Strong approximation and a central limit theorem for St. Petersburg sums

I. Berkes *

Abstract

The St. Petersburg paradox (Bernoulli 1738) concerns the fair entry fee in a game where the winnings are distributed as $P(X = 2^k) = 2^{-k}$, $k = 1, 2, \ldots$. The tails of X are not regularly varying and the sequence S_n of accumulated gains has, suitably centered and normalized, a class of semistable laws as subsequential limit distributions (Martin-Löf (1985), Csörgő and Dodunekova (1991)). This has led to a clarification of the paradox and an interesting and unusual asymptotic theory in past decades. In this paper we prove that S_n can be approximated by a semistable Lévy process $\{L(n), n \geq 1\}$ with a.s. error $O(\sqrt{n}(\log n)^{1+\varepsilon})$ and, surprisingly, the error term is asymptotically normal, exhibiting an unexpected central limit theorem in St. Petersburg theory.

MSC 2010. 60E07, 60F05, 60F17.

Keywords. St. Petersburg sums, semistable process, strong approximation, central limit theorem.

1 Introduction

Let $X, X_1, X_2, \ldots$ be i.i.d. r.v.'s with

$$P(X = 2^k) = 2^{-k}, \quad (k = 1, 2, \ldots) \tag{1.1}$$

and let $S_n = \sum_{k=1}^n X_k$. The asymptotic behavior of the sequence $\{S_n, n \geq 1\}$ has attracted considerable attraction in the literature in connection with the St. Petersburg paradox concerning the 'fair' entry fee in a game where the winnings are

*A. Rényi Institute of Mathematics, Reáltanoda u. 13-15, 1053 Budapest, Hungary, e-mail: berkes.istvan@renyi.mta.hu. Research supported by NFKIH grant K 125569.

distributed as X. We refer to Csörgő and Simons [10] for the history and bibliography of the problem. Feller [11] proved that

$$\lim_{n\to\infty} \frac{S_n}{n\log_2 n} = 1 \qquad \text{in probability}$$

(where $\log_2$ denotes logarithm with base 2) and Martin-Löf [16] showed

$$S_{2^k}/2^k - k \xrightarrow{d} G$$

where G is the infinitely divisible distribution function with characteristic function $\exp(g(t))$, where

$$g(t) = \sum_{l=-\infty}^{0} (e^{it2^l} - 1 - it2^l)2^{-l} + \sum_{l=1}^{\infty}(e^{it2^l} - 1)2^{-l}. \tag{1.2}$$

Let G_γ denote the distribution with characteristic function $\exp(\gamma g(t/\gamma) - it\log_2\gamma)$ and let $\gamma_n = n/2^{[\log_2 n]+1} \in [1/2, 1)$ be the parameter describing the location of n between two consecutive powers of 2, where $[y]$ denotes the (lower) integer part of $y \in \mathbb{R}$. Csörgő [6] proved that

$$\sup_x \left| P\left(\frac{S_n}{n} - \log_2 n \le x\right) - G_{\gamma_n}(x)\right| \longrightarrow 0 \quad \text{as } n \to \infty \tag{1.3}$$

and determined the precise convergence rate. It follows (and actually it was proved earlier in [8]) that the class of subsequential limit distributions of $S_n/n - \log_2 n$ is the class

$$\mathcal{G} = \{G_\gamma : 1/2 \le \gamma < 1\}.$$

If n runs through the interval $[2^k, 2^{k+1}]$, then G_{γ_n} moves through the distributions $G_{j/2^{k+1}}, 2^k \le j \le 2^{k+1}$ representing, in view of $G_{1/2} = G_1$, a "circular" path in $\mathcal{G}$. In view of (1.3), the distribution of $S_n/n - \log_2 n$ also describes approximately a circular path, a remarkable asymptotic behavior called in [6] *merging.*

From the merging theorem (1.3) and the results of [8] it also follows that given $\gamma \in (1/2, 1)$ and an increasing sequence (n_k) of integers, the limit relation

$$S_{n_k}/n_k - \log_2 n_k \xrightarrow{d} G_\gamma \tag{1.4}$$

holds iff $\gamma_{n_k} \to \gamma$ as $k \to \infty$. For $\gamma = 1/2$ this criterion breaks down and (1.4) holds iff the sequence γ_{n_k} has no other cluster points than 1/2 and 1.

