LMD konf. darbai, 33–38
1997, Vilnius

Sums of digits obey the Strassen law

E. Manstavičius* (VU)

This remark continues the author's investigation [7]. To start with, we note that $\mathbf{N}_0 = \mathbf{N}\cup\{0\}$ is free additive semigroup which basis is an arbitrary numeration system $U = \{u_j\},\ j \geqslant 0,\ u_j \in \mathbf{N}$. Thus each $m \in \mathbf{N}_0$ has the unique finite expression

$$m = \sum_{j\geqslant 0} d_j(m)u_j \tag{1}$$

with $d_j(m) \in \mathbf{N}_0$. As in [6], we confine ourselves to the mixed radix numeration system (Cantor system) U which (see [3] for more information) is determined by some sequence $b_j \geqslant 2,\ j \geqslant 1$ of natural numbers via $u_0 = 1,\ u_j = b_j u_{j-1}$ when $j \geqslant 1$. As it is proved in [3], the representation (1) of $m \in \mathbf{N}_0$ is unique if $0 \leqslant d_j(m) < b_{j+1},\ j \geqslant 0$. Denote

$$s_k(m) := \sum_{0\leqslant j\leqslant k} d_j(m), \qquad s_m(m) = s(m).$$

Let

$$\nu_n(A) := \frac{1}{n}\#\{0 \leqslant m < n,\ m \in A\}$$

be the probability measure defined on the subsets A of $\mathbf{N}_0$. Probabilistic properties of the sums of digits are fairly interesting. We name few of them in the case of q-adic numeration system, e.g. when $b_j \equiv q \geqslant 2$. H. Delange [1] obtained the following asymptotical formula for the mean value

$$\frac{1}{n}\sum_{m=0}^{n-1} s(m) = \frac{q-1}{2}N + F(N), \qquad N = \frac{\log n}{\log q},$$

where F is a suitable continuous and nowhere differentiable function of period 1. Exact bounds for the error $F(N)$ had been given in [2]. Denote

$$D_n(t) = \frac{1}{\lambda\sqrt{N}}\left(s_{Nt}(m) - \frac{(q-1)Nt}{2}\right), \qquad \lambda^2 = (q^2-1)/12, \qquad t \in [0,1].$$

We proved in [6] that the probability measure $\nu_n \cdot D_n^{-1}$ defined on the Borel σ-algebra of the space $\mathbf{D} = \mathbf{D}[0,1]$ weakly converges to the Wiener measure. In this case we

*The research is partially supported by Lithuanian Science and Studies Fund.

may assume that **D** is endowed with the topology defined by the supremum metric denoted in the sequel by ρ. We have from this result

$$\nu_n\left(\max_{k\leqslant N}\left|s_k(m)-\frac{(q-1)k}{2}\right|<x\lambda\sqrt{N}\right)$$
$$\longrightarrow\frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp\left\{-\frac{\pi^2(2k+1)^2}{8x^2}\right\},\qquad x>0.$$

Even the last relation gives little information about the bounds of the sequence $s_k(m)$ when $k\Longrightarrow\infty$ for "almost all m". Thus, the purpose of the present investigation is to fill up this gap by giving the exact order of growth for $s_k(m)$. We will extend this problem to a functional one and establish an analog of the Strassen functional law of iterated logarithm. In proving that we will use the author's approach [7] having its origin in the Kubilius' study [4]. The ideas of the paper [6] dealing with the strong invariance principle for additive arithmetic functions defined on the semigroup **N** will be also exploited.

Now denote

$$n=:a_Nu_N+\cdots+a_1u_1+a_0,\qquad N=N(n)=\max\{k:\ u_k\leqslant n\};$$

$$\bar d_j(m)=d_j(m)-(b_{j+1}-1)/2,\qquad \sigma_j^2=(b_{j+1}^2-1)/12$$

for $0\leqslant j\leqslant N-1$ and

$$\bar d_N(m)=d_N(m)-a_N/2,\qquad \sigma_N^2=a_N(a_N+2)/12.$$

Let

$$B_k=\sigma_0^2+\cdots+\sigma_{k-1}^2,\qquad \beta(k)=\sqrt{2B_k\,LLB_k},\qquad 0\leqslant k\leqslant N$$

where $Lu=\log\max\{u,e\}$ and

$$\bar S_k(m,t)=\frac{1}{\beta(k)}\sum_{\substack{j\geqslant 0\\ B_j\leqslant tB_k}}\bar d_j(m),\qquad k\leqslant N,\qquad t\in[0,1].$$

