whether the last integral (in which \( g \geq 0 \)) is convergent or divergent (\( = \infty \)). On the other hand, the definition of \( g(t) \) shows that, in the present case, (5) reduces to

\[
G(x) = b_0 \left( 1 - \cos \frac{1}{2}x \right)/x + \sum_{n=1}^{\infty} b_n \left[ \cos \left( n - \frac{1}{2} \right)x - \cos \left( n + \frac{1}{2} \right)x \right]/x.
\]

Since the bracket on the right is identical with \( 2 \sin \frac{1}{2}x \sin nx \), it is seen from (2) that

\[
G(x) = b_0 \left( 1 - \cos \frac{1}{2}x \right)/x + f(x) \left( \sin \frac{1}{2}x \right)/(\frac{1}{2}x).
\]

It follows therefore from (6) that, since \( 1 - \cos \frac{1}{2}x \sim \frac{1}{8}x^2 \),

\[
\frac{f(x)}{x} \to \int_{0}^{\infty} tg(t) \, dt - \frac{1}{8}b_0,
\]
as \( x \to 0 \). This proves (4), since, by the definition of \( g(t) \),

\[
\int_{0}^{\infty} tg(t) \, dt = b_0 \int_{0}^{\frac{1}{2}} t \, dt + \sum_{n=1}^{\infty} b_n \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} t \, dt = \frac{3}{8}b_0 + \sum_{n=1}^{\infty} nb_n.
\]

References.


The Johns Hopkins University,
U.S.A.

**ON CERTAIN SETS OF INTEGERS**

**K. F. Roth**

1. A set of positive integers \( u_1, u_2, \ldots \) will be called an \( \mathcal{A} \)-set if no three of the numbers are in arithmetic progression, so that \( u_h + u_k = 2u_l \) only if \( h = k = l \). Let \( A(x) \) denote the greatest number of integers that can be selected from \( 1, 2, \ldots, x \) to form an \( \mathcal{A} \)-set. We write \( a(x) = x^{-1} A(x) \). In a recent note† I proved that \( a(x) \to 0 \) as \( x \to \infty \), a result which had been conjectured for many years‡. The purpose of the present paper, which

* Received 9 April, 1952; read 17 April, 1952.
‡ For the literature of the problem, see a note by R. Salem and D. C. Spencer, *Nieuw Archief voor Wiskunde*, 23 (1950), 133–143.
will be self-contained, is partly to give a more detailed account of the
method, and partly to prove a more precise result, namely
\[ a(x) = O\left(\frac{1}{\log\log x}\right). \]  

I would like to thank Prof. Davenport for suggesting simplifications
of the method, which are incorporated both in the Comptes Rendus note
and in the present paper.

2. The following obvious remarks will be important later.

\( A(x) \) is also the greatest number of integers that can be selected from
\( x \) consecutive terms of an arithmetic progression to form an \( \mathcal{A} \)-set. For
if \( au_1 + b, au_2 + b, \ldots \) form an \( \mathcal{A} \)-set then so do \( u_1, u_2, \ldots \), and conversely.

Further, for any two positive integers \( x \) and \( y \) we have
\[ A(x+y) \leq A(x) + A(y). \]

Thus
\[ A(xy) \leq xA(y), \quad A(x) \leq A\left\{ \left(\left[ \frac{x}{y} \right] + 1 \right)y \right\} \leq \frac{x+y}{y} A(y); \]
so that
\[ a(xy) \leq a(y), \]  
\[ a(x) \leq (1+xy^{-1})a(y). \]

Finally we have the trivial inequality
\[ x^{-1} \leq a(x) \leq 1. \]

Throughout the paper, small Latin letters (other than \( c, e, h \)) denote
positive integers. \( h \) denotes any integer. \( c_1, c_2, \ldots \) are positive absolute
constants. The constants implied by the \( O \) notation are absolute.

3. Let \( u_1, u_2, \ldots, u_M \) be any \( \mathcal{A} \)-set selected from \( 1, 2, \ldots, M \). In this
section we investigate the exponential sum
\[ S = \sum_{k=1}^{\gamma} e(ux_k) \quad [e(\theta) = e^{2\pi i\theta}], \]
where \( x \) is a real number. For each \( x \) there exist \( h, q \) such that
\[ x = \frac{h}{q} + \beta, \quad (h, q) = 1, \quad q \leq M^4, \quad q|\beta| \leq M^{-4}. \]

Suppose \( m < M \), and put
\[ S' = a(m)q^{-1}\left( \sum_{r=1}^{q} e\left(\frac{rh}{q}\right) \right)\left( \sum_{n=1}^{M} e(\beta n) \right), \]
We prove that
\[ |S - S'| < Ma(m) - U + O(mM^4). \] 
(6)

To prove this we start from the relation
\[ S = \frac{1}{mq} \sum_{r=1}^{g} \sum_{n=1}^{M} \sum_{n \leq u_k < n + mq}^{u_k \equiv r \pmod{q}} e(\alpha u_k) + O(mq). \] 
(7)

This relation is obvious on noting that for given \( u_k, m, q \) there are exactly \( mq \) integers \( n \), satisfying
\[ n \leq u_k < n + mq, \]
and that these integers \( n \) also satisfy \( 1 \leq n \leq M \) provided that \( mq \leq u_k < M - mq \).

