Note

An Analogue of the Erdős-Ko-Rado Theorem for the Hamming Schemes $H(n, q)$

AERYUNG MOON

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210

Communicated by the Managing Editors

Received September 8, 1980

An analogue of the Erdős–Ko–Rado Theorem is proved for the Hamming scheme $H(n, q)$.

I. INTRODUCTION

Erdős, Ko and Rado [2] proved the following theorem:

Let $0 < r < k$. Let $F$ be a collection of $k$-subsets of an $n$-set $X$, and $n > n_0(k, r)$.

(i) If any two sets in $F$ intersect in at least $r$ elements, then $|F| \leq \binom{n}{k-r}$.

(ii) Equality holds if and only if $F$ consists of all $k$-subsets which contain a fixed $r$-subset of $X$.

Let $Q = \{0, 1, ..., q-1\}$ and $Q^n = \{x = (x_1, ..., x_n) : x_i \in Q\}$. We define $d(x, y)$, the distance between $x$ and $y$ ($x, y \in Q^n$), as the number of indices $i$ for which $x_i \neq y_i$. $Q^n$ with this distance is called the Hamming scheme $H(n, q)$. If $F, F' \subseteq Q^n$, we write $F \sim F'$ if there exists $\sigma \in S_n$ such that $\sigma F = \{\sigma x = (x_{\sigma(1)}, ..., x_{\sigma(n)}) : x = (x_1, ..., x_n) \in F\} = F'$, where $S_n$ is the symmetric group on $\{1, ..., n\}$.

In the language of association schemes (see [1]), the Erdős–Ko–Rado theorem gives an upper bound on the size of a family of diameter $k-r$ in the Johnson scheme $J(n, k)$, and characterizes the extremal case. Our theorem (or more precisely its specialization to $F_1 = F_2$) is an analogous result for $H(n, q)$.

0097-3165/82/030386-05$02.00/0
Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
THEOREM. Let $F_1, F_2 \subseteq Q^n$, and $q \geq r + 2$. If $|F||F_2| \geq q^{2(n-r)}$ and $d(x, y) \leq n - r$ for all $x \in F_1, y \in F_2$, then $F_1 = F_2 \sim \{(a_1, \ldots, a_r, x_{r+1}, \ldots, x_n): x_i \in Q \text{ for } i = r + 1, \ldots, n\}$, where $a_1, \ldots, a_r$ are fixed.

Remarks. 1. The case $r = 1$ and $F_1 = F_2$ has been proved by Livingston [4]. A new proof of Livingston's bound (though not of his characterization of the extremal cases) has been given by Stanton [6], using the technique discovered by Lovasz [5].

2. The bound for the case $F_1 = F_2$ has been proved by Frankl and Füredi [3], $r > 14$.

3. If $q < r + 2$, then there are counterexamples [3]: Let $n \geq r + 2$. Fix $a_1, \ldots, a_{r+2}$, and let $F_1 = F_2 = \{(x, \ldots, x_n) \in Q^n: x_i = a_i \text{ for at least } r + 1 \text{ }\}$s for $i \leq r + 2$. Then $|F_1| = |F_2| = q^{n-r-2}((r + 2)(q - 1) + 1)$, and $d(x, x') \leq n - r$ for all $x, x' \in F_1 = F_2$. If $q < r + 1$, then $|F_1| = |F_2| > q^{n-r-2}(q + 1)(q - 1) + 1 = q^{n-r}$. If $q = r + 1$, then $|F_1| = |F_2| = q^{n-r}$, but they don't have fixed $r$ coordinates.

II. PROOF OF THE THEOREM

To avoid trivialities we assume throughout that $0 < r < n$. If $x = (x_1, \ldots, x_n) \in Q^n$, then we let $F = (x_1, \ldots, x_n) \in Q^{n-1}$, and for $F \subseteq Q^n$, we let $F_2 = \{(x, \ldots, x_n) \in F\}$. We define $H_1 = \{x \in F_1: d(x, y) = n - r - 1 \text{ for all } y \in F_2\}$ and $H_2 = \{y \in F_2: d(x, y) = n - r - 1 \text{ for all } x \in F_1\}$. From now on, we assume the hypothesis of the Theorem and assume $|F_1||F_2|$ maximum and $|F_1| > |F_2|$. The next two lemmas say that if $x \in F_1$, then either $x \notin H_1$, in which case $x$ is unique; or $x \in H_1$, in which case $F_1$ may be assumed to contain each $z$ which differs from $x$ only in the first coordinate. The idea here is essentially the same as that used by Livingston [4].

LEMMA 1. If $x, z \notin H_1$ and $x = z$, then $x = z$.

Proof. Since $x, z \notin H_1$, there exists $y \in F_2$ such that $d(x, y) = d(z, y) = n - r$. Since both $d(x, y)$ and $d(z, y) \leq n - r$, we have $x_1 = y_1$ and $z_1 = y_1$. So, $x = z$.  

