Problems and results on block designs and set systems

Paul Erdős

Let $|S|=n$. A family F of subsets $A_i \subset S$, $1 \le k \le m$, is called an r -design if every r -tuple $\{x_1 ,\dots, x_r\}$ of S is contained in one and only one of the A 's. If r=2 we will just call these systems designs. (In the literature these 2 -designs are usually called partially balanced block designs.) Usually we will restrict ourselves to designs; i.e., to r=2. There are two rivial designs. The design with m=1, $|A_1|=n$ is completely trivial and will henceforth be ignored. Almost as trivial is the r -design with $m=1+\binom{n-1}{r-1}$ $|A_1|=n-1$, $|A_i|=r$, $2 \le i \le 1+\binom{n-1}{r-1}$. For r=2 this trivial design is sometimes called the near pencil. Henceforth, it will also be ignored.

An old theorem of de Bruijn and myself [1] states that for every design we have m=n with equality only for the finite geometries (we ignored the near pencil), $n=t^2+t+1$, $|A_i|=t+1$, $1 \le i \le n$.

Our theorem is a generalization of the older Fischer's inequality, which assumed that all the blocks A_i have the same size.

There are several simple proofs, both combinatorial and algebraic of our theorem. As far as I know there is no purely combinatorial proof of the following generalization due to Ryser [2]: Assume that every pair $\{x,y\}$ of S is contained in precisely λ of the sets A_k . Then $m \ge n$ again holds. It would be nice to find a purely combinatorial proof and to characterize those values of λ and n for which equality is possible.

As far as I know not much is known about extending the theorem of de Bruijn and myself for $r>2$. Let $m(n;r)$ be the smallest integer for which there is an r -design for $|S|=n$. $m(n;2)=n$ is the theorem of

CONGRESSUS NUMERANTIUM, Vol. 35 (1982), pp. 3-16

de Bruijn and myself. There are some results for $r=3$ due to Hanani and others - in particular, it follows from results of Hanani that

$$c_1 n^{3/2} < m(n;3) < c_2 n^{3/2} \quad .$$

Presumably we have

$$m(n;r) = (c_r + o(1))n^{r/2} \ ,$$

but as far as I know nothing much is known about $m(n;r)$ for $r\geq 4$.

Let us now assume $r=2$. First we discuss $m=n$. Unfortunately I can contribute nothing new here. It is well known that $m=n$ is possible only if $n=t^2+t+1$ and $|A_i|=t+1$, $1\leq i\leq n$ (we exclude the trivial cases $m=1$ and the near pencil). Further it is well known that such block designs - finite geometries - are possible if $t=p^\alpha$, p prime. It is one of the fundamental problems of combinatorial mathematics if such a finite geometry is possible if t is not a power of a prime. $t=10$ is the smallest value for which the existence of a finite geometry is open;i.e.,let $|S|=111$. Is there a system $|A_k|=11$, $1\leq k\leq 111$ of subsets of S so that every pair $\{x,y\}$ of S is contained in one and only one A_k ? Unfortunately I have nothing to contribute to this beautiful question. A classical result of Bruck and Ryser states that if $t\equiv 1$ or 2 (mod 4) then a finite geometry of t^2+t+1 elements can only exist if t is the sum of two squares. Suppose now that $n=t^2+t+1$. There are three possibilities:

I A finite geometry exists if and only if $t=p^\alpha$.

II A finite geometry always exists if the theorem of Bruck-Ryser does not exclude it.

III The "Truth" is somewhere in between.

Almost no progress has been made towards the resolution of this old problem.

Let $|A_i|=x_i$, and suppose the sets A_i , $(1\le i\le m)$ form a design. Then we must have

$$(1) \qquad \sum_{i=1}^{m} \binom{x_i}{2} = \binom{n}{2} ,$$

but of course (1) is not sufficient for the existence of a block design. It is perhaps not reasonable to expect to obtain a necessary and sufficient condition for the sequence $x_1\ge\ldots\ge x_m$ that there should be a block design on $|S|=n$ satisfying $|A_i|=x_i$. Denote by $F(n)$ the number of possible choices for $\{x_1 ,\ldots, x_m\}$ where both m and the x_i are variable. I conjectured several years ago that $(\exp z= e^z)$

$$(2) \qquad \exp(c_1\, n^{1/2} \log n) < F(n) < \exp(c_2\, n^{1/2} \log n) .$$

The upper bound in (2) is easy (it follows almost immediately from the fact that the number of indices k for which $|A_k|>10\, n^{1/2}$ is less than $n^{1/2}$), but I was not able to prove the lower bound. No doubt

$$(3) \qquad \log F(n)/_{n^{1/2} \log n} \to C$$

for a certain non-zero finite positive C . Perhaps the proof of (3) will not be difficult, but it will require more practice in constructing designs than I ever had.

