ACTA ARITHMETICA
XLIII (1984)

On the local behaviour of some arithmetical functions

by

R. R. HALL (York) and G. TENENBAUM (Nancy)

1. Introduction. Since the work of Erdös–Kac, Erdös–Wintner, Kubilius, Delange, Wirsing, Elliott and Halász, the limiting distribution of additive and multiplicative functions is fairly well known. However the local behaviour of these functions still presents difficulties. In this paper our aim is to gain some insight into this behaviour, on short intervals. Given any (real-valued) arithmetical function f we introduce two further functions,

$$f_k^+(n) := \max\{f(n+j): 0 \leqslant j < k\},$$

$$f_k^-(n) := \min\{f(n+j): 0 \leqslant j < k\}$$

which reflect the local irregularity of the values of f. Here $k = k(n)$ is a function of n, possibly constant; the case when k behaves like a power of $\log n$ is important. Our methods would apply to many additive or multiplicative functions but it seemed preferable to us to fix our attention on the simplest case, when $f(p)$ is constant for primes p. So we consider, typically $\omega(n) := \operatorname{card}\{p: p|n,\ p \text{ prime}\}$, $\tau(n) := \operatorname{card}\{d: d|n,\ d \in \mathbf{Z}^+\}$.

The starting point of this study was a double conjecture of Erdös, proved in [1], on the functions $\tau_k^+(n)$ and $\tau_k^-(n)$. The first part of the conjecture was that for each fixed k there holds

$$\sum_{n \leqslant x} \tau_k^+(n) \sim kx\log x.$$

In [4] it was shown that this remains valid if $k = k(x) \to \infty$ in such a way that

$$k < (\log x)^{\log 4 - 1} \exp\{-\xi(x)\sqrt{(\log\log x)}\}$$

where $\xi(x)$ is any function tending to infinity. The exponent $\log 4 - 1$ is sharp. On seeing this result Erdös conjectured that when

$$k = (\log x)^{\log 4 - 1} \exp\{(c + o(1))\sqrt{(\log\log x)}\}$$

then

$$\sum_{n\leqslant x}\tau_k^+(n)\sim F(c)kx\log x$$

where F is an appropriate distribution function decreasing from 1 to 0 on $\boldsymbol{R}$. This is a corollary of our first theorem which we now state. Throughout the paper we denote by $\phi(u)$ the function defined on $\boldsymbol{R}^+$ by $\phi(u) = u\log u - u + 1$.

THEOREM 1. *Let* $\alpha\in(0,1]$ *and* $\beta = 2^\alpha$. *If we write*

$$k = (\log x)^{\phi(\beta)}\exp\{a\log\beta\sqrt{(2\beta\log\log x)}\}$$

where $a = a(x,k)$ *is unrestricted, then*

$$\sum_{n\leqslant x}(\tau_k^+(n))^\alpha = H(\alpha)kx(\log x)^{\beta-1}\left\{\frac{1}{\sqrt{\pi}}\int_a^\infty e^{-w^2}dw + O\left(\frac{\log\log\log x}{a\sqrt{(\log\log x)}}\right)\right\}$$

uniformly for $k\in\boldsymbol{Z}^+$, $x>x_0$, *and* $0<\alpha\leqslant 1$, *where*

$$H(\alpha) = \frac{1}{\Gamma(\beta)}\prod_p\left(1-\frac{1}{p}\right)^\beta\left(1+\frac{2^\alpha}{p}+\frac{3^\alpha}{p^2}+\ldots\right).$$

Here and throughout the paper x_0 is an absolute constant (10000 will do) for which our iterated logarithms are positive, and we assume $x>x_0$ henceforth.

The problem of $\tau_k^-(n)$ is harder and its mean value is not known for any $k\geqslant 2$. The second part of Erdös' conjecture was that for each fixed k,

$$\sum_{n\leqslant x}\tau_k^-(n)\ll_k x(\log x)^{\alpha_k}$$

where α_k decreases towards log2 as k increases. In fact $\alpha_k = k(2^{1/k}-1)$ is the sharp exponent: the right hand side above, if it be divided by $(\log\log x)^{11k^2}$, persists as a lower bound. We now improve this lower bound to what we believe be a sharp result, that is we conjecture that this is the correct order of magnitude.

THEOREM 2. *For* $k\geqslant 1$ *and* $x>x_0$ *we have*

$$\sum_{n\leqslant x}\tau_k^-(n)\gg_k x(\log x)^{\alpha_k}(\log\log x)^{(1-k)/2}.$$

We cannot prove that the left hand side is $o\left(x(\log x)^{\alpha_k}\right)$ but it will be seen that the proof we give would work equally well for the upper bound, except for technical difficulties mostly related to the sieve.

Regarding $\omega_k^\pm(n)$, the most interesting problem is the normal order, from which one can easily deduce the average order. We shall see later that the normal order of $\tau_k^\pm(n)$ may also be derived. In [2] Erdös and Kátai gave a quite complicated proof of the following result.

THEOREM (Erdös–Kátai). *For almost all* n, *the following asymptotic formulae hold uniformly for* $k\in\boldsymbol{Z}^+$.

$$\omega_k^+(n) = \big(1+o(1)\big)u^+(\lambda)\log\log n,$$

$$\omega_k^-(n) = \big(u^-(\lambda)+o(1)\big)\log\log n,$$

where λ *is such that* $k = (\log n)^\lambda$, *and* $u^+(\lambda)$ *(resp.* $u^-(\lambda)$*) denote the root* $\geqslant 1$ *(resp.* $\leqslant 1$*) of the equation* $\phi(u) = \lambda$ *(resp.* $\phi(u) = \min(1,\lambda)$*).*

In the last section of the paper we give a new and simple proof of this theorem based on a variance argument. This provides some extra information on the number of j, $0\leqslant j<k$ for which $\omega(n+j)$ is near its extrema.

