c \cdot \frac{A'(x)}{A(x)} \) for \( x > 0 \). Then 8.2 holds.

Proof: We consider the function \( \varphi : \mathbb{R}_+ \to \mathbb{R} \) defined by \( \varphi(x) := 4m_2(\frac{x}{2}) - 2c m_2(x) \) for \( x \geq 0 \). From the convexity of \( m_2 \) (6.4) and 4.2 we obtain \( m_2(\frac{x}{2}) : \frac{A'(x/2)}{A(x/2)} \leq 2 \) and hence

\[
[2 \frac{A'(x)}{A(x)} - \frac{A'(x/2)}{A(x/2)}]m_2(\frac{x}{2}) \geq (2c - 1)\frac{A'(x/2)}{A(x/2)} m_2(\frac{x}{2}) \geq 4c - 2.
\]

An easy calculation yields

\[
L\varphi(x) = 4c - 2 + \left[ \frac{A'(x/2)}{A(x/2)} - 2 \frac{A'(x)}{A(x)} \right] m_2(\frac{x}{2}) \leq 0.
\]

Therefore the assumption \( \varphi'(x_0) < 0 \) leads to \( \varphi''(x_0) > 0 \) and hence \( \varphi'(0) < 0 \) which is a contradiction to \( \varphi'(0) = 2m_2(0) - 2cm_2(0) = 0 \). This implies \( m_2(\frac{x}{2}) > \frac{\delta}{2} m_2(x) \) for every \( x \geq 0 \).

Now let \( \varepsilon > 0 \). Then there is an \( n \in \mathbb{N} \) with \( 2^{-n} < \varepsilon \) and we obtain \( m_2(\varepsilon x) \geq m_2(2^{-n} x) \geq (\frac{\delta}{2})^n m_2(x) \) for every \( x \geq 0 \).
8.4. Theorem: Suppose that 8.2 holds. Let \((X_n : n \geq 1)\) be a series of independent random variables such that
\[
\sum_{n \geq 1} \frac{1}{m_2(n)} V_*(X_n) < \infty.
\]

Then \(\frac{1}{n} S_n \to 0\) \(P\)-almost surely.

Proof: Let \(\varepsilon > 0\). For every \(r \in \mathbb{N}\) we define \(A_r := \{\frac{1}{n} S_n \geq \varepsilon \text{ for some } n \in [2^r, 2^{r+1}]\}\). As in the proof of 7.4 we obtain
\[
P(A_r) \leq \frac{1}{m_2(\varepsilon \cdot 2^r)} \cdot \sum_{j=1}^{2^{r+1}-1} V_*(X_j).
\]

By 6.4 and 8.2 there exists \(b \geq 2\) such that \(2m_2(x) \leq m_2(2x) \leq bm_2(x)\) for every \(x \geq 0\). For every \(j \geq 1\) choose \(r_j \in \mathbb{N}\) with \(2^{-r_j} \leq j < 2^{-r_j+1}\). Then
\[
\sum_{2^{r+1} > j} \frac{1}{m_2(2^r)} \leq \sum_{r \geq r_j} 2^{-r-r_j} \frac{1}{m_2(2^r)} = \frac{2}{m_2(2^{r_j})} \leq \frac{2b}{m_2(2^{r_j+1})} \leq \frac{2b}{m_2(j)}.
\]

Now choose \(\delta\) according to \(\varepsilon\) in 8.2. Then
\[
\sum_{r \geq 0} P(A_r) \leq \sum_{r \geq 0} \frac{1}{m_2(\varepsilon 2^r)} \cdot \sum_{j=1}^{2^{r+1}-1} V_*(X_j)
\]
\[
\leq \frac{1}{\delta} \sum_{j=1}^{\infty} V_*(X_j) \cdot \sum_{2^{r+1} > j} \frac{1}{m_2(2^r)}
\]
\[
\leq \frac{2b}{\delta} \sum_{j=1}^{\infty} \frac{1}{m_2(j)} V_*(X_j) < \infty.
\]

The Borel-Cantelli lemma finally accomplishes the proof.

