A SHORT PROOF THAT ALL LINEAR CODES ARE WEAKLY ALGEBRAIC-GEOMETRIC USING A THEOREM OF B. POONEN

SRIMATHY SRINIVASAN

ABSTRACT. In this paper we give a simpler proof of a deep theorem proved by Pellikan, Shen and van Wee that all linear codes are weakly algebraic-geometric using a theorem of B.Poonen.

1. INTRODUCTION

Coding theory, the study of designing efficient codes that help in reliable data transmission, is an integral component of any communication system. Let $q$ be some power of a prime number. The basic idea is to “encode” the $q$-ary message to be transmitted by adding redundant information to immunize against errors and “decode” at the receiver thereby achieving reliability.

One way to encode is by linearly embedding the $k$ bit message space $\mathbb{F}_q^k$ into $\mathbb{F}_q^n$ using a $k \times n$ matrix called the generator matrix and transmit the vectors in the image of the embedding. We call the image a code and the vectors in it codewords. Minimum distance $d$, the minimum number of positions in which any two distinct codes differ, is an important parameter of a code that measures its error correcting capability. Typically, we need codes with relative parameters $\delta = \frac{d}{n}$ and $R = \frac{k}{n}$ to be as large as possible. In fact, one of the important problems in coding theory is to construct a sequence of codes with $n \to \infty$ such that the limit $(\delta, R)$ of relative parameters is non-zero. Such a sequence of codes is called a sequence of good codes.

It is well known that the maximal achievable $R$ for a given $\delta$, denoted as $\alpha_q(\delta)$, is a continuous decreasing function of $\delta$. Although the exact function $\alpha_q(\delta)$ is not known, there are many lower bounds for it. One significant bound is the Gilbert-Varshamov bound which was known to be the best until algebraic-geometric codes were discovered. These codes are constructed by picking rational points on smooth projective curves and evaluating global sections of a suitable line bundle at these points. The parameters of these codes are easily estimated using the Riemann-Roch theorem and good parameters can be obtained by choosing good curves with many rational points. One such example of a family of good curves is the family of modular curves. These curves have a lot of rational points thereby yielding a good family of codes that beat the Gilbert-Varshamov bound.

One can generalize the construction of codes on curves to codes on surfaces or any other higher dimensional variety in an analogous manner: Pick a set of $\mathbb{F}_q$ rational points on the variety and evaluate the global sections of a line bundle at these points to get a linear code. In this case, estimating the parameters of the code is harder than in the case of curves as we need Riemann-Roch in higher dimensions. In this paper we show that every linear code can be realized on a variety obtained by blowing up projective space at a finite set of points. Then we show that codes realized over smooth geometrically integral varieties can be realized as codes on curves using Bertini theorems of B.Poonen ([5], [6]) over finite fields. This gives an alternate proof of a deep theorem of Pellikan, Shen and van Wee ([4]) that every linear code is weakly algebraic-geometric (see Definition 2).

2. CODING THEORY BACKGROUND

An $(n, k, d)_q$-code is a $k$-dimensional subspace of an $n$-dimensional vector space over $\mathbb{F}_q$. The vectors in a code are called codewords. The parameter $d$ is called the minimum distance. It is the minimum of Hamming distances between any two distinct codewords. Since a code is a linear subspace, it is also
equal to the minimum of Hamming weights of all the non-zero codewords. Finding minimum distance given a basis for the code is NP-hard. For practical purposes we need the dimension \( k \) and the minimum distance \( d \) of a code to be large as possible for a given length. So for an \((n,k,d)\) - code \( C \), we define the relative dimension or code rate, \( R(C) = \frac{k}{n} \) and relative minimum distance, \( \delta(C) = \frac{d}{n} \). The pair \((\delta(C), R(C))\) denotes a point in \([0, 1] \times [0, 1]\). A sequence of codes \( \{C_i\} \) is said to be asymptotically good if \( \lim_{i \to \infty} \delta(C_i) > 0 \) and \( \lim_{i \to \infty} R(C_i) > 0 \).