As a consequence of the Erdös–Kátai theorem we have

THEOREM 3. *With the above notation we have for* $\lambda = O(1)$ *and almost all* n,

$$\tau_k^\pm(n) = (\log n)^{v^\pm(\lambda)+o(1)}$$

where $v^\pm = u^\pm\log 2$.

2. Some lemmas.

LEMMA 1. *Let* $x_0\geqslant x_1\geqslant\ldots\geqslant x_n\geqslant 0$, $y_0\geqslant y_1\geqslant\ldots\geqslant y_n\geqslant 0$. *Then*

$$(n+1)(x_0y_0+x_1y_1+\ldots+x_ny_n)\geqslant(x_0+x_1+\ldots+x_n)(y_0+y_1+\ldots+y_n).$$

This is a special case of an inequality of Chebyshev (cf. [5], p. 43).

LEMMA 2. *Let* $g:\boldsymbol{Z}^+\to\boldsymbol{R}^+$ *be multiplicative, and such that for each fixed prime* p, *the sequence* $\{g(p^\alpha): 0\leqslant\alpha<\infty\}$ *is decreasing.*

Let $G(n) = \sum\{g(d): d|n\}$. *Then for all* z,

$$\frac{1}{G(n)}\sum_{\substack{d|n\\ d<z}}g(d)\geqslant\frac{1}{\tau(n)}\sum_{\substack{d|n\\ d<z}}1. \tag{1}$$

Proof. This is by induction on $\omega(n)$, the number of distinct prime factors of n. Let $\omega(n) = 1$, $n = p^\alpha$. For $0\leqslant\beta\leqslant\alpha$ set

$$x_\beta = g(p^\beta),\qquad y_\beta = \begin{cases}1 & \text{if } p^\beta<z,\\ 0 & \text{else.}\end{cases}$$

Then (1) follows from Lemma 1 in this case. Now let $n = mp^\alpha$, $p\nmid m$, and suppose that (1) holds for all integers n' with $\omega(n')<\omega(n)$ and all z.

Since m is an n', we have

$$\frac{1}{G(n)}\sum_{\substack{d|n\\ d<z}} g(d) = \frac{1}{G(n)}\sum_{\beta=0}^{a} g(p^\beta)\sum_{\substack{t|m\\ t<zp^{-\beta}}} g(t)$$

$$\geqslant \frac{1}{G(p^a)}\sum_{\beta=0}^{a} g(p^\beta)\frac{1}{\tau(m)}\sum_{\substack{t|m\\ t<zp^{-\beta}}} 1 \geqslant \frac{1}{\tau(m)}\sum_{\substack{t|m\\ t<z}}\frac{1}{G(p^a)}\sum_{p^\beta<z/t} g(p^\beta)$$

$$\geqslant \frac{1}{\tau(m)}\sum_{\substack{t|m\\ t<z}}\frac{1}{\tau(p^a)}\sum_{p^\beta<z/t} 1 \geqslant \frac{1}{\tau(n)}\sum_{\substack{d|n\\ d<z}} 1.$$

Hence (1) holds for n and all z. This completes the induction and the result follows.

Lemma 3. *We have*

$$\sum_{n\leqslant x}\{\tau(n+i)\tau(n+j)\}^a \ll \frac{\sigma(|i-j|)}{|i-j|}\, x(\log x)^{2(2^a-1)}$$

uniformly for $0\leqslant a\leqslant 1$ *and* $0\leqslant i, j\leqslant x$, $i\neq j$.

Proof. We may write

$$\tau(n)^a = \sum_{d|n} g_a(d)$$

where g_a is a positive, multiplicative function such that $g_a(p^r) = (r+1)^a - r^a$. The function r^a is concave for each fixed $a\in[0,1]$, hence the sequence $\{g_a(p^r)\}_{r=0}^{\infty}$ is decreasing. We apply Lemma 2, with $g = g_a$, $G = \tau^a$, and $z = \sqrt{n}$. We have

$$\frac{1}{\tau(n)^a}\sum_{\substack{d|n\\ d\leqslant\sqrt{n}}} g_a(d) \geqslant \frac{1}{\tau(n)}\sum_{\substack{d|n\\ d\leqslant\sqrt{n}}} 1 \geqslant \frac{1}{2}.$$

Now set $z = \sqrt{(2x)} \geqslant \max\{\sqrt{(n+i)}, \sqrt{(n+j)}\}$ for every $n, i, j\leqslant x$. Then

$$\sum_{n\leqslant x}\{\tau(n+i)\tau(n+j)\}^a \leqslant 4\sum_{n\leqslant x}\sum_{\substack{d|n+i\\ d\leqslant z}}\sum_{\substack{d'|n+j\\ d'\leqslant z}} g_a(d)g_a(d')$$

$$\ll \sum_{\substack{d\leqslant z\\ (d,d')|i-j}}\sum_{d'\leqslant z} g_a(d)g_a(d')\left\{\frac{x}{[d,d']}+1\right\}$$

$$\ll x\sum_{\substack{d\leqslant z\\ (d,d')|i-j}}\sum_{d'\leqslant z}\frac{g_a(d)g_a(d')}{[d,d']} + z^2$$

since $g_a(d)\leqslant 1$. This is

$$\ll x\sum_{r|i-j}\frac{\varphi(r)}{r^2}\Big(\sum_{m\leqslant z/r}\frac{g_a(mr)}{m}\Big)^2 + x$$

$$\ll \frac{\sigma(|i-j|)}{|i-j|}\, x(\log x)^{2(2^a-1)}$$

since $g_a(mr)\leqslant g_a(m)$. This is the result stated.

