Linear Codes associated to Flag Varieties over Finite Fields

Sudhir R. Ghorpade
Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai 400076, India
http://www.math.iitb.ac.in/∼srg/

Based on joint work with Fernando Piñero and Prasant Singh

Coding Theory and Cryptography
A conference in honor of Joachim Rosenthal's 60th birthday
Zurich, Switzerland
July 11–15, 2022

Review of Coding Theory

$\mathbb{F}_q$: finite field with q elements.

- $[n,k]_q$-code: a k-dimensional subspace of $\mathbb{F}_q^n$.
- Hamming weight of $c = (c_1, \ldots, c_n) \in \mathbb{F}_{q^n}$: $\mathrm{wt}(c) := |\{i : c_i \neq 0\}|$.
- Hamming weight of a subcode D of C:

$$\mathrm{wt}(D) := |\{i : \exists\ c \in D \text{ such that } c_i \neq 0\}|.$$

- Minimum distance of a (linear) code C:

$$d(C) := \min\{\mathrm{wt}(c) :\ c \in C,\ c \neq 0\}|.$$

- The r^{th} generalized Hamming weight of C $(1 \leqslant r \leqslant k)$:

$$d_r(C) := \min\{\mathrm{wt}(D)\| : D \subseteq C,\ \dim D = r\}.$$

- C nondegenerate if $C \nsubseteq$ a coordinate hyperplane.
- Dual of C: The $[n, n-k]_q$-code $C^\perp := \{x \in \mathbb{F}_{q^n} : x \cdot c = 0 \text{ for all } c \in C\}$.

A Geometric Language for Coding Theory

Projective Systems á la Tsfasman-Vlăduţ

$\mathbb{P}^{k-1}$: $(k-1)$-dim projective space over $\mathbb{F}_q$.

$[n,k]_q$ projective system: collection $\mathcal{P}$ of n points in $\mathbb{P}^{k-1}$ (need not be distinct); this is nondegenerate if it is not contained in a hyperplane.

nondegenerate $[n,k]_q$-code $C \rightsquigarrow$ nondegenerate $[n,k]_q$-projective system $\mathcal{P}$

Conversely, a nondegenerate $[n,k]_q$-projective system $\mathcal{P}$ gives rise to a nondegenerate $[n,k]_q$-code $C_{\mathcal{P}}$, and the resulting correspondence is a bijection, up to equivalence. In this set-up,

$$d(C_{\mathcal{P}}) = n - \max\{|\mathcal{P} \cap H| : \ H \text{ hyperplane of } \mathbb{P}^{k-1}\}.$$

and GHWs corresponds to maximal linear sections, i.e., for $r = 1, 2 \ldots, k$,

$$d_r(C_{\mathcal{P}}) = n - \max\{|\mathcal{P} \cap E| : \ E \text{ linear subvariety of codim } r \text{ in } \mathbb{P}^{k-1}\}.$$

What is a good code?

What is a good code?

Often, one likes to construct $[n, k]_q$-codes that satisfy one or more of the following requirements (some of which may conflict with each other):

- $d = d(C)$ is large vis-a-vis n, e.g., it's nice if d/n is close to 1.
- k is also large vis-a-vis n, e.g., it's nice if k/n is close to 1.
- C admits "good encoding and decoding"
- The generalized Hamming weights $d_r(C)$ are known for $r = 1, \ldots, k$.
- "Automorphism group(s)" of C are known and have a fairly large size.
- The "spectrum" of C is known.
- C is a LDPC code; this is typically the case when $d(C^\perp)$ is small and the minimum weight codewords of $C^\perp$ generate $C^\perp$.

What is a good code?

