ON ENGEL'S AND SYLVESTER'S SERIES

By

P. ERDŐS, A. RÉNYI and P. SZÜSZ

Mathematical Institute, Roland Eötvös University, Budapest

(Received October 28, 1957)

Introduction

Every real number x $(0<x<1)$ can be expanded into Engel's series (called also "Engel's series of the first kind", see PERRON [1])

(1) $$x=\frac{1}{q_1}+\frac{1}{q_1q_2}+\cdots+\frac{1}{q_1q_2\ldots q_n}+\cdots$$

where the integers $q_n=q_n(x)$ are defined as follows:

We denote by $\mathbf{T}_1x$ the transformation

(2) $$\mathbf{T}_1x=x\left\{\frac{1}{x}\right\}-1 \qquad (0<x<1).$$

(Here and in what follows $\{z\}$ denotes the least integer which is $\geqq z$.) We define a sequence $r_n(x)$ by the recursion

(3) $$r_0(x)=x,\ r_{n+1}(x)=\mathbf{T}_1r_n(x) \qquad (n=0,1,\ldots)$$

and put

(4) $$q_{n+1}=q_{n+1}(x)=\left\{\frac{1}{r_n(x)}\right\} \qquad (n=0,1,\ldots).$$

It is easy to see that

$$2\leqq q_n\leqq q_{n+1} \qquad (n=1,2,\ldots).$$

Evidently, if x is given by (1), we have

(5) $$r_n(x)=\frac{1}{q_{n+1}}+\frac{1}{q_{n+1}q_{n+2}}+\cdots.$$

If x is rational, $x=\frac{a}{b}$, then $\mathbf{T}_1x=\frac{a'}{b'}$ with $a'<a$. Thus we have for some ν $r_\nu(x)=0$. Thus every rational number $\frac{a}{b}$ has a finite representation

$$\frac{a}{b}=\frac{1}{q_1}+\frac{1}{q_2q_2}+\cdots+\frac{1}{q_2q_2,\ldots,q_\nu}.$$

If x is irrational, then $r_n(x)>0$ for all values of n. It is easy to see that for irrational values of x one has $\lim_{n\to\infty}q_n(x)=+\infty$.

In the present paper we investigate the metrical properties[1] of the sequence $q_n(x)$. The results obtained may be characterized as follows. Let us consider the interval (0, 1) as the space of elementary events, and interpret the Lebesgue measure of a measurable subset of the interval (0, 1) as its probability. Then the random variables $x_1 = \log q_1$, $x_n = \log \frac{q_n}{q_{n-1}}$ $(n = 2, 3, \ldots)$ are in a certain sense almost independent, and almost identically distributed, and thus for

(6) $$y_n = \log q_n = x_1 + x_2 + \cdots + x_n$$

similar results are valid as for the partial sums of a sequence of independent and identically distributed random variables, e. g. the central limit theorem, the laws of large numbers, the law of the iterated logarithm, etc.

In §§ 1—5 we deal with Engel's series: in § 1 some fundamental identities are deduced; in § 2 we prove the central limit theorem for the sums (6), i. e. we prove that the distribution of $\frac{\log q_n - n}{\sqrt{n}}$ tends for $n \to \infty$ to the normal distribution (Theorem 2). In § 3 we prove the strong law of large numbers for the sum (6), i. e. that for almost all x $\lim_{n\to\infty} \sqrt[n]{q_n} = e$ (Theorem 3). In § 4 we give some inequalities which are used in § 5 to prove the law of the iterated logarithm for the sums (6), i. e. that for almost all x $\overline{\lim_{n\to\infty}} \frac{\log q_n - n}{\sqrt{2n \log\log n}} = +1$ and $\underline{\lim_{n\to\infty}} \frac{\log q_n - n}{\sqrt{2n \log\log n}} = -1$. (Theorem 4). Theorems 2, 3, and 4 are not new. Theorem 3 has been stated without proof in a short note by É. BOREL[2] in 1947 ([3]; see also [4]. In the same year, in his paper [5] P. LÉVY announced Theorems 2 and 4. P. LÉVY sketched also the proof of these theorems, as well as that of Theorem 3. He pointed out that if x is uniformly distributed in the interval (0, 1) the random variables $\xi_{n+1} = -\log[(q_n(x)-1)r_n(x)]$ $(n = 1, 2, \ldots)$ are exactly exponentially distributed with mean 1, and they are also almost (but not exactly) independent. As ξ_{n+1} is with probability near to 1 very near to $x_{n+1} = \log \frac{q_{n+1}(x)}{q_n(x)}$ for large n the same holds for these quantities too, and this is the real ground — as pointed out above — of the validity of Theorems 2, 3 and 4.

[1] In a recent paper [2] one of the authors considered the metrical theory of a general class of representations of real numbers, but the representations by means of Engel's and Sylvester's series do not belong to the class of representations considered in [2]. They belong, however, to the class of representations considered by L. BERG [16].

[2] BOREL called Engel's series of the first kind "développement unitaire normal".

However, P. LÉVY did not go into details. It seems to us that — owing to the fact that the variables ξ_{n+1} are not exactly independent, these details (especially in case of the law of the iterated logarithm) would be rather cumbersome. Therefore we thought it worth while to work out detailed proofs of these theorems. We have chosen a way which is different from that of LÉVY, as we made ample use of the explicit formulae given in § 1 for the probability distribution of q_n, resp. the conditional probability distribution of q_{n+m}, when the value of q_m is fixed. Besides these formulae we utilised also the remark, made in § 1 that the random variables q_n form a Markov chain.

In § 6 we consider Sylvester's series (called also "Engel's series of the second kind", see the first edition of [1]

Sylvester's series[3] of a real number x $(0<x<1)$ is

(7) $$x=\frac{1}{Q_1}+\frac{1}{Q_2}+\cdots+\frac{1}{Q_n}+\cdots$$

where $Q_1, Q_2, \ldots$ are positive integers, defined as follows:

We denote by $\mathbf{T}_2 x$ the transformation

(8) $$\mathbf{T}_2 x = x - \frac{1}{\left\{\frac{1}{x}\right\}} \qquad (0<x<1).$$

We define the sequence $R_n(x)$ by the recursion

(9) $$R_0(x)=x,\ R_{n+1}(x)=\mathbf{T}_2 R_n(x) \qquad (n=0,1,\ldots)$$

and put

(10) $$Q^*_{n+1}=Q_{n+1}(x)=\left\{\frac{1}{R_n(x)}\right\} \qquad (n=0,1,\ldots).$$

It is easy to see that $Q_1\geqq 2$ and $Q_{n+1}\geqq Q_n(Q_n-1)+1$ $(n=1,2,\ldots)$.

Clearly, if x is rational, $x=\frac{a}{b}$, then $\mathbf{T}_2 x=\frac{a'}{b'}$ with $a'<a$; thus $R_\nu(x)=0$ for some ν and therefore every rational number $\frac{a}{b}$ has a finite representation $\frac{a}{b}=\frac{1}{Q_1}+\frac{1}{Q_2}+\cdots+\frac{1}{Q_\nu}$. For irrational values of x we have $\lim_{n\to\infty} Q_n(x)=+\infty$ and

(11) $$R_n(x)=\frac{1}{Q_{n+1}}+\frac{1}{Q_{n+2}}+\cdots \qquad (n=0,1,\ldots).$$

Putting $X_1=\log Q_1$, $X_n=\log\frac{Q_n}{Q_{n-1}^2}$ $(n=2,3,\ldots)$ we shall see that the random variables X_n are in a certain sense almost independent and

[3] SYLVESTER [6] called the expansion (7) a "sorites". See also [7] for further bibliography.

almost identically distributed. Thus we obtain (Theorem 5.) that the central limit theorem holds for

$$Y_n = \log \frac{Q_n}{Q_1 Q_2 \dots Q_{n-1}} = X_1 + X_2 + \dots + X_n. \tag{12}$$

As regards Q_n we shall prove (Theorem 6) that the limit $\lim_{n\to\infty} \frac{\log Q_n}{2^n}$ exists for almost all x, but its value may depend on x.

These results concerning Q_n are according to our knowledge new.

In § 7, some number-theoretic questions concerning Engel's and Sylvester's series are discussed, and some unsolved problems are mentioned.

§ 1. Fundamental identities for Engel's series

In what follows we shall interpret the Lebesgue measure of the set of those real numbers x $(0 < x < 1)$ for which some relation concerning the sequence $q_n = q_n(x)$ holds, as the probability of the relation in question, and shall denote it by $P(\dots)$, where in the bracket the relation in question will be indicated. The conditional probability of A with respect to the condition B will be denoted by $P(A|B)$.