Using a decomposition idea of Le Page, Woodroofe and Zinn [15], in [3] a new representation of the limiting semistable variable of Petersburg sums was given, simplifying the theory considerably and leading to new asymptotic information. Let $\Psi(x)$ denote the function on $(0, \infty)$ which grows linearly from 1 to 2 on any interval $[2^k, 2^{k+1})$, $(k \in \mathbb{Z})$, let $\eta_1, \eta_2, \ldots$ be independent exponential random variables with mean 1 and let $Z_k = \sum_{j=1}^k \eta_k$. In [3], Lemma 2 it was proved that for any $1 \le \gamma < 2$ the series

$$Y^{(\gamma)} = \sum_{j=1}^{\infty} \left[\frac{1}{Z_j} \Psi\left(\frac{Z_j}{\gamma}\right) - \frac{1}{j} \Psi\left(\frac{j}{\gamma}\right) \right] \tag{1.5}$$

converges absolutely with probability 1 and the limit distribution G_γ above is identical with the distribution of $Y^{(\gamma)} + c_\gamma$, where

$$c_\gamma = \sum_{k=1}^{\infty} \frac{\{2^k \gamma\}}{2^k \gamma} - \log_2 \gamma.$$

We note that for each $\gamma \in [1/2, 1)$ we have

$$c_\gamma = \xi(\gamma) := 2 - \sum_{k=1}^{\infty} \frac{k\varepsilon_k}{2^k \gamma} - \log_2\, k$$

where the ε_k's are the dyadic digits of γ given by $\gamma = \sum_{k=1}^{\infty} \varepsilon_k 2^{-k}$ and the function ξ was introduced by Csörgő and Simons [9], see also Kern and Wiedrich [14]. In contrast to the representation of the limiting semistable variable of St. Petersburg theory as an infinite weighted sum of independent Poisson variables in [8], the terms of the sum (1.5) are dependent random variables. For an analogous representation of stable random variables, see [15]. A similar representation, implicit in [3], holds for the partial sums S_n, namely

$$\frac{1}{n} S_n - a_{n,\gamma_n} \stackrel{d}{=} (1+\varepsilon_n) \sum_{j=1}^{n} \left[\frac{1}{Z_j} \Psi\left(\frac{Z_j}{(1+\varepsilon_n)\gamma_n}\right) - \frac{1}{j} \Psi\left(\frac{j}{\gamma_n}\right) \right] + \varepsilon_n a_{n,\gamma_n} \tag{1.6}$$

where $\varepsilon_n = Z_{n+1}/n - 1$, $\gamma_n = n/2^{[\log_2 n]+1}$ is the dyadic location parameter introduced above and

$$a_{n,\gamma} = \sum_{j=1}^{n} \frac{\Psi(j/\gamma)}{j}. \tag{1.7}$$

For a simple proof, see Section 2. Note that the equality in (1.6) holds only in distribution and thus (1.6) yields an expansion of S_n in the sense of Strassen: S_n can be redefined on a suitable probability space together with a sequence (ε_n) of i.i.d. exponential random variables such that setting $Z_n = \sum_{k=1}^n \varepsilon_k$, (1.6) holds pointwise. This makes the formula easy to apply, in particular, (1.6) makes the asymptotic theory of St. Petersburg sums very transparent. (Note the difference between (1.6) and the Edgeworth expansion of S_n in [7], [17] giving an expansion of the distribution function of S_n. The expansion (1.6) is particularly convenient for problems involving almost everywhere convergence and asymptotics.) By the law of the iterated logarithm we have $\varepsilon_n = O(n^{-1/2}(\log\log n)^{1/2})$ a.s. and an easy calculation shows that replacing ε_n by 0 in (1.6) results in an error of $o_P(1)$ on the right hand side, and thus we get the result

$$\frac{1}{n} S_n - a_{n,\gamma_n} = Y^{(\gamma_n)} + o_P(1), \tag{1.8}$$

which is meant again in the sense that for each fixed n the variables S_n and $Y^{(\gamma_n)}$ can be defined on a common probability space such that (1.8) holds. Relation (1.8) thus yields a pointwise version of the merging result (1.3). The purpose of the present paper is to prove that actually much more is valid: the partial sum process of (X_n) can be approximated by a semistable Lévy process $\{L(t),\, t \geq 0\}$ with $L(1) \stackrel{d}{=} Y^{(1)}$ with a.s. error $O(\sqrt{n}(\log n)^{1+\varepsilon})$ and an asymptotically normal error term, establishing an unexpected central limit theorem in St. Petersburg theory.

