The Maximum Number of Lines Contained in Subsets of $PG(k,2)$

Hegang Chen
Department of Statistics
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0439

Abstract

Let M be an m-subset of $PG(k,2)$, the finite projective geometry of dimension k over $GF(2)$. We would like to know the maximum number of lines that can be contained in M. In this paper, we will not only give the maximum number of lines contained in m-subsets of $PG(k,2)$, but also construct an m-subset of $PG(k,2)$ containing the maximum number of lines.

1 Introduction

Let k be a positive integer. The space of all vectors $(x_0,...,x_k)$, x_i belonging to the finite field $GF(2)$ is called the projective geometry of dimension k over $GF(2)$ and is denoted by $PG(k,2)$. The zero vector $\mathbf{0}=(0,...,0)$ is the void space and we say that the void space has dimension -1. A point is determined by a vector $\mathbf{x}=(x_0,...,x_n)\neq\mathbf{0}$. There are $2^{k+1}-1$ distinct points in $PG(k,2)$. More generally, if $\mathbf{y}_0,...,\mathbf{y}_r$ are $r+1$ independent vectors, the set of all vectors $b_0\mathbf{y}_0+...+b_r\mathbf{y}_r$, $b_i\in GF(2)$ is a subspace of dimension r. One-dimensional subspaces of $PG(k,2)$ are called lines (for detailed discussion, see Dembowski(1968) or Hall(1986)).

Let M be a subset of m distinct points of $PG(k,2)$ and let $L(M)$ denote the number of lines in M. The rank of M, denoted by $rank(M)$, is the maximum number of independent points of M. In this paper, we study the problem of finding the maximum number of lines that can be contained in an m-subset of $PG(k,2)$. To search for an m-subset containing the maximum number of lines, the lemmas in Section 2 will show that we only need to consider m-subsets with the minimum possible rank. The main results of the paper will be presented in Section 3.

ARS COMBINATORIA 57(2000), pp. 103-110

2 Preliminaries

If M is a subset with rank $p+1$ ($\leq k+1$), then M can be represented as

$$M = H \cup \{\mathbf{a}, \mathbf{a}+\mathbf{b}_1, ..., \mathbf{a}+\mathbf{b}_f\}, \tag{1}$$

where H is a subset of $PG(p-1,2)$ (embedded in $PG(k,2)$) with rank p, $\mathbf{a} \in PG(k,2) \backslash PG(p-1,2)$ and $\mathbf{b}_1, ..., \mathbf{b}_f \in PG(p-1,2)$.

Lemma 1 *Let M be an m-subset of $PG(k,2)$ with rank $p+1$ and having a form as in (1). Then there exists an m-subset M' of $PG(k,2)$*

$$M' = H' \cup \{\mathbf{a}, \mathbf{a}+\mathbf{b}_1, ..., \mathbf{a}+\mathbf{b}_t\}, \tag{2}$$

where H' is a subset of $PG(p-1,2)$ with rank p, $\mathbf{a} \in PG(k,2) \backslash PG(p-1,2)$ and $\mathbf{b}_1, ..., \mathbf{b}_t \in H'$, such that M' has at least as many lines as M.

Proof. Let M have the representation as in (1), where $\mathbf{b}_1, ..., \mathbf{b}_t \in H$ and $\mathbf{b}_{t+1}, ..., \mathbf{b}_f \in PG(p-1,2) \backslash H$. Let $H_1^0 = \{\mathbf{a}, \mathbf{a}+\mathbf{b}_1, ..., \mathbf{a}+\mathbf{b}_t\}$, $H_2^0 = \{\mathbf{a}+\mathbf{b}_{t+1}, ..., \mathbf{a}+\mathbf{b}_f\}$ and $H^0 = H_1^0 \cup H_2^0$. Consider

$$M' = H' \cup \{\mathbf{a}, \mathbf{a}+\mathbf{b}_1, ..., \mathbf{a}+\mathbf{b}_t\},$$