Lemma 4. *We have*

$$\sum_{\substack{n\leqslant x\\ \omega(n)=k}}\tau(n)^a = H_a\Big(\frac{k-1}{2^a\log\log x}\Big)\frac{x}{\log x}\frac{(2^a\log\log x)^{k-1}}{(k-1)!}\Big(1+O\Big(\frac{k}{(\log\log x)^2}\Big)\Big)$$

uniformly for $a\in[0,1]$, $k\in\mathbf{Z}^+$, $x\geqslant 3$, *where*

$$H_a(z) = \frac{2^a}{\Gamma(2^a z+1)}\prod_p\Big(1-\frac{1}{p}\Big)^{2^a z}\Big(1+\frac{2^a z}{p}+\frac{3^a z}{p^2}+\dots\Big).$$

We omit the proof which follows that of Selberg [7].

Lemma 5. *We have*

$$\sum_{\substack{n\leqslant x\\ \omega(n)\leqslant l+c\sqrt{l}}}\tau(n)^a = H_a x(\log x)^{2^a-1}\Big(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c} e^{-v^2/2}dv + O\Big(\frac{1}{\sqrt{l}}\Big)\Big)$$

uniformly for $a\in[0,1]$, $c\in\mathbf{R}$ *where*

$$l = 2^a\log\log x,$$

$$H_a = H_a(1) = \frac{1}{\Gamma(2^a)}\prod_p\Big(1-\frac{1}{p}\Big)^{2^a}\Big(1+\frac{2^a}{p}+\frac{3^a}{p^2}+\dots\Big).$$

We omit the proof, which just involves Stirling's formula and the fact that $H_a'(z) = O(1)$ near $z = 1$.

3. Proof of Theorem 1. We put $l = 2^a\log\log x$ and

$$k = (\log x)^{2^a\log 2^a - 2^a + 1}\exp\{a\log 2^a\sqrt{(2l)}\},$$

$$M(x) = \varepsilon_a(x)(\log x)^{2^a\log 2}\exp\{a\log 2\sqrt{(2l)}\}.$$

We have

$$\sum_{n\leqslant x}\tau_k^+(n)^a \leqslant \sum_{n\leqslant x}\sum_{j<k}\chi(n+j)\tau(n+j)^a + \Big|\sum_{n\leqslant x}{}'\,\tau_k^+(n)^a$$

where $\sum'$ is restricted to integers n for which $\tau_k^+(n) \leqslant M(x)$ and χ denotes the characteristic function of the integers n such that $\tau(n) > M(x)$. Therefore the left hand side does not exceed

$$k \sum_{n \leqslant x+k} \chi(n)\tau(n)^\alpha + xM^\alpha(x).$$

We set $\varepsilon_a(x) = \exp\left(-\frac{1}{a}\log\log\log x\right)$ so that this does not exceed

$$k(S_1+S_2) + kx(\log x)^{2^\alpha-1}(\log\log x)^{-1}$$

where

$$S_1 = \sum_{n \leqslant x}\{\tau(n)^\alpha \colon 2^{\omega(n)} > M(x)\varepsilon_a(x)\},$$

$$S_2 = \sum_{n \leqslant x}\{\tau(n)^\alpha \colon \tau(n)/2^{\omega(n)} > \varepsilon_a^{-1}(x)\}.$$

The sum S_1 is treated by Lemma 5. We have

$$\omega(n) > l + a\sqrt{(2l)} - \frac{2}{\alpha\log 2}\log\log\log x$$

and so

$$S_1 = H_\alpha x(\log x)^{2^\alpha-1}\left(\frac{1}{\sqrt{\pi}}\int_a^\infty e^{-\omega^2}d\omega + O\left(\frac{\log\log\log x}{a\sqrt{(\log\log x)}}\right)\right).$$

Also

$$S_2 \leqslant \varepsilon_a(x)\sum_{n \leqslant x}\frac{\tau(n)^{\alpha+1}}{2^{\omega(n)}} \ll x(\log x)^{2^\alpha-1}(\log\log x)^{-1}.$$

Hence

$$\sum_{n \leqslant x}\tau_k^+(n)^\alpha \leqslant H_\alpha kx(\log x)^{2^\alpha-1}\left(\frac{1}{\sqrt{\pi}}\int_a^\infty e^{-\omega^2}d\omega + O\left(\frac{\log\log\log x}{a\sqrt{(\log\log x)}}\right)\right).$$

It remains to obtain a similar lower bound. Let χ now be the characteristic function of the integers n such that

$$\omega(n) > l + a\sqrt{(2l)} + \xi$$

where $\xi \to \infty$, $\xi = o\left(\sqrt{(\log\log x)}\right)$ is to be chosen later. Then

$$\sum_{n \leqslant x}\sum_{j<k}\chi(n+j)\tau(n+j)^\alpha \leqslant \sum_{n \leqslant x}\tau_k^+(n)^\alpha + \sum_{\substack{n \leqslant x \\ \lambda(n) \geqslant 2}}\sum_{j<k}\chi(n+j)\tau(n+j)^\alpha,$$

where

$$\lambda(n) = \operatorname{card}\{j \colon 0 \leqslant j < k,\ \chi(n+j) = 1\}.$$

