Proof. We shall prove somewhat more. Let $L = \log k - \gamma$.

By Theorem 7 to make $\gamma(r, k) = 0$ we must have

$$\psi(z) = -\log k, \quad z = r/k$$

or, by (17),

$$z\psi(1+z) = 1 - z\log k. \tag{26}$$

But by (19)

$$z\psi(1+z) = -\gamma z + \sum_{n=2}^{\infty} \zeta(n)(-z)^n.$$

Therefore (26) becomes

$$z = L^{-1}\{1 - \zeta(2)z^2 + \zeta(3)z^3 - \ldots\}. \tag{27}$$

Taking only the first term the theorem follows. Solving (27) by iteration we find

$$z = L^{-1} - \zeta(2)L^{-3} + \zeta(3)L^{-4} + [2\{\zeta(2)\}^2 - \zeta(4)]L^{-5} - \\ - [5\zeta(2)\zeta(3) + \zeta(5)]L^{-6} - \ldots \tag{28}$$

As an illustration we take $k = 100$. The terms of (28) become

$$.2483 - .0252 + .0046 + .0041 - .0025 = .2292.$$

By actual computation we find

$$\gamma(22, 100) = .00204268,$$
$$\gamma(23, 100) = -.00056747.$$

References

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Nat. Bur. Standards Appl. Math. Ser. 55 (1964), pp. 267–270.

[2] H. T. Davis, *Tables of the Higher Mathematical Functions*, vol. 1, Bloomington 1933, pp. 291–333.

[3] G. L. Dirichlet, *Recherches sur diverses applications de l'analyse infinitésimale à la théorie des nombres*, Journ. für Math. 21 (1840), pp. 134–155, *Werke*, vol. 1, p. 142.

[4] C. F. Gauss, *Disquisitiones generales circa seriem infinitam etc.*, Comment. Gotting. 2 (1813), pp. 1–46; *Werke*, vol. 3, 1876, pp. 122–162.

[5] O. Hölder, *Zur Theorie der Kreistheilungsgleichung*, Prace Mat. Fiz. 43 (1936), pp. 13–23.

[6] J. L. W. V. Jensen, *An elementary exposition of the theory of the Gamma function*, Ann. of Math., S.2, 17 (1915), pp. 124–166.

[7] V. A. Lebesgue, *Suite du Mémoire sur les applications de symbole* $(a|b)$, Journ. de Math., S.1, 15 (1850), pp. 215–237.

[8] I. Niven, *Irrational Numbers*, Carus Math. Monograph No. 11, 1956, pp. 36–41.

Received on 13. 6. 1973 (413)

ACTA ARITHMETICA
XXVII (1975)

On the distribution of additive arithmetic functions

by

GÁBOR HALÁSZ (Budapest)

Dedicated to the memory of Yu. V. Linnik

Let $g(n)$ be a real valued additive arithmetic function (i.e. $g(mn) = g(m) + g(n)$ if $(m, n) = 1$). The distribution of values of such functions has been extensively investigated. As a new direction, Erdös, Ruzsa and Sárközi [1] proposed to estimate

$$\max_{-\infty<a<\infty} N(a, x) \stackrel{\text{def}}{=} \max_{-\infty<a<\infty} \sum_{\substack{n\leqslant x\\ g(n)=a}} 1 \tag{1}$$

for general additive functions. They found bounds cx in various cases, often giving the best possible value of c. If, however, $g(n) = \omega(n)$, the number of prime divisors of n, then this quantity is about const $\frac{x}{\sqrt{\log\log x}}$ and they conjectured (oral communication) that this order of magnitude cannot be exceeded in any case, provided that $g(p) \neq 0$ for each prime p. The aim of this paper is to prove this conjecture in the following more precise form.

