On a combinatorial problem

N. G. DE BRUIJN and P. ERDÖS: *On a combinatioral problem.*

(Communicated at the meeting of November 27, 1948.)

Let there be given n elements $a_1, a_2, \ldots, a_n$. By $A_1, A_2, \ldots, A_m$ we shall denote combinations of the a's. We assume that we have given a system of $m > 1$ combinations $A_1, A_2, \ldots, A_m$ so that each pair (a_i, a_j) is contained in one and only one A. Then we prove

Theorem 1. We have $m \geq n$ [1], with equality occurring only if either the system is of the type $A_1 = (a_1, a_2, \ldots, a_{n-1})$, $A_2 = (a_1, a_n)$, $A_3 = (a_2, a_n) \ldots A_n = (a_{n-1}, a_n)$, or if n is of the form $n = k(k-1)+1$ and all the A's have k elements, and each a occurs in exactly k of the A's.

Corollary: If the elements a_i are points in the real projective plane the theorem can be stated as follows: Let there be given n points in the plane, not all on a line. Connect any two of these points. Then the number of lines in this system is $\geq n$. In this case equality occurs only if $n-1$ of the points are on a line.

This corollary can be proved independently of Theorem 1 by aid of the following theorem of GALLAI (= GRÜNWALD) [2]:

Let there be given n points in the plane, not all on a line. Then there exists a line which goes through two and only two of the points.

Remark: The points of inflexion of the cubic show that it is essential that the points should all be real, thus GALLAI's theorem permits no projective and a fortiori no combinatorial formulation. Also the result clearly fails for infinitely many points.

We now give GALLAI's ingenious proof: Assume the theorem false. Then any line through two of the points also goes through a third. Project one of the points, say a_1 to infinity, and connect it with the other points. Thus we get a set of parallel lines each containing two or more points a_i (in the finite part of the plane). Consider the system of lines connecting any two of these points, and assume that the line $(a_i\, a_j\, a_k)$ forms the smallest angle with the parallel lines. (This line again contains at least three points). But the line connecting a_j with a_1 (at infinity) contains at least another (finite) point a_r, and clearly (see figure) either the line $(a_i\, a_r)$

---

[1]) This was also proved by G. SZEKERES but his proof was more complicated.

[2]) This theorem was first conjectured by SYLVESTER, GALLAI's proof appeared in the Amer. Math. Montly as a solution to a problem by P. ERDÖS. The corollary to Theorem 1 also appeared as a problem in the Monthly.

See also H. S. M. COXETER, Amer. Math. Monthly 55, 26—28 (1948), where very simple proofs due to KELLY and STEINBERG are given.

or the line (a_r, a_k) forms a smaller angle with the parallel lines then $(a_i\, a_j\, a_k)$. This contradiction establishes the result.

Remark: Denote by $f(n)$ the minimum number of lines which go through exactly two points. It is not known whether $\lim f(n) = \infty$. All that we can show is that $f(n) \geq 3$.

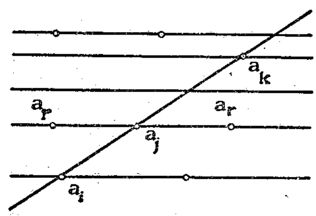

Fig. 1.

Now we prove the corollary as follows: We use induction. Assume that the number of lines determined by $n-1$ points, not all on a line, is $\geq n-1$. Then we shall prove that n points, not all on a line, determine at least n lines.

Let (a_1, a_2) be a line going through two points only. Consider the points $a_2, a_3, \ldots, a_n$. If they are all on a line, then (a_1, a_i), $i = 2, 3, \ldots, n$ and $(a_2, a_3, \ldots, a_n)$ clearly determine n lines. If they do not all lie on a line, then they determine at least $n-1$ lines, and (a_1, a_2) is clearly not one of these lines. Thus together with (a_1, a_2) we again get at least n lines. The same induction argument shows that we get exactly n lines only if $n-1$ of the points lie on a line, q.e.d.

Proof of theorem 1. For simplicity we shall call the elements $a_1, a_2, \ldots, a_n$ points and the sets $A_1, A_2, \ldots, A_m$ lines. Denote by k_i the number of lines passing through the point a_i, and by s_j the number of points on the line A_j. We evidently find (by counting the number of incidences in two ways)

$$\sum_{j=1}^{m} s_j = \sum_{i=1}^{n} k_i \quad \ldots\ldots\ldots\ldots \quad (1)$$

Further if A_j does not pass through a_i, then

$$s_j \leqq k_i \quad \ldots\ldots\ldots\ldots \quad (2)$$

(2) follows from the fact that a_i can be connected by a line (i.e. an A) to all the s_j points of A_j, and any two of these lines are different, since otherwise they would have two points in common.

Assume now that k_n is the smallest k_i and that $A_1, A_2, \ldots, A_{k_n}$ are the lines through a_n. We may suppose that each line contains at least two points, since otherwise it could be omitted. Also $k_n > 1$, for otherwise all the points are on a line. Thus we can find points a_i on A_i, $a_i \neq a_n$, $i = 1, 2, \ldots, k_n$. Also if $i \neq j$, $i \leq k_n$, $j \leq k_n$ then a_i is not on A_j (for otherwise A_i and A_j would have two points in common). Hence by (2) (putting $k_n = \nu$)

$$s_2 \leqq k_1, \quad s_3 \leqq k_2, \ldots, \quad s_\nu \leqq k_{\nu-1}, \quad s_1 \leqq k_\nu; \quad s_j \leqq k_n \text{ for } j > \nu. \quad (3)$$

From (1), (3) and the minimum property of k_n we obtain $m \geq n$, which proves the first part of Theorem 1.

We now determine the cases where $m = n$. If $m = n$, then all the inequalities of (3) have to be equalities. Consequently we can renumerate the points so that $s_1 = k_1$, $s_2 = k_2$, ..., $s_n = k_n$. We may suppose that $k_1 \geq k_2 \geq \ldots \geq k_n > 1$. There are two cases:

a) $k_1 > k_2$. Hence by $s_1 = k_1 > k_i \; (2 \leq i \leq n)$, (2) shows that all the a_i $(i \geq 2)$ lie on A_1. Of course a_1 does not lie on A_1 and we have the first case of Theorem 1.

b) $k_1 = k_2$. If no k_i is less than k_1 then clearly $k_i = s_j$ $(1 \leq i, j \leq n)$. We shall show that this is the only possibility. If $k_j < k_1$, then we have by (2) that a_j lies on both A_1 and A_2. Hence k_n is the only k which can be less than k_1. Now $s_n = k_n$ different lines contain a_n. Any line through a_n contains one further point and all but one contain two further points, since $k_1 = k_2 = \ldots = k_{n-1} > k_n \geq 2$. Thus there are at least two lines which do not contain a_n; for both of these lines we have by (2) $s_j \leq k_n$. This contradicts $s_1 = s_2 = \ldots = s_{n-1} > k_n$.

Apart from case a) we only have the case where $s_i = k_j = k$, $(1 \leq i, j \leq n)$. It is easily seen that then $n = k(k-1) + 1$, and also that any pair of lines has exactly one intersection point. For if A_i does not intersect A_j; and if a_l lies on A_i then we infer from (2) that $k_l \geq s_j + 1$ which is not possible since $k_l = s_j = k$. The two dimensional ~~projective~~ finite geometries with $k - 1 = p^\alpha$, p prime, are known to be systems of this type, but F. W. Levi [3]) constructed a non-~~projective~~ example with $k = 9$.

[3]) F. W. Levi, Finite geometrical systems, Calcutta 1942.

desarguian