Let $\widehat{S}_k(m,t)$ be the piecewise linear curve in the coordinate plane joining the points $(0,0),\ldots,(B_l/B_k,\bar S_k(m,B_l/B_k))$, $l=0,1,\ldots,k$. We recall that the Strassen set $\mathcal{K}$ is comprised by absolutely continuous functions g such that $g(0)=0$ and

$$\int_0^1\left(g'(t)\right)^2dt\leqslant 1.$$

For the processes X_k in **D** defined on some sequence of probability spaces $\{\Omega_n,\mathcal{F}_n,P_n\}$, say, for $k\leqslant N=N(n)$, we denote

$$X_k\Longrightarrow\mathcal{K}\qquad(P_n-\text{a.s.})\tag{2}$$

if the following two relations hold

$$\lim_{x\to\infty} \limsup_{n\to\infty} P_n\left(\max_{x\leqslant k\leqslant N} \rho(X_k, \mathcal{K}) \geqslant \varepsilon \right) = 0$$

and

$$\lim_{x\to\infty} \liminf_{n\to\infty} P_n\left(\min_{x\leqslant k\leqslant N} \rho(X_k, X) < \varepsilon \right) = 1$$

for each $\varepsilon > 0$ and each $X \in \mathcal{K}$. Here as usually, $\rho(X, \mathcal{A}) = \inf\{\rho(X, Y) : \ Y \in \mathcal{A}\}$. The measurability of this distance will be provided by the very construction of the probability spaces. If $P_n = P$ does not depend on n, our notion (2) coincides with the Strassen's invariance priciple, e.g., it means that the sequence X_k is relatively compact ant $\mathcal{K}$ is the set of its partial limits P almost surely.

THEOREM. *Let S_k be either of $\bar{S}_k$ or $\widehat{S}_k$ and*

$$b_j = o\left(\sqrt{B_j / LL\, B_j}\right) \tag{3}$$

as $j \to \infty$. Then

$$S_k(m, .) \Longrightarrow \mathcal{K} \qquad (\nu_n - a.s.). \tag{4}$$

Applying suitable continuous functionals in the relation (4), as in [6], we obtain

COROLLARY. *If the condition* (3) *of Theorem is satisfied, then we have ν_n-a.s.:*

i) $S_k(m, 1) \Longrightarrow [-1, 1]$;

ii) $(S_k(m, 1/2), S_k(m, 1)) \Longrightarrow \mathcal{L} := \{(u, v) : u^2 + (v - u)^2 \leqslant 1/2\}$;

iii) $S_k(m, 1/2) \Longrightarrow [-\sqrt{2}/2, \sqrt{2}/2]$;

iv) *if k_1 is the subsequence for which $S_{k_1}(m, 1/2) \to \sqrt{2}/2$, then $S_{k_1}(m, .) \to g_1$, where*

$$g_1(t) = \begin{cases} t\sqrt{2} & \text{if } 0 \leqslant t \leqslant 1/2, \\ \sqrt{2}/2 & \text{if } 1/2 \leqslant t \leqslant 1; \end{cases}$$

v) *if k_1 is the subsequence for which $S_{k_1}(m, 1/2) \to 1/2$ and $S_{k_1}(m, 1) \to 0$, then $S_{k_1}(m, t) \to g_2$, where*

$$g_2(t) = \begin{cases} t & \text{if } 0 \leqslant t \leqslant 1/2, \\ 1 - t & \text{if } 1/2 \leqslant t \leqslant 1. \end{cases}$$

Thus in the case of q-adic numeration system, we have

$$|s_k(m) - (q - 1)k/2| \leqslant (1 + \varepsilon)\lambda\sqrt{2kLL\, k}$$

"for almost all" m when $x \leqslant k \leqslant N$ with sufficiently large x and n. This estimate is sharp in the sense that ε can not be changed by $-\varepsilon$. In order to improve the remainder term $\varepsilon\sqrt{2kLLk}$, one had to apply the Feller's type law of iterated logarithm and the approach of the present remark. Going along these lines one can obtain the estimate

$$|s_k(m) - (q-1)k/2| \leqslant \lambda\Big(2k\big(L_2k + \frac{3}{2}L_3 k + L_4k + \cdots + (1+\varepsilon)L_pk\big)\Big)^{1/2}$$

valid for each $p \geqslant 4$ and $\varepsilon > 0$. Here $L_jk = L(L_{j-1}k)$.

