

On the distribution of digits

By I. KÁTAI and J. MOGYORÓDI (Budapest)

1. Introduction

Let $K(>1)$ be a fixed positive integer. Then any integer n can be uniquely represented as follows:

$$n = a_1 K^{n_1} + \dots + a_t K^{n_t}$$

where $n_1 > n_2 > n_3 > \dots > n_t \geq 0$ are integers; $1 \leq a_i \leq K-1$ ($i=1, \dots, t$). We set

$$\alpha(n) = \sum_{i=1}^t a_i.$$

R. BELLMAN and H. SHAPIRO proved the relation

$$\sum_{n \leq x} \alpha(n) = \frac{x \log x}{2 \log 2} + O(x \log \log x)$$

in the case $K=2$ [1]. S. C. TANG extended this result to the general case and discovered a better alternative residual, namely he proved [2], that for any positive integer

$$\sum_{n \leq x} \alpha(n) = \frac{K-1}{2} \frac{x \log x}{\log K} + O(x).$$

The first named author proved in [3], that assuming the validity of the density hypothesis concerning the Riemann zeta function

$$\sum \alpha(p) = \frac{K-1}{2} \frac{x}{\log K} + O\left(\frac{x}{(\log \log x)^{1/3}}\right)$$

holds, where p in the sum runs over all of the primes not exceeding x .

In the present paper we shall investigate the limit distribution of $\alpha(n)$ and of $\alpha(p)$, assuming the validity of the density hypothesis in the second case.

In what follows we use the following notations:

$$(1.1) - (1.2) \quad \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-t^2/2} dt, \quad \text{li } x = \int_2^x \frac{du}{\log u}.$$

Let us put

$$(1.3) - (1.4) \quad M_x = \frac{K-1}{2} \frac{\log x}{\log K}, \quad D_x^2 = \frac{K^2-1}{12} \frac{\log x}{\log K}$$

and let $N_x(y)$ be the number of those n -s for which

$$(1.5) \quad n \leq x \quad \text{and} \quad \alpha(n) < M_x + yD_x,$$

and similarly let $M_x(y)$ be the number of those primes p for which

$$(1.6) \quad p \leq x \quad \text{and} \quad \alpha(p) < M_x + yD_x$$

holds.

Let $\zeta(s)$ be the zeta function of Riemann and let $N(\sigma_0, T)$ denote the number of the zeros of $\zeta(s)$ in the rectangle $\sigma_0 \leq \sigma \leq 1, |t| \leq T, s = \sigma + it$.

We shall prove the following statements.

Theorem 1. *For every fixed K*

$$(1.7) \quad \frac{N_x(y)}{x} = \Phi(y) + O\left(\frac{\log \log x}{(\log x)^{1/2}}\right)$$

holds uniformly in y as x tends to infinity.

Theorem 2. *Assuming that*

$$N(\sigma, T) < cT^{2(1-\sigma)} \log^2 T, \quad \text{if } \frac{1}{2} \leq \sigma \leq 1, 1 \leq T < \infty$$

for a suitable constant c we have

$$\frac{M_x(y)}{\text{li } x} = \Phi(y) + O\left(\frac{1}{(\log \log x)^{1/3}}\right)$$

uniformly in y as x tends to infinity.

The proof of Theorem 1 goes with the application of the limit distribution theory for the sums of independent random variables.

For the proof of Theorem 2 we need a lemma concerning the distribution of prime numbers in small intervals.

2. Formulation and proof of Lemma 1.

In the sequel p denotes prime numbers, c a positive constant, not the same at every place.

Let $\Lambda(n)$ be Mangoldt's function i.e. $\Lambda(n) = \log p$, if n is a power of p and $\Lambda(n) = 0$ if n has two different prime divisors. Let

$$(2.1) - (2.2) \quad \psi(x) = \sum_{n \leq x} \Lambda(n); \quad \pi(x) = \sum_{p \leq x} 1,$$

i.e. $\pi(x)$ denotes the number of primes not exceeding x .

