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

$$G(x) = b_0(1-\cos\tfrac{1}{2}x)/x + \sum_{n=1}^{\infty} b_n\,[\cos(n-\tfrac{1}{2})x - \cos(n+\tfrac{1}{2})]/x.$$

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

$$G(x) = b_0(1-\cos\tfrac{1}{2}x)/x + f(x)(\sin\tfrac{1}{2}x)/(\tfrac{1}{2}x).$$

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

$$\frac{f(x)}{x} \to \int_0^\infty tg(t)\,dt - \tfrac{1}{8}b_0,$$

as $x \to 0$. This proves (4), since, by the definition of $g(t)$,

$$\int_0^\infty tg(t)\,dt = b_0\int_0^{\frac{1}{2}} t\,dt + \sum_{n=1}^{\infty} b_n \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} t\,dt = \tfrac{1}{8}b_0 + \sum_{n=1}^{\infty} nb_n.$$

References.

1. G. H. Hardy and W. W. Rogosinski, *Fourier Series* (Cambridge Tracts No. 38, 1944).
2. P. Hartman and A. Wintner, " On the behavior of Fourier sine transforms near the origin ", *Proc. American Math. Soc.*, 2 (1951), 398–400.
3. J. Karamata and M. Tomić, " Considérations géometriques relatives aux polynômes et séries trigonométriques ", *Publications de l'Institut Math., Académie Serbe des Sciences*, 2 (1948), 156–175.

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, ..., 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.

† Klaus Roth, *Comptes Rendus*, 234 (1952), 388–390.

‡ 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

$$a(x) = O\left(\frac{1}{\log\log x}\right). \tag{1}$$

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, ... form an $\mathcal{A}$-set then so do u_1, u_2, ..., and conversely.

Further, for any two positive integers x and y we have

$$A(x+y) \leqslant A(x)+A(y).$$

Thus

$$A(xy) \leqslant xA(y), \quad A(x) \leqslant A\left\{\left(\left[\frac{x}{y}\right]+1\right)y\right\} \leqslant \frac{x+y}{y}A(y);$$

so that

$$a(xy) \leqslant a(y), \tag{2}$$

$$a(x) \leqslant (1+yx^{-1})\,a(y). \tag{3}$$

Finally we have the trivial inequality

$$x^{-1} \leqslant a(x) \leqslant 1. \tag{4}$$

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

3. Let u_1, u_2, ..., u_U be any $\mathcal{A}$-set selected from 1, 2, ..., M. In this section we investigate the exponential sum

$$S = \sum_{k=1}^{U} e(\alpha u_k) \qquad [e(\theta) = e^{2\pi i\theta}],$$

where α is a real number. For each α there exist h, q such that

$$\alpha = \frac{h}{q}+\beta, \quad (h, q) = 1, \quad q \leqslant M^{\frac{1}{2}}, \quad q|\beta| \leqslant M^{-\frac{1}{2}}. \tag{5}$$

Suppose $m < M$, and put

$$S' = a(m)\,q^{-1}\left(\sum_{r=1}^{q} e\left(\frac{rh}{q}\right)\right)\left(\sum_{n=1}^{M} e(\beta n)\right),$$

(so that $S' = 0$ if $q > 1$). We prove that

$$|S-S'| < Ma(m)-U+O(mM^{\frac{1}{2}}). \tag{6}$$

To prove this we start from the relation

$$S = \frac{1}{mq} \sum_{r=1}^{q} \sum_{n=1}^{M} \sum_{\substack{n \leqslant u_k < n+mq \\ u_k \equiv r \pmod{q}}} e(\alpha u_k) + O(mq). \tag{7}$$