8.5. Remark: Even if the sequence \((X_n : n \geq 1)\) is i.i.d. with \(V_*(X_n) < \infty\) the condition of the preceding theorem not necessarily holds since \(\sum_{n \geq 1} \frac{1}{m_2(n)}\) may be infinite. Since it follows from 6.4 that \(m_2(x) \leq x^{A'(x)}/A(x)\) this happens for example if \(\sum_{n \geq 1} \frac{1}{n A'(n)} = \infty\).

8.6. Corollary: Suppose that \(\{x \cdot A'(x)/A(x) : x > 0\}\) is bounded. Then for every i.i.d. sequence \((X_n : n \geq 1)\) of integrable \(\mathbb{R}_+\)-valued random variables \(\frac{1}{n} S_n \to 0\) \(P\)-almost surely.

Proof: Let \(a > 0\). As in the proof of 7.6 we consider the truncated variables \(X_n^a := 1_{\{x < an\}} \cdot X_n\) and define \(S_0^a := 0, S_n^a := S_{n-1}^a + X_n^a\) for \(n \geq 1\). Let \(\gamma\) be defined as in Lemma 7.6.
6.5. Since $X_1$ is integrable,

$$
\sum_{n=1}^{\infty} \frac{1}{m_2(n)} V_*(X_n^a) \leq \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{j=0}^{\infty} E(1_{a j \leq X_n^a < a(j + 1)}) \cdot m_2(X_n^a)
$$

$$
\leq \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{j=0}^{n-1} P\{a j \leq X_1 < a(j + 1)\} \cdot a^2(j + 1)^2
$$

$$
= \frac{1}{\gamma} \sum_{j=0}^{\infty} P\{a j \leq X_1 < a(j + 1)\} \cdot a^2(j + 1)^2 \sum_{n=j+1}^{\infty} \frac{1}{n^2}
$$

$$
\leq \frac{2}{\gamma} \sum_{j=0}^{\infty} P\{a j \leq X_1 < a(j + 1)\} \cdot a^2(j + 1)
$$

$$
\leq \frac{2}{\gamma} (a E(X_1) + a^2) < \infty
$$

we obtain from 8.3.1 and the preceding theorem that $\frac{1}{n} S_n^a \to 0$ P-a.s. for every $a > 0$. The rest of the proof is identical with 7.6.

8.7. **Remark:** By the same argument as in 7.8 we see that 8.6 is valid even if the starting point of the random walk $(S_n : n \in \mathbb{N})$ is not 0 but arbitrary.

8.8. **Remark:** Suppose that $\{x \cdot \frac{A(x)}{A(x)} : x > 0\}$ is bounded and let $0 < \beta < 2$. Then for every i.i.d. sequence $(X_n : n \geq 1)$ of $\mathbb{R}_+$-valued random variables such that $E(X_1^\beta)$ is finite, $\frac{1}{n^{1/\beta}} S_n \to 0$ P-almost surely. A similar result has been proved by Gallardo and Ries [14].

**Proof:** It is a straightforward generalization of theorem 8.4 that for every independant sequence $(Y_n : n \geq 1)$ such that $\sum_{n=1}^{\infty} \frac{1}{n^{2/\beta}} V_*(Y_n) < \infty$ we obtain $\frac{1}{n^{1/\beta}} \sum_{j=1}^{n} Y_j \to 0$ P-almost surely. If we choose $Y_n$ to be the truncated variable $X_n 1_{X_n \leq a_n^{1/\beta}}$ it follows as in the proof of 8.6 that $\frac{1}{n^{1/\beta}} S_n$ and $\frac{1}{n^{1/\beta}} \sum_{j=1}^{n} Y_j$ tend to the same limit almost surely.

