Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 31 (3—4), (1978), pp. 355—386.

ON DIFFERENCE SETS OF SEQUENCES OF INTEGERS. III

By
A. SÁRKÖZY (Budapest)

1. Let $\mathscr{B}$ be a set of positive integers $b_1<b_2<\dots$. A set of positive integers $u_1<u_2<\dots$ will be called an $\mathscr{A}$-set relative to $\mathscr{B}$ if its difference set does not contain an element of $\mathscr{B}$; in other words, if

$$u_x-u_y=b_z \tag{1}$$

is not solvable in positive integers x, y, z.

L. Lovász conjectured that if $u_1<u_2<\dots$ is an $\mathscr{A}$-set relative to the set of the squares of the positive integers (i.e. $u_x-u_y=z^2$ is not solvable in positive integers x, y, z) then

$$\sum_{u_i\leqq x} 1 = o(x) \tag{2}$$

must hold. In Part I of this series (see [10]), I proved this conjecture in the following sharper form: if $u_1<u_2<\dots$ is an $\mathscr{A}$-set relative to the set of the squares then

$$\sum_{u_i\leqq x} 1 = O\left(x\frac{(\log\log x)^{2/3}}{(\log x)^{1/3}}\right). \tag{3}$$

I proved this theorem by adapting that version of the Hardy—Littlewood method which has been elaborated by K. F. ROTH in [4] and [5], in order to prove that if a set of positive integers $u_1<u_2<\dots$ does not contain an arithmetic progression of three terms, then (2) must hold, more exactly,

$$\sum_{u_i\leqq x} 1 = O\left(\frac{x}{\log\log x}\right). \tag{4}$$

(In Part II of this series, I gave a lower estimate for

$$\max \sum_{u_i\leqq x} 1$$

where the maximum is taken for those sets $u_1<u_2<\dots$ which form an $\mathscr{A}$-set relative to the set $1^2, 2^2, \dots, n^2, \dots$; see [11].)

In the case of the arithmetic progressions of three terms, we may use the following simple fact:

(i) A set $a+qu_1, a+qu_2, \dots, a+qu_t$ (where a is an integer and $t, q, u_1, u_2, \dots, u_t$ are positive integers) does not contain an arithmetic progression of three terms if and only if also the set $u_1, u_2, \dots, u_t$ has this property.

This fact plays a role of basic importance in the proof of (4). In the proof of (3), I could replace this assertion by the following one:

(ii) A set $a+q^2u_1,\ a+q^2u_2,\ \ldots,\ a+q^2u_t$ (where a is an integer and $t, q, u_1, u_2, \ldots, u_t$ are positive integers) is an $\mathscr{A}$-set relative to the set of the squares if and only if also the set $u_1, u_2, \ldots, u_t$ has this property.

(Note that here we have q^2 in place of q.)

Starting out from (3), one might like to show that (2) must hold also for sequences $u_1<u_2<\ldots$ which form an $\mathscr{A}$-set relative to certain other fixed set $b_1<b_2<\ldots$, e.g. relative to

$$b_i = i^k \tag{5}$$

(where $k\geqq 3$ is a fixed integer and $i=1, 2, \ldots$),

$$b_i = f(i) \tag{6}$$

(wher $f(x)$ is a fixed polynomial with integral coefficients and $i=1, 2, \ldots$) and

$$b_i = p_i \tag{7}$$

(where p_i denotes the i^{th} prime number and $i=1, 2, \ldots$), respectively.

The case (5) can be treated in the same way as the special case $k=2$; namely, the analogue of (ii) holds also in the general case $k\geqq 2$ with q^k in place of q^2. Thus it can be shown by the method used in [10] that if the set $u_1<u_2<\ldots$ forms an $\mathscr{A}$-set relative to the set (5) (also in case $k\geqq 3$) then (2) must hold.

On the other hand, in cases (6) and (7), simple counter examples can be given. Namely, let $f(x)=x^2+1$ and $u_1=6, u_2=12, \ldots, u_i=6i, \ldots$. Then (2) does not hold, however, $3\,|\,u_x-u_y$ and $6\,|\,u_x-u_y$ thus $u_x-u_y\neq b_z=z^2+1$ and $u_x-u_y\neq b_z=p_z$ (for $1\leqq y<x,\ z=1, 2, \ldots$).

P. Erdős raised the conjecture that if

$$b_i = i^2-1 \tag{8}$$

(i.e. $f(x)=x^2-1$ in (6)) respectively

$$b_i = p_i-1 \tag{9}$$

(for $i=1, 2, \ldots$), and $u_1<u_2<\ldots$ forms an $\mathscr{A}$-set relative to the set $b_1<b_2<\ldots$, then (2) must hold.

In both cases the difficulty is that an analogue of (i) or (ii) does not exist; thus we have to modify Roth's method. We shall be able to avoid this difficulty by using estimates for exponential sums of the form

$$\sum_{\substack{b_i\leqq x\\ q\,|\,b_i}} e(b_i\alpha) \tag{10}$$

where q is small in terms of x. (Throughout this paper, we use the notation $e^{2\pi i\alpha} = e(\alpha)$ where α is real.)

Since the cases (8) and (9) can be investigated analogously, we are going to discuss only the case (9). The remaining part of this paper will be devoted to the discussion of this case, i.e. the solvability of the equation

$$u_x - u_y = p_z - 1. \tag{11}$$

Consequently, we shall write briefly "$\mathscr{A}$-set" instead of "$\mathscr{A}$-set relative to the set $p_1-1, p_2-1, \ldots, p_i-1, \ldots$".

For $x=1, 2, \ldots,$ let $A(x)$ denote the greatest number of integers that can be selected from $1, 2, \ldots, x$ to form an $\mathscr{A}$-set and let us write

$$a(x) = \frac{A(x)}{x}.$$

We shall prove the following

THEOREM.

$$a(x) = O\left(\frac{(\log\log\log x)^3 (\log\log\log\log x)}{(\log\log x)^2}\right). \tag{12}$$

Throughout this paper, we use the following notations:

We denote the distance of the real number x from the nearest integer by $\|x\|$, i.e. $\|x\| = \min\{x-[x], [x]+1-x\}$. If a, b are real numbers and $b>0$ then we define the symbol $\min\left\{a, \frac{b}{0}\right\}$ by

$$\min\left\{a, \frac{b}{0}\right\} = a. \tag{13}$$

$C, c_1, c_2, \ldots, M_0, M_1, \ldots$ will denote (positive) absolute constants. We shall use also Vinogradov's notation: if f and g are two functions such that $g \geqq 0$ and there exists an absolute constant C satisfying $|f| \leqq Cg$ then we write $f \ll g$.

2. In this section, we estimate exponential sums of the form

$$S(\alpha) = S_N(\alpha) = \sum_{p \leqq N} (\log p) e((p-1)\alpha)$$

and

$$P(\alpha) = P_{M,q}(\alpha) = \sum_{\substack{\frac{p-1}{q} \leqq M \\ q \mid p-1}} (\log p) e\left(\frac{p-1}{q}\alpha\right). \tag{14}$$

(Here and in what follows, we shall leave the indices if this cannot cause confusion.)

LEMMA 1. *Let u be an arbitrary positive real number, M a positive integer for which $M \to +\infty$, and b, q, m integers satisfying*

$$1 \leqq b < (\log M)^u \tag{15}$$

and

$$1 \leqq q < (\log M)^u. \tag{16}$$

Then there exists an absolute constant $c_1>0$ *such that*

$$(17)\qquad \sum_{\substack{\frac{p-1}{q}\leqq M \\ q|p-1 \\ \frac{p-1}{q}\equiv m \pmod b}} \log p = \begin{cases} \dfrac{Mq}{\varphi(bq)} + O\left(Me^{-c_1\sqrt{\log M}}\right) & \text{for} \quad (mq+1, b)=1 \\ O\left(Me^{-c_1\sqrt{\log M}}\right) & \text{for} \quad (mq+1, b)>1 \end{cases}$$

(*where* c_1 *and the implicite constant in the error term may depend on* u *but not on* b, q, m).

PROOF. The conditions $q|p-1$ and $\dfrac{p-1}{q}\equiv m \pmod b$ can be rewritten in the equivalent form

$$(18)\qquad p \equiv mq+1 \pmod{bq}.$$

Thus for $(mq+1, bq)=1$, i.e. $(mq+1, b)=1$, we have to show that

$$\sum_{\substack{p\leqq Mq+1 \\ p\equiv mq+1 \pmod{bq}}} \log p = \frac{Mq}{\varphi(bq)} + O\left(Me^{-c_1\sqrt{\log M}}\right);$$

but this is a consequence of the prime number theorem of the arithmetic progressions of small $(<(\log M)^v)$ modulus (see e.g. [3], pp. 136 and 144).

For $(mq+1, bq)>1$, i.e. $(mq+1, b)>1$, (18) implies that $(mq+1, b)|p$. Hence, $(mq+1, b)$ is a prime number and $p=(mq+1, b)$. Thus in this case, the left hand side of (17) consists of the single term

$$\log p = \log (mq+1, b) \leqq \log b < \log (\log M)^u = u \log\log M = o\left(Me^{-c_1\sqrt{\log M}}\right)$$

which completes the proof of Lemma 1.

LEMMA 2. *Let* u *be an arbitrary positive real number,* M *a positive integer for which* $M\to+\infty$, *and* a, b, q *integers satisfying* (15), (16) *and* $(a, b)=1$. *Let us define the integer* $m_{b,q}$ *for* $(b, q)=1$ *by*

$$(19)\qquad m_{b,q}q+1 \equiv 0 \pmod b \quad \text{and} \quad 0 \leqq m_{b,q} \leqq b-1.$$

Then there exists an absolute constant $c_2>0$ *such that*

$$(20)\qquad P\left(\frac{a}{b}\right) = P_{M,q}\left(\frac{a}{b}\right) =$$

$$= \begin{cases} \dfrac{Mq}{\varphi(bq)}\mu(b)e\left(m_{b,q}\dfrac{a}{b}\right) + O\left(Me^{-c_2\sqrt{\log M}}\right) & \text{for} \quad (b, q)=1 \\ O\left(Me^{-c_2\sqrt{\log M}}\right) & \text{for} \quad (b, q)>1 \end{cases}$$

(*where* c_2 *and the implicite constant in the error term may depend on* u *but not on* a, b, q).

