Coding Theory in Projective Spaces

Tuvi Etzion
Department of Computer Science
Technion-Israel Institute of Technology
Haifa 32000, Israel
Email: etzion@cs.technion.ac.il

Alexander Vardy
Department of Electrical Engineering
University of California San Diego
La Jolla, CA 92093, USA
Email: avardy@ucsd.edu

***Abstract*— Given the n-dimensional space $\mathbb{F}_q^n$, the elements of the projective spaces are all subspaces of $\mathbb{F}_q^n$. Recently these codes have found application in error-correction of network coding. In this paper we examine a few interesting aspects of coding theory in projective spaces. We present codes and bounds in the projective spaces metric and prove that there are no perfect codes in this metric. We consider linear codes and complement codes in this metric and show some interesting phenomena. Lot of open interesting question arise from our discussion.**

I. Introduction

Given the n-dimensional space $\mathbb{F}_q^n$ ($\mathbb{F}_q = \mathrm{GF}(q)$), the *projective spaces metric* consists of all subspaces of $\mathbb{F}_q^n$ as elements; for two subspaces $\mathcal{X}$ and $\mathcal{Y}$ the distance between $\mathcal{X}$ and $\mathcal{Y}$ is defined as $d(\mathcal{X},\mathcal{Y}) = dim(\mathcal{X})+dim(\mathcal{Y})-2dim(\mathcal{X}\cap\mathcal{Y})$. Let $P(\mathbb{F}_q^n)$ denote the set of all subsets of subspaces from $\mathbb{F}_q^n$. A code $\mathcal{C}$ in the projective spaces is an element of $P(\mathbb{F}_q^n)$. For a subset $\mathcal{W} \in P(\mathbb{F}_q^n)$, let $Gr_k(\mathcal{W})$ denote the set of all k-dimensional subspaces of $\mathcal{W}$. Recently these codes have found application in error-correction of network coding [13]. This application is the motivation for our interest in these codes and for this research. All k-dimensional subspaces of $\mathbb{F}_q^n$ with the distance $d(\mathcal{X},\mathcal{Y}) = \frac{dim(\mathcal{X})+dim(\mathcal{Y})-2dim(\mathcal{X}\cap\mathcal{Y})}{2}$ for two given k-dimensional subspaces $\mathcal{X}$ and $\mathcal{Y}$ form what is known as the *Grassman scheme*. The Grassman scheme is akin to the well-known Johnson scheme and the projective spaces metric is akin to the Hamming scheme. We note that the distance in the Grassman scheme is half of the distance in the projective spaces metric. This is akin to the situation in the Johnson scheme and the Hamming scheme. As we will consider throughout this paper codes in the projective spaces metric, we will always take the distance as the one of this metric even if all codewords are from the Grassman scheme.

With respect to coding theory, the Grassman scheme was considered during the last twenty years only in connection of perfect codes and tilings [1], [6], [15], [17], but there were related works on intersecting families [11] and byte-correcting codes [10]. Specifically, the nonexistence of perfect codes in the Grassman scheme was proved first in [6] and later in [15]. In [1] it was proved that Steiner structures are the diameter perfect codes in this scheme. Properties of these structures were discussed in [17]. Koetter and Kschischang [13] discussed the applications of projective spaces codes as error-correcting codes in network coding. They discussed the Singleton bound, the sphere packing bound, and the Gilbert-Varshamov bound in the Grassman scheme and gave a construction of Reed-Solomon like codes. More applications for codes in the Grassman scheme were given in [24] in the context of authentication codes.

Subsets from the projective spaces were considered in the literature during the last thirty years in various aspects combinatorics, design theory, and extremal combinatorial problems. In many places they are called the q-analogues of the problems related to subsets of an n-set. Two interesting books which consider q-analogues are [2], [25]. Partitions of $\mathbb{F}_q^n$ into subspaces were considered in [3], [5], [7], [9], [12], [20]. Designs over the projective spaces were considered in [4], [16], [18], [19], [21], [22]. A q-analogue to the well known Sperner's Theorem is given in [25]. Many more problems are considered with a rich literature.