Define,

\[
U_q = \{(\delta, R) \mid \text{there exists a sequence of codes } \{C_i\} \text{ with } \lim_{i \to \infty} (\delta(C_i), R(C_i)) = (\delta, R)\}
\]

Then, it is know that there is a continuous decreasing function \( \alpha_q(\delta) \) such that \( U_q = \{(\delta, R) \mid 0 \leq R \leq \alpha_q(\delta)\} \). The exact function \( \alpha_q(\delta) \) is unknown although few upper and lower bounds for the function is known. One lower bound is given by the Gilbert-Varshamov(GV) bound:

\[
\alpha_q(\delta) \geq 1 - (\delta \cdot \log_q(q - 1) - \delta \cdot \log_q(\delta) - (1 - \delta) \cdot \log_q(1 - \delta))
\]

It was the best known bound for many years until the invention of algebraic-geometric (AG) Codes. When \( q \) is a square, one can construct algebraic-geometric codes on curves with good asymptotic behavior. This yields the algebraic geometry (AG) bound,

\[
\alpha_q(\delta) \geq 1 - \frac{1}{\sqrt{q} - 1} - \delta
\]

which beats the Gilbert-Varshamov bound in a region along \( \delta \) for \( q \geq 49 \).

3. ALGEBRAIC-GEOMETRIC CODES

Algebraic-geometric codes were first discovered by Goppa ([2]) and was further developed by Tsfasman,Vladut ([7]) and many others along the way. We start by briefly recalling the construction of algebraic-geometric codes. The text [3] by Hartshorne is a good reference for all the basic algebraic geometry and notations we use in this paper.

**Definition 1.** Let \( X \) be a smooth projective variety defined over \( \mathbb{F}_q \) and let \( \mathcal{P} = \{P_1, P_2, \cdots, P_n\} \subseteq X(\mathbb{F}_q) \). Let \( D \) be a divisor on \( X \) such that the support of \( D \) is disjoint from \( \mathcal{P} \). Define

\[
L(D) = \{f \in \mathbb{F}_q(X)^* | (f) + D \geq 0\} \cup \{0\}
\]
and consider the evaluation map:

\[ Ev_P : L(D) \rightarrow \mathbb{F}_q^n \]

\[ f \mapsto [f(P_1), f(P_2), \ldots, f(P_n)] \]

The image of the map gives a linear code \( C = C_L(X, \mathcal{P}, D) \) and we say that \( C \) is an algebraic-geometric code realized over \( X \).

Given the data \( X, \mathcal{P}, D \) as above, let \( \mathcal{L} \) denote the line bundle associated to \( D \) and \( H^0(X, \mathcal{L}) \) denote its global sections. Then we can get a code \( C(X, \mathcal{P}, \mathcal{L}) \) equivalent to \( C_L(X, \mathcal{P}, D) \) as follows. First note that the local ring \( \mathcal{L}_{P_i} \) modulo the maximal ideal of sections vanishing at \( P_i \) denoted by \( \mathcal{L}_{P_i} \) is isomorphic to \( \mathbb{F}_q \) by a choice of local trivialisation. Then the image of the germ map

\[ \alpha_P : H^0(X, \mathcal{L}) \rightarrow \bigoplus \mathcal{L}_{P_i} \cong \mathbb{F}_q^n \]

gives a linear code \( C(X, \mathcal{P}, \mathcal{L}) \) that is same as the code \( C_L(X, \mathcal{P}, D) \) up to monomial equivalence.

**Remark:** In Definition 1, if \( X \) is a smooth curve, we get the original construction of Goppa (Goppa codes [2]). In this case, the parameters of the code are easily estimated using the Riemann-Roch theorem. However, it is not so easy for codes over higher dimensional varieties as invoking Riemann-Roch brings higher cohomology groups come into picture.