PROOF. By (15) and Lemma 1,

$$P\left(\frac{a}{b}\right) = P_{M,q}\left(\frac{a}{b}\right) = \sum_{\substack{\frac{p-1}{q} \leqq M \\ q \mid p-1}} (\log p) e\left(\frac{p-1}{q} \cdot \frac{a}{b}\right) = \tag{21}$$

$$= \sum_{m=0}^{b-1} e\left(m\frac{a}{b}\right) \sum_{\substack{\frac{p-1}{q} \leqq M \\ q \mid p-1 \\ \frac{p-1}{q} \equiv m \pmod b}} \log p =$$

$$= \sum_{\substack{0 \leqq m \leqq b-1 \\ (mq+1, b)=1}} e\left(m\frac{a}{b}\right) \frac{Mq}{\varphi(bq)} + O\left(\sum_{m=0}^{b-1} M e^{-c_1\sqrt{\log M}}\right) =$$

$$= \frac{Mq}{\varphi(bq)} \sum_{\substack{0 \leqq m \leqq b-1 \\ (mq+1.\, b)=1}} e\left(m\frac{a}{b}\right) + O\left((\log M)^u M e^{-c_1\sqrt{\log M}}\right) =$$

$$= \frac{Mq}{\varphi(bq)} \sum_{\substack{0 \leqq m \leqq b-1 \\ (mq+1, b)=1}} e\left(m\frac{a}{b}\right) + O\left(M e^{-c_2\sqrt{\log M}}\right).$$

Here

$$\sum_{\substack{0 \leqq m \leqq b-1 \\ (mq+1, b)=1}} e\left(m\frac{a}{b}\right) = \sum_{m=0}^{b-1} e\left(m\frac{a}{b}\right) \sum_{d \mid (mq+1,\, b)} \mu(d) = \tag{22}$$

$$= \sum_{d \mid b} \mu(d) \sum_{\substack{0 \leqq m \leqq b-1 \\ d \mid mq+1}} e\left(m\frac{a}{b}\right).$$

Let m_0 denote the least non-negative integer m for which $d \mid mq+1$ holds. Then

$$m_0 q + 1 \equiv 0 \pmod d, \tag{23}$$

and $d \mid mq+1$ holds if and only if

$$(mq+1) - (m_0 q+1) = (m-m_0)q \equiv 0 \pmod d. \tag{24}$$

By (23),

$$(d, q) = 1. \tag{25}$$

(24) and (25) imply that $d|mq+1$ holds if and only if $m-m_0\equiv 0 \pmod d$. Thus with respect to $(a, b)=1$, the inner sum in (22) is

$$\sum_{\substack{0\leqq m\leqq b-1\\ d|mq+1}} e\left(m\frac{a}{b}\right)=\sum_{j=0}^{b/d-1} e\left((m_0+jd)\frac{a}{b}\right)=$$

$$=\begin{cases}\dfrac{b}{d}e\left(m_0\dfrac{a}{b}\right) & \text{for} \quad b|da \quad \text{i.e.} \quad b|d\\ e\left(m_0\dfrac{a}{b}\right)\dfrac{1-e\left(\dfrac{b}{d}\cdot d\dfrac{a}{b}\right)}{1-e\left(d\dfrac{a}{b}\right)}=0 & \text{for} \quad b\nmid d.\end{cases}$$

Hence, the inner sum in (22) is different from 0 only if $b|d$; but by $d|b$, this implies that $b=d$, and by (25), also $(b, q)=1$ must hold. Thus we obtain from (22) that

$$\sum_{\substack{0\leqq m\leqq b-1\\ (mq+1,\,b)=1}} e\left(m\frac{a}{b}\right)=$$

$$=\begin{cases}\mu(b)\displaystyle\sum_{\substack{0\leqq m\leqq b-1\\ b|mq+1}} e\left(m\frac{a}{b}\right)=\mu(b)\cdot\frac{b}{b}e\left(m_0\frac{a}{b}\right)=\mu(b)e\left(m_0\frac{a}{b}\right) & \text{for} \quad (b, q)=1,\\ 0 \quad \text{for} \quad (b, q)>1\end{cases}$$

where m_0 satisfies (23), i.e. $m_0q+1\equiv 0 \pmod b$; hence, $m_0=m_{b,q}$. Putting this into (21), we obtain (20) and the proof of Lemma 2 is complete.

LEMMA 3. *Let u be an arbitrary positive real number, M a positive integer for which $M\to+\infty$, a, b, q integers satisfying* (15), (16) *and* $(a, b)=1$, *finally, β any real number. Then*

$$P\left(\frac{a}{b}+\beta\right)=P_{M,q}\left(\frac{a}{b}+\beta\right)= \tag{26}$$

$$=\begin{cases}\dfrac{q}{\varphi(bq)}\mu(b)e\left(m_{b,q}\dfrac{a}{b}\right)\displaystyle\sum_{n=1}^{M}e(n\beta)+O\left((M|\beta|+1)Me^{-c_2\sqrt{\log M}}\right) & \textit{for} \quad (b, q)=1\\ O\left((M|\beta|+1)Me^{-c_2\sqrt{\log M}}\right) & \textit{for} \quad (b, q)>1\end{cases}$$

where $m_{b,q}$ is defined (*for* $(b, q)=1$) *by* (19).

PROOF. Applying Lemma 2, we obtain by partial summation that

$$P_{M,q}\left(\frac{a}{b}+\beta\right)=\sum_{\substack{\frac{p-1}{q}\leqq M\\ q\mid p-1}}(\log p)e\left(\frac{p-1}{q}\left(\frac{a}{b}+\beta\right)\right)= \tag{27}$$

$$=\sum_{\substack{\frac{p-1}{q}\leqq M\\ q\mid p-1}}\left\{(\log p)e\left(\frac{p-1}{q}\cdot\frac{a}{b}\right)\right\}e\left(\frac{p-1}{q}\beta\right)=$$

$$=\sum_{n=1}^{M}\left(P_{n,q}\left(\frac{a}{b}\right)-P_{n-1,q}\left(\frac{a}{b}\right)\right)e(n\beta)=$$

$$=\sum_{n=1}^{M}P_{n,q}\left(\frac{a}{b}\right)\left(e(n\beta)-e((n+1)\beta)\right)+P_{M,q}\left(\frac{a}{b}\right)e((M+1)\beta).$$

For $1\leqq n\leqq\sqrt{M}$,

$$\left|P_{n,q}\left(\frac{a}{b}\right)\right|=\left|\sum_{\substack{\frac{p-1}{q}\leqq n\\ q\mid p-1}}(\log p)e\left(\frac{p-1}{q}\cdot\frac{a}{b}\right)\right|\leqq$$

$$\leqq\sum_{k=1}^{n}\log(qn+1)=n\log(qn+1)<\sqrt{M}\log\left((\log M)^u\sqrt{M}+1\right)=O\left(\sqrt{M}\log M\right)$$

and

$$\left|\frac{nq}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\right|\leqq nq<\sqrt{M}(\log M)^u$$

(with respect to (16)).

For $\sqrt{M}<n\leqq M$, (16) implies that

$$1\leqq q<(\log M)^u<(\log n^2)^u=2^u(\log n)^u<(\log n)^{2u}$$

(if M is sufficiently large depending on u) thus Lemma 2 can be applied with $2u$ and n in place of u and M, respectively.

Summarizing, we obtain from (27) (using Lemma 2) that for $(b,q)=1$

$$P_{M,q}\left(\frac{a}{b}+\beta\right)=\left\{\sum_{n=1}^{M}\frac{nq}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\left(e(n\beta)-e((n+1)\beta)\right)+\right.$$

$$\left.+\frac{Mq}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)e((M+1)\beta)\right\}+$$

$$+\left\{\sum_{n=1}^{M}\left(P_{n,q}\left(\frac{a}{b}\right)-\frac{nq}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\right)\left(e(n\beta)-e((n+1)\beta)\right)+\right.$$

$$\left.+\left(P_{M,q}\left(\frac{a}{b}\right)-\frac{Mq}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\right)e((M+1)\beta)\right\}=$$

$$= \frac{q}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\sum_{n=1}^{M}e(n\beta)+$$

$$+\sum_{n=1}^{[\sqrt{M}]}O\left(\sqrt{M}\log M+\sqrt{M}(\log M)^{a}\right)\left|e(n\beta)-e\left((n+1)\beta\right)\right|+$$

$$+\sum_{n=[\sqrt{M}]+1}^{M}O\left(ne^{-c_2\sqrt{\log n}}\right)\left|e(n\beta)-e\left((n+1)\beta\right)\right|+O\left(Me^{-c_2\sqrt{\log M}}\right)=$$

$$= \frac{q}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\sum_{n=1}^{M}e(n\beta)+\sum_{n=1}^{[\sqrt{M}]}O\left(Me^{-c_2\sqrt{\log M}}|\beta|\right)+$$

$$+\sum_{n=[\sqrt{M}]+1}^{M}O\left(Me^{-c_2\sqrt{\log M}}|\beta|\right)+O\left(Me^{-c_2\sqrt{\log M}}\right)=$$

$$= \frac{q}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\sum_{n=1}^{M}e(n\beta)+O\left((M|\beta|+1)Me^{-c_2\sqrt{\log M}}\right)$$

while for $(b,q)>1$,

$$P_{M,q}\left(\frac{a}{b}+\beta\right)=\sum_{n=1}^{[\sqrt{M}]}P_{n,q}\left(\frac{a}{b}\right)\left(e(n\beta)-e\left((n+1)\beta\right)\right)+$$

$$+\sum_{n=[\sqrt{M}]+1}^{M}P_{n,q}\left(\frac{a}{b}\right)\left(e(n\beta)-e\left((n+1)\beta\right)\right)+P_{M,q}\left(\frac{a}{b}\right)e\left((M+1)\beta\right)=$$

$$=\sum_{n=1}^{[\sqrt{M}]}O\left(\sqrt{M}\log M\right)\left|e(n\beta)-e\left((n+1)\beta\right)\right|+$$

$$+\sum_{n=[\sqrt{M}]+1}^{M}O\left(ne^{-c_2\sqrt{\log n}}\right)\left|e(n\beta)-e\left((n+1)\beta\right)\right|+O\left(Me^{-c_2\sqrt{\log M}}\right)=$$

$$=\sum_{n=1}^{[\sqrt{M}]}O\left(Me^{-c_2\sqrt{\log M}}|\beta|\right)+\sum_{n=[\sqrt{M}]+1}^{M}O\left(Me^{-c_2\sqrt{\log M}}|\beta|\right)+$$

$$+O\left(Me^{-c_2\sqrt{\log M}}\right)=O\left((M|\beta|+1)Me^{-c_2\sqrt{\log M}}\right)$$

since

$$\left|e(n\beta)-e\left((n+1)\beta\right)\right|=\left|1-e(\beta)\right|=\left|e(-\beta/2)-e(\beta/2)\right|=2\left|\sin\pi\beta\right|\leqq 2\pi|\beta|$$

and the proof of Lemma 3 is complete.

LEMMA 4. *If a, b are integers such that $a\leqq b$, and β is an arbitrary real number then*

$$\left|\sum_{k=a}^{b}e(k\beta)\right|\leqq\min\left\{b-a+1,\frac{1}{2\|\beta\|}\right\}.$$

(*For $\|\beta\|=0$, the right hand side is defined by* (13).)

This lemma is identical to Lemma 1 in [10].

LEMMA 5. *Let u be an arbitrary positive real number. There exist constants M_0, $c_3>0$ (which may depend on u) such that if $M>M_0$, furthermore, a, b, q are integers satisfying* (15), (16) *and* $(a, b)=1$, *finally, β is a real number satisfying*

$$|\beta| \leqq \frac{e^{c_3\sqrt{\log M}}}{M}, \tag{28}$$

then

$$\left|P\left(\frac{a}{b}+\beta\right)\right| = \left|P_{M,q}\left(\frac{a}{b}+\beta\right)\right| < \begin{cases} 2\dfrac{Mq}{\varphi(b)\varphi(q)} & \text{for} \quad |\beta| \leqq \dfrac{1}{M} \\ \dfrac{q}{\varphi(b)\varphi(q)|\beta|} & \text{for} \quad \dfrac{1}{M} \leqq |\beta|. \end{cases} \tag{29}$$

PROOF. We are going to apply Lemma 3.

