Notes on the Applicability of Contraction Method for Stable Limit Laws

Diplomarbeit

Bastian Werth

Goethe Universität Frankfurt am Main
Fachbereich Informatik und Mathematik

Prof. Dr. Ralph Neininger

Mai 2010

Contents

SECTION 1

Introduction

The contraction method has its origin in the analysis of recursive algorithms, namely in the work of Rösler 1991[11], where he proofs a weak limit theorem for the complexity of Quicksort. For a general introduction we refer to Rösler and Rüschendorf 2004[12].

A typical approach in the analysis of algorithms in general is to define and analyze certain complexity measures, depending on the input kind and nature[1] of the problem that is to be solved. Usually, a certain quantity within the algorithm is identified, which can be seen as the main driver of the overall complexity. In the analysis of Quicksort in Rösler 1991[11], the number of key comparisons is taken as the main driver of overall time-complexity, and hence further established.

From an algorithmic problem of, for instance, divide-and-conquer or recursive type one can often quite naturally obtain equations for its complexity X_n of the form

$$\mathcal{L}\left(X_n\right) = \mathcal{L}\left(\sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}\right) , \tag{1.1}$$

where $(X_j^{(1)})_{j\geqslant 0}, \ldots, (X_j^{(K)})_{j\geqslant 0}, (A^{(n)}, I^{(n)}, b^{(n)})$ are independent, $I^{(n)} = (I_r^{(n)})$ is a random vector with values in $\{1, \ldots, n\}^K$, $A^{(n)} = (A_r^{(n)})$ denotes a random vector in $\mathbb{R}^K$, $\mathcal{L}(X_j^{(r)}) = \mathcal{L}(X_j)\ \forall\, 1 \leqslant j \leqslant n, 1 \leqslant r \leqslant K$, and$b^{(n)}$ is a real-valued random variable. Here, $\mathcal{L}(X)$ denotes the distribution of X.

A convergence of $\mathcal{L}(X_n)$ to a limit distribution μ is usually obtained from $d(\mathcal{L}(X_n), \mu) \to 0$ with an appropriate metric d. The recursive equation (1.1) can also be written as a contraction $T : M \to M$ on an appropriate metric space of distributions (M, d). Therefore, the overall approach is known as "Contraction Method".

To illustrate the basic concept of the contraction method we take an arbitrary distribution μ on $\mathfrak{B}(\mathbb{R})$ and define a sequence of distributions μ_n recursively by

$$\mu_1 = \mu, \quad \mu_{n+1} = \mathcal{L}\left(2^{-1/2}\, X_n + 2^{-1/2}\, X_n'\right) , \tag{1.2}$$

where X_n and X_n' are independent with $\mathcal{L}(X_n) = \mathcal{L}(X_n') = \mu_n$. We will see that μ_n tends to the standard normal distribution under some constraints.

[1] As an example take the length of a list of certain objects that the algorithm takes as input.

As metric we take the Zolotarev metric[2] of order 3, denoted by ζ_3. Since ζ_3 only depends on the marginal distributions we write, for real-valued X and Y, $\zeta_3(X,Y)$ instead of $\zeta_3(\mathcal{L}(X),\mathcal{L}(Y))$. ζ_3 is ideal of order three, i.e. for Z independent of (X,Y), $r \in \mathbb{R}$ we have

$$\zeta_3\left(X+Z, Y+Z\right) \leqslant \zeta_3\left(X,Y\right) \ ,$$

$$\zeta_3\left(rX, rY\right) = |r|^3\,\zeta_3\left(X,Y\right) \ .$$

Let W be a standard-normally distributed random variable. The convolution property of normal distributions states that

$$\mathcal{L}\left(W\right) = \mathcal{L}\left(2^{-1/2}\,W + 2^{-1/2}\,W'\right) \ ,$$

where W and W' are independent with $\mathcal{L}(W) = \mathcal{L}(W')$. Convergence in ζ_3 implies weak convergence, hence, we have to find out if $\zeta_3(X_n, W)$ tends to zero in order to obtain $\mathcal{L}(X_n) \to \mathcal{L}(W)$. Applying the above notations and properties it is straightforward to obtain

$$\begin{aligned}
\zeta_3\left(X_{n+1}, W\right) &= \zeta_3\left(2^{-1/2}\,X_n + 2^{-1/2}\,X_n', 2^{-1/2}\,W + 2^{-1/2}\,W'\right) \\
&= 2^{-3/2}\,\zeta_3\left(X_n + X_n', W + W'\right) \\
&\leqslant 2^{-3/2}\,\left(\zeta_3\left(X_n + X_n', W + X_n'\right) + \zeta_3\left(W + X_n', W + W'\right)\right) \\
&\leqslant 2^{-3/2}\,\left(\zeta_3\left(X_n, W\right) + \zeta_3\left(X_n', W'\right)\right) \\
&= 2^{-1/2}\,\zeta_3\left(X_n, W\right) \ . \qquad (1.3)
\end{aligned}$$

Writing out the recursion we clearly observe one crucial point in using contraction method,

$$\zeta_3\left(X_n, W\right) \leqslant 2^{-n/2}\,\zeta_3\left(X, W\right) \ . \qquad (1.4)$$

The left handside in (1.4) convergences to zero iff $\zeta_3\left(X,W\right)$ is finite. The finiteness of ζ_3 can be obtained using the following criterion. For real-valued random variables X and Y it holds that

$$\left.\begin{array}{l}\mathbb{E}\left[X^k - Y^k\right] = 0 \ \text{ for } \ k = 1,2 \\ \mathbb{E}\left[|X|^3 + |Y|^3\right] < \infty\end{array}\right\} \Rightarrow \zeta_3\left(X,Y\right) < \infty \ . \qquad (1.5)$$

Therefore, X is supposed to have a first moment of zero, a second moment of one and a finite absolute third moment. Note that the above estimate is an intuitive proof of a central limit law.

