The Erdős–Ko–Rado Theorem for Integer Sequences

Peter Frankl CNRS, ER 175 Combinatoire, 54 Bd Raspail, 75006 Paris, France

Norihide Tokushige College of Education, Ryukyu University, Nishihara, Okinawa, 903-0213, Japan hide@edu.u-ryukyu.ac.jp

Abstract

For positive integers n, q, t we determine the maximum number of integer sequences (a_1, \ldots, a_n) which satisfy $1 \le a_i \le q$ for $1 \le i \le n$, and any two sequences agree in at least t positions. The result gives an affirmative answer to a conjecture of Frankl and Füredi.

1 Introduction

Let n, q, t be positive integers with $q \ge 2$, $n \ge t$, and let $[q] := \{1, 2, \ldots, q\}$. Then $\mathcal{H} \subset [q]^n$ is a set of integer sequences $(a_1, \ldots, a_n), 1 \le a_i \le q$. We say that \mathcal{H} is *t*-intersecting if any two sequences intersects in at least *t* positions, more precisely, $|\{i : a_i = a'_i\}| \ge t$ holds for all $(a_1, \ldots, a_n), (a'_1, \ldots, a'_n) \in \mathcal{H}$. In this paper, we determine the exact value of the following function.

 $f(n,q,t) := \max\{|\mathcal{H}| : \mathcal{H} \subset [q]^n, \,\mathcal{H} \text{ is } t\text{-intersecting}\}.$

A family $\mathcal{A} \subset 2^{[n]}$ is called *t*-intersecting if $|A \cap A'| \ge t$ holds for all $A, A' \in \mathcal{A}$. Define a weighted size of \mathcal{A} by $w(\mathcal{A}) := \sum_{A \in \mathcal{A}} (q-1)^{n-|A|}$. Using a shifting technique, it is not difficult to check the following:

Lemma 1 (Proposition 2 in [5]) $f(n,q,t) = \max_{\mathcal{A}} w(\mathcal{A})$, where $\mathcal{A} \subset 2^{[n]}$ runs over all t-intersecting families.

If q = 2 then $w(\mathcal{A}) = |\mathcal{A}|$. Thus, f(n, 2, t) is simply the maximal size of t-intersecting family $\mathcal{A} \subset 2^{[n]}$, which is given by the Katona Theorem. This case was solved by Kleitman [7].

Let us define a *t*-intersecting family $\mathcal{A}_r \subset 2^{[n]}$ by

 $\mathcal{A}_r := \{ A \subset [n] : |A \cap [t+2r]| \ge t+r \}.$

In [5], Frankl and Füredi conjectured $f(n, q, t) = \max_{r \ge 0} w(\mathcal{A}_r)$. If $q \ge t + 1$ then the conjecture claims $f(n, q, t) = q^{n-t}$. They showed that this is true if $t \ge 15$.

Now we introduce the full Erdős–Ko–Rado theorem, which was conjectured by Frankl in [4], and proved by Ahlswede and Khachatrian in [1]. Set

$$AK(n, k, t, r) := |\{B \in {[n] \choose k} : |B \cap [t + 2r]| \ge t + r\}|.$$

Theorem 1 ([1]) Let $1 \le t \le k \le n$ and $\mathcal{B} \subset {\binom{[n]}{k}}$ be t-intersecting. If

$$(k-t+1)(2+\frac{t-1}{r+1}) \le n \le (k-t+1)(2+\frac{t-1}{r})$$

for some $r \in \mathbf{N}$, then $|\mathcal{B}| \leq \operatorname{AK}(n, k, t, r)$.

Using the above result, we prove the following in section 2.

Theorem 2 Let $q \ge 3$ and set $r := \lfloor \frac{t-1}{q-2} \rfloor$. Then $f(n,q,t) = w(\mathcal{A}_r)$ for $n \ge t+2r$. Note that

$$w(\mathcal{A}_{r}) = \sum_{j=0}^{n-t-2r} \sum_{i=t+r}^{t+2r} {\binom{t+2r}{i} \binom{n-t-2r}{j} (q-1)^{n-i-j}} = \sum_{j=0}^{n-t-2r} {\binom{n-t-2r}{j} (q-1)^{n-t-2r-j} \sum_{i=t+r}^{t+2r} {\binom{t+2r}{i} (q-1)^{t+2r-i}}} = q^{n-t-2r} \sum_{i=0}^{r} {\binom{t+2r}{i} (q-1)^{i}}.$$
(1)

In section 3, we prove the case $q \ge t + 1$ (and $t \ge 1$) directly.