This relation is obvious on noting that for given u_k, m, q there are exactly mq integers n, satisfying

$$n \leqslant u_k < n+mq,$$

and that these integers n also satisfy $1 \leqslant n \leqslant M$ provided that

$$mq \leqslant u_k < M-mq.$$

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

In the inner sum on the right-hand side of (7), we have

$$e(\alpha u_k) = e\left(\frac{rh}{q}\right) e(\beta n)+O(mq|\beta|).$$

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 $D \geqslant 0$. Hence

$$S = S'-\frac{1}{mq} \sum_{r=1}^{q} e\left(\frac{rh}{q}\right) \sum_{n=1}^{M} e(\beta n) D(n, m, q, r)+O(mq)+O(Mmq|\beta|).$$

If we put $\beta = 0$ and $h = 0$ (legitimate since we have not yet used $(h, q) = 1$), we obtain

$$U = Ma(m)-\frac{1}{mq} \sum_{r=1}^{q} \sum_{n=1}^{M} D(n, m, q, r)+O(mq).$$

Using this as an estimate for

$$\sum_{r=1}^{q} \sum_{n=1}^{M} D(n, m, q, r)$$

in the preceding relation, we obtain (6), on noting that $q \leqslant M^{\frac{1}{2}}$, $q|\beta| \leqslant M^{-\frac{1}{2}}$.

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 $u_1, u_2, \ldots, u_U$ be a maximal $\mathcal{A}$-set selected from $1, 2, \ldots, 2N$, so that $U = A(2N)$. Let $2v_1, 2v_2, \ldots, 2v_V$ be the even integers among the u_k. We note that

$$U = 2Na(2N) \tag{8}$$

and, by (2),

$$U \leqslant 2N a(m), \quad V \leqslant A(N) \leqslant Na(m). \tag{9}$$

Further, since the number of odd integers among the u_k does not exceed $A(N)$, we have, by (2),

$$V \geqslant A(2N) - A(N) \geqslant 2Na(2N) - Na(m). \tag{10}$$

We define

$$f_1(\alpha) = \sum_{k=1}^{U} e(\alpha u_k), \qquad f_2(\alpha) = \sum_{k=1}^{V} e(\alpha v_k);$$

$$F_1(\alpha) = a(m) \sum_{n=1}^{2N} e(\alpha n), \quad F_2(\alpha) = a(m) \sum_{n=1}^{N} e(\alpha n).$$

In view of (9), we have

$$f_r(\alpha) = O\Big(Na(m)\Big), \quad F_r(\alpha) = O\Big(Na(m)\Big); \quad r = 1, 2. \tag{11}$$

We now show that, for any α,

$$f_r(\alpha) - F_r(\alpha) = O\Big(N\{a(m) - a(2N)\} + N^{\frac{3}{4}}\Big); \quad r = 1, 2. \tag{12}$$

If $q = 1$ 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

$$F_r(\alpha) = O(N^{\frac{3}{4}}).$$

This inequality is in fact satisfied, since for any α,

$$\sum_{n=1}^{M} e(\alpha n) = O(\|\alpha\|^{-1}), \tag{13}$$

where $\|\alpha\|$ denotes the distance of α from the nearest integer, and $\|\alpha\| > M^{-\frac{3}{4}}$ if it is impossible to choose $q = 1$ in (5).

Further, using the inequality

$$|f_1 f_2^2 - F_1 F_2^2| = |f_1(f_2^2 - F_2^2) + F_2^2(f_1 - F_1)|$$
$$\leqslant |f_1(f_2 + F_2)(f_2 - F_2)| + |F_2^2(f_1 - F_1)|,$$

we obtain, by (11) and (12),

$$f_1(\alpha) f_2^2(-\alpha) - F_1(\alpha) F_2^2(-\alpha) = O\Big(\{Na(m)\}^2 \Big(N\{a(m) - a(2N)\} + N^{\frac{3}{4}}\Big)\Big). \tag{14}$$

Finally, by (12) and (13), if $0 < \eta < \alpha < 1 - \eta$ we have

$$f_1(\alpha) = O\Big(a(m)\eta^{-1} + N\{a(m) - a(2N)\} + N^{\frac{3}{4}}\Big). \tag{15}$$