The normal distribution is a member of the stable distributions family. A random variable W is called stable of order $\alpha \in (0,2]$ if

$$\mathcal{L}\left(W\right) = \mathcal{L}\left(n^{-1/\alpha}\,W_1 + \ldots n^{-1/\alpha}\,W_n\right) \ ,$$

[2]For a systematic use of the Zolotarev metric for distributional recurrences such as (1.1) see Neininger and Rüschendorf 2004[7].

for all $n \in \mathbb{N}$, (W_i) independent with $\mathcal{L}(W_i) = \mathcal{L}(W)$. Since the normal distribution is stable of order 2, the central limit law can also be called "stable limit law of order 2". In general, stable limit laws of every order $\alpha \in (0, 2]$ exist.

For stable limit laws of order less than 2 the finiteness of the recursion origin is more problematic, since stable distributions of order $\alpha < 2$ do not have absolute finite moments of order $\geqslant \alpha$. Hence, a generalized version of the criterion (1.5) does not help here. It is not clarified yet weather another criterion for the finiteness of the Zolotarev metric ζ_s, $s > \alpha$, is applicable in this setting. Note that $s > \alpha$ has to be chosen in order to obtain contraction properties.

After introducing the basic notations and definitions in Section 2, we use a result from Johnson and Samworth 2005[4] to proof the existence of $\lambda > \alpha$ such that $\zeta_\lambda(X, W) < \infty$, for W α-stable and X in the domain of strong normal attraction of W, in Section 3. In this case, a variation of the estimate (1.3) leads to the classical stable limit laws.

Nevertheless, a result of this kind is not applicable for the analysis of the algorithms complexity, since the measured complexity is always finite and only might "growth into" a stable distribution in the limit. Therefore, under the use of the minimal ℓ_p metric, in Section 4 we present a stable limit law for scaled and centered sums of independent random variables (Z_n), satisfying

$$\sum_{i=1}^{n} \ell_\beta \left(Z_i, W\right) = o_n \left(n^{1/\alpha}\right) \ ,$$

where W is α-stable distributed and $1 \leqslant \beta < \alpha$.

SECTION 2

Basic Notations and Definitions

The **Landau Notation** is widely-used for the description of the asymptotic size of a sequence or function in a certain limit. For real-valued sequences (a_n) and (b_n) the following symbols are defined.

i. $a_n = \mathcal{O}_n(b_n) :\Leftrightarrow \exists\, n_0 \in \mathbb{N},\ C \in \mathbb{R}_{>0} : |a_n| \leqslant C|b_n|\ \forall\, n \geqslant n_0$.

ii. $a_n = o_n(b_n) :\Leftrightarrow\ \forall\, \kappa \in \mathbb{R}_{>0}\ \exists\, n_0 \in \mathbb{N} : |a_n| < \kappa|b_n|\ \forall\, n \geqslant n_0$.

These symbols are also used for real functions, as

$$g(x) = \mathcal{O}_x\left(f(x)\right) \text{ at } \infty\ :\Leftrightarrow\ \exists\, K > 0, z > 0 : |g(x)| \leqslant K|f(x)|\ \forall\, x \geqslant z\ .$$

A function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is called **slowly varying** at infinity if

$$\lim_{x\to\infty} \frac{f(ax)}{f(x)} = 1\ \forall\ a > 0\ .$$

SECTION 2.1

Stable Distributions

For $\alpha \in (0, 2]$ a real-valued random variables W is defined as **stable of order** α if

$$W_1 + W_2 + ... + W_n \overset{D}{=} n^{\frac{1}{\alpha}} W \tag{2.1}$$

$\forall\ n \in \mathbb{N}$, (W_i) independent with $\mathcal{L}(W_i) = \mathcal{L}(W)$ The definition is applied in the same manner on distributions. The case of $\alpha = 2$ leads to the normal distribution and is special, since for the normal distribution absolute moments of all orders exist. Under these circumstances the contraction method can be used for stable limit laws of order 2, as presented in Section 1. We will focus on $\alpha < 2$ from now on.

The finiteness of both minimal ℓ_p and Zolotarev metric depends on the (asymptotical) tail behavior of the respective random variables. This will be seen in more details in the later sections. Therefore, it is important to understand the tail behavior of stable random variables. The following estimate is given by Mijnheer 1986[6] (see (2.2)). For $\alpha \in (0, 2)$ and W an α–stable random variable there exist constants $c_1, c_2 \in \mathbb{R}_{\geqslant 0}, c_1 + c_2 > 0$ such that

$$\mathbb{P}\left(W < x\right) = \frac{c_1}{|x|^\alpha} + \mathcal{O}_x\left(\frac{1}{|x|^{2\alpha}}\right) \text{ for } x \to -\infty \tag{2.2}$$

and

$$\mathbb{P}(W > x) = \frac{c_2}{x^\alpha} + \mathcal{O}_x\left(\frac{1}{x^{2\alpha}}\right) \quad \text{for} \quad x \to \infty \; . \tag{2.3}$$

$\alpha-$stable distributions are determined by a set of four parameters, the index of stability α, the centering constant b, and the constants c_1 and c_2 from (2.2) and (2.3) (see Appendix A). We denote the stable distribution for a fixed set of parameters by $\mathrm{St}(\alpha, b, c_1, c_2)$.

SECTION 2.2

Domains of Attraction

When establishing limit laws for scaled sums of iid random variables (X_n), $\mathcal{L}(X_i) = X$, one should ask about the distributions $\mathcal{L}(X)$ in scope of a limit distribution μ. More precisely, it needs to be examined if real-valued sequences (a_n) and (b_n) exist such that

$$\mathcal{L}\left(a_n^{-1} \sum_{i=1}^{n} X_i - b_n\right) \to \mu \; . \tag{2.4}$$

All the distributions $\mathcal{L}(X)$ satisfying (2.4) with appropriate (a_n) and (b_n) are said to build the **domain of attraction** of μ. It can be shown that only stable distributions do have a domain of attraction. We set

$$\mathrm{DoA}_\alpha := \Bigg\{\mathcal{L}(X) \mid \exists\, (a_n), (b_n) \subset \mathbb{R}, W\, \alpha-\text{ stable}: \\ \mathcal{L}\Bigg(a_n^{-1} \sum_{i=1}^{n} X_i - b_n\Bigg) \to \mathcal{L}(W)\Bigg\} \; .$$