Let further

$$(2.3) \quad \Delta_U(x) = \Delta(x) = \psi(x) - \psi(x - (U + 1)) - (U + 1),$$

$$(2.4) \quad \varrho_U(x) = \pi(x + U) - \pi(x) - \frac{U}{\log x},$$

where U is a positive number.

Let $h(x)$ be a monotonous non-decreasing function of x , which tends to infinity as $x \rightarrow +\infty$ and which satisfies the relation

$$(2.5) \quad 1 \leq h(x) < c \log x.$$

Lemma 1. If

$$(2.6) \quad N(\sigma, T) < cT^{2(1-\sigma)} \log^2 T \quad \text{in} \quad \frac{1}{2} \leq \sigma \leq 1,$$

and

$$(2.7) \quad U = [(\log x)^{7.5} h(x)],$$

then

$$(2.8) \quad \sum_{n \leq x} A_U^2(n) < c \frac{U^2 x}{h(x)},$$

and

$$(2.9) \quad \sum_{n \leq x} Q_U^2(n) < c \frac{U^2 x}{h(x) \log^2 x}.$$

We remark that a similar, stronger result was discovered by H. CRAMER [4] and improved by A. SELBERG [5] on the basis of the Riemann conjecture.

PROOF. First we prove that (2.9) follows from (2.8). We have

$$\begin{aligned} Q_U(n) &= (\log x)^{-1} \left\{ \sum_{n \leq p < n+U} \log x - U \right\} = (\log x)^{-1} \left\{ \sum_{n \leq v < n+U} A(v) - U \right\} + \\ &+ (\log x)^{-1} \sum_{n \leq p < n+U} \log \frac{x}{p} + (\log x)^{-1} \sum_{n \leq p^k < n+U} \log p, \quad k \geq 2 \end{aligned}$$

and so

$$|Q_U(n)| \leq \frac{1}{\log x} |A_U(n)| + \frac{\log \frac{x}{n}}{\log x} \left\{ |Q_U(n)| + \frac{U}{\log x} \right\} + 2 \sum_{n < p^k < n+U} 1, \quad k \geq 2$$

holds. Assuming that $n \geq x^{\frac{1}{2}}$, i.e. $\log \frac{x}{n} < \frac{1}{2} \log x$, we have

$$\frac{1}{2} |Q_U(n)| < \frac{|A_U(n)|}{\log x} + \frac{U}{\log^2 x} + 2 \sum_{n < p^k < n+U} 1, \quad k \geq 2.$$

Hence

$$(2.10) \quad \sum_{x^{1/2} \leq n < x} |Q_U(n)|^2 < \frac{c}{\log^2 x} \sum_{n \leq x} A_U^2(n) + \frac{cxU^2}{\log^4 x} + c \sum_{n \leq x} \left\{ \sum_{n < p^k < n+U} 1 \right\}^2, \quad k \geq 2$$

follows. Using the evident inequality

$$\left\{ \sum_{n \leq p^k < n+U} 1 \right\}^2 < U \sum_{n \leq p^k < n+U} 1$$

we obtain that the last sum on the right hand side of (2.10) is smaller than

$$U \sum_{n \leq x} \sum_{n \leq p^k < n+U} 1 < U^2 \sum_{p^k < x} 1 \leq U^2 (\pi(x^{1/2}) + \pi(x^{1/3}) + \dots) \quad k \geq 2$$

because of (2.5). Further using the inequalities

$$\frac{xU^2}{\log^4 x} < \frac{xU^2}{h(x)\log^2 x}, \quad \sum_{n \equiv x^{1/2}} |\varrho_V(n)|^2 < U^2 x^{1/2} < \frac{xU^2}{h(x)\log^2 x}$$

we obtain that

$$\sum_{n \equiv x} |\varrho_V(n)|^2 < \frac{c}{\log^2 x} \sum_{n \equiv x} |A_V(n)|^2 + c \frac{U^2 x}{h(x)\log^2 x}$$

and hence it follows that (2.9) is a consequence of (2.8).