The sum on the left is similar to S_1 above, indeed if we set $\xi = \frac{1}{\alpha\log 2}\times$ $\times\log\log\log x$ it has the same asymptotic formula. The second sum on the right does not exceed

$$2\sum_{0 \leqslant i<j<k}\chi(n+i)\chi(n+j)\tau(n+j)^\alpha \leqslant 2\sum_{0 \leqslant i<j<k}\frac{\tau(n+i)^\alpha\tau(n+j)^\alpha}{2^{\alpha l + \alpha a\sqrt{(2l)} + \alpha\xi}}$$

$$\ll \frac{1}{k(\log x)^{2^\alpha-1}\log\log x}\sum_{0 \leqslant i<j<k}\tau(n+i)^\alpha\tau(n+j)^\alpha.$$

By Lemma 3, the inner sum is

$$\ll k^2x(\log x)^{2(2^\alpha-1)}$$

and we deduce that

$$\sum_{n \leqslant x}\tau_k^+(n)^\alpha \geqslant H_\alpha kx(\log x)^{2^\alpha-1}\left(\frac{1}{\sqrt{\pi}}\int_a^\infty e^{-\omega^2}d\omega + O\left(\frac{\log\log\log x}{a\sqrt{(\log\log x)}}\right)\right)$$

as required. This completes the proof.

4. Proof of Theorem 2. This is a variation on the method worked out in [1]. We have

$$\sum_{n \leqslant x}\tau_k^-(n) \geqslant \sum_{v=0}^\infty 2^v T_k^*(x, v) \geqslant \sum_{v=1}^\infty 2^{v-1}T_k(x, v), \tag{2}$$

where

$$T_k^*(x, v) = \operatorname{card}\{n \leqslant x \colon \min_{0 \leqslant j<k}\omega(n+j) = v\},$$

$$T_k(x, v) = \operatorname{card}\{n \leqslant x \colon \min_{0 \leqslant j<k}\omega(n+j) \geqslant v\},$$

and we employ (2) in the range $[2^{1/k}\log\log x] \leqslant v \leqslant [2^{1/k}\log\log x] + \sqrt{(\log\log x)}$. Let $M \leqslant x^{1/3}$, $\mu(M) \neq 0$, $\omega(M) = kv$ and let $M = m_1m_2 \ldots m_k$ where $\omega(m_j) = v$ for each j. There exists N, $0 < N \leqslant M$, such that $N \equiv -j \pmod{m_j}$ for each j, and we put $N+j = a_jm_j$. For $l+1 \leqslant x/M$ put $q_j = q_j(l) = M_jl + a_j$ for each j, where $M_j = M/m_j$. Then $q_jm_j = Ml + N + j$ and $n = Ml + N$ is an integer contributing to $T_k(x, v)$.

We have to control the number of representations of n in the form described above: in fact we restrict M, l in such a way that the number of different representations in bounded. First, we suppose $(M, P(k)) = 1$ where $P(k)$ denotes the product of the primes not exceeding k. For each such prime p, let $\beta = \beta(p)$ denote the least integer such that $p^\beta > k$; thus β is bounded as a function of k, moreover so is $\prod\{p^\beta : p \leqslant k,\ \beta = \beta(p)\}$ which we denote by Q. Then we restrict l so that $Ml+N+j \not\equiv 0 \pmod{p^\beta}$ for any prime $p < k$. Notice that given M, N is determined uniquely and, as j varies, $0 \leqslant j < k$, just k residue classes $(\bmod\, p^\beta)$ are excluded for Ml. Since $p \nmid M$ and $p^\beta > k$, there is at least one residue class available for l. It follows by the Chinese Remainder Theorem that there is at least one residue class $(\bmod\, Q)$, say $l \equiv a \pmod Q$ such that the above condition is satisfied for each $p \leqslant k$. This means that for every j, $n+j$ has at most

$$b = \sum_{p \leqslant k} \big(\beta(p) - 1\big)$$

prime factors not exceeding k, counted according to multiplicity. Second, we suppose that every $q_j(l)$ is free of prime factors in the range $(k, z]$; here we use a sieve, as in [1], for a suitable z, which will be a small power of x, depending on k.

We may now estimate the number of representations of n in the form $Ml+N$. For each j, we have $n+j = q_j m_j$ where $q_j = q_j^- q_j^+$ say, and all the prime factors of q_j^- are $\leqslant k$, those of $q_j^+ > z$. The number of choices for q_j^-, q_j^+ do not exceed 2^b, 2^c respectively where b is given above and $c = \log(x+k)/\log z$. Hence the number of representations of n will not exceed $2^{k(b+c)}$. For each fixed k, this will be bounded.

It remains to construct our sieve. We suppose M given, $M \leqslant x^{1/3}$, $\mu(M) \neq 0$, $\big(M, P(k)\big) = 1$, $M = m_1 m_2 \ldots m_k$. Write $M_j = M/m_j$, and define

$$f(l) = \prod_{j=0}^{k-1} (M_j l + a_j) = \prod_{j=0}^{k-1} q_j(l),$$

$$F(r) = f(Qr+a) = \prod_{j=0}^{k-1} (M_j Q r + M_j a + a_j),$$

$$X = [(x/M - M)/Q] - 1, \qquad P = \prod\{p : k < p \leqslant z\}.$$

We seek a lower bound for

$$S = \operatorname{card}\{r : 1 \leqslant r \leqslant X,\ \big(F(r), P\big) = 1\}.$$