THEOREM. *Let $g(n)$ be an arbitrary real valued additive function and put*

$$E(x) = \sum_{\substack{p\leqslant x\\ g(p)\neq 0}} \frac{1}{p}.$$

Then there is a universal constant c_1 such that

$$N(a, x) = \sum_{\substack{n\leqslant x\\ g(n)=a}} 1 \leqslant c_1 \frac{x}{\sqrt{E(x)}}.$$

The result is sharp even in this more general form: The bound is attained if $g(p) = 0$ or 1 and $\sum_{\substack{p\leqslant x\\ g(p)=1}} 1/p = E(x)$ as is seen from [2] and [3] where much more detailed information is given in this special case. (For references to earlier works on the "local distribution" of additive functions see [2].) It can be shown that at least in a certain weaker sense this example is extremal even as far as the best value of c_1 is concerned. To this and other generalizations we intend to return elsewhere.

$c_1, c_2, \ldots$ will denote positive universal constants and the same is true of constants involved in the symbol $O(\ )$.

Proof. We think x large but fixed throughout. In order to simplify some detail, we begin by observing that $g(n)$ can be assumed to have integer values only.

To see this, let r be the maximal number of linearly independent values $g(n)$ $(n \leqslant x)$ over the rationals. We can then find r real numbers ϑ_i such that for each $n \leqslant x$ there is a unique representation

$$g(n) = \sum_{i=1}^{r} k_i(n)\vartheta_i \qquad (n \leqslant x)$$

with integers $k_i(n)$. We can further find integers b_i $(i = 1, \ldots, r)$ such that with

$$g_0(n) \overset{\text{def}}{=} \sum_{i=1}^{r} k_i(n) b_i \qquad (n \leqslant x)$$

$g_0(p) \neq 0$, whenever $g(p) \neq 0$ for $p \leqslant x$. With the definition

$$g_0(n) \overset{\text{def}}{=} \sum_{\substack{p^k \| n \\ p^k \leqslant x}} g_0(p^k)$$

we arrive at an integral valued additive function with the same $E(x)$ as for $g(n)$. For $n \leqslant x$ this definition is in accordance with the earlier one and this means that whenever $g(n)$ takes a value $a = \sum_{i=1}^{r} a_i\vartheta_i$ for $n \leqslant x$, $g_0(n)$ takes $\sum_{i=1}^{r} a_i b_i$ so that (1) is majorized by the corresponding quantity for $g_0(n)$ and our statement follows.

Preserving the notation $g(n)$, we consider, following Selberg [5] as we did in [2] and [3],

$$M(u) = M(u, x) = \sum_{n \leqslant x} e^{iug(n)},$$

giving for integral a

$$N(a, x) = \frac{1}{2\pi}\int_0^{2\pi} M(u)e^{-iau}\,du \leqslant \frac{1}{2\pi}\int_0^{2\pi} |M(u)|\,du.$$

Now, $f(n) = e^{iug(n)}$ is, for each u, a multiplicative function (i.e. $f(mn) = f(m)f(n)$ for $(m, n) = 1$) and we are led to investigate the mean values of such functions. In [4] we developed an analytic method that gives the desired bound in terms of the simple quantity

$$m(u, T) = m(u, T, x) = \min_{|t| \leqslant T} \sum_{p \leqslant x} \frac{1 - \operatorname{Re} f(p)p^{-it}}{p}$$

$$= \min_{|t| \leqslant T} \sum_{p \leqslant x} \frac{1 - \cos\big(ug(p) - t\log p\big)}{p}$$

as

$$|M(u)| \leqslant c_2 x \exp\{-\tfrac{1}{16} m(u, T)\},$$

e.g. with $T = \log x$. For the proof see Lemma later; now we use it in the integral to be estimated.