For the proof of the inequality (2.9) let

$$(2.11) \quad f(z) = \sum_{n=1}^{\infty} A(n)e^{-nz}; \quad z = u + iy, \quad 0 < u < 1, \quad -\pi \leq y \leq \pi.$$

Let

$$(2.12) \quad T(z) = \sum_{\varrho} z^{-\varrho} \Gamma(\varrho),$$

where the sum is extended over all the non-trivial zeros of $\zeta(s)$. Let further

$$(2.13) \quad \Delta = \left(\log \frac{1}{U} \right)^{-7},$$

$$(2.14) \quad g(z) \stackrel{\text{def}}{=} f(z)(1 + e^{-z} + \dots + e^{-Uz}) - (U+1) \frac{e^{-z}}{1 - e^{-z}}.$$

From (2.14) it follows that

$$(2.15) \quad g(z) = \sum_{n=1}^{\infty} A_V(n)e^{-nz},$$

and by the Parseval formula we have

$$\sum_{n=1}^{\infty} A^2(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(z)|^2 dy.$$

Let now assume that uU tends to zero for $u \rightarrow 0$. Then

$$1 + e^{-z} + \dots + e^{-Uz} = O\left(\frac{1}{|z|}\right) = O\left(\frac{1}{\Delta}\right), \quad \frac{e^{-z}}{1 - e^{-z}} = O\left(\frac{1}{y}\right),$$

if $|y| \geq \Delta$. Hence

$$\begin{aligned} \int_{|y| \geq \Delta} |g(z)|^2 dy &\leq 2 \int_{|y| \geq \Delta} |f(z)|^2 |1 + e^{-z} + \dots + e^{-Uz}|^2 dy + cU^2 \int_{|y| \geq \Delta} y^{-2} dy \leq \\ &\leq \frac{c}{\Delta^2} \int_{-\pi}^{\pi} |f(z)|^2 dy + c \frac{U^2}{\Delta} \end{aligned}$$

follows. Using the prime number theorem we have

$$\int_{-\pi}^{\pi} |f(z)|^2 dy < c \frac{1}{u} \log \frac{1}{u},$$

and so

$$(2.16) \quad \int_{|y| \leq A} |g(z)|^2 dy < \frac{c}{A^2 u} \log \frac{1}{u} + c \frac{U^2}{A}$$

holds.

We shall investigate the integral $\int_{|y| \leq A} |g(z)|^2 dy$. JU. V. LINNIK proved [6] that

$$(2.17) \quad f(z) = \frac{1}{z} - T(z) + O\left(\log^3 \frac{1}{u}\right),$$

and that under the assumptions of Lemma 1

$$(2.18) \quad \int_{-A}^A |T(z)|^2 dy < \frac{c}{u} \left(\log \frac{1}{u}\right)^{-1}$$

holds [7].

From (2.17) we obtain

$$\begin{aligned} \int_{|y| \leq A} |g(z)|^2 dy &\leq 3 \int_{-A}^A \left| \frac{1 + \dots + e^{-Uz}}{z} - (U+1) \frac{e^{-z}}{1 - e^{-z}} \right|^2 dy + \\ &+ 3 \int_{-A}^A |T(z)|^2 |1 + \dots + e^{-Uz}|^2 dy + O\left(\log^6 \frac{1}{u}\right) \int_{-A}^A |1 + \dots + e^{-Uz}|^2 dy. \end{aligned}$$

The first integral on the right hand side has the order $K^4 A$, because we have

$$\frac{e^{-z}}{1 - e^{-z}} = \frac{1}{z} + O(1), \quad 1 + \dots + e^{-Uz} = U + 1 + O(U^2 |z|).$$

The second term is smaller than $U^2 \frac{1}{u} \left(\log \frac{1}{u}\right)^{-1}$ (see 2.18)) and finally

$$\int_{-A}^A |1 + \dots + e^{-Uz}|^2 dy \leq (U+1)^2 2\pi.$$

Hence

$$\int_{-\pi}^{\pi} |g(z)|^2 dy < c \left\{ U^4 A + \frac{U^2}{u} \left(\log \frac{1}{u}\right)^{-1} + U \log^6 \frac{1}{u} + \frac{1}{A^2} \frac{1}{u} \left(\log \frac{1}{u}\right)^{-1} + \frac{U^2}{A} \right\}.$$

Let now $\frac{1}{u} = x$, $U = (\log x)^{7.5} h(x)$, $1 \leq h(x) \leq \log x$, then the right hand side is smaller than $c \frac{U^2 x}{h(x)}$.