Independently, Ahlswede and Khachatrian [2] obtained Theorem 2 as a diametric theorem in Hamming spaces. They used a different method. See [6] or [2] for the history of the problem.

2 Proof of the theorem

Throughout this section, we fix q and t and set

$$r := \lfloor \frac{t-1}{q-2} \rfloor = \frac{t-1}{q-2} - \delta.$$

Let us recall the following easy probabilistic result.

Lemma 2 (Proposition 3 in [5]) For every $\epsilon > 0$ the number of sequences $(a_1, \ldots, a_n) \in [q]^n$ which contain more than $(1 + \epsilon)(n/q)$ 1's or less than $(1 - \epsilon)(n/q)$ 1's is less than ϵq^n for $n > n_0(\epsilon)$.

Choose any sufficiently small positive ϵ , i.e., $0 < \epsilon < \epsilon_0(q, t)$, and set an open interval $I := ((1 - \epsilon)(n/q), (1 + \epsilon)(n/q))$. In view of Lemma 1, $f(n, q, t)q^{-n} = w(\mathcal{A})q^{-n}$ for some t-intersecting family \mathcal{A} . Moreover Lemma 2 gives that

$$f(n,q,t)q^{-n} < w(\mathcal{B})q^{-n} + \epsilon$$

where $\mathcal{B} := \{B \in \mathcal{A} : |B| \in I\}$. Set $\mathcal{B}(k) := \{B \in \mathcal{B} : |B| = k\}$.

Case I $0 < \delta < 1$.

Note that δ depends only on t and q.

Lemma 3 For $k \in I$ and sufficiently large n,

$$(k-t+1)\left(2+\frac{t-1}{r+1}\right) \le n \le (k-t+1)\left(2+\frac{t-1}{r}\right)$$
(2)

Proof (2) is equivalent to

$$(2 + (t-1)/r)^{-1}n + t - 1 \le k \le (2 + (t-1)/(r+1))^{-1}n + t - 1$$
(3)

Let us show the right half. Since $k < (1 + \epsilon)(n/q)$, it is sufficient to show

$$(1+\epsilon)(n/q) \le (2+(t-1)/(r+1))^{-1}n+t-1$$

or

$$(1+\epsilon)(2+(t-1)/(r+1)) < q.$$

This follows from $q = 2 + (t-1)/(r+\delta) > 2 + (t-1)/(r+1)$ and $\epsilon < \epsilon_0(q,t)$. One can prove the left half of (3) similarly.

Thus, by the Ahlswede–Khachatrian theorem we have $|\mathcal{B}(k)| \leq AK(n, k, t, r)$. Therefore,

$$\begin{split} f(n,q,t)q^{-n} &< q^{-n} \sum_{k \in I} w(\mathcal{B}(k)) + \epsilon \\ &\leq q^{-n} \sum_{k \in I} \operatorname{AK}(n,k,t,r)(q-1)^{n-k} + \epsilon \\ &= q^{-n} \sum_{k \in I} \sum_{j=t+r}^{t+2r} \binom{t+2r}{j} \binom{n-t-2r}{k-j} (q-1)^{n-k} + \epsilon \\ &< q^{-n} \sum_{j=t+r}^{t+2r} \binom{t+2r}{j} \sum_{k=j}^{n-t-2r+j} \binom{n-t-2r}{k-j} (q-1)^{n-k} + \epsilon \\ &= q^{-n} \sum_{j=t+r}^{t+2r} \binom{t+2r}{j} \sum_{i=0}^{n-t-2r} \binom{n-t-2r}{i} (q-1)^{(n-t-2r)-i} (q-1)^{t+2r-j} + \epsilon \\ &= q^{-n} \sum_{j=t+r}^{t+2r} \binom{t+2r}{j} q^{n-t-2r} (q-1)^{t+2r-j} + \epsilon \\ &= q^{-t-2r} \sum_{i=0}^{r} \binom{t+2r}{i} (q-1)^{i} + \epsilon. \end{split}$$