Moreover, all the distributions belonging to DoA_α with $\alpha < 2$ are exactly known by their tail behavior as follows. For constants $c_1, c_2 \in \mathbb{R}_{\geqslant 0}$ and h slowly varying at inifinity we have

$$\mathcal{L}(X) \in \mathrm{DoA}_\alpha \Leftrightarrow \begin{cases} \mathbb{P}(X < -x) = \frac{c_1 + o_x(1)}{|x|^\alpha} h(|x|) \text{ as } x \to -\infty \\ \mathbb{P}(X > x) = \frac{c_2 + o_x(1)}{x^\alpha} h(x) \text{ as } x \to \infty \end{cases} . \tag{2.5}$$

For more details see Petrov 1975[9] (Theorem 14). Using the formula

$$\mathbb{E}[X^z] = \int_0^\infty \mathbb{P}\left(X > x^{1/z}\right) dx$$

for non-negativ X and $z > 0$ we see that all random variables in the domain of attraction of an α-stable random variable do have finite absolute moments up to an order of $\alpha - \epsilon, \epsilon > 0$ only, i.e.

$$\mathcal{L}(X) \in \mathrm{DoA}_\alpha \;\Rightarrow\; \mathbb{E}[|X|^s] \begin{cases} < \infty & , s < \alpha \\ = \infty & , s \geqslant \alpha \end{cases} . \tag{2.6}$$

The treatment of tail behavior for random variables in DoA_α as given in (2.5) is complex and not obviously managable in our setting. Therefore, we focus on a subspace of DoA_α, the domain of strong normal attraction. By definition, a distribution $\mathcal{L}(X)$ belongs to this subspace iff, for $c_1, c_2 \in \mathbb{R}_{\geqslant 0}, \gamma > 0$,

$$\mathbb{P}\left(X < x\right) = \frac{c_1}{|x|^\alpha} + \mathcal{O}_x\left(\frac{1}{|x|^{\alpha+\gamma}}\right) \text{ as } x \to -\infty$$

and

$$\mathbb{P}\left(X > x\right) = \frac{c_2}{x^\alpha} + \mathcal{O}_x\left(\frac{1}{x^{\alpha+\gamma}}\right) \text{ as } x \to \infty\ . \tag{2.7}$$

Note that, as an implication of (2.2) and (2.3), every α-stable distribution belongs to the domain of strong normal attraction.

SECTION 2.3

Probability Metrics

As discussed in the Section 1, probability metrics are the major tool of contraction method. They are also needed to estimate convergence rates of limit theorems in general. For an overview on the field of probability metrics one can refer to Rachev 1991[10], especially the summary table on pages 464-477.

Basically, metrics can be defined on a subspace of distributions M as well as on a subspace of random variables $\mathfrak{X}$. These two types of probability metrics are linked, among others, by the following definitions. A metric d on M is said to be the **minimal metric** w.r.t. a metric d^* on $\mathfrak{X}$ if, for random variables X and Y,

$$d\left(\mathcal{L}\left(X\right), \mathcal{L}\left(Y\right)\right) = \inf_{(X,Y)} d^*\left(X, Y\right)\ .$$

Furthermore, a metric d^* on $\mathfrak{X}$ is called **simple** if it depends on the marginal distribution of respective random variables only. In that case d^* induces a metric d on M,

$$d\left(\mathcal{L}(X), \mathcal{L}(Y)\right) = d^*\left(X, Y\right)\ .$$

A metric d^* on $\mathfrak{X}$ is called **ideal** of order $s \in \mathbb{R}_{>0}$, if, for every $r \in \mathbb{R}$ and random variable Z independent from (X, Y),

$$\begin{aligned} d^*\left(X+Z, Y+Z\right) &\leqslant d^*\left(X, Y\right)\ , \\ d^*\left(rX, rY\right) &= |r|^s\, d^*\left(X, Y\right)\ . \end{aligned}$$

Here, two different metrics are used, the minimal ℓ_p metric and the Zolotarev metric ζ_s.

SECTION 2.3.1 Minimal ℓ_p Metric

The L_p metric is used in many applications, as it is complete on the space of absolute p-integrable functions. For $p \geqslant 1$ and random variables $X, Y : \Omega \to \mathbb{R}$ it is defined as

$$L_p(X,Y) := \left(\int_{\Omega\times\Omega} |x-y|^p \, d(\mathbb{P}_X, \mathbb{P}_Y)(x,y) \right)^{1/p} = \mathbb{E}\left[|X-Y|^p\right]^{1/p} . \tag{2.8}$$

Here, $(\mathbb{P}_X, \mathbb{P}_Y)$ denotes the product measure of $\mathbb{P}_X$ and $\mathbb{P}_Y$. On the space of all distributions on $\mathfrak{B}(\mathbb{R})$, the minimal ℓ_p metric is defined as the minimal metric w.r.t. L_p, i.e.

$$\ell_p(\mathcal{L}(X), \mathcal{L}(Y)) := \inf_{(X,Y)} L_p(X,Y) = \inf_{(X,Y)} \mathbb{E}\left[|X-Y|^p\right]^{1/p} . \tag{2.9}$$

This metric is known under many names. Rösler 1991[11] refers to it, for $p = 2$, as Wasserstein metric, Johnson and Samworth 2005[4] use the term Mallows distance, whereas Rachev 1991[10] also names it Kantorovich metric for the case of $p = 1$.

The minimal ℓ_p metric is ideal of order 1 for arbitrary p.

SECTION 2.3.2 Zolotarev Metric

Zolotarev 1976[14, 15] introduced a metric that is useful for the study of sums of independent random variables. For every $s = m + \alpha$ with $m \in \mathbb{N}_0$, $\alpha \in (0,1]$ and real valued random variables X, Y it is defined as

$$\zeta_s(X,Y) := \sup_{f\in C^s} \left|\mathbb{E}\left[f(X) - f(Y)\right]\right| ,$$

where $C^s = \left\{ f \in C^m(\mathbb{R}) \mid \ |f^{(m)}(x) - f^{(m)}(y)| \leqslant |x-y|^\alpha \ \forall x, y \right\}$.