Now our inequality (2.8) rapidly follows from the relation

$$\sum_{n \leq x} A^2(n) < c \sum_{n=1}^{\infty} A^2(n) e^{-2n/x}.$$

3. The proof of Theorem 1.

Let $\xi_1, \dots, \xi_n, \dots$ be independent random variables assuming the values $v=0, 1, 2, \dots, K-1$ with probability

$$(3.1) \quad P(\xi_i = v) = \frac{1}{K}, \quad i = 1, 2, \dots$$

Let $M(\xi)$ and $D(\xi)$ denote the mean-value and the variance of a random variable ξ respectively, further let for the sake of brevity

$$(3.2) \quad M = M(\xi_i) = \frac{K-1}{2}, \quad D^2 = D^2(\xi_i) = \frac{K^2-1}{12} \quad (i = 1, 2, \dots).$$

Let $\eta_n = \xi_1 + \dots + \xi_n$, so $M(\eta_n) = nM$, $D^2(\eta_n) = nD^2$. Let $F_n(y)$ be a distribution function defined by

$$(3.3) \quad F_n(y) = P(\eta_n < nM + y\sqrt{n}D).$$

In the theory of probability the following assertion is well known.

Lemma 2.

$$(3.4) \quad |F_n(y) - \Phi(y)| < \frac{c}{\sqrt{n}}$$

uniformly in $-\infty \leq y \leq \infty$.

For the proof see the book of B. V. GNEDENKO—A. N. KOLMOGOROV [8] (Theorem 1, p. 201).

Let $B_m(x)$ denote the number of those n -s, for which $n < x$, $\alpha(n) = m$ is satisfied.

It is evident, that for an integer $l \geq 1$

$$(3.5) \quad B_m(K^l) = K^l P(\eta_l = m).$$

Further, if $1 \leq A \leq K-1$, then

$$\begin{aligned} B_m(AK^l) &= \sum_{n < K^l} 1 + \dots + \sum_{(i-1)K^l \leq n < iK^l} 1 + \dots + \sum_{(A-1)K^l \leq n < AK^l} 1 = \\ &= B_m(K^l) + B_{m-1}(K^l) + \dots + B_{m-A}(K^l), \quad \alpha(n) = m. \end{aligned}$$

Now we assume:

$$x = A_1 K^{n_1} + \dots + A_t K^{n_t}, \quad n_1 > n_2 > \dots > n_t \geq 0,$$

$$1 \leq A_i \leq K-1, \quad i = 1, \dots, t.$$

Then

$$\begin{aligned} (3.6) \quad B_m(x) &= \sum_{i=0}^{A_1-1} B_{m-i}(K^{n_1}) + B_m(x - A_1 K^{n_1}) = \\ &= \sum_{i=0}^{A_1-1} B_{m-i}(K^{n_1}) + \sum_{i=A_1}^{A_1+A_2-1} B_{m-i}(K^{n_2}) + \dots + \sum_{i=A_1+\dots+A_{t-1}}^{A_1+\dots+A_t} B_{m-i}(K^{n_t}). \end{aligned}$$

Let now suppose that

$$(3.7) \quad n_1 - n_t > c \log \log x = L$$

with a sufficiently large constant c .

