Improved bounds for Erdős' Matching Conjecture

Peter Frankl

Rényi Institute of Mathematics, Budapest, Hungary

Abstract

The main result is the following. Let $F$ be a family of $k$-subsets of an $n$-set, containing no $s+1$ pairwise disjoint edges. Then for $n \geq (2s+1)k - s$ one has $|F| \leq \binom{n}{k} - \binom{n-s}{k}$. This upper bound is the best possible and confirms a conjecture of Erdős dating back to 1965. The proof is surprisingly compact. It applies a generalization of Katona’s Intersection Shadow Theorem.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction and notation

Let $[n] := \{1, 2, \ldots, n\}$ and let $F \subset \binom{[n]}{k}$, $n \geq k \geq 1$. The matching number $\nu(F)$ is the maximum number of pairwise disjoint members (edges) of $F$. One of the classical problems of extremal set theory is to determine $\max |F|$, for $\nu(F)$ fixed. Here are two easy constructions.

$$A(k, s) := \binom{k(s+1) - 1}{k}, \quad |A(k, s)| = \binom{k(s+1) - 1}{k},$$

$$A(n, 1, s) := \left\{ A \in \binom{[n]}{k} : A \cap [s] \neq \emptyset \right\}, \quad |A(n, 1, s)| = \binom{n}{k} - \binom{n-s}{k}.$$ 

The Matching Conjecture. (See Erdős [4] (1965).) If $F \subset \binom{[n]}{k}$, $\nu(F) = s$ and $n$ is at least $k(s+1) - 1$ then

$$|F| \leq \max \left\{ \binom{k(s+1)-1}{k}, \binom{n}{k} - \binom{n-s}{k} \right\}$$

holds.

E-mail address: peter.frankl@gmail.com.
The case \( s = 1 \) is the classical Erdős–Ko–Rado Theorem [6]. For \( k = 1 \) the conjecture holds trivially and for \( k = 2 \) it was proved by Erdős and Gallai [5]. Erdős [4] proved (1) for \( n > n_0(k, s) \). In [3] the bound on \( n_0(k, s) \) was lowered to \( 2sk^3 \). Recently, Huang, Loh and Sudakov [12] improved it to \( 3sk^2 \), which was slightly improved in [9]. On the other hand Füredi and the author proved \( n_0(k, s) \leq cks^2 \), however their result was never published. The aim of the present paper is to provide a completely new argument proving a bound simultaneously improving all known bounds.

**Theorem 1.1.** Let \( \mathcal{F} \subset \binom{[n]}{k} \), \( \nu(\mathcal{F}) = s \) and \( n \geq (2s + 1)k - s \) then

\[
|\mathcal{F}| \leq \left( \binom{n}{k} \right) - \left( \frac{n - s}{k} \right)
\]

with equality if and only if \( \mathcal{F} \) is isomorphic to \( \mathcal{A}(n, 1, s) \).

One of the principal tools in proving (2) is an extension of Katona’s Intersection Shadow Theorem [13]. For a family \( \mathcal{F} \subset \binom{[n]}{k} \) let us define its shadow \( \partial \mathcal{F} \) by

\[
\partial \mathcal{F} := \left\{ G \in \binom{[n]}{k-1} : \exists F \in \mathcal{F}, G \subset F \right\}.
\]

**Theorem 1.2.** Let \( \mathcal{F} \subset \binom{[n]}{k} \), \( \nu(\mathcal{F}) = s \), then

\[
s|\partial \mathcal{F}| \geq |\mathcal{F}|
\]

holds.

Let us note that for \( s = 1 \) the inequality (3) is a special case of Katona’s Intersection Theorem. The proof of Theorem 1.2 is by double induction on \( n \) and \( k \)—just imitating the original proof of Katona [13]. The starting case is \( \mathcal{A}(k, s) \), that is all \( k \)-subsets of an \( n \)-set where \( n = k(s+1) - 1 \). For \( \mathcal{A}(k, s) \) one has \( \partial \mathcal{A}(k, s) = \binom{\binom{k(s+1)-1}{k-1}}{k-1} \) and \( s \binom{k(s+1)-1}{k-1} = k(s+1)-1 \) showing that the factor \( s \) is the best possible. On the other hand it follows from the proof that (3) is strict unless \( \mathcal{F} \) is isomorphic to \( \mathcal{A}(k, s) \).

