Sh:21

ON PROBLEMS OF MOSER AND HANSON

P. Erdős
Imperial College
London, England

and

S. Shelah
Hebrew University
Jeruselem, Israel

The following problem is due to L. Moser: Let $A_1, \dots, A_n$ be any n sets. Take the largest subfamily $A_{i_1}, \dots, A_{i_r}$ which is <u>union-free</u>, i.e.,

$$A_{i_{j_1}} \cup A_{i_{j_2}} \neq A_{i_{j_3}}, \quad 1 \le j_1 \le r, \quad 1 \le j_2 \le r, \quad 1 \le j_3 \le r,$$

for every triple of distinct sets $A_{j_1}, A_{j_2}, A_{j_3}$. Put $f(n) = \min r$, where the minimum is taken over all families of n distinct sets. Determine or estimate $f(n)$. Riddel showed $f(n) > c\sqrt{n}$ and Erdős and Komlós [1] showed

$$\sqrt{n} \le f(n) \le 2\sqrt{2}\sqrt{n}. \tag{1}$$

We now show

$$\sqrt{2n} - 1 < f(n) < 2\sqrt{n} + 1 \tag{2}$$

and we conjecture that $f(n) = (2 + o(1))\sqrt{n}$.

Consider now the largest subfamily $A_{i_1}, \dots, A_{i_r}$ so that no four distinct sets satisfy

$$A_{i_{j_1}} \cup A_{i_{j_2}} = A_{i_{j_3}}, \quad A_{i_{j_1}} \cap A_{i_{j_2}} = A_{i_{j_4}}. \tag{3}$$

Put $F(n) = \min r$, where the minimum is taken over all families of distinct sets $A_1, \dots, A_n$. We prove

$$F(n) \le \frac{3}{2} n^{2/3}. \tag{4}$$

Probably $F(n) > c_2 n^{2/3}$, and in fact it seems likely that $F(n)/n^{2/3}$ tends to a limit, but we have not been able to show this.

Hanson posed the following problem: Let $|\mathcal{Q}| = n$, with $g(n)$ the smallest integer so that the subsets of $\mathcal{Q}$ can be split into $g(n)$ classes where each of the classes is union free. Hanson proved

$$c_3 \sqrt{n} < g(n) \le \frac{n}{2} + 2 \tag{5}$$

and he conjectured that the upper bound is substantially correct. We prove

$$g(n) > \frac{n}{4}. \tag{6}$$

Let $G(n)$ be the smallest integer so that the subsets of $\mathcal{Q}$ can be split into $G(n)$ classes so that no class contains four distinct sets A_1, A_2, A_3, A_4 satisfying (3). We prove

$$c_4\sqrt{n} < G(n) < c_5\sqrt{n}. \tag{7}$$

Probably $\lim_{n\to\infty} G(n)/n^{1/2}$ exists.

Now we prove (2). We use a slight improvement of the method of [1] to prove the upper bound. Let t be the least integer for which $[t^2/4] > n$. Our A's are the $[t^2/4]$ set of integers $A_{i,j} = \{x: i \le x \le j\}$, $1 \le i \le t/2 < j \le t$. We show that the largest union-free subfamily of the A's has at most t elements. To see this let A_{i_r,j_r}, $1 \le r \le \ell$, be a union-free subfamily of the A's. An endpoint i_r (or j_r) is called good if there is no other A_{i_s,j_s} of our family with $i_r = i_s$ and $j_s < j_r$ (or $j_r = j_s$, $i_s > i_r$). Clearly at least one endpoint of A_{i_r,j_r} must be good, for otherwise A_{i_r,j_r} would be the union of two A's of our family. But an integer can be a good endpoint of at most one A_{i_r,j_r}, which shows $\ell \le t$ and our assertion is proved. Now clearly

$$f(n) \le f([t^2/4]) \le t$$

Sh:21

or $f(n) \leq 2\sqrt{n}+1$. which proves the upper bound of (2).

We now prove the lower bound. Let $\{A_1,\dots,A_n\}$ be any family of n distinct sets. We define a union-free subfamily $\{A_{i_1},\dots,A_{i_r}\}$ as follows. A_{i_1} is any minimal A, i.e., contains no other as a proper subset. Suppose $A_{i_1},\dots,A_{i_s}$ have already been defined. Then $A_{i_{s+1}}$ is chosen to be a minimal member of $\{A_1,\dots,A_n\}\setminus\{A_{i_1},\dots,A_{i_s}\}$ which is not the union of two distinct members of $\{A_{i_1},\dots,A_{i_s}\}$. There is clearly a choice for $A_{i_{s+1}}$ if $n-s > \binom{s}{2}$. This process therefore defines a subfamily $\{A_{i_1},\dots,A_{i_r}\}$ of r sets, where $r+\binom{r}{2} \geq n$, i.e., $r \geq \sqrt{2n}-1$. To complete the proof it only remains to show that the family $\{A_{i_1},\dots,A_{i_r}\}$ is union-free.