For $(b, q)=1$, the main term in (26) in Lemma 3 can be estimated in the following way, by using Lemma 4 (and with respect to (28)):

$$\left|\frac{q}{\varphi(bq)}\mu(b)e\left(m_{b,q}\frac{a}{b}\right)\sum_{n=1}^{M}e(n\beta)\right| \leqq$$

$$\leqq \frac{q}{\varphi(bq)}\min\left\{M, \frac{1}{2\|\beta\|}\right\} = \frac{q}{\varphi(b)\varphi(q)}\min\left\{M, \frac{1}{2|\beta|}\right\} \leqq$$

$$\leqq \begin{cases} \dfrac{Mq}{\varphi(b)\varphi(q)} & \text{for} \quad |\beta| \leqq \dfrac{1}{M} \\ \dfrac{q}{2\varphi(b)\varphi(q)|\beta|} & \text{for} \quad \dfrac{1}{M} \leqq |\beta|. \end{cases}$$

Thus Lemma 3 yields that

$$\left|P_{M,q}\left(\frac{a}{b}+\beta\right)\right| < O\left((M|\beta|+1)Me^{-c_2\sqrt{\log M}}\right) + \begin{cases} \dfrac{Mq}{\varphi(b)\varphi(q)} & \text{for} \quad |\beta| \leqq \dfrac{1}{M} \\ \dfrac{q}{2\varphi(b)\varphi(q)|\beta|} & \text{for} \quad \dfrac{1}{M} \leqq |\beta|. \end{cases}$$

To obtain (29) from this inequality, it suffices to show that here the first term on the right (the $O(\ldots)$ term) is less than the second term. The first term is the greatest and the second is the least if $|\beta|$ is the possibly greatest, i.e. $|\beta|=e^{c_3\sqrt{\log M}}/M$. Then the first term is

$$O\left((e^{c_3\sqrt{\log M}}+1)Me^{-c_2\sqrt{\log M}}\right) = O\left(Me^{(c_3-c_2)\sqrt{\log M}}\right) \tag{30}$$

while the second term is (with respect to (16) and for large M)

$$\frac{qM}{2\varphi(b)\varphi(q)e^{c_3\sqrt{\log M}}} > \frac{M}{2be^{c_3\sqrt{\log M}}} > \frac{M}{2(\log M)^u e^{c_3\sqrt{\log M}}} > Me^{-2c_3\sqrt{\log M}}.$$

For $c_3=c_2/4$ and $M>M_1(u)$, the latter is greater than (30) and Lemma 5 is proved.

LEMMA 6. *If X, Y are real numbers, a, b integers and α a real number such that $Y \leqq b \leqq X/Y$, $1 \leqq Y \leqq X^{1/4}$, $(a, b)=1$ and*

$$\left|\alpha-\frac{a}{b}\right| \leqq \frac{1}{b^2}$$

then

$$|S_X(\alpha)| = \Big|\sum_{p \leqq X} (\log p) e((p-1)\alpha)\Big| = \Big|\sum_{p \leqq X} (\log p) e(p\alpha)\Big| \ll XY^{-1/2}(\log X)^{17}.$$

This is essentially a consequence of Theorems 1 and 3 of VINOGRADOV in [12], Chapter IX; see also MONTGOMERY [1], Chapter 16, and MONTGOMERY—VAUGHAN [2], Lemma 3.1.

LEMMA 7. *If $M(>0)$, q, a, b are integers and α is a real number satisfying*

(31) $$1 \leqq q \leqq \log M$$

and

(32) $$(a, b) = 1,$$

furthermore, writing

(33) $$Q = M(\log M)^{-41},$$

also

(34) $$2(\log M)^{40} \leqq b \leqq Q$$

and

(35) $$\left|\alpha-\frac{a}{b}\right| < \frac{1}{bQ}$$

hold then for large M,

(36) $$|P(\alpha)| = |P_{M,q}(\alpha)| \ll \frac{M}{(\log M)^2}.$$

PROOF.

(37) $$|P_{M,q}(\alpha)| = \Big|\sum_{\substack{p-1 \\ \overline{q} \leqq M \\ q \mid p-1}} (\log p) e\left((p-1)\frac{\alpha}{q}\right)\Big| =$$

$$= \Big|\sum_{p \leqq qM+1} (\log p) e\left((p-1)\frac{\alpha}{q}\right)\left\{\frac{1}{q}\sum_{j=0}^{q-1} e\left((p-1)\frac{j}{q}\right)\right\}\Big| =$$

$$= \frac{1}{q}\Big|\sum_{j=0}^{q-1}\sum_{p \leqq qM+1} (\log p) e\left((p-1)\frac{\alpha+j}{q}\right)\Big| =$$

$$= \frac{1}{q}\Big|\sum_{j=0}^{q-1} S_{qM+1}\left(\frac{\alpha+j}{q}\right)\Big| \leqq \frac{1}{q}\sum_{j=0}^{q-1}\Big|S_{qM+1}\left(\frac{\alpha+j}{q}\right)\Big|.$$

Let us write $\gamma=\frac{\alpha+j}{q}$. By Dirichlet's theorem, there exist integers A, B such that

$$(A, B) = 1, \tag{38}$$

$$1 \leqq B \leqq 2qQ \tag{39}$$

and

$$\left|\gamma-\frac{A}{B}\right| < \frac{1}{2BqQ}; \tag{40}$$

by (39) and (40), also

$$\left|\gamma-\frac{A}{B}\right| < \frac{1}{B^2} \tag{41}$$

holds.

We are going to show that these conditions imply that

$$B > \frac{1}{2}b. \tag{42}$$

Let us assume indirectly that

$$B \leqq \frac{1}{2}b. \tag{43}$$

By (35), γ can be written in the form

$$\gamma = \frac{\alpha+j}{q} = \frac{\frac{a}{b}+\frac{\theta_1}{bQ}+j}{q} = \frac{a+bj}{bq}+\frac{\theta_1}{bqQ} \tag{44}$$

where $|\theta_1|<1$. Let us define the integer U and the positive integer V by

$$\frac{a+bj}{bq} = \frac{U}{V}, \tag{45}$$

$$(U, V) = 1. \tag{46}$$

By (32), $(a+bj, b)=1$, thus

$$(a+bj, bq) \leqq q. \tag{47}$$

(45), (46) and (47) imply that

$$b \leqq V \leqq bq. \tag{48}$$

By (40), γ can be written in the form

$$\gamma = \frac{A}{B}+\frac{\theta_2}{2BqQ} \tag{49}$$

where $|\theta_2|<1$.

(44) and (49) yield that

$$\gamma = \frac{U}{V}+\frac{\theta_1}{bqQ} = \frac{A}{B}+\frac{\theta_2}{2BqQ},$$

hence, with respect to (34) and (48),

$$\left|\frac{U}{V}-\frac{A}{B}\right| \leqq \frac{|\theta_1|}{bqQ}+\frac{|\theta_2|}{2BqQ} < \frac{1}{bqQ}+\frac{1}{2BqQ} \leqq \tag{50}$$

$$\leqq \frac{1}{bqQ}+\frac{1}{2Bqb} \leqq \frac{1}{bqQ}+\frac{1}{2BV}.$$

On the other hand, we obtain from (38), (43), (46) and (48) that

$$\frac{U}{V} \neq \frac{A}{B},$$

thus

$$\left|\frac{U}{V}-\frac{A}{B}\right| = \frac{|UB-VA|}{VB} \geqq \frac{1}{VB}. \tag{51}$$

(50) and (51) yield that

$$\frac{1}{VB} < \frac{1}{bqQ}+\frac{1}{2BV}, \quad \frac{1}{2VB} < \frac{1}{bqQ},$$

hence, with respect to (34), (43) and (48),

$$1 < 2\frac{VB}{bqQ} = 2V\cdot\frac{1}{qQ}\cdot\frac{B}{b} \leqq 2bq\cdot\frac{1}{qQ}\cdot\frac{1}{2} = \frac{b}{Q} \leqq 1.$$

Thus the indirect assumption (43) leads to a contradiction, which proves (42).

Let us write $X=qM+1$, $Y=(\log M)^{40}$. Then for large M,

$$1 \leqq Y = (\log M)^{40} < (M+1)^{1/4} \leqq X^{1/4}, \tag{52}$$

furthermore, by (34) and (42),

$$B > \frac{1}{2}b \geqq (\log M)^{40} = Y, \tag{53}$$

finally, by (33) and (39),

$$B \leqq 2qQ = 2qM(\log M)^{-41} \leqq 2(qM+1)(\log M)^{-41} < \tag{54}$$

$$< (qM+1)(\log M)^{-40} = X/Y.$$

In view of (38), (41), (52), (53) and (54), Lemma 6 can be applied with $qM+1$, $(\log M)^{40}$, A, B and γ in place of X, Y, a, b and α. With respect to (31), we obtain that

$$|S_{qM+1}(\gamma)| = \left|S_{qM+1}\left(\frac{\alpha+j}{q}\right)\right| \ll (qM+1)\big((\log M)^{40}\big)^{-1/2}\big(\log(qM+1)\big)^{17} \leqq$$

$$\leqq \big((\log M)M+1\big)(\log M)^{-20}\big\{\log\big((\log M)M+1\big)\big\}^{17} \ll$$

$$\ll (\log M)M(\log M)^{-20}(\log M)^{17} = M(\log M)^{-2}.$$

Putting this into (37), we obtain that

$$|P_{M,q}(\alpha)| \ll \frac{1}{q}\sum_{j=0}^{q-1} M(\log M)^{-2} = \frac{M}{(\log M)^2}$$

which completes the proof of Lemma 7.

LEMMA 8. *There exists an absolute constant* $c_4(>0)$ *such that for* $n \geqq 3$,

$$\varphi(n) > c_4 \frac{n}{n \log\log n}.$$

This lemma is well-known; see e.g. [3], p. 24.

LEMMA 9. *Let* q, M *be positive integers,* R *a real number such that*

$$q \leqq \log M \tag{55}$$

and

$$3 \leqq R \leqq \log M. \tag{56}$$

Let $S_{R,M}$ *denote the set of those real numbers* α *for which* $0 \leqq \alpha \leqq 1$ *holds and there do not exist integers* a, b *such that*

$$(a, b) = 1, \tag{57}$$

$$1 \leqq b < R \tag{58}$$

and

$$\left|\alpha - \frac{a}{b}\right| < \frac{1}{M} \cdot \frac{R}{\log\log R}. \tag{59}$$

Then for $\alpha \in S_{R,M}$ *and large* M,

$$|P_{M,q}(\alpha)| \ll \frac{qM}{\varphi(q)} \cdot \frac{\log\log R}{R}. \tag{60}$$

PROOF. Let us define Q by (33). By Dirichlet's theorem, for all $\alpha \in S_{R,M}$, there exist integers A, B such that

$$(A, B) = 1, \tag{61}$$

$$1 \leqq B \leqq Q \tag{62}$$

and

$$\left|\alpha - \frac{A}{B}\right| < \frac{1}{BQ}. \tag{63}$$

If $2(\log M)^{40} \leqq B$, then Lemma 7 can be applied, with A and B in place of a and b, respectively. We obtain that

$$|P_{M,q}(\alpha)| \ll \frac{M}{(\log M)^2}. \tag{64}$$

By (56), the right hand side of (60) can be estimated in the following way:

$$(65) \qquad \frac{q}{\varphi(q)} \cdot M \cdot \frac{\log\log R}{R} \geqq M \frac{\log\log R}{R} \geqq M \frac{\log\log\log M}{\log M} > \frac{M}{(\log M)^2}$$

for sufficiently large M. (64) and (65) yield (60).