For this purpose we first try to bound the size of the set

$$S = S(m, T) = \{u\colon\ m(u, T) \leqslant m\}.$$

In other words, $u \in S(m, T)$ means that there is a $|t| \leqslant T$ depending on u with

$$\sum_{p \leqslant x} \frac{1 - \cos\big(ug(p) - t\log p\big)}{p} \leqslant m.$$

Taking into account that

$$1 - \cos a = 2\sin^2(a/2) \quad \text{and} \quad \sin^2\Big(\sum_{i=1}^{k} a_i\Big) \leqslant \Big(\sum_{i=1}^{k} |\sin a_i|\Big)^2 \leqslant k\sum_{i=1}^{k} \sin^2 a_i,$$

we see that $u \in S(k^2 m, kT)$ if $u = \sum_{i=1}^{k} u_i$ with $u_i \in S(m, T)$, the t corresponding to u can be chosen as $\sum_{i=1}^{k} t_i$ with the t_i corresponding to u_i. We shall prove for the sake of completeness that every real number u can be represented in this form provided that

$$k|S(m, T)| > 2\pi. \tag{2}$$

($|\ldots|$ stands for the measure inside $(0, 2\pi)$.) For such a k we have thus proved that for every real u there exists a $|t| \leqslant kT$ with

$$\sum_{p \leqslant x} \frac{1 - \cos\big(ug(p) - t\log p\big)}{p} \leqslant k^2 m. \tag{3}$$

From among the pairs (u, t) $(|t| \leqslant kT)$ satisfying this inequality let us choose one with maximal $|t|$. We denote it by (u_0, t_0). First we are going to prove $|t_0|$ "large".

Considering our quantity with $t = 0$,

$$\frac{1}{2\pi}\int_0^{2\pi}\sum_{p\leqslant x}\frac{1-\cos(ug(p))}{p}\,du = \sum_{\substack{p\leqslant x\\ g(p)\neq 0}}\frac{1}{p} = E(x).$$

Hence there must be a u' with

$$\sum_{p\leqslant x}\frac{1-\cos(u'g(p))}{p}\geqslant E(x).$$

Now, by the trigonometric inequality we have already used,

$$E(x)\leqslant\sum_{p\leqslant x}\frac{1-\cos(u'g(p))}{p}$$

$$\leqslant 2\sum_{p\leqslant x}\frac{1-\cos(u'g(p)-t\log p)}{p}+2\sum_{p\leqslant x}\frac{1-\cos(t\log p)}{p}.$$

Here

(4) $$\sum_{p\leqslant x}\frac{1-\cos(t\log p)}{p}\leqslant c_3\log(2+|t|\log x),$$

as we shall show later and for $|t|\leqslant\dfrac{e^{c_4E(x)}-2}{\log x}$ $(c_4 = 1/4c_3)$

$$\sum_{p\leqslant x}\frac{1-\cos(u'g(p)-t\log p)}{p}\geqslant\tfrac{1}{4}E(x).$$

This means that (3) cannot be satisfied for such a t with $u = u'$ if we assume that

(5) $$k^2m<\tfrac{1}{4}E(x)$$

and for the value t' that we know exists satisfying (3) (with $u = u'$) necessarily

$$|t'|>\frac{e^{c_4E(x)}-2}{\log x}\geqslant\frac{e^{c_5E(x)}}{\log x}$$

(since we can assume $E(x)\geqslant c_6$, otherwise our theorem is trivial) and t_0 being maximal,

$$|t_0|>\frac{e^{c_5E(x)}}{\log x}.$$

Now we manipulate with $(2u_0, 2t_0)$. Since (3) holds for the pair (u_0, t_0), by the much used inequality we have

$$\sum_{p\leqslant x}\frac{1-\cos(2u_0g(p)-2t_0\log p)}{p}\leqslant 4k^2m.$$

On the other hand, we know that to $2u_0$ there exists a $t = t_1$ $(|t_1|\leqslant kT)$ with

$$\sum_{p\leqslant x}\frac{1-\cos(2u_0g(p)-t_1\log p)}{p}\leqslant k^2m$$

and from the last two lines, again by the trigonometric inequality,

$$\sum_{p\leqslant x}\frac{1-\cos(t_2\log p)}{p}\leqslant 10k^2m$$

with $t_2 = 2t_0-t_1$. Here $|t_1|\leqslant|t_0|$, using the maximality of t_0, so that $|t_2|\geqslant|t_0|>e^{c_5E(x)}/\log x$. We shall show that an earlier inequality is sharp:

(6) $$\sum_{p\leqslant x}\frac{1-\cos(t_2\log p)}{p}\geqslant c_6\log(|t_2|\log x)\geqslant c_6c_5E(x),$$

provided that e.g. $|t_2|\leqslant T^2$ which is fulfilled if

(7) $$3kT\leqslant T^2.$$

Thus we get

$$c_6c_5E(x)\leqslant 10k^2m.$$

If the condition (5) fails, this holds automatically. Choose k in compliance with (2) as $k = [2\pi/|S|]+1$ implying

$$c_6c_5E(x)\leqslant 10\left(\frac{4\pi}{|S|}\right)^2m,\qquad |S|\leqslant c_7\sqrt{\frac{m}{E(x)}}.$$

It remains to dispose of condition (7) that with our definition of k takes the form

$$\frac{c_8}{|S|}\leqslant T,\qquad |S|\geqslant\frac{c_9}{T} = \frac{c_9}{\log x}.$$

However, our inequality is trivial otherwise at least for $m\geqslant 1$, $E(x)$ being $\ll\log\log x$. We have thus proved

$$|S(m, T)|\leqslant c_{10}\sqrt{\frac{m}{E(x)}}\quad\text{for}\quad m\geqslant 1.$$

Dividing the range of integration in

$$\frac{1}{2\pi}\int_0^{2\pi}|M(u)|\,du\leqslant\frac{c_2}{2\pi}x\int_0^{2\pi}\exp\{-\tfrac{1}{16}m(u, T)\}\,du$$

by the sets $\{u\colon 0\leqslant u\leqslant 2\pi,\, 2^{l-1}-1< m(u,T)\leqslant 2^l-1\}$ $(l=1,2,\ldots)$ and observing that this set is contained in $S(2^l-1,T)$, we get for the right hand side the upper bound

$$c_{11}x\sum_{l=1}^{\infty}\exp\{-\tfrac{1}{16}(2^{l-1}-1)\}\cdot c_{10}\sqrt{\frac{2^l-1}{E(x)}}=c_1\frac{x}{\sqrt{E(x)}},$$

what we had to prove.

As to the measure-theoretic result used, we prove it by induction in the following form: Let S be a closed set on the real line, periodic with period 2π, symmetric with respect to the origin and S_k be the set of points representable as $u=\sum_{i=1}^{k}u_i$ with $u_i\in S$. Then either S_k is the whole line or $|S_k|\geqslant k|S|$. (As before, $|\ldots|$ is the measure in an interval of length 2π; all these sets are easily seen to be closed, periodic and symmetric.)

Suppose that S_k is not the whole line and let $\{v_l\}$ be a sequence with $\lim v_l=v\in S_k$ but $v_l\notin S_k$. The latter implies that the two sets $v_l-S=v_l+S$ (meaning reflection and translation) and S_{k-1} are disjoint and letting $l\to\infty$, by a well-known continuity property of the Lebesgue-measure,

$$|(v+S)\cap S_{k-1}|=0.$$

On the other hand, $v\in S_k$ means $v=u'-u''$ $(u'\in S_{k-1},\, u''\in -S=S)$ and translation by u'' gives $|(u'+S)\cap(u''+S_{k-1})|=0$ and so

$$\begin{aligned}|(u'+S)\cup(u''+S_{k-1})|&=|u'+S|+|u''+S_{k-1}|=|S|+|S_{k-1}|\\&\geqslant|S|+(k-1)|S|=k|S|,\end{aligned}$$

using the inductive hypothesis. But obviously both $u'+S$ and $u''+S_{k-1}$ are contained in S_k and the proof is completed.