where $H' = H \cup \{\mathbf{b}_{t+1}, ..., \mathbf{b}_f\}$ is a subset of $PG(p-1,2)$ with rank p. All lines in M can be classified into two classes. One class consists of all lines contained in H. Clearly, M' preserves all these lines. The other class contains lines formed by two points from H^0 and one point from H. M' preserves all lines in M containing two points from H_1^0 and one point from H. Lines in M containing two points from H_2^0 and one point from H are counted by corresponding to lines in M' containing two points from $\{\mathbf{b}_{t+1}, ..., \mathbf{b}_f\}$ and one point from H. Lines in M containing one point from $H_1^0 \backslash \{\mathbf{a}\}$, one point from H_2^0 and one point from H are counted by corresponding to lines in M' with one point from $\{\mathbf{b}_1, ..., \mathbf{b}_t\}$, one point from $\{\mathbf{b}_{t+1}, ..., \mathbf{b}_f\}$ and one point from H. Thus, there are at least as many lines in M' as there are in M. Consequently, $L(M) \leq L(M')$.□

From Lemma 1, it suffices to consider the representation given in (2) for maximizing the number of lines. For $m = 2^r + q$ ($0 \leq q < 2^r$), the rank of any m–subset is at least $r+1$. However, if the rank exceeds $r+1$ then the following lemma gives additional information about M.

Lemma 2 *Let M be an m-subset of $PG(k,2)$, $m = 2^r + q$ and $0 \leq q < 2^r (r \leq k)$. If the rank of M is larger than r+1, then there is an m-subset of $PG(k,2)$ with smaller rank whose number of lines is greater than the number of lines in M.*

Proof. Let $rank(M) = p + 1 > r + 1$. By Lemma 1, we may assume that

$$M = H \cup \{\mathbf{a}, \mathbf{a} + \mathbf{b}_1, ..., \mathbf{a} + \mathbf{b}_t\},$$

where H is a subset of $PG(p-1,2)$ with rank p, $\mathbf{a} \in PG(k,2) \backslash PG(p-1,2)$ and $\mathbf{b}_1, ..., \mathbf{b}_t \in H$. Lines in M can be classified into two classes. One class consists of all lines contained in H and another class which contains all lines containing two points from $\{\mathbf{a}, \mathbf{a} + \mathbf{b}_1, ..., \mathbf{a} + \mathbf{b}_t\}$ and one point from H. Since $rank(M) > r + 1$, but $|M| < 2^{r+1}$, we find that $PG(p-1,2) \backslash H$ is non-empty, so there are at least two points $\mathbf{s}_0^1$, $\mathbf{s}_0^2 \in H$ such that $\mathbf{c}_0 = \mathbf{s}_0^1 + \mathbf{s}_0^2 \in PG(p-1,2) \backslash H$. Consider the following process:
Form the set

$$M_0 = H \cup \{\mathbf{c}_0, \mathbf{c}_0 + \mathbf{b}_1, ..., \mathbf{c}_0 + \mathbf{b}_t\},$$

with $\mathbf{c}_0, \mathbf{c}_0 + \mathbf{b}_1, ..., \mathbf{c}_0 + \mathbf{b}_{t_0} \notin H$ and $\mathbf{c}_0 + \mathbf{b}_{t_0+1}, ..., \mathbf{c}_0 + \mathbf{b}_t \in H$ for some t_0 ($\leq t$). Therefore, M_0 can be represented as

$$M_0 = H \cup \{\mathbf{c}_0, \mathbf{c}_0 + \mathbf{b}_1, ..., \mathbf{c}_0 + \mathbf{b}_{t_0}\}.$$

We observe that M_0 preserves all lines of the first class in M and its rank is p.

If $t = t_0$, we shall argue that M_0 has more lines than M. All lines of the second class in M are counted by corresponding to lines in M_0 which contain two points from $\{\mathbf{c}_0, \mathbf{c}_0 + \mathbf{b}_1, ..., \mathbf{c}_0 + \mathbf{b}_{t_0}\}$ and one point from H. Further, M_0 has at least one more line, namely, $\{\mathbf{s}_0^1, \mathbf{s}_0^2, \mathbf{c}_0\}$. Consequently, $L(M_0) > L(M)$.