Theorem 1.1 *Let $\{L(t),\ t \geq 0\}$ be the Lévy process defined by*

$$E(\exp(iuL(t)) = \exp(tg(u)). \tag{1.9}$$

where g is the function in (1.2). Then on a suitable probability space one can define the St. Petersburg sequence (X_n) and the process $\{L(t),\ t \geq 0\}$ jointly such that

$$\sum_{k=1}^{n} X_k = L(n) + O(\sqrt{n}\,(\log\, n)^{1+\varepsilon}) \qquad \text{a.s. for any } \varepsilon > 0 \tag{1.10}$$

and for some sequence $a_n \asymp (n\ \log\, n)^{1/2}$ we have

$$a_n^{-1}\left(\sum_{k=1}^{n} X_k - L(n)\right) \stackrel{d}{\longrightarrow} N(0,1). \tag{1.11}$$

Here $c_n \asymp d_n$ means that the ratio c_n/d_n lies between positive constants. For an explicit construction of a_n, see (2.22). Due to the irregular tail behavior of the random variables in our construction (see the proof of Lemma 2.1), it seems likely that $a_n \asymp (n \log n)^{1/2}$ in Theorem 1.1 cannot be replaced by $a_n \sim c(n \log n)^{1/2}$ with a constant c.

The process $L(t)$ was introduced by Martin-Löf [16] who proved the scaling relation

$$g(2^m t) = 2^m (g(t) - imt).$$

From this it follows that the transformation $t \longrightarrow 2t$ does not change the distribution of the process

$$\{L(t)/t - \log_2 t,\, t > 0\}. \tag{1.12}$$

In particular, $L(2)/2 - 1 \stackrel{d}{=} L(1)$, and since $L(2) \stackrel{d}{=} L(1) \star L(1)$, the distribution of $L(1)$ is semistable. In view of the atomic Lévy measure in the characteristic function of $Z(1)$, its distribution is not stable. It also follows that

$$L(n)/n - \log_2 n \stackrel{d}{=} L(\gamma_n)/\gamma_n - \log_2 \gamma_n \stackrel{d}{=} G_{\gamma_n}, \tag{1.13}$$

showing that $L(n)/n - \log_2 n$ exhibits the merging behavior (1.3) in an ideal way, with zero error. Thus Theorem 1.1 gives an invariance principle for the merging result (1.3) and actually, for a class of further limit theorems for (X_n). It shows also the surprising fact that the partial sum process of (X_n) can be represented as a semistable Lévy process with an asymptotically normal perturbation.

In a previous paper [1], a strong approximation of St. Petersburg sums with the weaker remainder term $O(n^{5/6+\varepsilon})$ and without the asymptotic normality of the error term was proved by a standard blocking argument. The proof in [1] works for a large class of i.i.d. sequences (X_n) in the domain of geometric partial attraction of a semistable law G. In contrast, the proof of Theorem 1.1 uses the structure of the St. Petersburg sequences in a substantial way and whether Theorem 1.1 remains valid for a larger class of i.i.d. sequences remains open.

Weak and strong approximation of partial sums of i.i.d. random variables (X_n) in the domain of attraction of stable laws were proved in Stout [21], Simons and Stout [20], Berkes and Dehling [2]. The remainder terms there are given in terms of the function $\beta(x) = x^\alpha |P(X_1 < x) - G(x)|$, where G is the limit distribution, and are rather complicated. In the case when $\beta(x)$ is a slowly varying function tending to 0, lower bounds for the remainder term (valid for any construction) are also given in [2], leaving only a small gap between the upper and lower bounds. However, in the case of the stable analogue of St. Petersburg sums, when G is a stable distribution with parameters $\alpha = 1$, $\beta = -1$ (see e.g. [13], p. 164), we have $\beta(x) = O(x^{-\gamma})$ for some $\gamma > 0$ and no lower bounds for the remainder term have been found. For the same reason, we do not have universal lower bounds for the remainder term in the St. Petersburg game and thus, even though Theorem 1.1 determines the precise stochastic order of magnitude of the error term for a specific construction, the question whether other constructions can give a better error term remains open.