LEMMA 2. $\{(i, x_2, \ldots, x_n): x = (x_1, \ldots, x_n) \in H_1, i = 0, \ldots, q - 1\} \subseteq F_1$.

Proof. Suppose that $x = (i, x_2, \ldots, x_n) \in H_1$. Then for any $j, 0 \leq j \leq q - 1$, and any $y \in F_2$, we have (since $d$ is a metric) $d((j, x_2, \ldots, x_n), y) \leq 1 + d(x, y) \leq n - r$, so that the maximality of $|F_2||F_2|$ implies $(j, x_2, \ldots, x_n) \in F_1$.  

Proof of the Theorem

We use induction on \( n \). When \( n = 1 \), it is trivially true. So we assume \( n \geq 2 \) and assume the Theorem is true for \( n' < n \). Using the induction hypothesis, we prove the following Lemma.

**Lemma 3.** Either \( H_i = F_i \) or \( |H_i| < |F_i|/(q - 1) \), \( i = 1, 2 \).

**Proof.** Consider \( \tilde{H}_1 \) and \( \tilde{F}_2 \). Then by induction hypothesis, \( |\tilde{H}_1| \leq q^{2(n'-r-1)} \leq |F_1||F_2|/q^2 \). But \( |\tilde{H}_1| = |H_1|/q \) and \( |\tilde{F}_2| = |\tilde{H}_2|/q + |F_2\setminus H_2| \). So we get

\[
|H_1| \cdot (q|F_2| + (1 - q)|H_2|) \leq |F_1||F_2|.
\]

Similarly

\[
|H_2| \cdot (q|F_1| + (1 - q)|H_1|) \leq |F_1||F_2|.
\]

Equation (1) can be written as

\[
|F_2|(q|H_1| - |F_1|) \leq (q - 1)|H_1||H_2|.
\]

If we multiply both sides of (2) by \( (q|H_1| - |F_1|) \), we get

\[
|H_2|(q|F_1| + (1 - q)|H_1|)(q|H_1| - |F_1|) \leq |F_1||F_2|(q|H_1| - |F_1|).
\]

(We may assume \( q|H_1| - |F_1| > 0 \).) By (3), we get

\[
(q|F_1| + (1 - q)|H_1|)(q|H_1| - |F_1|) \leq (q - 1)|H_1||F_1|,
\]

which gives

\[
[(q - 1)|H_1| - |F_1||][|H_1| - |F_1|] \geq 0.
\]

So either \( |H_1| = |F_1| \) or \( |H_1| \leq |F_1|/(q - 1) \). Similarly \( |H_2| = |F_2| \) or \( |H_2| \leq |F_2|/(q - 1) \). But by (1) and (2), we have \( |H_i| = |F_i| \) if and only if \( |H_2| = |F_2| \). If \( |H_i| = |F_i|/(q - 1) \) for \( i = 1, 2 \), then equality holds everywhere. So, by induction hypothesis, \( |\tilde{H}_i| = |\tilde{F}_i| = q^{n'-r-1} \) for \( i = 1, 2 \). Since \( q|\tilde{H}_i| = |H_i| \), we get \( H_i = F_i \) for \( i = 1, 2 \). So if \( H_i \neq F_i \), then we have \( |H_1| < |F_1|/(q - 1) \) for \( i = 1, 2 \). \( \blacksquare \)

If \( F_1 = H_1 \) (so \( F_2 = H_2 \)), then we consider \( \tilde{F}_1 \) and \( \tilde{F}_2 \), and by induction hypothesis, we are done. So, it is enough to show that if \( n > r \), then there is at least one coordinate \( j \) such that \( H_1(j) = F_1 \), where \( H_1(j) = \{ x \in F_1 : d(\tilde{x}, \tilde{y}) \leq n - r - 1 \text{ for all } y \in F_2 \}, \) \( \tilde{x} = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \). (Actually we are going to show that there are at least \( n - r \) such coordinates.) Let \( A_i = \{ x \in F_1 : x_i = i \} \) and \( B_i = \{ y \in F_2 : y_i = i \} \). The next lemma proves that if \( H_1 \neq F_1 \), then there is an \( i \) such that \( |A_i| > |F_1|/q \).
**Lemma 4.** The following are equivalent.

(i) \( |A_i| = \frac{|F_1|}{q} \) for all \( i \);

(ii) \( H_1 = F_1 \).