Denote by $f(n)$ the number of those block designs which can be obtained geometrically; i.e., the elements are points in the plane and the blocks the lines joining them. It is easy to see that $f(n) > \exp(c_3\, n^{1/2})$ and I conjectured many years ago that

$$(4) \qquad f(n) < \exp(c_4\, n^{1/2}) .$$

The proof of (4) (if it is correct) may be difficult (Szemerėdi and Trotter just proved (4).) Probably

$$(5) \qquad \log f(n)/_{n^{1/2}} \to c$$

holds $(0<c<\infty)$. It is unlikely that a simple and illuminating exact formula can be found for $f(n)$ and $F(n)$.

A few days ago (June 1982), I asked R. Wilson the following question: Let F be a family of k-element subsets of a set $|S|=n$. Assume that $k^2>n$ and for any two distinct A_i, A_j in F we have $|A_i \cap A_j| \le 1$. Determine or estimate $\max|F|$.

Wilson almost immediately proved

$$(6) \qquad |F| \le \frac{n(k-1)}{k^2-n} .$$

He observed that (6) implies that if $n=u^2+u+1$, $k=u+t$, $t>1$, then

$$(7) \qquad |F| \le \frac{1}{2t-1} n + \frac{1}{2} n^{1/2} .$$

(7) shows that if $t \ge 2$ then for every n

$$(8) \qquad |F| \le \frac{n}{3} + o(n^{1/2}) .$$

Observe that (8) shows that $\max F$ does drop a great deal from its possible maximum $|F|=n$ $u=p^\alpha$, $t=1$ (i.e., when there is a finite geometry).

Bollobás and I tried to get lower bounds for $|F|$ if $t=1$, $n=u^2+u+1$. We could not prove that in this case

$$(9) \qquad \lim \max |F|/_n = 1 .$$

In fact we doubted that (9) is true. Thus it is not yet clear whether Wilson's inequalities (7) and (8) can be significantly improved.

Wilson's results also imply that if $k=[c\, n^{1/2}]$, $c>1$ then

(10) $$\max|F| < \frac{2}{c^2-1} n^{1/2} .$$

I asked this question of Wilson many years ago and he told me (10) long ago.

He further observed that if we only assume $|A_i \cap A_j| \le \lambda$ then

(11) $$|F| < \frac{n(k-\lambda)}{k^2-n\lambda} .$$

Wilson never published any of these results since he tells me that all this is more or less well known (he refers to Johnson and Selmer "A new bound for error correcting codes" IEEE around 1965). In any case I now give Wilson's proofs with his permission.

The proof is really surprisingly simple. Let $|F|=b$, and denote by v_x the number of A's containing the element x of S . We evidently have

$$\sum_{x \in S} v_x = bk , \quad \sum_{x \in S} v_x(v_x-1) \le b(b-1) .$$

Thus

$$\sum_{x \in S} v_x^2 \le b(b-k+1) .$$

Hence

$$nb(b+k-1) \ge n \sum_{x \in S} v_x^2 \ge (\sum_{x \in S} v_x)^2 = b^2k^2$$

which proves (6). Unfortunately for $k \le \sqrt{n}$ this method only gives trivial results.

Let $r<k<n$, $|S|=n$. $f(n;k,r)$ is the largest integer for which there is a family F , $|F|=f(n;k,r)$, of k -element subsets of S for which $|A_i \cap A_j|<r$ for every A_i , A_j in F . Similarly $F(n;k,r)$ is the smallest integer for which there is a family F , $|F|=F(n;k,r)$ of

k -element subsets of S for which every r -tuple of S is contained in at least one A_i of F . Trivially

$$f(n;k,r) \leq \binom{n}{k} \Big/ \binom{k}{r} \leq F(n;k,r) . \tag{11}$$

One of the fundamental problems of combinatorial analysis is to determine the values of $f(n;k,r)$ and $F(n;k,r)$ as accurately as possible and in particular to determine the cases of equality in (11).

