On the iteration of multiplicative functions

By I. KÁTAI (Budapest)

Dedicated to Professor Zoltán Daróczy on his 50th birthday

1. Consider a completely multiplicative function $\theta(n)$ such that for each prime p,

$$\theta(p) = p+a \tag{1.1}$$

for a fixed integer $a>0$. Let K_a be the set of primes p such that, for some integer n, p occurs infinitely often as a divisor of the sequence of iterates $\theta^{(k)}(n)$, where $\theta^{(k)}(n)=\theta(\theta^{(k-1)}(n))$. This set K_a becomes a directed graph, with its elements as nodes, if we connect a prime $p\in K_a$ to each of the primes which divide $p+a$ (clearly these are also in K_a). These graphs K_a were introduced in [1], and it was proved that this set is finite, with largest element less than $2b^*a$, where b^* is the smallest prime not dividing a.

A more detailed study of these graphs has been given by R. M. POLLACK, H. N. SHAPIRO and G. H. SPARER in their very interesting paper [2]. They proved that the largest prime in K_a is at most $b^*+(b^*-1)a$. Primes greater than a occur in K_a only on strings $p\to p+a\to p+2a\to\dots$ out of a prime $p\leqq a$, where $p, p+a, \dots$ are primes. They have considered the directed infinite graph U_a in which all the primes are nodes, and a directed edge goes from the prime p to the prime q if and only if q divides $p+a$. They observed furthermore that K_a is the minimal complete core U_a (Theorem 2.1 in [2]).

Let $\varkappa(n)$ be the minimal nonnegative integer h for which all the prime divisors of $\theta^{(h)}(n)$ belong to K_a, where $\theta^{(0)}(n)=n$. In terms of the graphs U_a, $\varkappa(n)$ is the maximum over all primes p dividing n of the length of the path from p to K_a. It is obvious that $\varkappa(n)=\max_{q|n}\varkappa(q)$, and $\varkappa(q)=1+\max_{p|q+a}\varkappa(p)$ if $\varkappa(q)\neq 0$, where p, q run over the primes satisfying the conditions $p|q+a$, $q|n$, respectively. In [1] we proved that

$$\limsup_{n\to\infty}\varkappa(n)(\log\log n)^{-1} > 0 \tag{1.2}$$

and in [2] it was proved that

$$\varkappa(n) = \mathcal{O}(\log n), \tag{1.3}$$

for all fixed a.

To determine the exact maximal order of $\varkappa(n)$, or even the exact order of $\sum_1^N \varkappa(n)$

seems to be a very hard question. In this paper we shall prove that for almost all integers n, $\varkappa(n)>\eta \log\log n$, where η is a positive constant that may depend on a.

For a fixed a, let $\lambda(p)$ denote the length (i.e. the number of nodes) of the shortest path from the node p to K_a. In [2] it was proved that $\lambda(p)=\mathcal{O}(\log\log p)$ and the conjecture was stated that $\lambda(p)=\mathcal{O}((\log\log p)(\log\log\log p)^{-1})$ holds for almost all primes p. We can prove easily that much more is true concerning this last conjecture, namely that

(1.4) for almost all primes p, $\lambda(p)=1$ or 2.

Let $\mathscr{P}_1=\{q|\lambda(q)=1\}$, $\mathscr{P}_2=\{q|\lambda(q)\geqq 3\}$.

Let $Q_1, \ldots, Q_r$ be the set of the primes in K_a coprime to a. It is clear that this set is non-empty. Since $q\in\mathscr{P}_1$ if and only if $q\notin K_a$ and $q\equiv -a(\mathrm{mod}\, Q_j)$ for at least one $j\in\{1, \ldots, r\}$, from the prime number theorem we get that

$$\frac{1}{\Pi(x)}\#\{q\leqq x|\lambda(q)=1\}\to 1-\prod_{j=1}^{r}\left(1-\frac{1}{Q_j}\right)=\delta(>0). \tag{1.5}$$

Therefore the relative density of $\mathscr{P}_1$ (in the set of primes) is positive. Let us estimate the number of those primes p, for which $\lambda(p)\geqq 3$. If $\lambda(p)\geqq 3$, then $(q, p+a)=1$ for each $q\in\mathscr{P}_1$. But by using sieve theorems (see Lemma 1) we get that

$$\#\{p\leqq x|(p+a, q)=1\ \forall q\in\mathscr{P}_1\}\ll \pi(x)\frac{1}{(\log x)^{\delta}},$$

which implies (1.4). We have proved moreover, that

$$\frac{1}{\pi(x)}\#\{q\leqq x|\lambda(q)=1\}\to\prod_{j=1}^{r}\left(1-\frac{1}{Q_j}\right)\quad(x\to\infty). \tag{1.6}$$

