

## THE LIMITING DISTRIBUTION OF THE ST. PETERSBURG GAME

ILAN VARDI

(Communicated by Dennis A. Hejhal)

**ABSTRACT.** The St. Petersburg game is a well-known example of a random variable which has infinite expectation. Csörgő and Dodunekova have recently shown that the accumulated winnings do not have a limiting distribution, but that if measurements are taken at a subsequence  $b_n$ , then a limiting distribution exists exactly when the fractional parts of  $\log_2 b_n$  approach a limit. In this paper the characteristic functions of these distributions are computed explicitly and found to be continuous, self-similar, nowhere differentiable functions.

A single trial in the St. Petersburg game consists in tossing a fair coin until it falls heads and then the player receives  $2^k$  dollars if this happens at the  $k$ th throw. The expected value comes out to be

$$\frac{1}{1} + \frac{2}{2} + \cdots + \frac{2^k}{2^k} + \cdots = \infty,$$

so the St. Petersburg game gives an example of a random variable which has infinite expectation. However, the probability that this game last long enough to get large paybacks is quite low and the law of large numbers still holds. In fact, Feller showed [3, p. 237] that if one only allows a finite number  $n$  of trials, then for any fixed  $\varepsilon > 0$

$$\Pr \left\{ \left| \frac{S_n}{n \log_2 n} - 1 \right| > \varepsilon \right\} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where  $S_n$  denotes the accumulated winnings in  $n$  trials of the St. Petersburg game and  $\log_2 n = \frac{\log n}{\log 2}$ . In this sense the game can be given a "fair" entrance fee of  $n \log_2 n$  dollars.

It follows from Feller's result that a way of generating a fair entrance fee independent of the number of trials is to find a sequence  $\alpha_1, \alpha_2, \dots$  satisfying

$$\frac{\alpha_1 + \cdots + \alpha_n}{n \log_2 n} \rightarrow 1.$$

---

Received by the editors December 10, 1993.  
1991 *Mathematics Subject Classification.* Primary 60F05.

Now it was shown by Khinchin [7] that the sequence of continued fraction coefficients  $a_1(z), \dots, a_n(z)$  of real numbers  $0 < z < 1$  has the similar property

$$\lambda \left\{ z : \left| \frac{a_1(z) + \dots + a_n(z)}{n \log_2 n} - 1 \right| > \varepsilon \right\} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where  $\lambda$  is Lebesgue measure. This means that, with high probability, a fair entry fee for the St. Petersburg game can be chosen by picking a random number  $x$  and letting the entry fee at the  $n$ th trial be  $a_n(x)$ . For example, if one believes that  $\pi$  is a “pseudo-random” number, then its continued fraction digits could be used (the first 17 million digits have been computed by R.W. Gosper and they do seem to behave like those of a random real number).

It is interesting to see how well continued fraction coefficients model the behavior of the St. Petersburg game, so one compares the error distribution from the main term. The partial sums of continued fraction coefficients have a limiting error distribution (see [5] and [9]) given by the distribution function  $G(x)$  with characteristic function

$$g(t) = \int_{-\infty}^{\infty} e^{itx} dG(x) = \exp \left( -it \log |t| - \frac{\pi |t|}{2 \log 2} \right).$$

In other words, one has that

$$\left| \lambda \left\{ 0 < z < 1 : \frac{a_1(z) + \dots + a_n(z)}{n} - \log_2 n < x \right\} - G(x) \right| \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Note that  $G(x) = G_{1,1}(x, \pi/2 \log 2)$ , where  $G_{\alpha,\beta}(x, \gamma)$  is the stable distribution with characteristic function

$$g_{\alpha,\beta}(t, \gamma) = \int_{-\infty}^{\infty} e^{itx} dG_{\alpha,\beta}(x, \gamma)$$

whose logarithm is

$$\log g_{\alpha,\beta}(t, \gamma) = -\gamma |t|^\alpha (1 + i\beta \operatorname{sgn} t \omega(t, \alpha)),$$

where  $0 < \alpha \leq 2$ ,  $|\beta| \leq 1$ ,  $\gamma > 0$ , and  $\omega(t, \alpha) = \frac{2}{\pi} \log |t|$  for  $\alpha = 1$  (otherwise  $\tan(\alpha\pi/2)$ ); see [6].