In this paper we discuss two issues of code in projective spaces. Lower and upper bounds on the sizes of codes are discussed in Section II. Linear codes over projective spaces and complements of codes are discussed in Section III.

II. Basic definitions and bounds

In the same way that binomial coefficients play an important role in enumeration of subsets, Gaussian coefficients play an important role in enumeration computations for the projective spaces metric. The q-ary *Gaussian coefficient* $\begin{bmatrix} n \\ k \end{bmatrix}_q$ is defined by

$$\begin{bmatrix} n \\ 0 \end{bmatrix}_q = 1 \ ,$$

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n-1)(q^{n-1}-1)\cdots(q^{n-k+1}-1)}{(q^k-1)(q^{k-1}-1)\cdots(q-1)} \ .$$

A *Steiner structure* $S[t,k,n]_q$ is a collection $\mathcal{S}$ of k-dimensional subspaces from $\mathbb{F}_q^n$ such that each t-dimensional subspace of $\mathbb{F}_q^n$ is contained in exactly one element of $\mathcal{S}$. The following results were given in [17].

Lemma 1: If a Steiner structure $S[t,k,n]_q$ exists then a Steiner structure $S[t-1,k-1,n-1]_q$ exists.

Theorem 1: A necessary condition for a Steiner structure $S[t,k,n]_q$ to exist, is that the numbers $\frac{\begin{bmatrix} n-i \\ t-i \end{bmatrix}_q}{\begin{bmatrix} k-i \\ t-i \end{bmatrix}_q}$, must be integers, for all $0 \leq i \leq t$.

A *constant dimension code* $[n, 2\delta, k]_q$ is a code whose codewords are k-dimensional subspaces of $\mathbb{F}_q^n$ and its minimum distance is 2δ. An $[n,d]_q$ code is a code whose codewords are subspaces of $\mathbb{F}_q^n$ and its minimum distance is d. The code has dimension distribution $\mathcal{D}_0, \mathcal{D}_1, \cdots, \mathcal{D}_n$, where $\mathcal{D}_i$ is the number of codewords with dimension i. Finally, let $A_q[n, 2\delta, k]$ be the maximum number of codewords in an $[n, 2\delta, k]_q$ code, and let $A_q[n,d]$ be the maximum number of codewords in an $[n,d]_q$ code. For an $[n,d]_q$ code $\mathcal{C}$, the *dual code* $\mathcal{C}^\perp$ is defined by $\mathcal{C}^\perp = \{\mathcal{P}^\perp \ : \ \mathcal{P} \in \mathcal{C}\}$.

Koetter and Kschischang [13] presented an analog to the Singleton bound.

Theorem 2: $A_q[n, 2\delta, k] \leq \begin{bmatrix} n-\delta+1 \\ k-\delta+1 \end{bmatrix}_q$.

Xia and Fu [26] proved analogs for two of Johnson bounds [14]. The one of more interest is the following.

Theorem 3: $A_q[n, 2\delta, k] \leq \frac{q^n-1}{q^k-1} A_q[n-1, 2\delta, k-1]$.

Theorem 3 has a similar proof to the analog theorem in the Johnson scheme. An analog to the dual theorem in the Johnson scheme can be also given, although the two proofs we have for it are completely different from the proof of its analog.

Theorem 4: $A_q[n, 2\delta, k] \leq \frac{q^n-1}{q^{n-k}-1} A_q[n-1, 2\delta, k]$.

By using iterations of Theorem 3 we obtain [26].

Theorem 5:

$$A_q[n, 2\delta, k] \leq \lfloor \frac{q^n-1}{q^k-1} \lfloor \frac{q^{n-1}-1}{q^{k-1}-1} \cdots \lfloor \frac{q^{n-k+\delta}-1}{q^\delta-1} \rfloor \cdots \rfloor \rfloor \ .$$

The bound of Theorem 5 is always better than the bound of Theorem 2 [26].

