

Families in Which Disjoint Sets Have Large Union

N. ALON^a AND P. FRANKL^b

^a Department of Mathematics
Tel Aviv University
69978 Tel Aviv, Israel
and

Bell Communications Research
Morristown, New Jersey 07960

^b AT&T Bell Laboratories
Murray Hill, New Jersey 07971

INTRODUCTION AND STATEMENT OF THE RESULTS

Let $X = \{1, 2, \dots, n\}$ and F be a family of subsets of X , that is $F \subset 2^X$. For $1 \leq i \leq j \leq n$ set $[i, j] = \{i, \dots, j\}$. For integers k, m with $k \geq 2, 0 \leq m \leq n$, we say that F has property $P(k, m)$ if any k pairwise disjoint members of F have union of size greater than m . Thus $P(k, n)$ means simply that F contains no k pairwise disjoint sets.

Let us write m in the form $m = kt - r$, where $1 \leq r \leq k$. Define

$$F(n, k, m) = \{F \subseteq X : |F| + |F \cap [1, r-1]| \geq t\}.$$

It is easy to check that $F(n, m, k)$ has property $P(k, n)$. In fact, if F_1, \dots, F_k are pairwise disjoint members of F , then

$$|F_1 \cup \dots \cup F_k| = |F_1| + \dots + |F_k| \geq kt - \sum_{1 \leq i \leq k} |F_i \cap [1, r-1]| \geq kt - (r-1)$$

holds.

Note that for $m = kt - 1$ one has simply $F(n, k, m) = \{F \subseteq X : |F| \geq t\}$.

THEOREM 1: Suppose $F \subset 2^X$, F has $P(k, m)$. Then $|F| \leq |F(n, k, m)|$ holds in each of the following cases.

- (a) $m = kt - 1$.
- (b) $k = 2, m = 2t - 2$,
- (c) k, r arbitrary, $n > 2m^3$. Moreover, $|F| = |F(n, k, m)|$ is possible only if F is isomorphic to $F(n, k, m)$.

Let us mention that the condition $n > n_0(m)$ cannot be completely removed in (c). In fact, Kleitman [8] proved that for $n = m = kt - k$ the maximum size of a family having $P(n, k, n)$ is attained by $F = \{F \subseteq X : |F \cap \{1, 2, \dots, n-1\}| \geq t-1\}$.

Let us also note that if (c) holds for some triple (n, k, m) , then it also holds for all (n', k, m) with $n' > n$ —this will be clear from the inductive proof of (a) and (b).

The following old conjecture of Erdős is related to our problem.

CONJECTURE 1 [4]: Suppose $G \subset \binom{X}{t}$, $|X| \geq rt$, and G contain no r pairwise disjoint sets. Then

$$|G| \leq \max \left\{ \binom{n}{t} - \binom{n-r+1}{t}, \binom{rt-1}{t} \right\}. \tag{1}$$

The case $t = 2$ of the preceding conjecture is a theorem of Erdős and Gallai [5]. Erdős [4] proved that for $n > n_0(r, t)$ (1) holds; moreover, if $|G|$ is maximal, then for some R , $|R| = r - 1$ one has

$$G = \left\{ G \in \binom{X}{t} : G \cap R \neq \emptyset \right\}.$$

The case $r = 2$ is covered by the Erdős-Ko-Rado theorem (see the next section). In the case $n = rt$ the inequality

$$|G| \leq \frac{r-1}{r} \binom{rt}{t} = \binom{rt-1}{t}$$

is easy to prove.