The next step is to see whether the  $S_n/n - \log_2 n$  have a limiting distribution and whether it is the same as for continued fraction digits. This is answered by

**Theorem 1.** *Let  $F_n(x)$  be the distribution of*

$$\frac{S_n}{n} - \log_2 n.$$

*If  $b_n$  is any sequence of integers such that  $\lim_{n \rightarrow \infty} b_n = \infty$ , then  $F_{b_n}(x)$  approaches a limit distribution if and only if  $\lim_{n \rightarrow \infty} \{\log_2 b_n\}$  exists, where  $\{x\}$  is the fractional part of  $x$ . In this case let  $b = \lim \{\log_2 b_n\}$ . Then  $F_{b_n}(x)$  has*

a limit  $F(x, b)$  with characteristic function  $g(t, b)$ , where

$$\log g(t, b) = -it \log_2 |t| - \frac{\pi |t|}{2 \log 2} - \frac{it}{\log 2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \Gamma\left(-1 + \frac{2\pi ik}{\log 2}\right) e^{-\text{sign}(t)\pi^2 k / \log 2} e^{-2\pi ik(\log_2 |t| - b)}.$$

*Remarks.* Stirling’s asymptotic formula for  $\Gamma(z)$ , see (2) below, shows that the  $k$ th term of the above series is  $O(k^{-3/2})$ , so the series converges absolutely.

The fact that  $F_{b_n}(x)$  has a limit if and only if  $\{\log_2 b_n\}$  converges was shown in [1] using methods of probability theory. A preliminary result was also proved in [8].

Theorem 1 shows that when the limit exists, the first two terms of the limiting distribution are the same as the limiting distribution of continued fraction digits.

The function multiplying  $t$  in the last summand is a continuous, nowhere differentiable function which is periodic in  $\log_2 t$  and so “self-similar” in  $t$ . It is reminiscent of a function of Delange [2] and the proof below is similar to a proof of Delange’s result given in [4].

PROOF OF THEOREM 1

To study the distribution of the error in the St. Petersburg game one analyzes the characteristic function of the distribution after  $n$  steps. This is given by

$$\begin{aligned} g_n(t) &= \text{expectation } \{ \exp [it (S_n/n - \log_2 n)] \} \\ &= e^{-it \log_2 n} \sum_{1 \leq k_1, \dots, k_n < \infty} \frac{\exp[it(2^{k_1} + \dots + 2^{k_n})/n]}{2^{k_1 + \dots + k_n}} \\ &= e^{-it \log_2 n} \left( \sum_{k=1}^{\infty} \frac{e^{it2^k/n}}{2^k} \right)^n. \end{aligned}$$

There is a limiting distribution if and only if  $\lim_{n \rightarrow \infty} g_n(t)$  exists for each  $t$  and is continuous at  $t = 0$ . In the case at hand, one takes logarithms which reduces the problem to evaluating

$$\lim_{n \rightarrow \infty} n \log \left( \sum_{k=1}^{\infty} \frac{e^{it2^k/n}}{2^k} \right) - it \log_2 n.$$

One must therefore understand the asymptotic behavior of

$$p(x) = \sum_{k=1}^{\infty} \frac{e^{ix2^k}}{2^k}$$

as  $x \rightarrow 0$  and this is given by

**Proposition 1.**  $p(x)$  has the asymptotic series

$$p(x) = 1 - ix \log_2 |x| - \frac{\pi |x|}{2 \log 2} - \frac{ix}{\log 2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \Gamma\left(-1 + \frac{2\pi ik}{\log 2}\right) e^{-\text{sign}(x)\pi^2 k / \log 2} e^{-2\pi ik \log_2 |x|} + \sum_{n=2}^N \frac{i^n x^n}{(2^{1-n} - 1) n!} + O(x^{N+1}), \quad \text{as } x \rightarrow 0.$$

*Proof of Proposition 1.* The proof uses the Mellin transform

$$f^*(s) = \int_0^{\infty} f(x) x^{s-1} dx.$$

For example,

$$\Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1} dx.$$

The inversion formula is

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f^*(s) x^{-s} ds, \quad c > 0, \quad x > 0.$$

This gives the formal identity

$$e^{ix} = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Gamma(s) (-i)^{-s} x^{-s} ds,$$

where the integral fails to converge. Instead, one uses

**Proposition 2.**

$$(1) \quad e^{ix} - 1 = \frac{1}{2\pi i} \int_{-3/4-i\infty}^{-3/4+i\infty} \Gamma(s) e^{i\pi s/2} x^{-s} ds.$$