An important result on a dual code is the following.

Lemma 2: If $\mathcal{C}$ is an $[n,d]_q$ code then $\mathcal{C}^\perp$ is an $[n,d]_q$ code.

Xia and Fu [26] stated Lemma 2 for constant dimension codes and used it to prove.

Lemma 3: $A_q[n,d,k] = A_q[n,d,n-k]$.

For the projective spaces metric a Gilbert-Varshamov lower bound on $A_q[n,d]$ can be given by using the method of Tolhuizen [23].

A trivial upper bound on $A_q[n,d]$ is $A_q[n,d] \leq \sum_{k=0}^{n} A_q[n,d,k]$. This bound is improved by using linear programming [14]. We demonstrate it for $d=3$.

$$\text{maximize} \quad \sum_{k=0}^{n} \mathcal{D}_k \ ,$$

subject to

$$\frac{q^{n-i+1}-1}{q-1}\mathcal{D}_{i-1} + \mathcal{D}_i + \frac{q^{i+1}-1}{q-1}\mathcal{D}_{i+1} \leq \begin{bmatrix} n \\ i \end{bmatrix}_q, \ 0 \leq i \leq n,$$

$$\text{and} \quad \mathcal{D}_i \leq A_q[n,4,i], \ 0 \leq i \leq n,$$

where $\mathcal{D}_{-1}$ and $\mathcal{D}_{n+1}$ are defined to be 0.

Clearly, the target function maximize $\sum_{k=0}^{n} \mathcal{D}_k$ is an upper bound on $A_q[n,d]$.

By using several computational methods for finding codes with certain structures, certain interesting codes were found. As examples, we have $A_2[5,3] = 18$ and $A_2[9,3,4] \geq 5694$.

If k divides n then there exists a Steiner structure $S[1,k,n]_q$ [1], [17] which is equivalent to a perfect byte-correcting code [10], [17]. By Theorem 1, such structures can exist only when k divides n. When k does not divide n we have the following two results.

Theorem 6: $A_q[n, 2k, k] \leq \lfloor \frac{q^n-1}{q^k-1} \rfloor - 1$ if n is not divisible by k.

Theorem 7: If $n = sk + m$, where $m > k$, then $A_q[n, 2k, k] \geq \frac{q^m(q^{sk}-1)}{q^k-1} + 1$.

The nonexistence of perfect codes in the Grassman scheme is well known [6], [15]. The related graph is distance-regular and hence standard methods can be applied on it for a proof of such result. The related graph for the projective spaces metric is not distance-regular and hence the proof for the nonexistence of perfect codes in this metric requires different techniques. Anyway, the definition of a perfect code is given in the usual way. By using Theorem 1, Theorem 3, and Theorem 6 we obtain the following result.

Theorem 8: There is no nontrivial perfect code in the projective spaces metric over $\mathbb{F}_q$.

III. Linear codes and complements

There are many basic properties of codes in the Hamming scheme which are used for code construction. Linear codes are with no doubt the most basic one and over $\mathbb{F}_2$ complement is another one. In this section we will restrict ourself only to the binary case, even so most of the results can be extended for $\mathbb{F}_q$. If $\mathcal{C}$ is a binary code in the Hamming scheme, the complement of $\mathcal{C}$, $\bar{\mathcal{C}}$ has the same size as $\mathcal{C}$ and the same minimum distance. The complement of a code $\mathcal{C}$ consists of complements of its codewords. The usual definition of a complement for a k-dimensional subspace $\mathcal{P}$ of $\mathbb{F}_q^n$ is an $(n-k)$-dimensional subspace $\mathcal{Q}$ such that $\mathcal{P} \cap \mathcal{Q} = \{0\}$. But, in contrast to the binary Hamming scheme, this definition does not produce a unique complement over $\mathbb{F}_2^n$. This immediately raise the question whether there exists a definition of complements such that the complement of any code $\mathcal{C}$ over $\mathbb{F}_2^n$ is a code $\bar{\mathcal{C}}$ with the same size. We will prove that the answer to this question is positive, but any definition will suffer from two faults. The first one is that if $n = 2k$ then $\mathcal{Q}$ might be the complement of $\mathcal{P}$, while $\mathcal{P}$ is not a complement of $\mathcal{Q}$. The second fault is that the minimum distance of $\mathcal{C}$ and $\bar{\mathcal{C}}$ might be different.