Hence we have

$$g(q,t) := \lim_{n \to \infty} f(n,q,t) q^{-n} \le q^{-t-2r} \sum_{i=0}^{r} \binom{t+2r}{i} (q-1)^{i}.$$
 (4)

On the other hand, (1) implies

$$g(q,t) \ge q^{-t-2r} \sum_{i=0}^{r} \binom{t+2r}{i} (q-1)^{i}.$$
 (5)

By (4) and (5), we finally have

$$g(q,t) = q^{-t-2r} \sum_{i=0}^{r} {t+2r \choose i} (q-1)^{i}$$

Now suppose that for some t-intersecting family $\mathcal{A} \subset 2^{[n]}$ we have $w(\mathcal{A}) \geq q^n g(q, t) + 1$. Since $f(n+1, q, t) \geq qf(n, q, t)$ we have

$$f(n',q,t) \ge q^{n'-n} f(n,q,t) \ge q^{n'-n} w(\mathcal{A}) \ge q^{n'} (g(q,t) + q^{-n}),$$

which implies $\lim_{n'\to\infty} f(n',q,t)q^{-n'} \ge g(q,t) + q^{-n} > g(q,t)$, a contradiction. Thus we must have $w(\mathcal{A}) \le q^n g(q,t)$, and actually $w(\mathcal{A}_r) = q^n g(q,t)$. (We need $n \ge t+2r$ here.) This completes the proof of Case I.

Case II $\delta = 0$. In this case, we have $q = 2 + \frac{t-1}{r}$.

Lemma 4 For $k \in I$ and sufficiently large n,

$$(k-t+1)(2+\frac{t-1}{r+1}) \le n \le (k-t+1)(2+\frac{t-1}{r-1}).$$

In fact, one can prove

$$(2 + \frac{t-1}{r-1})^{-1}n + t - 1 \le (1-\epsilon)\frac{n}{q} < \frac{n}{q} + t - 1 < (1+\epsilon)\frac{n}{q} \le (2 + \frac{t-1}{r+1})^{-1}n + t - 1.$$

The proof is similar to the proof of Lemma 3 and we omit it. By this lemma, we have

$$|\mathcal{B}(k)| \le \max\{\mathrm{AK}(n,k,t,r), \mathrm{AK}(n,k,t,r-1)\}.$$

If n = q(k - t + 1) then AK(n, k, t, r) = AK(n, k, t, r - 1). Since

$$AK(n, k, t, r) = \sum_{j=0}^{r} {\binom{t+2r}{t+r+j} \binom{n-t-2r}{k-t-r-j}} \\ = {\binom{n-t-2r}{k-t-r}} \sum_{j=0}^{r} {\binom{t+2r}{t+r+j}} \prod_{i=1}^{j} \frac{k-t-r-i+1}{n-k-r+i},$$

we have

$$1 = \frac{\mathrm{AK}(n,k,t,r-1)}{\mathrm{AK}(n,k,t,r)} = \frac{(n-t-2r+2)(n-t-2r+1)}{(k-t-r+1)(n-k-r+1)} \frac{\sum_{j=0}^{r-1} {\binom{t+2r-2}{t+r+j-1}} \prod_{i=1}^{j} \frac{k-t-r-i+2}{n-k-r+i+1}}{\sum_{j=0}^{r} {\binom{t+2r}{t+r+j}} \prod_{i=1}^{j} \frac{k-t-r-i+2}{n-k-r+i}}.$$

