ON THE REPRESENTATION OF LARGE EVEN INTEGERS AS SUMS OF TWO ALMOST PRIMES. II

By

Saburô UCHIYAMA

In a previous paper [3] the writer has given with Miss A. Togashi an elementary proof for the fact that every sufficiently large even integer is representable as a sum of two almost primes, each of which has at most three prime factors, a result first obtained by A. I. Vinogradov. On the other hand, we are able to prove by a rather transcendental method that every large even integer is representable as a sum of a prime and an almost prime composed of at most four prime factors (see [4]). The aim in the present paper is to show that a somewhat weaker result than this can be obtained by an elementary argument. We shall prove the following[1)]

Theorem. *Every sufficiently large even integer N can be written in the form*

$$N = n_1 + n_2 ,$$

where $n_1 > 1$, $n_2 > 1$, $(n_1, n_2) = 1$ *and*

$$V(n_1) + V(n_2) \leqq 5 .$$

In other words, every large even integer N can be represented in the form $N = n_1 + n_2$, where $n_1 > 1$, $n_2 > 1$, $(n_1, n_2) = 1$ and either

$$V(n_1) = 1 , \qquad V(n_2) \leqq 4 ,$$

or

$$V(n_1) \leqq 2 , \qquad V(n_2) \leqq 3 .$$

Our method of proving this result is a refinement of that of proving the previous one, used in [3].

The writer wishes to express his gratitude to M. Uchiyama, Computation Centre, Hokkaidô University, for providing various numerical data to him in

1) Throughout in this paper, the letters i, j, k, m, n (with or without indices) represent positive integers, while p, q (with or without indices) represent prime numbers. We denote by $V(m)$ the total number of prime divisors of m.

the course of the present investigation.

1. Let N be a sufficiently large but fixed even integer. We consider as in [3] the $\varphi(N)$ integers $a_n=n(N-n)$ $(1\leqq n\leqq N,\ (n, N)=1)$.

Let $x\geqq 2$ be a fixed real number satisfying

$$N^{c_1}<x<N^{c_2},$$

where c_1 and c_2 are constants with $0<c_1<c_2<1$. We denote by $P(x)$ the number of those integers a_n $(1\leqq n\leqq N,\ (n, N)=1)$ which are not divisible by any prime $p\leqq x$. Using the sieve method of A. Selberg (cf. [5, Appendix]) we find that

$$P(x)\leqq\frac{\varphi(N)}{W}+R,$$

where

$$W=\sum_{\substack{m\leqq x^a\\ g(m)\leqq x}}\frac{\mu^2(m)}{f_1(m)}$$

and

$$R=O(B_N x^{2a}(\log\log x)^2),$$

$a>0$ being a constant[2]. In the following we shall suppose that $0<a\leqq 4$.

For brevity's sake we put, for $0<a\leqq 1$,

$$F(a)=F_0(a)\overset{\text{def}}{=}a^2;$$

for $1<a\leqq 2$,

$$F(a)=F_1(a)\overset{\text{def}}{=}(2a-1)^2-2a^2\log a;$$

for $2<a\leqq 3$,

$$F(a)=F_1(a)+F_2(a)+F_{2,1}(a)+F_{2,2}(a),$$

where

$$F_2(a)\overset{\text{def}}{=}(a-2)^2+2a^2\log^2\frac{a}{2}-\frac{1}{2}(a-2)(7a-2)\log\frac{a}{2}$$

and $F_{2,1}(a)$ and $F_{2,2}(a)$ are defined as follows: write

2) For the definitions of $f_1(m)$, $g(m)$, A_N, B_N, C_N we refer to [3].

$$a_i = 1 + \frac{a-2}{i+1} \qquad (i \geqq 1),$$

$$b_i = \frac{a}{2} - \frac{a-2}{4i+2} \qquad (i \geqq 1),$$

and let k_1, k_2 be arbitrary but fixed positive integers. Then

$$F_{2,1}(a) = \sum_{j=1}^{k_1} \left(8(a_{j+1}-1)(a_j - a_{j+1}) + 4a^2 \log a_{j+1} \log \frac{a-a_{j+1}}{a-a_j} \right.$$
$$-2(a_j - a_{j+1})(2a + a_j + a_{j+1}) \log a_{j+1}$$
$$\left. -2(a_{j+1}-1)(4a - a_{j+1} - 1) \log \frac{a-a_{j+1}}{a-a_j} \right),$$

$$F_{2,2}(a) = \sum_{j=1}^{k_2} \left(8(b_{j+1}-b_j)\left(\frac{a}{2} - b_{j+1}\right) + 4a^2 \log \frac{b_{j+1}}{b_j} \log \frac{2(a-b_{j+1})}{a} \right.$$
$$-\left(\frac{a}{2} - b_{j+1}\right)(5a + 2b_{j+1}) \log \frac{b_{j+1}}{b_j}$$
$$\left. -2(b_{j+1}-b_j)(4a - b_j - b_{j+1}) \log \frac{2(a-b_{j+1})}{a} \right).$$