Having explained the problems, that might occur, in the definition of complements, we turn to the more basic concept of linear codes. Linear codes are the most comfortable codes to handle in the Hamming scheme as they have simple representation and relatively simple encoding and decoding algorithms. Therefore, it is natural to ask whether there exist linear codes in the projective spaces metric? This natural

and simple question is in fact quite complicated. We start with seven definitions of the properties we require from a linear code and complements. We will show why these properties are necessary. After that we discuss which linear codes and complements exist in the projective spaces metric. Let $\mathcal{W}, \mathcal{W}' \in P(\mathbb{F}_2^n)$.

Definition 1: A function $+ : \mathcal{W} \times \mathcal{W} \rightarrow \mathcal{W}$ is a *group addition* on $\mathcal{W}$ if $\{\mathcal{W}, +\}$ is a group.

Definition 2: A function $+ : \mathcal{W} \times \mathcal{W} \rightarrow \mathcal{W}$ is an *abelian group addition* on $\mathcal{W}$ if $\{\mathcal{W}, +\}$ is an abelian group.

Definition 3: A function $+ : \mathcal{W} \times \mathcal{W} \rightarrow \mathcal{W}$ is a *linear addition* on $\mathcal{W}$ if $\{\mathcal{W}, +\}$ is vector space over $\mathbb{F}_2$ and the identity is the null-space, i.e. $\{0\}$.

Definition 4: The addition functions defined above are said to be *isometric* (so *isometric group addition*, *isometric linear addition*, etc) if the following condition holds:

For all $\mathcal{V}, \mathcal{U}_1, \mathcal{U}_2 \in \mathcal{W}$, $d(\mathcal{V} + \mathcal{U}_1, \mathcal{V} + \mathcal{U}_2) = d(\mathcal{U}_1, \mathcal{U}_2)$.

Definition 5: A function $f : \mathcal{W} \rightarrow \mathcal{W}'$ is a *complement* if the following conditions hold:

- For all $k = 0, 1, ..., n$, the restriction of f to $Gr_k(\mathcal{W})$ is a bijection between $Gr_k(\mathcal{W})$ and $Gr_{n-k}(\mathcal{W}')$.
- For all $\mathcal{U} \in \mathcal{W}$, $\mathcal{U} \cap f(\mathcal{U}) = \{0\}$.

Definition 6: A function $f : \mathcal{W} \rightarrow \mathcal{W}'$ is a *symmetric complement* on $\mathcal{W}$ if, in addition to the conditions of Definition 5, $\mathcal{W} = \mathcal{W}'$ and for all $\mathcal{U} \in \mathcal{W}$, $f(f(\mathcal{U})) = \mathcal{U}$.

Definition 7: A function $f : \mathcal{W} \rightarrow \mathcal{W}'$ is an *isometric complement* on $\mathcal{W}$ if, in addition to the conditions of Definition 5, for all $\mathcal{U}_1, \mathcal{U}_2 \in \mathcal{W}$, $d(f(\mathcal{U}_1), f(\mathcal{U}_2)) = d(\mathcal{U}_1, \mathcal{U}_2)$.

If $\mathcal{W} = \mathbb{F}_2^n$, can we have functions $+$ and f such that all the seven definitions are satisfied? Unfortunately, we don't have such functions which satisfy a smaller subset of the definitions. Therefore, we will examine now for which functions $+$ and f a large subset of the definitions are satisfied, when $\mathcal{W} = \mathbb{F}_2^n$? We will also examine what is the largest subset $\mathcal{W}$ of $\mathbb{F}_2^n$ for which there exist functions $+$ and f which satisfy the seven definitions (or subsets of these definitions)? Finally, we want to explain why we need all these seven definitions?