The above ratio tends to

$$\frac{q^2}{(q-1)} \frac{\sum_{j=0}^{r-1} {\binom{t+2r-2}{t+r+j-1}} (q-1)^{-j}}{\sum_{j=0}^r {\binom{t+2r}{t+r+j}} (q-1)^{-j}} = \frac{q^2}{(q-1)} \frac{\sum_{i=1}^r {\binom{t+2r-2}{i-1}} (q-1)^i}{\sum_{i=0}^r {\binom{t+2r-2}{i-1}} (q-1)^i}$$

as $n \to \infty$ for fixed q, t and n = q(k - t + 1). This proves

$$q^{2}\sum_{i=1}^{r} \binom{t+2r-2}{i-1} (q-1)^{i} = (q-1)\sum_{i=0}^{r} \binom{t+2r}{i} (q-1)^{i}$$
(6)

Now choose $k \in I$. (Here we do not assume n = q(k - t + 1).) Then,

$$\frac{\mathrm{AK}(n,k,t,r-1)}{\mathrm{AK}(n,k,t,r)} = \frac{(n-t-2r+2)(n-t-2r+1)}{(k-t-r+1)(n-k-r+1)} \frac{\sum_{j=0}^{r-1} \binom{t+2r-2}{t+r+j-1} \prod_{i=1}^{j} \frac{k-t-r-i+2}{n-k-r+i+1}}{\sum_{j=0}^{r} \binom{t+2r}{t+r+j} \prod_{i=1}^{j} \frac{k-t-r-i+1}{n-k-r+i}} \\
< \frac{n^2}{(1-\epsilon)(n/q)(1-(1+\epsilon)/q)n} \frac{\sum_{j=0}^{r-1} \binom{t+2r-2}{t+r+j-1} \prod_{i=1}^{j} \frac{(1+\epsilon)(n/q)}{(1-(1+\epsilon)/q)n}}{\sum_{j=0}^{r} \binom{t+2r}{t+r+j} \prod_{i=1}^{j} \frac{(1-\epsilon)(n/q)}{(1-(1-\epsilon)/q)n}} \\
= \frac{q^2}{(1-\epsilon)(q-1-\epsilon)} \frac{\sum_{i=1}^{r} \binom{t+2r-2}{i-1} (\frac{q-1-\epsilon}{1+\epsilon})^i}{\sum_{i=0}^{r} \binom{t+2r}{1-\epsilon} i}.$$

By (6), the above ratio tends to 1 as $\epsilon \to 0$. Thus for any $\epsilon' > 0$ we can conclude that

$$AK(n, k, t, r - 1) < (1 + \epsilon')AK(n, k, t, r)$$

if we choose ϵ sufficiently small and n sufficiently large, and $k \in I$. Finally we have

$$\begin{aligned} f(n,q,t)q^{-n} &< q^{-n} \sum_{k \in I} \max\{ \operatorname{AK}(n,k,t,r), \operatorname{AK}(n,k,t,r-1) \} (q-1)^{n-k} + \epsilon \\ &< (1+\epsilon')q^{-n} \sum_{k \in I} \operatorname{AK}(n,k,t,r) (q-1)^{n-k} + \epsilon \\ &< (1+\epsilon')q^{-t-2r} \sum_{i=0}^{r} \binom{t+2r}{i} (q-1)^{i} + \epsilon. \end{aligned}$$

Using the same argument in Case I, we have

$$g(q,t) := \lim_{n \to \infty} f(n,q,t)q^{-n} = q^{-t-2r} \sum_{i=0}^{r} \binom{t+2r}{i} (q-1)^{i},$$

and $f(n,q,t) = q^n g(q,t)$, which completes the proof of the theorem.

3 Another approach

In this section we give a direct proof for the case $q \ge t + 1$ using tools developed in [1].

Let $\mathcal{A} \subset 2^{[n]}$. A family $\mathcal{G} \subset 2^{[n]}$ is called a kernel of \mathcal{A} if $\mathcal{A} = \bigcup_{G \in \mathcal{G}} \mathcal{U}(G)$ where $\mathcal{U}(G) := \{F \subset [n] : G \subset F\}$. A rank of \mathcal{A} is defined by

$$\operatorname{rank}(\mathcal{A}) := \min\{|\bigcup_{G \in \mathcal{G}} G| : \mathcal{G} \text{ is a kernel of } \mathcal{A}\}.$$

Theorem 3 Let $\mathcal{A} \subset 2^{[n]}$ be a shifted t-intersecting family with $w(\mathcal{A}) = f(n, q, t)$. Then rank $(\mathcal{A}) \leq t + 2r$, where $r := \lfloor \frac{t-1}{q-2} \rfloor$.

