Annales Univ. Sci. Budapest., Sect. Comp. 27 (2007) 187-215

THE DISTRIBUTION OF AN ADDITIVE ARITHMETICAL FUNCTION ON THE SET OF SHIFTED INTEGERS HAVING k DISTINCT PRIME FACTORS

L. Germán (Budapest, Hungary)

To the memory of Professor Matukumalli Venkata Subbarao

Abstract. It is proven that an additive arithmetical function has a limit law on the set of shifted integers having k prime factors if and only if the Erdős-Wintner condition holds.

1. Notations and results

We shall use the following notations:

$\mathcal{P}=$ set of prime numbers; p, q with or without suffixes always denote prime numbers, $\omega(n)$ denote the number of distinct prime factors of n. $P(n)$, $p(n)$ denote the largest and the smallest prime divisor of n respectively. c denotes a constant, not certainly the same at different locations.

Let $\mathcal{P}_k=\{n : \omega(n)=k\}$, $\mathcal{P}_k(x)=\mathcal{P}_k\cap[1..x]$, and $\Pi_k(x)=\#\mathcal{P}_k(x)$. $\pi(x)=\#\{p\le x \mid p\in\mathcal{P}\}$,

$$\pi(x,k,l)=\sum_{\substack{p\le x\\ p\equiv l \pmod k}} 1.$$

Financially supported by the Hungarian National Foundation for Scientific Research under grant OTKA T46993 and partially by the DFG Project on the "Rates of convergence in limit theorems".

Throughout this paper π_k denote integers having k distinct prime factors.

Let F_x be a sequence of distribution functions, and let F be a distribution function. We say that F_x converges weakly to F, in notation $F_x \Rightarrow F$, if $\lim_{x\to\infty} F_x(z) = F(z)$ holds at all continuity points of F.

Let T be an assertion, A be a subset of $\mathbb{N}$ and let

$$\nu_x(n \in A \,:\, T(n)) := \frac{1}{|A \cap [1,x]|} \#\{n \le x, n \in A \,:\, T(n) \text{ holds true}\}.$$

Let f be a real additive function. We say that f possess a limit law on A, if for the sequence of distribution functions

$$F_x(z) = \nu_x(n \in A \,:\, f(n) \le z),$$

there is a distribution function say F, such that $F_x \Rightarrow F$.

Erdős and Wintner [3] proved that an additive arithmetical function f possesses a limit law on $\mathbb{N}$ if and only if the Erdős-Wintner condition holds, i.e. if and only if the three series

$$\sum_{|f(p)|>1} \frac{1}{p}, \qquad \sum_{|f(p)|\le 1} \frac{f(p)}{p}, \qquad \sum_{|f(p)|\le 1} \frac{f^2(p)}{p} \tag{1.1}$$

converge.

Kátai [6] proved about 40 years ago that the convergence of the 3 series (1.1) implies the existence of the limit distribution of f on the set $\mathcal{P}+1$.

The necessity of the convergence of (1.1) has been proved by Hildebrand [5] about 20 years ago.

The aim of this paper is to prove the following

Theorem 1. *Let f be a real additive function. Let $2 \le k \le \epsilon(x) \cdot \cdot \sqrt{\log\log x}$, $\epsilon(x) \to 0$ $(x \to \infty)$. Let*

$$F_{k,x}(z) = \nu_x(n \in \mathcal{P}_k \,:\, f(n+1) \le z).$$

Assume that there is a sequence $k = k_x$ and a distribution function F such that $F_{k_x,x} \Rightarrow F$. Then the 3 series given in (1.1) are convergent.

Conversely, assume that the series in (1.1) are convergent. Then with a distribution function G

$$\max_{2\le k\le \epsilon(x)\sqrt{\log\log x}} |F_{k,x}(y) - G(y)| \to 0, \quad (x \to \infty)$$

if y is a continuity point of G. Consequently $F = G$.

The characteristic function of F is given by

$$\varphi(t) = \prod_p (1 + h(p)),$$

where

$$h(p) = -\frac{1}{p-1} + \sum_{m=1}^{\infty} \frac{e^{itf(p^m)}}{p^m}. \tag{1.2}$$

Remark. We know from the theorem of Paul Lévy that F is of pure type, it is continuous if and only if

$$\sum_{f(p)\neq 0} \frac{1}{p}$$

diverges.

One can conclude the sufficiency part of Theorem 1 from the next

Theorem 2. *Let $1 \geq \epsilon > 0$ and $\varrho = \min\{\epsilon/4, 1/4\}$. Let f be a real additive function and assume that*

$$\sum_{|f(p)|>1} \frac{1}{p}, \qquad \sum_{|f(p)|\leq 1} \frac{f^2(p)}{p} \tag{1.3}$$

are convergent.

Let

$$A(x) := \sum_{\substack{|f(p)|\leq 1 \\ p\leq x}} \frac{f(p)}{p}, \tag{1.4}$$

$$a(m) := \sum_{\substack{|f(p)|\leq 1 \\ p|m}} \frac{f(p)}{p}. \tag{1.5}$$

Let $K_D(x) = \{Dp + 1 \leq x \mid p \in \mathcal{P}\}$,

$$F_{D,x}(y) := \nu_x \left(n \in K_D(x), f(n) - \left(A\left(\left(\frac{x-1}{D} \right)^{\varrho} \right) - a(D) \right) \leq z \right), \tag{1.6}$$

with characteristic function $\varphi_{D,x}(t)$, *and let*

$$\varphi_D(t) := \prod_{\substack{|f(p)|>1 \\ p\nmid D}} (1+h(p)) \prod_{\substack{|f(p)|\le 1 \\ p\nmid D}} (1+h(p))\, e^{-it\frac{f(p)}{p}}. \tag{1.7}$$

Then

$$\max_{1\le D\le x^{1-\epsilon}} |\varphi_{D,x}(t)-\varphi_D(t)| \to 0 \quad (x\to\infty) \tag{1.8}$$

uniformly for all bounded values of t, *i.e. if* $|t|<T$, T *is an arbitrary constant.*

2. Preliminaries

We need analogues of some-well known theorems related to prime numbers. In the proof of Theorem 1 it is allowed to drop not more than $o(\Pi_k(x))$ elements. By this the limit distribution function does not change.

Let

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}, \qquad p_1 < p_2 < \ldots < p_k$$

and

$$\pi_j = \pi_j(n) = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_j^{\alpha_j} \qquad (j=1,\ldots,k).$$

Let

$$\gamma_j(n) = \frac{\log p_{j+1}}{\log \pi_j} \quad (j=1,\ldots,k-1),$$

$$\Delta(n) = \min_{j=1,2,\ldots,k-1} \gamma_j(n).$$

First we investigate some sets, which are unimportant in our case. We have the following

Lemma 1. *Let* $\epsilon(x)\to 0$ *slowly. Let* $M=M_x$ *be defined by* $\log\log M = \sqrt{\log\log(x)}$. *Let* $2\le k\le \epsilon(x)\sqrt{\log\log x}$. *Then there exists a sequence* $A_x\to\infty$ $(x\to\infty)$ *such that*

1.

$$\Delta(n)\ge A_x,$$

and

2. $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 1$ and $p_1 \geq M$ holds for all but $\delta(x)\,\Pi_k(x)$ element of $\mathcal{P}_k(x)$, where $\delta(x) \to 0$. Let $U_k(x)$ be the set of those elements of $\mathcal{P}_k(x)$, for which these conditions hold true.

Proof. The following sets have zero relative density in $\mathcal{P}_k$:

1. $A_1 = \{n \in \mathcal{P}_k,\, n \leq x : \exists p^2 | n\}$. We have

$$\#A_1 \leq \sum_{\substack{p^\alpha \leq x^{1/2} \\ \alpha \geq 2}} \Pi_{k-1}\left(\frac{x}{p^\alpha}\right) + \sum_{\substack{p^\alpha > x^{1/2} \\ \alpha \geq 2}} \frac{x}{p^\alpha} \ll$$

$$\ll \Pi_k(x) \frac{k}{\log\log x} \sum_{\substack{p^\alpha \leq x^{1/2} \\ \alpha \geq 2}} \frac{1}{p^\alpha} + \mathcal{O}\left(x^{3/4}\right).$$

Here we used
$$\frac{\Pi_{k-1}(x)}{\Pi_k(x)} \sim \frac{k}{\log\log x}$$
which is a direct consequence of

$$\Pi_k(x) = \frac{x}{\log x} \frac{\log\log^{k-1} x}{(k-1)!} \left(1 + \mathcal{O}\left(\frac{1}{\log\log x}\right)\right) \tag{2.1}$$

(see e.g in [7]).