2 Proofs

We first prove (1.6). Clearly

$$P(X_1 > x) = \Psi(x)/x \qquad (x \geq 1). \tag{2.1}$$

Let F denote the distribution function of X_1 and let $F^{-1}(x) = \inf\{t : F(t) \geq x\}$ be its (generalized) inverse. Then

$$F^{-1}(x) = 2^k \quad \text{for} \quad x \in (1 - 2^{-(k-1)}, 1 - 2^{-k}], \quad k = 1, 2, \dots$$

and thus

$$F^{-1}(1 - x) = x^{-1}\Psi(x) \quad \text{for} \quad 0 < x < 1. \tag{2.2}$$

We also have

$$\Psi(2^{-k} x) = \Psi(x) \quad \text{for all } x \in \mathbb{R},\ k \in \mathbb{Z}.$$

As in the Introduction, let $\eta_1, \eta_2, \dots,$ be independent exp(1) random variables, $Z_k = \sum_{j=1}^k \eta_j$, $k = 1, 2, \dots$, put

$$X^*_{j,n} = F^{-1}\left(1 - \frac{Z_j}{Z_{n+1}}\right), \qquad 1 \leq j \leq n$$

and let $X_{1,n} \geq \ldots \geq X_{n,n}$ be the decreasing ordered sample of $X_1, \ldots, X_n$. By the well known representation of ordered samples (see e.g. [5], page 285 or [19], p. 335), the vector $(Z_1/Z_{n+1}, \ldots, Z_n/Z_{n+1})$ is distributed as the ordered sample $U_{n,1} \leq \ldots \leq U_{n,n}$ of i.i.d. uniform r.v.'s $U_1, \ldots, U_n$ in $(0,1)$ and thus the vectors $(X_{1,n}, \ldots, X_{n,n})$ and $(X^*_{1,n}, \ldots, X^*_{n,n})$ have the same distribution. By (2.2)

$$\begin{aligned} X^*_{j,n} &= F^{-1}\left(1 - \frac{Z_j}{Z_{n+1}}\right) = \frac{Z_{n+1}}{Z_j}\Psi\left(\frac{Z_j}{Z_{n+1}}\right) \\ &= (1+\varepsilon_n)\frac{n}{Z_j}\Psi\left(\frac{Z_j}{n}(1+\varepsilon_n)^{-1}\right) \end{aligned} \tag{2.3}$$

where $\varepsilon_n = Z_{n+1}/n - 1$. Now if $2^k \leq n < 2^{k+1}$, then $\gamma_n = n/2^{k+1}$ and thus from (2.3) we get

$$\frac{X^*_{j,n}}{n} = (1+\varepsilon_n)\frac{1}{Z_j}\Psi\left(\frac{Z_j}{\gamma_n 2^{k+1}}(1+\varepsilon_n)^{-1}\right) = (1+\varepsilon_n)\frac{1}{Z_j}\Psi\left(\frac{Z_j}{(1+\varepsilon_n)\gamma_n}\right) \tag{2.4}$$

which implies (1.6) immediately.

We turn now to the proof of Theorem 1.1, which uses, as in [2], [20], [21], a termwise approximation of partial sums. As it turns out (see Lemma 2.1 below), the termwise error in this approximation is determined by the second term of the expansion (1.5) whose tails were shown in [3] to be $\asymp x^{-2}$. This implies that the termwise error is in the domain of attraction of the normal law, explaining relation (1.11) in Theorem 1.1. The crucial influence of the second term of the expansion (1.5) in our approximation problem is similar to the convergence of Markov chains to the stationary distribution whose speed is determined by the second largest eigenvalue of the transition matrix.

Lemma 2.1 *A St. Petersburg variable X with distribution (1.1) and a random variable Y distributed as $Y^{(1)}$ in (1.5) can be jointly defined on a suitable probability space such that*

$$c_1 x^{-2} \leq P(|X - Y| > x) \leq c_2 x^{-2} \qquad (x \geq x_0) \tag{2.5}$$

for some positive constants c_1, c_2, x_0.