If $t > t_0$, the number of points in M_0 is less than m. We shall now consider the lines in the second class in M and relate some of these lines to those in M_0. The lines in the second class in M can be conveniently classified into the following two types:

(a) Lines formed by $\{\mathbf{a}, \mathbf{a}+\mathbf{b}_i, \mathbf{b}_i\}$, $i = 1, ..., t$

(b) Lines formed by $\{\mathbf{a}+\mathbf{b}_i, \mathbf{a}+\mathbf{b}_j, \mathbf{b}_i+\mathbf{b}_j\}$, where $\mathbf{b}_i+\mathbf{b}_j \in H$.

Let us now go back to M_0. There are two possibilities for $\mathbf{c}_0$.
(i): $\mathbf{c}_0$ is not a sum of two points from $\{\mathbf{b}_1, ..., \mathbf{b}_t\}$. All lines of type (a) are counted by corresponding to lines in M_0 formed by $\{\mathbf{c}_0, \mathbf{c}_0+\mathbf{b}_i, \mathbf{b}_i\}$. The lines of type (b) can be further classified into the following three cases,

(b1) Those lines $\{\mathbf{a} + \mathbf{b}_i, \mathbf{a} + \mathbf{b}_j, \mathbf{b}_i + \mathbf{b}_j\}$ with $1 \leq i, j \leq t_0$,

(b2) Those lines $\{\mathbf{a} + \mathbf{b}_i, \mathbf{a} + \mathbf{b}_j, \mathbf{b}_i + \mathbf{b}_j\}$ with $1 \leq i \leq t_0, t_0 + 1 \leq j \leq t$,

(b3) Those lines $\{\mathbf{a} + \mathbf{b}_i, \mathbf{a} + \mathbf{b}_j, \mathbf{b}_i + \mathbf{b}_j\}$ with $t_0 + 1 \leq i, j \leq t$.

The lines of type (b1) are counted by corresponding to distinct new lines $\{\mathbf{c}_0+\mathbf{b}_i, \mathbf{c}_0+\mathbf{b}_j, \mathbf{b}_i+\mathbf{b}_j\}$ in M_0. The lines of type (b3) have not been counted in M_0.

The lines of type (b2) are counted by corresponding to the new lines $\{\mathbf{c}_0+\mathbf{b}_i, \mathbf{c}_0+\mathbf{b}_j, \mathbf{b}_i+\mathbf{b}_j\}$ in M_0. However it is possible that two distinct lines of type (b2) may correspond to the same line in M_0. This causes us to lose one line of type (b2) for each case. Suppose that there are two lines of type (b2); $\{\mathbf{a}+\mathbf{b}_i, \mathbf{a}+\mathbf{b}_j, \mathbf{b}_i+\mathbf{b}_j\}$, $\{\mathbf{a}+\mathbf{b}_{i'}, \mathbf{a}+\mathbf{b}_{j'}, \mathbf{b}_{i'}+\mathbf{b}_{j'}\}$ ($1 \le i, i' \le t_0$ and $t_0+1 \le j, j' \le t$) corresponding to the same line in M_0, i.e., $\{\mathbf{c}_0+\mathbf{b}_i, \mathbf{c}_0+\mathbf{b}_j, \mathbf{b}_i+\mathbf{b}_j\} = \{\mathbf{c}_0+\mathbf{b}_{i'}, \mathbf{c}_0+\mathbf{b}_{j'}, \mathbf{b}_{i'}+\mathbf{b}_{j'}\}$. Since $\mathbf{c}_0+\mathbf{b}_j, \mathbf{b}_i+\mathbf{b}_j, \mathbf{c}_0+\mathbf{b}_{j'}, \mathbf{b}_{i'}+\mathbf{b}_{j'} \in H$ and $\mathbf{c}_0+\mathbf{b}_i, \mathbf{c}_0+\mathbf{b}_{i'} \in M_0 \backslash H$, we have $\mathbf{b}_i = \mathbf{b}_{i'}$, $\mathbf{c}_0+\mathbf{b}_j = \mathbf{b}_{i'}+\mathbf{b}_{j'}$ and $\mathbf{c}_0+\mathbf{b}_i = \mathbf{b}_j+\mathbf{b}_{j'} \in M_0 \backslash H$. Therefore, the indistinctness results in a new line $\{\mathbf{a}+\mathbf{b}_j, \mathbf{a}+\mathbf{b}_{j'}, \mathbf{b}_j+\mathbf{b}_{j'}\}$ ($\mathbf{b}_j+\mathbf{b}_{j'} \in M_0 \backslash H$) which did not exist in M. All such lines are classified as type (b3*). We will show that the lost lines of type (b2) resulting from indistinctness will be counted in the following process by taking the new lines of type (b3*) into account.

