Reprinted from the Publications of the Ramanujan Institute, Number 1, 1969

ON SOME APPLICATIONS OF GRAPH THEORY TO NUMBER THEORETIC PROBLEMS

DEDICATED TO THE MEMORY OF K. ANANDA RAU

P. Erdös

Let $a_1 < \ldots < a_k \leqslant n$ be a sequence of integers no one of which divides any other; then it is easy to see that [1] $\max k = \left[\frac{n+1}{2}\right]$. On the other hand, I proved [2] by a combination of number theoretic and graph theoretic methods that, if we assume $a_i \nmid a_j a_k$ $(i \neq j,\ i \neq k)$, then ($\pi(n)$ denotes the number of primes $\leqslant n$.)

$$\pi(n)+c_1 n^{2/3}/(\log n)^2 < \max k < \pi(n)+c_2 n^{2/3}/(\log n)^2. \qquad (1)$$

Further, if we only assume that the products $a_i a_j$ are all distinct, then [2]

$$\pi(n)+c_3 n^{3/4}/(\log n)^{3/2} < \max k < \pi(n)+c_4 n^{3/4}. \qquad (2)$$

In the present paper we prove that the lower estimation in (2) is sharp (apart from the value of the absolute constant c_3). In fact we prove the following

THEOREM. *Assume that $a_1 < \ldots < a_k \leqslant n$ is a sequence of integers for which the products $a_i a_j$ are all distinct. Then*

$$\pi(n)+c_3 n^{3/4}/(\log n)^{3/2} < \max k < \pi(n)+c_5 n^{3/4}/(\log n)^{3/2}. \qquad (3)$$

The proof of (3) will be similar to that of (2) and will use elementary results from number theory and graph theory. Before we prove our Theorem we would like to discuss a few related results. An old conjecture of Turán and myself states as follows: Let $b_1 < \ldots$ be an infinite sequence of integers. Denote by $f(n)$ the number of solutions of $n = b_i + b_j$. Then, if $f(n) > 0$ for all $n > n_0$, $\limsup_{n\to\infty} f(n) = \infty$. Probably the above conclusion also follows if we only assume that $b_k < c_6 k^2$ for all k. These conjectures are not yet settled and are probably quite deep. It is perhaps surprising that the multiplicative analogies of these conjectures have been settled. In fact I proved the following results [3]: *Let $a_1 < \ldots$ be an infinite sequence. Denote by $g(n)$ the number of solutions of $n = a_i a_j$. Then, if $g(n) > 0$ for all $n > n_0$, we have* $\limsup_{n\to\infty} g(n) = \infty$. In fact the following stronger result holds: *Assume that $a_1 < \ldots < a_k \leqslant n$, n sufficiently large and*

$$k > (1+\epsilon)n(\log\log n)^{l-1}/(l-1)!\log n.$$

Then for some m, $g(m) \geqslant 2^l$. The proof of these results uses combinatorial arguments on generalized graphs and is not quite simple.

Let finally $a_1 < \ldots < a_k \leqslant n$ and assume that all the products $\Pi a_i^{\alpha_i}$ are distinct. Then it is easy to see that $\max k = \pi(n)$. If we only assume that the products $\prod_{i=1}^{k} a_i^{\epsilon_i}$, $\epsilon_i = 0$ or 1 are all distinct, then [4]

$$\pi(n)+c_7 n^{1/2}/\log n < \max k < \pi(n)+c_8 n^{1/2}/\log n.$$

If we assume that all the products $a_{i_1} \ldots a_{i_r}$ (for fixed r) are all distinct, we probably have

$$\pi(n)+c_9\left(\frac{n^{1/2}}{\log n}\right)^{1+\frac{1}{r}} < \max k < \pi(n)+c_{10}\left(\frac{n^{1/2}}{\log n}\right)^{1+\frac{1}{r}}. \qquad (4)$$

Unfortunately I can prove (4) only if $r = 2$. (Then (4) becomes (3).) For $r = 3$ I can prove the right side of (4), in view of the incompleteness of this result I suppress the proof.

Now we prove our Theorem. The lower bound in (3) has already been proved in [2] by Miss E. Klein and myself; thus it suffices to prove the upper bound in (3). The method will be a refinement of the one used in [2]. We need two lemmas.

LEMMA 1. *Every integer $m \leqslant n$ can be written in the form*

$$m = uv, \ v \leqslant u \qquad (5)$$

where u is either a prime or is $\leqslant n^{2/3}$ and $v \leqslant n^{2/3}$.

The lemma is known [2].