Let for the sake of brevity

$$(3.8) \quad T_y = M_x + yD_x = \frac{\log x}{\log K} \cdot M + y \left(\frac{\log x}{\log K} \right)^{1/2} D.$$

Using the relations (3.5), (3.6) we have

$$(3.9) \quad N_x(y) = \sum_{\substack{\alpha(n) < T_y \\ n \leq x}} 1 = K^{n_1} \sum_{i=0}^{A_1-1} P(\eta_{n_1} < T_y - i) + K^{n_2} \sum_{i=A_1}^{A_1+A_2-1} P(\eta_{n_2} < T_y - i) + \\ + \dots + K^{n_t} \sum_{i=A_1+\dots+A_{t-1}}^{A_1+\dots+A_t} P(\eta_{n_t} < T_y - i).$$

Using the inequality (3.4) we have that for every pair of real numbers y_1, y_2 the inequality

$$(3.10) \quad |F_n(y_1) - F_n(y_2)| \leq |F_n(y_1) - \Phi(y_1)| + |\Phi(y_1) - \Phi(y_2)| + |\Phi(y_2) - F_n(y_2)| \leq \\ \leq \frac{2c}{\sqrt{n}} + \int_{y_1}^{y_2} e^{-\frac{u^2}{2}} du < \frac{2c}{\sqrt{n}} + |y_1 - y_2|$$

is satisfied.

Let i be an arbitrary integer value in the interval $0 \leq i \leq A_1 + \dots + A_t \leq Kt < cK \log \log x$ and let Θ be defined by

$$(3.11) \quad T_y - i = Mn_l + (y + \Theta)\sqrt{n_l}D,$$

l is an arbitrary integer in $1 \leq l \leq t$.

Using the inequalities $n_1 \leq \frac{\log x}{\log K} \leq n_1 + 1$ and (3.7) we have that

$$|\Theta| \leq \frac{M \left| \frac{\log x}{\log K} - n_l \right| + yD \left| \sqrt{n_l} - \left(\frac{\log x}{\log K} \right)^{1/2} \right|}{\sqrt{n_l}D} \leq \\ \leq c(\log x)^{-1/2} \log \log x \left(1 + \frac{|y|}{(\log x)^{1/2}} \right).$$

Hence for every term on the right hand side of (3.9)

$$|P(\eta_{n_l} < T_y - i) - F_{n_l}(y)| < c(\log x)^{-1/2} \log \log x \left(1 + \frac{|y|}{(\log x)^{1/2}} \right)$$

and so from (3.4)

$$(3.12) \quad |P(\eta_{n_l} < T_y - i) - \Phi(y)| < c(\log x)^{-1/2} \log \log x \left(1 + \frac{|y|}{(\log x)^{1/2}} \right)$$

follows.

So from (3. 9) we have

$$(3. 13) \quad N_x(y) = x\Phi(y) + O(x(\log x)^{-\frac{1}{2}} \log \log x (1 + |y|(\log x)^{-\frac{1}{2}})).$$

Let now x^* be an arbitrary integer, the K -adical representation of which is

$$x^* = A_1 K^{n_1} + \dots + A_u K^{n_u}. \quad \text{Let } n_1 \cong n_2 \cong \dots \cong n_t \cong L, L > n_{t+1} > \dots > n_u,$$

$$x = A_1 K^{n_1} + \dots + A_t K^{n_t}, \quad x = x^* + x_1.$$

So we have

$$(3. 14) \quad x_1 < K^{n_{t+1}+1} \cong K^{n_1} K^{-L} < x^* e^{-L \log K} < x^* / (\log x),$$

further from (3. 14)

$$(3. 15) \quad |N_{x^*}(y) - \Phi(y)x^*| \cong |N_x(y) - x\Phi(y)| + 2|x^* - x| \cong \\ \cong c(\log x^*)^{-1/2} \log \log x^* \left(1 + \frac{|y|}{(\log x^*)^{1/2}}\right)$$

follows.

So the relation (1. 7) is satisfied uniformly on the interval $|y| \cong c(\log x)^{\frac{1}{2}}$. Further, if $y > c(\log x)^{\frac{1}{2}}$, or $y < -c(\log x)^{\frac{1}{2}}$, then $N_x(y) = x$ or $N_x(y) = 0$, respectively, and in these cases the inequalities

$$0 \cong 1 - \Phi(y) \cong c \int_y^\infty e^{-u^2/2} du < (\log x)^{-1/2}, \quad y \cong c(\log x)^{1/2} \\ 0 \cong \Phi(y) < \int_{-\infty}^y e^{-u^2/2} du < (\log x)^{-1/2}, \quad y < -c(\log x)^{1/2}$$

hold.