As clearly $q_n = q_n(x) = k$ $(k = 2, 3, \dots)$ if and only if x is lying in some interval

$$\frac{1}{q_1} + \frac{1}{q_1 q_2} + \dots + \frac{1}{q_1 q_2 \dots q_{n-1} k} \leqq x < \frac{1}{q_1} + \frac{1}{q_1 q_2} + \dots + \frac{1}{q_1 q_2 \dots q_{n-1}(k-1)},$$

where $2 \leqq q_1 \leqq q_2 \leqq \dots \leqq q_{n-1} \leqq k$, and these intervals do not overlap, we have

$$p_n(k) = \mathbf{P}(q_n = k) = \frac{1}{k(k-1)} \sum_{2 \leqq q_1 \leqq q_2 \leqq \dots \leqq q_{n-1} \leqq k} \frac{1}{q_1 q_2 \dots q_{n-1}}. \tag{1.1}$$

Similarly we obtain that the conditional probability of the joint occurrence of $q_{n+r+1} = k_1, q_{n+r+2} = k_2, \dots, q_{n+r+s} = k_s$ under the condition that the values of $q_1, q_2, \dots, q_n$ are given, is

$$\mathbf{P}(q_{n+r+i} = k_i;\ 1 \leqq i \leqq s | q_1, \dots, q_n) = \tag{1.2}$$

$$= \frac{q_n - 1}{k_1 k_2 \dots k_s (k_s - 1)} \sum_{q_n \leqq q_{n+1} \leqq \dots \leqq q_{n+r} \leqq k_1} \frac{1}{q_{n+1} q_{n+2} \dots q_{n+r}}$$

if $2 \leqq q_1 \leqq q_2 \leqq \dots \leqq q_n \leqq k_1 \leqq k_2 \leqq \dots \leqq k_s$. (For $r = 0$ the empty sum is to be replaced by 1.)

As the conditional probability (1.2) does not depend on the values of $q_1, q_2, \dots, q_{n-1}$ (only on the value of q_n) and it does not depend on the number n either, *the sequence q_n, considered as a sequence of random vari-*

ables on the probability space furnished by the interval (0, 1) *the probability measure being the Lebesgue measure, is a homogeneous Markov chain.* The transition probabilities of this Markov chain can be obtained for $r=0$, $s=1$ from (1. 2) and are given by

$$\pi_{jk}=\mathbf{P}(q_{n+1}=k\,|\,q_n=j)=\frac{j-1}{k(k-1)}\qquad (k\geqq j\geqq 2). \tag{1.3}$$

It follows that the probabilities $p_n(k)$ can be obtained by the following recursion formulae:

$$\begin{aligned} p_1(k)&=\frac{1}{k(k-1)} && (k=2,3,\ldots),\\ p_n(k)&=\frac{1}{k(k-1)}\sum_{l=2}^{k}(l-1)p_{n-1}(l) && (k,n=2,3,\ldots). \end{aligned} \tag{1.4}$$

From (1. 1) we obtain

$$\sum_{n=1}^{\infty}p_n(k)x^{n-1}=\frac{1}{k(k-1)}\prod_{j=2}^{k}\frac{1}{\left(1-\frac{x}{j}\right)}. \tag{1.5}$$

Substituting $x=1$ into (1. 5) it follows that

$$\sum_{n=1}^{\infty}p_n(k)=\frac{1}{k-1}. \tag{1.6}$$

$\sum_{n=1}^{\infty}p_n(k)$ is clearly the mean value of the number of occurrences of the digit k in the sequence $q_1, q_2, \ldots, q_n, \ldots$. It is easy to determine also the probability ϱ_k that the number k occurs at least once in the sequence $q_1, q_2, \ldots, q_n, \ldots$. We have

$$\begin{aligned} \varrho_k&=\frac{1}{k(k-1)}\left(1+\sum_{n=2}^{\infty}\sum_{2\leqq q_1\leqq\ldots\leqq q_{n-1}\leqq k-1}\frac{1}{q_1q_2\ldots q_{n-1}}\right)=\\ &=\frac{1}{k(k-1)}\prod_{j=2}^{k-1}\frac{1}{1-\frac{1}{j}}=\frac{1}{k}. \end{aligned} \tag{1.7}$$

(In (1. 7) we considered the *first* occurrence of k in the sequence q_n, therefore we supposed $q_{n-1}\leqq k-1$ instead of $q_{n-1}\leqq k$.)

We may calculate similarly the probability $\varrho_k(r)$ that the digit k occurs exactly r times $(r=0, 1, \ldots)$ in the sequence q_n. We obtain

THEOREM 1. *The probability that the digit k occurs exactly r times in the sequence $q_n(x)$ is given by*

$$\varrho_k(r)=\frac{k-1}{k^{r+1}}\qquad (r=0,1,\ldots;\ k=2,3,\ldots). \tag{1.8}$$

Using (1.5) we may obtain an explicit formula for $p_n(k)$. Taking into account that

$$\prod_{j=2}^{k} \frac{1}{1-\frac{x}{j}} = \sum_{j=2}^{k} k(k-1)\binom{k-2}{j-2}\frac{(-1)^{j-2}}{j-x}, \tag{1.9}$$

we obtain

$$p_n(k) = \sum_{j=2}^{k} \binom{k-2}{j-2}\frac{(-1)^{j-2}}{j^n}. \tag{1.10}$$

As

$$\frac{1}{j^n} = \frac{1}{(n-1)!}\int_0^\infty e^{-uj} u^{n-1} du, \tag{1.11}$$

it follows from (1.10) that

$$p_n(k) = \frac{1}{(n-1)!}\int_0^\infty u^{n-1} e^{-2u}(1-e^{-u})^{k-2} du. \tag{1.12}$$

Putting

$$W_n(k) = \sum_{l>k} p_n(l) \tag{1.13}$$

and

$$S_n(k) = \sum_{l\le k} p_n(l), \tag{1.14}$$

we obtain from (1.12)

$$W_n(k) = \frac{1}{(n-1)!}\int_0^\infty u^{n-1} e^{-u}(1-e^{-u})^{[k]-1} du \tag{1.15}$$

and

$$S_n(k) = \frac{1}{(n-1)!}\int_0^\infty u^{n-1} e^{-u}[1-(1-e^{-u})^{[k]-1}] du. \tag{1.16}$$

Similar formulae can be found for the conditional probabilities

$$\begin{aligned} p_n(k|j) &= \mathbf{P}(q_{m+n}=k|q_m=j)= \\ &= \frac{j-1}{k(k-1)} \sum_{j\le l_1\le l_2\le\cdots\le l_{n-1}\le k} \frac{1}{l_1 l_2 \ldots l_{n-1}} \end{aligned} \tag{1.17}$$

($p_n(k|j)$ does not depend on m according to the homogeneity of the Markov chain q_n).

We obtain

$$\sum_{n=1}^{\infty} p_n(k|j) x^{n-1} = \frac{j-1}{k(k-1)} \prod_{h=j}^{k} \frac{1}{\left(1-\frac{x}{h}\right)}. \tag{1.18}$$

As

$$\prod_{h=j}^{k} \frac{1}{\left(1-\frac{x}{h}\right)} = j\binom{k}{j}\sum_{h=j}^{k}\binom{k-j}{h-j}\frac{(-1)^{h-j}}{h-x}, \tag{1.19}$$

we obtain

$$p_n(k|j) = \binom{k-2}{j-2}\sum_{h=j}^{k}\binom{k-j}{h-j}\frac{(-1)^{h-j}}{h^n}. \tag{1.20}$$

It follows by (1.11) that

$$p_n(k|j) = \frac{1}{(n-1)!}\binom{k-2}{j-2}\int_0^\infty u^{n-1}e^{-ju}(1-e^{-u})^{k-j}\,du \tag{1.21}$$

(we have evidently $p_n(k|2) = p_n(k)$; therefore putting $j=2$ (1.20) resp. (1.21) reduce to (1.10) resp. (1.12)).