2. $A_2 = \{n \in \mathcal{P}_k,\, n \leq x : p(n) < M\}$. We have

$$\#A_2 \leq$$

$$\leq \sum_{\substack{p^\alpha \leq x^{1/2} \\ p < M}} \Pi_{k-1}\left(\frac{x}{p^\alpha}\right) + \sum_{\substack{p^\alpha > x^{1/2} \\ \alpha \geq 2}} \frac{x}{p^\alpha} \ll \Pi_k(x) \frac{k}{\log\log x} \sum_{p<M} \frac{1}{p} + \mathcal{O}\left(x^{3/4}\right).$$

By means of these last two steps we can assume that $p(n) > M$, and n is squarefree.

Observe that, if $n \in U_k(x)$ and $pq|n$, $p < q$, then $p^{A_x} < q$, which is an immediate consequence of the assumption $\Delta(n) \geq A_x$. The number of integers $n \in \mathcal{P}_k(x)$ for which $n \notin A_1$ or $n \notin A_2$ or there are prime divisors $p < q$, such that $q < p^{A_x}$ is not more than

$$\sum_{\substack{pq\pi_{k-2} \leq x \\ M<p<q<p^{A_x} \\ p<x^{1/2A_x}}} 1 + \sum_{\substack{pq\pi_{k-2} \leq x \\ M<p<q<p^{A_x} \\ p \geq x^{1/2A_x}}} 1.$$

The first sum is at most

$$\sum_{\substack{M<p<q<p^{A_x}\\ p<x^{1/2A_x}}} \Pi_{k-2}\left(\frac{x}{pq}\right) \ll \frac{x}{\log x}\frac{(\log\log x)^{k-3}}{(k-3)!}\sum_{\substack{M<p<q<p^{A_x}\\ p<x^{1/2A_x}}}\frac{1}{pq} \ll$$

$$\ll \Pi_k(x)\frac{k^2}{\log\log x}\log A_x,$$

so if A_x tends to infinity suitably slowly, then this last expression is $\Pi_k(x)\,o(1)$.

For the second sum we use the sieve estimation for the number of integers not exceeding x, for which $P(n) > z$. The number of these integers is smaller than

$$\frac{x}{\log z} + c\frac{z^2}{\log^2 z}$$

if $z > z_0$. This estimation can be found for example in [4] (Theorem 3.6). Using this, we get that the second sum does not exceed

$$\sum_{\substack{pq\pi_{k-2}\le x\\ M<p<q<p^{A_x}\\ p>x^{1/2A_x}}} 1 \ll$$

$$\ll \sum_{\pi_{k-2}<x^{1-\frac{1}{A_x}}}\ \sum_{\substack{n\le\frac{x}{\pi_{k-2}}\\ P(n)>x^{1/2A_x}}} 1 \ll A_x \sum_{\pi_{k-2}<x^{1-\frac{1}{A_x}}} \frac{x}{\pi_{k-2}\log x} + A_x^2\frac{x^{1/A_x}}{\log^2 x},$$

which is $\Pi_k(x)\,o(1)$ with appropriate A_x. Here we used that

$$\sum_{\pi_{k-1}\le x}\frac{1}{\pi_{k-1}} = \frac{\log\log^{k-1} x}{(k-1)!}(1+o(1)).$$

The proof is ready.

Let us introduce the notations

$$B_k(x) := \#U_k(x)$$

and

$$B_k(x,d,l) = \sum_{\substack{n\in U_k(x)\\ n\equiv l \pmod d}} 1\,.$$

We shall prove some inequalities for the distribution of the elements of $U_k(x)$ in arithmetic progressions.

We have

Lemma 2. *Let $x \geq 2$, $d < \log^A x$, where A is a fixed positive number, and let $(d,l) = 1$. We have*

$$B_k(x,d,l) = \frac{B_k(x)}{\varphi(d)}\left(1 + \mathcal{O}\left(e^{-c\sqrt{\log x}}\right)\right)$$

uniformly for $2 \leq k \leq \epsilon(x)\sqrt{\log\log x}$, where $\epsilon(x) \to 0$ $(x \to \infty)$.

Proof. Let S_x be the set of those π_{k-1}, for which there exists at least one prime $p > P(\pi_{k-1})$ such that $\pi_{k-1}p \in U_k(x)$. Let $p^* = p_{\pi_{k-1}}$ be the smallest p with this property. Then $\pi_{k-1}p \in U_k(x)$ for all $p^* \leq p \leq \dfrac{x}{\pi_{k-1}}$. Let $\pi_{k-1}p \equiv l (\text{mod } d)$. Then, using Lemma 1, $\pi_{k-1} < x^\lambda$, with an appropriate $\lambda < 1/4$, $P(\pi_{k-1}) < p$ and $p(\pi_{k-1}) > \log^A x$, when x is larger than x_A say. With this conditions we have $(\pi_{k-1}, d) = 1$. We get $p \equiv l_{\pi_{k-1}} (\text{mod } d)$ with a unique $l_{\pi_{k-1}}$ $(\text{mod } d)$.

We have that

$$\left|B_k(x,d,l) - \frac{B_k(x)}{\varphi(d)}\right| \ll \tag{2.2}$$

$$\sum_{\pi_{k-1}\in S_x}\left|\left\{\pi\left(\frac{x}{\pi_{k-1}}, d, l_{\pi_{k-1}}\right) - \pi\left(P(\pi_{k-1}), d, l_{\pi_{k-1}}\right)\right\} -\right.$$

$$\left. -\frac{1}{\varphi(d)}\left\{\pi\left(\frac{x}{\pi_{k-1}}\right) - \pi\left(P(\pi_{k-1})\right)\right\}\right|.$$

The Theorem of Siegel-Walfisz for prime numbers is applicable for

$$\sum_{\pi_{k-1}\in S_x}\left|\pi\left(\frac{x}{\pi_{k-1}}, d, l_{\pi_{k-1}}\right) - \frac{1}{\varphi(d)}\pi\left(\frac{x}{\pi_{k-1}}\right)\right|.$$

We get that the right hand side of (2.2) does not exceed

$$\frac{c}{\varphi(d)}\sum_{\pi_{k-1}\in S_x}\frac{x}{\pi_{k-1}}e^{-c\sqrt{\log x}} + c\sum_{\pi_{k-1}\in S_x}P(\pi_{k-1}) \ll$$

$$\ll \frac{1}{\varphi(d)}xe^{-c\sqrt{\log x}}\frac{(\log\log x)^{k-1}}{(k-1)!} + \mathcal{O}\left(x^{1/2}\right),$$

because $\sum_{\pi_{k-1}<x} \frac{1}{\pi_{k-1}} \ll \frac{(\log\log x)^{k-1}}{(k-1)!}$.

Lemma 3. *Let $x \geq 2$, $d < x^a$ with a fixed $0 \leq a < 1$. Then*

$$B_k(x,d,l) < c(a)\frac{B_k(x)}{\varphi(d)},$$

if $(d,l)=1$ and $2 \leq k \leq \epsilon(x)\sqrt{\log\log x}$, where $\epsilon(x) \to 0$ $(x \to \infty)$. Here $c(a)$ depends only on a.