LEMMA 2. *Let G be a graph having t_1 vertices $x_1, \ldots, x_{t_1}$ and $C(G)$ edges. Assume that each edge of G is incident to one of the vertices x_i, $1 \leqslant i < t_2 < t_1$, and that G contains no rectangle (i.e. no circuit of four edges). Then*

$$C(G) < t_1+t_1\left[\frac{t_2}{t_1^{1/2}}\right]+t_2^2\left(1+\left[\frac{t_2}{t_1^{1/2}}\right]\right)^{-1}. \qquad (6)$$

Denote by $v(x_j)$ the number of x_i, $1 \leqslant i \leqslant t_2$, joined to x_j. By our assumption we have

$$C(G) \leqslant \sum_{j=1}^{t_1} v(x_j). \qquad (7)$$

Now we split the vertices of G into two classes. In the first class are the vertices for which

$$v(x_j) \leqslant \left[\frac{t_2}{t_1^{1/2}}\right]+1 = l \qquad (8)$$

and in the second class are the vertices with $v(x_j) > l \geqslant 1$. We evidently have by (8)

$$\sum{}' v(x_j) \leqslant t_1\left(1+\left[\frac{t_2}{t_1^{1/2}}\right]\right) \qquad (9)$$

where in Σ' the summation is extended over the x_j of the first class.

Let now $x_{j_1}, \ldots, x_{j_s}$ be the vertices of the second class. It is easy to see that

$$\sum_{r=1}^{s} \binom{v(x_{j_r})}{2} \leqslant \binom{t_2}{2}. \qquad \qquad (10)$$

To prove (10) observe that since G has no rectangle no two x_{j_r} can be joined to the same two x_i's, $1 \leqslant i \leqslant t_2$. One can clearly form $\binom{v(x_{j_r})}{2}$ pairs from the x_i's, $1 \leqslant i \leqslant t_2$, which are joined to x_{j_r} and these $\sum_{r=1}^{s} \binom{v(x_{j_r})}{2}$ pairs are all distinct. Since there are $\binom{t_2}{2}$ pairs formed from the x_i's, $1 \leqslant i \leqslant t_2$, (10) clearly follows.

From (10) and $v(x_{j_r}) > l$ we have

$$\sum_{r=1}^{s} v(x_{j_r}) \leqslant \frac{2}{l}\binom{t_2}{2} < t_2^2\left(1+\left[\frac{t_2}{t_1^{1/2}}\right]\right)^{-1}. \qquad \qquad (11)$$

(6) follows from (9) and (11), hence Lemma 2 is proved.

Now we can prove our Theorem. Let $a_1 < \ldots < a_k \leqslant n$ be a sequence of integers for which the products $a_i a_j$ are all distinct. We can assume that none of the a's are squares since the number of squares $\leqslant n$ is $[n^{1/2}]$ which can be absorbed in the error term in (3). Put

$$a_i = u_i v_i, \;\; v_i < u_i, \qquad \qquad (12)$$

where u_i and v_i satisfy Lemma 1 and v_i is minimal. Now we associate with our sequence $a_1 < \ldots < a_k \leqslant n$ a graph G having $\pi(n)+n^{2/3}-\pi(n^{2/3})$ vertices and k edges. The vertices of our graph are the integers $\leqslant n^{2/3}$ and the primes $\leqslant n$. Each a_i we represent in the form (12) and we make correspond to it the edge joining the vertices u_i and v_i. The fact that the products $a_i a_j$ are all distinct implies that G has no rectangle. For, if the edges $a_1 = u_1 v_1$, $a_2 = u_1 v_2$, $a_3 = u_2 v_2$ and $a_4 = u_2 v_1$ would form a rectangle, we would have

$$a_1 a_3 = a_2 a_4 = u_1 u_2 v_1 v_2.$$

Using Lemma 2 we now estimate k from above. We split the a's into three classes. In the first class are the a's for which $v_i \leqslant n^{1/3}$. In the second class are the a's for which

$$n^{1/3} < v_i \leqslant n^{1/2}/2^{10 \log \log n}$$

and in the third class are the a's for which

$$\frac{n^{1/2}}{2^{10 \log \log n}} < v_i < n^{1/2}.$$

Consider now the subgraph of G corresponding to the a's of the first class. We apply Lemma 2. We have here

$$t_1 = \pi(n) + [n^{2/3}] - \pi(n^{2/3}), \quad t_2 = [n^{1/3}]. \qquad (13)$$

Thus by (6) the number of a's of the first class is less than

$$\pi(n) + 2n^{2/3}. \qquad (14)$$

(14) follows from the fact that by (13) $\left[t_2/t_1^{1/2}\right] = 0$.

The a's of the second class we split into several subclasses. Put $[10 \log \log n] = L$. In the r-th subclass are the a's for which

$$\frac{n^{1/2}}{2^{r+L}} < v_i \leqslant \frac{n^{1/2}}{2^{r+L-1}}. \qquad (15)$$

If a_i is in the r-th subclass, we have from (15)

$$u_i < 2^{r+L} n^{1/2}. \qquad (16)$$

Now we again apply Lemma 2. By (15) and (16) we have

$$t_1 < 2^{r+L} n^{1/2}, \; t_2 < \frac{n^{1/2}}{2^{r+L-1}}. \qquad (17)$$

Hence from (17) and (6) the number of a's of the r-th subclass is less than

$$2^{r+L} n^{1/2} + 4(2^{r+L} n^{1/2})^{1/2} \frac{n^{1/2}}{2^{r+L-1}} < n^{2/3} + 8n^{3/4}/2^{(L+r)/2}. \qquad (18)$$

To prove (18) observe that $t_1 < 2^{L+r} n^{1/2} < n^{2/3}$ since otherwise a_i would have belonged to the first class.