ζ_s is ideal of order s and simple. Moreover, convergence in the Zolotarev metric implies weak convergence. A criterion for the finiteness of the Zolotarev metric is necessary to use it in the context of the contraction method. The following condition is easy to verify.

$$\left.\begin{array}{ll} (a) & \mathbb{E}\left[X^k - Y^k\right] = 0 \ \forall\, k = 1, \dots, m \\ (b) & \mathbb{E}\left[|X|^s + |Y|^s\right] < \infty \end{array}\right\} \Rightarrow \zeta_s(X,Y) < \infty . \tag{2.10}$$

SECTION 3

Stable Limit Laws via the Zolotarev Metric

The following theorem ensures that, as pointed out in Section 1, the contraction method together with the Zolotarev metric is applicable for a proof of stable limit laws for scaled sums of random variables in the domain of strong normal attraction.

Theorem 3.1. *$\alpha \in (1,2), c_1, c_2 \in \mathbb{R}_{\geqslant 0}, c_1 + c_2 > 0, \gamma > 0$, X a random variable with expectation μ satisfying*

$$\mathbb{P}(X < x) = \frac{c_1}{|x|^\alpha} + \mathcal{O}_x\left(\frac{1}{|x|^{\alpha+\gamma}}\right) \text{ as } x \to -\infty$$

and

$$\mathbb{P}(X > x) = \frac{c_2}{x^\alpha} + \mathcal{O}_x\left(\frac{1}{x^{\alpha+\gamma}}\right) \text{ as } x \to \infty\ . \tag{3.1}$$

Then, for the stable distribution $\mathcal{L}(W) = St(\alpha, \mu, c_1, c_2)$, there exists $\lambda > \alpha$ such that

$$\zeta_\lambda(X, W) < \infty\ .$$

The proof is given below. As a direct corollary of Theorem 3.1 we obtain the classical stable limit laws.

Corollary 3.2. *X and W as in Theorem 3.1. Then, for an iid sequence (X_n) with $\mathcal{L}(X_i) = \mathcal{L}(X)$,*

$$\mathcal{L}\left(n^{-1/\alpha}\left(\sum_{i=1}^n X_i - n\mu\right)\right) \to \mathcal{L}(W - \mu) \text{ as } n \to \infty\ . \tag{3.2}$$

Proof. We assume w.l.o.g. $\mu = 0$. Let $Y_n := n^{-1/\alpha}(\sum_{i=1}^n X_i)$. Y_{2^n} satisfies a recursive equation,

$$\begin{aligned} Y_{2n} &= (2n)^{-1/\alpha}\left(\sum_{i=1}^{2n} X_i\right) \\ &= (2n)^{-1/\alpha}\left(\sum_{i=1}^{n} X_i\right) + (2n)^{-1/\alpha}\left(\sum_{i=n+1}^{2n} X_i\right) \end{aligned}$$

$$\overset{D}{=} 2^{-1/\alpha}\, Y_n + 2^{-1/\alpha}\, Y_n' \ . \tag{3.3}$$

Here, $Y_n' := n^{-1/\alpha}(\sum_{i=1}^{n} X_i')$ with $(X_i') \overset{D}{=} (X_i)$ and (X_i') independent of (X_i). An equation similar to (3.3) holds true for W. From the convolution property (2.1) of stable distributions it is known that

$$W \overset{D}{=} 2^{-1/\alpha}\, W + 2^{-1/\alpha}\, W' \ , \tag{3.4}$$

where $\mathcal{L}(W) = \mathcal{L}(W')$ and W' is independent of W. Theorem 3.1 ensures that there exists $\lambda > \alpha$ with

$$\zeta_\lambda (X, W) < \infty \ . \tag{3.5}$$

Using both (3.3) and (3.4) we find, similar to (1.3),

$$\begin{aligned}
\zeta_\lambda (Y_{2n}, W) &= \zeta_\lambda \left(2^{-1/\alpha}\, Y_n + 2^{-1/\alpha}\, Y_n', 2^{-1/\alpha}\, W + 2^{-1/\alpha}\, W'\right) \\
&= 2^{-\lambda/\alpha}\, \zeta_\lambda \left(Y_n + Y_n', W + W'\right) \\
&\leqslant 2^{-\lambda/\alpha} \, \left(\zeta_\lambda \left(Y_n + Y_n', Y_n + W'\right) + \zeta_\lambda \left(Y_n + W', W + W'\right)\right) \\
&\leqslant 2^{-\lambda/\alpha} \, \left(\zeta_\lambda \left(Y_n, W\right) + \zeta_\lambda \left(Y_n', W'\right)\right) \\
&\leqslant 2^{1-\lambda/\alpha}\, \zeta_\lambda \left(Y_n, W\right) \ .
\end{aligned} \tag{3.6}$$

From (3.6) together with (3.5) it follows that, on account of $2^{1-\lambda/\alpha} < 1$,

$$\zeta_\lambda (Y_{2^n}, W) \leqslant \left(2^{1-\lambda/\alpha}\right)^n \zeta_\lambda (X, W) \to 0 \ \text{ as } \ n \to \infty \ . \tag{3.7}$$

Convergence in ζ_λ implies weak convergence, and thus, (3.2) follows directly from (3.7).

∎

The proof of Theorem 3.1 uses a result of Johnson and Samworth 2005[4], which states that there exists $\beta > \alpha$ such that $\ell_\beta(\mathcal{L}(X), \mathcal{L}(W)) < \infty$. Then, $\zeta_s(X, W)$ is bounded in dependency of ℓ_β and absolute $(\alpha - \epsilon)$-th moments. The following preparatory lemma is a variation of Lemma 5.7 in Drmota et al. 2008[2].