Here and in what follows we shall denote by $\mathbf{M}(\zeta)$ the mean value of ζ, i. e. we put $\mathbf{M}(\zeta) = \int_0^1 \zeta(x)\,dx$ for $\zeta = \zeta(x)$. We shall further denote by $\mathbf{M}(\zeta|B)$ the conditional mean value of the random variable ζ with respect to the condition B. We shall now prove

LEMMA 1. $\mathbf{M}(\log q_n) = n - \gamma + o(1)$ *where* γ *is Euler's constant.*[4]

PROOF. Let us consider $\zeta_n = 1 + \frac{1}{2} + \cdots + \frac{1}{q_n - 2}$. We have by (1.12)

$$\mathbf{M}(\zeta_n) = \frac{1}{(n-1)!}\int_0^\infty u^{n-1}e^{-2u}\left(\sum_{i=1}^{\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{i}\right)(1-e^{-u})^i\right)du = n. \tag{1.22}$$

As q_n tends with probability 1 to $+\infty$, and as well known,

$$\log N = 1 + \frac{1}{2} + \cdots + \frac{1}{N-2} - \gamma + o(1),$$

it follows that

$$\mathbf{P}(\lim_{n\to\infty}(\log q_n - \zeta_n) = -\gamma) = 1. \tag{1.23}$$

(1.22) and (1.23) together prove Lemma 1.

[4] It is clear from (1.12) that $\mathbf{M}(q_n) = +\infty$.

§ 2. The central limit theorem for Engel's series

In this § we shall need the following

LEMMA 2. $\sum_{k=2}^{\infty} \frac{p_n(k)}{k+1} \leqq \left(\frac{3}{4}\right)^n \quad (n=1,2,\ldots).$

PROOF. Taking into account that by virtue of (1.4)

$$\sum_{k=2}^{\infty} \frac{p_n(k)}{k+1} = \sum_{l=2}^{\infty} \frac{p_{n-1}(l)}{l+1} \cdot \left(\frac{l+1}{2l}\right) \quad \text{and} \quad \frac{l+1}{2l} \leqq \frac{3}{4} \quad \text{for} \quad l \geqq 2,$$

we obtain that

$$\sum_{k=2}^{\infty} \frac{p_n(k)}{k+1} \leqq \frac{3}{4} \sum_{l=2}^{\infty} \frac{p_{n-1}(l)}{l+1}.$$

As $\sum_{k=2}^{\infty} \frac{p_1(k)}{k+1} \leqq \frac{3}{4}$, our Lemma follows. Now we can prove

THEOREM 2 (P. LÉVY). *For any real y we have*

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\log q_n - n}{\sqrt{n}} < y\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{t^2}{2}} dt.$$

By other words $\frac{\log q_n - n}{\sqrt{n}}$ *is in the limit normally distributed for* $n \to \infty$.

PROOF. Let us put

(2. 1) $$x_1 = \log q_1, \; x_n = \log \frac{q_n}{q_{n-1}} \qquad (n=2,3,\ldots)$$

and

(2. 2) $$y_n = \log q_n = x_1 + x_2 + \cdots + x_n,$$

further

(2. 3) $$\varphi_n(t) = \mathbf{M}(e^{ity_n}) = \int_0^1 e^{it\log q_n(x)} dx.$$

To prove Theorem 2 it suffices to show (see e. g. [8]) that

(2. 4) $$\lim_{n\to\infty} \varphi_n\left(\frac{t}{\sqrt{n}}\right) e^{-it\sqrt{n}} = e^{-\frac{t^2}{2}}$$

for any real t. Let us put

(2. 5) $$\psi(t) = \int_1^{\infty} \frac{e^{it\log x}}{x^2} dx = \frac{1}{1-it}.$$

First we shall show that

(2. 6) $$|\varphi_n(t) - \psi(t)\varphi_{n-1}(t)| \leqq (3+2|t|)\left(\frac{3}{4}\right)^{n-1} \qquad (n=2,3,\ldots).$$

This can be obtained as follows. We have

$$\varphi_n(t)=\sum_{l=2}^{\infty}p_{n-1}(l)\mathbf{M}(e^{it\log q_n}|q_{n-1}=l) \tag{2.7}$$

and thus by (1.3)

$$\varphi_n(t)=\sum_{l=2}^{\infty}p_{n-1}(l)e^{it\log l}\left(\sum_{k=l}^{\infty}e^{it\log\frac{k}{l}}\frac{l-1}{k(k-1)}\right). \tag{2.8}$$

As

$$\left|\sum_{k=l}^{\infty}e^{it\log\frac{k}{l}}\cdot\frac{(l-1)}{k(k-1)}-\psi(t)\right|\leqq\frac{3}{l+1}+\sum_{k=l+1}^{\infty}\int_{\frac{k-1}{l}}^{\frac{k}{l}}\left|e^{it\log\frac{k}{l}}-e^{it\log x}\right|\frac{dx}{x^2} \tag{2.9}$$

and as for $\frac{k-1}{l}\leqq x\leqq\frac{k}{l}$ and $k\geqq l+1$ we have

$$\left|e^{it\log\frac{k}{l}}-e^{it\log x}\right|\leqq\frac{|t|}{k-1}\leqq\frac{2|t|}{l+1},$$

it follows that

$$\left|\sum_{k=l}^{\infty}e^{it\log\frac{k}{l}}\frac{(l-1)}{k(k-1)}-\psi(t)\right|\leqq\frac{3+2|t|}{l+1}. \tag{2.10}$$

Thus, taking into account that

$$\varphi_{n-1}(t)=\sum_{l=2}^{\infty}p_{n-1}(l)e^{it\log l}, \tag{2.11}$$

it follows from (2.8) and (2.10) that

$$|\varphi_n(t)-\psi(t)\varphi_{n-1}(t)|\leqq(3+2|t|)\sum_{l=2}^{\infty}\frac{p_{n-1}(l)}{l+1}. \tag{2.12}$$

Applying Lemma 2 we obtain (2.6). Let us apply (2.6) for $n-r$ instead of n and multiply it by $\psi^r(t)$. It follows

$$|\psi^r(t)\varphi_{n-r}(t)-\psi^{r+1}(t)\varphi_{n-r-1}(t)|\leqq(3+2|t|)\left(\frac{3}{4}\right)^{n-r-1} \tag{2.13}$$

Adding (2.13) for $r=0, 1, \ldots, n-m-1$ we obtain

$$|\varphi_n(t)-\psi^{n-m}(t)\varphi_m(t)|\leqq(12+8|t|)\left(\frac{3}{4}\right)^m. \tag{2.14}$$

Now we have clearly for any fixed value of m

$$\lim_{n\to\infty}\psi\left(\frac{t}{\sqrt{n}}\right)^{n-m}e^{-it\sqrt{n}}=e^{-t^2/2} \tag{2.15}$$

and

$$\lim_{n\to\infty}\varphi_m\left(\frac{t}{\sqrt{n}}\right)=1 \tag{2.16}$$

and therefore

(2. 17) $$\limsup_{n\to\infty} \left| \varphi_n\left(\frac{t}{\sqrt{n}}\right) e^{-it\sqrt{n}} - e^{-t^2/2} \right| \leqq 12\left(\frac{3}{4}\right)^m.$$

As we may choose the value of m arbitrarily large, we obtain (2. 4). Thus Theorem 1 is proved.

It can be seen from the proof that the random variables x_n behave approximately as if they were independent and distributed according to the exponential distribution with mean 1. The latter assertion can be expressed also by saying that $\frac{q_{n-1}}{q_n}$ is for $n \to \infty$ in the limit uniformly distributed in the interval (0, 1). This result can be deduced also directly from the remark of P. LÉVY mentioned in the introduction, that the random variable $(q_n(x)-1)r_n(x)$ is exactly uniformly distributed in the interval (0, 1).

§ 3. The strong law of large numbers for Engel's series

In this § we give a short proof of the following Theorem 3, which has been announced without proof by É. BOREL [3]. Though Theorem 3 is contained in Theorem 4 (the law of the iterated logarithm), we thought it worth while to give a direct proof of Theorem 3 because the proof of Theorem 4 is rather complicated.

THEOREM 3. (É. BOREL). *For almost all* x $\lim_{n\to\infty} \sqrt[n]{q_n} = e$.