Unfortunately I have nothing to contribute towards the solution of this beautiful problem and want just to make a few historical remarks. The case of the so called Steiner triples ($k=3$, $r=2$) is completely solved. There is equality in (11) if and only if $n \equiv 1$ or 3 (mod 6) . Steiner triples should really be called Kirkman triples since Kirkman completely solved the problem of their existence more than 10 years before Steiner posed the problem. I wonder if Steiner could have seen Kirkman's paper? I hope the reader will forgive me if I seem to doubt the honesty of a great mathematician. I quote from the very interesting book of Felix Klein Development of Mathematics in the 19-th century: When Steiner became old and his productivity suffered "fallen out with God and Himself" he claimed that he found results which he in fact read mainly from English authors."

After the case $k=3$, $r=2$ was settled almost nothing happened for nearly a century. Then Hanani settled the cases $k=4$, $r=3$; $k=4$, $r=2$; $k=5$, $r=2$.

In a series of brilliant papers Wilson nearly completely settled $r=2$. He proved that if $n>n_0(k,2)$ there is equality in (11) if and only if n satisfies certain congruence conditions. If $r>2$ there are no general results other than $r=3$, $k=4$ settled by Hanani.

Very recently Rödl proved that for every k and r

$$(12) \qquad f(n;k,r) = (1+ o(1)) \binom{n}{k} \Big/ \binom{k}{r} \quad .$$

Special cases of (12) were proved earlier by Rényi, Hanani and myself.

Jean Larson and I investigated the following problem: Let $A(n)$ be the largest integer for which there is a design on the set $|S|=n$, each block of which has size $\geq A(n)$. Observe that if $n=p^{2\alpha}+p^{\alpha}+1$ then there is a finite geometry and we have $A(n)=p^{\alpha}+1$; i.e., $A(n)=[n^{1/2}]+1$ for infinitely many n. We proved $A(n) > n^{1/2}-n^{1/4+c}$ where $c>0$ is a small but positive constant. If we make plausible but hopeless assumptions on the distribution of primes we obtain $A(n)>n^{1/2} - (\log n)^{c}$. We have a fascinating unsolved problem: Is it true that there is an absolute constant C so that for all n

$$(13) \qquad A(n) > n^{1/2} - C \quad ?$$

We thought that (14) is more likely false. I offer 250 dollars for the proof or disproof of (13). Our paper has just appeared in the volume dedicated to Mendelsohn.

The following somewhat vague conjecture should be true: Let $|S|=n$, $\{A_i\}$ $1\leq i\leq m$ the blocks of a design. Assume that $m>n$ but that $m-n$ is "small". Then there is a finite geometry on $|S_1|\geq m$, $|S_1|-m$ small elements so that we obtain our design from the finite geometry by omitting $|S_1|-n$ suitably chosen elements.

Perhaps it would be worthwhile to try to extend our work with Jean Larson for r-designs, perhaps this is not hopeless.

In a forthcoming paper V. T. Sós, R. Wilson and I investigated the following question: Let $|S|=q^2+q+1$ where we assume that q is such that there is a finite geometry in S (i.e. a design with $|B_i|=q+1$,

$1 \le i \le q^2+q+1$). Let now $\{A_i\}$, $1 \le i \le m$ be a design for which $m > q^2+q+1$. Then $m \ge q^2+2q+1$. In other words: no design can have m blocks for $q^2+q+1 < m < q^2+2q+1$. We further show that unless our design is obtained by breaking up a block of our finite geometry we in fact have $m > q^2+(2+c)q$, where c is a positive absolute constant. Probably c can be chosen to be 1 but at present we can not prove this. Answering a question of Doyen we determine with fairly good accuracy the set of possible values of m . Some of these questions could be investigated for r -designs but as far as I know nothing is known.

In a paper which will soon appear in Discrete Mathematics, R. C. Mullin, V. T. Sós, D. R. Stinson and I investigated the following related problem.

Let $|S|=v$ and $b=B(v)$ be the smallest integer for which there is a non-trivial design on v elements and b blocks. Such a design will be called minimal. We prove that for $v \ge 5$ we have

$$(14) \qquad B(v) \ge \begin{cases} n^2+n+1 & \text{if } n^2+2 \le v \le n^2+n+1 \\ n^2+n & \text{if } n^2-n+3 \le v \le n^2+1 \\ n^2+n-1 & \text{if } v = n^2-n+1 . \end{cases}$$

Equality holds if there is a projective plane on n^2+n+1 elements.