The total number of subclasses of the second class is clearly less than $\log n/\log 2$. Thus from (18) the number of a's of the second class is less than ($L = [10 \log \log n]$)

$$\frac{n^{2/3} \log n}{\log 2} + 8n^{3/4} \sum_{r=0}^{\infty} \frac{1}{2^{(L+r)/2}} = o(n^{3/4}/(\log n)^{3/2}). \qquad (19)$$

To estimate the number of integers of the third class we need

LEMMA 3. *Let $p_1 < \ldots < p_s \leqslant n$ be a sequence of primes. Then the number of integers $m \leqslant n$ which are not divisible by any of the p_i is less than*

$$c_{11} n \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right).$$

Lemma 3 follows immediately by Bruns method and is well known. (See eg. [5].)

Further, we will need the following classical result of Mertens.

$$c_{12}/\log n < \prod_{p<n} \left(1 - \frac{1}{p}\right) < c_{13}/\log n. \qquad (20)$$

The a's of the third class we again divide into subclasses. In the r-th subclass are the a's for which

$$\frac{n^{1/2}}{2^r} \leqslant v_i < \frac{n^{1/2}}{2^{r-1}}. \qquad \qquad (21)$$

Now we use for the first time the minimality property of v_i. v_i cannot have a prime factor $p < n^{1/7}$. For, if $v_i = pv_i'$, $p < n^{1/7}$, then (since by (21) $u_i < 2^r n^{1/2} < 2^{10 \log \log n}\, n^{1/2}$)

$$a_i = v_i' \,.\, pu_i \text{ and } pu_i < n^{2/3} \qquad \qquad (22)$$

or the representation (22) satisfies Lemma 1 which contradicts the minimality property of v_i.

Thus by Lemma 3 and (20) the number of possible choices for the v_i belonging to the a's of the r-th subclass is less than

$$c_{11} \frac{n^{1/2}}{2^{r-1}} \prod_{p < n^{1/7}} \left(1 - \frac{1}{p}\right) < c_{14} \frac{n^{1/2}}{2^r \log n}. \qquad \qquad (23)$$

Let us next estimate the number of possible choices of the u_i which belong to the a's of the r-th subclass. These u_i cannot have a prime factor in the interval $(2^{2r+2}, n^{1/7})$. To see this assume $p \mid u_i$, $2^{2r+2} < p < n^{1/7}$. From (21) and $v_i < u_i$ we have $\frac{n^{1/2}}{2^r} < u_i < 2^{r+1} n^{1/2}$. Put now

$$a_i = \frac{u_i}{p} \,.\, pv_i.$$

From (21) we evidently have

$$pv_i < n^{2/3}, \quad \frac{u_i}{p} < v_i$$

which contradicts the minimality property of v_i.

Hence by Lemma 3 and (20) the number of possible choices of the u_i belonging to the a's of the r-th subclass is less than

$$2^{r+1} n^{1/2} \prod_{2^{2r+2} < p < n^{1/7}} \left(1 - \frac{1}{p}\right) < c_{15} r \, 2^r \frac{n^{1/2}}{\log n}. \qquad \qquad (24)$$

Now we apply Lemma 2. From (24) and (23) we have here

$$t_1 < c_{15} r \, 2^r \frac{n^{1/2}}{\log n}, \quad t_2 < c_{14} \frac{n^{1/2}}{2^r \log n},$$

thus from Lemma 2 we have that the number of a's of the r-th subclass is less than

$$c_{15} r \, 2^r n^{1/2} / \log n + c_{16} \frac{n^{3/4}}{(\log n)^{3/2}} \frac{r^{1/2}}{2^{r/2}}. \qquad \qquad (25)$$

From (25) we obtain that the number of a's of the third class is less than

$$c_{17} n^{3/4} / (\log n)^{3/2}. \qquad \qquad (26)$$

(14), (19) and (26) imply the upper bound in (3) and hence the proof of our Theorem is complete.

REFERENCES

1. Problem 3739, *Amer. Math. Monthly* 44 (1937), 120.
2. P. Erdös, 'On sequences of integers no one of which divides the product of two others and on some related problems', *Tomsk. Gos. Univ. Učen Zap.* 2 (1938), 74–82.
3. ———, 'On the multiplicative representation of integers', *Israel J. Math.* 2 (1964), 251–261.
4. ———, 'Extremal problems in number theory II', *Mat. Lapok* 17 (1966), 135–155 [Hungarian].
5. H. Halberstam and K. F. Roth, *Sequences* (Oxford, 1966).

UNIVERSITY OF READING
READING
ENGLAND

(*Received 7 July 1967*)