Often, one likes to construct $[n, k]_q$-codes that satisfy one or more of the following requirements (some of which may conflict with each other):

- $d = d(C)$ is large vis-a-vis n, e.g., it's nice if d/n is close to 1.
- k is also large vis-a-vis n, e.g., it's nice if k/n is close to 1.
- C admits "good encoding and decoding"
- The generalized Hamming weights $d_r(C)$ are known for $r = 1, \ldots, k$.
- "Automorphism group(s)" of C are known and have a fairly large size.
- The "spectrum" of C is known.
- C is a LDPC code; this is typically the case when $d(C^\perp)$ is small and the minimum weight codewords of $C^\perp$ generate $C^\perp$.

Brief Take-home Message from this Talk: We review the class of Grassmann codes and argue that this is a good code. Then we discuss a natural generalization of Grassmann codes, called flag codes, which is potentially a good code and also a source of some interesting and challenging problems.

Grassmann Varieties: A Quick Introduction

V : vector space of dimension m over a field $\mathbb{F}$.

Grassmann variety: For $1 \leqslant \ell \leqslant m$,

$$G_{\ell,m} = G_\ell(V) = \{\ell\text{-dimensional subspace of } V\}$$

For example, $G_{1,m} = \mathbb{P}^{m-1}$.

Grassmann Varieties: A Quick Introduction

V : vector space of dimension m over a field $\mathbb{F}$.

Grassmann variety: For $1 \leqslant \ell \leqslant m$,

$$G_{\ell,m} = G_\ell(V) = \{\ell\text{-dimensional subspace of } V\}$$

For example, $G_{1,m} = \mathbb{P}^{m-1}$.

Plücker embedding: $G_{\ell,m} \hookrightarrow \mathbb{P}(\bigwedge^\ell V) = \mathbb{P}^{\binom{m}{\ell}-1}$ defined by

$$W = \langle w_1, \ldots, w_\ell \rangle \longmapsto [w_1 \wedge \cdots \wedge w_\ell].$$

In terms of coordinates,

$$W = \langle w_1, \ldots, w_\ell \rangle \in G_\ell(V) \longmapsto p(W) = (p_\alpha(W)) \in \mathbb{P}\Big(\bigwedge^\ell V\Big)$$

where $A_W = (a_{ij})$ is the $\ell \times m$ matrix whose rows are (the coordinates of) a basis of W and $p_\alpha(W)$ is the α^{th} minor of A_W, viz., $\det\left(a_{i\alpha_j}\right)_{1\leqslant i,j\leqslant \ell}$.

Some Basic Facts about Grassmann Varieties

- $G_{\ell,m}$ is a projective algebraic variety given by the common zeros of certain quadratic homogeneous polynomials in $\binom{m}{\ell}$ variables. As a projective algebraic variety $G_{\ell,m}$ is nondegenerate, irreducible, smooth, and rational.

Some Basic Facts about Grassmann Varieties

- $G_{\ell,m}$ is a projective algebraic variety given by the common zeros of certain quadratic homogeneous polynomials in $\binom{m}{\ell}$ variables. As a projective algebraic variety $G_{\ell,m}$ is nondegenerate, irreducible, smooth, and rational.
- When $\mathbb{F} = \mathbb{F}_q$,

$$|G_{\ell,m}(\mathbb{F}_q)| = \begin{bmatrix} m \\ \ell \end{bmatrix}_q := \frac{(q^m-1)(q^m-q)\dots(q^m-q^{\ell-1})}{(q^\ell-1)(q^\ell-q)\dots(q^\ell-q^{\ell-1})} = \sum_{\nu=0}^{\delta} b_\nu q^\nu$$

where $\delta := \ell(m-\ell)$, and b_ν are positive integers. In fact, for $0 \leqslant \nu \leqslant \delta$, $b_\nu = \dim H^{2\nu}(G_{\ell,m}, \mathcal{C})$. Moreover, b_ν is the number of partitions of ν into at most ℓ parts, each part $\leqslant m-\ell$, Note also that