**Proof.** Statement (ii) clearly implies (i). On the other hand, suppose 
\( |A_i| = \frac{|F_1|}{q} \) and \( |B_i| = \frac{|F_2|}{q} \) for all \( i \). It is enough to show \( \mathcal{A}_0 = \mathcal{A}_1 \). Let us consider \( \mathcal{A}_0 \) and \( \mathcal{B}_2 \) (we used \( q > 2 \)). Then since \( d(x, y) \leq n - r - 1 \) for all \( x \in \mathcal{A}_0, \ y \in \mathcal{B}_2 \) and \( |\mathcal{A}_0| |\mathcal{B}_2| \geq q^{2(n-r-1)} \), \( \mathcal{A}_0 = \mathcal{B}_2 \) by induction hypothesis. Similarly, \( \mathcal{A}_1 = \mathcal{B}_2 \). So \( \mathcal{A}_0 = \mathcal{A}_1 \).

We may assume \( |H_j| < \frac{|F_j|}{(q - 1)} \) for \( j = 1, 2 \). So by Lemma 4, we may assume \( |A_0| > \frac{|F_1|}{q} \). If \( |B_0| \leq \frac{|F_2|}{q} \), then there is \( i \neq 0 \) such that \( |B_i| \geq \frac{|F_2|}{q} \). By considering \( \mathcal{A}_0 \) and \( \mathcal{B}_i \) and using induction hypothesis, we get a contradiction. So, we have \( |B_0| > \frac{|F_2|}{q} \). The next lemma proves that both \( A_0 \) and \( B_0 \) will be dominating \( F_1 \) and \( F_2 \).

**Lemma 5.** \( |A_0| > \frac{|F_1|}{(1 - 1/q)} \) and \( |B_0| > \frac{|F_2|}{(1 - 1/q)} \).

**Proof.** We consider \( \mathcal{B}_0 \) and \( \mathcal{F}_1 \setminus A_0 \). Then \( |\mathcal{B}_0| |\mathcal{F}_1 \setminus A_0| \leq \frac{|F_1|}{|F_2|} q^2 \). Since \( |\mathcal{F}_1 \setminus A_0| = \left( (\mathcal{F}_1 \setminus A_0) \cap \mathcal{H}_1 \right) \cup \left( (\mathcal{F}_1 \setminus A_0) \setminus \mathcal{H}_1 \right) \),

\[
|\mathcal{F}_1 \setminus A_0| = \frac{|H_1|}{q} + \left( \frac{|F_1| - |A_0| - \frac{q-2}{q} |H_1|}{q} \right) = \frac{|H_1|}{q} - \frac{|A_0| - \frac{q-2}{q} |H_1|}{q} = \frac{|F_1|}{q} \left( 1 - \frac{q-2}{q(q-1)} \right) - |A_0|.
\]

So, we get

\[
\frac{q |B_0|}{|F_2|} \left( q - 1 + \frac{1}{q-1} - \frac{|A_0|}{|F_1|} \right) < 1.
\]

It implies

\[
\frac{|A_0|}{|F_1|} > q - 1 + \frac{1}{q-1} - \frac{|F_2|}{q |B_0|} . \quad (1)
\]

Similarly

\[
\frac{q |B_0|}{|F_2|} > q - 1 + \frac{1}{q-1} - \frac{|F_1|}{q |A_0|} . \quad (2)
\]

Let \( x = q |A_0| / |F_1|, \ y = q |B_0| / |F_2| \). Let \( x \geq y \), (If \( y \)-symmetric argument works.) Then by (1), \( x > q - 1 + 1/(q - 1) - 1/y \geq q - 1 + 1/(q - 1) - 1/x \).
Consider the function \( f(x) = x - (q - 1 + 1/(q - 1) - 1/x) \). (By assumption, \( x > 1 \).) Since \( f''(x) > 0 \) for \( x > 1 \), \( f(x) \) is strictly increasing. Since \( f(q - 1) = 0 \) and \( f(x) > 0, \ x > q - 1 \). Then by (2), \( y > q - 1 \). It implies \( |A_0| > |F_1|(1 - 1/q) \) and \( |B_0| > |F_2|(1 - 1/q) \).

**Lemma 6.** There are at least \( n - r \) coordinates \( i \) such that \( H_1(i) = F_1 \).

**Proof.** Suppose it is false. Then, there are at least \( r + 1 \) coordinates, say, \( i = 1, 2, \ldots, r + 1 \) such that \( |H_1(i)| < |F_1|(q - 1) \). By Lemma 5, there are \( a_1, \ldots, a_{r+1} \) such that \( |C_i| = |\{x \in F_1 : x_i = a_i\}| > |F_1|(1 - 1/q) \) for \( i = 1, \ldots, r + 1 \). So, \( |\bigcap_{i=1}^{r+1} C_i| > |F_1|(1 - (r + 1)/q) \geq q^{n-r}(1 - (r + 1)/q) \). But \( |\bigcap_{i=1}^{r+1} C_i| \leq q^{n-r-1} \), because the first \( r + 1 \) coordinates are fixed. So, we get \( q(1 - (r + 1)/q) < 1 \). It implies \( q < r + 2 \), which gives a contradiction.

As we noted before, Lemma 6 finishes the proof of the Theorem.

**References**