If

$$(66) \qquad B < 2(\log M)^{40}$$

and M is large then we may apply Lemma 5 with $a=A$, $b=B$, $\beta = \alpha - \frac{A}{B}$ and $u=41$. Namely, for large M, (15) and (16) hold by (55) and (66). Furthermore, by (63),

$$|\beta| = \left|\alpha - \frac{A}{B}\right| < \frac{1}{BQ} \leqq \frac{1}{Q} = \frac{(\log M)^{41}}{M},$$

which implies (28) for sufficiently large M. Thus, in fact, all the assumptions in Lemma 5 hold. Applying Lemma 5, we obtain that for large M,

$$(67) \qquad |P_{M,q}(\alpha)| < \begin{cases} 2\dfrac{Mq}{\varphi(B)\varphi(q)} & \text{for} \quad |\beta| \leqq \dfrac{1}{M} \\ \dfrac{q}{\varphi(B)\varphi(q)|\beta|} & \text{for} \quad \dfrac{1}{M} \leqq |\beta|. \end{cases}$$

The right hand side is maximal for $|\beta| \leqq \frac{1}{M}$. Thus for $R \leqq B$, we obtain by applying Lemma 8 that

$$|P_{M,q}(\alpha)| < \frac{2Mq}{\varphi(B)\varphi(q)} \ll \frac{\log\log B}{B} \cdot \frac{Mq}{\varphi(q)} \ll \frac{\log\log R}{R} \cdot \frac{qM}{\varphi(q)}$$

(with respect to $R \geqq 3$).

Finally, if $B < R$ then $\alpha \in S_{R,M}$ implies that

$$(68) \qquad |\beta| = \left|\alpha - \frac{A}{B}\right| \geqq \frac{1}{M} \cdot \frac{R}{\log\log R}$$

which yields also $|\beta| > \frac{1}{M}$ since it can be shown easily that

$$\frac{R}{\log\log R} > 1$$

for $R \geqq 3$. Thus we obtain from (67) and (68) that

$$|P_{M,q}(\alpha)| < \frac{q}{\varphi(B)\varphi(q)|\beta|} \leqq \frac{q}{\varphi(q)} \cdot \frac{1}{|\beta|} \leqq \frac{q}{\varphi(q)} \cdot M \cdot \frac{\log\log R}{R}$$

which completes the proof of Lemma 9.

3. For arbitrary positive integers M, q, let

$$u_1 q, u_2 q, \ldots, u_T q \tag{69}$$

be *a* maximal $\mathscr{A}$-set selected from $q, 2q, \ldots, Mq$, and let

$$F(\alpha) = F_{M,q}(\alpha) = \sum_{k=1}^{T} e(u_k \alpha). \tag{70}$$

In this section, we estimate this function $F_{M,q}(\alpha)$.

For an integer b and positive integers m, x, let $A_{(b,m)}(x)$ denote the greatest number of integers that can be selected from $b+m, b+2m, \ldots, b+xm$ to form an $\mathscr{A}$-set (so that $A_{(0,1)}(x)=A(x)$).

LEMMA 10. *For any integers b, d and positive integers m, x, we have*

$$A_{(b,m)}(x) = A_{(d,m)}(x).$$

PROOF. This follows trivially from the fact that the numbers $b+u_1 m, b+u_2 m, \ldots, b+u_k m$ form an $\mathscr{A}$-set if and only if also the numbers $d+u_1 m, d+u_2 m, \ldots, d+u_k m$ do.

By Lemma 10, we may simplify the notation $A_{(b,m)}(x)$ in the following way: let us write $A_m(x)$ instead of $A_{(b,m)}(x)$, i.e. let

$$A_m(x) = A_{(b,m)}(x) \quad (\text{for } b = 0, \pm 1, \pm 2, \ldots).$$

Furthermore, let

$$a_m(x) = \frac{A_m(x)}{x},$$

so that $A(x)=A_1(x)$ and $a_1(x)=a(x)$; moreover, $T=A_q(M)$ in (69) and (70), thus

$$F(\alpha) = F_{M,q}(\alpha) = \sum_{k=1}^{A_q(M)} e(u_k \alpha). \tag{71}$$

Lemmas 11 and 12 follow trivially from the definitions of the functions $A_m(x)$ and $a_m(x)$, respectively.

LEMMA 11. *If m, x and y are positive integers such that $x \leqq y$ then $A_m(x) \leqq A_m(y)$.*

LEMMA 12. *For arbitrary positive integers m and x, we have $a_m(x) \leqq 1$.*

LEMMA 13. *For arbitrary positive integers m, x and y, we have*

$$A_m(x+y) \leqq A_m(x) + A_m(y), \tag{72}$$

$$A_m(xy) \leqq x A_m(y), \tag{73}$$

$$a_m(xy) \leqq a_m(y), \tag{74}$$

$$a_m(x) \leqq \left(1+\frac{y}{x}\right) a_m(y). \tag{75}$$

PROOF. By Lemma 10, the greatest number of integers that can be selected from $m, 2m, \dots, xm$ and $(x+1)m, (x+2)m, \dots, (x+y)m$ to form an $\mathscr{A}$-set, is $A_m(x)$ and $A_m(y)$, respectively; thus the greatest number of integers that can be selected from $m, 2m, \dots, xm, (x+1)m, (x+2)m, \dots, (x+y)m$ to form an $\mathscr{A}$-set, is $\leqq A_m(x)+A_m(y)$ which proves (72).

(73) is a consequence of (72).

Dividing (73) by xy, we obtain (74).

Finally, by Lemma 11 and (73),

$$A_m(x) \leqq A_m\left(\left(\left[\frac{x}{y}\right]+1\right)y\right) \leqq \left(\left[\frac{x}{y}\right]+1\right)A_m(y) \leqq$$

$$\leqq \left(\frac{x}{y}+1\right)A_m(y) = (x+y)\frac{A_m(y)}{y}.$$

Dividing by x, we obtain (75).

LEMMA 14. *Let q, b, t, M be positive integers, a an integer, α, β real numbers such that*

$$\alpha - \frac{a}{b} = \beta. \tag{76}$$

Let

$$F^*(\alpha) = F^*_{M,q}(\alpha) = \frac{a_{bq}(t)}{b}\left(\sum_{s=1}^{b} e\left(\frac{as}{b}\right)\right)\left(\sum_{j=1}^{M} e(\beta j)\right),$$

so that if $(a, b)=1$ then

$$F^*_{M,q}(\alpha) = \begin{cases} a_q(t)\sum_{j=1}^{M} e(\beta j) & \text{for} \quad b=1 \\ 0 \quad \text{for} \quad b>1 & (\text{where} \quad (a,b)=1). \end{cases} \tag{77}$$

Then there exists an absolute constant c_5 such that

$$|F_{M,q}(\alpha) - F^*_{M,q}(\alpha)| \leqq \left(a_{bq}(t) - a_q(M)M\right) + c_5\left(|\beta|Ma_{bq}(t) + a_q(t)\right)tb. \tag{78}$$

PROOF. We are going to show at first that

$$F_{M,q}(\alpha) = \frac{1}{tb}\sum_{s=1}^{b}\sum_{j=1}^{M}\sum_{\substack{j \leqq u_k < j+tb \\ u_k \equiv s \pmod b}} e(\alpha u_k) + O(a_q(t)tb). \tag{79}$$

Let us investigate the coefficient of $e(\alpha u_k)$ on the right hand side.

If $tb \leqq u_k \leqq M$ then we account $e(\alpha u_k)$ exactly tb times, namely for the following values of j:

$$j = u_k - tb + 1, u_k - tb + 2, \dots, u_k.$$

Thus the coefficient of $e(\alpha u_k)$ is

$$tb \cdot \frac{1}{tb} = 1$$

in this case (and its coefficient is the same on the left hand side).

If

$$1 \leqq u_k < tb \tag{80}$$

then we account $e(\alpha u_k)$ on the right of (79) for $j=1, 2, \ldots, u_k$, thus its coefficient is

$$(0 \leqq) \quad u_k \cdot \frac{1}{tb} < tb \cdot \frac{1}{tb} = 1$$

on the right and 1 on the left of (79). For the numbers u_k satisfying (80), the numbers $u_k q$ form an $\mathscr{A}$-set selected from $q, 2q, \ldots, tbq$ thus in view of (73) in Lemma 13, their number is

$$\leqq A_q(tb) \leqq A_q(t)b = a_q(t)tb.$$

These facts yield that, in fact, the error term in (79) is $O(a_q(t)tb)$.

The term $e(\alpha u_k)$ in the inner sum in (79) can be rewritten in the following way:

$$e(\alpha u_k) = e\left(\left(\frac{a}{b}+\beta\right)u_k\right) = e\left(\frac{au_k}{b}\right)e(\beta u_k) =$$

$$= e\left(\frac{as}{b}\right)e(\beta j)\,e(\beta(u_k-j)) = e\left(\frac{as}{b}\right)e(\beta j)(1+O(|\beta(u_k-j)|)) =$$

$$= e\left(\frac{as}{b}\right)e(\beta j)+O(|\beta(u_k-j)|) = e\left(\frac{as}{b}\right)e(\beta j)+O(|\beta|tb)$$

since $|u_k-j|<tb$ in the inner sum, and

$$|e(\gamma)-1| = |e(\gamma/2)-e(-\gamma/2)| = |2\sin\pi\gamma| \leqq 2|\pi\gamma| = 2\pi|\gamma|$$

for any real number γ.