·
9. The addition of small random variables

The last three paragraphs of this article are dealing with different forms of the central limit theorem for random walks on \((\mathbb{R}_+ , \ast)\). The main difference to the situation on \((\mathbb{R} , +)\) is the fact that the distributivity of the ordinary addition on \(\mathbb{R} \) is not valid, \(a \cdot (X + Y) = a \cdot X + a \cdot Y\), does not hold for \(\mathbb{R}^\ast\) except in the case of the Kingman hypergroup (see 3.2.3 and 3.2.4). In all other cases \(a \cdot (X \mathbb{R}^\ast Y)\) is not equal to \((aX) \mathbb{R}^\ast (aY)\) not even asymptotically as can be seen by comparing theorems 7.6, 9.3, 10.2, and 11.5: If \(\varrho = 0\) then \(\sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\) converges in distribution to a Gaussian distribution while \(\sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\) converges to a Rayleigh distribution. If \(\varrho > 0\) then \(\sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\) converges to a Gaussian distribution while \(\sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\) approaches \(+\infty\) and \(\frac{1}{\sqrt{n}}(\sum_{j=1}^{n} X_j - nE_\ast(X_1))\) converges to a normal distribution (in all cases the sequence \((X_n : n \geq 1)\) is supposed to be i.i.d. with finite variance).

We begin our discussion with the study of the behaviour of \(\sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\). This result is a special case of Trimèche [23], théorème 2. In this case it makes no difference whether \(\varrho > 0\) or \(\varrho = 0\).

9.1. Definition: Let \((\mathbb{R}_+ , \ast)\) be a Chébli-Trimèche hypergroup and \(\mathcal{F}\) its Fourier transform. The Gaussian distribution on \((\mathbb{R}_+ , \ast)\) with parameter \(t \geq 0\) is the unique probability measure \(\alpha_t\) on \(\mathbb{R}_+\) with

\[
\mathcal{F}\alpha_t(\lambda) = e^{-\frac{t}{2}(\lambda^2 + \lambda^2 \varrho)} \quad \text{for} \quad \lambda \in \mathbb{R}_+ \cup i[0, \varrho].
\]

9.2. Remark: The existence of \(\alpha_t\) is a consequence of theorem 9.3. Although \(\alpha_t\) is uniquely determined for every given hypergroup \((\mathbb{R}_+ \ast)\), a different hypergroup will in general have different Gaussian measures. The family of Gaussian measures \((\alpha_t : t \geq 0)\) forms a convolution semigroup. It is easily calculated that the \(\ast\)-expectation of \(\alpha_t\) is \(\varrho t\) and the \(\ast\)-variance is \(t\). Explicit formulas for the density of a Gaussian distribution can be found in 11.2 and Karpelevich, Tutubalin, Shur [20].

9.3. Theorem (Trimèche): Let \((X_n : n \geq 1)\) be a sequence of i.i.d. random variables such that \(E(X_1^2) < \infty\). Then \(Y_n := \sum_{j=1}^{n} \frac{1}{\sqrt{n}} X_j\) converges in distribution to the Gaussian distribution \(\alpha_t\) with parameter \(t = \frac{1}{\alpha + 1} E(X_1^2)\).

Proof: It is easily checked in both cases \(\varrho > 0\) and \(\varrho = 0\) that conditions (i), (ii), and (iii) of Trimèche [23], théorème 2 are valid for \(\mu_{nj} := P_{\mathbb{R}^\ast X_j}\). Since it is easier to prove this theorem directly than to verify these conditions we give an outline of the proof in the following.