Lemma 3.3. *Let $s \in (1, 2)$, $\epsilon > 0$ and $\epsilon' = \frac{\epsilon}{s+\epsilon-1}$. Then, for all real-valued random variables X and Y with equal first moments,*

$$\zeta_s (X, Y) \leqslant \left(\mathbb{E}\left[|X|^{s-\epsilon'}\right]^{\frac{s-1}{s-\epsilon'}} + \mathbb{E}\left[|Y|^{s-\epsilon'}\right]^{\frac{s-1}{s-\epsilon'}}\right) \ \ell_{s+\epsilon} \left(\mathcal{L}(X), \mathcal{L}(Y)\right) \ .$$

Proof. We set $Z := X - Y$. For every $f \in C^s$ with $f'(0) = 0$ the mean value theorem implies, for appropriate $0 \leqslant \xi \leqslant 1$,

$$\begin{aligned}
|f(X) - f(Y)| &= |f(Y + Z) - f(Y)| \\
&= \left|f(Y) + f'(Y + \xi Z)\, Z - f(Y)\right|
\end{aligned}$$

$$
\begin{aligned}
&= \left|\left(f'\left(Y+\xi Z\right)-f'\left(0\right)\right) Z\right| \\
&\leqslant \left|Y+\xi Z\right|^{s-1}\left|Z\right| \\
&= \left|Y\left(1-\xi\right)+\xi X\right|^{s-1}\left|Z\right| \\
&\leqslant \left(\left|Y\right|^{s-1}+\left|X\right|^{s-1}\right)\left|Z\right| \ .
\end{aligned}
\tag{3.8}
$$

Since X and Y do have equal first moments we find

$$
\begin{aligned}
\sup_{f\in C^s}\left|\mathbb{E}\left[f\left(X\right)-f\left(Y\right)\right]\right| &= \sup_{f\in C^s}\left|\mathbb{E}\left[f\left(X\right)-f'\left(0\right)X-\left(f\left(Y\right)-f'\left(0\right)Y\right)\right]\right| \\
&= \sup_{f\in C^s, f'(0)=0}\left|\mathbb{E}\left[f\left(X\right)-f\left(Y\right)\right]\right| \ .
\end{aligned}
\tag{3.9}
$$

Substituting (3.8) in (3.9) yields

$$
\sup_{f\in C^s}\left|\mathbb{E}\left[f\left(X\right)-f\left(Y\right)\right]\right| \leqslant \mathbb{E}\left[\left(\left|Y\right|^{s-1}+\left|X\right|^{s-1}\right)\left|Z\right|\right] \ . \tag{3.10}
$$

Next we apply Hölder's inequality to (3.10). For $p, q > 1$, $\frac{1}{p}+\frac{1}{q}=1$ it states that, for real-valued random variables U and V,

$$
\mathbb{E}\left[\left|UV\right|\right] \leqslant \mathbb{E}\left[\left|U\right|^p\right]^{1/p}\mathbb{E}\left[\left|V\right|^q\right]^{1/q} \ . \tag{3.11}
$$

By definition we have $\epsilon' = \frac{\epsilon}{s+\epsilon-1}$, thus

$$
s\epsilon' + \epsilon\epsilon' = \epsilon + \epsilon' \ ,
$$

and therefore,

$$
\frac{s-1}{s-\epsilon'} + \frac{1}{s+\epsilon} = \frac{s^2-s+s\epsilon-\epsilon+s-\epsilon'}{s^2+s\epsilon-s\epsilon'-\epsilon\epsilon'} = \frac{s^2+s\epsilon-\quad(\epsilon+\epsilon')}{s^2+s\epsilon-(s\epsilon'+\epsilon\epsilon')} = 1 \ .
$$

Hence, we can use (3.11) with $p=\frac{s-\epsilon'}{s-1}$ and $q=s+\epsilon$ to bound (3.10) by

$$
\zeta_s\left(X,Y\right) \leqslant \mathbb{E}\left[\left(\left|Y\right|^{s-1}+\left|X\right|^{s-1}\right)^{\frac{s-\epsilon'}{s-1}}\right]^{\frac{s-1}{s-\epsilon'}} \mathbb{E}\left[\left|Z\right|^{s+\epsilon}\right]^{\frac{1}{s+\epsilon}} \ . \tag{3.12}
$$

Bounding the first term on the right hand side in (3.12) with Minkowski's inequality yields

$$
\zeta_s\left(X,Y\right) \leqslant \left(\mathbb{E}\left[\left|Y\right|^{s-\epsilon'}\right]^{\frac{s-1}{s-\epsilon'}} + \mathbb{E}\left[\left|X\right|^{s-\epsilon'}\right]^{\frac{s-1}{s-\epsilon'}}\right) \mathbb{E}\left[\left|X-Y\right|^{s+\epsilon}\right]^{\frac{1}{s+\epsilon}} \ . \tag{3.13}
$$

Both X and Y are arbitrary chosen. Therefore, we can build the infimum over all random variables with distribution $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ respectively on (3.13) and conclude, using that ζ_s is simple,

$$
\zeta_s\left(X,Y\right) \leqslant \left(\mathbb{E}\left[\left|Y\right|^{s-\epsilon'}\right]^{\frac{s-1}{s-\epsilon'}} + \mathbb{E}\left[\left|X\right|^{s-\epsilon'}\right]^{\frac{s-1}{s-\epsilon'}}\right) \ell_{s+\epsilon}\left(\mathcal{L}(X),\mathcal{L}(Y)\right) \ .
$$

■

Proof of theorem 3.1. From Lemma 5.1 in Johnson and Samworth 2005[4] we know that there exists $\beta > \alpha$ such that

$$\ell_\beta\left(\mathcal{L}(X), \mathcal{L}(W)\right) < \infty \,. \tag{3.14}$$

In (3.14), $\beta < 2$ can be chosen w.l.o.g.. We set

$$\delta := \frac{1}{2}\left(\beta - \alpha\right) \;, \quad \lambda := \beta - \delta \quad \text{and} \quad \delta' := \frac{\delta}{\lambda + \delta - 1} \,.$$

Note that $\lambda > \alpha$. Using Lemma 3.3 we bound $\zeta_\lambda(X, W)$ by

$$\zeta_\lambda\left(X, W\right) \leqslant \left(\mathbb{E}\left[|X|^{\lambda-\delta'}\right]^{\frac{\lambda-1}{\lambda-\delta'}} + \mathbb{E}\left[|W|^{\lambda-\delta'}\right]^{\frac{\lambda-1}{\lambda-\delta'}}\right) \ell_{\lambda+\delta}\left(\mathcal{L}(X), \mathcal{L}(W)\right) \,. \tag{3.15}$$