Thus the coefficient of \( e(\alpha u_k) \) on the right-hand side of (7) is unity, except when \( u_k < mq \) or \( M - mq \leq u_k \leq M \); these terms being compensated by the error term.

In the inner sum on the right-hand side of (7), we have
\[ e(\alpha u_k) = e\left(\frac{\tau h}{\alpha}\right) e(\beta n) + O(mq) + O(Mmq|\beta|). \]

The number of terms in this inner sum is at most \( A(m) \), by a remark of the previous section, and is therefore \( A(m) - D(n, m, q, r) \), where \( D \geq 0 \).

Hence
\[ S = S' - \frac{1}{mq} \sum_{r=1}^{g} e\left(\frac{\tau h}{\alpha}\right) \sum_{n=1}^{M} e(\beta n) D(n, m, q, r) + O(mq) + O(Mmq|\beta|). \]

If we put \( \beta = 0 \) and \( h = 0 \) (legitimate since we have not yet used \( (h, q) = 1 \)), we obtain
\[ U = Ma(m) - \frac{1}{mq} \sum_{r=1}^{g} \sum_{n=1}^{M} D(n, m, q, r) + O(mq). \]

Using this as an estimate for
\[ \sum_{r=1}^{g} \sum_{n=1}^{M} D(n, m, q, r) \]
in the preceding relation, we obtain (6), on noting that \( q \leq M^4, q|\beta| \leq M^{-i} \).

4. In this section we use an adaptation of the Hardy-Littlewood method to obtain a functional inequality for the function \( a(x) \).

Let \( m \) be an even integer, \( 2N = m^4 \), and let now \( u_1, u_2, \ldots, u_{2r} \) be a maximal \( \mathcal{A} \)-set selected from \( 1, 2, \ldots, 2N \), so that \( U = A(2N) \). Let \( 2v_1, 2v_2, \ldots, 2v_{2r} \) be the even integers among the \( u_k \). We note that
\[ U = 2Na(2N) \] 
(8)
and, by (2),
\[ U \leq 2Na(m), \quad V \leq A(N) \leq Na(m). \] (9)

Further, since the number of odd integers among the \( u_k \) does not exceed \( A(N) \), we have, by (2),
\[ V \geq A(2N) - A(N) \geq 2Na(2N) - Na(m). \] (10)

We define
\[ f_1(\alpha) = \sum_{k=1}^{U} e(\alpha u_k), \quad f_2(\alpha) = \sum_{k=1}^{V} e(\alpha v_k); \]
\[ F_1(\alpha) = a(m) \sum_{n=1}^{2N} e(\alpha n), \quad F_2(\alpha) = a(m) \sum_{n=1}^{N} e(\alpha n). \]

In view of (9), we have
\[ f_r(\alpha) = O\left( Na(m) \right), \quad F_r(\alpha) = O\left( Na(m) \right); \quad r = 1, 2. \] (11)

We now show that, for any \( \alpha \),
\[ f_r(\alpha) - F_r(\alpha) = O\left( N \{ a(m) - a(2N) \} + N^2 \right); \quad r = 1, 2. \] (12)

If \( g = 1 \) in (5), this follows at once from (6) (with \( M = 2N \) or \( M = N \) according as \( r = 1 \) or 2) in view of (8) and (10). On the other hand, if it is impossible to choose \( q = 1 \) in (5) [so that \( S' = 0 \) in (6)], (6) (with \( M = 2N \) or \( N \)) will imply (12) provided that
\[ F_r(\alpha) = O(N^2). \]

This inequality is in fact satisfied, since for any \( \alpha \),
\[ \sum_{n=1}^{M} e(\alpha n) = O\left( \| \alpha \|^{-1} \right), \] (13)

where \( \| \alpha \| \) denotes the distance of \( \alpha \) from the nearest integer, and \( \| \alpha \| > M^{-1} \) if it is impossible to choose \( q = 1 \) in (5).

Further, using the inequality
\[ |f_1 f_2 - F_1^2| = |f_1(\alpha f_2^2 - F_2^2) + F_2^2(f_1 - F_1)| \leq |f_1 f_2 - F_2^2| + |F_2^2(f_1 - F_1)|, \]
we obtain, by (11) and (12),
\[ f_1(\alpha) f_2^2(-\alpha) - F_1(\alpha) F_2^2(-\alpha) = O\left( \{ Na(m) \}^2 \{ N \{ a(m) - a(2N) \} + N^2 \} \right). \] (14)

Finally, by (12) and (13), if \( 0 < \eta < a < 1 - \eta \) we have
\[ f_1(\alpha) = O\left( a(m) \eta^{-1} + N \{ a(m) - a(2N) \} + N^2 \right). \] (15)
The fact that \( u_1, u_2, \ldots, u_\sigma \) form an \( \mathcal{A} \)-set implies that \( u_k = v_k + \nu_l \) if and only if \( k = l \) and \( u_h = 2v_k \). This can be expressed, following the method of Hardy and Littlewood, by

\[
\int_{-\eta}^{1-\eta} f_1(x) f_2^2(-x) \, dx = V \leq Na(m). \tag{16}
\]