Let \(\lambda \in \mathbb{R}_+ \cup i[0, \varrho]\) be fixed. Since there exists \(\gamma > 0\) such that \(1 \geq \varphi_\lambda(x) \geq 1 - \gamma x^2\) for every \(x \geq 0\) it follows easily that \(n \cdot (\varphi_\lambda(X_1/\sqrt{n}) - 1 - \frac{1}{2}\varphi''_\lambda(0) \cdot X_1^2 / n)\) is bounded by a multiple of \(X_1^2\).
(which by assumption is integrable) and hence by Taylor’s formula

\[ n \, E \left( \varphi_{\lambda}(X_1/\sqrt{n}) - 1 - \frac{1}{2} \varphi''_{\lambda}(0) X_1^2 / n \right) \to 0 \quad \text{as } n \to \infty. \]

But from \( E(\varphi_{\lambda}(X_1/\sqrt{n})) = 1 + \frac{1}{2} \varphi''_{\lambda}(0) \frac{E(X_1^2)}{n} + o\left( \frac{1}{n} \right) \) we conclude

\[
\mathcal{F} P_{\lambda_n}(\lambda) = E(\varphi_{\lambda}(X_1/\sqrt{n}))^n \\
= \left( 1 + \frac{1}{2} \varphi''_{\lambda}(0) \frac{E(X_1^2)}{n} + o\left( \frac{1}{n} \right) \right)^n \\
\to e^{\frac{1}{2} \varphi''_{\lambda}(0) E(X_1^2)} \\
= e^{-t(\lambda^2 + \epsilon^2)} \\
= \mathcal{F} \alpha_t(\lambda)
\]

for \( \lambda \in \mathbb{R}_+ \cup \{i[0, \epsilon]\} \). Using the continuity theorem we obtain the desired result.  \[ \blacksquare \]
10. The central limit theorem in the case of exponential growth

We now return to the study of the asymptotic behaviour of the random variable $S_n$ as defined in 7.2 and will suppose that the variables $X_1, X_2, \ldots$ are i.i.d. with finite variance. In this paragraph we will deal with the case $\rho > 0$. We already know (Theorem 7.4) that the random variable $S_n$ will be close to $m_1^{-1}(n \cdot E_x(X_1))$. Because of $V_x(S_n) = n V_x(X_1)$ we expect the standard deviation to be of order $\sqrt{n}$. Therefore it is not too surprising that $\frac{S_n - m_1^{-1}(n \cdot E_x(X_1))}{\sqrt{n} V_x(X_1)}$ converges in distribution; but it is perhaps unexpected that the limiting distribution is the normal law on $\mathbb{R}$, independent of which hypergroup $(\mathbb{R}_+, \ast)$ had been considered (as long as $\rho > 0$). In order to prove this we need some information about the asymptotic behaviour of the characters $\varphi_{i \tilde{e} + \lambda}$ as $\lambda \to 0$.

10.1. Lemma: Let $a > 0$. Then

$$\sup_{t \in [0, na]} |\varphi_{i a - \lambda / \sqrt{n}}(t) - e^{i \lambda \frac{1}{\sqrt{n}} m_1(t)}| \to 0 \quad \text{as } n \to \infty.$$ 

Proof: Let $r = \sqrt{n}$ and consider the function $\psi_r : \mathbb{R}_+ \to \mathbb{C}$ defined by

$$\psi_r(x) := \varphi_{i a - \lambda / r}(rx) \cdot e^{-i \lambda \frac{m_1(rx)}{r}} \quad (x \in \mathbb{R}_+).$$

It is easily derived from the differential equation 2.2 for $\varphi_{i \tilde{e} - \lambda / r}$ that $\psi_r$ satisfies

$$\psi''_r(x) + \left( r \frac{A'(rx)}{A(rx)} + 2i \lambda m'_1(rx) \right) \psi'_r(x) + \lambda^2 (1 - m'_1(rx)^2) \psi_r(x) = 0$$

for $x \in \mathbb{R}_+$ as well as $\psi_r(0) = 1, \psi'_r(0) = 0$.