Since the proof is almost the same as the proof of Lemma 6 in [1], we omit the details.

Proof (Outline) Choose a shifted, inclusion minimal (i.e., antichain) kernel $\mathcal{G} \subset 2^{[n]}$ of \mathcal{A} satisfying rank $(\mathcal{A}) = |\bigcup_{G \in \mathcal{G}} G|$. Assume that $\delta > 0$ and $M := t + 2r + \delta =$ rank(\mathcal{A}). Let $\mathcal{G} = \mathcal{G}_0 \cup \mathcal{G}_1$, $\mathcal{G}_0 := \{G \in \mathcal{G} : M \in G\}$, $\mathcal{G}_1 := \mathcal{G} - \mathcal{G}_0$, and let

$$\mathcal{G}_0 = \mathcal{R}_{t+1} \cup \cdots \cup \mathcal{R}_{M-1},$$

where $\mathcal{R}_i := \mathcal{G}_0 \cap {\binom{[M]}{i}}$. Set

$$\mathcal{R}'_i := \{E - \{M\} : E \in \mathcal{R}_i\} \subset \binom{[M-1]}{i-1}.$$

Then, $E \in \mathcal{R}'_i, E' \in \mathcal{R}'_j$ and $i+j \neq M+t$ imply $|E \cap E'| \geq t$. Thus we may assume that $\mathcal{R}_i \neq \emptyset$, $\mathcal{R}_j \neq \emptyset$, i + j = M + t for some i, j.

Case I $i \neq j$. Define

$$\begin{aligned} \mathcal{F}_1 &:= & \mathcal{G}_1 \cup (\mathcal{G}_0 - (\mathcal{R}_i \cup \mathcal{R}_j)) \cup \mathcal{R}'_i, \\ \mathcal{F}_2 &:= & \mathcal{G}_1 \cup (\mathcal{G}_0 - (\mathcal{R}_i \cup \mathcal{R}_j)) \cup \mathcal{R}'_j, \\ \mathcal{B}_i &:= & \mathcal{U}(\mathcal{F}_i). \end{aligned}$$

Then we have

$$\mathcal{A} - \mathcal{B}_1 = \{ R \cup S : R \in \mathcal{R}_j, S \in 2^{[M+1,n]} \},\$$

$$\mathcal{B}_1 - \mathcal{A} = \{ R \cup S : R \in \mathcal{R}'_i, S \in 2^{[M+1,n]} \},\$$

and hence

$$w(\mathcal{A} - \mathcal{B}_1) = |\mathcal{R}_j|(q-1)^{M-j}q^{n-M},$$

$$w(\mathcal{B}_1 - \mathcal{A}) = |\mathcal{R}_i|(q-1)^{M-i+1}q^{n-M}.$$

If $w(\mathcal{A}) \geq w(\mathcal{B}_1)$ and $w(\mathcal{A}) \geq w(\mathcal{B}_2)$ then

$$\begin{aligned} |\mathcal{R}_j|(q-1)^{M-j} &\geq |\mathcal{R}_i|(q-1)^{M-i+1}, \\ |\mathcal{R}_i|(q-1)^{M-i} &\geq |\mathcal{R}_j|(q-1)^{M-j+1}. \end{aligned}$$

Thus $1 \ge (q-1)^2$, a contradiction.