From (3.14) it holds that $\ell_{\lambda+\delta}(\mathcal{L}(X), \mathcal{L}(W))$ is finite, so it remains to proof the finiteness of the first term on the right hand side of (3.15). β, δ and δ' are chosen such that $\delta' = \delta/(\beta - 1) > \delta$, thus,

$$\lambda - \delta' = \beta - \delta - \delta' < \beta - 2\delta = \beta - 2\,\frac{1}{2}\left(\beta - \alpha\right) = \alpha \,. \tag{3.16}$$

Both X and W belong to DoA_α by definition. Therefore, (2.6) under use of (3.16) holds that

$$\mathbb{E}\left[|X|^{\lambda-\delta'}\right] < \infty \quad \text{and} \quad \mathbb{E}\left[|W|^{\lambda-\delta'}\right] < \infty \,. \tag{3.17}$$

Substituting (3.14) and (3.17) into (3.15) yields

$$\zeta_\lambda\left(X, W\right) < \infty \,.$$

∎

SECTION 4

Stable Limit Laws Including "Growing Into" Sequences

Until now, the ideality of the Zolotarev metric was one key argument in using contraction method for stable limit laws[1]. For a stable limit law of order α, ζ_s with $s > \alpha$ was used, since it is ideal of order s. The minimal ℓ_p metric is ideal of order 1 only, nevertheless, it can be used to proof stable limit laws, as shown by Johnson and Samworth 2005[4] (Theorem 1.2).

We now use their ideas to proof a stable limit law for the scaled sum of a "growing into" sequence. A sequence of random variables Z_n is said to grow into a limit W, if $Z_n < \infty$ for all n and $d(Z_n, W) \to 0$ for a metric d. These sequences are of higher interest in the analysis algorithms as the distribution of measured complexity is surely not an element of the domain of strong normal attraction, as needed by Corollary 3.2. The measured complexity is finite for all n, while it might converge to a stable distribution in the limit.

Bahr and Esseen 1965[13] proved some inequalities for the absolute r-th moment of a sum of independent random variables. There, Theorem 2 implies that, for $1 \leqslant r \leqslant 2$ and a sequence of independent, not needfully identically distributed random variables X_i with zero mean, the absolute rth moment of the nth partial sum can be bound by

$$\mathbb{E}\left[|X_1 + \cdots + X_n|^r\right] \leqslant 2 \sum_{i=1}^{n} \mathbb{E}\left[|X_i|^r\right] \ , \tag{4.1}$$

Using this inequality give a stable limit law in two different versions; one formulated for scaled sums and the other for recursively defined sequences.

SECTION 4.1

Proof for Scaled Sums

Theorem 4.1. *Let $\alpha \in (1,2)$, $1 \leqslant \beta < \alpha$, W α-stable with zero mean, Z_n a sequence of independent random variables with zero mean such that*

$$\sum_{i=1}^{n} \ell_\beta\left(\mathcal{L}(Z_i), \mathcal{L}(W)\right) = o_n\left(n^{1/\alpha}\right) \ . \tag{4.2}$$

[1] See the estimates (1.3) and (3.6).

Then, the scaled sum

$$Y_n := n^{-1/\alpha} \sum_{i=1}^{n} Z_i \tag{4.3}$$

converges to W *in distribution, i.e.*

$$Y_n \xrightarrow{D} W \ \textit{ as } \ n \to \infty \; .$$

Proof. We chose independent pairs of random variables (Z_i^*, W_i^*) such that, for all $i = 1, \ldots, n$,

$$\mathcal{L}(Z_i^*) = \mathcal{L}(Z_i) \; , \quad \mathcal{L}(W_i^*) = \mathcal{L}(W)$$

and

$$\mathbb{E}\left[|Z_i^* - W_i^*|^\beta\right]^{1/\beta} \leqslant \ell_\beta\left(\mathcal{L}(Z_i), \mathcal{L}(W)\right) + i^{-2} \; . \tag{4.4}$$

Using both (4.3) and the convolution property of stable distributions (2.1) we find

$$\begin{aligned} \ell_\beta\left(\mathcal{L}(Y_n), \mathcal{L}(W)\right) &\leqslant \mathbb{E}\left[\left|n^{-1/\alpha}\sum_{i=1}^{n} Z_i^* - n^{-1/\alpha}\sum_{i=1}^{n} W_i^*\right|^\beta\right]^{1/\beta} \\ &= n^{-1/\alpha}\,\mathbb{E}\left[\left|\sum_{i=1}^{n} Z_i^* - W_i^*\right|^\beta\right]^{1/\beta} \; . \end{aligned} \tag{4.5}$$

The particular summands $(Z_i^* - W_i^*)$ are independent and have a zero mean, thus, we can apply (4.1) on (4.5) and obtain

$$\ell_\beta\left(\mathcal{L}(Y_n), \mathcal{L}(W)\right) \leqslant 2n^{-1/\alpha}\,\mathbb{E}\left[\sum_{i=1}^{n} |Z_i^* - W_i^*|^\beta\right]^{1/\beta} \; . \tag{4.6}$$

Using the inequality $(a+b)^y \leqslant a^y + b^y$ for $a, b \geqslant 0$ and $y \leqslant 1$ we bound (4.6) by

$$\ell_\beta\left(\mathcal{L}(Y_n), \mathcal{L}(W)\right) \leqslant 2n^{-1/\alpha} \sum_{i=1}^{n} \mathbb{E}\left[|Z_i^* - W_i^*|^\beta\right]^{1/\beta} \; . \tag{4.7}$$

Applying the bounds assumed in (4.4) to (4.7) yields

$$\begin{aligned} \ell_\beta\left(\mathcal{L}(Y_n), \mathcal{L}(W)\right) &\leqslant 2n^{-1/\alpha} \sum_{i=1}^{n} \left(\ell_\beta\left(\mathcal{L}(Z_i), \mathcal{L}(W)\right) + i^{-2}\right) \\ &= 2n^{-1/\alpha} \sum_{i=1}^{n} \ell_\beta\left(\mathcal{L}(Z_i), \mathcal{L}(W)\right) + 2n^{-1/\alpha} \sum_{i=1}^{n} i^{-2} \; . \end{aligned} \tag{4.8}$$

