ACTA ARITHMETICA
VI (1960)

Remarks on number theory III

On addition chains

by

P. Erdös (Budapest)

Consider a sequence $a_0 = 1 < a_1 < a_2 < \ldots < a_k = n$ of integers such that every a_l $(l \geqslant 1)$ can be written as the sum $a_i + a_j$ of two preceding elements of the sequence. Such a sequence has been called by A. Scholz ([1]) an *addition chain*. He defines $l(n)$ as the smallest k for which there exists an addition chain $1 = a_0 < a_1 < \ldots < a_k = n$.

Clearly $l(n) \geqslant \log n / \log 2$, the equality occurring only if $n = 2^u$. Scholz conjectured that

$$\lim_{n\to\infty} l(n) \frac{\log 2}{\log n} = 1 \tag{1}$$

and A. Brauer ([2]) proved (1). In fact Brauer proved that

$$l(n) \leqslant \min_{1\leqslant r\leqslant m} \left\{\left(1+\frac{1}{r}\right)\frac{\log n}{\log 2} + 2^r - 2\right\} \tag{2}$$

where $2^m \leqslant n < 2^{m+1}$. From (2) by choosing $r = \left[(1-\varepsilon)\dfrac{\log\log n}{\log 2}\right]$ it follows that

$$l(n) < \frac{\log n}{\log 2} + \frac{\log n}{\log\log n} + o\left(\frac{\log n}{\log\log n}\right). \tag{3}$$

In the present note I am going to prove that (3) is the best possible. In fact I shall prove the following

Theorem. *For almost all n (i. e. for all n except a sequence of density 0)*

$$l(n) = \frac{\log n}{\log 2} + \frac{\log n}{\log\log n} + o\left(\frac{\log n}{\log\log n}\right).$$

([1]) Jahresbericht der Deutschen Math. Vereinigung 47 (1937), p. 41.

([2]) Bull. Amer. Math. Soc. 45 (1939), p. 736-739.

In view of (3) it will suffice to prove that for every ε the number of integers m satisfying

$$(4) \qquad \frac{n}{2} < m < n, \quad l(m) < \frac{\log n}{\log 2} + (1-\varepsilon)\frac{\log n}{\log\log n}$$

is $o(n)$. In fact we shall prove that the number of integers satisfying (4) is less than $n^{1-\eta}$ for some $\eta = \eta(\varepsilon) > 0$.

To prove our assertion we shall show (as the stronger result) that the number of addition chains $1 = a_0 < a_1 < \ldots < a_k$ satisfying

$$(5) \qquad \frac{n}{2} < a_k < n, \quad k < \frac{\log n}{\log 2} + (1-\varepsilon)\frac{\log n}{\log\log n}$$

is less than $n^{1-\eta}$ for some $\eta > 0$ ($\eta = \eta(\varepsilon)$).

An addition chain is clearly determined by its length k and by a mapping $\psi(i)$, $1 \leqslant i \leqslant k-1$, which associates with i two indices $j_1^{(i)}$ and $j_1'^{(i)}$ not exceeding i. To such a mapping there corresponds an addition chain if and only if for every i, $a_{j_1^{(i)}} + a_{j_1'^{(i)}} > a_i$.

We split the indices i, $2 \leqslant i \leqslant k-1$, into three classes. In the first class are the indices i for which $a_{i+1} = 2a_i$. In the second class are the i's for which $a_{i+1} < 2a_i$ and $a_{i+1} \geqslant (1+\delta)^r a_{i+1-r}$ for every $r > 0$ ($\delta = \delta(\varepsilon)$ is a sufficiently small positive number). In the third class are the i's for which $a_{i+1} < 2a_i$ and $a_{i+1} < (1+\delta)^r a_{i+1-r}$ for some $r > 0$. Denote the number of i's in the classes by u_1, u_2, u_3, $u_1 + u_2 + u_3 = k-1$.