$$\lim_{q\to 1} \begin{bmatrix} m \\ \ell \end{bmatrix}_q = \binom{m}{\ell}.$$

Grassmann code

Fix $1 \leqslant \ell \leqslant m$. The Grassmann code $C(\ell, m)$ is the $[n, k]_q$-code corresponding to the projective system $G_{\ell,m}(\mathbb{F}_q) \hookrightarrow \mathbb{P}^{k-1}$, where

$$n = \begin{bmatrix} m \\ \ell \end{bmatrix}_q = \frac{(q^m - 1)(q^m - q) \ldots (q^m - q^{\ell-1})}{(q^\ell - 1)(q^\ell - q) \ldots (q^\ell - q^{\ell-1})} \quad \text{and} \quad k = \binom{m}{\ell}$$

Grassmann code

Fix $1 \leqslant \ell \leqslant m$. The Grassmann code $C(\ell, m)$ is the $[n, k]_q$-code corresponding to the projective system $G_{\ell,m}(\mathbb{F}_q) \hookrightarrow \mathbb{P}^{k-1}$, where

$$n = \begin{bmatrix} m \\ \ell \end{bmatrix}_q = \frac{(q^m - 1)(q^m - q) \dots (q^m - q^{\ell-1})}{(q^\ell - 1)(q^\ell - q) \dots (q^\ell - q^{\ell-1})} \quad \text{and} \quad k = \binom{m}{\ell}$$

- $d\left(C(\ell, m)\right) = q^\delta$ where $\delta := \ell(m - \ell)$. [Ryan (1990), Nogin (1996)]
- The GHWs $d_r\left(C(\ell, m)\right)$ are known for $1 \leqslant r \leqslant \mu$ and $k - \mu \leqslant r \leqslant k$, where $\mu := 1 + \max\{\ell, m - \ell\}$. [Nogin (1996), G-Lachaud (2000), Hansen-Johnsen-Ranestad (2007)]. More is known when $\ell = 2$.
- Automorphism group(s) of $C(\ell, m)$ are known. [G-Kaipa (2013)]
- $d\left(C(\ell, m)^\perp\right) = 3$ and $C(\ell, m)^\perp$ is generated by its minimum weight codewords. [Beelen-Piñero (2016)]
- A majority-logic decoder that can correct (asymptotically) up to $d/2^{\ell+1}$ errors is known. [Beelen-Singh (2021)]

Codes associated to (generalized) flag varieties

Fix an m-dimensional vector space V and a sequence

$$\underline{\ell} = (\ell_1, \ldots, \ell_s) \in \mathbb{Z}^s \text{ with } 0 \leq \ell_1 \leq \cdots \leq \ell_s < m.$$

We can consider the variety of (generalized) partial flags:

$$\mathcal{F}_{\underline{\ell}}(V) := \{(V_1, \ldots, V_s) : V_1 \subseteq \cdots \subseteq V_s, \ \dim V_i = \ell_i, \ i = 1, \ldots, s\}.$$

Thanks to Plücker and Segre,

$$\mathcal{F}_{\underline{\ell}}(V) \hookrightarrow \prod_{i=1}^{s} G_{\ell_i,m} \hookrightarrow \prod_{i=1}^{s} \mathbb{P}^{k_i-1} \hookrightarrow \mathbb{P}^{(k_1\cdots k_s)-1} = \mathbb{P}(\bigotimes_{i=1}^{s} \bigwedge^{\ell_i} V) =: T_{\underline{\ell}}(V)$$

where $k_i = \binom{m}{\ell_i}$. As before,

$$\mathcal{F}_{\underline{\ell}}(V)\,(\mathbb{F}_q) \rightsquigarrow [n_{\underline{\ell}}, k_{\underline{\ell}}]_q\text{-code } C(\underline{\ell}; m).$$

We may call $C(\underline{\ell}; m)$ as the flag code or s-step flag code corresponding to $\underline{\ell}$.