The fact that $u_1, u_2, \ldots, u_U$ form an $\mathscr{A}$-set implies that $u_h = v_k + v_l$ if and only if $k = l$ and $u_h = 2v_k$. This can be expressed, following the method of Hardy and Littlewood, by

$$\int_{-\eta}^{1-\eta} f_1(\alpha) f_2^2(-\alpha)\, d\alpha = V \leqslant Na(m). \tag{16}$$

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

$$0 < \eta < \tfrac{1}{2}. \tag{17}$$

Since
$$\int_0^1 |f_2(\alpha)|^2\, d\alpha = V \leqslant Na(m),$$

we have, by (15),

$$\int_{\eta}^{1-\eta} f_1(\alpha) f_2^2(-\alpha)\, d\alpha = O\Big(\big\{a(m)\eta^{-1} + N\big(a(m) - a(2N)\big) + N^{\frac{3}{4}}\big\}\, Na(m)\Big). \tag{18}$$

Further, by (14), we have

$$\int_{-\eta}^{\eta} f_1(\alpha) f_2^2(-\alpha)\, d\alpha$$
$$= \int_{-\eta}^{\eta} F_1(\alpha)\, F_2^2(-\alpha)\, d\alpha + O\Big(\eta \{Na(m)\}^2 \big(N\{a(m) - a(2N)\} + N^{\frac{3}{4}}\big)\Big). \tag{19}$$

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

$$\int_{-\eta}^{\eta} F_1(\alpha)\, F_2^2(-\alpha)\, d\alpha = \int_{-\frac{1}{2}}^{\frac{1}{2}} F_1(\alpha)\, F_2^2(-\alpha)\, d\alpha + O\big(a^3(m)\, \eta^{-2}\big). \tag{20}$$

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 $n \leqslant 2N$, $n' \leqslant N$, $n'' \leqslant N$. This number is N^2. Thus, collecting together the results (16), (18), (19), (20), we obtain

$$a^2(m) = \{N^2 a(m)\}^{-1} \int_{-\frac{1}{2}}^{\frac{1}{2}} F_1(\alpha)\, F_2^2(-\alpha)\, d\alpha$$
$$= O\big(a^2(m)\, N^{-2} \eta^{-2} + \{\eta Na(m) + 1\}\{a(m) - a(2N) + N^{-\frac{1}{4}}\} + a(m)\, N^{-1} \eta^{-1}\big).$$

Hence writing
$$\delta = (N\eta)^{-1}, \tag{21}$$

we obtain, noting that $2N = m^4$,

$$a^2(m) < c_1 \big\{a(m)\,\delta + a^2(m)\,\delta^2 + \big(\delta^{-1} a(m) + 1\big)\big(a(m) - a(m^4) + m^{-1}\big)\big\}. \tag{22}$$

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

We now write

$$m = 2^{4^x}, \quad b(x) = a(m),$$

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 $b(x) \leqslant 1$ by (4), we have

$$c_1\{b(x)\,\delta + b^2(x)\,\delta^2\} \leqslant b^2(x)\left\{\tfrac{1}{2}+\frac{1}{4c_1}\right\} < \tfrac{3}{4}b^2(x).$$

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

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

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

$$Pb^2(2P) \leqslant \sum_{x=P}^{2P-1} b^2(x) < c_5\left(b(P)-b(2P)+\frac{4c_5}{2P}\right) \quad (25)$$

for all integers $P > c_4$.

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

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

This clearly implies (by a backward induction) that if $c_4 < 2^{t_0} < 2^t$, then

$$2^t b(2^t) \leqslant \max\left(4c_5,\; 2^{t_0}b(2^{t_0})\right),$$

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

$$2^{4^x} < y \leqslant 2^{4^{x+1}}.$$

Then, by (3), we have

$$a(y) \leqslant 2a(2^{4^x}) = 2b(x),$$

so that (26) implies (1).

University College,
London.