## 4. All Linear Codes are Weakly Algebraic-Geometric

In this section we give a shorter proof of a theorem of Pellikkan, Shen, van Wee ([4]). Recall the definition of weakly algebraic-geometric code from [4]:

**Definition 2.** A q-ary linear code \( C \) is said to be weakly algebraic geometric if there exists a projective non-singular absolutely irreducible curve \( X \) defined over \( \mathbb{F}_q \), \( n \) distinct points \( \mathcal{P} = \{ P_1, P_2, \ldots, P_n \} \) on \( X \) and a divisor \( D \) with support disjoint from \( \mathcal{P} \) such that \( C = C_L(X, \mathcal{P}, D) \).

We now show that algebraic-geometric codes are ubiquitous in the sense that every linear code can be realized over some smooth variety. In fact we have the following stronger result.

**Theorem 1.** Let \( C \) be a linear code. Then \( C = C_L(X, \mathcal{P}, D) \) where \( X \) is the blow up of some projective space at finitely many points, \( \mathcal{P} \) is a finite set of distinct \( \mathbb{F}_q \)-points in \( X \) and \( D \) is a divisor such that the support of \( D \) is disjoint from \( \mathcal{P} \).

**Proof.** Let \( C \) be a \((n, k, d)_q\) linear code with \( k \times n \) generator matrix \( G \). Then the columns \( C_1, C_2, \ldots, C_n \) of \( G \) form (not necessarily distinct) points of \( \mathbb{A}^k \). Then we can find an integer \( r \geq 2 \) and \( n \) distinct points \( P_1, P_2, \ldots, P_n \) in \( \mathbb{A}^{r+k} \) such that the projection map

\[ \phi : \mathbb{A}^{r+k} \rightarrow \mathbb{A}^k \]

\[ [y_1, y_2, \ldots, y_r, x_1, x_2, \ldots, x_k] \rightarrow [x_1, x_2, \ldots, x_n] \]

takes \( P_i \) to \( C_i \). Let \( y_0, y_1, \ldots, y_r, x_1, x_2, \ldots, x_k \) denote the coordinates of \( \mathbb{P}^{r+k} \). Identify \( \mathbb{A}^{r+k} \) with the open affine set \( y_0 = 1 \) in \( \mathbb{P}^{r+k} \). For \( 1 \leq i \leq r \), let \( V_i \) denote the point in \( \mathbb{P}^{r+k} \) with \( y_0 = y_i = 1 \) and all other coordinates 0. By choosing \( r \) large enough we can assume that \( V_i \neq P_j \ \forall i, j \). Let \( X \) be the smooth geometrically integral variety obtained via the blow up \( \pi : X \rightarrow \mathbb{P}^{r+k} \) at the points \( V_i \) with the corresponding exceptional divisor \( E_i \). Denote by \( H \) the hyperplane section \( y_0 = 0 \) in \( \mathbb{P}^{r+k} \). Then the global sections of the line bundle associated to the divisor \( D = \mathcal{L}(\pi^*H - \sum_i E_i) \) is generated by \( x_0, x_1, \ldots, x_k \). It is easy to see that code \( C = C(X, \mathcal{P}, D) \) where \( \mathcal{P} \) is the set \( \{ \pi^{-1}P_1, \pi^{-1}P_2, \ldots, \pi^{-1}P_n \} \). \( \square \)
Let us now restate the results on Bertini theorems over finite fields due to B. Poonen. We refer the reader to Theorem 1.1 in [6] and remarks below Theorem 3.3 in [5] for more details.

**Theorem 2** (Poonen). Let $X$ be a smooth, projective geometrically integral variety of $\mathbb{P}^n$ of dimension $m \geq 2$ over $\mathbb{F}_q$, and let $\mathcal{P} \subset X$ be a finite set of closed points. Then, given any integer $d_0$, there exists a hypersurface $H \subset \mathbb{P}^n$ of degree $d \geq d_0$ such that $Y = H \cap X$ is smooth, projective and geometrically integral of dimension $m - 1$ and contains $\mathcal{P}$.

**Example.** Consider $X = \mathbb{P}^2$ over $\mathbb{F}_2$. Let $\mathcal{P}$ be the set of all 7 $\mathbb{F}_2$-points. Then, the curve $Y = y^3z + y^2z^2 + x^2y^2 + x^3z$ is a smooth curve passing through $\mathcal{P}$. In fact, one can show that there are 24 smooth curves of degree 4 passing through $\mathcal{P}$.