Assume now that (5) is satisfied, we are going to estimate the number of addition chains satisfying (5). First we show that (5) implies

$$(6) \qquad u_2 + u_3 = o(k).$$

To prove (6) observe that if $a_{i+1} \neq 2a_i$ then $a_{i+1} \leqslant a_i + a_{i-1}$. Thus from $a_i \leqslant 2a_{i-1}$ we obtain

$$(7) \qquad a_{i+1} \leqslant 3a_{i-1}.$$

Thus from (5) and (7), since there are at least $\frac{1}{2}[(u_2+u_3)] = [\frac{1}{2}(k-u_1-1)]-1$ intervals $(i-1, i+1)$, $1 \leqslant i \leqslant k-1$, which are disjoint half-open (i. e. open to the left) and for which i is in the second or third class, we have

$$\frac{n}{2} < a_k < 2^{u_1+1}3^{(k-u_1)/2} = 2^k \cdot \frac{2}{(\frac{4}{3})^{(k-u_1)/2}} < 2^{k-(u_2+u_3)/100}$$

or $k > \dfrac{\log n}{\log 2}\left(1 + \dfrac{u_2+u_3}{100}\right) - 1$, which contradicts (4) if (6) is not satisfied.

The number of ways in which we can split the indices i into three classes having u_1, u_2, u_3 elements $(u_1+u_2+u_3 = k-1)$ equals $\binom{k-1}{u_2+u_3}\times$ $\times\binom{u_2+u_3}{u_2}$. Now since $u_2+u_3 = o(k)$, $\binom{u_2+u_3}{u_2} < 2^{u_2+u_3} = (1+o(1))^k$, also $\binom{k}{u_2+u_3}\binom{k}{u_2+u_3} = \binom{k}{o(k)} = (1+o(1))^k$. Further for u_2 and u_3 we have at most k^2 choices. Thus the total number of ways of splitting the indices into three classes is $(1+o(1))^k$. Henceforth we consider a fixed splitting of the indices into three classes.

For the i's of the first class $a_{i+1} = 2a_i$, and thus a_{i+1} is uniquely determined. If i belongs to the second class then from $a_{i+1} \geqslant (1+\delta)^r a_{i+r-1}$ it clearly follows that there are at most $c_1 = c_1(\delta)$ a's in the interval $(\delta a_i, a_i)$. From $a_{i+1} \geqslant (1+\delta)a_i$ it follows that only the a_j's of the interval $(\delta a_i, a_i)$ have to be considered in defining a_{i+1}. Thus there are at most c_1^2 choices for a_{i+1}, and hence for the number of addition chains satisfying (5) the contribution of the i's of the second class it at most $c_1^{2u_2} = (1+o(1))^k$.

The number of possible choices given by the u_3 indices of the third class is less than $\binom{k^2}{u_3}$. To see this observe that the indices $i_1, i_2, \ldots, i_{u_3}$ which belong to the third class have already been fixed and our sequence is completely determined if we fix the indices $j_1^{(i_1)}, j_1'^{(i_1)}; j_2^{(i_2)}, j_2'^{(i_2)}, \ldots, j_{u_3}^{(i_{u_3})}, j_{u_3}'^{(i_{u_3})}$ which define $a_{i_1+1}, a_{i_2+1}, \ldots, a_{i_{u_3}+1}$. Because of $a_{i_1+1} < a_{i_2+1} < \ldots < a_{i_{u_3}+1}$ their order is determined uniquely (this is easy to see by induction). The total number of pairs (u, v), $1 \leqslant u \leqslant v \leqslant k$, equals $\binom{k}{2}+k < k^2$, whence the result.

Thus we have proved that the number of addition chains satisfying (5) is less than

$$\sum_k (1+o(1))^k \sum_{u_3} \binom{k^2}{u_3}, \tag{8}$$

where the summation is extended over all possible choices of k and u_3, satisfying (5). Now we show

$$u_3 < \left(1-\frac{\varepsilon}{2}\right)\frac{\log n}{\log\log n}. \tag{9}$$

To prove (9) observe that if i is in the third class then for some $r_i > 0$

$$a_{i+1} < a_{i+1-r_i}(1+\delta)^{r_i}. \tag{10}$$

The intervals $(i+1-r_i, i+1)$ cover all the i's of the third class. From these intervals we form (in a unique way) a set of non-overlapping intervals (u_s, v_s), $s = 1, 2, \ldots, t$, which contain all the intervals $(i+1-r_i, i+1)$, where i is in the third class.

