JOURNAL OF NUMBER THEORY 2, 379-385 (1970)

On Sums of Primes and l-th Powers of Small Integers

K. PRACHAR

Institut für Mathematik der Hochschule Bodenkultur, 1180 Vienna, Austria

Communicated by P. T. Bateman

DEDICATED TO PROF. TH. SCHNEIDER ON THE OCCASION OF HIS 60TH BIRTHDAY

Received May 5, 1969

Given a positive integer l, this paper establishes the existence of constants $\eta > 1$ and $\delta > 0$ such that for large N at least δN of the positive integers up to N are not expressible in the form $p + m^l$, where p is a prime and m is a positive integer not exceeding $\eta \log N$.

Using methods of Linnik [2] [3], the author has shown [4] that almost all integers are representable as a sum of a prime p and of the l-th power m^l of an integer m with $m^l < p^\lambda$, where l is a fixed integer $\geqslant 2$ and λ is a suitably chosen constant satisfying $0 < \lambda < 1$. If the Riemann-hypothesis is true for all L-functions (or even only the so called density hypothesis), p^λ can be replaced by a certain power of $\log p$. (For the case $l = 1$ this was proved by A. Selberg [5] and follows also from Linnik's work cited above; the theorem holds also for $l \geqslant 2$ with the value $\lambda = 19/77$ given by A. Selberg.) In 1964 Erdös proposed to show that there are more than $o(N)$ integers $\leqslant N$ not representable in the form $p + m^l$, $1 \leqslant m \leqslant \log N$, p a prime, and gave a proof* for $l = 1$. With the help of a new method of Bombieri and Davenport [1] we prove Erdös's conjecture for $l \geqslant 2$ and also a result concerning the existence of "small" differences $p' - p$ which are of the form $m_2{}^l - m_1{}^l$.

THEOREM 1. *There are infinitely many pairs of primes p, p' satisfying*

$$p' - p = m_2{}^l - m_1{}^l$$

with $0 < m_1 < m_2 < (\frac{1}{2} + \epsilon) \log p$, where $\epsilon > 0$ is an arbitrary small fixed number.

* Oral communication by Professor Erdös. Selberg states (*loc. cit.*) the same with $1 \leqslant m \leqslant K \log N$, K an arbitrary large positive constant. But, as far as the author knows, no proof of this result has been published.

641/2/4-1

THEOREM 2. *There is a constant $\eta > 1$ and a positive constant δ so that for large N at least δN integers $\leqslant N$ are not representable in the form*

$$p + m^l, \qquad 1 \leqslant m \leqslant \eta \log N.$$

Here η and δ may depend on l.

Proof of Theorem 1. Let k be any positive integer and

$$T(\alpha) = \sum_{n=-k}^{k} t(n)\, e(2n\alpha) \qquad (e(\vartheta) = e^{2\pi i\vartheta}) \tag{1}$$

any trigonometric polynomial with real coefficients $t(n)$ satisfying $t(-n) = t(n)$, which is nonnegative for all real α. Define

$$Z(N; 2n) = \sum_{\substack{p\leqslant N,\, p'\leqslant N \\ p'-p=2n}} (\log p)(\log p'). \tag{2}$$

Then Bombieri and Davenport in their paper cited above proved: If $k < (\log N)^C$ for some fixed C, then for any fixed positive ϵ we have

$$\sum_{n=1}^{k} t(n)\, Z(N; 2n) > 2N \sum_{n=1}^{k} t(n)\, H(n) - (\tfrac{1}{4} + \epsilon)\, t(0)\, N \log N, \tag{3}$$

provided N is sufficiently large, where

$$H(n) = H \prod_{\substack{p\mid n \\ p>2}} \frac{p-1}{p-2} \tag{4}$$

and

$$H = \prod_{p>2} \{1 - (p-1)^{-2}\}. \tag{5}$$

We put

$$T(\alpha) = \Big| \sum_{g=1}^{P} e((2g)^l\, \alpha) \Big|^2 + \Big| \sum_{g=1}^{P} e((2g-1)^l\, \alpha) \Big|^2, \tag{6}$$

where $P = [\frac{1}{2}\eta \log N]$ for some constant η to be fixed later. Then $t(0) = 2P$ and, for $1 \leqslant n < (2P)^l$, $t(n)$ is the number of representations of $2n$ in the form

$$2n = m_2{}^l - m_1{}^l, \qquad 0 < m_1 < m_2 \leqslant 2P, \qquad m_1 \equiv m_2 \bmod 2. \tag{7}$$