Using condition (4.2) we find for the first summand on the right handside of (4.8) the following,

$$2n^{-1/\alpha} \sum_{i=1}^{n} \ell_\beta \left(\mathcal{L}(Z_i), \mathcal{L}(W)\right) = 2n^{-1/\alpha}\, o_n\left(n^{1/\alpha}\right) \to 0 \text{ as } n \to \infty \,. \tag{4.9}$$

The second summand on the right-handside of (4.8) is established using the well known fact that $\sum_{i=1}^{\infty} i^{-2} < \infty$,

$$2n^{-1/\alpha} \sum_{i=1}^{n} i^{-2} \leqslant 2n^{-1/\alpha} \sum_{i=1}^{\infty} i^{-2} \to 0 \text{ as } n \to \infty \,. \tag{4.10}$$

Substituting (4.9) and (4.10) in (4.8) we conclude

$$\ell_\beta \left(\mathcal{L}(Y_n), \mathcal{L}(W)\right) \to 0 \text{ as } n \to \infty \,.$$

■

SECTION 4.2

Formulation as a Degenerate Recursion

In most cases, the recursive equation itself contains certain information about the limit distribution. As an example consider the recursive equation (1.2) in Section 1,

$$X_1 = X, \quad X_{n+1} \stackrel{D}{=} 2^{-1/2}\, X_n + 2^{-1/2}\, X_n' \,.$$

It is obvious that, if a limit distribution μ exists, it must fulfill

$$\mu = \mathcal{L}\left(2^{-1/2}\, Y + 2^{-1/2}\, Y'\right) \,,$$

where Y and Y' are independent with $\mathcal{L}(Y) = \mathcal{L}(Y') = \mu$. In this paradigm the limit distribution μ is asured to be a normal distribution.

There exists recursive equations which, in the limit, become

$$X \stackrel{D}{=} X \,,$$

and thus give no information about the limit distribution. Recursions of that type are called **degenerate**.

Neininger and Rüschendorf 2004[8] prove a normal limit law for degenerate recursive equations under certain constraints using the Zolotarev metric. A natural question arises if this result can be generalized to stable limit laws of arbitrary orders[2].

[2]Such a stable limit law of order 1 is shown for a particular setting with analytical methods in Drmota et al. 2009[1].

In order to give an idea of how such a stable limit law may look like we modify the conditions in Theorem 4.1. Let (Z_n) and W as in Theorem 4.1. Instead of defining Y_n directly we can determine it by a degenerate recursion,

$$Y_1 := Z_1 \ , \quad Y_{n+1} := \left(\frac{n}{n+1}\right)^{1/\alpha} Y_n + \left(\frac{1}{n+1}\right)^{1/\alpha} Z_{n+1} \ . \tag{4.11}$$

By a simple induction we see that

$$Y_n = n^{-1/\alpha} \sum_{i=1}^{n} Z_i \ ,$$

and hence, as Theorem 4.1 ensures, Y_n converges to W in distribution. The case $n = 1$ is fulfilled by definition, for $n + 1$ it holds

$$\begin{aligned} Y_{n+1} &= \left(\frac{n}{n+1}\right)^{1/\alpha} Y_n + \left(\frac{1}{n+1}\right)^{1/\alpha} Z_{n+1} \\ &= \left(\frac{n}{n+1}\right)^{1/\alpha} \left(\left(\frac{1}{n}\right)^{1/\alpha} \sum_{i=1}^{n} Z_i\right) + \left(\frac{1}{n+1}\right)^{1/\alpha} Z_{n+1} \\ &= \left(\frac{1}{n+1}\right)^{1/\alpha} \sum_{i=1}^{n} Z_i + \left(\frac{1}{n+1}\right)^{1/\alpha} Z_{n+1} \\ &= \left(\frac{1}{n+1}\right)^{1/\alpha} \sum_{i=1}^{n+1} Z_i \ . \end{aligned}$$

SECTION 5

Conclusion

We presented a proof for the classical stable limit laws under use of contraction method in combination with the Zolotarev metric. Furthermore, a stable limit law was proved for scaled sums of growing into sequences. This limit law was alternatively formulated for sequences of random variables defined by a simple degenerate recursion.

APPENDIX A

Stable Distributions and Characteristic Functions

The content of this section is mainly taken from Feller 1966[3], all references in this section correspond to the same book[1]. Let μ be a probability distribution on $\mathfrak{B}(\mathbb{R})$, its **characteristic function** $\phi_\mu : \mathbb{R} \to \mathbb{C}$ is defined as

$$\phi_\mu(t) = \int_{-\infty}^{\infty} e^{-itx}\, d\mu(x) \ .$$

The characteristic function is finite by definition, and thus can be used for the description of arbitrary distributions. Distinct distributions do have distinct characteristic functions, and every characteristic function defines uniquely a distribution. One further important property of the characteristic function is that it holds the following formula for the sum of independent random variables X and Y, i.e.

$$\phi_{\mathcal{L}(X+Y)} = \phi_{\mathcal{L}(X)}\phi_{\mathcal{L}(Y)} \ . \tag{A.1}$$

A distribution μ is called **infinitely divisible** if, for any $n \in \mathbb{N}$, there exists a distribution μ_n such that

$$\mu = \mathcal{L}\left(X_{n,1} + X_{n,2} + \cdots + X_{n,n}\right) \ ,$$

where the $X_{n,i}$ are independent with $\mathcal{L}(X_{n,i}) = \mu_n$. This property can as well be written under the use of characteristic functions, on account of (A.1). μ is infinitely divisible if, for any $n \in \mathbb{N}$, there exists a characteristic function ϕ_n such that

$$\phi_\mu = \phi_n^n \ .$$

Note that stable distribution are infinitely divisible. In terms of characteristic functions, the convolution property (2.1) can be written as, for W α-stable distributed,

$$\phi_{\mathcal{L}(n^{1/\alpha}W)} = \phi^n_{\mathcal{L}(W)} \ . \tag{A.2}$$

The characteristic function of an infinitely divisible distribution μ matches the following unique form (see XVII.2 (2.9)),

$$\phi_\mu(t) = \exp\left\{ \int_{-\infty}^{\infty} \frac{e^{itx} - 1 - it\sin x}{x^2}\, dM(x) \right\} \ , \tag{A.3}$$