So Theorem 1 is proved.

4. Proof of Theorem 2.

Let

$$(4. 1) \quad U = [(\log x)^{7.5} h(x)],$$

where $h(x)$ is a function of x tending monotonically to infinity as $x \rightarrow +\infty$, and satisfying the relation $h(x) = O(\log x)$. Let l be a natural number satisfying the inequality

$$(4. 2) \quad U \cong K^l < KU.$$

Hence

$$(4. 3) \quad (0 <) c_1 < \frac{l}{\log \log x} < c_2$$

follows.

Let further A_j denote the set of integers in the interval

$$(4.4) \quad [K^l j, K^l(j+1)] \quad \text{for } j=0, 1, \dots, j_0,$$

where

$$(4.5) \quad j_0 = \left\lfloor \frac{x}{K^l} \right\rfloor.$$

It is evident that these sets are disjoint and their union contains any natural number smaller than x .

Further, if $n \in A_j$, then

$$(4.6) \quad \alpha(j) \leq \alpha(n) \leq \alpha(j) + (K-1)l.$$

Let $\lambda(n, y)$ be defined as follows:

$$(4.7) \quad \lambda(n, y) = \begin{cases} 1, & \text{if } \alpha(n) < T_y, \\ 0, & \text{if } \alpha(n) \geq T_y. \end{cases}$$

From the definition of $M_x(y)$

$$(4.8) \quad M_x(y) = \sum_{p \leq x} \lambda(p, y),$$

further from (4.4)

$$M_x(y) = \sum_{j=0}^{j_0} \sum_{p \in A_j} \lambda(p, y) - \sum_{\substack{x < p \\ p \in A_{j_0}}} \lambda(p, y)$$

follows.

Using the inequality

$$(4.9) \quad \pi(x+y) - \pi(x) < c \frac{y}{\log y} \quad (x \geq 1, y > 1),$$

which can be obtained by the sieve-method of Selberg [9], we have

$$(4.10) \quad \sum_{x < p} \lambda(p, y) < c \frac{U}{\log U} < c \frac{U}{\log \log x}, \quad p \in A_{j_0}.$$

Further

$$\lambda(p, y) = \lambda(j, y)$$

if

$$(4.11) \quad T_y - lK \leq j < T_y.$$

So

$$(4.12) \quad M_x(y) = \sum_{j=0}^{j_0} \lambda(j, y) (\pi(K^l(j+1)) - \pi(K^l j)) + O(\Sigma_1) + O(\Sigma_2),$$

where

$$(4.13) \quad \Sigma_1 = \sum'_{j \geq j_0} (\pi(K^l(j+1)) - \pi(K^l j)),$$

$$(4.14) \quad |\Sigma_2| \leq \frac{U}{\log \log x},$$

and the dash in (4.13) means that we sum over those j -s for which (4.11) is satisfied.

Let now

$$(4.15) \quad V = \sum_{j \equiv j_0} 1, \quad T_y - IK < \alpha(j) < T_y.$$

Then

$$(4.16) \quad \Sigma_1 \leq V \frac{U}{\log x} + O(\Sigma_3),$$

where

$$(4.17) \quad \Sigma_3 \stackrel{\text{def}}{=} \sum_{j \equiv j_0} |\varrho_{K^l}(K^l j)|.$$

Further we have

$$(4.18) \quad \sum_{j \equiv j_0} \lambda(j, y) (\pi(K^l(j+1)) - \pi(K^l j)) = \frac{K^l}{\log x} \sum_{j \equiv j_0} \lambda(j, y) + O(\Sigma_3),$$

and so from (4.12) we obtain

$$(4.19) \quad M_x(y) = \frac{K^l}{\log x} \sum_{j \equiv j_0} \lambda(j, y) + O\left(V \frac{U}{\log x}\right) + O\left(\frac{U}{\log \log x}\right) + O(\Sigma_3).$$

First we shall estimate Σ_3 .