PROOF. Let us choose an arbitrary $\varepsilon > 0$. We start from the formula (1. 15). We have evidently for $\tau > 0$

(3. 1) $$W_n(k) \leqq e^{-e^{\tau}} + \frac{1}{(n-1)!} \int_{\log(k-1)-\tau}^{\infty} u^{n-1} e^{-u}\, du.$$

Thus if $k = e^{n(1+\varepsilon)}$, $\tau = \log\log k$ and $n \geqq n_1(\varepsilon)$ we obtain

$$W_n(k) \leqq \frac{1}{k} + \frac{1}{(n-1)!} \int_{n\left(1+\frac{\varepsilon}{2}\right)}^{\infty} u^{n-1} e^{-u}\, du.$$

As $\frac{1}{(n-1)!} < \left(\frac{e}{n-1}\right)^{n-1}$, and ve^{1-v} is decreasing and < 1 for $v > 1$, it follows

$$\frac{1}{(n-1)!} \int_{n\left(1+\frac{\varepsilon}{2}\right)}^{\infty} u^{n-1} e^{-u}\, du \leqq n\left(\left(1+\frac{\varepsilon}{2}\right) e^{-\frac{\varepsilon}{2}}\right)^{n-2}.$$

Thus we obtain

LEMMA 3. *There exists for any* $\varepsilon > 0$ *a number* $q_1(\varepsilon)$ *for which* $0 < q_1(\varepsilon) < 1$ *and*

$$W_n(e^{n(1+\varepsilon)}) \leqq (q_1(\varepsilon))^n \quad \text{for} \quad n \geqq n_1(\varepsilon). \tag{3.2}$$

Clearly for $q_1(\varepsilon)$ we may take any number satisfying

$$\max\left(\frac{1}{e}, \left(1+\frac{\varepsilon}{2}\right)e^{-\varepsilon/2}\right) < q_1(\varepsilon) < 1.$$

Similarly we prove

LEMMA 4. *To any* ε $(0 < \varepsilon < 1)$ *there exists a number* $q_2(\varepsilon)$ *such that* $0 < q_2(\varepsilon) < 1$ *and*

$$S_n(e^{n(1-\varepsilon)}) < q_2^n(\varepsilon) \quad \text{for} \quad n \geqq n_2(\varepsilon). \tag{3.3}$$

As a matter of fact we have by (1.16) for $\tau > 0$

$$S_n(k) \leqq \frac{1}{(n-1)!} \int_0^{\log k+\tau} u^{n-1} e^{-u} du + e^{-\tau}. \tag{3.4}$$

Choosing $\tau = \frac{n\varepsilon}{2}$, as ve^{1-v} is increasing and < 1 for $0 < v < 1$ it follows that

$$S_n(e^{n(1-\varepsilon)}) \leqq 2en\left(\left(1-\frac{\varepsilon}{2}\right)e^{\varepsilon/2}\right)^{n-1} \quad \text{if} \quad n \geqq n_2(\varepsilon),$$

which proves (3.3).

It follows from Lemma 3 resp. Lemma 4 that the series

$$\sum_{n=1}^{\infty} \mathbf{P}(\sqrt[n]{q_n} > e^{1+\varepsilon}) \tag{3.5}$$

resp. the series

$$\sum_{n=1}^{\infty} \mathbf{P}(\sqrt[n]{q_n} < e^{1-\varepsilon}) \tag{3.6}$$

are convergent for any $\varepsilon > 0$. Therefore for every $\varepsilon > 0$ for almost all x the inequalities

$$e^{1-\varepsilon} < \sqrt[n]{q_n} < e^{1+\varepsilon} \tag{3.7}$$

are valid, except for a finite number of values of n. This proves Theorem 3.

§ 4. Some inequalities

In what follows $c_1, c_2, \ldots$ denote positive absolute constants.

LEMMA 5.

$$p_n(k|j) \geqq \frac{\left(\sum_{l=j}^{k} \frac{1}{l}\right)^{n-1} (j-1)}{k(k-1)(n-1)!} \quad \textit{for} \quad k \geqq j \geqq 2. \tag{4. 1}$$

PROOF. We have

$$\frac{1}{(n-1)!}\left(\sum_{l=j}^{k} \frac{1}{l}\right)^{n-1} = \sum_{\alpha_j+\cdots+\alpha_k=n-1} \frac{1}{\alpha_j!\ldots\alpha_k!} \frac{1}{j^{\alpha_j}\ldots k^{\alpha_k}} \leqq \tag{4. 2}$$

$$\leqq \sum_{j \leqq l_1 \leqq \cdots \leqq l_{n-1} \leqq k} \frac{1}{l_1 l_2 \ldots l_{n-1}}.$$

With respect to (1. 17) this proves Lemma 5.

LEMMA 6. *If* $k = e^{n+x\sqrt{n}}$ *where* $0 < x < n^{1/7}$ *we have*

$$\frac{c_1 e^{-\frac{x^2}{2}}}{x} \leqq W_n(k) \leqq \frac{c_2 e^{-\frac{x^2}{2}}}{x}. \tag{4. 3}$$

LEMMA 7. *If* $k = e^{n-x\sqrt{n}}$ *where* $0 < x < n^{1/7}$, *we have*

$$\frac{c_3 e^{-\frac{x^2}{2}}}{x} \leqq S_n(k) \leqq \frac{c_4 e^{-\frac{x^2}{2}}}{x}. \tag{4. 4}$$

PROOF OF LEMMAS 6 AND 7. We have from (1. 15) resp. (1. 16)

$$c_5 \frac{1}{(n-1)!} \int_{\log k}^{\infty} u^{n-1} e^{-u} du \leqq W_n(k), \tag{4. 5}$$

resp.

$$c_6 \frac{1}{(n-1)!} \int_{0}^{\log k} u^{n-1} e^{-u} du \leqq S_n(k). \tag{4. 6}$$

From (3. 1), (3. 4), (4. 5) and (4. 6) the assertions of Lemmas 6 and 7 follow easily taking into account that by the method of Laplace we obtain

$$\frac{1}{(n-1)!} \int_{n+x\sqrt{n}}^{\infty} u^{n-1} e^{-u} du \sim \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-u^2/2} du \tag{4. 7}$$

for $n \to \infty$ and $0 < x < n^{1/7}$, further

$$\frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-u^2/2} du \sim \frac{1}{\sqrt{2\pi}\, x} e^{-x^2/2} \quad \text{for} \quad x \to +\infty. \tag{4. 8}$$

LEMMA 8. *For* $\left|\log \frac{k}{j} - n\right| \leqq \sqrt{n \log n}$

$$W_n(k|j) \geqq c_7 W_n\left(\frac{k}{j}\right). \tag{4.9}$$

LEMMA 9. *For* $\left|\log \frac{k}{j} - n\right| < \sqrt{n \log n}$

$$S_n(k|j) \geqq c_8 S_n\left(\frac{k}{j}\right). \tag{4.10}$$

PROOF OF LEMMAS 8 AND 9. By Lemma 5

$$W_n(k|j) \geqq \frac{(j-1)}{(n-1)!} \sum_{h=k}^{\infty} \frac{\left(\sum_{l=j}^{h} \frac{1}{l}\right)^{n-1}}{h(h-1)} \geqq \frac{c_9}{(n-1)!} \int_{\frac{k}{j}}^{\infty} \frac{(\log x)^{n-1}}{x^2} dx = \tag{4.11}$$

$$= \frac{c_9}{(n-1)!} \int_{\log \frac{k}{j}}^{\infty} u^{n-1} e^{-u} du$$

and thus by (3. 1) the assertion of Lemma 9 follows. Similarly we obtain

$$S_n(k|j) \geqq \frac{c_{10}}{(n-1)!} \int_{0}^{\log \frac{k}{j}} u^{n-1} e^{-u} du, \tag{4.12}$$

from which by (3. 4) the assertion of Lemma 9 follows.

The asymptotic behaviour of $p_n(k)$ resp. $p_n(k|j)$ has been considered more thoroughly by A. BÉKÉSSY [9]. He proved e. g. that

$$p_n(k) \sim \Gamma\left(2 - \frac{n-1}{\log k}\right) \frac{(\log k)^{n-1}}{k^2(n-1)!} \tag{4.13}$$

for $\frac{n-1}{2 \log k} \leqq q < 1$. For our purposes, however, the estimates given in this § are sufficient.