Finally, we set for $3 < a \leqq 4$

$$F(a) = F_1(a) + F_2(a) + F_{2,1}(a) + F_{2,2}(a) - F_3(a),$$

where

$$F_3(a) \overset{\text{def}}{=} \frac{4}{9}(a-3)^2 \log \frac{a}{3} \cdot \left(3 \log^2 \frac{a}{3} \right.$$
$$\left. + 4 \log \frac{a}{3} \log \frac{3(a-2)}{a} + \log^2 \frac{2a-3}{a} \right).$$

(Note that the function $F(a)$ thus defined is positive and continuous for $0 < a \leqq 4$.) We have then, for $0 < a \leqq 4$,

$$W \geqq F(a) C_N \log^2 x + O(\log N \log \log N).$$

Indeed, it is easy to verify this relation for $0 < a \leqq 2$ (cf. [3]). For $2 < a \leqq 3$ we have

$$W = \sum_{m \leqq x^a} \frac{\mu^2(m)}{f_1(m)} - \sum_{x < p \leqq x^a} \sum_{\substack{m \leqq x^a \\ m \equiv 0\ (p)}} \frac{\mu^2(m)}{f_1(m)} + \sum_{\substack{x < p_1 < p_2 \\ p_1 p_2 \leqq x^a}} \sum_{\substack{m \leqq x^a \\ m \equiv 0\ (p_1 p_2)}} \frac{\mu^2(m)}{f_1(m)},$$

where the last double summation is found to be

$$\geqq \frac{1}{2} \sum_{\substack{x<p_1\leqq x^{a/2} \\ x<p_2\leqq x^{a/2}}} \sum_{\substack{m\leqq x^a \\ m\equiv 0\,(p_1p_2)}} \frac{\mu^2(m)}{f_1(m)} + O\left(\frac{\log^2 x}{x}\right)$$

$$+\sum_{j=1}^{k_1} \sum_{\substack{x<p_1\leqq x^{a_{j+1}} \\ x^{a-a_j}<p_2\leqq x^{a-a_{j+1}}}} \sum_{\substack{m\leqq x^a \\ m\equiv 0\,(p_1p_2)}} \frac{\mu^2(m)}{f_1(m)} + \sum_{j=1}^{k_2} \sum_{\substack{x^{b_j}<p_1\leqq x^{b_{j+1}} \\ x^{a/2}<p_2\leqq x^{a-b_{j+1}}}} \sum_{\substack{m\leqq x^a \\ m\equiv 0\,(p_1p_2)}} \frac{\mu^2(m)}{f_1(m)}$$

$$= (F_2(a)+F_{2,1}(a)+F_{2,2}(a))\, C_N \log^2 x + O(\log N \log\log N)\,,$$

and for $3<a\leqq 4$ we have

$$W = \sum_{m\leqq x^a} \frac{\mu^2(m)}{f_1(m)} - \sum_{x<p\leqq x^a} \sum_{\substack{m\leqq x^a \\ m\equiv 0\,(p)}} \frac{\mu^2(m)}{f_1(m)}$$

$$+ \sum_{\substack{x<p_1<p_2 \\ p_1p_2\leqq x^a}} \sum_{\substack{m\leqq x^a \\ m\equiv 0\,(p_1p_2)}} \frac{\mu^2(m)}{f_1(m)} - \sum_{\substack{x<p_3<p_4<p_5 \\ p_3p_4p_5\leqq x^a}} \sum_{\substack{m\leqq x^a \\ m\equiv 0\,(p_3p_4p_5)}} \frac{\mu^2(m)}{f_1(m)}\,,$$

where the last double summation is (in absolute value)