To count the remaining lines in (b3) and the new lines of type (b3*) resulting from indistinctness of the corresponding lines in M_0 of the lines of type (b2), we add $(t-t_0)$ more point(s) to M_0 such that the new m-subset with rank p not only counts those lines, but also has at least one more additional line than those in M. Since $rank(M_0) = p \ge r+1$, $PG(p-1,2) \backslash M_0 \neq \emptyset$, there are at least two points $\mathbf{s}_1^1, \mathbf{s}_1^2 \in M_0$ such that $\mathbf{s}_1^1+\mathbf{s}_1^2 \in PG(p-1,2) \backslash M_0$. Let $\mathbf{c}_1 = \mathbf{s}_1^1+\mathbf{s}_1^2+\mathbf{b}_{t_0+1}$, and now build M_1 from M_0 by

$$M_1 = M_0 \cup \{\mathbf{c}_1+\mathbf{b}_{t_0+1}, ..., \mathbf{c}_1+\mathbf{b}_t\}.$$

If all $\mathbf{c}_1+\mathbf{b}_{t_0+1}, ..., \mathbf{c}_1+\mathbf{b}_t$ are in $PG(p-1,2) \backslash M_0$, the lines of type (b3) can be counted by corresponding to the lines in M_1 formed by $\{\mathbf{c}_1+\mathbf{b}_i, \mathbf{c}_1+\mathbf{b}_j, \mathbf{b}_i+\mathbf{b}_j\}$ where $\mathbf{b}_i+\mathbf{b}_j \in H$ and $t_0 < i, j \le t$. Meanwhile the lines of type (b3*) which are used to compensate for the lost lines resulted from indistinctness are counted by corresponding to the lines formed by $\{\mathbf{c}_1+\mathbf{b}_{i'}, \mathbf{c}_1+\mathbf{b}_{j'}, \mathbf{b}_{i'}+\mathbf{b}_{j'}\}$ where $\mathbf{b}_{i'}+\mathbf{b}_{j'} \in M_0 \backslash H$ and $t_0 < i', j' \le t$. In this case, M_1 has at least one more additional line, namely, $\{\mathbf{s}_1^1, \mathbf{s}_1^2, \mathbf{c}_1+\mathbf{b}_{t_0+1}\}$, than M. Otherwise, say, $\mathbf{c}_1+\mathbf{b}_{t_0+1}, ..., \mathbf{c}_1+\mathbf{b}_{t_1} \in PG(p-1,2) \backslash M_0$ and $\mathbf{c}_1+\mathbf{b}_{t_1+1}, ..., \mathbf{c}_1+\mathbf{b}_t$ are in M_0 or zero. In this case, we build M_1 as follows:

$$M_1 = M_0 \cup \{\mathbf{c}_1+\mathbf{b}_{t_0+1}, ..., \mathbf{c}_1+\mathbf{b}_{t_1}\}.$$