Flag varieties and associated codes

- The variety $\mathcal{F}_{\underline{\ell}}(V)$ of (generalized) partial flags can also be viewed as

$$\mathcal{F}_{\underline{\ell}}(V) = \left\{ \left[\bigotimes_{i=1}^{s} v_1 \wedge \cdots \wedge v_{\ell_i} \right] : v_1, \ldots, v_{\ell_s} \in V \text{ linearly indep.} \right\}.$$

- $\mathcal{F}_{\underline{\ell}}(V)$ is a projective variety of dimension

$$\delta(\underline{\ell}) := \sum_{i=1}^{s} (\ell_i - \ell_{i-1})(m - \ell_i).$$

- Let

$$T_{\underline{\ell}}(V) = \bigotimes_{i=1}^{s} \bigwedge^{\ell_i} V \quad \text{and} \quad T^*_{\underline{\ell}}(V) = \bigotimes_{i=1}^{s} \bigwedge^{m-\ell_i} V \simeq T_{\underline{\ell}}(V)^*$$

- The code $C(\underline{\ell}; m)$ corresponding to $\mathcal{F}_{\underline{\ell}}(V)$ can be viewed as the image of

$$\mathrm{Ev} : T^*_{\underline{\ell}}(V) \to \mathbb{F}_q^{n_{\underline{\ell}}} \quad \text{given by} \quad \mathrm{Ev}(f) := \big(f(P_1), \ldots, f(P_{\underline{\ell}})\big)$$

where $P_1, \ldots, P_{\underline{\ell}}$ are fixed representatives in $T_{\underline{\ell}}(V)$ of points of $\mathcal{F}_{\underline{\ell}}(V)(\mathbb{F}_q)$.

Special partial flags: Line-Hyperplane Incidence

Theorem (Rodier, 2003)

If $s = 2$ *and* $\underline{\ell} = (1, m-1)$, *then*

$$n_{\underline{\ell}} = \frac{(q^m - 1)(q^{m-1} - 1)}{(q-1)^2} \quad \text{and} \quad k_{\underline{\ell}} = m^2 - 1.$$

Moreover,

$$d(C(\underline{\ell}; m)) = q^{2m-3} - q^{m-2} = q^{m-2}(q^{m-1} - 1).$$

Theorem (Hana, 2010)

$$q^{\delta(\underline{\ell})} \left(\frac{q-1}{q}\right)^{\ell_s - 1} \leq d(C(\underline{\ell}; m)) \leq q^{\delta(\underline{\ell})}.$$

Further, if $s = 2$, $\ell_1 < \ell_2$ *and* $\ell_1 + \ell_2 \leq m$, *then*

$$d(C(\underline{\ell}; m)) \leq q^{\ell_2(m-\ell_2)-\ell_1^2}(q^{\ell_2} - 1)(q^{\ell_2} - q) \cdots (q^{\ell_2} - q^{\ell_1 - 1}).$$

The length $n_{\underline{\ell}}$ of $C(\underline{\ell}, m)$

$$n_{\underline{\ell}} = \begin{bmatrix} m \\ \ell_1,\ \ell_2 - \ell_1, \ldots, \ell_{s+1} - \ell_s \end{bmatrix} = \prod_{i=1}^{s} \begin{bmatrix} m - \ell_{i-1} \\ \ell_i - \ell_{i-1} \end{bmatrix}$$

where, by convention, $\ell_0 := 0$ and $\ell_{s+1} := m$.

Equivalently, the length $n_{\underline{\ell}}$ is given by

$$n_{\underline{\ell}} = \sum_{\sigma \in W_{\underline{\ell}}} q^{\mathrm{inv}(\sigma)} = \sum_{\tau \in M_{\underline{\ell}}} q^{\mathrm{inv}(\tau)}$$

where $W_{\underline{\ell}}$: permutations $\sigma \in S_m$ satisfying

$$\sigma(\ell_{i-1}+1) < \sigma(\ell_{i-1}+2) < \cdots < \sigma(\ell_i),$$

for $i = 1, \ldots, s+1$, and $M_{\underline{\ell}}$: permutations of the multiset

$$\{1^{\ell_1}, 2^{\ell_2 - \ell_1}, \ldots, s^{\ell_s - \ell_{s-1}}, (s+1)^{m - \ell_s}\}$$

and inv denotes the number of inversions.