We also consider the embeddability of minimal designs into projective planes and prove several re ults. E.g., if $v=n^2-\alpha$, $\alpha > 0$ and $\alpha^2+\alpha(2n-3)-(2n^2-2n) \le 0$ then a minimal ıesign on v elements can be imbedded into a projective plane of n^2+n+1 elements. Further if $v=$ $=n^2-n+2$ $(v>8)$ and $B(v)=n^2-n-1$ then the design can be embedded.

Assume that $n^2-n+2 \le v \le n^2+n+1$ and that the minimal design on v elements can not be embedded into the projective plane on n^2+n+1 elements.

We then prove

$$(15) \qquad B(v) \geq v + \frac{3-\sqrt{5}}{2} n .$$

The constant in (15) is no doubt very far from being best possible. Many unsolved problems remain; e.g., is it true that

$$(16) \qquad \limsup_{v\to\infty} \frac{B(v) - v}{v^{1/2}} = \infty \quad ?$$

(16) seems to us a very interesting and intriguing problem. It is related to (13).

To end the paper I state a few disconnected problems and results. V. T. Sós and I proved that if $|S|=v$ then in a design on S there always is an element $x\varepsilon S$ which is not contained in at least $v-\sqrt{v}-1$ blocks. Equality holds if and only if the design is a finite geometry.

A geometric version of this result is stated as follows. Let there be given n points in the plane not all on a line. Then there always is a point x_0 so that there are at least $n-2$ lines which are not incident to x_0. It is easy to see that this is best possible.

We further considered the following two problems. Let $v=n^2+n+1$ and assume that there is a finite geometry on $|S|=v$. Let $F(v)$ be the smallest integer so that if $F(n)$ subsets A_h , $1\leq h\leq F(n)$, $|A_h|=n+1$ of S are given then there always is a subfamily $A_{i_1},\ldots,A_{i_t}$, $t=n^2+n+1$ which forms a finite geometry. It is easy to see that for some $0<c<1$

$$F(n) > c\binom{v}{n+1}$$

and perhaps in fact

$$F(n) = (1+o(1))\binom{v}{n+1} .$$

Denote by $H(n)$ the smallest integer so that every $G^{(3)}(n;H(n))$ contains a Fano plane. $G^{(3)}(n;t)$ is a three uniform hypergraph of n vertices and t edges (i.e. triples), the Fano plane is the finite geometry of 7 points and 7 lines. We conjectured that

$$H(2n) = n^2(n-1) + 1 ,$$

and the extremal graph is the well known triple system of Turán. $S=S_1 \cup S_2$, $|S_1|=|S_2|=n$, $S_1 \cap S_2=\emptyset$. The edges are the triples which have a non-empty intersection with both S_1 and S_2 . Clearly this triple system contains no Fano plane since it is two-chromatic and the Fano plane is known to have chromatic number three. It is well known and easy to see that all finite geometries except the Fano plane have chromatic number two.

A very recent paper of Wallis led me to the following question: Let $G(n;t_n)$ be a graph of n vertices and t_n edges whose complementary graph can be covered by $n-1$ or less edge disjoint complete graphs of size not exceeding $n-2$. By the theorem of de Bruijn and myself $t_n \geq 1$. Determine or estimate the smallest possible value of t_n . Perhaps $t_n \geq cn$.

Let F be a 2 -design on p^2+p+1 elements. Assume that F contains two blocks A_1 , A_2 satisfying $|A_1|=|A_2|=p+1$, $A_1 \cap A_2 = \emptyset$. Determine or estimate $\min|F|$. Our Lemma C gives $|F|>p^2+3p$ this probably can be improved and and a non-trivial upper bound can be obtained.

Magliveras conjectured that if there is a finite geometry on n^2+n+1 elements and $n>n_0$ then one can partition the $\binom{n^2+n+1}{n+1}$ subsets of size $n+1$ of $|S|=n^2+n+1$ into $\binom{n^2+n+1}{n+1}(n^2+n+1)^{-1}$ finite geometries. My first idea was to try to disprove this, but I got nowhere with this plan, and at the moment I can not attack this attractive conjecture.

Let $|S|=m$ and let $F_1, \ldots F_{t_m}$ be a family of designs on the set S, no two of which have a block in common and which are essentially different; i.e., there is no permutation of the elements of S which carries F_i into F_j. Denote by $h_1(m) = \max t_m$.