Therefore, if $n \geqslant 1$, we have $t(n) \neq 0$ only if n has the form

$$n = \tfrac{1}{2}(m_2^l - m_1^l),$$

where

$$0 < m_1 < m_2 \leqslant 2P, \qquad m_1 \equiv m_2 \bmod 2. \tag{8}$$

In the following considerations we always impose on m_1 and m_2 the above conditions (8) without saying so explicitly. We get from (3)

$$\sum_{m_1}\sum_{m_2} Z(N; m_2^l - m_1^l) > 2N \sum_{m_1}\sum_{m_2} H(\tfrac{1}{2}(m_2^l - m_1^l)) - (\tfrac{1}{4} + \epsilon)\, 2PN \log N. \tag{9}$$

From the definition (4) of $H(n)$ we conclude that

$$H(\tfrac{1}{2}(m_2^l - m_1^l)) \geqslant H(\tfrac{1}{2}(m_2 - m_1)),$$

since $(m_2 - m_1) \mid (m_2^l - m_1^l)$, and therefore

$$\sum_{m_1}\sum_{m_2} H(\tfrac{1}{2}(m_2^l - m_1^l)) \geqslant \sum_{m_1 \equiv m_2 \equiv 0 \bmod 2}\sum H(\tfrac{1}{2}(m_2 - m_1)) + \sum_{m_1 \equiv m_2 \equiv 1 \bmod 2}\sum H(\tfrac{1}{2}(m_2 - m_1)). \tag{10}$$

Both sums on the right hand side are equal to

$$\sum_{0 < g_1 < g_2 \leqslant P}\sum H(g_2 - g_1). \tag{11}$$

This is the sum arising from the trigonometric polynomial

$$\left| \sum_{g=1}^{P} e(2g\alpha) \right|^2 = \left(\frac{\sin 2\pi P\alpha}{\sin 2\pi\alpha} \right)^2$$

which was treated by Bombieri and Davenport. They proved that

$$\sum_{0 < g_1 < g_2 \leqslant P}\sum H(g_2 - g_1) > \tfrac{1}{2}(1 - \epsilon)\, P^2. \tag{12}$$

Substituting this in (10) and (9) we get

$$\sum_{m_1}\sum_{m_2} Z(N; m_2^l - m_1^l) > 2(1 - \epsilon)\, P^2 N - (\tfrac{1}{4} + \epsilon)\, 2PN \log N. \tag{13}$$

With $P = [(\frac{1}{4} + 3\epsilon) \log N]$, this gives as in Bombieri-Davenport *loc. cit.*

$$\sum_{m_1} \sum_{m_2} Z(N; m_2^l - m_1^l) > 2\epsilon P^2 N.$$

The number of terms on the left hand side is $P^2 - P < P^2$. Then there is some number $2n = m_2^l - m_1^l$ for which

$$Z(N; 2n) = \sum_{\substack{p \leqslant N,\, p' \leqslant N \\ p' - p = 2n}} (\log p)(\log p') > 2\epsilon N.$$

It is impossible that all the terms in the sum should have $p \leqslant N(\log N)^{-2}$, for then the sum would be less than

$$\sum_{p \leqslant N(\log N)^{-2}} (\log N)^2 = O(N(\log N)^{-1}).$$

Hence, for all sufficiently large N, there is a pair of primes p, p' with

$$N(\log N)^{-2} < p \leqslant N, \qquad p' \leqslant N,$$
$$p' - p = m_2^l - m_1^l, \qquad 0 < m_1 < m_2 \leqslant 2[(\tfrac{1}{4} + 3\epsilon) \log N],$$

which, because of $\log N \sim \log p$ for large N, implies Theorem 1.

Proof of Theorem 2. (13) gives, with $P = [\frac{1}{2}\eta \log N]$,

$$\sum_{m_1} \sum_{m_2} Z(N; m_2^l - m_1^l) > \left(\frac{\eta^2}{2} - \frac{\eta}{4} - 4\epsilon\right) N \log^2 N, \tag{14}$$

if $\eta < 2$, say. Now let

$$Z_1(N; 2n) = \sum_{\substack{p \leqslant N,\, p' \leqslant N \\ p' - p = 2n}} 1.$$

Then

$$Z(N; 2n) \leqslant Z_1(N; 2n) \log^2 N$$

and we have by (14)

$$\sum_{m_1} \sum_{m_2} Z_1(N; m_2^l - m_1^l) > \left(\frac{\eta^2}{2} - \frac{\eta}{4} - 4\epsilon\right) N. \tag{15}$$

Now we define

$$r(n) = \sum_{\substack{p \leqslant N, m \leqslant 2P \\ p + m^l = n}} 1.$$

It is easily seen from the prime number theorem that

$$\sum_{n\leqslant N} r^2(n) = 2\sum_{m_1}\sum_{m_2} Z_1(N; m_2^l - m_1^l) + \eta N + O(N(\log N)^{-1}), \quad (16)$$

the last two terms arising from the case $m_1 = m_2$. Thus

$$\sum_{n\leqslant N} r^2(n) > \left(\eta^2 + \frac{\eta}{2} - 9\epsilon\right) N. \quad (17)$$