For the proof of (c) we need a strengthening of (1), which was obtained by Bollobás *et al.* [3]. First let us define the family

$$E_t(n, r) = \left\{ E \in \binom{X}{t} : E \cap [1, r-2] \neq \emptyset \right\} \\ \cup \left\{ E \in \binom{X}{t} : (r-1) \in E, E \cap [r, r+t-1] \neq \emptyset \right\} \cup \{[r, r+t-1]\}.$$

It is not hard to check that $E_t(n, r)$ contains no r pairwise disjoint members,

$$|E_t(n, r)| = \binom{n}{t} - \binom{n-r+1}{t} - \binom{n-r-t+1}{t-1} + 1.$$

THEOREM 2 [3]: Suppose $F \subset \binom{X}{t}$, F contains no r pairwise disjoint members, $|F| > |E_t(n, r)|$ and $n > 2t^3(r-1)$, then for some $R \subset \binom{X}{r-1}$ one has $F \cap R \neq \emptyset$ for all $F \in F$.

Let us call a family F k -times dense on Y if for all $Y_0 \subset Y$ there exist $F_1, F_2, \dots, F_k \in F$, so that $F_i \cap Y = Y_0$ for $1 \leq i \leq k$, $F_1 - Y, \dots, F_k - Y$ partition $X - Y$. For $0 \leq s \leq n$ let $d(n, k, s)$ denote the maximum size of F subject to the assumption that there is no s -element set Y on which F is k -times dense.

Also, set $f(n, k, m) = \max\{|F| : F \subset 2^X, F \text{ has } P(k, m)\}$.

THEOREM 3:

$$d(n, k, s) = f(n, k, n - s).$$

Remember that 2-times dense families were used by Alon and Milman [2] in connection with embedding problems of Banach spaces. They proved the case $k = 2$ of our theorems. Actually Theorem 3 will be easily derived using the compression techniques of [1] and [7].

SHIFTING AND SOME RELATED RESULTS

One of the basic results in extremal set theory is the following.

ERDÖS-KO-RADO THEOREM [6]: Suppose $F \subset \binom{X}{l}$, $n \geq 2l$ and $F \cap F' \neq \emptyset$ holds for all $F, F' \in F$. Then

$$|F| \leq \binom{n-1}{l-1}.$$

For the proof of this result Erdős, Ko, and Rado introduced an important operation; the (i, j) -shift S_{ij}

$$S_{ij}(F) = \{S_{ij}(F) : F \in F\},$$

where

$$S_{ij}(F) = \begin{cases} (F - \{j\}) \cup \{i\} & ; i \notin F, \quad j \in F \text{ and } ((F - \{j\}) \cup \{i\}) \notin F \\ F, & \text{otherwise.} \end{cases}$$

Note that the (i, j) -shift just replaces element j by i in those sets that contain j but not i and for which the new set was not already in the family.

The importance of this operation lies in the following.

PROPOSITION 1: Suppose F has property $P(k, m)$. Then $|S_{ij}(F)| = |F|$ and $S_{ij}(F)$ has $P(k, m)$, too.

Proof: Take pairwise disjoint sets $A_1, A_2, \dots, A_k \in S_{ij}(F)$ and suppose for contradiction $|A_1 \cup \dots \cup A_k| \leq m$. Let B_l be the inverse image of A_l , that is $S_{ij}(B_l) = A_l$. Since F has $P(k, m)$, we may assume $B_1 \neq A_1$, and consequently $i \in A_1, j \notin A_1, i \notin B_1, j \in B_1$. Since the A_l are pairwise disjoint, $i \notin A_l$ for $l \geq 2$. Then $j \in B_2, i \notin B_2$ and $S_{ij}(B_2) = A_2$. Why? The only possibility is that $B_2^* = (B_2 - \{j\}) \cup \{i\}$ is in F . However, $B_1, B_2^*, B_3, \dots, B_k$ are pairwise disjoint sets in F with

$$|B_1 \cup B_2^* \cup B_3 \cup \dots \cup B_k| = |A_1 \cup \dots \cup A_k| \leq m,$$

a contradiction. \square

During his recent visit to Japan, Erdős suggested that the following might be true.

THEOREM 4: Suppose $n \geq (r + 1)k - 1$, $\binom{X}{r} = F_1 \cup F_2$. Then either F_1 or F_2 contains k pairwise disjoint sets.