If there is a $\mathbf{b}_i$ such that $\mathbf{c}_1+\mathbf{b}_i=0$, say $i = t_1+1$, then $\mathbf{b}_{t_1+1} = \mathbf{c}_1 = \mathbf{s}_1^1+\mathbf{s}_1^2+\mathbf{b}_{t_0+1}$. Since $\mathbf{b}_{t_1+1}+\mathbf{b}_i = \mathbf{c}_1+\mathbf{b}_i \in M_1 \backslash M_0$ for $t_0+1 \le i \le t_1$, there are no remaining lines of the form $\{\mathbf{a}+\mathbf{b}_i, \mathbf{a}+\mathbf{b}_{t_1+1}, \mathbf{b}_{t_1+1}+\mathbf{b}_i\}$. Thus it does not affect the process. Clearly, $|M_1| > |M_0|$ (M_1 contains at least one more point, namely, $\mathbf{s}_1^1+\mathbf{s}_1^2 = \mathbf{c}_1+\mathbf{b}_{t_0+1}$). The lines uncounted in

M_1 are similar to those at the first step, i.e., the lines containing two points from $\{\mathbf{a}+\mathbf{b}_{t_1+1}, ..., \mathbf{a}+\mathbf{b}_t\}$ and one point from M_0 or $M_1 \backslash M_0$ (which result from indistinctness of corresponding lines in M_1).

To count all these lines, we repeat the same process on M_1, and iterate until we have added $t - t_0$ points to M_0. As we always add at least one new point to M_0 at each step the process is guaranteed to stop, and the fact that we never lose any lines ensures that the final step produces a set M_v with m points and rank p, which has at least as many lines as M.
(ii): $\mathbf{c}_0$ is a sum of two points from $\{\mathbf{b}_1, ..., \mathbf{b}_t\}$. $\mathbf{c}_0$ has to be a sum of two points from $\{\mathbf{b}_{t_0+1}, ..., \mathbf{b}_t\}$ and $t > t_0 + 1$. It is possible that there are several pairs $\{\mathbf{b}_{1_i}, \mathbf{b}_{2_i}\}$, $(i = 1, ..., h)$, such that the sum of each pair is equal to $\mathbf{c}_0$ (all these pairs are distinct). Since $\mathbf{c}_0 = \mathbf{b}_{1_i} + \mathbf{b}_{2_i}$, two lines of type (a) $\{\mathbf{a}, \mathbf{a}+\mathbf{b}_{1_i}, \mathbf{b}_{1_i}\}$ and $\{\mathbf{a}, \mathbf{a}+\mathbf{b}_{2_i}, \mathbf{b}_{2_i}\}$ in M correspond to one line $\{\mathbf{c}_0, \mathbf{c}_0 + \mathbf{b}_{1_i}, \mathbf{b}_{1_i}\}$=$\{\mathbf{c}_0, \mathbf{c}_0 + \mathbf{b}_{2_i}, \mathbf{b}_{2_i}\}$ in M_0. Similarly, the indistinctness results in a new line $\{\mathbf{a}+\mathbf{b}_{1_i}, \mathbf{a}+\mathbf{b}_{2_i}, \mathbf{b}_{1_i}+\mathbf{b}_{2_i}\}$ which did not exist in M. As discussed in (i), the lost lines in the first step will be compensated in the following process by taking the new lines $\{\mathbf{a}+\mathbf{b}_{1_i}, \mathbf{a}+\mathbf{b}_{2_i}, \mathbf{b}_{1_i}+\mathbf{b}_{2_i},\}$ $(i = 1, ..., h)$ into account. By the same argument as in (i), the lemma is established.□

3 Main Results

From Lemma 2, we can see that m-subsets containing the maximum number of lines must have the minimum rank. To search for the m-subsets containing the maximum number of lines, we only need to consider all m-subsets of $PG(r, 2)$ $(2^r \leq m < 2^{r+1})$.

An m-subset M with minimum rank $r+1$ can be obtained by deleting $2^{r+1} - 1 - m$ points from $PG(r, 2)$. Without loss of generality, we can represent all points of $PG(r, 2)$ as

$$\underbrace{\mathbf{a}_1, ..., \mathbf{a}_m}_{M}, \underbrace{\mathbf{a}_{m+1}, ..., \mathbf{a}_{2^{r+1}-1}}_{\overline{M}}, \tag{3}$$

where, the first m points are all points of M, and $\overline{M}$ denotes all points of $PG(r, 2) \backslash M = \{\mathbf{a}_{m+1}, ..., \mathbf{a}_{2^{r+1}-1}\}$.