It is well known (cf. for example [7]) that in proving both theorems one can assume that \( \mathcal{F} \) is stable. That is, for all \( 1 \leq i < j \leq n \) and \( F \in \mathcal{F} \), the conditions \( i \notin F \), \( j \in F \) imply that \( F \cup \{i\} - \{j\} \) is in \( \mathcal{F} \) as well. The only other ingredient of the proof is the following version of the König–Hall Theorem.

**König–Hall Theorem.** (Cf. [14].) Let \( G \) be a bipartite graph with \( \nu(G) = s \). Then there exists a subset \( T \) of the vertices with \( |T| = s \), such that all edges of \( G \) are incident to at least one vertex of \( T \).

2. **Proof of Theorem 1.2**

Assume that \( \mathcal{F} \subset \binom{[n]}{k} \) is a stable family with \( \nu(\mathcal{F}) \leq s \). Let us first prove the statement for all \( k \) and \( s \) with \( (s + 1)k - 1 \geq n \). Let us construct a bipartite graph with partite sets \( \mathcal{F} \) and \( \partial \mathcal{F} \) where we put an edge connecting \( F \) and \( G \) if and only if \( G \in \partial \mathcal{F} \). It is immediate that each \( F \in \mathcal{F} \) has degree \( k \), and each \( G \in \partial \mathcal{F} \) has degree at most \( n - |G| = n - k - 1 \). Since \( sk \geq n - k + 1 \) for \( n \leq (s + 1)k - 1 \), (3) holds in the above range. Moreover, equality can hold only if \( n = (s + 1)k - 1 \) and each \( G \in \partial \mathcal{F} \) has degree \( ks \), so \( G \cup \{y\} \in \mathcal{F} \) for \( y \notin G \in \partial \mathcal{F} \). It follows that \( G - \{x\} + \{y\} \) also should be a member of \( \partial \mathcal{F} \) (for \( x \in G \), \( y \notin G \)) so \( \partial \mathcal{F} \) is the complete \((k-1)\)-uniform hypergraph on \([(s + 1)k - 1]\) and \( \mathcal{F} = \binom{[(s + 1)k - 1]}{k} \) follows.

From now on, we suppose that \( n \geq (s + 1)k \), \( k \geq 2 \) and (3) holds for \( n - 1 \) for both \( k \) and \( k - 1 \). Let us use the usual notation \( \mathcal{F}(\bar{n}) := \{ F \in \mathcal{F} : n \notin F \}, \mathcal{F}(n) := \{ F - \{n\} : F \in \mathcal{F}, n \in F \} \). These are the two families for which we want to use the induction hypothesis. Here \( \nu(\mathcal{F}(\bar{n})) \leq s \) is obvious. The inequality \( \nu(\mathcal{F}(n)) \leq s \) follows from stability using the following standard argument (cf. [7]). If one
has $s + 1$ disjoint sets $F_i - \{n\} \in \mathcal{F}(n)$ (where $F_i \in \mathcal{F}$, $1 \leq i \leq s + 1$), then $n - 1 \geq (s + 1)(k - 1) + s$ implies that there are elements $1 \leq x_1 < \cdots < x_k \leq n - 1$ disjoint to each $F_i$. Then stability implies that the sets $F_i - \{n\} \cup \{x_i\} \in \mathcal{F}$ (here $1 \leq i \leq s$) together with $F_{s+1}$ form a matching of size $s + 1$ in $\mathcal{F}$, a contradiction.