Proof. With the notations of Lemma 2 we have

$$B_k(x,d,l) \leq c \sum_{\pi_{k-1}\in S_x} \pi\left(\frac{x}{\pi_{k-1}}, d, l_{\pi_{k-1}}\right).$$

Applying the Brun-Titchmarsh Theorem, the above sum does not exceed

$$c\frac{x}{\varphi(d)\log x} \sum_{\pi_{k-1}\in S_x} \frac{1}{\pi_{k-1}} \ll \frac{B_k(x)}{\varphi(d)}.$$

Let

$$B_k(x|d) = \sum_{\substack{n\in U_k(x)\\ (n,d)=1}} 1\,.$$

Lemma 4. *Let $x \geq 2$, $A > 0$ an arbitrary number and $\alpha < 1/2$. For $2 \leq k \leq \epsilon(x)\sqrt{\log\log x}$, where $\epsilon(x) \to 0$ $(x \to \infty)$, we have*

$$\sum_{d<x^\alpha} \max_{(l,d)=1} \max_{z\leq x} \left| B_k(z,d,l) - \frac{B_k(z|d)}{\varphi(d)} \right| \ll B_k(x)\log^{-A} x.$$

The constant implied by $\ll$ does not depend on k.

Proof. With the notations of Lemma 2 we have for $x > x_0$ that

$$\left| B_k(x,d,l) - \frac{B_k(x|d)}{\varphi(d)} \right| \leq$$

$$\leq \sum_{\substack{\pi_{k-1}\in S_x\\ (\pi_{k-1},d)=1}} \left| \pi\left(\frac{x}{\pi_{k-1}}, d, l_{\pi_{k-1}}\right) - \frac{\pi\left(\frac{x}{\pi_{k-1}}|d\right)}{\varphi(d)} \right| + \mathcal{O}\left(x^{1/3}\right).$$

Applying the Bombieri-Vinogradov theorem we get that the right hand side of this last inequality does not exceed

$$c\frac{x}{\log x}\log^{-A}x\sum_{\pi_{k-1}\in S}\frac{1}{\pi_{k-1}}+\mathcal{O}\left(x^{2/3}\right)\ll B_k\left(x\right)\log^{-A}x.$$

Here we used, that $\pi(x)-\pi(x|d)\leq\omega(d)\leq\log x$.

A more general version of this lemma was established by Wolke and Zhang [9].

Lemma 5. (Wolke-Zhang) *For any given $A>0$ and $0<\epsilon<1/2$ there exist $\eta>0$ such that*

$$\sum_{d\leq x^{1/2-\epsilon}}\max_{(a,d)=1}\max_{y\leq x}\left|\Pi_k\left(y,d,a\right)-\frac{\Pi_k\left(y|d\right)}{\varphi\left(d\right)}\right|\ll\Pi_k\left(x\right)\log^{-A}x$$

holds uniformly for $k\leq\eta\log x/\log\log^2x$, and the constant implied in $\ll$ depends on A and ϵ only.

Corollary 1. *Let $0\leq\eta<1/4$. We have*

$$\sum_{x^\eta<q<x^{2\eta}}B_k\left(x,q,-1\right)=\sum_{x^\eta<q<x^{2\eta}}\frac{B_k\left(x\right)}{q-1}+\mathcal{O}\left(\frac{B_k\left(x\right)}{\log^A x}\right).$$

Proof. From Lemma 4 we have

$$\sum_{x^\eta<q<x^{2\eta}}B_k\left(x,q,-1\right)=\sum_{x^\eta<q<x^{2\eta}}\frac{B_k\left(x|q\right)}{q-1}+\mathcal{O}\left(\frac{B_k\left(x\right)}{\log^A x}\right).$$

On the other hand

$$B_k(x)\geq B_k\left(x|q\right)\geq B_k\left(x\right)-\sum_{l=1}^{\infty}\sum_{\substack{aq^l<x\\ a\in P_{k-1}(x)\\ (a,q)=1}}1\geq B_k\left(x\right)+\mathcal{O}\left(\frac{x}{q}\right).$$

Putting it together

$$\sum_{x^\eta<q<x^{2\eta}}\frac{B_k\left(x|q\right)}{q-1}=\sum_{x^\eta<q<x^{2\eta}}\frac{B_k\left(x\right)}{q-1}+\mathcal{O}\left(\sum_{x^\eta<q<x^{2\eta}}\frac{x}{q^2}\right)=$$

$$=\sum_{x^\eta<q<x^{2\eta}}\frac{B_k\left(x\right)}{q-1}+\mathcal{O}\left(x^{1-\eta}\right).$$

3. Proof of Theorem 2

We shall investigate the sequence of the characteristic functions. Let y_x be tending to infinity so slowly that $y_x \le \frac{1}{2}\log\log\log\log x$, and $x \ge c(\epsilon)$, say. Let us define the additive function

$$f_0(p^\alpha) = \begin{cases} f(p^\alpha) & \text{if } p^\alpha \le y_x, \\ 0 & \text{else,} \end{cases}$$

and consider the following distribution function:

$$G_{x,D}(z) := \nu_x\left(n \in K_D(x), f_0(n) - (A(y_x) - a(D)) \le z\right).$$

Then the characteristic function of $G_{x,D}$ is

$$\psi_{x,D}(t) = \frac{1}{\#K_D(x)} e^{-it(A(y_x)-a(D))} \sum_{n \in K_D(x)} e^{itf_0(n)}. \tag{3.1}$$

We define furthermore g_x by

$$g_x = \mu * e^{itf_0}, \tag{3.2}$$

where μ is the Mbius function. We have that $g_x(p^\alpha) = e^{itf_0(p^\alpha)} - e^{itf_0(p^{\alpha-1})}$, and by the inequality $\pi(y_x) < 2\frac{y_x}{\log y_x}$ we get that $g_x(n) = 0$ for $n > e^{2y_x}$, i.e. if $n > \log\log\log x$. These together imply that

$$\sum_{n=1}^{\infty} \frac{g_x(n)}{\varphi(n)} = \tag{3.3}$$

$$= \prod_{p \le y_x}\left(1 + \sum_{\alpha \ge 1} \frac{e^{itf_0(p^\alpha)} - e^{itf_0(p^{\alpha-1})}}{p^{\alpha-1}(p-1)}\right) = \prod_{p \le y_x}\left(1 - \frac{1}{p-1} + \sum_{\alpha \ge 1} \frac{e^{itf_0(p^\alpha)}}{p^\alpha}\right).$$

It is clear that

$$\sum_{n \in K_D(x)} e^{itf_0(n)} = \sum_{\substack{d \le x \\ (D,d)=1}} g_x(d)\pi\left(\frac{x-1}{D}, d, l_d\right), \tag{3.4}$$

where l_d is defined by $Dl_d + 1 \equiv 0 (\mathrm{mod}\ d)$. Since $g_x(d) = 0$ if $d > \epsilon \log x$, we can apply the Siegel-Walfisz Theorem, whence the right hand side of (3.4) is

$$\pi\left(\frac{x-1}{D}\right) \sum_{\substack{d=1\\(D,d)=1}}^{\infty} \frac{g_x(d)}{\varphi(d)} + \mathcal{O}\left(\pi\left(\frac{x-1}{D}\right) e^{-c\sqrt{\log x}} \sum_{\substack{d=1\\(D,d)=1}}^{\infty} \frac{|g_x(d)|}{\varphi(d)}\right).$$

In the error term

$$\sum_{\substack{d=1\\(D,d)=1}}^{\infty} \frac{|g_x(d)|}{\varphi(d)} \leq$$

$$\leq \prod_{p \leq y_x} \left(1 + \sum_{\alpha=1}^{\infty} \frac{|g_x(p^\alpha)|}{\varphi(p^\alpha)}\right) \leq \prod_{p \leq y_x} \left(1 + \frac{2p}{(p-1)^2}\right) \ll \log\log\log x,$$

say. Furthermore

$$\sum_{\substack{d=1\\(d,D)=1}}^{\infty} \frac{g_x(d)}{\varphi(d)} = \prod_{\substack{p \leq y_x\\ p \nmid D}} \left(1 - \frac{1}{p-1} + \sum_{\alpha \geq 1} \frac{e^{itf_0(p^\alpha)}}{p^\alpha}\right) = (1 + o_x(1)) \prod_{\substack{p \leq y_x\\ p \nmid D}} (1 + h(p)),$$

where the implied constant in $o_x(1)$ is absolute, it does not depend on t, and $h(p)$ is defined by (1.2). Thus we obtain, that

$$\psi_{x,D}(t) = \prod_{\substack{p \leq y_x\\ |f(p)|>1\\ p \nmid D}} (1 + h(p)) \prod_{\substack{p \leq y_x\\ |f(p)| \leq 1\\ p \nmid D}} (1 + h(p)) e^{-it\frac{f(p)}{p}} + o_x(1)$$

uniformly in $t \in \mathbb{R}$, and the constant implied by $o_x(1)$ is absolute.