Now we prove (4) and (6) introducing the Riemann zeta function

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_p\frac{1}{1-\frac{1}{p^s}}=e^{\sum_p\log\frac{1}{1-1/p^s}}=e^{\sum_p\sum_{k=1}^{\infty}\frac{1}{kp^{ks}}}$$

with $s=\sigma+it$, $\sigma=1+\frac{1}{\log x}$, although (4) is elementary. (For full detail concerning the simple elementary steps that follow see (6) and (7) in [2]; here some hints will suffice.)

We drop terms with $k\geqslant 2$, making an error $O(1)$ in the exponent. Dropping also terms with $p>x$, $k=1$, we make the same error since

$$\sum_{p>x}\frac{1}{p^\sigma}=O(1)\qquad\Big(\sigma=1+\frac{1}{\log x}\Big),$$

(dividing the range by x^{2^l} $(l=1,\ldots)$ and using $\sum_{y<p\leqslant y^2}1/p=O(1)$) and also when replacing σ by 1 in the remaining terms $p\leqslant x$, $k=1$, owing to

$$\sum_{p\leqslant x}\Big(\frac{1}{p}-\frac{1}{p^\sigma}\Big)=O(1)\qquad\Big(\sigma=1+\frac{1}{\log x}\Big),$$

(using $1-\frac{1}{p^{\sigma-1}}\leqslant(\sigma-1)\log p$ and $\sum_{p\leqslant x}\frac{\log p}{p}=O(\log x)$). Therefore we can express our quantity as

$$\sum_{p\leqslant x}\frac{1-\cos(t\log p)}{p}=\log\zeta(\sigma)-\log|\zeta(\sigma+it)|+O(1).$$

For $|t|\leqslant 1$, by the first order pole at $s=1$, this is

$$\log\frac{|\sigma-1+it|}{\sigma-1}+O(1)=\log(2+|t|\log x)+O(1)$$

and for $1\leqslant|t|\leqslant T^2=\log^2x$

$$\log\frac{1}{\sigma-1}+O(\log\log|t|)=\log\log x+o(\log\log x)$$

by the well-known estimation (see these e.g. in [6]) and our statements follow in any case.

Lemma. *Let $f(n)$ be a multiplicative function with $|f(n)|\leqslant 1$ and set*

$$m(f,T)=m(f,T,x)=\min_{|t|\leqslant T}\sum_{p\leqslant x}\frac{1-\mathrm{Re}f(p)p^{-it}}{p}.$$

Then

$$\Big|\sum_{n\leqslant x}f(n)\Big|\leqslant c_{11}x\exp\{-\tfrac{1}{16}m(f,T)\}$$

e.g. with $T=\log x$. ($\frac{1}{16}$ can be replaced by the sharp factor 1 but we do not need it here.)

Proof. This is the proof of Satz 1′ in [4] with straightforward modifications and we shall rely on our paper [4].

We first assume $f(n)$ to be completely multiplicative, i.e. $f(mn)=f(m)f(n)$ for each pair (m,n). Denoting by $\sum'$ summation over integers having all their prime divisors $\leqslant x$, let

$$\begin{aligned}F(s)&=\sum{}'\frac{f(n)}{n^s}=\prod_{p\leqslant x}\Big(1+\frac{f(p)}{p^s}+\frac{f(p^2)}{p^{2s}}+\ldots\Big)=\prod_{p\leqslant x}\frac{1}{1-f(p)/p^s}\\&=\exp\Big\{\sum_{p\leqslant x}\sum_{k=1}^{\infty}\frac{f(p^k)}{kp^{ks}}\Big\}.\end{aligned}$$

(This corresponds to the $F^*(s)$ of [4].) Performing the same changes as before in the case of the zeta function, we also get

$$(8)\qquad F(s) = \exp\Big\{\sum_{p\le x}\frac{f(p)p^{-it}}{p}+O(1)\Big\}\qquad \Big(\sigma = 1+\frac{1}{\log x}\Big).$$