[1]For a more modern view we refer to Klenke 2008[5], Chapter 16.

where M, the so called **canonical measure**, attributes finite masses to finite intervals and satisfies

$$M^+(x) = \int_x^\infty y^{-2}\, dM(y) < \infty \quad \text{and} \quad M^-(x) = \int_{-\infty}^x y^{-2}\, dM(y) < \infty \,. \qquad \text{(A.4)}$$

For stable distributions, according to (A.2), the canonical measure M_{stable} is given by (compare to XVII.3 (3.17))

$$M_{\text{stable}}\{\overline{-x,0}\} = c_1\, x^{2-\alpha} \quad \text{and} \quad M_{\text{stable}}\overline{\{0,x\}} = c_2\, x^{2-\alpha}\,, \qquad \text{(A.5)}$$

where $c_1, c_2 \geqslant 0, c_1 + c_2 > 0$. Therefore, the quantities M^+_{stable} and M^-_{stable} as defined in (A.4) can be calculated as follows.

$$M^-_{\text{stable}}(x) = \int_{-\infty}^x y^{-2}\, dM_{\text{stable}}(y) = c_1\, x^{-a}\,,$$

and similarly

$$M^+_{\text{stable}}(x) = \int_x^\infty y^{-2}\, dM_{\text{stable}}(y) = c_2\, x^{-a}\,.$$

The tail behavior of stable distributions asymptotically equals M^-_{stable} and M^+_{stable} (see XVII.4 (d)), therefore, the constants c_1 and and c_2 are the same as in (2.2) and (2.3), respectively. The general form of the characteristic function of a stable distribution $\mathfrak{W}$ is obtained by substituting (A.5) in (A.3) and adding a centering constant b. This results in

$$\phi_{\mathfrak{W}}(t) = \begin{cases} \exp\{ibt + |t|^\alpha \frac{\Gamma(3-\alpha)}{\alpha(\alpha-1)}((c_2+c_1)\cos\frac{\pi\alpha}{2} + \text{sign}(t)\, i\,(c_2-c_1)\sin\frac{\pi\alpha}{2})\} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , \alpha \neq 1 \\ \exp\{ibt - |t|\,(c_2+c_1)\,(\frac{\pi}{2} + \text{sign}(t)\, i\,(c_2-c_1)\log|t|)\} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , \alpha = 1 \end{cases}. \qquad \text{(A.6)}$$

Hence, every stable distributions is uniquely determined by a set of four parameters (α, b, c_1, c_2). The above form (A.6) is often referred as

$$\phi_{\mathfrak{W}}(t) = \begin{cases} \exp\{ibt + |t|^\alpha\, C\frac{\Gamma(3-\alpha)}{\alpha(\alpha-1)}(\cos\frac{\pi\alpha}{2} + \text{sign}(t)\, i\gamma\sin\frac{\pi\alpha}{2})\} & , \alpha \neq 1 \\ \exp\{ibt - |t|\, C(\frac{\pi}{2} + \text{sign}(t)\, i\gamma\log|t|)\} & , \alpha = 1 \end{cases},$$

with $C = c_2 + c_2$ and $\gamma = (c_2 - c_1)/(c_2 + c_1)$.

Bibliography

[1] DRMOTA, M., IKSANOV, A., MOEHLE, M., AND ROESLER, U. A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures Algorithms 34*, 3 (2009), 319–336.

[2] DRMOTA, M., JANSON, S., AND NEININGER, R. A functional limit theorem for the profile of search trees. *Ann. Appl. Probab. 18*, 1 (2008), 288–333.

[3] FELLER, W. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons Inc., New York, 1966.

[4] JOHNSON, O., AND SAMWORTH, R. Acknowledgement of priority: "Central limit theorem and convergence to stable laws in Mallows distance" [Bernoulli **11** (2005), no. 5, 829–845; mr2172843]. *Bernoulli 12*, 1 (2006), 191.

[5] KLENKE, A. *Wahrscheinlichkeitstheorie*, vol. 2. Springer-Verlag, Heidelberg, 2008.

[6] MIJNHEER, J. On the rate of convergence to a stable limit law. II. *Litovsk. Mat. Sb. 26*, 3 (1986), 482–487.

[7] NEININGER, R., AND RÜSCHENDORF, L. A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab. 14*, 1 (2004), 378–418.

[8] NEININGER, R., AND RÜSCHENDORF, L. On the contraction method with degenerate limit equation. *Ann. Probab. 32*, 3B (2004), 2838–2856.

[9] PETROV, V. V. *Sums of independent random variables*. Springer-Verlag, New York, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

[10] RACHEV, S. T. *Probability metrics and the stability of stochastic models*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1991.

[11] RÖSLER, U. A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl. 25*, 1 (1991), 85–100.

[12] RÖSLER, U., AND RÜSCHENDORF, L. The contraction method for recursive algorithms. *Algorithmica 29*, 1-2 (2001), 3–33. Average-case analysis of algorithms (Princeton, NJ, 1998).

[13] VON BAHR, B., AND ESSEEN, C.-G. Inequalities for the rth absolute moment of a sum of random variables, $1 \leqslant r \leqslant 2$. *Ann. Math. Statist 36* (1965), 299–303.

[14] ZOLOTAREV, V. M. Ideal metrics in the problem of approximating the distributions of sums of independent random variables. *Teor. Verojatnost. i Primenen. 22*, 3 (1977), 449–465.

[15] ZOLOTAREV, V. M. Letter to the editors: "Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces" (Teor. Verojatnost. i Primenen. 21 (1976), no. 4, 741–758). *Teor. Verojatnost. i Primenen. 22*, 4 (1977), 901.