Now using the convergence of (1.3), one can easily see that

$$\max_{1 \leq D \leq x^{1-\epsilon}} |\psi_{x,D}(t) - \varphi_D(t)| \to 0 \quad (x \to \infty) \tag{3.5}$$

uniformly for all $|t| < T$, where $\varphi_D(t)$ is given by (1.7).

Using Lemma 1.11 in [1] it follows immediately, that $G_{x,D} \Rightarrow F_D$ for all fixed D.

Next we define another two additive functions as follows: let

$$f_1(p^\alpha) = \begin{cases} f(p) & \text{if } y_x < p \leq \left(\frac{x-1}{D}\right)^\varrho,\ \alpha = 1 \text{ and } |f(p)| \leq 1, \\ 0 & \text{else,} \end{cases}$$

and let

$$f_2(p^\alpha) = \begin{cases} f(p) & \text{if } \left(\frac{x-1}{D}\right)^{\varrho} < p \le \left(\frac{x-1}{D}\right)^{1-\vartheta_x},\ \alpha = 1 \text{ and } |f(p)| \le 1, \\ 0 & \text{else,} \end{cases}$$

where $\varrho < \min(\epsilon/4, 1/4)$ and ϑ_x is tending to zero slowly.

First we show that

$$\nu_x(n \in K_D(x),\ f(n) \ne f_0(n) + f_1(n) + f_2(n)) = o(1)\pi\left(\frac{x-1}{D}\right) \tag{3.6}$$

with an appropriate ϑ_x.

To do this, let $B = \{q \in \mathcal{P} \mid |f(q)| > 1\}$, and $B_y = \{q \in \mathcal{P} \mid q > y,\ q \in B\}$. Then

$$\#\{n \in K_D(x) \mid \exists q | n,\ q \in B_{y_x}\} = \delta(y_x)\pi\left(\frac{x-1}{D}\right), \tag{3.7}$$

where $\delta(y_x) \to 0$ $(x \to \infty)$. To see this we split the numbers in the set in (3.7) into two, not necessary distinct parts. In the first part we take numbers which have a prime divisor q with $y_x < q \le \left(\frac{x-1}{D}\right)^{\varrho}$. The other part contains the numbers, which have a prime divisor q such that $\left(\frac{x-1}{D}\right)^{\varrho} < q \le \left(\frac{x-1}{D}\right)^{1-\vartheta_x}$. Let us denote the number of integers in this two parts by Σ_1 and Σ_2, respectively.

We have that

$$\Sigma_1 \le \sum_{Dp+1 \le x} \sum_{\substack{q | Dp+1 \\ y_x < q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ |f(q)| > 1}} 1 \le \sum_{\substack{y_x < q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ |f(q)| > 1}} \pi\left(\frac{x-1}{D}, q, l_q\right),$$

and similarly

$$\Sigma_2 \le \sum_{\substack{\left(\frac{x-1}{D}\right)^{\varrho} < q \le \left(\frac{x-1}{D}\right)^{1-\vartheta_x} \\ |f(q)| > 1}} \pi\left(\frac{x-1}{D}, q, l_q\right).$$

Using the Brun-Titchmarsh theorem for Σ_1 we get,

$$\Sigma_1 \le c\pi\left(\frac{x-1}{D}\right) \sum_{\substack{y_x < q \\ |f(q)| > 1}} \frac{1}{q}.$$

Using sieve estimates for Σ_2 we get,

$$\Sigma_2 \le c \sum_{\substack{\left(\frac{x-1}{D}\right)^{\varrho} < q \le \left(\frac{x-1}{D}\right)^{1-\vartheta_x} \\ |f(q)|>1}} \frac{x-1}{qD\log\frac{x-1}{qD}} \le c\frac{1}{\vartheta_x}\frac{x-1}{D\log\left(\frac{x-1}{D}\right)} \sum_{\substack{\left(\frac{x-1}{D}\right)^{\varrho} < q \\ |f(q)|>1}} \frac{1}{q}.$$

With the choice

$$\vartheta_x^2 = \max\left(\sum_{\substack{\left(\frac{x-1}{D}\right)^{\varrho} < q \\ |f(q)|>1}} \frac{1}{q}, \sum_{\substack{\left(\frac{x-1}{D}\right)^{\varrho} < q \\ |f(q)|\le 1}} \frac{f^2(q)}{q} \right), \tag{3.8}$$

we get $\Sigma_1 + \Sigma_2 = o(1)\pi\left(\frac{x-1}{D}\right)$.

The next assertion holds:

$$\nu_x\left(n \in K_D(x); \exists q|n\ ,\ q > \left(\frac{x-1}{D}\right)^{1-\vartheta_x} \right) = o(1)\pi\left(\frac{x-1}{D}\right), \tag{3.9}$$

which can be obtained using sieve estimates. We have that the left hand side of (3.9) is at most

$$\sum_{Dp+1\le x} \sum_{\substack{Dp+1=aq \\ \left(\frac{x-1}{D}\right)^{1-\vartheta_x} < q}} 1 \le \sum_{a \le \left(\frac{x-1}{D}\right)^{\vartheta_x}} \sum_{Dp+1=aq} 1,$$

which does not exceed

$$c\sum_{a<\left(\frac{x-1}{D}\right)^{\vartheta_x}} \frac{x}{aD\log^2\frac{x}{aD}} \le c\frac{x}{D\log^2\frac{x}{D}} \sum_{a<\left(\frac{x-1}{D}\right)^{\vartheta_x}} \frac{1}{a},$$

which is $o(1)\pi\left(\frac{x-1}{D}\right)$.

Similarly

$$\#\{n \in K_D(x) \mid \exists q^2|n,\ q>y\} \le \tag{3.10}$$

$$\le \sum_{y<q<\left(\frac{x-1}{D}\right)^{a}} \pi\left(\frac{x-1}{D}, q^2, l_q\right) + \frac{x-1}{D} \sum_{q\ge\left(\frac{x-1}{D}\right)^{a}} \frac{1}{q^2} = \delta(y)\pi\left(\frac{x-1}{D}\right),$$

where $\delta(y) \to 0$ $(y\to\infty)$.

(3.7) and (3.9) and (3.10) imply (3.6).

Let

$$A_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)=\sum_{\substack{y_x<p\le\left(\frac{x-1}{D}\right)^{\varrho}\\ |f_1(p)|\le 1\\ p\nmid D}}\frac{f_1(p)}{p}.$$

Next we prove a Turn-Kubilius type inequality, namely that

(3.11)

$$\sum_{Dp+1\le x}\left|f_1(Dp+1)-A_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\right|^2\le c\pi\left(\frac{x-1}{D}\right)\sum_{\substack{y_x<p\\ |f(p)|\le 1}}\frac{f^2(p)}{p},$$

and

$$\sum_{Dp+1\le x}f_2^2(Dp+1)\le c\pi\left(\frac{x-1}{D}\right)\sum_{\substack{\left(\frac{x-1}{D}\right)^{\varrho}<p\\ |f(p)|\le 1}}\frac{f^2(p)}{p}. \tag{3.12}$$

Proof of (3.11). Let

$$B^2\left(\left(\frac{x-1}{D}\right)^{\varrho}\right):=\sum_{\substack{y_x<q\le\left(\frac{x-1}{D}\right)^{\varrho}\\ q\nmid D}}\frac{f_1^2(q)}{q}.$$