§ 5. The law of the iterated logarithm

Now we are in the position to prove

THEOREM 4. (P. LÉVY). *For almost all* x $(0 < x < 1)$ *we have*

$$\limsup_{n \to +\infty} \frac{\log q_n - n}{\sqrt{2n \log \log n}} = 1 \tag{5.1}$$

and

$$\liminf_{n\to+\infty} \frac{\log q_n - n}{\sqrt{2n \log\log n}} = -1. \tag{5.2}$$

PROOF. We prove first that for almost all x

$$\limsup_{n\to+\infty} \frac{\log q_n - n}{\sqrt{2n \log\log n}} \leqq 1 \tag{5.3}$$

and

$$\liminf_{n\to+\infty} \frac{\log q_n - n}{\sqrt{2n \log\log n}} \geqq -1. \tag{5.4}$$

For this purpose it suffices to show that the inequalities

$$\frac{\log q_n - n}{\sqrt{2n \log\log n}} > 1 + \delta \tag{5.5}$$

resp.

$$\frac{\log q_n - n}{\sqrt{2n \log\log n}} < -1 - \delta \tag{5.6}$$

are satisfied for almost all x for a finite number of values of n only, for any $\delta > 0$. The proof of this given below follows essentially that of the paper [10] for the ordinary law of the iterated logarithm.[5] Let us put $m_n = [(1+\varepsilon)^n]$ where $0 < \varepsilon < \frac{\delta}{2}$. It follows by Lemmas 6 and 7 that the series

$$\sum_{n=1}^{\infty} \mathbf{P}\left(\frac{\log q_{m_n} - m_n}{\sqrt{2m_n \log\log m_n}} > 1 + \eta\right) \tag{5.7}$$

resp.

$$\sum_{n=1}^{\infty} \mathbf{P}\left(\frac{\log q_{m_n} - m_n}{\sqrt{2m_n \log\log m_n}} < -1 - \eta\right) \tag{5.8}$$

are convergent if $\eta > 0$.

We shall prove in detail only the assertion concerning the inequality (5.5) as the proof for (5.6) is exactly the same. Let us denote by $A_k(\delta)$ the event $\frac{\log q_k - k}{\sqrt{2k \log\log k}} > 1 + \delta$. Let us denote for any event A by $\overline{A}$ the event contrary to A; put $B_{m_n}^{(n)} = A_{m_n}(\delta)$, further

$$B_k^{(n)} = \overline{A_{m_n}(\delta)} \ldots \overline{A_{k-1}(\delta)} \cdot A_k(\delta) \quad \text{for} \quad k > m_n.$$

[5] As, however, we do not consider here for an arbitrary function $\varphi(n)$ whether the inequality $\log q_n - n > \sqrt{n\varphi(n)}$ is satisfied for almost all x for a finite or an infinite number of values of n, but restrict ourselves to the case $\varphi(n) = 2(1 \pm \delta) \log\log n$, we may take $m_n = [(1+\varepsilon)^n]$ instead of $m_n = [e^{\frac{n}{\log n}}]$ needed in the general case.

(Here and in what follows the product of events denotes the joint occurrence of the events in question.) The events $B_k^{(n)}$ $(k \geqq m_n)$ clearly exclude each other, as $B_k^{(n)}$ means that k is the first index $\geqq m_n$ for which $A_k(\delta)$ takes place. If the sum of events means that at least one of the events occurs, then we have evidently for any $l \geqq m_n$

$$\sum_{k=m_n}^{l} B_k^{(n)} = \sum_{k=m_n}^{l} A_k(\delta).$$

If our assertion concerning (5.5) would not hold, i. e. if (5.5) would be satisfied for an infinity of values of n for all x belonging to a set having positive measure, we could find a constant $c > 0$ and for any positive integer M an other integer $N > M$ such that

$$\mathbf{P}\left(\sum_{k=m_M}^{m_N-1} A_k(\delta)\right) \geqq c > 0$$

(c does not depend on M). Let us denote for $m_n \leqq k < m_{n+1}$ by $D_{n,k}$ the event that

$$\log q_{m_{n+2}} - m_{n+2} \geqq \log q_k - k.$$

Clearly the joint occurrence of the events $B_k^{(n)}$ and $D_{n,k}$ implies the occurrence of $A_{m_{n+2}}(\eta)$ if $(1+\delta) > (1+\eta)(1+\varepsilon)$.

If $0 < \delta < 1$ and we choose $0 < \varepsilon < \frac{\delta}{2}$ and $0 < \eta < \frac{\delta}{4}$, then this condition is clearly satisfied and we obtain

$$(5.9) \qquad \mathbf{P}(A_{m_{n+2}}(\eta)) \geqq \sum_{k=m_n}^{m_{n+1}-1} \mathbf{P}(D_{n,k} B_k^{(n)}) = \sum_{k=m_n}^{m_{n+1}-1} \mathbf{P}(B_k^{(n)})\mathbf{P}(D_{n,k} | B_k^{(n)}).$$

But, as q_k is a Markov chain,

$$(5.10) \qquad \mathbf{P}(D_{n,k} | B_k^{(n)}) \geqq \operatorname*{Min}_{\log j > k + (1+\delta)\sqrt{2k \log\log k}} \mathbf{P}(D_{n,k} | q_k = j)$$

and by Lemma 8

$$(5.11) \qquad \mathbf{P}(D_{n,k} | q_k = j) \geqq c_7 W_{m_{n+2}-k}(e^{m_{n+2}-k}).$$

Thus

$$(5.12) \qquad \mathbf{P}(D_{n,k} | B_k^{(n)}) \geqq c_{11} > 0$$

and therefore by (5.9)

$$(5.13) \qquad \sum_{n=M}^{N-1} \mathbf{P}(A_{m_{n+2}}(\eta)) \geqq c_{11} \sum_{k=m_M}^{m_N-1} \mathbf{P}(B_k^{(n)}) = c_{11} \mathbf{P}\left(\sum_{k=m_M}^{m_N-1} A_k(\delta)\right).$$

It follows that for any M there can be found a number N such that

$$(5.14) \qquad \sum_{n=M}^{N-1} \mathbf{P}(A_{m_{n+2}}(\eta)) \geqq c c_{11} > 0.$$

But this is a contradiction because the series (5. 7) is convergent. Thus our assertion that for almost all x (5. 5) is satisfied only for a finite number of values of n is proved.

The corresponding statement for (5.6) is proved similarly. The only difference consists in that we need here Lemma 9 instead of Lemma 8.

Now we turn to the proof of the other part of the theorem, i. e. we prove that

$$\frac{\log q_n - n}{\sqrt{2n \log\log n}} > 1-\delta \tag{5.15}$$

resp.

$$\frac{\log q_n - n}{\sqrt{2n \log\log n}} < -1+\delta \tag{5.16}$$

are both satisfied for any infinity of values of n, for almost all x, if $\delta > 0$, which implies

$$\overline{\lim_{n\to\infty}} \frac{\log q_n - n}{\sqrt{2n \log\log n}} \geqq 1 \tag{5.17}$$

resp.

$$\underline{\lim_{n\to\infty}} \frac{\log q_n - n}{\sqrt{2n \log\log n}} \leqq -1 \tag{5.18}$$

for almost all x.

We choose a $q>1$, the value of which will be fixed later, and put $m_n = [q^n]$. We denote by $C_n(\delta)$ the event that the inequality

$$-1-\delta < \frac{\log q_n - n}{\sqrt{2n \log\log n}} < 1-\delta \tag{5.19}$$

does not hold. Let us consider the probability

$$\mathbf{P}\left(\prod_{k=M}^{N} \overline{C_{m_k}(\delta)}\right) = \mathbf{P}(\overline{C_{m_M}(\delta)}) \prod_{n=M}^{N-1} \mathbf{P}\left(\overline{C_{m_{n+1}}(\delta)} \,\middle|\, \prod_{k=M}^{n} \overline{C_{m_k}(\delta)}\right). \tag{5.20}$$

If we can prove that this probability may be made arbitrarily small for any M by choosing N sufficiently large, this implies that the measure of the set of those x for which

$$-1-\delta < \frac{\log q_{m_n} - m_n}{\sqrt{2m_{n+1} \log\log m_n}} < 1-\delta \quad \text{for} \quad n \geqq M, \tag{5.21}$$

is equal to 0 for any M.

As we already know that the set of those x for which

$$\frac{\log q_{m_n} - m_n}{\sqrt{2m_n \log\log m_n}} < -1-\delta$$

for $n \geqq M$ has the measure 0, it follows that the set for which

$$\frac{\log q_{m_n} - m_n}{\sqrt{2m_n \log\log m_n}} < 1-\delta$$

for $n \geqq M$ has also the measure 0, and this is what we want to prove.