$$\leqq \left(\frac{1}{6} \sum_{\substack{x<p_3\leqq x^{a/3} \\ x<p_4\leqq x^{a/3} \\ x<p_5\leqq x^{a/3}}} + \frac{1}{2} \sum_{\substack{x<p_3\leqq x^{a/3} \\ x<p_4\leqq x^{a/3} \\ x^{a/3}<p_5\leqq x^{a-2}}} + \frac{1}{2} \sum_{\substack{x<p_3\leqq x^{a/3} \\ x^{a/3}<p_4\leqq x^{(2a/3)-1} \\ x^{a/3}<p_5\leqq x^{(2a/3)-1}}} \right) \sum_{\substack{m\leqq x^a \\ m\equiv 0\,(p_3p_4p_5)}} \frac{\mu^2(m)}{f_1(m)}$$

$$+ O\left(\frac{\log^2 x}{x}\right) = F_3(a)\, C_N \log^2 x + O(\log N \log\log N)\,.$$

This proves our assertion. Hence:

Lemma 1. *For* $0<a\leqq 4$ *we have*

$$P(x) \leqq \frac{2e^{2C}}{F(a)} A_N \frac{\varphi(N)}{\log^2 x} + O\left(\frac{\varphi(N)(\log\log N)^5}{\log^3 N}\right) + O(B_N x^{2a} (\log\log N)^2)\,.$$

2. We now evaluate $P(N^{\frac{1}{u}})$ for some values of u $(\geqq 2)$. The result to be obtained will be of the form either

$$P(N^{\frac{1}{u}}) \leqq A(u) A_N \frac{\varphi(N)}{\log^2 N} + O\left(\frac{\varphi(N)(\log\log N)^5}{\log^3 N}\right)$$

or

$$P(N^{\frac{1}{u}}) \geqq a(u) A_N \frac{\varphi(N)}{\log^2 N} + O\left(\frac{\varphi(N)(\log\log N)^5}{\log^3 N}\right).$$

Here, it is clear that $0\leqq a(u) \leqq A(u) < \infty$ $(u\geqq 2)$: moreover, we may assume without loss of generality that each of the coefficients $A(u)$ and $a(u)$ is, as a function of u, monotone non-decreasing for $u \geqq 2$.

Lemma 2. *We have*

$$a(10) = 98.0 .$$

This is in substance identical with [1, Lemma 1]. The result is obtained by simply applying the sieve method of Viggo Brun. We note a rather better result than the above, viz. $a(10)=99.9818$, is known (see [2]): but this is not necessary for our purpose.

Lemma 3. *For* $8 \leqq u \leqq 9$ *we have*

$$A(u) = 1.0541 u^2 .$$

If we apply Lemma 1 with $x = N^{\frac{1}{u}}$ $(8 \leqq u \leqq 9)$ and $a = 3.95$, then we get

$$P(N^{\frac{1}{u}}) \leqq \frac{2e^{2C}u^2}{F(3.95)} A_N \frac{\varphi(N)}{\log^2 N} + O\left(\frac{\varphi(N)(\log\log N)^5}{\log^3 N}\right),$$

and the result follows from this at once, since we have

$$\frac{2e^{2C}}{F(3.95)} < 1.0541 ,$$

where we have taken $k_1 = 20$ and $k_2 = 5$.

Lemma 4. *Suppose that* $3 \leqq u \leqq u_1 \leqq 10$. *If we set*

$$a_1(u) = \max\left(a(u),\ a(u_1) - 2\int_{u-1}^{u_1-1} A(v)\frac{v+1}{v^2}\,dv\right),$$

then the coefficient $a(u)$ *can be replaced by the new one,* $a_1(u)$.

This is [1, Theorem 1].

Now, it is known that $a(9)=75.58$ (see [1, p. 385]), while we find using the results in Lemmas 2 and 3 that

$$a(10) - 2\int_8^9 A(v)\frac{v+1}{v^2}\,dv = a(10) - 2\cdot 1.0541\int_8^9 (v+1)\,dv$$
$$= 98.0 - 20.0279 = 77.9721 .$$

It follows from Lemma 4 (and the definition of $a(u)$) that $a_1(9)=77.9721$, and we thus have proved the following

Lemma 5. *We have*

$$S_1 \overset{\text{def}}{=} P(N^{\frac{1}{9}}) \geqq 77.9721\, A_N \frac{\varphi(N)}{\log^2 N} + O\left(\frac{\varphi(N)(\log\log N)^5}{\log^3 N}\right).$$