Let

$$\sum_{Dp+1\le x}\left|f_1(Dp+1)-A_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\right|^2=$$

$$=S_1-2A_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)S_2+A_{D,y_x}^2\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right),$$

where

$$S_1=\sum_{Dp+1\le x}\left(\sum_{q|Dp+1}f_1(q)\right)^2=\sum_{\substack{y_x<q\le\left(\frac{x-1}{D}\right)^{\varrho}\\ q\nmid D}}f_1^2(q)\,\pi\left(\frac{x-1}{D},q,l_q\right)+$$

$$+\sum_{\substack{y_x\le q\le\left(\frac{x-1}{D}\right)^{\varrho}\\ q\nmid D}}\ \sum_{\substack{y_x\le q'\le\left(\frac{x-1}{D}\right)^{\varrho}\\ q'\nmid D\\ q\ne q'}}f_1(q)\,f_1(q')\,\pi\left(\frac{x-1}{D},qq',l_{qq'}\right)=\Sigma_{11}+\Sigma_{12},$$

and

$$S_2 = \sum_{\substack{y_x < q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q \nmid D}} f_1(q)\, \pi\left(\frac{x-1}{D}, q, l_q\right).$$

Since $\varrho < \epsilon/4$, thus we can estimate Σ_{11} using the Brun-Titchmarsh theorem, and we get

$$\Sigma_{11} < cB^2\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right).$$

Moreover Σ_{12} equals

$$\pi\left(\frac{x-1}{D}\right)\sum_{\substack{y_x \le q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q \nmid D}}\ \sum_{\substack{y_x \le q' \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q' \nmid D \\ q \ne q'}} \frac{f_1(q)\, f_1(q')}{\varphi(qq')} +$$

$$+\sum_{\substack{y_x \le q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q \nmid D}}\ \sum_{\substack{y_x \le q' \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q' \nmid D \\ q \ne q'}} f_1(q)\, f_1(q')\left(\pi\left(\frac{x-1}{D}, qq', l_{qq'}\right) - \frac{\pi\left(\frac{x-1}{D}\right)}{\varphi(qq')}\right) =$$

$$= \zeta + E$$

such that

$$\Sigma_{12} \le A^2_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right) + E.$$

An application of the Cauchy-Schwarz inequality shows that E^2 is at most

$$\sum_{\substack{y_x < q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q \nmid D}}\ \sum_{\substack{y_x < q' \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q' \nmid D \\ q \ne q'}} \frac{f_1^2(q)\, f_1^2(q')}{\varphi(q)\,\varphi(q')} \times$$

$$\times \sum_{\substack{y_x < q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q \nmid D}}\ \sum_{\substack{y_x < q' \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q' \nmid D \\ q \ne q'}} \varphi(qq')\left|\frac{\pi\left(\frac{x-1}{D}\right)}{\varphi(qq')} - \pi\left(\frac{x-1}{D}, qq', l_{qq'}\right)\right|^2.$$

Since $qq' < \left(\frac{x-1}{D}\right)^{1/2}$, the Brun-Titchmarsh theorem is applicable, and we get that E is at most

$$B^2\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\left(\pi\left(\frac{x-1}{D}\right)\right)^{1/2} \times$$

$$\times \left(\sum_{\substack{y_x < q \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q \nmid D}} \; \sum_{\substack{y_x < q' \le \left(\frac{x-1}{D}\right)^{\varrho} \\ q' \nmid D \\ q \ne q'}} \left| \frac{\pi\left(\frac{x-1}{D}\right)}{\varphi(qq')} - \pi\left(\frac{x-1}{D}, qq', l_{qq'}\right) \right| \right)^{1/2}.$$

The Bombieri-Vinogradov theorem is also applicable, and we get

$$\Sigma_{12} \le$$

$$\le A^2_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right) + \mathcal{O}(1)B^2\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right)\log x^{-A},$$

where A is an arbitrary big positive constant. We have

$$S_1 = A^2_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right) + \mathcal{O}(1)B^2\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right).$$

To estimate S_2 we note, that using the Cauchy- Schwarz inequality, we have

$$A_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right) =$$

$$= \sum_{\substack{y_x < p \le \left(\frac{x-1}{D}\right)^{\varrho} \\ p \nmid D}} \frac{f_1(p)}{p} \le \left(\sum_{\substack{y_x < p \le \left(\frac{x-1}{D}\right)^{\varrho} \\ p \nmid D}} \frac{f_1^2(p)}{p} \sum_{\substack{y_x < p \le \left(\frac{x-1}{D}\right)^{\varrho} \\ p \nmid D}} \frac{1}{p} \right)^{1/2} \ll$$

$$\ll B\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)(\log\log x)^{1/2},$$

such that using the above method one can easily see, that

$$S_2 = 2A^2_{D,y_x}\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right) + o(1)\,B^2\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\pi\left(\frac{x-1}{D}\right),$$

which implies (3.11).

Proof of (3.12). Since a positive integer $n \le x$ can have only a bounded number of distinct prime divisors $q > \left(\frac{x-1}{D}\right)^{\varrho}$, we have that (3.12) does not exceed

$$c \sum_{\left(\frac{x-1}{D}\right)^{\varrho} < q \le \left(\frac{x-1}{D}\right)^{1-\vartheta_x}} f_2^2(q)\pi\left(\frac{x-1}{D}, q, l_q\right).$$

Using sieve estimates this is

$$c\frac{1}{\vartheta_x}\frac{x-1}{D\log\frac{x-1}{D}}\sum_{\left(\frac{x-1}{D}\right)^{\varrho}<q\le\left(\frac{x-1}{D}\right)^{1-\vartheta_x}}\frac{f_2^2(q)}{q},$$

which thanks to the choice (3.8) implies (3.12).

Using (3.11) and (3.12) we get that

$$\frac{1}{K_D(x)}\sum_{n\in K_D(x)}\left|e^{it\left(f_0(n)+f_1(n)+f_2(n)-\left(A\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)\right)-a(D)\right)}-\right.$$

$$\left.-e^{it(f_0(n)-(A(y_x)-a(D)))}\right|^2=o_x(1)$$

uniformly for all $|t|<T$, and $1\le D\le x^{1-\epsilon}$.

Using this and (3.5) we have proved that

$$\sup_{D\le x^{1-\epsilon}}\left|\frac{1}{\#K_D(x)}e^{-it\left(A\left(\left(\frac{x-1}{D}\right)^{\varrho}\right)-a(D)\right)}\sum_{n\in K_D(x)}e^{itf(n)}-\varphi_D(t)\right|\to 0\quad(x\to\infty)$$

uniformly as $|t|\le T$, T is an arbitrary constant.

4. Proof of Theorem 1

4.1. Concluding the sufficiency part of Theorem 1 from Theorem 2

Consider first the following sequence of distribution functions:

$$F_{k,x}(z)=\nu_x\left(n\in U_k(x)\;:\;f(n+1)-A(x)\le z\right),$$

where

$$A(x)=\sum_{\substack{p\le x\\|f(p)|\le 1}}\frac{f(p)}{p}.$$

With the notations of Lemma 2, an application of Lemma 1 shows that the characteristic function of F_x is

$$\frac{1}{B_k(x)} e^{-itA(x)} \sum_{\pi_{k-1}\in S_x} \sum_{P(\pi_{k-1})<p\le \frac{x}{\pi_{k-1}}} e^{itf(\pi_{k-1}p+1)}.$$

Since $A(x)$ is convergent, using Theorem 2 and Lemma 1, we can express it in the following form

$$\frac{1}{B_k(x)} \sum_{\pi_{k-1}\in S_x} \pi\left(\frac{x}{\pi_{k-1}}\right)\left(\varphi_{\pi_{k-1}}(t)+o_x(1)\right)-$$

$$-\frac{1}{B_k(x)} \sum_{\pi_{k-1}\in S_x} \pi\left(P(\pi_{k-1})\right)\left(\varphi_{\pi_{k-1}}(t)+o_x(1)\right), \tag{4.1}$$

for all $|t|<T$. We use the estimation $\pi\left(P(\pi_{k-1})\right)\le x^{1/A_x}$ in the second term, and we get that this is

$$\mathcal{O}(x^{2/A_x}).$$

Using the identity

$$\varphi(t)=\varphi_D(t)\,K_D(t),$$

where

$$K_D(t)=\prod_{\substack{p|D\\ |f(p)|>1}}(1+h(p))\prod_{\substack{p|D\\ |f(p)|\le 1}}(1+h(p))\,e^{-it\frac{f(p)}{p}}$$

we get that the main term in (4.1) is

$$\Sigma_1=\varphi(t)\frac{1}{B_k(x)}\sum_{\pi_{k-1}\in S_x}\pi\left(\frac{x}{\pi_{k-1}}\right)+ \tag{4.2}$$

$$+\frac{1}{B_k(x)}\sum_{\pi_{k-1}\in S_x}\pi\left(\frac{x}{\pi_{k-1}}\right)\varphi_{\pi_{k-1}}(t)\left(1-K_{\pi_{k-1}}(t)\right)+$$

$$+o(1)\frac{1}{B_k(x)}\sum_{\pi_{k-1}\in S_x}\pi\left(\frac{x}{\pi_{k-1}}\right).$$