*Proof of Proposition 2.* The first thing to note is that the integral converges absolutely due to Stirling's formula

$$\log \Gamma(s) = \left(s - \frac{1}{2}\right) \log s - s + \frac{1}{2} \log 2\pi + O(|s|^{-1}),$$

for  $-\pi + \delta < \arg s < \pi - \delta$ , implying that

$$(2) \quad \Gamma(s) = O(|t|^{\sigma_0-1/2} e^{-\pi|t|/2}), \quad \sigma_1 \leq \sigma \leq \sigma_0, \quad \sigma_0, \sigma_1 \text{ fixed},$$

with the usual convention that  $s = \sigma + it$ . The identity (1) is checked by moving the line of integration to the left and picking up residues. What this

means is that the integral

$$I(T, N) = \frac{1}{2\pi i} \int_{C(T, N)} \Gamma(s) e^{ins/2} x^{-s} ds$$

is considered, where  $C(T, N)$  is the counterclockwise rectangular path of integration consisting of the four line segments

$$C_1 = \left\{ -\frac{3}{4} + it, -T \leq t \leq T \right\}, \quad C_2 = \left\{ \sigma + iT, -N - \frac{3}{4} \leq \sigma \leq -\frac{3}{4} \right\},$$

$$C_3 = \left\{ -N - \frac{3}{4} + it, -T \leq t \leq T \right\}, \quad C_4 = \left\{ \sigma - iT, -N - \frac{3}{4} \leq \sigma \leq -\frac{3}{4} \right\}.$$

By the above remark, the integral on  $C_1$  converges absolutely to the right-hand side of (1).

To bound the remaining three integrals, recall the recursion formula  $\Gamma(s + 1) = s\Gamma(s)$  which gives

$$|\Gamma(s - k)| \leq \frac{2^k |\Gamma(s)|}{k!}, \quad \text{for } -1 - \frac{3}{4} \leq \sigma \leq -\frac{3}{4}, \quad s \neq -1.$$

Thus, the integral along  $C_2$  is bounded using

$$I_{C_2} = \sum_{k=1}^N \int_{-k+1-3/4}^{-k-3/4} \Gamma(\sigma + iT) e^{i\pi(\sigma+iT)/2} x^{-(\sigma+iT)} d\sigma$$

$$\leq \sum_{k=1}^N \frac{2^k}{k!} \int_{-3/4}^{-1-3/4} |\Gamma(\sigma + iT) e^{i\pi(\sigma+iT)/2} x^{-(\sigma+iT)}| d\sigma$$

$$= o(1), \quad \text{as } T \rightarrow \infty, \text{ independent of } N,$$

since the bound of (2) shows that the integral goes to zero as  $T \rightarrow \infty$ .

The exact same argument applies to show that  $I_{C_4} \rightarrow 0$  as  $T \rightarrow \infty$ , and a similar argument applies to the integral on  $C_3$  showing that it is  $O(2^N/N!)$ , independent of  $T$ .

Combining these estimates shows that  $\lim_{T, N \rightarrow \infty} I(T, N)$  converges to the right-hand side of (1). On the other hand, Cauchy's formula can be used to evaluate the contour integral  $I(T, N)$ . Since  $\Gamma(s)$  has poles at  $n = -1, -2, \dots$ , with residues  $(-1)^n/n!$ , this gives

$$I(T, N) = \sum_{n=1}^N \frac{(-1)^n}{n!} e^{-in\pi/2} x^n = \sum_{n=1}^N \frac{(ix)^n}{n!},$$

which converges to the left side of (1) as  $N \rightarrow \infty$ .  $\square$

Continuing the proof of Proposition 1, note that the integral formula (1) gives

$$\frac{e^{i\alpha x} - 1}{\alpha} = \frac{1}{2\pi i} \int_{-3/4-i\infty}^{-3/4+i\infty} \Gamma(s) e^{ins/2} \alpha^{-s-1} x^{-s} ds$$

and summing over  $\alpha = 2^k$ ,  $k = 1, 2, \dots$ , gives

$$p(x) = 1 + \frac{1}{2\pi i} \int_{-3/4-i\infty}^{-3/4+i\infty} \frac{\Gamma(s)e^{ins/2}x^{-s}}{2^{s+1}-1} ds.$$

Exactly as in the proof of (1), this integral can be moved over to the left of  $\text{Re}(s) = -1$  and the difference will be given by the residues of the integrand at its poles, i.e., at

$$s = -1, \quad s = -1 + \frac{2\pi ik}{\log 2}, \quad k = \pm 1, \pm 2, \dots,$$

and

$$s = -n, \quad n = 2, 3, \dots$$

At  $s = -1$  one has

$$\begin{aligned} \Gamma(s) &= \frac{-1}{s+1} + O(1), \\ e^{ins/2} &= e^{-\pi i/2} \left( 1 + \frac{i\pi(s+1)}{2} + O((s+1)^2) \right), \\ x^{-s} &= x(1 - (s+1)\log_2 x + O((s+1)^2)), \\ \frac{1}{2^{s+1}-1} &= \frac{1}{\log 2(s+1) + O((s+1)^2)}, \end{aligned}$$

so the residue is

$$-ix \log_2 x - \frac{\pi x}{2 \log 2}.$$

The residues at  $-1 + 2\pi ik/\log 2$  are

$$-\frac{ix \Gamma(-1 + 2\pi ik/\log 2)}{\log 2} e^{-\pi^2 k/\log 2} e^{-2\pi ik \log_2 x}.$$