Proof. Put

$$W_1 = \frac{\Psi(Z_1)}{Z_1} - 1, \qquad W_2 = \frac{\Psi(1-e^{-Z_1})}{1-e^{-Z_1}}, \qquad W_3 = \sum_{k=1}^{\infty}\left(\frac{\Psi(Z_k)}{Z_k} - \frac{\Psi(k)}{k}\right). \tag{2.6}$$

We show that (2.5) holds with $X = W_2$, $Y = W_3$. Clearly, the distribution function of Z_1 is $G(x) = 1 - e^{-x}$, $(x \geq 0)$ and thus $U = G(Z_1) = 1 - e^{-Z_1}$ has distribution $U(0,1)$. Next we observe that for any $k \in \mathbb{Z}$ the function $\Psi(u)/u$ equals 2^k for

$u \in [2^{-k}, 2^{-k+1})$ and thus for a fixed $x \in [2^\ell, 2^{\ell+1})$, $\ell \in \mathbb{Z}$, the inequality $\Psi(u)/u > x$ holds iff $u < 2^{-\ell}$. Therefore for $x \in [2^\ell, 2^{\ell+1})$ we have

$$P(W_2 > x) = P\left(\Psi(U)/U > x\right) = P(U < 2^{-\ell}). \tag{2.7}$$

If $x \geq 1$, then $\ell \geq 0$ and thus the last probability in (2.7) equals $2^{-\ell}$; otherwise $\ell < 0$ and the last probability in (2.7) equals 1. Thus W_2 is a St. Petersburg variable. On the other hand, W_3 has distribution $Y^{(1)}$ in (1.5) and thus to prove Lemma 2.1 it suffices to show that

$$c_1 x^{-2} \leq P(|W_2 - W_3| > x) \leq c_2 x^{-2} \qquad (x \geq x_0). \tag{2.8}$$

We first prove that

$$c_3 x^{-2} \leq P(|W_1 - W_2| > x) \leq c_4 x^{-2} \qquad (x \geq x_1) \tag{2.9}$$

with some positive constants c_3, c_4, x_1. As already noted, for any $k \in \mathbb{Z}$ the function $\Psi(x)/x$ equals 2^k on the interval $I_k = [2^{-k}, 2^{-k+1})$. Let now $k \geq 2$ and assume $Z_1 \in I_k$. Then $0 < Z_1 < 1$, $0 < 1 - e^{-Z_1} < Z_1$ and

$$Z_1 - \frac{1}{2} Z_1^2 < 1 - e^{-Z_1} < Z_1 - \frac{1}{2} Z_1^2 + \frac{1}{6} Z_1^3. \tag{2.10}$$

Thus if $Z_1 \in I_k$, then for $k \geq 2$ we have by (2.10)

$$2^{-k+1} \geq Z_1 \geq 1 - e^{-Z_1} \geq Z_1 - \frac{1}{2} Z_1^2 \geq 2^{-k} - 2^{-2k+1} \geq 2^{-(k+1)}$$

showing that $1 - e^{-Z_1} \in I_k$ or $1 - e^{-Z_1} \in I_{k+1}$. Thus

$$\Delta = |W_1 - W_2 + 1| = \left| \frac{\Psi(Z_1)}{Z_1} - \frac{\Psi(1 - e^{-Z_1})}{1 - e^{-Z_1}} \right| \tag{2.11}$$

equals 0 or $2^{k+1} - 2^k = 2^k$ according as $1 - e^{-Z_1}$ belongs to I_k or I_{k+1}. In view of (2.10) the second alternative implies

$$2^{-k} \leq Z_1 < 1 - e^{-Z_1} + \frac{1}{2} Z_1^2 \leq 2^{-k} + \frac{1}{2} 2^{-2k+2}$$

and thus Z_1 is closer to the left endpoint of I_k than 2^{-2k+1}. But then $Z_1 \sim 2^{-k}$, $\frac{1}{2} Z_1^2 - O(Z_1^3) \sim \frac{1}{2} 2^{-2k}$ as $k \to \infty$ and thus by (2.10)

$$1 - e^{-Z_1} = Z_1 - \frac{1}{2} Z_1^2 + O(Z_1^3) = Z_1 - \left(\frac{1}{2} + o_k(1)\right) 2^{-2k}.$$