Since

$$\frac{1}{B_k(x)}\sum_{\pi_{k-1}\in S_x}\pi\left(\frac{x}{\pi_{k-1}}\right)\to 1 \quad (x\to\infty), \tag{4.3}$$

we get

$$\Sigma_1 = \varphi(t)(1 + o(1)).$$

To see this last identity we calculate the second term in (4.2). If $p|\pi_{k-1}$, then $p > M_x$, so for big enough x

$$\log K_{\pi_{k-1}}(t) = \sum_{\substack{p|\pi_{k-1} \\ |f(p)|>1}} \log(1 + h(p)) + \sum_{\substack{p|\pi_{k-1} \\ |f(p)|\le 1}} \log(1 + h(p)) - \frac{itf(p)}{p} =$$

$$= \sum_{\substack{p|\pi_{k-1} \\ |f(p)|>1}} \frac{e^{itf(p)} - 1}{p} + \mathcal{O}\left(\sum_{\substack{p|\pi_{k-1} \\ |f(p)|>1}} \frac{1}{p^2}\right) +$$

$$+ \sum_{\substack{p|\pi_{k-1} \\ |f(p)|\le 1}} \frac{e^{itf(p)} - 1 - itf(p)}{p} + \mathcal{O}\left(\sum_{\substack{p|\pi_{k-1} \\ |f(p)|\le 1}} \frac{1}{p^2}\right) =$$

$$= \mathcal{O}\left(\sum_{\substack{p|\pi_{k-1} \\ |f(p)|>1}} \frac{1}{p}\right) + \mathcal{O}\left(\sum_{\substack{p|\pi_{k-1} \\ |f(p)|>1}} \frac{1}{p^2}\right) +$$

$$+\mathcal{O}(1)|t|^2\left(\sum_{\substack{p|\pi_{k-1} \\ |f(p)|\le 1}} \frac{f^2(p)}{p}\right) + \mathcal{O}\left(\sum_{\substack{p|\pi_{k-1} \\ |f(p)|\le 1}} \frac{1}{p^2}\right),$$

so

$$K_{\pi_{k-1}}(t) - 1 = o_x(1)$$

for all $|t| < T$, which together with the convergence of $A(x)$ implies our assertion.

4.2. Proof of the necessity part of Theorem 1

In the proof we shall use some ideas of Hildebrand [5] combining these with ours.

Corollary 2. *Let f be an additive function. Assume that $f(n) = c\log n + g(n)$ and the series*

$$\sum_{|g(p)|>1} \frac{1}{p-1}, \qquad \sum_{|g(p)|\le 1} \frac{g^2(p)}{p-1}$$

converge. Define

$$U(x) = c\log x + \sum_{\substack{p\le x\\ |g(p)|\le 1}} \frac{g(p)}{p-1}.$$

Then $\nu_x(n\in\mathcal{P}_k,\ n\le x:\ f(n+1) - U(x)\le z)$ *converge weakly to a limit distribution as* $x\to\infty$. *The characteristic function of the limit law is*

$$\chi(t) := \frac{1}{1+itc}\prod_{|g(p)|>1}\left(1-\frac{1}{p-1}+\sum_{m\ge 1}\frac{e^{itg(p^m)}}{p^m}\right)\times$$

$$\times\prod_{|g(p)|\le 1}\left(1-\frac{1}{p-1}+\sum_{m\ge 1}\frac{e^{itg(p^m)}}{p^m}\right)e^{-it\frac{g(p)}{p}},$$

and the limit distribution is continuous if and only if

$$\sum_{f(p)\ne 0}\frac{1}{p}$$

diverges.

Proof. Let

$$A(x) = \sum_{\substack{p\le x\\ |g(p)|\le 1}} \frac{g(p)}{p-1}.$$

Then this last lemma shows, that the distributions

$$\nu_x(n\in\mathcal{P}_k,\ n\le x:\ g(n+1) - A(x)\le z)$$

possess a limit law with characteristic function $\xi(t)$, say. Let

$$\varphi_x(t) = \frac{1}{\Pi_k(x)}\sum_{\pi_k\le x} e^{itg(\pi_k+1)}.$$

We have

$$\varphi_x(t)\,e^{-itA(x)} \to \xi(t)\quad (x\to\infty).$$

Consider next the following sum

$$\frac{1}{\Pi_k(x)}\sum_{\pi_k\le x} e^{itf(\pi_k+1)} = \frac{1}{\Pi_k(x)}\sum_{\pi_k\le x}(\pi_k+1)^{itc}\,e^{itg(\pi_k+1)}. \tag{4.4}$$

Remember that k may depend on x. Let us introduce the following notation

$$\Psi_{k,x}(y,t) := \sum_{\pi_{k_x} \le y} e^{itg(\pi_k+1)}.$$

Applying the Abel summation formula for the right hand side for (4.4) we get

$$\varphi_x(t)\, x^{itc} - itc \frac{1}{\Pi_k(x)} \int_1^x y^{itc-1} \Psi_{k,x}(y,t)\, dy. \tag{4.5}$$

Since for $x^\lambda < y < x$ using (2.1), we have

$$\lim_{x\to\infty} \frac{1}{\Pi_{k_x}(y)} \Psi_{k,x}(y,t) e^{-itA(y)} = \xi(t),$$

thus for this y

$$\Psi_{k,x}(y,t) = \xi(t)\, e^{itA(x)} + \xi(t)\, e^{itA(y)} \left\{1 - e^{it(A(x)-A(y))}\right\} + o(1).$$

Since

$$A(x) - A(y) = o(1)$$

if $x^\lambda < y < x$, we have

$$\frac{1}{\Pi_{k_x}(y)} \Psi_{k,x}(y,t) = \xi(t)\, e^{itA(x)} + o(1). \tag{4.6}$$

Using this we have that the integral in (4.5) equals

$$\int_1^{x^\lambda} y^{itc-1} \Psi_{k,x}(y,t)\, dy \, + \, \xi(t)\, e^{itA(x)} \int_{x^\lambda}^x y^{itc-1} \Pi_k(y)\; dy \, + \, o(x). \tag{4.7}$$

The first term is

$$\mathcal{O}(x^\lambda) = \Pi_k(x) o(1).$$

Using the estimation (2.1) we get that the second term in (4.7) is

$$\xi(t)\, e^{itA(x)} \int_{x^\lambda}^x y^{itc-1} \left(\frac{y}{\log y} \frac{\log\log^{k-1} y}{(k-1)!} (1 + o(1)) \right) dy,$$

which is

$$\frac{1}{(itc+1)}\xi\left(t\right)e^{itA(x)}\Pi_k(x)+\Pi_k(x)o(1).$$

Here we used, that k depends only upon the upper bound of the integration, and that

$$\left(\frac{x^{itc+1}}{\log x}\frac{\log\log^{k-1}x}{(k-1)!}\right)'=$$

$$=(itc+1)\frac{x^{itc}}{\log x}\frac{\log\log^{k-1}x}{(k-1)!}+\mathcal{O}\left(\frac{\log\log^{k-2}x}{\log^2 x}\frac{1}{(k-2)!}\right).$$

We had shown, that (4.5) equals

$$\xi\left(t\right)\frac{e^{itA(x)}x^{itc}}{1+itc}\;+\;o\left(1\right).$$

We get

$$\frac{1}{\Pi_k\left(x\right)}\sum_{\pi_k\le x}e^{it(f(\pi_k+1))}=\frac{x^{ict}e^{itA(x)}}{1+ict}\xi\left(t\right)+o\left(1\right),$$

and so

$$\frac{1}{\Pi_k\left(x\right)}\sum_{\pi_k\le x}e^{it(f(\pi_k+1)-U(x))}=\frac{\xi\left(t\right)}{1+ict}+o\left(1\right),$$

and our lemma immediately follows.