This is similar to the sieve in [1] (Lemma 9), and we do not give the details. The conclusion is that

$$S \geqslant C_1(k) X (\log X)^{-k} \qquad (X \geqslant X_0(k)),$$

where $z = X^{1/3k}$. Now $\log Q \ll k$, $M \leqslant x^{1/3}$ and so $X > \sqrt{x}$ for sufficiently

large x. Thus $c < 7k$ is bounded for each fixed k as required. Indeed

$$S \geqslant C_2(k) \frac{x}{M} (\log x)^{-k},$$

and

$$T_k(x, v) \geqslant C_3(k) \frac{x}{(\log x)^k} \sum_M \frac{\nu(M)}{M}$$

where $\nu(M)$ denotes the number of ways of writing $M = m_1 m_2 \ldots m_k$ with $\omega(m_j) = v$ for every j. Thus $\nu(M) = (kv)!/(v!)^k$ and

$$T_k(x, v) \gg_k \frac{x}{(\log x)^k} \frac{\big(\log\log x + O(1)\big)^{kv}}{(v!)^k}$$

whence by Stirling's formula, for $v = [2^{1/k}\log\log x] + r$, $r \leqslant \sqrt{(\log\log x)}$,

$$2^v T_k(x, u) \gg_k \frac{x e^{kv}}{l^{k/2} (\log x)^k} \left(1 + \frac{r}{2^{1/k}\log\log x}\right)^{-kv}$$
$$\gg_k \frac{x e^{k(v-r)}}{l^{k/2} (\log x)^k} \gg l^{-k/2} x (\log x)^{\alpha_k}.$$

Summing over r, we obtain the result stated.

5. Proof of Theorem 3 and new proof of the Erdös–Kátai Theorem. Theorem 3 readily follows from the result of Erdös and Kátai [2] stated in the introduction, since for almost all n, $(\log n)^{-\varepsilon} < \tau(n)/2^{\omega(n)} < (\log n)^\varepsilon$ for any fixed $\varepsilon > 0$. We now prove their theorem: we remark that the original result on this problem, due to Kátai [6], gave one-sided estimates: for almost all n and uniformly for $k \geqslant 1$, we have

$$\omega_k^+(n) \leqslant \big(1 + o(1)\big) u^+(\lambda) \log\log n,$$

$$\omega_k^-(n) \geqslant \big(u^-(\lambda) + o(1)\big) \log\log n$$

where $\lambda = \log k/\log\log n$. For completeness we also prove this in the appendix. Next, with the aim of reducing the problem, we note

(i) As $u^\pm(0) = 1$ and $\omega_k^-(n) \leqslant \omega(n) \leqslant \omega_k^+(n)$, in the case $\log k = o(\log\log n)$ the result follows from Hardy and Ramanujan theorem that the normal order of $\omega(n)$ is $\log\log n$.

(ii) Regarding $\omega_k^+(n)$, we may assume that $\log k \leqslant (\log\log n)^{\psi(n)}$ where $\psi(n)$ is any function tending to infinity with n. Else we have

$$\omega_k^+(n) \geqslant \omega_k^+(1) \sim \frac{\log k}{\log\log k} \sim u^+(\lambda) \log\log n.$$

(iii) Since $u^-(\lambda) = 0$ for $\lambda \geqslant 1$, we may assume $\lambda \leqslant 1$ in the case of $\omega_k^-(n)$.

We now need some more notation. First set $\omega(n, z) = \operatorname{card}\{p: p|n, p < z\}$ and define the corresponding functions $\omega_k^+(n, z)$, $\omega_k^-(n, z)$. Next, for each y, z define a multiplicative function $f(n, y)$ such that

$$\sum_{d|n} f(d, y) = y^{\omega(n,z)},$$

in fact $f(p^\nu, y) = y-1$ if $\nu = 1$ and $p < z$, $= 0$ else. We set

$$E(z) = \sum_{p<z} \frac{1}{p}, \quad M(y, z) = \prod_{p<z} \left(1 + \frac{y}{p}\right) = \sum_{d=1}^{\infty} \frac{f(d, y+1)}{d}$$

and for any pair (y, z) we let χ be the characteristic function of the sequence of integers n such that

$$|\omega(n, z) - yE(z)| \leqslant \xi(z) \sqrt{yE(z)}$$

where $\xi \to \infty$ as $z \to \infty$, $\xi(z) = o(\sqrt{\log\log z})$. For every n we set

$$F(n; y, k) = \sum_{j<k} \chi(n+j) y^{\omega(n+j,z)}.$$

Finally, we denote by $P^+(n)$ the largest prime factor of n. We state

THEOREM 4. *Provided* $k < x^{1/3}$, $y \geqslant y_0 > 0$, $y^3 \log z = o\left(\dfrac{\log x}{\log\log x}\right)$ *we have*

$$\sum_{n<x} \big(F(n; y, k) - kM(y-1, z)\big)^2 \leqslant xk^2 M^2(y-1, z) R,$$

where

$$R = R(x, y, z, k, b) \ll_b k^{-1/2} + (\log x)^{-b} + \\ + \exp\big((-\tfrac{1}{2} + o(1)) \xi^2(z)\big) + \frac{(\log x)^{\sigma(y)}}{k} \exp\{\xi(z)|\log y| \sqrt{(y \log\log x)}\}$$

uniformly (*except for the dependence on* b *implied by* $\ll_b$).

This is rather technical, but we check that it implies the result of Erdös and Kátai, taking the case of $\omega_k^+(n)$ as the other is easier.