Note that $\partial \mathcal{F}(n)$ provides us with sets in $\partial \mathcal{F}$ which do not contain $n$. At the same time, adjoining $n$ to any member of $\partial \mathcal{F}(n)$ provides us with a member of $\partial \mathcal{F}$ which contains $n$. This proves $|\partial \mathcal{F}| \geq |\partial \mathcal{F}(n)| + |\partial \mathcal{F}(n)|$. Using the induction hypothesis yields

\[ s|\partial \mathcal{F}| \geq s|\partial \mathcal{F}(n)| + s|\partial \mathcal{F}(n)| \geq |\mathcal{F}(n)| + |\mathcal{F}(n)| = |\mathcal{F}| \]

as desired. □

3. A general inequality

The families $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1}$ are called nested if $\mathcal{F}_{s+1} \subset \mathcal{F}_s \subset \cdots \subset \mathcal{F}_1$ holds. The families $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1}$ are called cross-dependent if there is no choice of $F_i \in \mathcal{F}_i$ such that $F_1, \ldots, F_{s+1}$ are pairwise disjoint.

**Theorem 3.1.** Let $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1} \subset \binom{Y}{s}$, be nested, cross-dependent families, $|Y| \geq t \ell$. Suppose further $t \geq 2s + 1$, then

\[ |\mathcal{F}_1| + |\mathcal{F}_2| + \cdots + |\mathcal{F}_s| + (s + 1)|\mathcal{F}_{s+1}| \leq s \left( \frac{|Y|}{\ell} \right). \]

**Proof.** Let us choose randomly (according to uniform distribution) $t$ pairwise disjoint sets $B_1, \ldots, B_t \in \binom{Y}{s}$ and define $B = \{B_1, \ldots, B_t\}$. Since the probability $p(B_j \in \mathcal{F}_i) = |\mathcal{F}_i|/\binom{|Y|}{s}$, the expected size $M(|B \cap \mathcal{F}_i|)$ is $t|\mathcal{F}_i|/\binom{|Y|}{s}$. Let us prove a lemma.

**Lemma 3.2.** For every choice of $B$ one has

\[ |B \cap \mathcal{F}_1| + \cdots + |B \cap \mathcal{F}_s| + (s + 1)|B \cap \mathcal{F}_{s+1}| \leq st. \]

**Proof.** Define a bipartite graph $\mathcal{G}$ with partite sets $B$ and $\{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1}\}$ where we join $B_j$ and $\mathcal{F}_i$ by an edge if and only if $B_j \in \mathcal{F}_i$. The fact that $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1}$ are cross-dependent translates to $v(\mathcal{G}) \leq s$. Applying the König–Hall Theorem we can find a subset $T$ of the vertices, $|T| = s$ such that all edges are incident to some element of $T$.

Let $T$ have $x$ elements in $B$ and $s - x$ elements in $\{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1}\}$. Let us estimate the total number of edges incident to $T$. For $\mathcal{F}_i$ there can be at most $t$ incident edges. This gives an upper bound $(s - x)t$ for the $s - x$ vertices from $\{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1}\}$. The $x$ vertices in $B$ can be adjacent to $(s + 1) - (s - x) = x + 1$ additional vertices each. This gives the upper bound

\[ (s - x)t + x(x + 1) = x^2 - (t - 1)x + st. \]

So far we have not used that $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{s+1}$ are nested. If $B_j \in \mathcal{F}_{s+1}$ then $B_j \in \mathcal{F}_i$ follows for all $1 \leq i \leq s$ as well. That is, $B_j$ has degree $s + 1$ in $\mathcal{G}$. Consequently, $B_j \in T$.