If we multiply this equation by $2\psi'_r(x)$, take the real part, and integrate, we obtain

$$|\psi'_r(x)|^2 + 2r \int_0^x \frac{A'(rt)}{A(rt)} |\psi'_r(t)|^2 \, dt + 2\lambda^2 \int_0^x (1 - m'_1(rt)^2) \Re(\psi_r(t)\psi'_r(t)) \, dt = 0.$$ 

It follows from the Laplace representation 2.3 that the modulus of $\varphi_{i \tilde{e} - \lambda / r}$ is bounded by 1 and hence $\| \psi_r \|_{\infty} \leq 1$. Therefore $\Re(\psi_r(t)\psi'_r(t)) \geq -|\psi'_r(t)|$ and we conclude

$$\int_0^x |\psi'_r(t)|^2 \, dt \leq \frac{1}{4r} \left\{ |\psi'_r(x)|^2 + 2r \int_0^x \frac{A'(rt)}{A(rt)} |\psi'_r(t)|^2 \, dt \right\} \leq \frac{\lambda^2}{2r} \int_0^x (1 - m'_1(rt)^2) |\psi'_r(t)| \, dt \leq \frac{\lambda^2}{2r} \int_0^x (1 - m'_1(rt)^2) |\psi'_r(t)| \, dt \leq \frac{\lambda^2}{2r} \left( \int_0^x (1 - m'_1(rt))^2 \, dt \cdot \int_0^x |\psi'_r(t)|^2 \, dt \right)^{1/2}.$$ 

30
From this we obtain by another application of Hölder's inequality

\[ |1 - \psi_r(x)| \leq \int_0^x |\psi_r'(t)| dt \]

\[ \leq \left\{ x \cdot \int_0^x |\psi_r'(t)|^2 dt \right\}^{1/2} \]

\[ \leq \sqrt{\pi} \cdot \frac{\lambda^2}{pr} \cdot \left\{ \int_0^x (1 - m_1^r(rt))^2 dt \right\}^{1/2} \]

\[ \leq \frac{\lambda^2}{pr} \cdot \left\{ \int_0^x (1 - m_1^r(rt)) dt \right\}^{1/2} \]

\[ = \frac{\lambda^2}{p} \cdot \sqrt{\pi} \cdot \sqrt{\frac{x}{m_1^r(rx)}}. \]

Therefore

\[ \sup_{t \in [0, a]} |\psi_r(t) - 1| \leq \frac{\lambda^2}{p} \sup_{u \in [0, a]} \frac{\sqrt{ru}}{r} \sqrt{ru - \frac{1}{r} m_1(r^2 u)} \]

\[ = \frac{\lambda^2}{p} \sup_{u \in [0, a]} u \cdot \sqrt{1 - \frac{1}{r^2 u} m_1(r^2 u)} \]

which converges to 0 by 5.7 and Dini's theorem, and the assertion follows since \(e^{i\lambda n_1} \) is of modulus one.

10.2. Theorem: Suppose that \(\sigma^2 := V_*(X_1)\) is finite and let \(\mu := E_*(X_1)\).

Then \(Z_n := \frac{S_n - n^{-1}(n\mu)}{\sqrt{n}}\) converges in distribution to the normal law \(N(0, \sigma^2)\).

Proof: Let \(\lambda \in \mathbb{R}\). Then by 4.12

\[ e^{-i\lambda \mu \sqrt{n}} \cdot E(\varphi_{i\lambda - \lambda \sqrt{n}}(S_n)) \]

\[ = e^{-i\lambda \mu \sqrt{n}} \cdot E(\varphi_{i\lambda - \lambda \sqrt{n}}(X_1)) \]

\[ = \left[(1 - \frac{i\lambda \mu}{\sqrt{n}} - \frac{\lambda^2 \mu^2}{2n} + o(\frac{1}{n})) \cdot (1 + \frac{i\lambda}{\sqrt{n}} E(m_1(X_1)) - \frac{\lambda^2}{2n} E(m_2(X_1)) + o(\frac{1}{n})) \right]^n \]

\[ = \left[(1 - \frac{i\lambda \mu}{\sqrt{n}} - \frac{\lambda^2 \mu^2}{2n} + o(\frac{1}{n})) \cdot (1 + \frac{i\lambda \mu}{\sqrt{n}} - \frac{\lambda^2}{2n}(\sigma^2 + \mu^2) + o(\frac{1}{n})) \right]^n \]

\[ = \left[1 - \frac{\lambda^2}{2n} \cdot \sigma^2 + o(\frac{1}{n}) \right]^n \]

and hence

\[ \lim_{n \to \infty} e^{-i\lambda \mu \sqrt{n}} E(\varphi_{i\lambda - \lambda \sqrt{n}}(S_n)) = e^{-\lambda^2 \frac{\sigma^2}{2n}}. \]