Let η be a small positive constant and set $k_m = [(\log x)^{m\eta}]$, $m = 1, 2, 3, \ldots$; by the monotonicity of $\omega_k^+(n)$ as a function of k, the continuity of $u^+(\lambda)$ and remark (ii) above, it will be sufficient to prove that for all but $o(x)$ integers $n \leqslant x$ and for any $\varepsilon > 0$ we have

$$\omega_k^+(n) \geqslant (1-\varepsilon) u^+(m\eta) \log\log x,$$

where $k = k_m$ and $m \leqslant \exp\{(\log\log\log x)^2\}$. We choose z such that $\log z = (\log x) \exp\{-(\log\log\log x)^3\}$, and set $\xi(z) = (\log\log x)^{1/3}$, $y_m = (1-\varepsilon/2) u^+(m\eta)$ in our theorem. For almost all n, and all the m above, we see that

$$F(n; y_m, k_m) \sim k_m M(y_m - 1, z)$$

and so

$$\sum_{j<k_m} \chi(n+j) \geqslant k_m (\log x)^{-(1+o(1))\sigma(y_m)}.$$

The right hand side tends to ∞, which is all we need.

Now we embark on the proof of Theorem 4. We begin with

LEMMA 6. *Provided* $k < x^{1/3}$, $y \geqslant y_0 > 0$, $y^3 \log z = o(\log x/\log\log x)$, *we have*

$$\sum_{n \leqslant x} F(n; y, k) = kxM(y-1, z)\big\{1 + O\big(\exp\big(-\tfrac{1}{2} + o(1)\big)\xi^2(z)\big)\big\}. \tag{3}$$

Proof. We write

$$\sum_{n \leqslant x} F(n; y, k) = \sum_{n \leqslant x} \sum_{j<k} y^{\omega(n+j,z)} + O\Big(k \sum_{n \leqslant x} \big(1-\chi(n)\big) y^{\omega(n,z)}\Big)$$
$$= k \sum_{d \leqslant x} f(d, y)[x/d] +$$
$$+ O\Big(k \sum_{i=1}^{2} \sum_{n \leqslant x} (u_i y)^{\omega(n,z)} \exp\big\{-yE(z)\log u_i - \xi(z)|\log u_i| \sqrt{(yE(z))}\big\}\Big)$$

provided $u_1 \leqslant 1 \leqslant u_2$. For example, if $\omega(n, z) > yE(z) + \xi(z)\sqrt{(yE(z))}$ then

$$1 - \chi(n) = 1 \leqslant u_2^{\omega(n,z)} \exp\big\{-yE(z)\log u_2 - \xi(z) \log u_2 \sqrt{(yE(z))}\big\}.$$

The sums involving u_i may be estimated from above by the inequality of Halberstam and Richert [3], and we obtain

$$\sum_{n \leqslant x} F(n; y, k) = kxM(y-1, z) + O(k\{R_1 + R_2 + S_1 + S_2\})$$

with

$$R_1 = \sum_{d \leqslant x} |f(d, y)|, \quad R_2 = x \sum_{d > x} |f(d, y)|/d,$$

and for $i = 1, 2$,

$$S_i = xM(y-1, z) \prod_{p<z} \left(1 + \frac{(u_i - 1)y}{p+y-1}\right) \exp\big\{-yE(z)\log u_i - \\ - \xi(z)|\log u_i| \sqrt{(yE(z))}\big\}.$$

If we choose

$$u_i = 1+(-1)^i\xi(z)/\sqrt{(yE(z))}$$

then we find that

$$S_1+S_2 \ll xM(y-1, x)\exp\{(-\tfrac{1}{2}+o(1))\,\xi^2(z)\}.$$

Set $a = (\log z)^{-1}$. We also have

$$R_1+R_2 \leqslant x^{1-a}\sum_{d=1}^{\infty}\frac{|f(d,y)|}{d^{1-a}} \leqslant x^{1-a}\prod_{p<z}\left(1+\frac{e(y-1)}{p}\right)$$

$$\ll xM(y-1, z)\exp\{(-\tfrac{1}{2}+o(1))\,\xi^2(z)\}.$$

This completes the proof. Obviously the crucial step in the proof of the theorem lies in the evaluation of the second moment of $F(n; y, k)$.

LEMMA 7. *Provided* $k < x^{1/3}$, $y \geqslant y_0 > 0$, $y^3\log z = o(\log x\log\log x)$, *we have*

$$\sum_{n\leqslant x} F^2(n; y, k) = xk^2M^2(y-1, z)\{1+O(R)\}$$

where R is defined in the statement of the theorem.

Proof. Set

$$A(x; i, j) = \sum_{n\leqslant x} y^{\omega(n+i,z)+\omega(n+j,z)},$$

$$B(x; i, j) = \sum_{n\leqslant x}(1-\chi(n+i))y^{\omega(n+i,z)+\omega(n+j,z)},$$

$$C(x) = \sum_{n\leqslant x}\chi(n)y^{2\omega(n,z)}$$

so that the sum to be evaluated is

$$2\sum_{i<j<k}A(x; i, j)+O\Big(\sum_{i<j<k}B(x; i, j)+B(x; j, i)+kC(x)\Big)$$

and we note that

$$C(x) \leqslant y^{yE(z)+\operatorname{sgn}(y-1)\xi(z)\sqrt{(yE(z))}}\sum_{n\leqslant x}y^{\omega(n,z)}$$

$$\ll xM(y-1, z)\exp\{(y\log y)E(z)+|\log y|\,\xi(z)\sqrt{(yE(z))}\}$$

$$\ll xM^2(y-1, z)(\log x)^{\phi(y)}\exp\{|\log y|\,\xi(z)\sqrt{(y\log\log x)}\}.$$