The residues at  $-n$ ,  $n = 2, 3, \dots$ , are

$$\frac{(-1)^n (-i)^n x^n}{n! (2^{1-n} - 1)}.$$

This gives the asymptotic formula

$$\begin{aligned} p(x) &= 1 - ix \log_2 x - \frac{\pi x}{2 \log 2} \\ (3) \quad &- \frac{ix}{\log 2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \Gamma\left(-1 + \frac{2\pi ik}{\log 2}\right) e^{-\pi^2 k/\log 2} e^{-2\pi ik \log_2 x} \\ &+ \sum_{n=2}^N \frac{i^n x^n}{(2^{1-n} - 1) n!} + O(x^{N+1}) \end{aligned}$$

for  $x > 0$ .

If  $x < 0$ , then one has a corresponding integral formula

$$e^{ix} - 1 = \frac{1}{2\pi i} \int_{-3/4-i\infty}^{-3/4+i\infty} \Gamma(s)e^{-i\pi s/2}(-x)^{-s} ds,$$

which combined with (3) proves Proposition 1.  $\square$

*Proof of Theorem 1.* Substituting the asymptotic formula for  $p(x)$  into the formula for  $g_n(t)$  and using the fact that

$$\log(1 + x) = x + O(x^2)$$

for small  $x$ , gives

$$\begin{aligned} g_n(t) &= n \log p(t/n) - it \log n \\ &= -it \log_2 |t| - \frac{\pi |t|}{2 \log 2} \\ &\quad - \frac{it}{\log 2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \Gamma\left(-1 + \frac{2\pi ik}{\log 2}\right) e^{-\text{sign}(t)\pi^2 k / \log 2} e^{-2\pi ik(\log_2 |t| - \log_2 n)} + o(1) \end{aligned}$$

where the  $o(1)$  term depends on  $t$ . It follows from the fact that equal functions have the same Fourier coefficients that  $g_{b_n}(t)$  approaches a limit as  $n \rightarrow \infty$  if and only if  $e^{2\pi i \log_2 b_n}$  approaches a limit, i.e., exactly when  $\{\log_2 b_n\}$  converges. In this case let  $b = \lim\{\log_2 b_n\}$ ; then for any fixed  $t$

$$\begin{aligned} \lim_{n \rightarrow \infty} g_{b_n}(t) &= -it \log_2 |t| - \frac{\pi |t|}{2 \log 2} \\ &\quad - \frac{it}{\log 2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \Gamma\left(-1 + \frac{2\pi ik}{\log 2}\right) e^{-\text{sign}(t)\pi^2 k / \log 2} e^{-2\pi ik(\log_2 t - b)}, \end{aligned}$$

completing the proof of Theorem 1.  $\square$

ACKNOWLEDGMENT

I would like to thank P. Diaconis, D. Hejhal, and S. Csörgő for helpful remarks.

REFERENCES

1. S. Csörgő and R. Dodunekova, *Limit theorems for the Petersburg game*, Sums, Trimmed Sums and Extremes (M. G. Hahn, D. M. Mason, and D. C. Weiner, eds.), Progr. Probab., vol. 23, Birkhauser, Boston, 1991, pp. 285–315.
2. H. Delange, *Sur la fonction sommatoire de la fonction "somme des chiffres"*, Enseign. Math. (2) **21** (1975), 31–47.
3. W. Feller, *An introduction to probability theory and its applications*, Vol. I, 2nd ed., Wiley, New York, 1964.
4. P. Flajolet, P. Grabner, and P. Kirschenhofer, *Mellin transforms and asymptotics: Digital sums*, Theoretical Computer Science (in press).

5. L. Heinrich, *Rates of convergence in stable limit theorems for sums of exponentially  $\psi$ -mixing random variables with applications to metric theory of continued fractions*, *Math. Nachr.* **131** (1987), 149–165.
6. I. A. Ibragimov and Yu. V. Linnik, *Independent and stationary sequences of random variables*, Noordhoff, Groningen, 1975.
7. A. Ya. Khinchin, *Metrische Kettenbruchprobleme*, *Compositio Math.* **1** (1935), 361–382.
8. A. Martin-Löf, *A limit theorem which clarifies the 'Petersburg paradox'*, *J. Appl. Probab.* **22** (1985), 634–643.
9. W. Philipp, *Limit theorems for sums of partial quotients of continued fractions*, *Monatsh. Math.* **105** (1988), 195–206.

MATHEMATICAL SCIENCES RESEARCH INSTITUTE, 1000 CENTENNIAL DR., BERKELEY, CALIFORNIA 94720

*E-mail address:* ilan@msri.org