It follows from the preceding lemma that \(|\varphi_{i\lambda n_1(S_n)} - \varphi_{i\lambda - \lambda \sqrt{n}}(S_n)|\) converges to 0 uniformly on \(\{S_n \leq n(\mu + 1)\}\). By the law of large numbers 7.4 the probability of this event approaches 1. Since
$e^{i\lambda \sqrt{n}}$ is of modulus 1 we conclude that

$$
\lim_{n \to \infty} E\left( e^{i\frac{\lambda}{\sqrt{n}}(m_1(S_n) - n\mu)} \right) = \lim_{n \to \infty} e^{-i\lambda \mu \sqrt{n}} \cdot E\left( e^{i\frac{\lambda}{\sqrt{n}}m_1(S_n)} 1_{\{S_n \leq n(\mu + 1)\}} \right)
$$

$$
= \lim_{n \to \infty} E\left( e^{i\frac{\lambda}{\sqrt{n}}\left(\varphi_{n}(S_n) - \lambda \sqrt{n}\right)} 1_{\{S_n \leq n(\mu + 1)\}} \right)
$$

$$
= \lim_{n \to \infty} e^{-i\lambda \mu \sqrt{n}} \cdot E\left( e^{i\frac{\lambda}{\sqrt{n}}\varphi_{n}(S_n)} \right)
$$

$$
= e^{-\lambda^2 \sigma^2}
$$

$$
=N(0, \sigma^2)(\lambda).
$$

By the continuity theorem for the Fourier transformation $\frac{1}{\sqrt{n}}(m_1(S_n) - n\mu)$ converges in distribution to $N(0, \sigma^2)$.

Finally we use the inequality

$$
|(m_1(S_n) - S_n) - (n\mu - m_1^{-1}(n\mu))| \leq \left( (m_1^{-1})'\left(\min\{m_1(S_n), n\mu\}\right) - 1\right) |m_1(S_n) - n\mu|
$$

and $(m_1^{-1})'(x) \searrow 1$ as $x \to \infty$ to conclude that $\frac{1}{\sqrt{n}}(S_n - m_1^{-1}(n\mu))$ converges to $N(0, \sigma^2)$ in distribution.  \[T\]

10.3. Remark: If $I(\mathbb{R}, *)$ is the hypergroup with $A(x) = (\cosh x)^2$ (this is not a Chébli-Trimèche hypergroup, see Zeuner [28]) then the assertions of 7.4, 7.6, 7.7, and 10.2 remain valid. In this case $\varrho = 1$, $m_1(x) = x \tanh x$, $m_2(x) = x^2$, and $\varphi_\lambda(x) = \frac{\cosh \lambda x}{\cosh x}$ ($x \geq 0$, $\lambda \in \mathbb{C}$) and it is easily checked that the facts used in the proofs of 7.4, 7.6, 7.7, and 10.2 also hold in this situation.
11. The central limit theorem in the case of polynomial growth

As in the proof of the law of large numbers we have to impose an additional condition on the hypergroup \((\mathbb{R}_+\times, \cdot)\) in order to obtain a central limit theorem in the case \(\varrho = 0\). This results from the fact that the asymptotic behaviour of \(\varphi_\lambda (x)\) is strongly affected by the properties of \(\frac{A'(x)}{A(x)}\) as \(x \to \infty\) if \(\varrho = 0\). Unlike in the case \(\varrho > 0\) where the limiting distribution of the normalised random variable did not depend on \(A\), we obtain as limiting distribution in the case \(\varrho = 0\) the Rayleigh distribution \(\varrho_\alpha\) if \(2\alpha + 1 = \lim_{x \to \infty} x \cdot \frac{A'(x)}{A(x)}\) (the existence of this limit is at the same time the additional condition to be required; it implies that the hypergroup is of polynomial growth).