With the same u_1, u_2 as before, set

$$G_l(x; i, j) = \sum_{n\leqslant x}(u_l y)^{\omega(n+i,z)}y^{\omega(n+j,z)}, \quad l = 1, 2$$

so that

$$B(x; i, j) \leqslant \sum_{l=1}^{2} G_l(x; i, j)\exp\{-yE(z)\log u_l - \xi(z)|\log u_l|\sqrt{(yE(z))}\}.$$

If η denotes the characteristic function of the sequence of integers n having a divisor d such that $d \geqslant x^{1/10}$, $P^+(d) < z$, $\mu(d) \neq 0$, we have

$$G_l(x; i, j) = \sum_{n\leqslant x}\sum_{\substack{d|n+i\\ d<x^{1/10}}}\sum_{\substack{d'|n+j\\ d'<x^{1/10}}} f(d, u_l y)f(d', y)+O(R_3) \tag{4}$$

where

$$R_3 = \sum_{n\leqslant x}\eta(n)(1+|u_l y-1|)^{\omega(n+i,z)}(1+|y-1|)^{\omega(n+j,z)}$$

and estimate the mean value of $\eta(n)$ by Rankin's method, viz:

$$\sum_{n\leqslant x}\eta(n) \leqslant x^{-a/10}\sum_{n\leqslant x}\sum_{\substack{d|n\\ P^+(d)<z}}|\mu(d)|d^a \leqslant x^{1-a/10}\prod_{p<z}{}'(1+p^{a-1})$$

with $a = 1/\log z$ as before. Hölder's inequality gives

$$R_3 \ll_b xM^2(y-1, z)(\log x)^{-b}.$$

Denote the main term in (4) by $G_l^*(x; i, j)$. Then

$$G_l^*(x; i, j) = \sum_{d,d'<x^{1/10}} f(d, u_l y)f(d', y)\operatorname{card}\{n \leqslant x: d|n+i, d'|n+j\}$$

$$= \sum_{s|j-i}\sum_{\substack{m,m'<x^{1/10}\\ (m,m')=1}} f(ms, u_l y)f(m's, y)\left\{\frac{x}{mm's}+O(1)\right\}$$

$$= \sum_{s|j-i}\frac{1}{s}\sum_{t<x^{1/10}s^{-1}}\mu(t)t^{-2}\sum_{d,d'<x^{1/10}(ts)^{-1}} f(dts, u_l y)f(d'ts, y)/dd'+O(x^{1/4}).$$

Let h be multiplicative, $h(p^\nu, y) = f(p^\nu, y)\left(1+\frac{y-1}{p}\right)^{-1}$. Then

$$\sum_{d=1}^{\infty}\frac{f(rd, y)}{d} = M(y-1, z)h(r, y).$$

We split off from the sum above terms with $ts \geqslant x^{1/20}$, and then add on all the terms with $ts < x^{1/20}$ but d or $d' \geqslant x^{1/10}/ts$. The contribution of both sets of terms is estimated by Rankin's method with $a = 1/\log z$, and we get

$$G_l^*(x;i,j) = xM(u_ly-1,z)M(y-1,z)\sum_{s|j-i}\frac{1}{s}\sum_{ts<x^{1/20}}\frac{\mu(t)}{t^2}h(ts,u_ly)h(ts,y)+$$
$$+O_b\big(xM^2(y-1,z)(\log x)^{-b}\big)$$

and so

$$\sum_{i<j<k} G_l(x;i,j) = xM(u_ly-1,z)M(y-1,z)K_l+O_b\big(xk^2M^2(y-1,z)(\log x)^{-b}\big)$$

where

$$K_l = \sum_{n=1}^{\infty}\frac{h(n,u_ly)h(n,y)}{n^2}\sum_{s|n}s\mu(n/s)\sum_{r<k}\left[\frac{r}{s}\right]$$

because

$$\sum_{\substack{i<j<k\\ s|j-i}}1 = \sum_{r<k}\left[\frac{r}{s}\right].$$

We consider two cases according as $y\leqslant 2$ or not. In the first case, the inner sum is $k^2/2s+O(k)$ and the term involving k^2 vanishes when $2\leqslant n<k$. If we estimate the remaining terms by Rankin's method we get

$$K_l = \tfrac{1}{2}k^2+O(k^{3/2}).$$

If $y>2$, $h(n,y)$ and $h(n,u_ly)$ are non-negative and an upper bound for K_l is available, indeed

$$K_l = \sum_{s=1}^{\infty}s\sum_{r<k}\left[\frac{r}{s}\right]\sum_{m=1}^{\infty}\frac{\mu(m)h(ms,u_ly)h(ms,y)}{m^2s^2}$$
$$\leqslant\prod_p\left(1-\frac{h(p,u_ly)h(p,y)}{p^2}\right)\sum_{r<k}r\sum_{s=1}^{\infty}\frac{g(s)}{s^2} = \tfrac{1}{2}k(k-1),$$

g being the multiplicative function such that

$$g(p^\nu) = \frac{h(p^\nu,u_ly)h(p^\nu,y)}{1-h(p^\nu,u_ly)h(p^\nu,y)p^{-2}}.$$

We have now shown that for all y satisfying our hypotheses, we have

$$\sum_{i<j<k}G_l(x;i,j)\leqslant\tfrac{1}{2}k^2xM(u_ly-1,z)M(y-1,z)\big\{1+O_b\big(k^{-1/2}+(\log x)^{-b}\big)\big\}$$

whence, recalling the definition of u_1 and u_2,

$$\sum_{i<j<k}B(x;i,j)\ll xk^2M^2(y-1,z)\exp\big\{\big(-\tfrac{1}{2}+o(1)\big)\xi^2(z)\big\},$$

a similar estimate being valid for $B(x;j,i)$. If we substitute $u_l=1$ in the above we also obtain

$$\sum_{i<j<k}A(x;i,j)\leqslant\tfrac{1}{2}k^2xM^2(y-1,z)\big\{1+O_b\big(k^{-1/2}+(\log x)^{-b}\big)+$$
$$+\exp\big(\big(-\tfrac{1}{2}+o(1)\big)\xi^2(z)\big)\big\}$$

and so we have

$$\sum_{n\leqslant x}F^2(n;y,k)\leqslant k^2xM^2(y-1,z)\big(1+O(R)\big)$$

and the opposite inequality also holds — we need only apply Cauchy's inequality to the first moment. This completes the proof of Lemma 7, and together with Lemma 6 this gives Theorem 4.