From now on we shall suppose that \( \eta = \eta(m) \) satisfies

\[
0 < \eta < \tfrac{1}{\delta}. \tag{17}
\]

Since

\[
\int_{0}^{1} |f_2(x)|^2 \, dx = V \leq Na(m),
\]

we have, by (15),

\[
\int_{-\eta}^{1-\eta} f_1(x) f_2^2(-x) \, dx = O\left( \{a(m)\eta^{-1} + N(a(m) - a(2N)) + N^\frac{1}{4}\} Na(m) \right). \tag{18}
\]

Further, by (14), we have

\[
\int_{-\eta}^{\eta} f_1(x) f_2^2(-x) \, dx = \int_{-\eta}^{\eta} f_1(x) F_2^2(-x) \, dx + O\left( \{Na(m))^2 \left( N(a(m) - a(2N)) + N^\frac{1}{4}\right) \right). \tag{19}
\]

Finally, by (13) and (17), we have

\[
\int_{-\eta}^{\eta} F_1(x) F_2^2(-x) \, dx = \int_{-\eta}^{\eta} F_1(x) F_2^2(-x) \, dx + O\left( \{N^3(a(m))^2 \right). \tag{20}
\]

Now the integral on the right represents \( a^3(m) \) times the number of solutions of \( n = n' + n'' \) in integers \( n, n', n'' \) satisfying \( n \leq 2N, n' \leq N, n'' \leq N \). This number is \( N^3 \). Thus, collecting together the results (16), (18), (19), (20), we obtain

\[
a^2(m) = \left\{ N^2 a(m) \right\}^{-1} \int_{-\eta}^{\eta} F_1(x) F_2^2(-x) \, dx
\]

\[
= O\left( a^2(m) N^{-2} \eta^{-2} + \{N a(m)+1\} a(m) a(2N) + N^{-1} \right). \tag{21}
\]

Hence writing \( \delta = (N\eta)^{-1} \),

\[
\delta = (N\eta)^{-1},
\]

we obtain, noting that \( 2N = m^4 \),

\[
a^2(m) < c_1 \left[ a(m) \delta + a^2(m) \delta^2 + \left( \delta^{-1} a(m) + 1 \right) a(m) - a(m) + m^{-1} \right]. \tag{22}
\]

Here \( \delta = \delta(m) \) is subject only to the restriction implied by (17).

We now write

\[
m = 2^e, \quad b(x) = a(m),
\]
where $x$ is any positive integer, so that (22) becomes

\[ b^2(x) < c_1 \left\{ b(x) \delta + b^2(x) \delta^2 + \left( \delta^{-1} b(x) + 1 \right) \left( b(x) - b(x+1) + 2^{-4^x} \right) \right\}. \quad (23) \]

5. In this section we shall deduce (1) from (2), (3), (4) and (23).

We assume $c_1 > 1$ ($c_1$ can be so chosen), and write $\delta = (2c_1)^{-1} b(x)$.

Then, noting that $b(x) \leq 1$ by (4), we have

\[ c_1 \left\{ b(x) \delta + b^2(x) \delta^2 \right\} \leq b^2(x) \left( \frac{1}{\delta} + \frac{1}{4c_1} \right) < \frac{3}{4} b^2(x). \]

Further, by (21) and (4)

\[ \eta = N^{-1} \delta^{-1} = c_2 m^{-4} \{a(m)\}^{-1} < c_2 m^{-3}, \]

so that (17) is satisfied for large $x$. Thus (23) implies

\[ b^2(x) < c_3 \left( b(x) - b(x+1) + 2^{-4^x} \right) \quad \text{for } x > c_4. \quad (24) \]

Now $b(x)$ is a decreasing function, by (2), and hence

\[ P b^2(2P) \leq \sum_{x=P}^{2P-1} b^2(x) < c_6 \left( b(P) - b(2P) + \frac{4c_5}{2P} \right) \quad (25) \]

for all integers $P > c_4$.

Hence, if $P > c_4$ and $2P b(2P) > 4c_5$, we have

\[ 2P b(2P) < \frac{1}{4c_5} \left( 2P b(2P) \right)^2 < P \left( b(P) - b(2P) + \frac{4c_5}{2P} \right) < P b(P). \]

This clearly implies (by a backward induction) that if $c_4 < 2^a < 2^t$, then

\[ 2^t b(2^t) \leq \max \left( 4c_5, 2^a b(2^a) \right), \]

so that

\[ b(2^t) = O(2^{-t}); \]

and hence, since $b(x)$ is a decreasing function,

\[ b(x) = O(x^{-1}). \quad (26) \]

Finally, corresponding to any large integer $y$ we may choose $x$ to satisfy

\[ 2^{4^x} < y \leq 2^{4^{x+1}}. \]

Then, by (3), we have

\[ a(y) \leq 2a(2^{4^x}) = 2b(x), \]

so that (26) implies (1).

University College,
London.