Let $L(M)$ and $L(\overline{M})$ be the numbers of lines in M and $\overline{M}$ respectively. We shall now study the relationship between $L(M)$ and $L(\overline{M})$. Let $L(PG(r, 2))$ be the number of lines of $PG(r, 2)$, we have (see Hirschfeld(1979))

$$L(PG(r,2)) = \frac{(2^{r+1}-1)(2^{r+1}-2)}{(2^2-1)(2^2-2)} = \frac{(2^{r+1}-1)(2^{r+1}-2)}{6}$$

All lines of $PG(r, 2)$ can be classified as one of the following three types:
(T1) Those lines containing one point from $\overline{M}$ and two points from M,

(T2) Those lines containing one point from M and two points from $\overline{M}$,
(T3) Those lines containing three points from $\overline{M}$ or from M.

Each pair of m points in M determines a line, but these $\begin{pmatrix} m \\ 2 \end{pmatrix}$ lines are not all distinct. Indeed, $\begin{pmatrix} 3 \\ 2 \end{pmatrix} L(M)$ pairs out of $L(PG(r,2))$ lines of $PG(r,2)$ in M are duplicated. Therefore, the number of lines of type (T1) is

$$\begin{pmatrix} m \\ 2 \end{pmatrix} - 3L(M). \tag{4}$$

Similarly, the number of lines of type (T2) is

$$\begin{pmatrix} 2^{r+1}-1-m \\ 2 \end{pmatrix} - 3L(\overline{M}), \tag{5}$$

and the number of lines of type (T3) is

$$L(M) + L(\overline{M}). \tag{6}$$

Since the sum of (4), (5) and (6) is equal to the total number of lines in $PG(r,2)$, i.e.,

$$\begin{pmatrix} m \\ 2 \end{pmatrix} - 3L(M) + \begin{pmatrix} 2^{r+1}-1-m \\ 2 \end{pmatrix} - 3L(\overline{M}) + L(M) + L(\overline{M}) = \frac{(2^{r+1}-1)(2^{r+1}-2)}{6},$$

we can conclude the following relation between $L(M)$ and $L(\overline{M})$

$$L(M) = \frac{1}{2}\left[\begin{pmatrix} m \\ 2 \end{pmatrix} + \begin{pmatrix} 2^{r+1}-1-m \\ 2 \end{pmatrix} - \frac{(2^{r+1}-1)(2^{r+1}-2)}{6}\right] - L(\overline{M}). \tag{7}$$

From (7), we have the following lemma.

Lemma 3 *Let M be an m-subset of $PG(r,2)$ and $\overline{M} = PG(r,2)\backslash M$. Then M contains the maximum number of lines among all m-subsets of $PG(r,2)$ if and only if $\overline{M}$ contains the minimum number of lines among all $(2^{r+1}-1-m)$-subsets of $PG(r,2)$.*

By applying Lemma 3, the following theorem provides one structure of an m-subset containing the maximum number of lines.

Theorem 1 *Let $m = 2^r + q$ and $0 \leq q < 2^r$ $(r \leq k)$. Then the maximum number of lines in an m-subset of $PG(k,2)$ is*

$$\frac{(2^r-1)(2^r-2)}{6} + \begin{pmatrix} q+1 \\ 2 \end{pmatrix}. \tag{8}$$

One structure of an m-subset of $PG(k,2)$ containing the maximum number of lines is

$$M = PG(r-1,2) \cup \{\mathbf{a}, \mathbf{a}+\mathbf{a}_{2^r-q}, ..., \mathbf{a}+\mathbf{a}_{2^r-1}\}, \tag{9}$$

where $PG(r-1,2) = \{\mathbf{a}_1, ..., \mathbf{a}_{2^r-1}\}$, $\mathbf{a}_i \in PG(r-1,2)$ $(i = 2^r-q, ..., 2^r-1)$, and $\mathbf{a} \notin PG(r-1,2)$.