Proof of Theorem. The main auxilliary results have been obtained by the author in [6]. To quote them, we need some notation. Let $\{I_l\}$ and $\{J_l\}$ be arbitrary systems of subsets of the sets $\{0, 1, \dots, r\}$ and $\{0, 1, \dots, N\}$ respectively, where $1 \leqslant l \leqslant p$, $p \geqslant 1$ and $1 \leqslant r \leqslant N-1$. Put

$$s_I(m) = \sum_{j\in I} d_j(m), \qquad S_I = \sum_{j\in I} \xi_j,$$

where ξ_j, $0 \leqslant j \leqslant N$ are independent random variables (i.r.vs) defined on some probability space $\{\Omega, \mathcal{F}, P\}$ by

$$P\big(\xi_j = d\big) = 1/b_{j+1}, \qquad 0 \leqslant d < b_{j+1}$$

for $0 \leqslant j \leqslant N-1$ and

$$P\big(\xi_N = d\big) = 1/(a_N + 1), \qquad 0 \leqslant d \leqslant a_N.$$

Put $\bar{\xi}_j = \xi_j - \mathbf{E}\xi_j$, $0 \leqslant j \leqslant N$.

LEMMA. *There exist a probability space and the i.r.vs such that*

$$\nu_n\big((s_{I_1}(m), \dots, s_{I_p}(m)) \in B\big) = P\big((S_{I_1}, \dots, S_{I_p}) \in B\big) + \frac{2\Theta u_{r+1}}{n}, \tag{5}$$

where $|\Theta| \leqslant 1$, *and*

$$\nu_n\big((s_{J_1}(m), \dots, s_{J_p}(m)) \in B\big) \leqslant 2P\big((S_{J_1}, \dots, S_{J_p}) \in B\big) \tag{6}$$

uniformly in $B \subset \mathbf{R}^p$.

At first, we note that the assertions of Theorem for the processes $\bar{S}_k(m, .)$ and $\widehat{S}_k(m, .)$ are equivalent. This follows from the estimate

$$\nu_n\Big(\max_{x\leqslant k\leqslant N} \rho(\bar{S}_k(m, .), \widehat{S}_k(m, .)) \geqslant \varepsilon\Big)$$
$$\leqslant \nu_n\Big(\max_{x\leqslant k\leqslant N} \beta(k)^{-1} \max_{0\leqslant l\leqslant k} \big|\bar{d}_l(m)\big| \geqslant \varepsilon\Big) = o(1)$$

for each $\varepsilon > 0$ as $n \to \infty$ and later $x \to \infty$. In the last step we have used $d_l(m) \leqslant b_{l+1}$ and the condition (3) of Theorem.

Let in the sequel $r = N - K - 2$, $N = N(n) > K$ be sufficiently large, and $S_k^r(m, t)$ be defined from $\bar{S}_k(m, t)$ by adding the extra condition $j \leqslant r$ for the summation index j. Similarly, from the process

$$Y_k(t) := \frac{1}{\beta(k)} \sum_{\substack{j \geqslant 0 \\ B_j \leqslant t B_k}} \bar{\xi}_j, \qquad k \leqslant N,\ t \in [0, 1]$$

by adding the same bound for j we define $Y_k^r(t)$. Now we have to prove the relations

$$\nu_n(\varepsilon) := \nu_n\left(\max_{x \leqslant k \leqslant N} \rho(\bar{S}_k(m, .), S_k^r(m, .)) \geqslant \varepsilon \right) = o(1) \tag{7}$$

and

$$P(\varepsilon) := P\left(\max_{x \leqslant k \leqslant N} \rho(Y_k, Y_k^r) \geqslant \varepsilon \right) = o(1) \tag{8}$$

for each $\varepsilon > 0$ as $n \to \infty$ and $x \to \infty$.