Let A be a natural number $< K^l$. For the integers u in $1 \leq u \leq A$ we have

$$|\varrho_{K^l}(n+u) - \varrho_{K^l}(n)| \leq \pi(n+K^l+A) - \pi(n+K^l) + \pi(n+A) - \pi(n) \leq c \frac{A}{\log A}$$

if $n \leq x$ (see (4.9)).

Hence the inequality

$$|\varrho_{K^l}(K^l j)| \leq \frac{1}{A+1} \sum_{u=0}^A |\varrho_{K^l}(K^l j + u)| + c \frac{A}{\log A}$$

follows, and so

$$(4.20) \quad \Sigma_3 \leq \frac{1}{A+1} \sum_{j=0}^{j_0} \sum_{u=0}^A |\varrho_{K^l}(K^l j + u)| + c j_0 \frac{A}{\log A}$$

holds.

Since $A < K^l$, any natural number n can be represented in the form $n = K^l j + u$ ($j=0, \dots, j_0$; $u=0, \dots, A$) at most once, and the number of the represented numbers equals $(A+1)(j_0+1)$. Using the Hölder-inequality and Lemma 1 we have that the double sum is smaller than

$$(4.21) \quad \frac{1}{A+1} \left\{ \sum_{n=K^l j+u} 1 \right\}^{1/2} \left\{ \sum_{n \leq 2x} |\varrho_{K^l}(n)|^2 \right\}^{1/2} \leq \frac{c}{A} \left(\frac{x A}{K^l} \right)^{1/2} \left(\frac{U^2 x}{h(x) \log^2 x} \right)^{1/2} \leq \\ \leq c \frac{x}{\log x} \left(\frac{U}{A h(x)} \right)^{1/2}.$$

Further

$$(4.22) \quad j_0 \frac{A}{\log A} \leq \frac{x}{U} \frac{A}{\log A}.$$

Let us now choose $h(x) = \log x$, $A = \frac{U(\log \log x)^{2/3}}{\log x}$, so the right hand sides of the inequalities (4. 21), (4. 22) are smaller than

$$\frac{cx}{\log x (\log \log x)^{1/3}},$$

and so

$$(4. 23) \quad \Sigma_3 < \frac{cx}{\log x (\log \log x)^{1/3}}$$

holds.

We shall now estimate V .

Let

$$T_y = M \frac{\log \frac{x}{K^l}}{\log K} + y_1 D \left(\frac{\log \frac{x}{K^l}}{\log K} \right)^{1/2} \stackrel{\text{def}}{=} T_{y_1}^*$$

$$T_y - lK = M \frac{\log \frac{x}{K^l}}{\log K} + y_2 D \left(\frac{\log \frac{x}{K^l}}{\log K} \right)^{1/2} \stackrel{\text{def}}{=} T_{y_2}^*.$$

Hence

$$(4. 24) \quad |y_1 - y_2| \leq c \frac{lK}{(\log x)^{1/2}} \leq c \frac{\log \log x}{(\log x)^{1/2}}$$

follows.

Further using Theorem 1 we have

$$(4. 25) \quad V = \sum_{T_{y_1} \leq \alpha(j) < T_{y_2}} N_{x/K^l}(y_2) - N_{x/K^l}(y_1) \leq$$

$$\leq \frac{x}{K^l} |\Phi(y_1) - \Phi(y_2)| + \left| N_{x/K^l}(y_1) - \frac{x}{K^l} \Phi(y_1) \right| + \left| N_{x/K^l}(y_2) - \frac{x}{K^l} \Phi(y_2) \right| \leq$$

$$\leq c \frac{x}{K^l} (\log \log x) (\log x)^{-1/2}, \quad j \leq j_0.$$

Further

$$T_y = M \frac{\log \frac{x}{K^l}}{\log K} + (y + \Theta) D \left(\frac{\log \frac{x}{K^l}}{\log x} \right)^{1/2},$$

where

$$|\Theta| \leq \frac{\log \log x}{(\log x)^{1/2}} \left(1 + \frac{|y|}{(\log x)^{1/2}} \right).$$

So, using similar argumentation as in the estimation of V , we have

$$\sum_{\alpha(j) < T_y} 1 = N_{x/K^l}(y + \Theta) = \frac{x}{K^l} \{ \Phi(y) + O((\log x)^{-1/2} \log \log x) \}, \quad j \leq j_0$$

uniformly in $|y| < c (\log x)^{1/2}$.