Clearly if there exists a group $\{G, +\}$, $|G| = g$, then there exists an addition function $+$ which satisfies Definition 1 for any $\mathcal{W}$ such that $|\mathcal{W}| = g$. The same is true if $\{G, +\}$ is an abelian group and Definition 2. Any bijection between G and $\mathcal{W}$, will imply a group addition which satisfies Definition 1 (or Definition 2).

Theorem 9: There exists a subset $\mathcal{W}$ of $\mathbb{F}_2^n$, $|\mathcal{W}| = 2^n$, and functions $+$ and f, which satisfy Definitions 1 through 7.

We conjecture that there is no subset $\mathcal{W}$ of $\mathbb{F}_2^n$, $|\mathcal{W}| > 2^n$, and functions $+$ and f, which satisfy Definitions 1 through 7.

It is easy to define an addition which satisfies Definitions 1 through 3 if $\mathcal{W}$ has cardinality 2^m, $\{0\} \in \mathcal{W}$. Definition 4 is the one which makes the difference. Definition 4 implies that the minimum distance of a linear code in projective spaces is preserved under addition of a constant subspace from $\mathcal{W}$. It implies several other results such as:

Lemma 4: If $\mathcal{S}_1$ and $\mathcal{S}_2$ are two subspaces in $\mathcal{W}$ then $d(\mathcal{S}_1, \mathcal{S}_2) = dim(\mathcal{S}_1 + \mathcal{S}_2)$, i.e., $dim(\mathcal{S}_1 + \mathcal{S}_2) = dim(\mathcal{S}_1) + dim(\mathcal{S}_2) - 2dim(\mathcal{S}_1 \cap \mathcal{S}_2)$.

Lemma 5: If $\mathcal{S}_1$ and $\mathcal{S}_2$ two disjoint subspaces in $\mathcal{W}$ then $S_1 + S_2 = span(S_1, S_2)$.

Lemma 6: The number of 1-dimensional subspaces of $\mathbb{F}_2^n$ in a set $\mathcal{W}$ with a function $+$ which satisfy Definitions 1 through 4 is at most n.

Can Lemma 6 be extended for any k-dimensional subspaces, i.e., is there a set $\mathcal{W}$ with a function $+$ which satisfy Definitions 1 through 4 with more than $\binom{n}{k}$ k-dimensional subspaces? We believe that the answer is NO if $\mathbb{F}_2^n \in \mathcal{W}$, i.e., when there exists a function f such that Definitions 5 through 7 are also satisfied. In this case we beleive that all codes are isomorphic to the one constructed in Theorem 9. Otherwise, the answer is YES whenever $n \geq 3$.

By using a recursive construction we are able to prove the following theorem.

Theorem 10: There is a function f which satisfies Definition 5 when $\mathcal{W} = \mathbb{F}_2^n$.

Finally, we ask whether Definitions 5 through 7 can be satisfied? For this we have the following results.

Lemma 7: There is no function f which satisfies Definition 6, when $\mathcal{W} = Gr_k(\mathbb{F}_2^{2k})$.

Lemma 8: There is no function f which satisfies Definitions 5 and 7, when $\mathcal{W} = Gr_2(\mathbb{F}_2^4)$.

Lemma 9: There is no function f which satisfies Definitions 5 and 7, when $\mathcal{W} = Gr_1(\mathbb{F}_2^3)$.

ACKNOWLEDGMENT

This work was supported in part by the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel, under Grant 2006097.

REFERENCES

[1] R. Ahlswede, H. K. Aydinian, and L. H. Khachatrian, "On perfect codes and related concepts", *Designs, Codes, and Cryptography,* vol. 22, 221–237, 2001.

[2] G. E. Andrews, "The Theory of Partitions", *Cambridge: Cambridge University Press*, 1984.