2. In this section we shall estimate $\varkappa(n)$. We shall state Theorem 4.2 ([3]) as

Lemma 1. *Let $F(n)$ be a polynomial of degree $g(\geqq 1)$ with integer coefficients. Then, for any set $\mathscr{B}$ of primes,*

$$\#\{p: p\leqq x|(F(p), \mathscr{B})=1\}\leqq c_g\prod_{\substack{p<x\\ p\in\mathscr{B}}}\left(1-\frac{\varrho(p)}{p}\right)\prod_{\substack{p<x\\ p\in\mathscr{B}\\ p\nmid F(0)}}\left(1-\frac{1}{p}\right)^{-1}\frac{x}{\log x},$$

where c_g depends only on g, and $\varrho(p)$ denotes the number of solutions of $F(n)\equiv\equiv\mathcal{O}(\mathrm{mod}\, p)$.

We shall apply this Lemma for $F(p)=p+a$, under the condition $(\mathscr{B}, a)=1$. So we have

$$\#\{p\leqq x|(p+a, \mathscr{B})=1\}\leqq c\prod_{\substack{p<x\\ p\in\mathscr{B}}}\left(1-\frac{1}{p}\right)\frac{x}{\log x}. \tag{2.1}$$

Let $\eta>1$, z_1 be given, $z_n=z_1^{\eta^n}$, $\zeta=\log\eta$, $y_n=z_{2n-1}$, $x_n=z_{2n}$. The values η, z_1 will be determined later. $\mathscr{A}_1, \mathscr{A}_2, \ldots$ are some finite subsets of primes, defined as follows:

(1) for all n the elements of $\mathscr{A}_n$ belong to (y_n, x_n),

(2) $\mathscr{A}_1$ is an arbitrary set,
(3) $p\in\mathscr{A}_{n+1}$ if and only if there exists $q\in\mathscr{A}_n$, such that $q|p+a$.

Let $S_n=\sum_{p\in\mathscr{A}_n} 1/p$, $T_n=\prod_{p\in\mathscr{A}_n}\left(1-\frac{1}{p}\right)$. It is clear that

$$T_n < e^{-S_n}(\leqq 1). \tag{2.2}$$

Lemma 2. *Let a be fixed. Then there exist positive numbers z^*, θ, such that for every $\mathscr{A}_1$ satisfying $z_1>z^*$, $T_1\leqq\theta$, the inequality*

$$T_n \leqq \theta \tag{2.3}$$

holds.

PROOF. It is enough to prove (2.3) for $n=2$. Let $x\geqq y_2$. From Lemma 1 we get

$$\sum_{\substack{p\in\mathscr{A}_2\\ x<p\leqq 2x}} 1/p \geqq \sum_{x<p\leqq 2x}\frac{1}{p}\sum_{\substack{x<p\leqq 2x\\ p\notin\mathscr{A}}} 1 \geqq \sum_{x<p\leqq 2x}\frac{1}{p}-2cT_1(\log x)^{-1}.$$

Applying this inequality with $x=2^k y_2$ $(k=0, 1, \ldots, k_0)$, where k_0 is the largest integer for which $2^{k_0+1}y_2\leqq x_2$, we get that

$$S_2 \geqq \sum_{y_2<p\leqq 2^{k_0+1}y_0}\frac{1}{p}-2cT_1\sum_{k=0}^{k_0}\frac{1}{(\log 2^k y_2)}.$$

Since the first sum is $\log\frac{\log x_2}{\log y_2}+\mathcal{O}\left(\frac{1}{\log y_2}\right)=\zeta+\mathcal{O}\left(\frac{1}{\eta^3\log z_1}\right)$, and the second is $\zeta+\mathcal{O}\left(\frac{1}{\log y_2}\right)=\zeta+\mathcal{O}\left(\frac{1}{\eta^3\log z_1}\right)$, with an absolute constant in the order terms, we get

$$S_2 \geqq (1-2cT_1)\zeta-\frac{c^*}{\eta^3\log z_1}, \tag{2.4}$$

with an absolute constant c^*. If z_1 is so chosen that $c^*/(\eta^3\log z_1)<1/2$, then from (2.2), (2.4) we get that

$$T_2 \leqq e^{-(1-2cT_1)\zeta}\left(1+\frac{2c^*}{\eta^3\log z_1}\right) < 2\cdot e^{-(1-2cT_1)\zeta}. \tag{2.5}$$

Let now be $\theta=\frac{1}{4c}$, ζ so large that $\zeta>\log 64$, and $4c\leqq e^{\zeta/3}$. Then the condition $T_1\leqq\theta$ implies that $T_2\leqq\theta$. Indeed, from (2.5) we get

$$T_2 < 2\cdot e^{-\zeta/2} < e^{-\zeta/3} \leqq \theta.$$

By this the proof of our lemma is completed.