Proof: Suppose for contradiction neither F_1 nor F_2 contains k pairwise disjoint members. We may assume $F_2 = \binom{X}{r} - F_1$. Applying the (i, j) -shift to F_1 means to apply the (j, i) -shift to $\binom{X}{r} - F_1$. Thus, by Proposition 1, we may apply the (i, j) -shift repeatedly to F_1 for all $1 \leq i < j \leq n$. Since $\sum_{F \in F_1} \sum_{i \in F} i$ is not increasing during this process and is strictly decreasing with each nontrivial shift, after finite time we get a family G_1 (and thus $G_2 = \binom{X}{r} - G_1$), which is *shifted*, that is, $S_{ij}(G_1) = G_1$ holds for all $1 \leq i < j \leq n$. Moreover, neither G_1 nor G_2 contains k pairwise disjoint sets. However, the set $A_0 = \{k, 2k, \dots, k\}$ must be either in G_1 or in G_2 . Consider now the possible cases:

- (a) $A_0 \in G_1$. Since G_1 is shifted $A_i = \{k - i, \dots, rk - i\} \in G_1$ follows for $i = 0, 1, \dots, k - 1$ (because $S_{lk-i, lk}(\mathbf{G}) = \mathbf{G}$ for all $\mathbf{G} \in G_1, l = 1, \dots, r$). However, A_0, A_1, \dots, A_{k-1} are pairwise disjoint, a contradiction.
- (b) $A_0 \in G_2$. Since G_2 is shifted $A_i = \{k + i, 2k + i, \dots, rk + i\} \in G_2$ for $i = 0, 1, \dots, k - 1$. However, A_0, \dots, A_{k-1} are pairwise disjoint, a contradiction. \square

REMARK 1: Note that for $n = (r + 1)k - 2$, neither

$$F_1 = \binom{\{1, \dots, kr - 1\}}{r} \quad \text{nor} \quad F_2 = \binom{X}{r} - F_1$$

contain k pairwise disjoint sets. Thus Theorem 4 is the best possible.

PROOF OF THEOREM 1(A) AND (B)

We apply induction on n . The case $m = n$ was proved by Kleitman [8]. Thus, assume $m < n$. Suppose F has $P(k, m)$. In view of Proposition 1, just as in the proof of Theorem 4, we may suppose that F is shifted, that is, $S_{ij}(F) = F$ holds for all $1 \leq i < j \leq n$. Define

$$F(n) = \{F \subset \{1, 2, \dots, n - 1\} : (F \cup \{n\}) \in F\},$$

$$F(\bar{n}) = \{F \subset \{1, 2, \dots, n - 1\} : F \in F\}.$$

Clearly $|F| = |F(n)| + |F(\bar{n})|$. Suppose $m = kt - 1$ or $k = 2$ and $m = 2t - 2$. Note that if $F = F(n, k, m), n > m \geq k$, then $F(n) = F(n - 1, k, n - k), F(\bar{n}) = F(n - 1, k, m)$ hold. Thus the statement will follow from the induction hypothesis as soon as we show $F(\bar{n})$ has $P(k, m)$ and $F(n)$ has $P(k, m - k)$. The first is obvious. To prove the second, suppose for contradiction A_1, \dots, A_k are pairwise disjoint sets in $F(n)$ with $|A_1 \cup \dots \cup A_k| \leq m - k$. Since $n > m$, we can find elements i_1, \dots, i_k such that $(A_1 \cup \dots \cup A_k) \cap \{i_1, \dots, i_k\} = \emptyset$. Since $S_l(F) = F$ for $l = 1, \dots, k, B_l = (A_l \cup \{i_l\}) \in F$ follows. However, B_1, \dots, B_k are pairwise disjoint and $|B_1 \cup \dots \cup B_k| = |A_1 \cup \dots \cup A_k| + k \leq m$, a contradiction.