Thus setting $b := |B \cap \mathcal{F}_{s+1}|$, we infer $x \geq b$. Now (6) is a quadratic polynomial in $x$ with main term $x^2$. Therefore the maximum of (6) in the range $b \leq x \leq s$ is attained either for $x = b$ or $x = s$. We infer

\[ |\mathcal{G}| = |B \cap \mathcal{F}_1| + \cdots + |B \cap \mathcal{F}_{s+1}| \leq \max\{b^2 - (t - 1)b + st, s^2 - (t - 1)s + st\}. \]

To prove (5) we need to show that here the right hand side is at most $st - sb$. Let us check it separately for both terms. The inequality $b^2 - (t - 1)b + st \leq st - sb$ is equivalent to $b(t - 1 - s - b) \geq 0$ which is true because of $b \leq s$, $t \geq 2s + 1$. The inequality $s^2 - (t - 1)s + st \leq st - sb$ is equivalent to $s(t - 1 - s - b) \geq 0$ which is true for the same reason. □
Let us return to the proof of the Theorem 3.1. Since the lemma holds for all choices of $B$, the same inequality must hold for the expected values as well, yielding
\[
\frac{t|\mathcal{F}_1|}{\binom{|Y|}{\ell}} + \cdots + \frac{t|\mathcal{F}_s|}{\binom{|Y|}{\ell}} + (s + 1)\frac{t|\mathcal{F}_{s+1}|}{\binom{|Y|}{\ell}} \leq ts,
\]
or equivalently
\[
|\mathcal{F}_1| + \cdots + |\mathcal{F}_s| + (s + 1)|\mathcal{F}_{s+1}| \leq s\left(\binom{|Y|}{\ell}\right)
\]
as desired. □

**Remark.** Changing the requirement $t \geq 2s + 1$ one can prove similar inequalities where the coefficient of $|\mathcal{F}_{s+1}|$ is changing in function of $t$ and $s$.

4. The proof of Theorem 1.1

Let $\mathcal{F} \subseteq \binom{[n]}{s}$ be a stable family with $v(\mathcal{F}) = s$, $n \geq (2s + 1)k - s$. We want to prove $|\mathcal{F}| \leq |A(n, 1, s)|$. Let us write $A$ for short instead of $A(n, 1, s)$ throughout the proof. Let us partition both families according to the intersection of their edges with $[s + 1]$: For a subset $Q \subseteq [s + 1]$ define
\[
\mathcal{F}(Q) := \{ F \in \mathcal{F}: F \cap [s + 1] = Q \},
\]
\[
A(Q) := \{ A \in A: A \cap [s + 1] = Q \}.
\]
Note that for $|Q| \geq 2$, $|A(Q)| = \binom{n - s - 1}{k - 1}$ implying $|\mathcal{F}(Q)| \leq |A(Q)|$. For $1 \leq i \leq s$, $|A((i))| = \binom{n - s - 1}{k - 1}$ and $A((s + 1)) = A(\emptyset) = \emptyset$. Thus all we need to show is
\[
|\mathcal{F}(\emptyset)| + \sum_{1 \leq i \leq s+1} |\mathcal{F}((i))| \leq s\left(\binom{n - s - 1}{k - 1}\right).
\]
We prove (7) in two steps. First we prove
\[
|\mathcal{F}(\emptyset)| \leq s|\mathcal{F}([s + 1])|.
\]
As a matter of fact, for every $H \in \partial \mathcal{F}(\emptyset)$ stability of $\mathcal{F}$ implies $(H \cup [s + 1]) \in \mathcal{F}([s + 1])$. Now (8) is a direct consequence of Theorem 1.2. Plugging (8) into (7) we see that the inequality to prove is
\[
|\mathcal{F}((1))| + \cdots + |\mathcal{F}([s])| + (s + 1)|\mathcal{F}([s + 1])| \leq s\left(\binom{n - s - 1}{k - 1}\right).
\]
To apply Theorem 3.1 set $\mathcal{F}_i := \{ F - (i): F \in \mathcal{F}((i))\}$. Since $\mathcal{F}$ is stable, $\mathcal{F}_1, \ldots, \mathcal{F}_{s+1}$ are nested. Also, since $v(\mathcal{F}) = s$, $\mathcal{F}_1, \ldots, \mathcal{F}_{s+1}$ are cross-dependent. Setting $\ell = k - 1$, $Y = [s + 2, n]$, $|Y| = n - s - 1 \geq (2s + 1)(k - 1)$, all conditions of Theorem 3.1 are satisfied for $t = 2s + 1$. Thus (9) follows from (4), completing the proof.