To deal with the case $|y| > c (\log x)^{1/2}$ one has to repeat the argument used for the proof of Theorem 1.

Taking into account the relations (4. 19), (4. 23), (4. 25) the proof is completed.

5. Remarks

Let $\beta_v(n)$ denote the number of v -s among the digits of the K -adical representation of n , i.e. when

$$K^U \leq n < K^{U+1}, \quad n = \sum_{j=0}^U \varepsilon_j(n) K^j, \quad (\varepsilon_j(n) = 0, 1, \dots, K-1),$$

then

$$\beta_v(n) = \sum_{j=0}^U 1, \quad \varepsilon_j(n) = v.$$

Let further $U_v(x, y)$ be the number of those n -s for which

$$n \leq x, \quad \beta_v(n) < \frac{\log x}{K \log K} + \frac{(K-1)^{1/2}}{K} \left(\frac{\log x}{\log K} \right)^{1/2} y$$

is satisfied, and similarly let $V_v(x, y)$ be the number of those primes p for which

$$p \leq x, \quad \beta_v(p) < \frac{\log x}{K \log K} + \frac{(K-1)^{1/2}}{K} \left(\frac{\log x}{\log K} \right)^{1/2} y$$

holds.

Using the same methods as used for the proof of Theorems 1, 2 we can prove the following assertions.

Theorem 3. For every $v=0, 1, \dots, K-1$

$$\frac{1}{x} U_v(x, y) = \Phi(y) + O\left(\frac{\log \log x}{(\log x)^{1/2}}\right)$$

uniformly in y as x tends to infinity.

Theorem 4. Assuming the conditions of Theorem 2 we have

$$\frac{1}{\text{li } x} V_v(x, y) = \Phi(y) + O\left(\frac{1}{(\log \log x)^{1/3}}\right), \quad v = 0, 1, \dots, K-1$$

uniformly in y as x tends to infinity.

References

- [1] R. BELLMAN and H. SHAPIRO, On a problem in additive number theory, *Ann. of Math.* (2) **49** (1948), 333—340.
- [2] S. C. TANG, An improvement and generalization of Bellman-Shapiro's theorem on a problem in additive number theory, *Proc. Amer. Math. Soc.*, **14** (1963), 199—204.
- [3] I. KÁTAI, On the sum of digits of prime numbers, *Annales Univ. Budapest*, **10** (1967), under press.
- [4] H. CRAMER, On the order of magnitude of the difference between consecutive prime numbers, *Acta Arithm.* **2** (1937), 23—46.
- [5] A. SELBERG, On the normal density of primes in small intervals and the difference between consecutive primes, *Arch. Math. Naturvid.*, **47** (1943), 87—105.
- [6] JU. V. LINNIK, Some conditional theorems concerning binary problems with prime numbers, *Dokl. Akad. Nauk. SSSR*, **77** (1951), 15—18. (in Russian).
- [7] JU. V. LINNIK, Some conditional theorems concerning the binary problem of Goldbach, *Izv. Akad. Nauk SSSR Ser. Mat.* **16** (1952), 503—520. (in Russian).
- [8] B. V. GNEDENKO and A. N. KOLMOGOROV, Limit distributions for sums of independent random variables (translated from the Russian) *Cambridge*.
- [9] A. SELBERG, On elementary methods in prime number theory and their limitations, Den 11-te Skandinaviske Matematikerkongress, 1952. 13—22.

(Received February 1, 1967.)