On the other side we have

$$\sum_{n\leqslant N} r(n) = \eta N + O(N(\log N)^{-1}). \quad (18)$$

We denote by $A_k(N)$ the number of n's with $n \leqslant N$, $r(n) = k$, that is

$$A_k(N) = \sum_{n\leqslant N, r(n)=k} 1, \qquad k = 1, 2,\ldots .$$

(Trivially we have $k \leqslant 2P$.) The sum $\sum_{n\leqslant N+(2P)^l} r^3(n)$ is the number of solutions of the system

$$p_1 + m_1^l = p_2 + m_2^l = p_3 + m_3^l$$

with $p_j \leqslant N$, $m_j \leqslant 2P \leqslant \eta \log N$. By Brun's method† we conclude for every integer $L \geqslant 2$

$$\sum_{n\leqslant N, r(n)>L} r^2(n) \leqslant \frac{1}{L} \sum_{n\leqslant N} r^3(n)$$

$$< \frac{c}{L} \frac{N}{\log^3 N} (\eta \log N)^3 = \frac{c}{L} \eta^3 N, \quad (19)$$

where c is a constant which may depend on l. This gives with (17), since $A_1(N) \leqslant N$,

$$4A_2(N) + 9A_3(N) + \cdots + L^2 A_L(N)$$

$$> \left(\eta^2 + \frac{\eta}{2} - 1 - \frac{c}{L}\eta^3 - 9\epsilon\right) N. \quad (20)$$

† Brun's method gives for the number or primes p less than N with $p + a$ and $p + b$ also prime the upper bound

$$C \frac{N}{\log^3 N} \prod_{p\mid ab} \left(1 + \frac{1}{p}\right) \prod_{p\mid (a-b)} \left(1 + \frac{1}{p}\right)$$

Putting $p_2 = p$, $a = m_2^l - m_1^l$, $b = m_2^l - m_3^l$, and summing over m_1, m_2, m_3, one gets the estimate used in (19) (see for example the analogous considerations in [4], p. 419).

Further, we write

$$\sum_{n\leqslant N, r(n)\neq 0} 1 = A_1(N) + A_2(N) + \cdots + A_L(N) + \cdots$$

$$= \{A_1(N) + 2A_2(N) + \cdots + LA_L(N) + \cdots\}$$

$$- \{A_2(N) + 2A_3(N) + \cdots + (L-1)\, A_L(N) + \cdots\}$$

$$= \sum_{n\leqslant N} r(n) - \{A_2(N) + \cdots + (L-1)\, A_L(N) + \cdots\}$$

$$\leqslant \eta N - \{A_2(N) + \cdots + (L-1)\, A_L(N)\} + O(N(\log N)^{-1}). \tag{21}$$

By (20) we get the estimate

$$A_2(N) + \cdots + (L-1)\, A_L(N) \geqslant \frac{1}{2L}\,(4A_2(N) + \cdots + L^2 A_L(N))$$

$$> \frac{1}{2L}\left(\eta^2 + \frac{\eta}{2} - 1 - \frac{c}{L}\,\eta^3 - 9\epsilon\right) N.$$

Expanding the coefficient of N with respect to powers of $\eta - 1$ and putting $L = 4c$, where we can suppose that c is an integer, we find by (21), if $1 \leqslant \eta < \frac{3}{4}$,

$$\sum_{n\leqslant N, r(n)\neq 0} 1 < \left(1 - \frac{1}{32c} + (\eta - 1)\left(1 - \frac{7}{32c}\right) + 4\epsilon\right) N.$$

The coefficient of N is less than one, if we put, for example,

$$\eta = 1 + \frac{1}{32c},$$

and Theorem 2 is proved.

Acknowledgment

The author wishes to express his gratitude to Professor Bateman and the referee for their help and corrections.

References

1. E. Bombieri and H. Davenport. Small differences between prime numbers. *Proc. of the Royal Soc. Ser. A*, **293** (1966), 1–18.
2. Ju. V. Linnik. Some conditional theorems connected with the binary Goldbach problem. *Isvestija Akad. Nauk SSSR*, Ser. Mat. **16** (1952), 503–520 (in Russian).

3. JU. V. LINNIK. Composition of prime numbers with powers of one and the same number. *Mat. Sb.* **32** (1953), 3–60 (in Russian).
4. K. PRACHAR. Über Zahlen, die sich als Summe einer Primzahl und einer "kleinen" Potenz darstellen lassen. *Monatsh. Math.* **68** (1964), 409–420, und **69** (1965), 62–68.
5. A. SELBERG. On the normal density of primes in small intervals, and the difference between consecutive primes. *Arch. Math. og Naturv.* B **47** (1943), No. 6, 87–105.