Let further be $F'_1, \ldots F'_{t'_m}$ be another family of designs no two of which have a common block. Denote by $x_1^{(i)} \geq x_2^{(i)} \geq \ldots$, the sizes of the blocks of F'_i and assume that the sets $\{x_1^{(i)}, \ldots\}$ and $\{x_1^{(j)}, \ldots\}$ are not identical. This condition clearly implies that there is no permutation on the elements of S which carries F'_i into F'_j. Put $\max t'_m = h_2(m)$. Clearly $h_1(m) \geq h_2(m)$. Estimate $h_1(m)$ and $h_2(m)$ as well as possible from above and below. I am not even sure if they are of polynomial growth. Though perhaps here I overlook a trivial point.

Several of my old problems in combinatorial geometry have recently been settled by Szemerêdi-Trotter and Beck. Szemerêdi and Trotter proved that if there are n points in the plane and k lines where $n^{1/2} < k \leq \binom{n}{2}$ then the number of incidences between the points and the lines is less than $c\, n^{1/3}\, k^{2/3}$. This remarkable inequality is essentially best possible and they proved (4) by using this inequality. Beck (independently) proved a weaker inequality from which he deduced another conjecture of mine: Let there be given n points in the plane no $n-k$ on a line. Then they determine at least $c\,k\,n$ lines. The value which Beck's proof gives for the i^c is no doubt very far from being best possible. Perhaps if only $o(n)$ points are on a line then the points determine at least $(1+o(1))\,\frac{n^2}{6}$ lines. It follows from a result of Sylvester that if true this is best possible.

I conjectured that any set of n points determines at least $cn/\sqrt{\log n}$ distinct distances and the lattice points show that if true this is best possible. In fact I thought that perhaps one of our n points has the property that there are at least $c\,n/\sqrt{\log n}$ distinct distances from it. 30 years ago L. Moser proved these conjectures with $n^{2/3}$ instead of $cn/\sqrt{\log n}$. A few months ago F. Chung proved that n points determine at least $c\,n^{5/7}$ distinct distances and very recently Beck proved that the number of distinct distances from one of the points is $> n^{2/3+\varepsilon}$ where $\varepsilon > 0$ is small but is independent of n. Perhaps one of the points, say x_i, has the property that every circle of center x_i contains $o(n^{\varepsilon})$ of the other x_i's.

Very recently V. T. Sôs, Trotter and I considered the following problems: Let $\{A_1, A_2, \ldots, A_m\}$ be a design on n elements. $|A_1| \geq |A_2| \geq \ldots$. Denote by

$$(17) \qquad F(n;k) = \max \sum_{i=1}^{k} |A_i|$$

where in (17) the maximum is extended over all possible designs of n elements and

$$(18) \qquad f(n;k) = \max \sum_{i=1}^{k} |L_i|$$

where in (18) the maximum is extended over all designs which can be obtained from n points in the plane. Clearly $F(n;k) \geq f(n;k)$ and first of all it would be of interest to determine the smallest $k=k(n)$ for which $f(n;k) < F(n;k)$. It is easy to see that

$$(19) \qquad F(n;k) < \max\ (c_1 n\ k^{1/2}\ ,\ c_2\ n^{1/2}\ k\ ).$$

(19) can probably be replaced by an asymptotic formula, but we have not yet done this. Several further related questions can be asked but we leave this to the interested reader.

Finally we thought that perhaps the following problems might be of interest (we only thought of these problems recently, thus we apologize to the reader if the problems turn out to be trivial). Let $\{A_i\}$ be a design on the elements $x_1, \ldots, x_n$ and denote by v_i the number of A_i 's containing x_i. Let $|v_1| \geq \ldots \geq |v_n|$. Denote by $G(n)$ the number of possible choices of $\{v_1, \ldots, v_n\}$. Denote by $g(n)$ the number of possible choices of $\{v_1, \ldots, v_n\}$ which where the x_i are points in the plane. How do G_n and $g(n)$ compare to $F(n)$ and $f(n)$? (See (3) and (4).) If the sizes of the $\{A_i\}$ are given can we have many choices of the $\{v_i\}$? Is there any hope of getting a reasonable condition for $v_1 \geq \ldots \geq v_n$ so that there should be a design for which v_i blocks contain x_i, $1 \leq i \leq n$?

References

[1] N. G. de Bruijn and P. Erdős, On a combinatorial problem, Nederl. Akad. Wetensch. Proc 51 (1948), 1277-1279.

[2] H. J. Ryser, On extension of a theorem of de Bruijn and Erdős on combinatorial designs, J. Algebra 10, 246-261.

[3] R. M. Wilson, An existence theory for pairwise balanced designs III: Proof of the existence conjectures. J. C. T. (A) 18 (1975), 71-79. This paper has extensive references.