These distributions also occur in the central limit theorem of L. Gallardo on the hypergroup on \(\mathbb{N}\) defined by the Gegenbauer polynomials (see Gallardo [13]).

11.1. Notation: For every \(\beta \geq 0\) let \(\alpha := \frac{\beta-1}{2}\) and \(j_\alpha\) be the modified Bessel function on \(\mathbb{R}_+\) defined by \(j_\alpha (x) := \frac{\Gamma (\alpha+1)}{\Gamma (\alpha+1/2)} J_\alpha (x)\) for \(x > 0\), \(j_\alpha (0) := 0\) where \(J_\alpha\) is the usual Bessel function in the sense of Watson [26], p. 40. \(j_\alpha\) is the solution of the initial value problem

\[
j''_\alpha (x) + \frac{\beta}{x} j'_\alpha (x) + j_\alpha (x) = 0 \quad (x > 0), \quad j_\alpha (0) = 1, \quad j'_\alpha (0) = 0
\]

and therefore the functions \(x \mapsto j_\alpha (\lambda x) (x \geq 0)\) for \(\lambda \in \mathbb{R}_+\) constitute the dual of the Kingman hypergroup with \(A(x) = x^\beta\).

11.2. Notation: The Rayleigh distribution \(\varrho_\alpha\) with parameter \(\alpha = \frac{\beta-1}{2} \geq -\frac{1}{2}\) is defined by its Lebesgue density \(x \mapsto c_\alpha x^{2\alpha+1} e^{-x^2/2}\) on \(\mathbb{R}_+\) where \(c_\alpha = \frac{1}{\Gamma (\alpha+1/2)}\). In the case that \(\beta\) is an integer, \(\varrho_\alpha\) is the radial part of the \(\beta\)-dimensional symmetric normal distribution. From the integral representation of the Bessel function (Watson [26], p. 394) it is easily obtained that \(\int j_\alpha (\lambda x) \varrho_\alpha (dx) = e^{-\lambda^2/2}\) for every \(\lambda \geq 0\) (see Finckh [10], 9.8.1) and hence \(\varrho_\alpha\) is a Gaussian distribution for the Kingman hypergroup.

As in the previous paragraph we need information about the asymptotic behaviour of \(j_\alpha\) in order to prove a central limit theorem.

11.3. Lemma: Assume that \(\beta := \lim_{x \to \infty} x \cdot \frac{A'(x)}{A(x)}\) exists and put \(\alpha := \frac{\beta-1}{2}\). Then for every \(a > 0\)

\[
\sup_{t \in [0,a]} |\varphi_{1/r} (x) - j_\alpha \left( \frac{x}{r} \right) | \to 0 \text{ as } r \to \infty.
\]

Proof: For every \(r > 0\) we consider the function \(\delta_r : x \mapsto \varphi_{1/r} (xr) - j_\alpha (x)\) on \([0,a]\). It
follows from $\varphi''_{1/r} + \frac{A'}{A} \varphi'_{1/r} + \frac{1}{r^2} \varphi_{1/r} = 0$ and the differential equation (11.1) for $j_\alpha$ that

$$\delta''(x) + r \frac{A'(rx)}{A(rx)} \delta'(x) + \delta(x) = \frac{j_\alpha(x)}{x} \cdot \left( r x \frac{A'(rx)}{A(rx)} - \beta \right)$$

and

$$\delta(0) = \delta'(0) = 0.$$

If we denote the right hand side of the differential equation by $f_r(x)$, it follows from the asymptotic behaviour of $A'/A$ near the origin (see 2.2) that $\{f_r : r > 0\}$ is uniformly bounded on $[0, a]$ and hence the assumption of the lemma implies $f_r \to 0$ in the $L^1([0, a])$-norm as $r \to \infty$.