Theorem 1. *Let a be a positive integer, $\varkappa(n)$ defined as above. Then there exists a constant $\lambda>0$ such that*

$$\frac{1}{\pi(x)}\#\{p\leqq x|\varkappa(p)<\lambda\log\log p\}\to 0 \quad \text{as} \quad x\to\infty.$$

PROOF. Let $\mathscr{A}_1$ be such a subset of primes for which the conditions of Lemma 2 are satisfied, and assume that $z_1 > b^* + (b^*-1)a$. The last condition guarantees that $\varkappa(p) \geqq 1$ for all the elements of $\mathscr{A}_1$. Then for each n, $\varkappa(p) \geqq n$ if $p \in \mathscr{A}_n$. This is obvious from the definition of $\mathscr{A}_n$ and from the inequality $\varkappa(p) \geqq 1 + \varkappa(q)$ if $q | p+a$ and $\varkappa(p) \neq 0$.

Let now X be a large number, and assume that $z_n < x \leqq z_{n+1}$. Let $b \geqq 2$ be a fixed integer, $k \geqq 1$. Then by Lemmas 1 and 2 for $\mathscr{B} = \mathscr{B}_{b,k} = \bigcup_{l=0}^{k} \mathscr{A}_{n-b-l}$ we have

$$\#\{p | p \in [X, 2X], (p, \mathscr{B}_{b,k}) = 1\} \leqq \tag{2.6}$$

$$\leqq c T_{n-b} \dots T_{n-b-k} \frac{x}{\log x} \leqq c\theta^k \frac{x}{\log x}.$$

If $(p, \mathscr{B}_{b,k}) > 1$, then $\varkappa(p) \geqq n-b-k$. Let k be so large that $c\theta^k < 1/4$. Since

$$n \geqq \frac{1}{2} \frac{\log\log X}{\log \eta},$$

with a suitable constant $\lambda_1 > 0$ we have

$$\#\{p | p \in [X, 2X], \varkappa(p) > \lambda_1 \log\log p\} > \frac{1}{2} \frac{x}{\log x}, \tag{2.7}$$

say.

Let now

$$\mathscr{B} = \mathscr{B}_x = \{q | w < q < x,\ \varkappa(q) > \lambda_1 \log\log w\},$$

where $w = e^{(\log x)^{1/2}}$. If p is such a prime for which there exists q with $q | p+a$, then $\varkappa(p) \geqq \lambda_1 \log\log w \geqq \frac{\lambda_1}{2} \log\log x$. By Lemma 1,

$$\#\{p \leqq x | (p+a, \mathscr{B}_x) = 1\} < c \prod_{q \in \mathscr{B}} \left(1 - \frac{1}{q}\right) \frac{x}{\log x}.$$

Since $\Pi\left(1 - \frac{1}{q}\right)$ in the right hand side tends to zero, Theorem 1 has been proved. ∎

3. Let $h(a)$ denote the smallest prime which does not belong to K_a, and let $H(a)$ be the largest prime which belongs to K_a.

It is clear that K_a contains all prime divisors of a. In [2] it was proved that $2 \in K_a$ for each $a \in \mathbf{N}$.

Let $Q_1 < Q_2 < \dots < Q_R$ be the elements of K_n, and assume that $p \notin K_n$. It is clear that if Q is a prime, $Q | Q_j + n$, then $Q \in K_n$. Consequently $p \nmid Q_j + n$. Let $n \equiv l \pmod p$. If there is a prime π such that $\pi | n + Q_i$ for at least one Q_i $(i = 0, 1, \dots, R)$, $Q_0 = 0$, and $\pi \equiv -l \pmod p$, then $p \in K_n$. Indeed, if $\pi | n + Q_i$, then $\pi \in K(n)$, since $p | \pi + n$, therefore $p \in K_n$.

Consequently, if $n \equiv l \pmod p$, then the numbers $n, n+Q_1, \dots, n+Q_R$ do not contain prime divisors from the arithmetic progression $\equiv -l \pmod p$.