To prove the uniqueness of the extremal families apply induction again. From the proof we know $|F(n)| = |F(n - 1, k, m - k)|, n - 1 > m - k$; thus, $F(n) = F(n - 1,$

$k, m - k$. Since F is shifted, $(F \cup \{j\}) \in F$ for all $F \in F(n)$ and $1 \leq j \leq n$. This gives $F \supseteq F(n, k, m)$.

Thus we proved that for $n > m$, $|F| = |F(n, k, m)|$ implies $F = F(n, k, m)$ if F is shifted.

To conclude the proof of the uniqueness we must show that if G has $P(k, m)$, and for some $1 \leq i < j \leq n$ one has $S_{ij}(G) = F(n, k, m)$, then G is isomorphic to $F(n, k, m)$.

As for all

$$\mathbf{G} \in G, |S_{ij}(\mathbf{G})| = |\mathbf{G}|, \binom{X}{l} \subset G \text{ follows for } t \leq l \leq n.$$

This concludes the proof for the case (a). In the case (b) we have to deal with $G' = \{\mathbf{G} \in G : |\mathbf{G}| = t - 1\}$.

Again $S_{ij}(G) = F(n, k, m)$ implies

$$|G'| = \binom{n-1}{t-2}.$$

As G has $P(2, 2t - 2)$, G' contains no two disjoint sets. That is, G' is an extremal family for the Erdős-Ko-Rado theorem ($l = t - 1, n > 2l$). Consequently, for some $x \in X$ one has

$$G' = \left\{ \mathbf{G} \in \binom{X}{t-1} : x \in \mathbf{G} \right\},$$

concluding the proof. \square

REMARK 2: Actually, the same proof would work word for word in case (c) as well, except that the starting case ($m = n$) of the induction is missing.

REMARK 3: We outline here an alternative proof of Theorem 1 for the case $m = kt - 1$, which does not use shifting. Suppose F has property $P(k, m)$. If $A \in F$ is of size $j \leq m$, then A is contained in $\binom{n-j}{m-j}$ m -subsets of X . It is easy to check that if $|F| > |F(n, k, t)|$ and $m = kt - 1$, this implies that there is an m -subset of X containing more than $|F(m, k, t)|$ members of F . The result now follows from the starting case of the induction: $n = m$.

PROOF OF THEOREM 1(C)

We suppose again that F is shifted, $|F|$ is maximal, and F has property $P(k, m)$. Apply induction on m . Suppose $r < k$.

CLAIM 1. F has $P(r, rt - r)$.

Suppose for contradiction $A_1, \dots, A_r \in F, A_i \cap A_j = \emptyset$ and $|A_1 \cup \dots \cup A_r| \leq rt - r$. Using shiftedness and the maximality of $|F|$, we may assume $A_1 \cup \dots \cup A_r = [1, rt - r]$. Define $F^* = \{F \in F : F \cap [1, rt - r] = \emptyset\}$. Then F^* has property

$P(k-r, (k-r)(t+1) - (k-r))$. By the induction hypothesis

$$\begin{aligned} |F^*| &\leq |F(n - (rt-r), (k-r), (k-r)(t+1) - (k-r))| \\ &< 2^{n-rt+r} - \binom{n-rt+r-(k-r)}{t}. \end{aligned}$$

Consequently,

$$|F| < 2^n - \binom{n-r(t-2)-k}{t} < 2^n - \binom{n}{t-1} < |F(n, k, m)|,$$

a contradiction for, for example, $n > 2mt$.

Thus F has $P(r, rt-r)$ and by the induction assumption $|F| \leq |F(n, r, rt-r)| = |F(n, k, kt-r)|$ follows, together with the uniqueness of the extremal configurations.

Finally we have to consider the case $r = k$, that is $m = kt - k$.

CLAIM 2: F has $P(n, k-j, (k-j)(t-1) - j)$ for all $1 \leq j < k$.