Now we have clearly, as q_n is a Markov chain,

$$\mathbf{P}\left(\overline{C_{m_{n+1}}(\delta)} \,\middle|\, \prod_{k=M}^{n} \overline{C_{m_k}(\delta)}\right) \leqq \lambda_{n+1}, \tag{5.22}$$

where

$$\lambda_{n+1} = \operatorname*{Max}_{-1-\delta < \frac{\log j_n - m_n}{\sqrt{2m_n \log\log m_n}} < 1-\delta} \mathbf{P}(\overline{C_{m_{n+1}}(\delta)} \mid q_{m_n} = j_n). \tag{5.23}$$

We now try to obtain an estimate from above for λ_{n+1}. We have clearly, putting

$$k_n = e^{m_{n+1} + (1-\delta)\sqrt{2m_{n+1} \log\log m_{n+1}}}, \tag{5.24}$$

$$\mathbf{P}(C_{m_{n+1}}(\delta) \mid q_{m_n} = j_n) \geqq W_{m_{n+1}-m_n}(k_n \mid j_n). \tag{5.25}$$

We shall give an estimate from below for the right hand side of (5.25). By Lemma 8 we obtain, putting $v_n = m_{n+1} - m_n$

$$\operatorname*{Min}_{-1-\delta < \frac{\log j_n - m_n}{\sqrt{2m_n \log\log m_n}} < 1-\delta} W_{v_n}(k_n \mid j_n) \geqq c_7 W_{v_n}\left(\frac{k_n}{e^{m_n - (1+\delta)\sqrt{2m_n \log\log m_n}}}\right). \tag{5.26}$$

If q is sufficiently large, then

$$\log k_n - m_n + (1+\delta)\sqrt{2m_n \log\log m_n} < v_n + \left(1 - \frac{\delta}{2}\right)\sqrt{2v_n \log\log v_n}$$

and in this case

$$\lambda_{n+1} \leqq 1 - c_7 W_{v_n}\left(e^{v_n + \left(1-\frac{\delta}{2}\right)\sqrt{2v_n \log\log v_n}}\right). \tag{5.27}$$

As it follows by some easy calculation from Lemma 6 that the series

$$\sum_{n=1}^{\infty} W_{v_n}\left(e^{v_n + \left(1-\frac{\delta}{2}\right)\sqrt{2v_n \log\log v_n}}\right)$$

is divergent, our assertion that

$$\lim_{N\to\infty} \mathbf{P}\left(\prod_{k=M}^{N} \overline{C_{m_k}(\delta)}\right) = 0$$

follows. This proves (5.17); clearly (5.18) can be proved in the same way. Thus Theorem 4 is completely proved.

We do not consider here the obvious generalizations of Theorem 4, though they may be treated in the same way (see footnote [5]).

§ 6. Sylvester's series

In this § we consider Sylvester's series (Engel's series of the second kind) for a real number x $(0 < x < 1)$

$$x = \frac{1}{Q_1} + \frac{1}{Q_2} + \cdots + \frac{1}{Q_n} + \cdots. \tag{6.1}$$

We investigate some metrical properties of the denominators $Q_n = Q_n(x)$. We have clearly

$$Q_{n+1} \geqq Q_n(Q_n - 1) + 1$$

and

$$\mathbf{P}(Q_1 = k_1, \ldots, Q_n = k_n) = \frac{1}{k_n(k_n - 1)}, \tag{6.3}$$

provided that $k_1 \geqq 2$ and $k_{i+1} \geqq k_i(k_i - 1) + 1$ $(i = 1, 2, \ldots, n-1)$. Thus if these inequalities are satisfied,

$$\begin{aligned} &\mathbf{P}(Q_n = k_n \mid Q_1 = k_1, \ldots, Q_{n-1} = k_{n-1}) = \\ &= \mathbf{P}(Q_n = k_n \mid Q_{n-1} = k_{n-1}) = \frac{k_{n-1}(k_{n-1} - 1)}{k_n(k_n - 1)}. \end{aligned} \tag{6.4}$$

Thus *the sequence $Q_n(x)$ is a homogeneous Markov chain,* whose transition probabilities are given by

$$\Pi_{jk} = \mathbf{P}(Q_n = k \mid Q_{n-1} = j) = \frac{j(j-1)}{k(k-1)} \tag{6.5}$$

if $j \geqq 2$ and $k \geqq j(j-1) + 1$.

It follows that putting $P_n(k) = \mathbf{P}(Q_n = k)$ we have the recursion relations

$$P_n(k) = \sum_{j(j-1)+1 \leqq k} \frac{j(j-1)}{k(k-1)} P_{n-1}(j). \tag{6.6}$$

Therefore

$$\sum_{k=2}^{\infty} \frac{P_n(k)}{k} = \sum_{j=2}^{\infty} P_{n-1}(j) \sum_{k=j(j-1)+1}^{\infty} \frac{j(j-1)}{k^2(k-1)} \leqq \sum_{j=2}^{\infty} \frac{P_{n-1}(j)}{j(j-1)+1} \tag{6.7}$$

and as $\dfrac{j}{j(j-1)+1} \leqq \dfrac{2}{3}$ for $j \geqq 2$, it follows

$$\sum_{k=2}^{\infty} \frac{P_n(k)}{k} \leqq \frac{2}{3} \sum_{k=2}^{\infty} \frac{P_{n-1}(k)}{k}. \tag{6.8}$$

Therefore as $P_1(k) = \dfrac{1}{k(k-1)}$, we obtain

LEMMA 10.

$$\sum_{k=2}^{\infty} \frac{P_n(k)}{k} \leqq \left(\frac{2}{3}\right)^n.$$

Let us put now

(6.9) $\Phi_1(t) = M(e^{it\log Q_1})$ and $\Phi_n(t) = \mathbf{M}\left(e^{it\log\frac{Q_n}{Q_1\cdots Q_{n-1}}}\right)$ $(n = 2, 3, \ldots)$.

By (6.3), putting

$$\psi_j(t) = \sum_{k \geqq j(j-1)+1} \frac{e^{it\log\frac{k}{j^2}}}{k(k-1)} j(j-1), \tag{6.10}$$

we have

$$\Phi_n(t) = \sum \frac{e^{it\log\frac{k_{n-1}}{k_1k_2\ldots k_{n-2}}}}{k_{n-1}(k_{n-1}-1)} \psi_{k_{n-1}}(t). \tag{6.11}$$

We obtain easily, similarly as in § 2

$$\left|\psi_j(t) - \int_1^{\infty} \frac{e^{it\log x}}{x^2} dx\right| \leqq \frac{c_{12}(1+|t|)}{j}; \tag{6.12}$$

thus, putting again $\psi(t) = \int_1^{\infty} \frac{e^{it\log x}}{x^2} dx$, we have by Lemma 10

$$|\Phi_n(t) - \psi(t)\Phi_{n-1}(t)| \leqq c_{12}\left(\frac{2}{3}\right)^n (1+|t|). \tag{6.13}$$

Now we can apply the same method as in the proof of Theorem 2, and obtain thus

THEOREM 5.

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\log\frac{Q_n}{Q_1\ldots Q_{n-1}} - n}{\sqrt{n}} < y\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{u^2}{2}} du$$

for any real y, i. e. $\log\frac{Q_n}{Q_1\ldots Q_{n-1}}$ *is in the limit normally distributed.*[6]

The result which we obtained is a consequence of the facts which have been proved implicitly above, that the random variables $X_1 = \log Q_1$,

[6] Theorem 5 implies that $\sqrt[n]{\frac{Q_n}{Q_1\ldots Q_{n-1}}}$ tends in measure to e. More is true, namely that $\sqrt[n]{\frac{Q_n}{Q_1\ldots Q_{n-1}}}$ tend also *almost everywhere* to e. The proof will be published elsewhere.

$X_n = \log \frac{Q_n}{Q_{n-1}^2}$ $(n=2,3,\ldots)$ are in a certain sense almost independent and the distribution of X_n tends for $n \to \infty$ to the exponential distribution with mean 1. The latter fact can be also deduced from the remark that $Q_n(Q_n - 1)R_n$ is for every n exactly uniformly distributed in the interval $(0, 1)$.

As regards Q_n itself, we can prove the following

THEOREM 6.

$$\lim_{n\to\infty} \frac{\log Q_n(x)}{2^n}$$

exists and is finite and positive for almost all x.

PROOF. We have

$$\frac{\log Q_n(x)}{2^n} = \sum_{k=1}^{n} \frac{X_k}{2^k}$$

and as

$$\mathbf{M}(|X_k|) \leqq c_{13} \qquad (k=1,2,\ldots),$$

it follows by the theorem of B. LEVI [11], that the limit in question exists for almost all x. It is easy to see that the limit is always positive. As a matter of fact, if the sequence S_n is defined by $S_1 = 2$, $S_{n+1} = S_n(S_n - 1) + 1$, then (as has been shown already by SYLVESTER) $S_n \geqq 2^{2^{n-2}}$. But $Q_n(x) \geqq S_n$ and therefore $\frac{\log Q_n}{2^n} \geqq \frac{1}{4} \log 2$ $(n = 1, 2, \ldots)$.