By using sieve results (see [3]) we get that

$$\#\{n\in[X, 2X]|n\equiv s(\mathrm{mod}\, p), \pi\nmid n \quad \text{if} \quad \pi\equiv -l(\mathrm{mod}\, p)\}\leqq \tag{3.1}$$

$$\leqq c\frac{x}{p}\prod_{\pi\equiv -l(\mathrm{mod}\, p)}\left(1-\frac{1}{\pi}\right)\leqq c\frac{x}{p(\log x)^{1/(p-1)}}$$

if $(s, p)=(l, p)=1$, $p<x$.

Let

$$\mathscr{B}_p=\{n|\exists l \bmod p,\ (l, p)=1,\ \pi\nmid n \text{ if } \pi\equiv l(\mathrm{mod}\, p)\}.$$

From (3.1) it follows that

$$\#\{n\in[X, 2X]|n\in\mathscr{B}_p\}\leqq c\frac{x}{(\log x)^{1/(p-1)}}. \tag{3.2}$$

Summing (3.2) for all the primes $p\leqq(\log\log x)(\log\log\log x)^{-1}$, the right hand side is $o(x)$. Consequently we have

Theorem 2. *For almost all n,*

$$h(n)\geqq\frac{\log\log n}{\log\log\log n}. \tag{3.3}$$

By using the Turán—Kubilius inequality, hence we easily get the following

Theorem 3. *Let $\eta>0$ be an arbitrary positive constant. Then*

$$H(n)\geqq n^{1/2-\eta} \tag{3.4}$$

holds for almost all n.

Proof. Let us consider the integers n in $\left[\frac{x}{2}, x\right]$. Let $q_1<q_2<\ldots<q_s$ be all the primes less than $(\log\log x)(\log\log\log x)^{-1}$. Let $S(n)=\prod_{j=1}^{s}(n+q_j)$, let $\mathscr{P}$ be the set of all the primes in $[x^{1/2-\eta}, x^{1/2}]$, and

$$f(n):=\sum_{\substack{p\mid S(n)\\ p\in\mathscr{P}}}1.$$

Then, by the Turán—Kubilius inequality we have

$$\sum_{(x/2)\leqq n\leqq x}(f(n)-sA)^2\leqq cxsA,$$

$$A=\sum_{p\in\mathscr{P}}\frac{1}{p}=\log\frac{\log 1/2}{\log\left(\frac{1}{2}-\eta\right)}+\mathcal{O}\left(\frac{1}{\log x}\right)$$

with a suitable constant c. Hence we get that

$$\sum_{\substack{f(n)=0\\ (x/2)\leqq n\leqq x}}1=o(x)\quad(x\to\infty).$$

If for an $n\in\left[\frac{x}{2}, x\right]$ (3.3) holds and $f(n)>0$, then $K(n)$ contains all the q_i, and consequently all the primes p that divide $S(n)$. But $p|S(n)$, $p\in\mathscr{P}$ implies that $p>x^{1/2-\eta}$. This completes the proof of Theorem 3.

From the Turán—Kubilius inequality we can get an estimate for the cardinality of K_n as well. Let $\mathscr{P}^*$ be the set of the primes in (q_s, x),

$$g(n)=\sum_{\substack{p|S(n)\\ p\in\mathscr{P}^*}} 1, \quad B:=\sum_{p\in\mathscr{P}^*}\frac{1}{p}=\log\log x-\log\log q_s+o(1).$$

Now

$$\sum_{n\leqq x}(g(n)-sB)^2\leqq cxBs, \quad s=(\log\log x)(\log\log\log x)^{-1},$$

therefore

$$\#\{n\leqq x, g(n)<sB-w\sqrt{sB}\}\leqq\frac{cx}{w}.$$

If $q_1, \ldots, q_s\in K_n$, then $p|S(n)$, and $p\in\mathscr{P}^*$ belongs also to K_n. So we have

Theorem 4. *Let $\varepsilon>0$ be an arbitrary positive constant. Then*

$$\operatorname{card}(K_n)>(1-\varepsilon)(\log\log n)^2(\log\log\log n)^{-1}$$

holds for almost all n.

References

[1] I. KÁTAI, Some problems on the iteration of multiplicative number-theoretical functions, *Acta Math. Acad. Sci. Hung.*, **19** (1968), 441—450.

[2] R. M. POLLACK, H. N. SHAPIRO and G. H. SPARER, On the graphs of I. Kátai, *Communications on Pure and Applied Mathematics*, **27** (1974), 669—713.

[3] H. HALBERSTAM and H. E. RICHERT, Sieve methods, *London* 1974, *Academic Press*.

I. KÁTAI
EÖTVÖS LORÁND UNIVERSITY
COMPUTER CENTER
BUDAPEST, H—1117
BOGDÁNFY U. 10/B

(Received January 27, 1987)