In case of equality $\mathcal{F}(\emptyset) = \emptyset$ is immediate through Theorem 1.2. Then $\mathcal{F}([s + 1]) = \emptyset$ follows, leading to $\mathcal{F} \subseteq A$. □

5. Concluding remarks

The situation with Erdős’ Matching Conjecture was dormant for two decades. There was a sudden increase of interest during the last two years. It was mainly caused by the fact that through the works of Alon, Frankl, Huang, Rödl, Ruciński, and Sudakov [2] and Alon, Huang and Sudakov [1] it was shown that the Matching Conjecture is relevant in the proof of some seemingly unrelated problems. This motivated the research of Huang, Loh and Sudakov [12] and Frankl, Rödl and Ruciński [10] improving the old bounds of Erdős [4] and Bollobás, Daykin and Erdős [3]. Also it led to the complete solution

The present proof comes within a factor of two to of covering the full range, i.e., \( n \leq (s + 1)k - 1 \). However, a full solution does not seem possible along these lines. On the other hand some improvements are possible. Let us mention just two of them.

If \( k \leq s + 1 \) then \( |k| \subset [s + 1] \) implies that \( \nu(\mathcal{F}(\emptyset)) \leq s - 1 \). Using this fact the same proof yields that the Matching Conjecture is true already for \( n \geq 2sk - s \) and even earlier for the case that \( k \) is substantially smaller than \( s \).

For \( \mathcal{F}(\emptyset) \) we used that its matching number is at most \( s \). However, the much stronger statement \( \nu(\mathcal{F}(\emptyset)) \leq s \) follows from the stability of \( \mathcal{F} \). Using this property and the same inductive argument, the factor \( 1/s \) can be replaced by the larger \( \left( \frac{k}{(k-1)s-1} \right) \). The only reason that we did not prove and use this version is that for fixed \( s \) and \( k \) large, the ratio is approaching \( 1/s \) which does not permit an improvement of our bounds in general.

Let us conclude this paper by mentioning a Hilton–Milner-type extension of Erdős' Theorem. Hilton and Milner [11] determined the size of the largest intersecting subfamily \( \mathcal{F} \subset \binom{[n]}{k} \) with the property \( \bigcap \mathcal{F} = \emptyset \). We generalize their construction for all \( s \geq 1 \) by defining a family \( \mathcal{H} \) with \( \nu(\mathcal{H}) = s \) with the property that for every element \( x \in [n] \) one still has \( \nu(\mathcal{H}(\{x\})) = s \) (i.e., \( \mathcal{H} \) is \( s \)-stable). Let \( x_0, \ldots, x_{s-1} \) be elements and \( T_1, \ldots, T_s \) be disjoint \( k \)-subsets of \([n]\) such that \( x_i \in T_i \), \( i = 1, \ldots, s - 1 \) but \( x_0 \) is not contained in any of \( T_i \), \( i = 1, \ldots, s \). Define the family

\[
\mathcal{H}(n, s, k) := \left\{ H \in \binom{[n]}{k} : \text{there is an } i, \ 0 \leq i \leq s - 1, \ x_i \in H \right\}
\]

and then \( H \cap (T_{i+1} \cup \cdots \cup T_s) \neq \emptyset \right\} \cup \{T_1, \ldots, T_s\} \).

\[\text{Theorem 5.1. If } \mathcal{F} \subset \binom{[n]}{k} \text{ satisfies } \nu(\mathcal{F}) = s, \nu(\mathcal{F}(x)) = s \text{ for every } x \in [n], k \geq 4 \text{ and } n \geq n_1(s, k) \text{ then } |\mathcal{F}| \leq |\mathcal{H}(n, s, k)| \text{ holds with equality if and only if } \mathcal{F} \text{ is isomorphic to } \mathcal{H}(n, s, k).\]

The proof of this theorem together with a similar result for \( k = 3 \) will appear in a forthcoming paper.

References