See, for example, [G-Lachaud, 2002].

The case of 2-step flags

- For $0 \leq t \leq m$, set $\mathbb{I}(t,m) := \{I = (i_1, \ldots, i_t) : 1 \leq i_1 < \ldots < i_t \leq m\}$.
- Fix a basis $\{e_1, \ldots, e_m\}$ of V. For $I = (i_1, \ldots, i_t) \in \mathbb{I}(t,m)$, put $e_I := e_{i_1} \wedge \cdots \wedge e_{i_t} \in \wedge^t V$.
- For any $0 \leq t_2 \leq t_1 \leq m$ and $I \in \mathbb{I}(t_1,m)$ and $J \in \mathbb{I}(t_2,m)$, define

$$I = (i_1, \ldots, i_{t_1}) \leq J = (j_1, \ldots, j_{t_2}) \iff i_r \leq j_r \text{ for } r = 1, \ldots, t_1$$

Theorem (G, Singh, Piñero, 2017)

Let $s = 2$. A basis for the flag code $C(\underline{\ell}; m) = \mathrm{Ev}(T^*_{\underline{\ell}}(V))$ is given by

$$\{\mathrm{Ev}(e_I \otimes e_J) : (I,J) \in \mathbb{I}(m-\ell_1, m) \times \mathbb{I}(m-\ell_2, m),\ I \leq J\}.$$

Consequently, the dimension of the 2-step flag code $C(\underline{\ell}; m)$ is given by

$$k_{\underline{\ell}} = \binom{m}{\ell_1}\binom{m}{\ell_2} - \binom{m}{\ell_1 - 1}\binom{m}{\ell_2 + 1}.$$

Minimum Distance for 2-step flags

Lemma

Assume that $s = 2$ *and* $\ell_1 + \ell_2 \leq m$. *Then*

$$d(C(\underline{\ell}; m)) \leq q^{\ell_2(m-\ell_2)-\ell_1^2} |GL_{\ell_1}(\mathbb{F}_q)||G_{\ell_1,\ell_2}(\mathbb{F}_q)|.$$

Theorem (G, Singh, Piñero, 2017)

Assume that $s = 2$ *and* $\ell_1 + \ell_2 \leq m$. *Write* $\underline{\ell} = (\ell_1, \ell_2)$ *and* $\underline{\ell'} = (\ell_1, \ell_1)$. *If the minimum distance of the code* $C(\underline{\ell'}; m)$ *is* $q^{\ell_1(m-\ell_1)-\ell_1^2}|\mathrm{GL}_{\ell_1}(\mathbb{F}_q)|$, *then*

$$d(C(\underline{\ell}; m)) = q^{\ell_2(m-\ell_2)-\ell_1^2} |GL_{\ell_1}(\mathbb{F}_q)||G_{\ell_1,\ell_2}(\mathbb{F}_q)|.$$

Corollary

If $s = 2$, $1 \leq \ell < m$, *and* $\underline{\ell} = (1, \ell)$, *then*

$$d(C(\underline{\ell}; m)) = q^{\ell(m-\ell)-1}(q^\ell - 1).$$

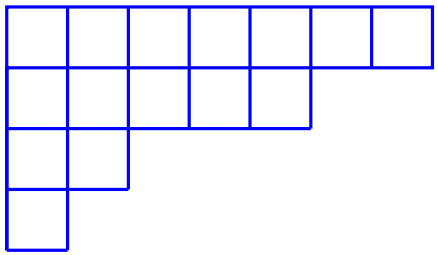
A General Formula for the dimension $k_{\underline{\ell}}$ of $C(\underline{\ell}, m)$