3. Let q, $(q, N)=1$, be a fixed prime number in the interval $z<q\leqq z_1$, where

$$z = N^{\frac{1}{9}}, \qquad z_1 = N^{\frac{5}{9}},$$

and let $S(q)$ denote the number of those integers $a_n = n(N-n)$ $(1\leqq n\leqq N,\ (n, N)=1)$ which are not divisible by any prime $p\leqq z$ and are divisible by the prime q. Then we find as in [3] that

$$S(q) \leqq \frac{2\varphi(N)}{qW_q} + R,$$

where

$$W_q \geqq \sum_{\substack{m\leqq z^a \\ g(m)\leqq z}} \frac{\mu^2(m)}{f_1(m)} \quad \text{with} \quad a = 4.5\left(1-2\varepsilon-\frac{\log q}{\log N}\right)$$

and

$$R_q = O\left(\frac{N^{1-\varepsilon}(\log\log N)^2}{q}\right),$$

ε being a sufficiently small, fixed positive real number. Clearly we have $0<a<4$. Hence we may repeat the argument in the proof of Lemma 1 to obtain

$$W_q \geqq F(a)\, C_N \log^2 z + O(\log N \log\log N),$$

so that

$$S(q) \leqq \frac{C_\varepsilon(t_q)}{q} A_N \frac{\varphi(N)}{\log^2 N} + O\left(\frac{\varphi(N)(\log\log N)^5}{q\log^3 N}\right),$$

where we have put $t_q = (\log N)/\log q$ and

$$C_\varepsilon(t) = \frac{324e^{2C}}{F(a)} \quad \text{with} \quad a = 4.5\left(1-2\varepsilon-\frac{1}{t}\right).$$

Now, if we denote by S_2 the number of those integers a_n $(1\leqq n\leqq N,\ (n, N)=1)$ which are not divisible by any prime $p\leqq z$ and are divisible by at least four distinct primes q, $(q, N)=1$, in the interval $z<q\leqq z_1$, then

$$S_2 \leqq \frac{1}{4} \sum_{\substack{z<q\leqq z_1 \\ (q,N)=1}} S(q).$$

Since we have

$$\sum_{\substack{z<q\leqq z_1 \\ (q,N)=1}} \frac{C_\varepsilon(t_q)}{q} \leqq \int_{1.8}^{9} \frac{C_\varepsilon(t)}{t} dt + O\left(\frac{1}{\log^{1/2} N}\right),$$

we obtain the following lemma:

Lemma 6. *We have*

$$S_2 \leqq \frac{1}{4} \int_{1.8}^{9} \frac{C_\varepsilon(t)}{t} dt\, A_N \frac{\varphi(N)}{\log^2 N} + O\left(\frac{\varphi(N)(\log\log N)^2}{\log^{5/2} N}\right).$$

4. We shall show that for some sufficiently samll $\varepsilon>0$ we have

$$I(\varepsilon) \overset{\text{def}}{=} \frac{1}{4} \int_{1.8}^{9} \frac{C_\varepsilon(t)}{t} dt \leqq 77.5008 .$$

To do this it will obviously suffice to prove that

$$I(0) < 77.500719 ,$$

since the integral defining $I(\varepsilon)$ is, as a function of ε, continuous at $\varepsilon=0$.

We have

$$I(0) = \frac{1}{4} \int_{1.8}^{9} \frac{C_0(t)}{t} dt = 81e^{2C} \int_{2}^{4} \frac{ds}{(4.5-s)F(s)},$$

on substituting $s=4.5\left(1-\frac{1}{t}\right)$. It is not difficult to verify—by a simple but rather tedious calculation—that the function

$$k(s) = \frac{1}{(4.5-s)F(s)}$$

is convex on either of the intervals $2\leqq s\leqq 3$ and $3\leqq s\leqq 4$ (indeed, $k(s)$ may be convex throughout on the interval $2\leqq s\leqq 4$). Hence, in order to estimate from above the value of the integral of $k(s)$ over the interval $2\leqq s\leqq 4$, we may apply the trapezoidal rule with an arbitrary set of division points including the point $s=3$. Thus

$$\int_2^4 k(s)ds \leqq 0.05\left(\frac{1}{2}k(2.00) + \sum_{j=1}^{39} k(2.00+0.05j) + \frac{1}{2}k(4.00)\right).$$

Taking again $k_1=20$ and $k_2=5$, we find that:

$k(2.00)<0.115781$,	$k(2.05)<0.114122$,
$k(2.10)<0.112731$,	$k(2.15)<0.111591$,
$k(2.20)<0.110689$,	$k(2.25)<0.110014$,
$k(2.30)<0.109555$,	$k(2.35)<0.109306$,
$k(2.40)<0.109262$,	$k(2.45)<0.109417$,
$k(2.50)<0.109769$,	$k(2.55)<0.110317$,
$k(2.60)<0.111063$,	$k(2.65)<0.111961$,
$k(2.70)<0.113156$,	$k(2.75)<0.114535$,
$k(2.80)<0.116084$,	$k(2.85)<0.117881$,
$k(2.90)<0.119914$,	$k(2.95)<0.122197$,
$k(3.00)<0.124746$,	$k(3.05)<0.127505$,
$k(3.10)<0.130723$,	$k(3.15)<0.134199$,
$k(3.20)<0.138042$,	$k(3.25)<0.142288$,
$k(3.30)<0.146983$,	$k(3.35)<0.152180$,
$k(3.40)<0.157943$,	$k(3.45)<0.164347$,
$k(3.50)<0.171494$,	$k(3.55)<0.179491$,
$k(3.60)<0.188478$,	$k(3.65)<0.198634$,
$k(3.70)<0.210176$,	$k(3.75)<0.223384$,
$k(3.80)<0.238618$,	$k(3.85)<0.256350$,
$k(3.90)<0.277210$,	$k(3.95)<0.302060$,
$k(4.00)<0.332108$.	

The above table was computed on the electronic digital computer, HIPAC 103.

These data will give

$$\int_2^4 k(s)ds < 0.301618\,,$$

and consequently

$$I(0) < 81e^{2C}\cdot 0.301618 < 77.500719\,.$$

This is the result which was to be shown.

5. Let us fix $\varepsilon>0$ so small as to satisfy $I(\varepsilon)\leqq 77.5008$, and suppose that $N\geqq N_0=N_0(\varepsilon)$ be a sufficiently large even integer. Let S denote the number of those integers $a_n=n(N-n)$ $(1\leqq n\leqq N,\ (n, N)=1)$ which are divisible by no primes $p\leqq z$, by at most three primes q, $(q, N)=1$, in the interval $z<q\leqq z_1$, and by no integers of the form q^2 with q in the interval $z<q\leqq z_1$, where, as before,

$$z=N^{\frac{1}{9}}\,,\qquad z_1=N^{\frac{5}{9}}\,.$$

Since the number S_3 of those integers a_n ($1 \leqq n \leqq N$, $(n, N)=1$) which are not divisible by any prime $p \leqq z$ and are divisible by some integer q^2 with q, $(q, N)=1$, in $z<q \leqq z_1$ is of $O(N^{\frac{8}{9}})$, we have, by virtue of Lemmas 5 and 6,

$$\begin{aligned} S &\geqq S_1 - S_2 - S_3 \\ &\geqq (77.9721 - 77.5008) A_N \frac{\varphi(N)}{\log^2 N} + O\left(\frac{\varphi(N)(\log\log N)^2}{\log^{5/2} N}\right) \\ &> 0.4712 A_N \frac{\varphi(N)}{\log^2 N} > 2 . \end{aligned}$$

This implies the existence of at least one integer n with $1<n<N-1$, $(n, N)=1$, such that $V(a_n) \leqq 5$, i.e.

$$V(n) + V(N-n) \leqq 5 ,$$

which completes the proof of our theorem, since $N=n+(N-n)$.

References

[1] A. A. BUHŠTAB: New improvements in the sieve method of Eratosthenes, Matem. Sbornik, vol. 4 (46) (1938), pp. 375–387 (in Russian).

[2] A. A. BUHŠTAB: On the representation of even numbers as the sum of two numbers with a bounded number of prime factors, Dokl. Akad. Nauk SSSR, vol. 29 (1940), pp. 544–548 (in Russian).

[3] A. TOGASHI and S. UCHIYAMA: On the representation of large even integers as sums of two almost primes, I, J. Fac. Sci., Hokkaidô Univ., Ser. I, vol. 18 (1964), pp. 60–68.

[4] M. UCHIYAMA and S. UCHIYAMA: On the representation of large even integers as sums of a prime and an almost prime, Proc. Japan Acad., vol. 40 (1964), pp. 150–154.

[5] S. UCHIYAMA: On the distribution of almost primes in an arithmetic progression, J. Fac. Sci., Hokkaidô Univ., Ser. I, vol. 18 (1964), pp. 1–22.

Department of Mathematics,
Hokkaidô University

(Received April 1, 1964)