Appendix. In this section we prove the one-sided estimates of Kátai. We may assume $k<n$, else $\omega_k^-(n)=1$ and

$$\omega_k^+(n)\leqslant\omega_k^+(1)\sim(\log k)/\log\log k\sim u^+(\lambda)\log\log n,\quad \text{as } n\to\infty.$$

Next, fix A so large that for $z=x^{1/A}$, we have $E(z)\leqslant\log\log x$. Since $\omega(n)-\omega(n,z)\leqslant A$ we need only consider $\omega_k^+(n,z)$, $\omega_k^-(n,z)$. For $y>0$ and $k<x$ we have

$$(5)\qquad \sum_{n\leqslant x}y^{\omega_k^{\pm}(n,z)}\leqslant\sum_{n\leqslant x}\sum_{j<k}y^{\omega(n+j,z)}\ll kx(\log x)^{y-1}.$$

Recall the definition of the numbers k_m. Since $\omega_k^{\pm}$ are monotonic functions of k and $u^{\pm}(\lambda)$ are continuous, we need only prove that for each fixed $\varepsilon>0$ we have $N^+(x)+N^-(x)=o(x)$, where

$$N^+(x) = \operatorname{card}\left\{n\leqslant x:\exists m<\frac{\log x}{\eta\log\log x}:\omega_{k_m}^+(n,z)>(1+\varepsilon)u^+(\eta m)\log\log x\right\}$$

and

$$N^-(x) = \operatorname{card}\left\{n\leqslant x:\exists m<\frac{\log x}{\eta\log\log x}:\omega_{k_m}^-(n,z)<(1-\varepsilon)u^-(\eta m)\log\log x\right\}.$$

We apply (5) with $k=k_m$, $y=y_m^+$, y_m^- where $y_m^+=(1+\varepsilon)u^+(\eta m)$, $y_m^-=(1-\varepsilon)u^-(\eta m)$. This gives the result stated.

References

[1] P. Erdös and R. R. Hall, *Values of the divisor function on short intervals*, J. Number Theory 12 (1980), pp. 176–187.
[2] P. Erdös and I. Kátai, *On the growth of some additive functions on small intervals*, Acta Math. Hungar. 33 (1979), pp. 345–359.
[3] H. Halberstam and H.-E. Richert, *On a result of R. R. Hall*, J. Number Theory 11 (1979), pp. 76–89.
[4] R. R. Hall, *The maximum value of the divisor function on short intervals*, Bull. London Math. Soc. 13 (1981), pp. 121–124.
[5] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge 1934.
[6] I. Kátai, *Local growth of the number of prime divisors of consecutive integers*, Publ. Math. Debrecen 18 (1971), pp. 171–175.
[7] A. Selberg, *Note on a paper by L. G. Sathe*, J. Indian Math. Soc. 18 (1954), pp. 83–87.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF YORK
York, Y01 5DD, England
UNIVERSITÉ DE NANCY I
UER SCIENCES MATHÉMATIQUES
Boîte Postale 239
54506 Vandoeuvre les Nancy, France

Received on 21.10.1982
and in revised form on 17.1.1983 (1323)

ACTA ARITHMETICA
XLIII (1984)

The joint distribution of the binary digits of integer multiples

by

JOHANNES SCHMID (Ulm)

1. Problem and results. Each number $n \in N_0$ has a unique binary expansion

$$n = \sum_{\nu=0}^{\infty} d_\nu 2^\nu, \quad d_\nu \in \{0, 1\}.$$

We consider the "digital sum", i.e. the total

$$B(n) := \sum_{\nu=0}^{\infty} d_\nu$$

of digits 1 in this expansion.

Elementarily, $B(n)$ is binomially distributed and hence its approximate distribution is given by the central limit theorems, e.g. in the simplest form:

$$\frac{1}{2^m} \# \left\{0 \leqslant n < 2^m; \frac{B(n) - m/2}{\sqrt{m/4}} \leqslant \xi\right\} \xrightarrow[(m \to \infty)]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-t^2/2} dt \quad \text{for } \xi \in \boldsymbol{R}.$$

It is not difficult to show (elementarily) that the right hand term describes the limiting distribution of $B(kn)$, $k \in N$, too.

Thence, it should be interesting to know the joint distribution of $\big(B(k_1 n), B(k_2 n), \ldots, B(k_s n)\big)$ (,where, w.l.o.g., the k_ν are different and odd).

A special question of this type was asked by Stolarsky and Muskat [9] who obtained the upper estimate:

$$\# \{1 \leqslant n < x; B(kn) - B(n) = a\} < c_1(k) \frac{x}{\sqrt{\log x}} + c_2(k) \frac{x}{\log x}$$

with explicit but rather large $c_1(k)$, $c_2(k)$. Further, they conjectured tha