Consequently, the relation $1 - e^{-Z_1} \in I_{k+1}$, or, equivalently, $1 - e^{-Z_1} < 2^{-k}$ holds iff $Z_1 \in J_k$, where $J_k = [2^{-k}, 2^{-k} + u_k)$ with $u_k \sim \frac{1}{2} 2^{-2k}$. Since the density e^{-x} of Z_1 is $1 + O(2^{-k})$ on J_k, we have $P(Z_1 \in J_k) \sim \frac{1}{2} 2^{-2k}$ as $k \to \infty$. We thus proved that the difference Δ in (2.11) equals 2^k on a set A_k in the probability space, where the A_k are disjoint for $k \ge k_0$, $P(A_k) \sim \frac{1}{2} 2^{-2k}$ and otherwise $\Delta = 0$. Hence

$$P(\Delta > x) \sim \sum_{2^k > x} \frac{1}{2} 2^{-2k} = \frac{2}{3} 4^{-k_0} \qquad \text{as } x \to \infty$$

where $k_0 = k_0(x)$ denotes the smallest integer such that $2^{k_0} > x$. Thus if x runs in the interval $(2^s, 2^{s+1})$ for some integer $s \ge 1$, then $x^2 P(\Delta > x)$ runs from $\frac{1}{6} + o_s(1)$ to $\frac{4}{6} + o_s(1)$ as $s \to \infty$, which proves (2.9) and we also see that $x^2 P(\Delta > x)$ and thus $x^2 P(|W_1 - W_2| > x)$ fluctuates between positive constants, without a limit.

Next we observe that

$$W_3 - W_1 = \sum_{k=2}^{\infty} \left(\frac{\Psi(Z_k)}{Z_k} - \frac{\Psi(k)}{k} \right)$$

is a tail sum of the series representing $Y^{(1)}$ in (1.5) whose tail behavior is described by Theorem 5 of [3]; in particular we have

$$c_5 x^{-2} \le P(W_3 - W_1 > x) \le P(|W_3 - W_1| > x) \le c_6 x^{-2} \tag{2.12}$$

with suitable positive constants c_5, c_6. Theorem 5 of [3] also shows that $x^2 P(|W_1 - W_3| > x)$ has no limit as $x \to \infty$. Now (2.9) and (2.12) imply

$$P(|W_2 - W_3| > x) \le P(|W_2 - W_1| > x/2) + P(|W_1 - W_3| > x/2) \le c_7 x^{-2}, \tag{2.13}$$

proving the upper half of (2.8). To prove the lower half, we note that

$$\begin{aligned} P(|W_2 - W_3| > x) &\ge P(W_3 - W_2 > x) \ge P(W_3 - W_1 > 3x/2,\ W_1 - W_2 > -x/2) \\ &= P(W_3 - W_1 > 3x/2) - P(W_3 - W_1 > 3x/2,\ W_1 - W_2 \le -x/2) \\ &\ge P(W_3 - W_1 > 3x/2) - P(W_3 - W_1 > 3x/2,\ |W_1 - W_2| \ge x/2). \end{aligned} \tag{2.14}$$

For any $t \ge 0$, set

$$V_t = \sum_{k=2}^{\infty} \left(\frac{\Psi(t + Z_k^*)}{t + Z_k^*} - \frac{\Psi(k)}{k} \right), \tag{2.15}$$

where $Z_k^* = \sum_{j=2}^{k} \eta_j$ for $k \ge 2$. We claim that there exists a positive constant C such that for any $0 \le t \le 1$ we have

$$E|V_t| \le C. \tag{2.16}$$

Since the sequence (Z_k^*) has the same distribution as (Z_k), for $t = 0$ relation (2.16) follows from Lemma 2 of [3]. As inspection shows, the properties of (Z_k) used in the proof in [3] remain valid for the sequence $(Z_k + t)$ for any fixed $t \ge 0$ and moreover, the inequalities in [3] hold uniformly for $0 \le t \le 1$, proving (2.16). Now, conditionally on $Z_1 = t$, $W_3 - W_1$ becomes V_t in (2.15) which is independent of $\eta_1 = Z_1$ and thus of $\Delta = |W_1 - W_2 + 1|$ in (2.11) and consequently for $x \ge x_0$

$$\begin{aligned} &P(W_3 - W_1 > 3x/2,\, |W_1 - W_2| \ge x/2 \mid Z_1 = t) \\ &\le P(W_3 - W_1 > 3x/2,\, \Delta \ge x/4 \mid Z_1 = t) \\ &= P(V_t > 3x/2) I\{\Delta(t) \ge x/4\} \\ &\le \frac{2}{3x} E|V_t| I\{\Delta(t) \ge x/4\} \le C x^{-1} I\{\Delta(t) \ge x/4\}, \end{aligned} \tag{2.17}$$