Proof. From the equation (7), it can be checked that the maximum number (8) is an upper bound on the number of lines. By Lemma 2, we only need to consider all m-subsets of $PG(r,2)$ for finding the maximum number of lines. The points of $PG(r,2)$ can be partitioned into the sets:

$$PG(r-1,2), \mathbf{a}, \mathbf{a}+PG(r-1,2) \tag{10}$$

where $PG(r-1,2) = \{\mathbf{a}_1, ..., \mathbf{a}_{2^r-1}\}$, $\mathbf{a}+PG(r-1,2) = \{\mathbf{a}+\mathbf{a}_i | \mathbf{a}_i \in PG(r-1,2)\}$ and $\mathbf{a} \notin PG(r-1,2)$. An m-subset M of $PG(r,2)$ can be obtained by deleting 2^r-1-q points in (10). Let $\overline{M} = PG(r,2) \backslash M$, $\overline{M} = \{\mathbf{a}+\mathbf{a}_1, ..., \mathbf{a}+\mathbf{a}_{2^r-1-q}\}$. $\overline{M}$ contains no lines, i.e., $L(\overline{M}) = 0$. By Lemma 3, M in (9) contains the maximum number of lines among all m-subsets of $PG(k,2)$. There are $(2^r-1)(2^r-2)/6$ lines in $PG(r-1,2)$, others lines in M are those containing two points from $\{\mathbf{a}, \mathbf{a}+\mathbf{a}_{2^r-q}, ..., \mathbf{a}+\mathbf{a}_{2^r-1}\}$ and one point from $PG(r-1,2)$. The number of lines of the latter type in M is $\begin{pmatrix} q+1 \\ 2 \end{pmatrix}$. The result (8) is the sum of these two numbers.□

Remark.(9) is a structure of an m-subset of $PG(k,2)$ containing the maximum number of lines. However, this structure is not unique. For convenience, a point of $PG(k,2)$ is denoted by $i_1 i_2 \cdots i_l$ if the i_1th, i_2th,...,i_lth coordinates of this point are 1 and all others are zeros. For example, when $k = 3$, the following two 10-subsets both contain the maximum number of lines,

$$M_{10} = \{1, 2, 3, 123, 12, 23, 4, 34, 234, 1234\}$$

and

$$M_{10}^* = \{1, 2, 3, 13, 12, 23, 123, 4, 14, 1234\}.$$

M_{10}^* has a structure as in (9), and M_{10} does not have a structure as in (9).

In projective space, a *cap* is a set of points with the property that no three points are collinear. In particular, a cap of $PG(k,2)$ contains no line. From Theorem 1, we have the following theorem.

Theorem 2 *Let M be an m-subset of PG(k,2), with $m = 2^r + q$ and $0 \leq q < 2^r$. Then M contains the maximum possible number of lines of all m-subsets if and only if $rank(M) = r + 1$ and the complement of M in the copy of $PG(r,2)$ which M generates is a cap.*

Proof. If M contains the maximum possible number of lines of all m-subsets, then M must have the minimum rank, i.e., $rank(M) = r + 1$. By Theorem 1, the maximum number of lines in M is (8). Without loss of generality, say $PG(r,2)$ is generated by M, and the complement of M in $PG(r,2)$ is $\overline{M} = PG(r,2) \backslash M$. From equation (7), it can be calculated that $L(\overline{M}) = 0$. Therefore, $\overline{M}$ is a cap. By using the similar argument, it is not difficult to show that these conditions are also sufficient.□

Acknowledgment: I would like to thank the referee for his truly helpful suggestions which led to a much improved presentation of the paper.

References

[1] P. Dembowski, Finite Geometries (Springer-Verlag, New York, 1968).

[2] M. Hall Jr., Combinatorial Theory (John Wiley & Sons, New York, 1986).

[3] J. W. P. Hirschfeld, Projective Geometries over Finite Fields (Oxford University Press, Oxford, 1979).