A simple argument shows by (10) and the construction of the intervals (u_s, v_s) that

$$a_{v_s} \leqslant a_{u_s}(1+\delta)^{2(v_s-u_s)}. \tag{11}$$

The intervals $u_s < x \leqslant v_s$, $1 \leqslant s \leqslant t$ cover all the i's of the third class. Thus

$$\sum_{s=1}^{t}(v_s-u_s) \geqslant u_3. \tag{12}$$

From (5), (11), (12) and $a_{i+1} \leqslant 2a_i$ we infer that

$$\frac{n}{2} \leqslant a_k \leqslant 2^{k-u_3}(1+\delta)^{2u_3} < 2^{k-u_3(1-\varepsilon/2)} \tag{13}$$

for sufficiently small $\delta = \delta(\varepsilon)$. Thus from (13)

$$k-u_3\left(1-\frac{\varepsilon}{2}\right) > \frac{\log n}{\log 2}-1. \tag{14}$$

(14) and (5) clearly implies (9).

From (5), (9) and (8) we infer that the number of addition chains satisfying (5) is less than

$$\big(1+o(1)\big)^{\log n}\binom{A}{B}, \tag{15}$$

where

$$A = \left[\left(\frac{\log n}{\log 2}+(1-\varepsilon)\frac{\log n}{\log\log n}\right)^2\right], \qquad B = \left[\left(1-\frac{\varepsilon}{2}\right)\frac{\log n}{\log\log n}\right].$$

Now

$$\begin{aligned}\binom{A}{B} < \left(\frac{A}{B}\right)^B e^B &= \big(1+o(1)\big)^{\log n}\left(\frac{A}{B}\right)^B \\ &= \big(1+o(1)\big)^{\log n}(\log n)^{B(1+o(1))} = n^{1-\varepsilon/2+o(1)}.\end{aligned} \tag{16}$$

From (15) and (16) we finally infer that the number of addition chains satisfying (5) is less than $n^{1-\varepsilon/2+o(1)} < n^{1-\eta}$ for $\eta < \varepsilon/2$, which completes the proof of our Theorem.

It would be of interest to obtain a more accurate estimation of $l(n)$ and in particular to try to obtain an asymptotic distribution function for $l(n)$, but I have not succeeded in making any progress in this direction.

We can modify the definition of an addition chain as follows: a sequence $1 = a_1 < a_2 < \ldots < a_k = n$ is said to be an *addition chain of* *order* r if each a_j is the sum of r or fewer a_i's where the indices do not exceed j. Denote by $l_r(n)$ the length of the shortest addition chain of order r with $a_k = n$. Using a modification of the method of Brauer and of this note we can prove that for all n

$$l_r(n) < \frac{\log n}{\log r} + \frac{\log n}{(r-1)\log\log n} + o\left(\frac{\log n}{\log\log n}\right),$$

and that for almost all n

$$l_r(n) = \frac{\log n}{\log r} + \frac{\log n}{(r-1)\log\log n} + o\left(\frac{\log n}{\log\log n}\right).$$

Peter Ungár in a letter has asked me the followig question: Define $l'(n)$ as the smallest k for which there exists a sequence $a_0 = 1, a_1, a_2, \ldots, a_k = n$ where for each j, $a_j = a_u \pm a_v$, $u \leqslant j$, $v \leqslant j$ ($a_1 < a_2 < \ldots$ is not assumed here). The problem has arisen in trying to compute x^n with the smallest number of multiplications and divisions. Clearly $l'(n) \leqslant l(n)$ and it can be shown that our Theorem holds for $l'(n)$ too.

Reçu par la Rédaction le 20. 8. 1959