Let us now define $\Delta_r(x) := \max \{|\delta'_r(t)| : t \in [0, x]\}$. If we multiply the differential equation for $\delta_r$ by $2\delta'_r(x)$ we obtain by integration

$$\Delta_r(x)^2 \leq \max \{\delta'_r(t)^2 + 2 \int_0^t r \frac{A'(ry)}{A(ry)} \cdot \delta'_r(y)^2 \, dy + \delta_r(t)^2 : t \in [0, x]\}$$

$$= \max \{2 \int_0^t \delta'_r(y) f_r(y) \, dy : t \in [0, x]\}$$

$$\leq 2 \cdot \Delta_r(x) \cdot \int_0^x |f_r(y)| \, dy$$

and therefore

$$|\delta'_r(x)| \leq 2 \int_0^a |f_r(y)| \, dy \quad \text{for every } x \in [0, a].$$

This implies that $\delta_r \to 0$ uniformly on $[0, a]$ and thus the lemma is proved.

11.4. **Theorem**: Suppose that $\beta := \lim_{x \to \infty} \frac{A'(x)}{A(x)}$ exists and that $X_1, X_2, \ldots$ is a sequence of i.i.d. non-negative random variables such that $\sigma^2 := V_\sigma(X_1) < \infty$. Then $\frac{1}{\sigma \sqrt{n}} \cdot S_n$ converges in distribution to the Rayleigh measure $\varrho_\alpha$ (where $\alpha = \frac{\beta - 1}{2}$).

**Proof**: Let $\lambda \geq 0$. It follows from 4.7 that

$$E(\varphi_{\lambda / \sqrt{n}}(S_n)) = E(\varphi_{\lambda / \sqrt{n}}(X_1))^n$$

$$= (1 - \frac{\lambda^2}{2n} \sigma^2 + o(\frac{1}{n}))^n$$

$$\to e^{-\frac{\lambda^2 \sigma^2}{2}}.$$ 

Let $\varepsilon > 0$. By 6.5 we may choose $r > 0$ such that, uniformly in $n$,

$$P\{S_n > r \sqrt{n}\} = P\{m_2(S_n) > m_2(r \sqrt{n})\}$$

$$\leq \frac{V_\sigma(S_n)}{m_2(r \sqrt{n})}$$

$$\leq \frac{n V_\sigma(X_1)}{\gamma r^2 n}$$

$$\leq \varepsilon$$
where $\gamma$ has been defined in 6.5. It follows from 11.3 that

$$|\varphi_{\lambda/\sqrt{n}}(x) - j_\alpha(\frac{\lambda}{\sqrt{n}} x)| < \varepsilon \text{ for } x \in [0, r\sqrt{n}]$$

if $n$ is large enough. Hence

$$|E(\varphi_{\lambda/\sqrt{n}}(S_n)) - E(j_\alpha(\frac{\lambda}{\sqrt{n}} S_n))|$$

$$\leq E\left(|\varphi_{\lambda/\sqrt{n}}(S_n) - j_\alpha(\frac{\lambda}{\sqrt{n}} S_n)| \cdot 1_{\{S_n \leq r\sqrt{n}\}}\right) + 2P\{S_n > r\sqrt{n}\}$$

$$\leq 3\varepsilon$$

for large $n$ and this implies

$$\lim_{n \to \infty} E(j_\alpha(\frac{\lambda}{\sqrt{n}} S_n)) = \lim_{n \to \infty} E(\varphi_{\lambda/\sqrt{n}}(S_n)) = e^{-\lambda^2/2}.$$

By the continuity theorem for the Hankel transformation (which is the Fourier transformation of the Kingman hypergroup (see Kingman [21], p. 22, Finckh [10], 3.6)) and 11.2, $\frac{1}{\sqrt{n}} S_n$ converges to $\varphi_\alpha$ in distribution. 

$\blacksquare$
Bibliography