The processes differ from their truncated versions in the interval $B_r/B_k \leqslant t \leqslant 1$ only. Thus from the definitions we obtain

$$\rho(\bar{S}_k(m, .), S_k^r(m, .)) \leqslant \max_{r \leqslant l \leqslant k} \beta(k)^{-1} \left| \sum_{r < j \leqslant l} \bar{d}_j(m) \right|.$$

Hence by (6), Kolmogorov's inequality, and (3),

$$\nu_n(\varepsilon) \leqslant 2P\left(\max_{r \leqslant l \leqslant N} \left| \sum_{r < j \leqslant l} \bar{\xi}_j \right| \geqslant \varepsilon\beta(r) \right) \leqslant \frac{2(B_N - B_r)}{\varepsilon^2 \beta(r)^2} = o_K\left(\frac{1}{LL\, B_N} \right)$$

as $n \to \infty$ and $x \to \infty$. Even more simple arguments and Kolmogorov's inequality yield (8).

The main probabilistic ingredient is the following assertion.

LEMMA 2. *If the condition* (3) *holds, then*

$$Y_k^r \Longrightarrow \mathcal{K} \qquad (P - a.s.).$$

Proof. The assertion of Lemma 2 for the processes Y_k follows from Major's [5] result. The passage to Y_k^r is provided by (8).

Further, we note that the quantities $\rho(S_k^r(m, .), g)$, $g \in \mathcal{K}$ are determined by m with the digits having indeces $0 \leqslant j \leqslant r$, and that the events $\{m\colon\ \rho(S_k^r(m, .), g) > \varepsilon\}$

or $\{m\colon\ \rho(S^r_k(m,.),g) < \varepsilon\}$ can be expressed by that considered in Lemma 1. Its assertion (5) and Lemma 2 imply

$$\nu_n\Big(\max_{x\leqslant k\leqslant r}\rho(S^r_k(m,.),\mathcal{K})\geqslant\varepsilon\Big) = P\Big(\max_{x\leqslant k\leqslant r}\rho(Y^r_k,\mathcal{K})\geqslant\varepsilon\Big)+O\big(2^{-K}\big)$$
$$= O\big(2^{-K}\big)$$

and

$$\nu_n\Big(\min_{x\leqslant k\leqslant r}\rho(S^r_k(m,.),g)<\varepsilon\Big) = P\Big(\min_{x\leqslant k\leqslant r}\rho(Y^r_k,g)<\varepsilon\Big)+O\big(2^{-K}\big)$$
$$= 1+O\big(2^{-K}\big)$$

for each $\varepsilon > 0$ and $g \in \mathcal{K}$ when $n \to \infty$ and $x \to \infty$. Since $K > 1$ is arbitry, Theorem follows from the estimate (7).

The approach just exposed allows to extend our Theorem to general U additive functions. Applying other forms of the iterated logarithm laws for sums of i.r.vs one can obtain their analogs for arithmetic functions related to systems of numerations.

REFERENCES

[1] H. Delange, Sur la fonction sommatoire de la fonction 'somme des chiffres', *Enseignement Math.*, **221** (1975), 31–47.

[2] M. P. Drazin and J. S. Griffiths, On the decimal representations of integers, *Proc. Camb. Phil. Soc.*, **48**(4) (1952), 555–565.

[3] A. S. Fraenkel, Systems of numeration, *American Math. Monthly*, **92**(2) (1985), 105–114.

[4] J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Translations of Math. Monographs, Vol. 11, Amer. Math. Soc., Providence, RI, 1964,

[5] P. Major, A note on Kolmogorov's law of iterated logarithm, *Studia Scient. Math. Hung.*, **12** (1977), 161–167.

[6] E. Manstavičius, Functional approach in the divisor distribution problems, *Acta Math. Hungar.*, **66**(3) (1995), 343–359.

[7] E. Manstavičius, Probabilistic theory of additive functions related to systems of numeration, *New Trends in Probability and Statistics*, Vol. 4, A. Laurinčikas et al. (Eds), VSP/TEV, Utrecht/Vilnius, 1997.

Skaitmenų sumos paklūsta Strassen'o dėsniui

E. Manstavičius

Darbe Strassen'o kartotinio logaritmo dėsnis įrodomas sveikųjų neneigiamų skaičių skaitmenų Kantoro skaičiavimo sistemoje sumoms. Naudojamasi [6] ir [7] straipsnių idėjomis.