Proof: Suppose for contradiction $A_1, \dots, A_{k-j} \in F$ are pairwise disjoint and A is a set of size $(k-j)(t-1) - j$ containing $A_1 \cup \dots \cup A_{k-j}$. Define

$$F^* = \{F \in F: F \cap A = \emptyset\}.$$

Then F^* has $P(n - |A|, j, j(t+1) - j)$, and this leads to a contradiction in the same way as in the case of Claim 1. \square

Let us define

$$F^{(i)} = \{F \in F: |F| = i\}, f^{(i)} = |F^{(i)}|.$$

In view of Claim 2 there are no $k-1$ pairwise disjoint members in $F^{(i)}$ for $i < t-1$. This yields

$$f^{(i)} < (k-2) \binom{n-1}{i-1}.$$

And $|F| \geq |F(n, k, kt-k)|$ implies

$$\sum_{i < t} f^{(i)} \geq \binom{n}{t-1} - \binom{n-k+1}{t-1}.$$

These two inequalities lead to

$$f^{(t-1)} > \left(k - \frac{3}{2}\right) \binom{n-1}{t-1}$$

for, for example, $n > 2k(t-1)^3$. Since $F^{(t-1)}$ contains no k pairwise disjoint sets, from Theorem 2 and $n > 2m^3 > 2k(t-1)^3$, it follows that there exists $T \subset X$, $|T| = k-1$, so that

$$F^{(t-1)} \subset \left\{ F \in \binom{X}{t-1}: F \cap T \neq \emptyset \right\}.$$

Consequently, for each $x \in T$ there are at least $\frac{1}{2} \binom{n-1}{t-1}$ sets $F \in F^{(t-1)}$ with $F \cap T = \{x\}$. In particular, there exist kt sets F_x^1, \dots, F_x^{kt} , so that $F_x^i \cap T = \{x\}$ and $F_x^i \cap F_x^j = \{x\}$ for $1 \leq i \neq j \leq kt$.

CLAIM 3: For all $G \in F^{(t-i)}$ one has $|G \cap T| \geq i$.

Proof: The statement holds voidly for $i \leq 0$, and we just proved it for $i = 1$. For $i \geq k$ Claim 2 implies $F^{(t-i)} = \emptyset$. Thus we may assume $2 \leq i < k$. Suppose for contradiction $G \in F^{(t-i)}, |G \cap T| < i$. Let x_1, \dots, x_{k-i} be distinct elements of $T - G$. We want to find successively sets F_1, \dots, F_{k-i} so that G, F_1, \dots, F_{k-i} are pairwise disjoint, $F_j \in F^{(t-1)}, F_j \cap T = \{x_j\}, j = 1, \dots, k-i$.

Suppose F_1, \dots, F_{j-1} were already chosen, $j \leq k-i$. Then $|G \cup F_1 \cup \dots \cup F_{j-1}| < jt$; therefore, we can choose one out of the kt sets $F_{x_j}^1, \dots, F_{x_j}^{kt}$ so that it is disjoint from $G \cup F_1 \cup \dots \cup F_{j-1}$.

However, $|G \cup F_1 \cup \dots \cup F_{k-i}| = (t-i) + (k-i)(t-1) = (k-i+1)t - k$, contradicting Claim 2.

Now the proof is finished because, by maximality, we must have

$$F = \{F \subseteq X : |F \cap T| + |F| \geq t\}. \quad \square$$

A REDUCTION LEMMA FOR k -TIMES DENSE FAMILIES

For $F \subset 2^X$ and $i \in X$, let us define the following shifting-type operation C_i :

$$C_i(F) = \{C_i(F) : F \in F\}$$

where

$$C_i(F) = \begin{cases} F \cup \{i\}, & \text{if } i \in F, (F \cup \{i\}) \notin F \\ F, & \text{otherwise.} \end{cases}$$

LEMMA 1: Suppose F is a family that is not k -times dense on any s -element subset of X . Then $C_i(F)$ has the same property as well.