Lemma 6 (Sieve estimate). *Let q, q' be prime numbers, and $x>x_0$. We have*

$$A=\#\{n\le x: qn+1, q'n+1\in U_k(x)\}\ll$$

(4.8)
$$\ll\frac{x}{\log^2 x}\left(\frac{(\log\log x)^{k-1}}{(k-1)!}\right)^2\Psi\left(|q-q'|\right),$$

where

$$\Psi\left(n\right)=\prod_{\substack{p|n\\p>2}}\frac{p-1}{p-2}.$$

Proof. Using Lemma 1 we need only to count the elements of the set

$$A=\#\{n\le x:\; qn+1=\pi_{k-1}p,\; q'n+1=\pi'_{k-1}p' \text{ and } \pi_{k-1}\le x^\beta,\; \pi'_{k-1}\le x^\beta\}.$$

Recall that $P(qn+1) = p$, $P(q'n+1) = p'$ and $p(\pi_{k-1}) \geq M_x$, $p(\pi'_{k-1}) \geq M_x$. We have

$$qn + 1 = \pi_{k-1} p,$$
$$q'n + 1 = \pi'_{k-1} p',$$

so by the Chinese remainder theorem $n \equiv l_{\pi_{k-1}\pi'_{k-1}} (\bmod\ \pi_{k-1}\pi'_{k-1})$ with a unique $l_{\pi_{k-1}\pi'_{k-1}}$. So $n = t\pi_{k-1}\pi'_{k-1} + l_{\pi_{k-1}\pi'_{k-1}}$, where $t \leq \frac{x}{\pi_{k-1}\pi'_{k-1}}$. For p, p' we have

$$p = qt\pi'_{k-1} + \frac{ql+1}{\pi_{k-1}}, \qquad t \leq \frac{x}{\pi_{k-1}\pi'_{k-1}}.$$
$$p' = q't\pi_{k-1} + \frac{q'l+1}{\pi'_{k-1}},$$

Using sieve estimates we have, that the number of such p, p' is not more than

$$A \ll$$

$$\sum_{\substack{\pi_{k-1} \leq x^\beta \\ \pi'_{k-1} \leq x^\beta}} \Psi\left(q\pi'_{k-1} q'\pi_{k-1} \left| q\pi'_{k-1} \frac{q'l+1}{\pi'_{k-1}} - q'\pi_{k-1} \frac{ql+1}{\pi_{k-1}} \right| \right) \frac{x}{\log^2 x} \frac{1}{\pi_{k-1}\pi'_{k-1}} \ll$$

$$\ll \Psi\left(|q - q'|\right) \frac{x}{\log^2 x} \sum_{\substack{\pi_{k-1} \leq x^\beta \\ \pi'_{k-1} \leq x^\beta}} \Psi\left(\pi'_{k-1}\pi_{k-1}\right) \frac{1}{\pi_{k-1}\pi'_{k-1}}. \tag{4.9}$$

Here we observed, that

$$q\pi'_{k-1} \frac{q'l+1}{\pi'_{k-1}} - q'\pi_{k-1} \frac{ql+1}{\pi_{k-1}} = q - q'.$$

The sum in the last expression of (4.9) is

$$\sum_{\pi_{k-1} \leq x^\beta} \prod_{p_i | \pi_{k-1}} \frac{p_i - 1}{(p_i - 2)p_i} \sum_{\pi'_{k-1} \leq x^\beta} \prod_{p'_i | \pi'_{k-1}} \frac{p'_i - 1}{(p'_i - 2)p'_i} \ll$$

$$\ll \sum_{\pi_{k-1} \leq x^\beta} \prod_{p_i | \pi_{k-1}} \frac{1}{(p_i - 2)} \sum_{\pi'_{k-1} \leq x^\beta} \prod_{p'_i | \pi'_{k-1}} \frac{1}{(p'_i - 2)} \ll$$

$$\ll \left(\frac{1}{(k-1)!} \left(\sum_{p \leq x^\beta} \frac{1}{p} + \sum_{p > M_x} \frac{1}{p^2} \right)^{k-1} \right)^2,$$

which implies our assertion.

The next theorem can be found in [5] for shifted primes.

Lemma 7. *With suitable constant* δ_1 *and* c_1*, and multiplicative* $g:\mathbb{N}\to\mathbb{C}$ *such that* $|g|=1$*, and*

$$\max_{1\le l\le c_1}\left|\frac{1}{x}\sum_{n\le x} g(n)^l\right|\le \delta_1$$

we have

$$\left|\frac{1}{B_k(x)}\sum_{n\in U_k(x)} g(n+1)\right|\le 1-\delta_1. \tag{4.10}$$

Proof. We use the ideas of Hildebrand without any important changes. Therefore we shall give an outline of the proof, only.

It is enough to prove that if the conditions hold, then

$$1-\left|\frac{1}{B_k(x)}\sum_{\pi_k\in U_k(x)} g(\pi_k+1)\right|\gg 1.$$

Some computation shows that

$$1-\left|\frac{1}{B_k(x)}\sum_{\pi_k\in U_k(x)} g(\pi_k+1)\right|\ge\frac{1}{2B_k(x)}\sum_{\pi_k\in U_k(x)}|1-wg(\pi_k+1)|^2 \tag{4.11}$$

with an appropriate complex w, with absolute value 1. Setting

$$R(Q)=\frac{1}{B_k(x)}\sum_{\pi_k\in U_k(x)}|1-wg(\pi_k+1)|^2\sum_{\substack{q\mid \pi_k+1\\ Q<q<2Q}}1,$$

we get with $0<\eta<1/4$ and $x\ge 2^{1/\eta}$ after some computation that the right hand side of (4.11) is at least

$$\frac{\eta^2}{2\log 2}\log x\min_{x^\eta\le Q\le x^{2\eta}}R(Q),$$

so it is enough to prove that $R(Q) \gg \frac{1}{\log x}$ uniformly with $x > x_0$, $x^\eta \le Q \le$ $\le x^{2\eta}$. Let $\delta < 1/4$. There is an ω complex number with absolute value 1 such that

$$\sum_{\substack{Q<q<2Q \\ |g(q)-\omega|\le\delta}} 1 \ge \frac{\delta}{10}\frac{Q}{\log Q}. \tag{4.12}$$

Set

$$S = \sum_{\pi_k\in U_k(x)} {\sum_{\substack{q|\pi_k+1 \\ Q<q<2Q}}}' 1.$$

The $'$ in the inner sum means the restriction to prime numbers q for which $|g(q)-\omega| \le \delta$ with this appropriate ω. We get a lower bound for this sum applying Lemma 4 and Corollary 1.

$$S = {\sum_{Q<q<2Q}}' B_k(x,q,-1) \ge \frac{B_k(x)}{4Q} {\sum_{Q<q<2Q}}' 1 + \mathcal{O}\left(\frac{B_k(x)}{\log^A x}\right).$$

If $\pi_k + 1 = nq$ and $q^2 \nmid \pi_k + 1$, then after some computation we have

$$|1 - wg(\pi_k+1)| \ge |g(n) - \overline{w\omega}| - |g(q) - \omega|.$$

Thus $|g(q)-\omega| \le \delta$ implies $|g(n) - \overline{w\omega}| \le 2\delta$ or $|1 - wg(\pi_k+1)| > \delta$.