[3] A. Beutelspacher, "Partitions of finite vector spaces: An application of the Frobenius number in Geometry", *Arch Math.,* vol. 31, 202–208, 1978.

[4] M. Braun, A. Kerber, and R. Laue, "Systematic construction of q-analogs of $t-(v,k,\lambda)$-designs", *Designs, Codes, and Cryptography,* vol. 34, 55–70, 2005.

[5] T. Bu, "Partitions of a vector space", *Discrete Mathematics,* vol. 31, 79–83, 1980.

[6] L. Chihara, "On the zeros of the Askey-Wilson polynomials, with applications to coding theory", *SIAM Journal Math. Anal.,* vol. 18, 191–207, 1987.

[7] W. Clark and L. Dunning, "Partial partitions of vector spaces arising from the construction of byte error control codes", *Ars Combinatoria,* vol. 33, 161–177, 1992.

[8] P. Delsarte, "An algebraic approach to association schemes of coding theory", *Philips J. Res.,* vol. 10, 1–97, 1973.

[9] S. I. El=Zanati, G. F. Seelinger, P. A. Sissokho, L. E. Spence, and C. Vanden Eynden, "Partitions of finite vector spaces into subspaces", *Journal of Combinatorial Designs,* to appear.

[10] T. Etzion, "Perfect byte-correcting codes", *IEEE Transactions on Information Theory,* vol. IT-44, 3140–3146, 1998.

[11] P. Frankl and R. M. Wilson, "The Erdos-Ko-Rado theorem for vector spaces", *Journal Combinatorial Theory, Series A,* vol. 43, 228–236, 1986.

[12] O. Heden, "On partitions of finite vector spaces of small dimensions", *Arch math.,* vol. 43, 507–509, 1984.

[13] R. Kötter and F. R. Kschischang, "Coding for errors and erasures in random network coding", *IEEE Trans. on Inform. Theory*, submitted.
[14] F. J. MacWilliams and N. J. A. Sloane, "The Theory of Error-Correcting Codes", *Amsterdam: North-Holland*, 1977.
[15] W. J. Martin and X. J. Zhu, "Anticodes for the Grassman and bilinear forms graphs", *Designs, Codes, and Cryptography,* vol. 6, 73–79, 1995.
[16] M. Miyakawa, A. Munemasa, and S. Yoshiara, "On a class of small 2-designs over GF(q)", *Journal Combinatorial Designs,* vol. 3, 61–77, 1995.
[17] M. Schwartz and T. Etzion, "Codes and anticodes in the Grassman graph", *Journal Combinatorial Theory, Series A,* vol. 97, 27–42, 2002.
[18] H. Suzuki, "2-designs over GF(2^m)", *Graphs and Combinatorics,* vol. 6, 293–296, 1990.
[19] H. Suzuki, "2-designs over GF(q)", *Graphs and Combinatorics,* vol. 8, 381–389, 1992.
[20] P. Tannenbaum, "Partitions of $\mathbb{Z}_2^n$", *SIAM Journal Alg. Discrete Meth.,* vol. 4, 22–29, 1983.
[21] S. Thomas, "Designs over finite fields", *Geometriae Dedicata,* vol. 21, 237–242, 1987.
[22] S. Thomas, "Designs and partial geometries over finite fields", *Geometriae Dedicata,* vol. 63, 247–253, 1995.
[23] L. M. G. M. Tolhuizen, "The generalized Gilbert-Varshamov bound is implied by Turán theorem", *IEEE Transactions on Information Theory,* vol. IT-43, 1605–1606, 1997.
[24] H. Wang, C. Xing, R. Safavi-Naimi, "Linear authentication codes: bounds and constructions", *IEEE Transactions on Information Theory,* vol. IT-49, 866–872, 2003.
[25] J. H. van Lint and R. M. Wilson, "A course in Combinatorics", *Cambridge University Press*, 1992.
[26] S.-T. Xia and F.-W. Fu, "Johnson type bounds on constant dimension codes", *Designs, codes, and cryptography,* submitted.