The well-known coefficient formula applied to the series representation of $F'(s)$ gives

$$(9)\qquad \Big|\sum_{n\le y} f(n)\log n\log\frac{y}{n}\Big| = \Big|\frac{1}{2\pi i}\int_{(\sigma)}\frac{y^sF'(s)}{s^2}\,ds\Big| \le \frac{ey}{2\pi}\int_{(\sigma)}\Big|\frac{F'(s)}{s^2}\Big|\,dt$$

for $y\le x$, choosing $\sigma = 1+\dfrac{1}{\log x}$. By Schwarz's inequality

$$\int_{(\sigma)}\Big|\frac{F'(s)}{s^2}\Big|\,dt \le \sqrt{\int_{(\sigma)}\Big|\frac{F'}{F}(s)\Big|^2\frac{dt}{|s|^{3/2}}\int_{(\sigma)}\frac{|F(s)|^2}{|s|^{5/2}}\,dt}.$$

Owing to (17) of [4] the first integral is $O\left(\dfrac{1}{\sigma-1}\right) = O(\log x)$. As to the second,

$$\int_{\substack{(\sigma)\\ |t|\le T}}\frac{|F(s)|^2}{|s|^{5/2}}\,dt \le \max_{\substack{(\sigma)\\ |t|\le T}}|F(s)|^{1/2}\int_{(\sigma)}\frac{|F(s)|^{3/2}}{|s|^{5/2}}\,dt = \max_{\substack{(\sigma)\\ |t|\le T}}|F(s)|^{1/2}\,O\left(\frac{1}{(\sigma-1)^{1/2}}\right)$$

by (19) and (20) of [4] and

$$\int_{\substack{(\sigma)\\ |t|\ge T}}\frac{|F(s)|^2}{|s|^{5/2}}\,dt \le \frac{1}{T^{1/2}}\int_{(\sigma)}\frac{|F(s)|^2}{|s|^2}\,dt = \frac{1}{T^{1/2}}\,O\left(\frac{1}{\sigma-1}\right)$$

by the formula in the middle of page 379 in [4]. Here, by (8)

$$F(s) = O\left(\exp\Big\{\sum_{p\le x}\frac{\mathrm{Re} f(p)p^{-it}}{p}\Big\}\right) = O\big(\log x\exp\{-m(f,T)\}\big)$$

$$\Big(\sigma = 1+\frac{1}{\log x},\ |t|\le T\Big)$$

implying

$$\int_{(\sigma)}\frac{|F(s)|^2}{|s|^{5/2}}\,dt = O\big(\log x\exp\{-\tfrac12 m(f,T)\}\big)+O(\sqrt{\log x})$$
$$= O\big(\log x\exp\{-\tfrac14 m(f,T)\}\big),$$

$m(f,T)$ being at most $2\sum_{p\le x}1/p = 2\log\log x+O(1)$. Putting our estimates into (9),

$$\sum_{n\le y} f(n)\log n\log\frac{y}{n} = O\big(y\log x\exp\{-\tfrac18 m(f,T)\}\big)$$

for $y\le x$. Applying this also with $y/1+\delta$ and subtracting

$$\sum_{n\le y} f(n)\log n\log(1+\delta)+O\Big(\sum_{y/1+\delta<n\le y}\log n\log(1+\delta)\Big)$$
$$= O\big(y\log x\exp\{-\tfrac18 m(f,T)\}\big),$$

$$\gamma(y) \overset{\text{def}}{=} \sum_{n\le y} f(n)\log n = O\left(\frac{y\log x\exp\{-\frac18 m(f,T)\}}{\log(1+\delta)}\right)+O(\delta y\log x)\quad (y\le x).$$

The optimal choice $\delta = \exp\{-\frac{1}{16}m(f,T)\}$ yields the bound

$$y\log x\exp\{-\tfrac{1}{16}m(f,T)\}$$

and by partial integration

$$\sum_{n\le y} f(n) = 1+\int_2^y\frac{d\gamma(u)}{\log u} = 1+\frac{\gamma(y)}{\log y}+\int_2^y\frac{\gamma(u)}{u\log^2u}\,du$$
$$= O\left(y\,\frac{\log x}{\log y}\exp\{-\tfrac{1}{16}m(f,T)\}\right).$$

This with $y = x$ is the statement for completely multiplicative functions.