Assume

$$A_{i_j} \cup A_{i_k} = A_{i_\ell} \quad (j \neq \ell,\ k \neq \ell). \tag{8}$$

We cannot have $\ell > j$ and $\ell > k$ by the construction. So we can assume $k > \ell$. But this is also impossible, since A_{i_ℓ} was chosen ad a minimal member of $\{A_1,\dots,A_n\}\setminus\{A_{i_1},\dots,A_{i_{\ell-1}}\}$. Hence (8) cannot hold and the proof of (2) is complete.

It is not difficult to improve the lower bound of (2) slightly to show that $f(n) > (1+c)\sqrt{2n}$. However, we cannot show that $f(n) = (2+o(1))\sqrt{n}$.

To prove (4) we use an idea due to Folkman. Let t be the least integer for which $t^3 \geq n$. Consider the t^3 sets $A_{i,j} = \{n: i \leq n \leq j\}$, $(1 \leq i \leq t < j \leq t^2+t)$. Thus the sets $A_{i,j}$ correspond to the edges of the complete bipartite graph $K(t,t^2)$. A simple argument shows that every subgraph of $K(t,t^2)$ having $t^2 + \binom{t}{2} + 1$ edges contains a rectangle; this is false for $t^2+\binom{t}{2}$ edges. A rectangle corresponds to four distinct sets A satisfying (3). Thus

$$F(n) \leq F(t^3) \leq t^2 + \binom{t}{2},$$

which proves (4).

Instead of (3) we could consider other systems of equations with sets as unknowns, but in view of the fact that we did not succeed in getting a satisfactory lower bound of $F(n)$ we do not investigate this question at present.

To prove (6) consider again the family of $[n^2/4]$ sets $A_{r,s}$ used in the proof of (2).

We already showed that the largest union-free subfamily of our sets has n elements. Thus

$$g(n) \ge [\tfrac{n^2}{4}]/n \ge \tfrac{n}{4}.$$

Hanson suggested that a more careful analysis of this family would in fact give $g(n) \ge n/3$.

Now finally we prove (7). Consider again the sets $A_{r,s}$ used in the proof of (2). As stated before the number of these sets is $[n^2/4]$ and the sets $A_{r,s}$ correspond to a complete bipartite graph of n vertices with $[n/2]$ white and $[(n+1)/2]$ black vertices. If a subfamily $\{A_{r_i,s_i}\}$ is such that no four distinct elements of it satisfy (3), then as stated the corresponding bipartite graph (we join the vertices r_i and s_i) has no rectangle. By a theorem of Reiman [2] such a graph can have at most

$$(1 + o(1))n^{3/2}/2\sqrt{2}$$

edges; thus we immediately obtain

$$G(n) > (1 + o(1))n^{1/2}/2^{1/2}.$$

Now we prove the upper bound of (7). Let q be the smallest prime power for which $q^2 + q + 1 \ge n$. By a well known result of Singer [3], there are $q + 1$ residues mod (q^2+q+1), $a_1, \ldots, a_{q+1}$, so that all non zero residue classes have a unique representation in the form $a_j - a_i$.

Now we split the subsets of a set Q of $q^2 + q + 1 \ge n$ elements into $q + 1$ classes so that the sets of none of the classes contain four sets satisfying (3). To see this put in the i-th class $(1 < i < q + 1)$ all sets having $\overline{a_j - a_i}$ $(1 \le j \le q + 1,\ i \ne j)$ elements, where $\overline{a_j - a_i}$ is the least positive integer $\equiv (a_j - a_i) \pmod{q^2+q+1}$.

Sh:21

If four distinct sets A_1, A_2, A_3, A_4 of the i-th class satisfy (3) we would have

$$|A_1| + |A_2| = |A_3| + |A_4|$$

or

$$a_{j_1} - a_i + a_{j_2} - a_i = (a_{j_3} - a_i + a_{j_4} - a_i) \pmod{q^2+q+1}.$$

Hence, $a_{j_1} - a_{j_3} = (a_{j_4} - a_{j_2}) \pmod{q^2+q+1}$, which is impossible.

Thus

$$G(n) \le G(q^2+q+1) \le q + 1 \le \sqrt{n} + 1,$$

which completes the proof of (7).

References

1. P. Erdös and J. Komlós, On a problem of Moser, Combinatorial Theory and its Applications I, (Edited by P. Erdös, A. Rényi, and V. T. Sós), North Holland Publishing Company, Amsterdam, 1969.

2. I. Reiman, Über ein Problem von K. Zarankiewicz, Acta Math. Acad. Sci. Hungarica 9 (1958), 269-273.

3. J. Singer, A theorem in finite projective geometry and some applications to number theory, Trans. Amer. Math. Soc. 43 (1938), 377-385.