Thus the inner sum in (79) can be estimated in the following way:

$$\sum_{\substack{j\leqq u_k<j+tb\\ u_k\equiv s \,(\mathrm{mod}\, b)}} e(\alpha u_k) = \sum_{\substack{j\leqq u_k<j+tb\\ u_k\equiv s \,(\mathrm{mod}\, b)}} \left(e\left(\frac{as}{b}\right)e(\beta j)+O(|\beta|tb)\right) = \tag{81}$$

$$= \left(e\left(\frac{as}{b}\right)e(\beta j)+O(|\beta|tb)\right)\sum_{\substack{j\leqq u_k<j+tb\\ u_k\equiv s \,(\mathrm{mod}\, b)}} 1.$$

Let us define the integer v by

$$v < j \leqq v+b, \quad v \equiv s \pmod b.$$

Then for the numbers u_k satisfying $j\leqq u_k<j+tb$ and $u_k\equiv s \,(\mathrm{mod}\, b)$, the numbers $u_k q$ form an $\mathscr{A}$-set selected from $vq+bq, vq+2bq, \ldots, vq+tbq$. Thus by Lemma 10,

$$\sum_{\substack{j\leqq u_k<j+tb\\ u_k\equiv s \,(\mathrm{mod}\, b)}} 1 \leqq A_{(vq,\,bq)}(t) = A_{bq}(t) = a_{bq}(t)t.$$

Hence, defining $D(j, t, b, s)$ by

$$\sum_{\substack{j\leqq u_k<j+tb\\ u_k\equiv s \,(\mathrm{mod}\, b)}} 1 = a_{bq}(t)t - D(j, t, b, s),$$

we have $D(j, t, q, s) \equiv 0$. Putting this into (81):

$$\sum_{\substack{j \leqq u_k < j+tb \\ u_k \equiv s \pmod b}} e(\alpha u_k) = \left(e\left(\frac{as}{b}\right) e(\beta j) + O(|\beta| tb)\right)(a_{bq}(t)t - D(j, t, b, s)) =$$

$$= e\left(\frac{as}{b}\right) e(\beta j)(a_{bq}(t)t - D(j, t, b, s)) + O(|\beta| a_{bq}(t) t^2 b).$$

Thus (79) yields that

$$(82) \quad F_{M,q}(\alpha) = \frac{1}{tb} \sum_{s=1}^{b} \sum_{j=1}^{M} \left\{ e\left(\frac{as}{b}\right) e(\beta j)(a_{bq}(t)t - D(j, t, b, s)) + O(|\beta| a_{bq}(t) t^2 b) \right\} +$$

$$+ O(a_q(t) tb) = \frac{a_{bq}(t)}{b} \left(\sum_{s=1}^{b} e\left(\frac{as}{b}\right) \right) \left(\sum_{j=1}^{M} e(\beta j) \right) -$$

$$- \frac{1}{tb} \sum_{s=1}^{b} \sum_{j=1}^{M} e\left(\frac{as}{b}\right) e(\beta j) D(j, t, b, s) +$$

$$+ O\left(\frac{1}{tb} \cdot b \cdot M \cdot |\beta| a_{bq}(t) t^2 b\right) + O(a_q(t) tb) =$$

$$= F^*_{M,q}(\alpha) - \frac{1}{tb} \sum_{s=1}^{b} \sum_{j=1}^{M} e\left(\frac{as}{b}\right) e(\beta j) D(j, t, b, s) + O((|\beta| M a_{bq}(t) + a_q(t)) tb).$$

Putting here $\alpha = \beta = a = 0$, we obtain that

$$a_q(M) M = A_q(M) = a_{bq}(t) M - \frac{1}{tb} \sum_{s=1}^{b} \sum_{j=1}^{M} D(j, t, b, s) + O(a_q(t) tb),$$

hence

$$\frac{1}{tb} \sum_{s=1}^{b} \sum_{j=1}^{M} D(j, t, b, s) < (a_{bq}(t) - a_q(M)) M + c_6 a_q(t) tb.$$

Thus (82) yields that

$$|F_{M,q}(\alpha) - F^*_{M,q}(\alpha)| <$$

$$< \frac{1}{tb} \sum_{s=1}^{b} \sum_{j=1}^{M} D(j, t, b, s) + c_7(|\beta| M a_{bq}(t) + a_q(t)) tb <$$

$$< ((a_{bq}(t) - a_q(M)) M + c_6 a_q(t) tb) + c_7(|\beta| M a_{bq}(t) + a_q(t)) tb <$$

$$< (a_{bq}(t) - a_q(M)) M + c_8(|\beta| M a_{bq}(t) + a_q(t)) tb$$

which proves Lemma 14.

4. (12) will be deduced from a lower estimate for

$$a^*(t) = \max_{1 \leqq b \leqq R} a_{bq}(t)$$

in terms of $a_q(M)$ where $t = o(M)$ and $R \to +\infty$, however, t is relatively large, R is small in terms of M.

LEMMA 15. *Let t, M, q be positive integers, R a real number such that*

$$t|M, \tag{83}$$

$$q \leqq \log M, \tag{84}$$

$$3 \leqq R \leqq \log M. \tag{85}$$

Then there exist absolute constants c_9, c_{10} such that for sufficiently large M,

$$(a_q(M))^2 \leqq c_9\Big\{(a^*(t)-a_q(M))^2 R\log R + \tag{86}$$

$$+a^*(t)(a^*(t)-a_q(M))+(a^*(t))^2\left(\frac{t}{M}\log R+\frac{t^2}{M^2}\frac{R^5}{(\log\log R)^2}\right)+$$

$$+a^*(t)\left(e^{-c_{10}\sqrt{\log M}}+\frac{\log\log R}{R}\right)\Big\}.$$

PROOF. We are going to use a modification of that version of the Hardy—Littlewood method which has been elaborated by K. F. ROTH in [4] and [5].

$P(\alpha)$, $F(\alpha)$ and $F^*(\alpha)$ will denote the functions defined by (14), (71) and (77). (We recall that $u_1, u_2, \ldots, u_{A_q(M)}$ in (71) denote integers such that $u_1 q, u_2 q, \ldots, u_{A_q(M)} q$ form a maximal $\mathscr{A}$-set selected from $q, 2q, \ldots, Mq$.) Then

$$\int_0^1 F(\alpha)F(-\alpha)P(\alpha)\,d\alpha = \tag{87}$$

$$= \int_0^1 \sum_{y=1}^{A_q(M)} e(u_y\alpha) \sum_{x=1}^{A_q(M)} e(-u_x\alpha) \sum_{\substack{\frac{p-1}{q}\leqq M \\ q|p-1}} (\log p)\, e\left(\frac{p-1}{q}\alpha\right) d\alpha =$$

$$= \sum_{\substack{x,y,p \\ u_y-u_x+\frac{p-1}{q}=0}} \log p = 0,$$

namely,

$$u_y - u_x + \frac{p-1}{q} = 0$$

or in equivalent form,

$$u_x q - u_y q = p-1$$

is not solvable, since the numbers $u_1 q, u_2 q, \ldots, u_{A_q(M)} q$ form an $\mathscr{A}$-set.

Let us write

$$\delta = \frac{1}{M}\frac{R}{\log\log R}; \tag{88}$$

then by (85),

$$\frac{1}{M} < \delta \leqq \frac{\log M}{M\log\log\log M}\left(< \frac{1}{4}\right) \tag{89}$$

for large M.

By (87),

$$\int_{-\delta}^{+\delta} |F(\alpha)|^2 P(\alpha)\, d\alpha = -\int_{+\delta}^{1-\delta} |F(\alpha)|^2 P(\alpha)\, d\alpha,$$

hence,

$$\left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 P(\alpha)\, d\alpha\right| = \left|\int_{+\delta}^{1-\delta} |F(\alpha)|^2 P(\alpha)\, d\alpha\right|,$$

$$\left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 P(\alpha)\, d\alpha\right| \leqq \int_{+\delta}^{1-\delta} |F(\alpha)|^2 |P(\alpha)|\, d\alpha. \tag{90}$$

We are going to give a lower estimate for the left hand side and an upper estimate for the right hand side.

In order to estimate $P(\alpha)$ for $|\alpha| \leqq \delta$, we apply Lemma 3 with $u=1$, $a=0$, $b=1$ ((16) holds by (84)). Then $m_{b,q}=0$ in Lemma 3, thus we obtain with respect to (89) that there exists an absolute constant c_{10} such that for large M and $|\alpha| \leqq \delta$,

$$\left|P(\alpha) - \frac{q}{\varphi(q)} \sum_{n=1}^{M} e(n\alpha)\right| = O\left((M\delta+1)\, M e^{-c_2\sqrt{\log M}}\right) =$$

$$= O\left(\frac{M \log M}{\log\log\log M}\, e^{-c_2\sqrt{\log M}}\right) < M e^{-c_{10}\sqrt{\log M}}.$$

Thus we obtain applying Parseval's formula that

$$\left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 P(\alpha)\, d\alpha\right| = \tag{91}$$

$$= \left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 \cdot \frac{q}{\varphi(q)} \left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha + \int_{-\delta}^{+\delta} |F(\alpha)|^2 \left(P(\alpha) - \frac{q}{\varphi(q)} \sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| \geqq$$

$$\geqq \left|\frac{q}{\varphi(q)} \int_{-\delta}^{+\delta} |F(\alpha)|^2 \left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| - \int_{-\delta}^{+\delta} |F(\alpha)|^2 \left|P(\alpha) - \frac{q}{\varphi(q)} \sum_{n=1}^{M} e(n\alpha)\right| d\alpha >$$

$$> \frac{q}{\varphi(q)} \left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 \left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| - \int_{-\delta}^{+\delta} |F(\alpha)|^2 M e^{-c_{10}\sqrt{\log M}}\, d\alpha >$$

$$> \frac{q}{\varphi(q)} \left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 \left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| - M e^{-c_{10}\sqrt{\log M}} \int_{0}^{1} |F(\alpha)|^2\, d\alpha =$$

$$= \frac{q}{\varphi(q)} \left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 \left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| - a_q(M)\, M^2 e^{-c_{10}\sqrt{\log M}} \geqq$$

$$\geqq \frac{q}{\varphi(q)} \left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 \left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| - a^*(t)\, M^2 e^{-c_{10}\sqrt{\log M}}$$

since

$$a_q(M) \leqq a_q(t) \leqq a^*(t) \tag{92}$$

by (74), (83) and the definition of the function $a^*(t)$.