Case II $i = j = \frac{M+t}{2} = t + r + \frac{\delta}{2}$. In this case δ is even and $\delta \ge 2$. Using the same argument in Case I, we may assume that $\mathcal{R}_{\alpha} = \emptyset$ for all $\alpha \neq i$, and $\mathcal{G} = \mathcal{R}_i \cup \mathcal{G}_1$. The average degree \overline{d} of $\mathcal{R}'_i \subset {\binom{[M-1]}{i-1}}$ is given by $\overline{d} = (i-1)|\mathcal{R}_i|/(M-1)$. Therefore we can find $\ell \in [M-1]$ such that $\deg_{\mathcal{R}'_{\ell}}(\ell) \leq \bar{d}$. Define a *t*-intersecting family \mathcal{T} as follows:

$$\mathcal{T} := \{ E \in \mathcal{R}'_i : \ell \notin E \} \subset \binom{[M-1] - \{\ell\}}{i-1}.$$

Then $|\mathcal{T}| \geq |\mathcal{R}'_i| - \bar{d} = \frac{M-i}{M-1} |\mathcal{R}_i|$. Let $\mathcal{A} = \mathcal{D}_1 \cup \mathcal{D}_2$ where $\mathcal{D}_1 := \mathcal{U}(\mathcal{G}_1), \mathcal{D}_2 := \mathcal{U}(\mathcal{R}_i) - \mathcal{D}_1$, and let $\mathcal{U}(\mathcal{T} \cup \mathcal{G}_1) = \mathcal{D}_1 \cup \mathcal{D}_3$ where $\mathcal{D}_3 := \mathcal{U}(\mathcal{T}) - \mathcal{D}_1$. Then we have

$$w(\mathcal{D}_2) = |\mathcal{R}_i|(q-1)^{M-i}q^{n-M},$$

$$w(\mathcal{D}_3) = |\mathcal{T}|(q-1)^{M-i}q^{n-M+1} \ge \frac{M-i}{M-1}|\mathcal{R}_i|(q-1)^{M-i}q^{n-M+1}.$$

If $w(\mathcal{D}_2) \ge w(\mathcal{D}_3)$ then $1 \ge \frac{M-i}{M-1} \cdot q$. Since $M = t + 2r + \delta$ and $i = t + r + \frac{\delta}{2}$, we have

$$t + 2r + \delta - 1 \ge \frac{2r + \delta}{2}q,$$

or equivalently,

$$r \le \frac{t-1 - (q/2 - 1)\delta}{q - 2} = \frac{t-1}{q - 2} - \frac{\delta}{2}.$$

Since $\frac{\delta}{2} \ge 1$ we have $r \le \frac{t-1}{q-2} - 1$, which contradicts a definition of r.

Corollary 1 If $q \ge t+1$ then $f(n,q,t) = q^{n-t}$.

Proof Suppose that $\mathcal{A} \subset 2^{[n]}$ is *t*-intersecting and $w(\mathcal{A}) = f(n, q, t)$. By Theorem 3, we may assume $\operatorname{rank}(\mathcal{A}) \leq t + 2r$, $r := \lfloor \frac{t-1}{q-2} \rfloor$. If $q \geq t+2$ then r = 0, and $f(n, q, t) \leq w(\mathcal{A}_0) = q^{n-t}$.

If q = t + 1 then r = 1 and $f(n, q, t) \leq \max\{w(\mathcal{A}_0), w(\mathcal{A}_1)\}$. In this case we have $w(\mathcal{A}_0) = w(\mathcal{A}_1) = q^{n-t}$.

References

- R. Ahlswede, L.H. Khachatrian. The complete intersection theorem for systems of finite sets. *European J. of Combinatorics*, 18:125–136, 1997.
- [2] R. Ahlswede, L.H. Khachatrian. The diametric theorem in Hamming spaces Optimal anticodes. Adv. in Appl. Math., 20:429–449, 1998.
- [3] P. Erdős, C. Ko and R. Rado. Intersection theorems for systems of finite sets. Quart. J. Math. Oxford (2), 12:313–320, 1961.
- [4] P. Frankl. The Erdős-Ko-Rado theorem is true for n = ckt. Collq. Math. Soc. János Bolyai, 18:364–375, 1978.
- [5] P. Frankl, Z. Füredi. The Erdős–Ko–Rado theorem for integer sequences. SIAM J. Alg. Disc. Math., 1:376–381, 1980.
- [6] P. Frankl, Z. Füredi. Beyond the Erdős–Ko–Rado theorem. J. Combin. Theory (A), 56:182–194, 1991.
- [7] D. J. Kleitman. On a combinatorial conjecture of Erdős. J. Combin. Theory, 1:209–214, 1966.