Using Theorem 2 we now show that codes realized over higher dimensional varieties can be realized over curves. Although there are papers [1] that apply Poonen’s theorem to algebraic-geometric codes, there does not seem to be any literature that states this result.

**Theorem 3.** Let $\mathcal{C} = C(X, \mathcal{P}, \mathcal{L})$ be a code on a geometrically integral smooth projective variety $X \subseteq \mathbb{P}^k$ of dimension $m \geq 2$ over $\mathbb{F}_q$. Then $\mathcal{C}$ can be realized over a smooth projective geometrically integral curve. In particular, there exists a geometrically integral smooth projective curve $Z$ containing $\mathcal{P}$ such that $\mathcal{C} = C(Z, \mathcal{P}, \mathcal{L}|_Z)$.

**Proof.** Let $Y = H \cap X$ be as Theorem 2 with degree $d$ of $H$ large enough and let $i : Y \hookrightarrow X$ denote the inclusion morphism. Then, we have the following short exact sequence on $X$:

$$0 \longrightarrow i_*\mathcal{O}_Y \longrightarrow \mathcal{O}_X \longrightarrow \mathcal{I}_Y \longrightarrow 0$$

where $\mathcal{I}_Y = \mathcal{O}_X(-d)$. Tensoring with $\mathcal{L}$ we get

$$0 \longrightarrow \mathcal{L} \otimes \mathcal{I}_Y \longrightarrow \mathcal{L} \longrightarrow \mathcal{L} \otimes i_*\mathcal{O}_Y \longrightarrow 0$$

(Here tensoring is over $\mathcal{O}_X$). The above short exact sequence gives rise to a long exact sequence in cohomology on $X$

$$0 \longrightarrow H^0(\mathcal{L} \otimes \mathcal{I}_Y, X) \longrightarrow H^0(\mathcal{L}) \longrightarrow H^0(\mathcal{L} \otimes i_*\mathcal{O}_Y, X) \longrightarrow H^1(\mathcal{L} \otimes \mathcal{I}_Y, X) \longrightarrow \cdots$$

By duality, we have

$$H^i(\mathcal{L} \otimes \mathcal{I}_Y, X) \simeq H^{m-i}(\omega_X \otimes \mathcal{I}_Y^\vee \otimes \mathcal{L}^\vee, X) \simeq H^{m-i}(\omega_X \otimes \mathcal{L}^\vee \otimes \mathcal{O}_X(d), X)$$

where $\omega_X$ is the canonical sheaf on $X$. Since $\mathcal{O}_X(1)$ is ample, for large enough $d$, $H^0(\mathcal{L} \otimes \mathcal{I}_Y, X)$ and $H^1(\mathcal{L} \otimes \mathcal{I}_Y, X)$ vanishes and we get a canonical isomorphism obtained via restriction

$$H^0(\mathcal{L}) \longrightarrow H^0(\mathcal{L} \otimes i_*\mathcal{O}_Y, X) \simeq H^0(\mathcal{L}|_Y \otimes \mathcal{O}_Y, Y) \simeq H^0(\mathcal{L}|_Y, Y)$$

Inducting the above argument by replacing the $m$-dimensional variety $X$ with $(m - 1)$-dimensional variety $Y$ and $\mathcal{L}$ with $\mathcal{L}|_Y$ we get the result.

Hence, given a code over a geometrically integral smooth projective variety, we have realized it over a geometrically integral smooth projective curve. As a consequence we get:

**Corollary 4** (Pellikan, Shen, van Wee). All linear codes are weakly algebraic-geometric.

**Proof.** This easily follows from Theorem 1 and Theorem 3.

ACKNOWLEDGEMENTS. I would like to thank Prof. Patrick Brosnan and Prof. Lawrence Washington for their valuable comments and discussions. I would also like to thank Richard Rast for writing a computer program to verify some results.
REFERENCES


DEPARTMENT OF MATHEMATICS, 1301 MATHEMATICS BUILDING, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742-4015, USA