For any complex numbers u, v, we have

$$\big||u|^2-|v|^2\big| = |u\bar{u}-v\bar{v}| = |(u-v)\bar{u}+v(\bar{u}-\bar{v})| \leqq$$

$$\leqq |u-v||\bar{u}|+|v||\bar{u}-\bar{v}| = |u-v|(|u|+|v|) =$$

$$= |u-v|\big(|(u-v)+v|+|v|\big) \leqq |u-v|(|u-v|+2|v|) =$$

$$= |u-v|^2+2|u-v||v|.$$

Thus

$$\left|\int_{-\delta}^{+\delta}\big(|F(\alpha)|^2-|F^*(\alpha)|^2\big)\Big(\sum_{n=1}^{M}e(n\alpha)\Big)d\alpha\right| \leqq \tag{93}$$

$$\leqq \int_{-\delta}^{+\delta}\big||F(\alpha)|^2-|F^*(\alpha)|^2\big|\left|\sum_{n=1}^{M}e(n\alpha)\right|d\alpha \leqq$$

$$\leqq \int_{-\delta}^{+\delta}\big(|F(\alpha)-F^*(\alpha)|^2+2|F(\alpha)-F^*(\alpha)||F^*(\alpha)|\big)\left|\sum_{n=1}^{M}e(n\alpha)\right|d\alpha.$$

For $a=0$, $b=1$, Lemma 14 yields with respect to (92) that

$$|F(\alpha)-F^*(\alpha)| \leqq \big(a_q(t)-a_q(M)\big)M+c_5\big(|\alpha|Ma_q(t)+a_q(t)\big)t \leqq$$

$$\leqq \big(a^*(t)-a_q(M)\big)M+c_5\big(|\alpha|M+1\big)a^*(t)t \leqq$$

$$\leqq \begin{cases}\big(a^*(t)-a_q(M)\big)M+c_{11}a^*(t)t & \text{for} \quad |\alpha| \leqq 1/M\\ \big(a^*(t)-a_q(M)\big)M+c_{12}|\alpha|a^*(t)tM & \text{for} \quad 1/M \leqq |\alpha| \leqq \delta.\end{cases}$$

Thus using also Lemma 4, we obtain from (93) (with respect to (88), (89), (92) and the inequality

$$(A+B)^2 \leqq 2A^2+2B^2 \tag{94}$$

where A, B are arbitrary real numbers) that

$$(95)\qquad \left|\int_{-\delta}^{+\delta}\left(|F(\alpha)|^2-|F^*(\alpha)|^2\right)\left(\sum_{n=1}^{M}e(n\alpha)\right)d\alpha\right| \ll$$

$$\ll \int_{|\alpha|\leqq 1/M}\left\{\left((a^*(t)-a_q(M))M+a^*(t)t\right)^2+\right.$$

$$\left.+\left((a^*(t)-a_q(M))M+a^*(t)t\right)a^*(t)M\right\}Md\alpha+$$

$$+\int_{1/M\leqq|\alpha|\leqq\delta}\left\{\left((a^*(t)-a_q(M))M+|\alpha|a^*(t)tM\right)^2+\right.$$

$$\left.+\left((a^*(t)-a_q(M))M+|\alpha|a^*(t)tM\right)a^*(t)\frac{1}{|\alpha|}\right\}\frac{1}{|\alpha|}d\alpha \ll$$

$$\ll (a^*(t)-a_q(M))^2M^2+(a^*(t))^2t^2+a^*(t)(a^*(t)-a_q(M))M^2+(a^*(t))^2tM+$$

$$+\left\{(a^*(t)-a_q(M))^2M^2+(a^*(t))^2tM\right\}\int_{1/M\leqq|\alpha|\leqq\delta}\frac{1}{|\alpha|}d\alpha+$$

$$+(a^*(t))^2t^2M^2\int_{1/M\leqq|\alpha|\leqq\delta}|\alpha|\,d\alpha+a^*(t)(a^*(t)-a_q(M))M\int_{1/M\leqq|\alpha|\leqq\delta}\frac{1}{|\alpha|^2}d\alpha \ll$$

$$\ll (a^*(t)-a_q(M))^2M^2+(a^*(t))^2tM+a^*(t)(a^*(t)-a_q(M))M^2+$$

$$+\left\{(a^*(t)-a_q(M))^2M^2+(a^*(t))^2tM\right\}\log M\delta+$$

$$+(a^*(t))^2t^2M^2\delta^2+a^*(t)(a^*(t)-a_q(M))M^2 \ll$$

$$\ll (a^*(t)-a_q(M))^2M^2+(a^*(t))^2tM+a^*(t)(a^*(t)-a_q(M))M^2+$$

$$+\left\{(a^*(t)-a_q(M))^2M^2+(a^*(t))^2tM\right\}\log R+(a^*(t))^2t^2\frac{R^2}{(\log\log R)^2} \ll$$

$$\ll (a^*(t)-a_q(M))^2M^2\log R+(a^*(t))^2\left(tM\log R+t^2\frac{R^2}{(\log\log R)^2}\right)+$$

$$+a^*(t)(a^*(t)-a_q(M))M^2.$$

By Lemma 4 and Parseval's formula, we have

$$\int_{-\delta}^{+\delta}|F^*(\alpha)|^2\left(\sum_{n=1}^{M}e(n\alpha)\right)d\alpha =$$

$$=\int_{-\delta}^{1-\delta}|F^*(\alpha)|^2\left(\sum_{n=1}^{M}e(n\alpha)\right)d\alpha-\int_{+\delta}^{1-\delta}|F^*(\alpha)|^2\left(\sum_{n=1}^{M}e(n\alpha)\right)d\alpha \geqq$$

$$\geqq (a_q(t))^2\sum_{\substack{1\leqq x,y,z\leqq M\\ x-y+z=0}}1-2\int_{+\delta}^{1/2}(a_q(t))^2\cdot\frac{1}{4\alpha^2}\cdot\frac{1}{2\alpha}d\alpha.$$

Here for large M,

$$\sum_{\substack{1\leqq x,y,z\leqq M\\ x-y+z=0}} 1 \geqq \left[\frac{M}{2}\right]^2 > \frac{M^2}{5}$$

since $1\leqq x\leqq\left[\frac{M}{2}\right]$, $1\leqq z\leqq\left[\frac{M}{2}\right]$ and $y=x+z$ satisfy the conditions $1\leqq x, y, z\leqq M$, $x-y+z=0$. Thus with respect to (85) and (92),

$$\int_{-\delta}^{+\delta} |F^*(\alpha)|^2\left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha > \tag{96}$$

$$> (a_q(t))^2\cdot\left(\frac{M^2}{5}-\frac{1}{4}\int_{+\delta}^{+\infty}\frac{1}{\alpha^3}\,d\alpha\right) = (a_q(t))^2\left(\frac{M^2}{5}-\frac{1}{8\delta^2}\right) =$$

$$= (a_q(t))^2 M^2\left(\frac{1}{5}-\frac{1}{8}\left(\frac{\log\log R}{R}\right)^2\right) > \frac{1}{10}(a_q(t))^2 M^2 \geqq \frac{1}{10}(a_q(M))^2 M^2.$$

(91), (95) and (96) yield that

$$\left|\int_{-\delta}^{+\delta} |F(\alpha)|^2 P(\alpha)\,d\alpha\right| > \tag{97}$$

$$> \frac{q}{\varphi(q)}\left|\int_{-\delta}^{+\delta} |F(\alpha)|^2\left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| - a^*(t) M^2 e^{-c_{10}\sqrt{\log M}} \geqq$$

$$\geqq \frac{q}{\varphi(q)}\left|\int_{-\delta}^{+\delta} |F^*(\alpha)|^2\left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| -$$

$$-\frac{q}{\varphi(q)}\left|\int_{-\delta}^{+\delta}\left(|F(\alpha)|^2-|F^*(\alpha)|^2\right)\left(\sum_{n=1}^{M} e(n\alpha)\right) d\alpha\right| -$$

$$-a^*(t) M^2 e^{-c_{10}\sqrt{\log M}} > \frac{1}{10}\cdot\frac{q}{\varphi(q)}(a_q(M))^2\cdot M^2 -$$

$$-c_{13}\frac{q}{\varphi(q)}\left\{(a^*(t)-a_q(M))^2 M^2\log R+(a^*(t))^2\left(tM\log R+t^2\frac{R^2}{(\log\log R)^2}\right)+\right.$$

$$\left.+a^*(t)(a^*(t)-a_q(M))M^2+a^*(t)M^2 e^{-c_{10}\sqrt{\log M}}\right\}.$$

Now we are going to give an upper estimate for the right hand side of (90).

If a, b are integers such that $0\leqq a\leqq b-1$, $1\leqq b\leqq R$ and $(a, b)=1$ then let us denote the interval

$$\left[\frac{a}{b}-\delta, \frac{a}{b}+\delta\right] = \left[\frac{a}{b}-\frac{1}{M}\frac{R}{\log\log R}, \frac{a}{b}+\frac{1}{M}\frac{R}{\log\log R}\right]$$

by $I_{a,b}$ (so that $I_{0,1}=[-\delta\ +\delta]$) and define the set $S_{R,M}$ in the same way as in Lemma 9. Then obviously,

$$[\delta, 1-\delta] \subset \left\{\bigcup_{2\leqq b\leqq R}\ \bigcup_{\substack{1\leqq a\leqq b-1\\ (a,b)=1}} I_{a,b}\right\}\cup S_{R,M}$$

thus

$$\int_{+\delta}^{1-\delta} |F(\alpha)|^2 |P(\alpha)|\, d\alpha \leqq \tag{98}$$

$$\leqq \sum_{b=2}^{[R]} \sum_{\substack{1 \leqq a \leqq b-1 \\ (a,b)=1}} \int_{I_{a,b}} |F(\alpha)|^2 |P(\alpha)|\, d\alpha + \int_{S_{R,M}} |F(\alpha)|^2 |P(\alpha)|\, d\alpha =$$

$$= \sum_{b=2}^{[R]} \sum_{\substack{1 \leqq a \leqq b-1 \\ (a,b)=1}} E_{a,b} + E_S .$$

For $\alpha \in I_{a,b}$, we use Lemma 14 to estimate $|F(\alpha)|$, while $|P(\alpha)|$ can be estimated by applying Lemma 5 with $u=2$, $\alpha = \frac{a}{b}+\beta$, since (15) and (16) hold by (84), (85) and $b \leqq R$, and also (28) holds for large M by $|\beta| \leqq \delta$ and (89). Applying these lemmas, we obtain with respect to (92) that if $\alpha \in I_{a,b}$ (where $1 < b \leqq R$) then

$$|F(\alpha)| \leqq \left(a_{bq}(t) - a_q(M)\right) M + c_5 \left(|\beta| M a_{bq}(t) + a_q(t)\right) tb \leqq$$

$$\leqq \left(a^*(t) - a_q(M)\right) M + c_5 \left(|\beta| M + 1\right) a^*(t)\, tb \leqq$$

$$\leqq \begin{cases} \left(a^*(t) - a_q(M)\right) M + 2c_5 a^*(t)\, tb & \text{for} \quad |\beta| \leqq \dfrac{1}{M} \\ \left(a^*(t) - a_q(M)\right) M + 2c_5 |\beta| M a^*(t)\, tb & \text{for} \quad |\beta| > \dfrac{1}{M} \end{cases}$$

and (29) hold. Thus in view of (85), (88), (89) and (94),

$$E_{a,b} = \int_{|\beta| \leqq \frac{1}{M}} \left|F\left(\frac{a}{b}+\beta\right)\right|^2 \left|P\left(\frac{a}{b}+\beta\right)\right| d\beta +$$

$$+ \int_{\frac{1}{M} \leqq |\beta| \leqq \delta} \left|F\left(\frac{a}{b}+\beta\right)\right|^2 \left|P\left(\frac{a}{b}+\beta\right)\right| d\beta \ll$$

$$\ll \int_{|\beta| \leqq \frac{1}{M}} \left\{\left(a^*(t) - a_q(M)\right)^2 M^2 + \left(a^*(t)\right)^2 t^2 b^2\right\} \frac{Mq}{\varphi(b)\varphi(q)}\, d\beta +$$

$$+ \int_{\frac{1}{M} \leqq |\beta| \leqq \delta} \left\{\left(a^*(t) - a_q(M)\right)^2 M^2 + |\beta|^2 M^2 \left(a^*(t)\right)^2 t^2 b^2\right\} \frac{q}{\varphi(b)\varphi(q)|\beta|}\, d\beta \ll$$

$$\ll \left\{\left(a^*(t) - a_q(M)\right)^2 M^2 + \left(a^*(t)\right)^2 t^2 b^2\right\} \frac{q}{\varphi(b)\varphi(q)} +$$