In the general case, as is shown e.g. in [4] (pp. 368–370),

$$f(n) = \sum_{d|n} h(d)f^*\left(\frac{n}{d}\right)$$

where $\sum_{d=1}^{\infty}|h(d)|/d^{3/4}\le c_{12}$ (1) and $f^*(n)$ is completely multiplicative with $f^*(p) = f(p)$ (hence also $m(f^*,T) = m(f,T)$). This implies

$$\sum_{n\le x} f(n) = \sum_{d\le x} h(d)\sum_{k\le x/d} f^*(k)$$
$$= O\Big(\sum_{d\le\sqrt{x}}|h(d)|\frac{x}{d}\exp\{-\tfrac{1}{16}m(f^*,T)\}\Big)+O\Big(\sum_{d>\sqrt{x}}|h(d)|\frac{x}{d}\Big)$$
$$= O\big(x\exp\{-\tfrac{1}{16}m(f,T)\}\big)+O(x^{7/8})$$

and the proof is complete.

(1) In [4] with 1 in place of 3/4, but the same proof applies.

References

[1] P. Erdös, I. Ruzsa, Jr., and A. Sárközi, *On the number of solutions of $f(n) = a$ for additive functions*, Acta Arith. 24 (1973), pp. 1–9.
[2] G. Halász, *On the distribution of additive and the mean values of multiplicative functions*, Studia Sci. Math. Hung. 6 (1971), pp. 211–233.
[3] — *Remarks to my paper "On the distribution of additive and the mean values of multiplicative functions"*, Acta Math. Acad. Sci. Hung. 23 (1972), pp. 425–432.
[4] — *Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen*, Acta Math. Acad. Sci. Hung. 19 (1968), pp. 365–403.
[5] A. Selberg, *Note on a paper by L. G. Sathe*, J. Indian Math. Soc. 18 (1954), pp. 83–87.
[6] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, Oxford 1951.

Received on 15. 6. 1973 (414)

ACTA ARITHMETICA
XXVII (1975)

On the difference between consecutive prime numbers

by

S. Uchiyama (Okayama)

Yuriĭ Vladimirovič Linnik in memoriam

Let p_n denote the nth prime number, and define

$$E = \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\log p_n}.$$

The crude estimate $E \leqslant 1$ follows, as is easily seen, from the fact that $p_n \sim n\log n$ $(n\to\infty)$, which is equivalent to the prime number theorem. The long-standing conjecture that states that $E = 0$, which is obviously the case if there exist infinitely many pairs of primes p, q with a fixed non-zero difference, remains still unproved. The best result on the size of E that is known so far is due to G. Z. Pil'tjaĭ [2], who showed that

$$E \leqslant \tfrac{1}{4}(2\sqrt{2}-1) = 0.457106\ldots \tag{1}$$

improving a previous result of E. Bombieri and H. Davenport [1],

$$E \leqslant \tfrac{1}{8}(2+\sqrt{3}) = 0.466506\ldots$$

The purpose of the present article is to make a further improvement on these results. Indeed, we shall prove the following

THEOREM.(1) *We have*

$$E \leqslant \frac{9-\sqrt{3}}{16} = 0.454246\ldots \tag{2}$$

An inspection of Pil'tjaĭ's paper [2] suggests a possibility of ameliorating the estimate (1) for E by an alternative choice of the various parameters therewith concerned. Our proof of (2) is thus a slight modification

(1) After the present paper had been submitted the writer learned from a kind letter of Prof. A. Schinzel that M. N. Huxley obtained, by improving Pil'tjaĭ's argument, the inequality $E \leqslant (4+\pi)/16 = 0.446349\ldots$, which supersedes (2).