§ 7. Some number-theoretical questions

Let a and b be positive integers, $0 < \frac{a}{b} < 1$. It is well-known that $\frac{a}{b}$ can be represented in the form

$$\text{(7.1)} \qquad \frac{a}{b} = \frac{1}{S_1} + \frac{1}{S_2} + \cdots + \frac{1}{S_n},$$

where $S_1 < S_2 < \cdots < S_n$ are positive integers. Such representations of rational numbers have been considered already by the Egyptians, more than 3500 years ago. Denote by $f(a, b)$ the smallest value of n, i. e. the length of the shortest representation of $\frac{a}{b}$ in the form (7.1). If we choose S_1 to be the smallest integer with $\frac{1}{S_1} \leqq \frac{a}{b}$, we have

$$\text{(7.2)} \qquad \frac{a}{b} - \frac{1}{S_1} = \frac{a'}{bS_1}$$

with $a' = aS_1 - b < a$. Thus $f(a, b) \leqq a$. P. ERDŐS proved [12] that

$$f(a, b) < \frac{c_{13} \log b}{\log \log b}, \tag{7.3}$$

but very likely (7.3) can be very much improved; perhaps $f(a, b) < c_{14} \log \log b$. More can not be true, as it is known [8] that $f(b-1, b) > \log \log b - 1$.

It is known [13] that for infinitely many b-s $f(3, b) = 3$. STRAUSS and ERDŐS conjectured that $f(4, b) < 4$ and SCHINZEL and SIERPINSKI conjectured that $f(a, n) \leqq 3$ for all $n > n_0(a)$.

Consider now various special representations of the form (7.1). First of all consider Sylvester's series of $\frac{a}{b}$. Denote by $E_2(a, b)$ the number of terms occurring in this series. From (7.2) it follows that $E_2(a, b) \leqq a$ and it is easy to see that this is best possible, since $E_2(a, a!+1) = a$. We know of no good estimation of $E_2(a, b)$ in terms of b; the trivial estimation $E_2(a, b) < b$ is no doubt very far from being best possible.

We remark that $E_2(a, b) \neq f(a, b)$ e. g. $\frac{9}{20} = \frac{1}{4} + \frac{1}{5} = \frac{1}{3} + \frac{1}{9} + \frac{1}{180}$, i. e. $f(9, 20) = 2$ but $E_2(9, 20) = 3$.

In § 6 we proved that $\frac{\log Q_n}{2^n}$ tends for almost all x to a limit, where Q_n is the n-th denominator of the Sylvester's series of x. It follows from the divergence of the harmonic series that no function $F(x, n)$ can be given so that for any representation $x = \frac{1}{S_1} + \frac{1}{S_2} + \cdots + \frac{1}{S_n} + \cdots$, where S_n is a positive integer and $S_n < S_{n+1}$, we should have $S_n \leqq F(x, n)$. But it seemed possible that for any such representation

$$S_n \leqq Q_n(x) \tag{7.4}$$

infinitely often. Now we show that (7.4) is not always true. Let n_k tend to $+\infty$ sufficiently fast, and put

$$x = \sum_{k=1}^{\infty} \left(\frac{1}{2n_k} + \frac{1}{2n_k + 1} \right).$$

Let $\frac{1}{2n_k} + \frac{1}{2n_k+1} = \frac{1}{S_{k1}} + \cdots + \frac{1}{S_{kl_k}}$ be the Sylvester's series of $\frac{1}{2n_k} + \frac{1}{2n_k+1}$. A simple computation shows that $l_k > 2$ if $n_k > 1$ and if $2n_{k+1} > (S_{kl_k})^2$ then $\sum_{k=1}^{\infty} \sum_{l=1}^{l_k} \frac{1}{S_{kl}}$ is the Sylvester's series of x and clearly here $Q_n(x) < S_n$ for all sufficiently large n. It seems that (7.4) fails for almost all x.

It follows from the fact that a given x can not be approximated arbitrarily well by rational numbers with denominator $\leqq y$ that there exists for every x a function $G(x, n)$ so that if $x = \sum_{k=1}^{\infty} \frac{1}{S_k}$, then $S_n < G(x, n)$ infinitely often. It is easy to find such a function $G(x, n)$ for almost all x; e. g. for almost all x $G(x, n) = (3+\varepsilon)^{2^n}$ has the mentioned property if $\varepsilon > 0$, but it seems difficult to give a good estimation for the smallest such function. Thus in particular we could not decide whether for almost all x there exists a series $x = \sum_{k=1}^{\infty} \frac{1}{S_k}$ (S_n positive integer, $S_n < S_{n+1}$) so that

$$\lim_{n\to\infty} \frac{\log S_n}{2^n} = +\infty.$$

Similar algorithms like that leading to Sylvester's series can be defined by replacing the harmonic series by some other series of positive terms (see e. g. [14]). For these algorithms similar questions can be asked.

Now we consider Engel's series of the first kind of rational numbers, that is the representation

$$\frac{a}{b} = \frac{1}{q_1} + \frac{1}{q_1 q_2} + \cdots + \frac{1}{q_1 q_2 \cdots q_n}.$$

Put $E_1(a, b) = n$. We have no non-trivial estimation of $E_1(a, b)$. Clearly $q_n \leqq b$ and the same value of q_i can not occur too often; in this way one can obtain a very poor upper bound for $E_1(a, b)$. Here too it would be interesting to estimate $E_1(a, b)$ in terms of both a and b. $E_1(a, b) \leqq a$ can be proved as follows:

$$\frac{a}{b} - \frac{1}{q_1} = \frac{a'}{b q_1} \quad \text{where} \quad q' = a q_1 - b < a.$$

As $r_1\left(\frac{a}{b}\right) = \frac{a'}{b}$, it is clear that $E_1(a, b) \leqq a$.

Often $E_1(a, b) = E_2(a, b)$, e. g. $E_1(3, 4) = E_2(3, 4) = 2$ but $E_1(21, 32) = 3$, $E_2(21, 32) = 4$ and $E_1(5, 6) = 3$, $E_2(5, 6) = 2$; thus in general there is no simple inequality between the two numbers.

Denote by $D(a, b)$ the smallest n for which $\frac{a}{b}$ has a representation

$$\frac{a}{b} = \frac{1}{d_1} + \frac{1}{d_1 d_2} + \cdots + \frac{1}{d_1 d_2 \ldots d_n}, \tag{7.5}$$

where $d_1, d_2, \ldots, d_n$ are integers $d_n \geqq 2$ $(n = 1, 2, \ldots)$. Often $D(a, b) < E_1(a, b)$.

If x is irrational and

$$x = \sum_{n=1}^{\infty} \frac{1}{d_1 d_2 \ldots d_n}, \tag{7.6}$$

where d_n is an integer $d_n \geqq 2$ $(n=1,2,\ldots)$, it is easy to give a function $H(x,n)$ so that for almost all x $d_1 d_2 \ldots d_n < H(x,n)$ for all $n > n_0(x)$, but it seems hard to give a good estimation for a function $H(n)$ so that for almost all x and $n > n_0$ we have $d_1 d_2 \ldots d_n < H(n)$.

It is not difficult to see that every x for which $\frac{1}{k} \leqq x \leqq 1$ ($k \geqq 3$ integer) can be written in the form (7.6) with $2 \leqq d_n \leqq k+1$.

As a matter of fact, if $\frac{1}{\nu} \leqq x < \frac{1}{\nu-1}$ where ν is an integer $2 \leqq \nu \leqq k$, then let us put $d_1(x) = \nu+1$ if $\nu > 2$ or if $\nu = 2$ and $\frac{1}{2} \leqq x < \frac{1}{2} + \frac{1}{2k}$ and $d_1(x) = 2$ if $\frac{1}{2} + \frac{1}{2k} \leqq x \leqq 1$. Then putting $r_1(x) = d_1(x)x - 1$, we have always $\frac{1}{k} \leqq r_1(x) \leqq 1$; we define $d_2(x) = d_1(r_1(x))$ and $r_2(x) = d_2(x) r_1(x) - 1$ and so on. Thus we obtain the representation

$$x = \frac{1}{d_1(x)} + \frac{1}{d_1(x)\, d_2(x)} + \cdots + \frac{1}{d_1(x) \ldots d_n(x)} + \cdots, \tag{7.7}$$

where $2 \leqq d_n(x) \leqq k+1$.