$$+ \frac{q}{\varphi(b)\varphi(q)} \left\{\left(a^*(t) - a_q(M)\right)^2 M^2 \int_{\frac{1}{M} \leqq |\beta| \leqq \delta} \frac{1}{|\beta|}\, d\beta + M^2 \left(a^*(t)\right)^2 t^2 b^2 \int_{\frac{1}{M} \leqq |\beta| \leqq \delta} |\beta|\, d\beta\right\} \ll$$

$$\ll \frac{q}{\varphi(b)\varphi(q)} \left\{\left(a^*(t) - a_q(M)\right)^2 M^2 + \left(a^*(t)\right)^2 t^2 b^2 +\right.$$

$$+\left(a^*(t)-a_q(M)\right)^2 M^2 \log M\delta + M^2\left(a^*(t)\right)^2 t^2 b^2 \delta^2\Big\} \ll$$

$$\ll \frac{q}{\varphi(b)\varphi(q)}\left\{\left(a^*(t)-a_q(M)\right)^2 M^2 \log R + \left(a^*(t)\right)^2 t^2 b^2 \left(\frac{R}{\log\log R}\right)^2\right\},$$

hence

$$(99)\qquad \sum_{b=2}^{[R]} \sum_{\substack{1\leqq a\leqq b-1\\(a,b)=1}} E_{a,b} \ll \sum_{b=2}^{[R]} \sum_{\substack{1\leqq a\leqq b-1\\(a,b)=1}} \frac{q}{\varphi(b)\varphi(q)}\left\{\left(a^*(t)-a_q(M)\right)^2 M^2 \log R + \right.$$

$$\left. + \left(a^*(t)\right)^2 t^2 b^2 \left(\frac{R}{\log\log R}\right)^2\right\} \ll$$

$$\ll \frac{q}{\varphi(q)}\left\{\left(a^*(t)-a_q(M)\right)^2 M^2 R \log R + \left(a^*(t)\right)^2 t^2 \frac{R^5}{(\log\log R)^2}\right\}.$$

Finally, to estimate E_S, we use Lemma 9 and Parseval's formula:

$$(100)\qquad E_S = \int_{S_{R,M}} |F(\alpha)|^2 |P(\alpha)|\, d\alpha \leqq \sup_{\alpha\in S_{R,M}} |P(\alpha)| \int_{S_{R,M}} |F(\alpha)|^2\, d\alpha \ll$$

$$\ll \frac{qM}{\varphi(q)} \frac{\log\log R}{R} \int_0^1 |F(\alpha)|^2\, d\alpha =$$

$$= \frac{qM}{\varphi(q)} \frac{\log\log R}{R} a_q(M) M \leqq \frac{q}{\varphi(q)} a^*(t) M^2 \frac{\log\log R}{R}$$

(with respect to (92)).

(90), (97), (98), (99) and (100) yield that

$$\frac{1}{10}\frac{q}{\varphi(q)}\left(a_q(M)\right)^2 M^2 - c_{13}\frac{q}{\varphi(q)}\left\{\left(a^*(t)-a_q(M)\right)^2 M^2 \log R + \right.$$

$$+\left(a^*(t)\right)^2\left(tM\log R + t^2\frac{R^2}{(\log\log R)^2}\right) + a^*(t)\left(a^*(t)-a_q(M)\right)M^2 +$$

$$\left. + a^*(t) M^2 e^{-c_{10}\sqrt{\log M}}\right\} \ll$$

$$\ll \frac{q}{\varphi(q)}\left\{\left(a^*(t)-a_q(M)\right)^2 M^2 R\log R + \left(a^*(t)\right)^2 t^2 \frac{R^5}{(\log\log R)^2}\right\} +$$

$$+ \frac{q}{\varphi(q)} a^*(t) M^2 \frac{\log\log R}{R}$$

or in equivalent form,

$$\left(a_q(M)\right)^2 \ll \left(a^*(t)-a_q(M)\right)^2 R\log R +$$

$$+ a^*(t)\left(a^*(t)-a_q(M)\right) + \left(a^*(t)\right)^2\left(\frac{t}{M}\log R + \frac{t^2}{M^2}\frac{R^5}{(\log\log R)^2}\right) +$$

$$+ a^*(t)\left(e^{-c_{10}\sqrt{\log M}} + \frac{\log\log R}{R}\right)$$

(with respect to (92)) which completes the proof of Lemma 15.

5. In this section, we will complete the proof of our theorem by showing that Lemma 15 implies (12).

C will denote a large enough (but fixed) constant and x will be an arbitrary integer which is sufficiently large in terms of C.

Let us write

$$Z = \left[\frac{1}{6}\frac{\log\log x}{\log\log\log x}\right]$$

and define the positive integer N by

$$[(\log\log x)^5]^Z | N \tag{101}$$

and

$$N \leqq x < N + [(\log\log x)^5]^Z, \tag{102}$$

so that

$$N = \left[\frac{x}{[(\log\log x)^5]^Z}\right][(\log\log x)^5]^Z.$$

For $x \to +\infty$,

$$Z \sim \frac{1}{6}\frac{\log\log x}{\log\log\log x}, \tag{103}$$

hence

$$\log[(\log\log x)^5]^Z = Z\log[(\log\log x)^5] \sim$$

$$\sim 5Z\log\log\log x \sim 5\cdot\frac{1}{6}\frac{\log\log x}{\log\log\log x}\log\log\log x =$$

$$= \frac{5}{6}\log\log x$$

thus for large x,

$$[(\log\log x)^5]^Z < e^{\log\log x} = \log x. \tag{104}$$

(102) and (104) imply that for large x,

$$x \geqq N > x - \log x. \tag{105}$$

Let us define the positive integers $t_0, t_1, \ldots, t_{Z-1}, t_Z$ in the following way: for $k=0, 1, \ldots, Z$, let

$$t_k = \frac{N}{[(\log\log x)^5]^{Z-k}},$$

so that $t_Z = N$. (In fact, these numbers are positive integers by (101).) Furthermore, (104) and (105) imply that for large x,

$$x \geqq N = t_Z > t_{Z-1} > \ldots > t_1 > t_0 = \frac{N}{[(\log\log x)^5]^Z} > \tag{106}$$

$$> \frac{x - \log x}{\log x} = \frac{x}{\log x} - 1 \quad (> \sqrt{x}).$$

For $u \geqq 3$, let us define the function $f(u)$ by

$$f(u) = \frac{\log u \log\log u}{u^2}$$

and for $k=0, 1, \ldots, Z-1$, let

$$R_k = (f(C))^{1/2}(f(k+C))^{-1}\log\log(f(k+C))^{-1}.$$

Finally, we define the positive integers $q_0, q_1, \ldots, q_{Z-1}, q_Z$ by the following backward recursion:

Let $q_Z=1$. If $q_Z, q_{Z-1}, \ldots, q_{k+1}$ have been defined (where $0 \leqq k \leqq Z-1$) then let q_k denote a positive integer for which

$$q_{k+1} | q_k \tag{107}$$

and

$$1 \leqq \frac{q_k}{q_{k+1}} \leqq R_k \tag{108}$$

hold and $a_{q_k}(t_k)$ is maximal; i.e. using the notations of Lemma 15 (with t_k, q_{k+1} and R_k in place of t, q and R, respectively), let us define q_k by (107), (108) and

$$a_{q_k}(t_k) = a^*(t) = \max_{1 \leqq b \leqq R_k} a_{bq_{k+1}}(t_k). \tag{109}$$

We are going to show by straight induction that if C is large enough and x is sufficiently large in terms of C then for $k=0, 1, \ldots, Z$,

$$a_{q_k}(t_k) \leqq \frac{f(k+C)}{f(C)}. \tag{110}$$

For $k=0$, (110) can be written in the form $a_{q_0}(t_0) \leqq 1$ but this holds trivially by Lemma 12 (independently of C).

Now let us suppose that (110) holds for some positive integer k, satisfying $0 \leqq k \leqq Z-1$. We have to show that this implies that also

$$a_{q_{k+1}}(t_{k+1}) \leqq \frac{f(k+1+C)}{f(C)}$$

holds.

Let us assume indirectly that

$$a_{q_{k+1}}(t_{k+1}) > \frac{f(k+1+C)}{f(C)}. \tag{111}$$

We are going to deduce a contradiction from this indirect assumption by using Lemma 15. For this purpose, we need some estimates for the function $f(u)$ and the parameters Z and R_k.

Obviously, for large u, the function $f(u)$ is decreasing and

$$\lim_{u \to +\infty} f(u) = 0. \tag{112}$$

Furthermore, if $u \to +\infty$ and $\frac{u}{v} \to 1$ then

$$f(u) \sim f(v) \quad \left(\text{for} \quad u \to +\infty, \ \frac{u}{v} \to 1\right). \tag{113}$$

For $u \to +\infty$,

$$\log\big(f(u)\big)^{-1} \sim \log u^2 = 2 \log u \quad (\text{for} \quad u \to +\infty) \tag{114}$$

and

$$\log\log\big(f(u)\big)^{-1} \sim \log\log u \quad (\text{for} \quad u \to +\infty). \tag{115}$$

By Lagrange's mean value theorem, for $u \geqq 3$, there exists a real number v such that

$$f(u)-f(u+1) = -f'(v) \quad \text{and} \quad u \leqq v \leqq u+1.$$

Thus for $u \to +\infty$, we obtain with respect to (113) that

$$f(u)-f(u+1) = -f'(v) = \frac{-\log\log v - 1 + 2(\log v)(\log\log v)}{v^3} \sim \tag{116}$$

$$\sim 2\frac{(\log v)(\log\log v)}{v^3} = 2\frac{f(v)}{v} \sim 2\frac{f(u)}{u} \quad (\text{for} \quad u \to +\infty).$$