where $\Delta(t)$ is the expression in (2.11) with Z_1 replaced by t. If Z_1 is bounded away from 0, then $|W_1 - W_2|$ is bounded above, or putting differently, if $|W_1 - W_2|$ is large, then Z_1 is near 0. Thus integrating (2.17) over $0 \le t \le 1$ with respect to $P(Z_1 \in dt)$ we get

$$P(W_3 - W_1 > 3x/2,\, |W_1 - W_2| \ge x/2) \le C x^{-1} P(|\Delta| > x/2) \le C' x^{-3} \tag{2.18}$$

for sufficiently large x, where in the last step we used (2.9). Now using (2.12), (2.14) and (2.18) we get the lower half of (2.8).

Proof of Theorem 1.1. For the vector (X, Y) in Lemma 2.1, let H denote the distribution function of $X - Y$ and put

$$U(x) = \int_{|t| \le x} t^2 dH(t).$$

Using Lemma 2.1 and integration by parts we get

$$\begin{aligned} U(x) &= -x^2(1 - H(x) + H(-x)) + \int_0^x 2t(1 - H(t) + H(-t))dt \\ &= O(1) + \int_0^x 2t(1 - H(t) + H(-t))dt \end{aligned} \tag{2.19}$$

provided that x and $-x$ are continuity points of H. Using Lemma 2.1 again for the last integral it follows that

$$c_8 \log x \le U(x) \le c_9 \log x \quad \text{and} \quad U(2x) - U(x) = O(1) \quad \text{as } x \to \infty \tag{2.20}$$

with suitable positive constants c_8 and c_9. Thus $\lim_{x\to\infty} U(2x)/U(x) = 1$, i.e. the nondecreasing function U is slowly varying. Further, (2.5) implies that H has a finite expectation. Let now (X_n, Y_n) be i.i.d. copies of the vector (X, Y) in Lemma 2.1. By the slow variation of U, $X - Y$ is in the domain of attraction of the normal law, specifically we have

$$\frac{1}{a_n} \sum_{k=1}^{n} (X_k - Y_k - c) \xrightarrow{d} N(0, 1) \tag{2.21}$$

where $c = E(X - Y)$ and

$$a_n = \inf\{x \in (0, \infty) : n x^{-2} U(x) \le 1\}. \tag{2.22}$$

(See e.g. [12], p. 580, Theorem 3 and the comment after (5.23) on page 579.) Using (2.22) and the first relation of (2.20), we get by a simple calculation

$$c_{10}(n\log n)^{1/2} \le a_n \le c_{11}(n\log n)^{1/2} \tag{2.23}$$

with suitable constants c_{10}, c_{11}. Recall now that along the sequence $n = 2^k$ we have

$$\frac{1}{n}\sum_{k=1}^{n} X_k - \log_2 n \stackrel{d}{\longrightarrow} G, \quad \frac{1}{n}\sum_{k=1}^{n} Y_k - \log_2 n \stackrel{d}{\longrightarrow} G \tag{2.24}$$

where $G = G_{1/2}$ is the semistable distribution defined after (1.2). The first relation here follows from (1.4) and the second from (1.13), since $\sum_{k=1}^{n} Y_k \stackrel{d}{=} L(n)$. Relation (2.23) shows that replacing $1/a_n$ by $1/n$ in (2.21), the left hand side will converge to 0 in probability and adding the second relation of (2.24) yields

$$\frac{1}{n}\sum_{k=1}^{n} X_k - \log_2 n - c \stackrel{d}{\longrightarrow} G$$

which, together with the first relation of (2.24), implies $c = 0$ and thus (2.21) yields

$$\frac{1}{a_n}\sum_{k=1}^{n}(X_k - Y_k) \stackrel{d}{\longrightarrow} N(0,1). \tag{2.25}$$

Since the process $\{\sum_{k=1}^{n} Y_k,\ n \ge 1\}$ has the same distribution as $\{L(n),\ n \ge 1\}$ where L is the Lévy process defined by (1.9), relation (1.11) is proven.