Proof: Suppose for contradiction that $C_i(F)$ is k -times dense on $S \in \binom{X}{s}$. Let T be an arbitrary subset of S . We want to show that there exist $F_1(T), \dots, F_k(T) \in F$ so that $F_j(T) \cap S = T$ and the sets $F_j(T) - S, j = 1, \dots, k$ partition $X - S$.

Suppose first $i \notin S$ and let $G_1(T), \dots, G_k(T) \in C_i(F)$ satisfy the preceding assumptions. If $G_j(T) \in F$ for $j = 1, \dots, k$, then we have nothing to prove. Suppose $G_1(T) \notin F$. Then $i \in G_1(T), F_1(T) = G_1(T) - \{i\}$ is in F . Consider $G_2(T) \in C_i(F)$. How could it happen that $i \notin G_2(T)$? The only explanation is that $F_2(T) = G_2(T) \cup \{i\}$ is also in F . Now choosing $F_j(T) = G_j(T)$ for the remaining values $j = 3, \dots, k$ we are done.

Suppose next $i \in S$ and set $\tilde{T} = T - \{i\}$.

As $C_i(F)$ is k -times dense on S , there exist $G_j(\tilde{T}) \in C_i(F), j = 1, \dots, k$, with $G_j(\tilde{T}) \cap S = \tilde{T}$ and the sets $G_j(\tilde{T}) - S$ forming a partition of $X - S$.

Since $i \notin \bar{T}$, we infer that both $G_j(\bar{T})$ and $G_j(\bar{T}) \cup \{i\}$ are in F for $j = 1, \dots, k$. This completes the proof of the lemma. \square

Proof of Theorem 3: Suppose $F \subset 2^X$, F is not k -times dense on any $S \in \binom{X}{s}$, $|F| = d(n, k, s)$. Repeatedly applying the operation C_i , for $i = 1, \dots, n$, to F , leads to a family G that is not k -times dense on any $S \in \binom{X}{s}$ either and that satisfies $C_i(G) = G$, that is, G is a *monotone* family ($G \in G, G \subset H \subseteq X$, imply $H \in G$). We claim that G has $P(n, k, n - s)$. Suppose the contrary, that is, there exist pairwise disjoint sets $G_1, \dots, G_k \in G$ with $|G_1 \cup \dots \cup G_k| \leq n - s$. Let S be an arbitrary s -element subset of $X - (G_1 \cup \dots \cup G_k)$.

Since G is monotone, for every $T \subseteq S$, the k -sets $G_1 \cup T, G_2 \cup T, \dots, G_{k-1} \cup T$ and $(X - (G_1 \cup \dots \cup G_{k-1})) \cup T$ are in G , showing that G is k -times dense on S .

As $|G| = |F|$, $|F| = d(n, k, s) \leq f(n, k, n - s)$ follows. The opposite inequality is trivial; if F has $P(n, k, n - s)$, $S \in \binom{X}{s}$, then consider $T = \emptyset$ to show that F is not k -times dense on S . \square

REFERENCES

1. ALON, N. 1983. On the density of sets of vectors. *Discrete Math.* **46**: 199–202.
2. ALON, N. & V. D. MILMAN. 1983. Embedding of ℓ_∞^k in finite dimensional Banach spaces. *Israel J. Math.* **45**: 265–280.
3. BOLLOBÁS, B., D. E. DAYKIN & P. ERDŐS. 1976. Sets of independent edges of a hypergraph. *Quart. J. Math. Oxford* **21**: 25–32.
4. ERDŐS, P. 1965. A problem of independent r -tuples. *Ann. Univ. Budapest* **8**: 93–95.
5. ERDŐS, P. & T. GALLAI. 1959. On the maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hung.* **10**: 337–357.
6. ERDŐS, P., C. KO & R. RADO. 1961. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford* **12**: 313–320.
7. FRANKL, P. 1983. On the trace of finite sets. *J. Comb. Theory A* **34**: 41–45.
8. KLEITMAN, D. J. 1968. Maximal number of subsets of a finite set no k of which are pairwise disjoint. *J. Comb. Theory A* **5**: 157–163.