A simple modification of the above argument shows that every x $(0 < x < 1)$ can be written in the form

$$x = \sum_{n=1}^{\infty} \frac{1}{d_1 d_2 \ldots d_n} \quad \text{with} \quad 2 \leqq d_n \leqq 4 \text{ for } n \geqq 2. \tag{7.8}$$

As a matter of fact, if $\frac{1}{\nu} \leqq x < \frac{1}{\nu-1}$ where $\nu \geqq 2$ is an integer, then two cases are possible: either $\nu \leqq 3$, then as shown above, x has a representation (7.8) with $2 \leqq d_n \leqq 4$, for $n = 1, 2, \ldots$; on the other hand if $\nu > 3$ then $\frac{1}{3} < (2\nu-2)x - 1 \leqq 1$ and thus x has the representation (7.8) with $2 \leqq d_n \leqq 4$ $(n = 2, 3, \ldots)$ and $d_1 = 2\nu - 2$.

If we require in (7.6) $d_n = 2$ or $d_n = 3$ for $n = 1, 2, \ldots$ it is easy to see that the measure of the set of those numbers for which such a representation exists, is 0.

Let us consider namely all numbers of the form (7.5) where $d_k = 2$ or 3 $(k = 1, 2, \ldots)$. As for each x which has the representation (7.6) with $d_n = 2$ or 3 $(n = 1, 2, \ldots)$ where $d_1, d_2, \ldots, d_N$ are fixed, is contained in an interval of length $(2 d_1 \ldots d_N)^{-1}$, the set of all numbers x which have such a representation is covered by a set of intervals, the sum of length of which does not exceed $\left(\frac{1}{2} + \frac{1}{3}\right)^N$. Thus this set has the measure 0.

It would be interesting to determine the greatest value of β such that for almost all x and any representation of x in the form (7.8), putting $D_n = d_1 d_2 \ldots d_n$ we have $\varliminf_{n\to\infty} \sqrt[n]{D_n} \geqq \beta$. A modification of the above argument shows that there exists such a $\beta > 2$; more exactly we shall show that we may take $\beta = 2^{12/13} \cdot 3^{1/13}$. This can be shown as follows. The set of those numbers x, which have a representation of the form (7.8) where $d_1, \ldots, d_{N+1}$ are fixed, is contained in an interval of the length $\frac{2}{3 d_1 \ldots d_{N+1}}$. Thus all those numbers x, which have a representation of the above form such that $d_1 = j$ and between the numbers $d_2, \ldots, d_{N+1}$ not more than cN are different from 2, where $c < 1$, are covered by a set of intervals with total length

$$L_N(c) = \frac{2}{3j}\left(\frac{13}{12}\right)^N \sum_{k \leqq cN} \binom{N}{k} \left(\frac{7}{13}\right)^k \left(\frac{6}{13}\right)^{N-k}.$$

Now it is easy to show (see e. g. [15], p. 405) that if $0 < p < 1$, $q = 1 - p$ and $0 < \varepsilon < p$, then

$$\sum_{k \leqq (p-\varepsilon)N} \binom{N}{k} p^k q^{N-k} \leqq e^{-\frac{\varepsilon^2 N}{2pq\left(1+\frac{\varepsilon}{2pq}\right)^2}}.$$

Thus it follows that for $c = \frac{1}{13}$

$$L_N\left(\frac{1}{13}\right) \leqq \left(\frac{85293}{87924}\right)^N \cdot \frac{2}{3j}.$$

Thus the set of those numbers x which have a representation (7.8) in which putting $D_n = d_1 d_2 \ldots d_n$, we have $\varliminf_{n\to\infty} \sqrt[n]{D_n} < 3^{1/13} 2^{12/13}$ has measure 0. It follows that for almost all x $\varliminf_{n\to\infty} \sqrt[n]{D_n} \geqq 2^{12/13} 3^{1/13}$, which was to be proved.

Now we construct an x $(0 < x < 1)$ for which

$$x = \sum_{n=1}^{\infty} \frac{1}{D_n}, \tag{7.9}$$

where $D_n \geqq 2$ is an integer, D_n / D_{n+1} and $D_{n+1} \geqq D_n^2$ i. e. (7.9) is the *Engel's* series and at the same time the *Sylvester's* series of x and is such that for every k

$$\min_{2 \leqq S_1 \leqq S_2 \leqq S_k,\ \sum_{i=1}^{k} \frac{1}{S_i} < x} \left(x - \sum_{i=1}^{k} \frac{1}{S_i}\right) = x - \sum_{i=1}^{k} \frac{1}{D_i}. \tag{7.10}$$

From this it will be easy to deduce that if

$$x = \sum_{n=1}^{\infty} \frac{1}{S_n} \tag{7.11}$$

is any other representation of x such that the S_n are integers, $S_{n+1} \geqq S_n$, then $S_n \leqq D_n$ infinitely often (we showed previously that this is not true for all x).

We construct the D_n inductively. Put $D_1 = 2$, let $D_2 = kD_1$ be so large that $\frac{1}{D_1} + \frac{1}{D_2}$ is much less than every $\frac{1}{a} + \frac{1}{b} > \frac{1}{D_1}$ with $a \neq D_1$ („much less" means $\frac{1}{D_1} + \frac{2}{D_2} < \frac{1}{a} + \frac{1}{b}$); then choose $D_3 = lD_2$ so large that $\frac{1}{D_1} + \frac{1}{D_2} + \frac{1}{D_3}$ is much less than every $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} > \frac{1}{D_1} + \frac{1}{D_2}$ with $a \neq D_1$ or $b \neq D_2$.

This construction clearly gives an x with the required properties.

Let

$$x = \sum_{n=1}^{\infty} \frac{1}{D_n} \tag{7.12}$$

be Sylvester's series of x and (7.11) another representation of x. Is it true that for almost all x (7.10) holds for infinitely many k? (Our construction gives only a set of measure 0 of such numbers x for which this is true.)

Literature

[1] O. Perron, *Irrationalzahlen,* De Gruyter, Berlin u. Leipzig, 1st edition, 1921. 2nd. ed., 1939, 116—122.

[2] A. Rényi, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hung.,* **8** (1957), 477—493.

[3] É. Borel, Sur les développements unitaires normaux, *Comptes Rendus,* **225** (1947), 51.

[4] É. Borel, Sur les développements unitaires normaux, *Ann. Soc. Polonaise de Math.,* **21** (1948), 74—79.

[5] P. Lévy, Remarques sur un théorème de M. Émile Borel, *Comptes Rendus,* **225** (1947), 918—919.

[6] J. J. Sylvester, On a point in the theory of vulgar fractions, *Amer. Journal of Math.* **3** (1880), 332—335., postscript ibid. 388—389. (Reprinted in J. J. Sylvester, *The collected mathematical papers,* Cambridge, 1909. *Vol. 3.,* 440—443.)

[7] H. E. Salzer, Further remarks on the approximation of numbers as sums of reciprocals. *Amer. Math. Monthly,* **55** (1948), 350—356.

[8] H. Cramér, *Mathematical methods of statistics,* Princeton, 1946, 96.

[9] A. Békéssy, Bemerkungen zur Engelschen Darstellung reeller Zahlen. *Annales Univ. Sci. Budapest,* Sectio Math., **1** (1958), 143—152.

[10] P. Erdős, On the law of the iterated logarithm, *Annals of Math.*, **43** (1942), 419—436.
[11] F. Riesz—B. Szőkefalvi-Nagy, *Leçons d'analyse fonctionnelle*, 2nd. ed. Akadémiai Kiadó, Budapest, 1953, 36.
[12] P. Erdős, On a diophantine equation (in hungarian, with english and russian summary). *Matematikai Lapok*, **1** (1950), 192—210.
[13] N. Nakayama, On the decomposition of a rational number into "Stammbrüche". *Tôkohu Math. Journal*, **46** (1940), 1—21.
[14] É. Borel, Sur les systèmes généraux de numeration. *Comptes Rendus*, **226** (1948), 1405—1407.
[15] Rényi A., *Valószinűségszámitás*, Budapest, Tankönyvkiadó, 1954.
[16] L. Berg, Allgemeine Kriterien zur Maßbestimmung linearer Punktmengen, *Mathematische Nachrichten*, **14** (1955), 263—285.