Let

$$S_1 = \sum_{\pi_k\in U_k(x)} {\sum_{Q<q<2Q}}' {\sum_{\substack{n \\ qn=\pi_k+1}}}'' 1,$$

where $''$ in the inner sum means the restriction to integers for which $|g(n) -$ $-\overline{w\omega}| \le 2\delta$. Let

$$S_2 = {\sum_{\pi_k\in U_k(x)}}^* {\sum_{\substack{q|\pi_k+1 \\ Q<q<2Q}}}' 1,$$

where $*$ in the outer sum means the restriction to the prime numbers for which $|1 - wg(\pi_k+1)| > \delta$. Let

$$S_3 = \sum_{\pi_k\in U_k(x)} {\sum_{\substack{q^2|\pi_k+1 \\ Q<q<2Q}}}' 1.$$

We have $S \leq S_1 + S_2 + S_3$. It is easy to see that $S_3 \leq x^{1-\eta}$. We use $R(Q)$ to estimate S_2.

$$S_2 \leq \sum_{\pi_k \in U_k(x)} \frac{|1 - wg(\pi_k + 1)|^2}{\delta^2} \sum_{\substack{q \mid \pi_k+1 \\ Q<q\leq 2Q}} 1 = \frac{B_k(x)}{\delta^2} R(Q).$$

Putting it all together it is enough to prove that $S_1 \leq \frac{B_k(x)}{8Q} \sum'_{Q<q<2Q} 1$.

We have

$$S_1 \leq \sum''_{n \leq \frac{x+1}{Q}} \sum'_{\substack{Q<q<2Q \\ qn-1=\pi_k}} 1.$$

Applying the Cauchy-Schwarz inequality we get

$$(S_1)^2 \leq \underbrace{\left(\sum''_{n \leq \frac{x+1}{Q}} 1\right)}_{S_{11}} \underbrace{\left\{\sum_{n \leq \frac{x+1}{Q}} \left(\sum'_{\substack{Q<q<2Q \\ qn-1=\pi_k}} 1\right)^2\right\}}_{S_{12}}.$$

We get

$$S_{12} = \sum''_{Q<q,q'<2Q} \sum'_{\substack{n \leq \frac{x+1}{Q} \\ qn-1=\pi_k \\ q'n-1=\pi'_k}} 1.$$

If $q = q'$ then this sum is $\mathcal{O}(x)$. For the other case we can use Lemma 6, and we get

$$S_{12} \ll \frac{x}{Q \log^2 x} \left(\frac{(\log\log x)^{k-1}}{(k-1)!}\right)^2 \sum'_{Q<q<2Q} \sum'_{Q<q'<q} \Psi(q - q') + x.$$

After some computation we get that the inner sum does not exceed

$$c \left(\frac{\delta}{10}\right)^{-1/2} \sum'_{Q<q<2Q} 1$$

(see Hildebrand).

We have

$$S_{12} \ll \delta^{-1/2} \frac{x}{Q \log^2 x} \left(\frac{(\log\log x)^{k-1}}{(k-1)!}\right)^2 \left(\sum'_{Q<q<2Q} 1\right)^2 + x.$$

It is not so hard to prove that if the conditions of the lemma hold, then $S_{11} \le \frac{x}{Q}\delta \log \frac{1}{\delta}$.

Putting it together we get

$$S_1 \ll \left(\delta^{1/5}\frac{\Pi_k(x)}{Q} + \frac{x}{\delta}\frac{\log Q}{Q^{3/2}}\right) \sideset{}{'}\sum_{Q<q<2Q} 1.$$

Choosing small δ, we finished the outline of the proof.

The following two lemmas can be found in [5], thus we give only remarks accordingly to our case.

Lemma 8. *Assume that*

$$\min_{\substack{1\le l\le c_1\\ \tau\le c_2}} \operatorname{Re} \sum_{p\le x} \frac{1-g(p)^l p^{-i\tau}}{p} > c_2$$

with a suitable $c_2 > 0$. *Then (4.10) holds.*

Remark. One can prove that if this condition holds, then the condition of Lemma 7 holds, too, with a suitable large c_2.

Lemma 9. *There are constants* δ_3, c_3 *such that for fixed* $x \ge 2$, $h > 0$ *and*

$$\max_{a\in\mathbb{R}}\{n \in U_k(x) \,:\, f(n+1) \in [a, a+h]\} \ge (1-\delta_3)\, B_k(x)$$

we have

$$\min_{|\lambda|\le c_3 h^2} \sum_{p\le x} \frac{\left(\min\left(h, f(p) - \lambda \log p\right)\right)^2}{p} \le c_3 h^2.$$

Proof. We must show that the condition of the previous lemma is violated if $h = 1$. We have

$$\left|\frac{1}{B_k(x)} \sum_{\pi_k\in U_k(x)} e^{itf(\pi_k+1)}\right| = \left|\frac{1}{B_k(x)} \sum_{\pi_k\in U_k(x)} e^{it(f(\pi_k+1)-a)}\right| \ge 1 - |t| - 2\delta_3.$$

Set $\delta_3 \le \frac{\delta_1}{4}$ and $|t| \le \frac{\delta_1}{2}$. The result of Lemma 7 is violated. We have from Lemma 8 that for $k = k(t) \le c_1$ and $\tau = \tau(t)$, $|\tau| \le c_2$

$$\sum_{p\le x} \frac{\operatorname{Re}\left(1 - e^{itf(p^k)} p^{-i\tau}\right)}{p} \le c_2.$$

The following lemma can be found in [1] Lemma 1.9.

Lemma 10. *Let F_x be a sequence of distribution functions, and let α, β be real functions. If $F_x(z+\alpha(x))$ and $F_x(z+\beta(x))$ converge weakly to a limit distribution with $x \to \infty$, then $\lim_{x\to\infty}(\alpha(x)-\beta(x))$ exists, and is finite.*

Proof of necessity part of Theorem 1. Assume f has a limit distribution. Then

$$\#\{n \leq x,\, n \in \mathcal{P}_k :\, f(n+1) \in [-z, z]\} \geq (1-\delta_3)\,\Pi_k(x)$$

holds for some z real number. We get using Lemma 9 with $h = 2z$

$$\sum_{p \leq x} \frac{(\min(h, f(p) - \lambda_x \log p))^2}{p} \leq c_3 h^2$$

with $|\lambda_x| \leq c_3 h^2$. It follows that

$$\sum_{p} \frac{(\min(h, f(p) - \lambda \log p))^2}{p}$$

converge for some λ. We get immediately that for $g(n) = f(n) - \lambda \log n$ the series

$$\sum_{|g(p)|>1} \frac{1}{p-1}, \qquad \sum_{|g(p)|\leq 1} \frac{g(p)^2}{p-1}$$

converge. From Corollary 2 and Lemma 10 we get that

$$\lambda \log x + \sum_{\substack{p \leq x \\ |g(p)| \leq 1}} \frac{g(p)}{p-1}$$

converge. This implies $\lambda = 0$ and the convergence of

$$\sum_{|f(p)|\leq 1} \frac{f(p)}{p-1},$$

and the proof of Theorem 1 is completed.

Acknowledgement. The author is grateful to Professor Kátai, who read carefully this paper and constructively commented it.

References

[1] **Elliot P.D.T.A.,** *Probabilistic number theory I.,* Springer Verlag, 1979.

[2] **Elliot P.D.T.A.,** *Probabilistic number theory II.,* Springer Verlag, 1980.

[3] **Erdős P. and Wintner A.,** Additive arithmetical functions and statistical independence, *Amer. J. Math.,* **61** (1939), 713-721.

[4] **Halberstam H. and Richert H.-E.,** *Sieve methods,* Acad. Press, London, 1974.

[5] **Hildebrand A.,** Additive and multiplicative functions on shifted primes, *Proc. London Math. Soc.,* **53** (1989), 209-232.

[6] **Kátai I.,** On the distribution of arithmetical functions on the set of primes plus one, *Compositio Math.,* **19** (1968), 278-289.

[7] **Kubilius J.,** *Probabilistic methods in the theory of numbers,* American Mathematical Society, Povidence, Rhode Island, 1964.

[8] **Prachar K.,** *Primzahlverteilung,* Springer Verlag, 1957.

[9] **Wolke D. and Zhan T.,** On the distribution of integers with a fixed number of prime factors, *Math. Z.,* **213** (1993), 133-147.

(Received February 19, 2007)

L. Germán
Department of Computer Algebra
Eötvös Loránd University
Pázmány Péter sétány 1/C
H-1117 Budapest, Hungary
german@compalg.inf.elte.hu