To prove (1.10), let $b_n = \sqrt{n}(\log n)^{1+\varepsilon}$, $\varepsilon > 0$. Then using Lemma 2.1, (2.20) and integration by parts one can easily verify the relations

$$\sum_{n=1}^{\infty} \frac{1}{b_n^2}\int_{|x|<b_n} x^2 dH(x) < +\infty, \quad \frac{1}{b_n}\sum_{k=1}^{n}\int_{|x|<b_n} x dH(x) \longrightarrow 0 \tag{2.26}$$

and

$$\sum_{n=1}^{\infty} P(|X-Y| \ge b_n) < +\infty.$$

(In the case of the second relation of (2.26) replace the domain $\{|x| < b_n\}$ of integration by $\{|x| \ge b_n\}$ in view of $E(X-Y) = 0$.) Thus using Theorem 6.8 in Petrov [18], p. 211 we get

$$\frac{1}{b_n}\sum_{k=1}^{n}(X_k - Y_k) \longrightarrow 0 \quad \text{a.s.},$$

yielding (1.10).

References

[1] Berkes, I.: Strong approximation of the St. Petersburg game. *Statistics* 51 (2017), 3–10.

[2] Berkes, I. and Dehling, H.: Almost sure and weak invariance principles for random variables attracted by a stable law. *Prob. Theory Rel. Fields* 83 (1989), 331–353.

[3] Berkes, I., Győrfi, L. and Kevei, P.: Tail probabilities of St. Petersburg sums, trimmed sums, and their limit. *J. Theor. Probability* 30 (2017), 1104–1129.

[4] Bernoulli, D.: Specimen theoriae novae de mensura sortis. Comm. Acad. Sci. Imp. Petropolitanae 5 (1738), 175-192. German translation: A. Pringsheim: Versuch einer neuen Theorie der Wertbestimmung von Glücksfällen, Verlag von Duncker & Humblot, Leipzig, 1896.

[5] Breiman, L.: Probability. Second edition, SIAM, 1992.

[6] Csörgő, S.: Rates of merge in generalized St. Petersburg games. *Acta Sci. Math.* (Szeged) 68 (2002), 815–847.

[7] Csörgő, S.: Merging asymptotic expansions in generalized St. Petersburg games. *Acta Sci. Math.* (Szeged) 73 (2007), 297–331.

[8] Csörgő, S. and Dodunekova, R.: Limit theorems for the Petersburg game, in: Trimmed sums and extremes (M. G. Hahn, D. M. Mason, D. C. Weiner, eds.) pp. 285–315. *Progress in Probability*, Vol. 23, Birkhäuser, Boston 1991.

[9] Csörgő, S., Simons, G.: On Steinhaus' resolution of the St. Petersburg paradox. Probab. Math. Stat. 14 (1993), 157-172.

[10] Csörgő, S., Simons, G.: A bibliography of the St. Petersburg paradox. University of Szeged, 2014.

[11] Feller, W.: Note on the law of large numbers and "fair" games. *Ann. Math. Statist.* 16 (1945), 301–304.

[12] Feller, W.: An introduction to probability theory and its applications, Vol II, second edition. Wiley, 1971.

[13] Gnedenko, B.V. and Kolmogorov, A.N.: Limit distributions of sums of independent random variables. Addison-Wesley, 1954.

[14] Kern, P. and Wedrich, L.: Dimension results related to the St. Petersburg game. Probab. Math. Stat. 34 (2014), 97–117.

[15] LePage, R., Woodroofe, M. and Zinn, J.: Convergence to a stable distribution via order statistics. *Ann. Probab.* 9 (1981), 624–632.

[16] Martin-Löf, A.: A limit theorem which clarifies the 'Petersburg paradox'. *J. Appl. Probab.* 22 (1985), 634–643.

[17] Pap, Gy.: The accuracy of merging approximation in generalized St. Petersburg games. *J. Theoret. Probab.* 24 (2011), 240–270.

[18] Petrov, V.V.: Limit theorems of probability theory. Oxford University Press, 1995.

[19] Shorack, G. and Wellner, J.: Empirical processes with applications to statistics. Wiley, 1986.

[20] Simons, G. and Stout, W.: A weak invariance principle with applications to domains of attraction. *Ann. Probab.* 6 (1978), 294–315.

[21] Stout, W.F: Almost sure invariance principles when $EX_1^2 = \infty$. *Z. Wahrsch. Verw. Gebiete* 49 (1979), 23–32.