(103) implies that

$$\log Z \sim \log\log\log x \tag{117}$$

and

$$\log\log Z \sim \log\log\log\log x$$

(for $x \to +\infty$). Thus with respect to (103) and (113), we have

$$f(Z+C) \sim f(Z) = \frac{\log Z \log\log Z}{Z^2} \sim \tag{118}$$

$$\sim 36\frac{(\log\log\log x)^3(\log\log\log\log x)}{(\log\log x)^2}.$$

Finally, if C is large enough and $k = 0, 1, \ldots, Z-1$ then with respect to (115),

$$R_k = \big(f(C)\big)^{1/2}\big(f(k+C)\big)^{-1}\log\log\big(f(k+C)\big)^{-1} < \tag{119}$$

$$< \big(f(C)\big)^{1/2}\frac{(k+C)^2}{\log(k+C)\log\log(k+C)} \cdot 2\log\log(k+C) =$$

$$= 2\big(f(C)\big)^{1/2}\frac{(k+C)^2}{\log(k+C)}$$

and

$$R_k > \big(f(C)\big)^{1/2}\big(f(k+C)\big)^{-1} \cdot \frac{1}{2}\log\log(k+C) = \tag{120}$$

$$= \frac{1}{2}\big(f(C)\big)^{1/2}\frac{(k+C)^2}{\log(k+C)}.$$

Furthermore, by (112) and since $f(u)$ is decreasing for large u, we have also

$$R_k < \big(f(k+C)\big)^{-1} \log\log\big(f(k+C)\big)^{-1}$$

and

$$R_k \geqq \big(f(k+C)\big)^{1/2}\big(f(k+C)\big)^{-1} \log\log\big(f(k+C)\big)^{-1} =$$

$$= \big(f(k+C)\big)^{-1/2} \log\log\big(f(k+C)\big)^{-1}$$

for large enough C. Hence, in view of (112), (114) and (115), we obtain for large C and $k=0, 1, \ldots, Z-1$ that

$$\frac{1}{2}\log(k+C) < \log R_k < 3\log(k+C) \tag{121}$$

and

$$\frac{1}{2}\log\log(k+C) < \log\log R_k < 2\log\log(k+C). \tag{122}$$

We are ready to show that if C is large enough and x is sufficiently large (in terms of C) then Lemma 15 can be applied with t_k, t_{k+1}, q_{k+1} and R_k in place of t, M, q and R. In fact, (83) holds obviously by the definition of the numbers $t_0, t_1, \ldots, t_Z$. Also, $R \geqq 3$ holds trivially for large C by (121). Furthermore,

$$q_{k+1} = q_Z \prod_{j=k+1}^{Z-1} \frac{q_j}{q_{j+1}} = \prod_{j=k+1}^{Z-1} \frac{q_j}{q_{j+1}} \leqq \prod_{j=0}^{Z-1} R_j,$$

thus to prove that both (84) and (85) hold, it suffices to show that

$$\prod_{j=0}^{Z-1} R_j \leqq \log t_{k+1} (= \log M)$$

or in equivalent form,

$$\sum_{j=0}^{Z-1} \log R_j \leqq \log\log t_{k+1}. \tag{123}$$

By (106),

$$\log\log t_{k+1} > \log\log \sqrt{x} > \frac{5}{6}\log\log x \tag{124}$$

for large x. On the other hand, by (103), (117) and (121), we have

$$\sum_{j=0}^{Z-1} \log R_j < 3 \sum_{j=0}^{Z-1} \log(j+C) < 3Z\log(Z+C) <$$

$$< 4Z\log Z < 5 \cdot \frac{1}{6} \frac{\log\log x}{\log\log\log x} \log\log\log x = \frac{5}{6}\log\log x \tag{125}$$

for large C and x. (124) and (125) yield (123). Thus in fact, Lemma 15 can be applied; we obtain that (86) holds. To deduce a contradiction from (86), we have to estimate $a_q(M)$ and $a^*(t) - a_q(M)$.

Using the notations of Lemma 15, (110) and (111) can be rewritten in the form

$$a^*(t) \leqq \frac{f(k+C)}{f(C)} \tag{126}$$

and

$$a_q(M) > \frac{f(k+1+C)}{f(C)}. \tag{127}$$

By (74) in Lemma 13, $t=t_k/t_{k+1}=M$ implies that

$$0 \leqq a_{q_{k+1}}(t_k)-a_{q_{k+1}}(t_{k+1}) = a_q(t)-a_q(M) \leqq a^*(t)-a_q(M). \tag{128}$$

With respect to (113), (126), (127) and (128) imply for large C that

$$a^*(t) \geqq a_q(M) > \frac{1}{2}a^*(t). \tag{129}$$

Furthermore, (126) and (127) yield with respect to (113), (116) and (129) that for large C,

$$a^*(t)-a_q(M) < \frac{f(k+C)}{f(C)} - \frac{f(k+1+C)}{f(C)} < \frac{3}{f(C)}\frac{f(k+C)}{k+C} < \tag{130}$$

$$< \frac{4}{k+C}\frac{f(k+1+C)}{f(C)} < \frac{4}{k+C}a_q(M) \leqq \frac{4}{k+C}a^*(t).$$

By (118), (127) and (129), we have

$$a^*(t) \geqq a_q(M) > \frac{f(k+1+C)}{f(C)} \geqq \frac{f(Z+C)}{f(C)} > \tag{131}$$

$$> \frac{35}{f(C)}\frac{(\log\log\log x)^3(\log\log\log\log x)}{(\log\log x)^2}$$

for large x, while in view of (106),

$$e^{-c_{10}\sqrt{\log M}} = e^{-c_{10}\sqrt{\log t_{k+1}}} \leqq e^{-c_{10}\sqrt{\log t_0}} < \tag{132}$$

$$< e^{-c_{10}\sqrt{\log\sqrt{x}}} = e^{-c_{14}\sqrt{\log x}} = o\left(\frac{(\log\log\log x)^3(\log\log\log\log x)}{(\log\log x)^2}\right)$$

for $x\to+\infty$. (131) and (132) yield that for fixed C and large x,

$$e^{-c_{10}\sqrt{\log M}} < f(C)a^*(t). \tag{133}$$

Finally, by (113), (120), (122), (127) and (129), we have

$$\frac{\log\log R}{R} < \frac{2\log\log(k+C)}{\frac{1}{2}(f(C))^{1/2}\frac{(k+C)^2}{\log(k+C)}} = \tag{134}$$

$$= 4(f(C))^{-1/2}f(k+C) = 4(f(C))^{1/2}\frac{f(k+C)}{f(C)} <$$

$$< 5(f(C))^{1/2}\frac{f(k+1+C)}{f(C)} < 5(f(C))^{1/2}a_q(M) \leqq 5(f(C))^{1/2}a^*(t)$$

for large C.

With respect to (119), (121), (122), (128), (129), (130), (133) and (134), (86) yields that

$$\left(\frac{1}{2}a^*(t)\right)^2 < c_9\left\{\left(\frac{4}{k+C}a^*(t)\right)^2\cdot 2(f(C))^{1/2}\frac{(k+C)^2}{\log(k+C)}\cdot 3\log(k+C)+\right.$$

$$+a^*(t)\cdot\frac{4}{k+C}a^*(t)+(a^*(t))^2\left(\frac{1}{[(\log\log x)^5]}\cdot 3\log(k+C)+\right.$$

$$\left.\left.+\frac{1}{[(\log\log x)^5]^2}\cdot\left(2(f(C))^{1/2}\frac{(k+C)^2}{\log(k+C)}\right)^5\cdot\frac{1}{\left(\frac{1}{2}\log\log(k+C)\right)^2}\right)+\right.$$

$$\left.+a^*(t)(f(C)a^*(t)+5(f(C))^{1/2}a^*(t))\right\}.$$

Dividing by $(a^*(t))^2$ and with respect to (103), (112) and (117), we obtain that if C is large enough and x is sufficiently large depending on C then

$$\frac{1}{4} < 96c_9(f(C))^{1/2}+\frac{4c_9}{k+C}+c_9\frac{2}{(\log\log x)^5}\cdot 3\log(Z+C)+$$

$$+c_9\frac{2}{(\log\log x)^{10}}\cdot 2^5(f(C))^{5/2}(Z+C)^{10}+c_9f(C)+5c_9(f(C))^{1/2} <$$

$$< \frac{1}{30}+\frac{1}{30}+\frac{7c_9}{(\log\log x)^5}\log Z+\frac{2^{16}c_9(f(C))^{5/2}}{(\log\log x)^{10}}\cdot Z^{10}+\frac{1}{30}+\frac{1}{30} <$$

$$< \frac{2}{15}+\frac{8c_9}{(\log\log x)^5}(\log\log\log x)+\frac{2^{16}c_9(f(C))^{5/2}}{(\log\log x)^{10}}\left(\frac{1}{5}\frac{\log\log x}{\log\log\log x}\right)^{10} <$$

$$< \frac{2}{15}+\frac{1}{30}+\frac{2^{16}c_9(f(C))^{5/2}}{5^{10}}\cdot\frac{1}{(\log\log\log x)^{10}} < \frac{2}{15}+\frac{1}{30}+\frac{1}{30} = \frac{1}{5}.$$

Thus in fact, the indirect assumption (111) leads to a contradiction which proves that (110) holds for $k=0, 1, \ldots, Z$.

Applying (110) with $k=Z$, we obtain with respect to (118) that

$$a_{qz}(t_Z) = a_1(N) = a(N) \leqq \frac{f(Z+C)}{f(C)} < \tag{135}$$

$$< \frac{37}{f(C)} \frac{(\log\log\log x)^3 (\log\log\log\log x)}{(\log\log x)^2},$$

provided that x is sufficiently large.

Finally, (135) yields by (75) in Lemma 13 and (105) that

$$a(x) \leqq \left(1+\frac{N}{x}\right) a(N) \leqq 2a(N) < \frac{74}{f(C)} \frac{(\log\log\log x)^3 (\log\log\log\log x)}{(\log\log x)^2}$$

which completes the proof of our theorem.

6. In [6]—[9], K. F. ROTH generalized the method developed in [4] and [5], in order to investigate the solvability of systems of equations of the form

$$\sum_{j=1}^{\nu} \alpha_{ij} u_{x_j} = 0 \quad (i = 1, 2, \ldots, \mu)$$

where the numbers α_{ij} are integers satisfying $\sum_{j=1}^{\nu} \alpha_{ij} = 0$, and $u_1 < u_2 < \ldots$ is an arbitrary "dense" set of positive integers.

By using that extension of Roth's method which has been elaborated in this paper, one may investigate also the solvability of systems of equations of the more general form

$$\sum_{j=1}^{\nu} \alpha_{ij} u_{x_j} = \sum_{k=1}^{\varkappa} \beta_{ik} b_{y_k}^{(k)} \quad (i = 1, 2, \ldots, \mu)$$

where the numbers α_{ij} and β_{ik} are integers $\left(\text{again, } \sum_{j=1}^{\nu} \alpha_{ij} = 0\right)$, $u_1 < u_2 < \ldots$ is an arbitrary "dense" set of positive integers and the sets $b_1^{(k)} < b_2^{(k)} < \ldots$ (where $k = 1, \ldots, \varkappa$) are fixed sets of positive integers.

References

[1] H. L. MONTGOMERY, *Topics in Multiplicative Number Theory*, Springer Verlag, 1971.

[2] H. L. MONTGOMERY and R. C. VAUGHAN, The exceptional set in Goldbach's problem, *Acta Arithmetica*, **27** (1975), 353—370.

[3] K. PRACHAR, *Primzahlverteilung*, Springer Verlag, 1957.

[4] K. F. ROTH, Sur quelques ensembles d'entiers, *Comptes Rendus*, **234** (1952), 388—390.

[5] K. F. ROTH, On certain sets of integers, *J. London Math. Soc.*, **28** (1953), 104—109.

[6], [7], [8], [9] K. F. ROTH, Irregularities of sequences relative to arithmetic progressions, I—IV, *Math. Ann.*, **169** (1967), 1—25, *Math. Ann.*, **174** (1967), 41—52, *J. Number Theory*, **2** (1970), 125—142 and *Periodica Math. Hung.*, **2** (1972), 301—326.

[10], [11] A. SÁRKÖZY, On difference sets of sequences of integers, I—II, *Acta Math. Acad. Sci. Hungar.*, **31** (1978), 125—149 and *Annales Univ. Sci. Budapest, Sectio Math.*, to appear.

[12] I. M. VINOGRADOV, *The Method of Trigonometrical Sums in the Theory of Numbers*, Interscience (New York, 1954).

(Received March 31, 1977)

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15.