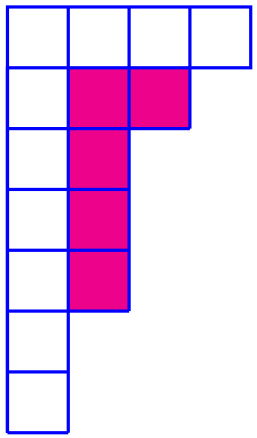
Given $\underline{\ell} = (\ell_1, \ldots, \ell_s)$, consider the partition

$$m - \ell_1 \geq m - \ell_2 \geq \cdots \geq m - \ell_s \geq 1$$

and let λ be the conjugate partition.

For example if $m = 8$ and $\underline{\ell} = (1, 3, 6, 7)$, then the associated partition is $(7, 5, 2, 1)$ and the conjugate partition is $\lambda = (4, 3, 2, 2, 2, 1, 1)$. These partitions can be viewed as folllows.





Description of $k_{\underline{\ell}}$ Contd.

For each box (i,j) in the (Young diagram of) λ, let $h_{(i,j)}$ be the hook length at (i,j), that is, the number of boxes in the hook at (i,j). For example the hook at $(2,2)$ in the partition $\lambda = (4,3,2,2,2,1,1)$ is shown by shaded boxes and we have $h_{(2,2)} = 5$.

Theorem

The dimension $k_{\underline{\ell}}$ of $C(\underline{\ell}, m)$ is given by

$$k_{\underline{\ell}} = \prod_{(i,j)\in\lambda} \frac{m+j-i}{h_{(i,j)}}.$$

Idea: Use the connection between flag varieties and representations of linear groups together with classical results from Combinatorial Representation Theory. However, one should be careful in applying classical results to the case of finite ground field $\mathbb{F}_q$; in particular, it is better to assume $s \leq q$.

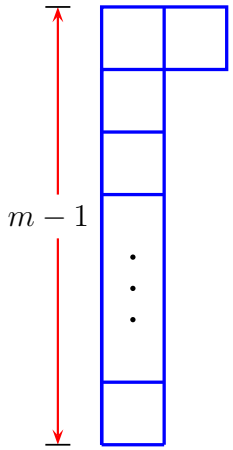
Example (2-Step Flag Code of Rodier): If $\underline{\ell} = (1, m-1)$, then

$$\lambda = \text{conjugate of } (m-1, 1) = (2, \underbrace{1, 1, \ldots, 1}_{m-2 \text{ times}}).$$

Hence by the above formula

$$\begin{aligned} k_{\underline{\ell}} &= \frac{m(m+1)(m-1)(m-2)\cdots(m-(m-2))}{m(1)(m-2)(m-3)\cdots 1} \\ &= (m+1)(m-1) = m^2 - 1, \end{aligned}$$

as is to be expected.



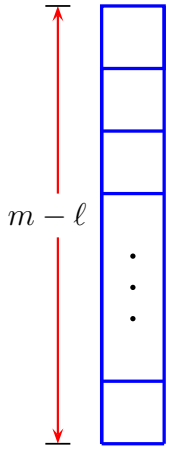
Another Example (Grassmann codes): If $\underline{\ell} = (\ell)$, then

$$\lambda = \text{conjugate of } (m-\ell) = (\underbrace{1, 1, \ldots, 1}_{m-\ell \text{ times}}).$$

Hence by the above formula

$$\begin{aligned} k_{\underline{\ell}} &= \frac{m(m-1)(m-2)\cdots(m-(m-\ell)+1)}{(m-\ell)(m-\ell-1)\cdots 1} \\ &= \frac{m!}{\ell!(m-\ell)!} = \binom{m}{\ell} \end{aligned}$$

as is to be expected.



Thank you!