Linear codes from Schubert Varieties

James S. Wolper
Department of Mathematics
Idaho State University
Pocatello, ID 83209
wolperj@pequod.isu.edu

12 January 1996

Abstract: A family of linear "Algebraic–Geometric" codes is constructed from line bundles on Schubert varieties and analyzed using techniques from Representation Theory.

AMS(MOS) 1991 Subject Classification (Primary) 94B27 (Secondary) 14M99, 20G10.

Research partially supported by the US Air Force Office of Scientific Research.

Research partially supported by the Isaac Newton Institute for Mathematical Sciences, University of Cambridge, UK. Preprint number NI96001.

1. Introduction

Recent work in the constructions of Error–Correcting Codes has exploited techniques from algebraic geometry, especially from the geometry of varieties over a finite field; [TV] contains a complete description of this work. Most of the codes constructed from varieties are of the *Reed–Müller* or *Goppa* type. In these, one starts with a variety X over $GF(q)$, the field with q elements, and chooses a subset $P \subset X$ and a line bundle (locally free sheaf of rank 1) $L \to X$. There is a linear transformation $H^0(X, L) \to V$, where V is a $GF(q)$ vector space of dimension $|P|$, essentially defined by "evaluating" all of the sections in $H^0(X, L)$ (Čech cohomology) at points of P. The codewords are the image of $H^0(X, L)$ under this map. (The value of a section is not well–defined, so a choice must be made in evaluation.) Høholdt, van Lint, and Pellikaan [HLP] have presented a general theory of such codes.

Most of the codes so constructed have used a curve X; this is natural since so much is known about the Picard group (= group of line bundles) of a curve. Comparatively little work has been done with varieties of higher dimension, notable examples being the work of Chakravarti [C] and Ryan and Ryan [RR]. In the current work, a family of codes is constructed from Schubert subvarieties of flag manifolds G/B, where G is a semisimple algebraic group defined over $GF(2)$ (ie, a Chevalley group) and B is a Borel subgroup. Line bundles over G/B correspond to irreducible representations of G. Each such a representation can, by the Bott–Borel–Weil Theorem ([D]), be constructed as $H^0(G/B, L_\chi)$ where L_χ is the line bundle associated to some character χ of B. The dimension of $H^0(G/B, L_\chi)$ (*ie*, the degree of the representation) can be determined from the Weyl Character Formula, and many of the properties of these codes can be deduced from the combinatorial machinery of representation theory.

Before the reader finds the abstract nature of the derivation too daunting, it should be pointed out that most of the contents reduce to simple linear algebra computations, mostly taking determinants of minors of matrices. The general picture is useful, however, because the best of these codes come from the Schubert cells, and because it shows that this is a very large family of codes.

Here is an outline of the rest of this paper. The next section contains a brief résumé of results about algebraic groups; all of this material is standard. Section 3 discusses Schubert varieties and is also standard. In section 4, the codes are constructed, and various parameters are estimated. It works out that the most interesting codes are constructed from fundamental representations of $SL(n)$, and codes so constructed are discussed in section 5; we obtain good information about the parameters. Section 6 has some examples; the examples are easier than the theory! The final section describes some unanswered questions.

2. Algebraic Groups

In this section, k is a field of any characteristic. The material here can be found in [Bo].

A *Linear Algebraic Group G* over k is a (reduced, irreducible) variety over k which is also a group; the group operations are k–morphisms; and there exists an embedding $G \to GL(n, k)$ for some n. A torus T is a linear algebraic group which is isomorphic to a product of $GL(1, k)$s. Suppose T is a *maximal* torus in G (such T exist for dimension reasons). A subgroup $B \subset G$ is a *Borel* subgroup if it is maximal among the connected, solvable subgroups. It is a fundamental fact that all Borel subgroups are conjugate. The quotient space G/B is the *flag variety*. It is smooth. The *Weyl group* is $W = N(T)/T$, where $N(T)$ is the normalizer of T in G. If G is semisimple, W is finite.

When $G = SL(n, k)$ or $GL(n, k)$, take T to be the subgroup of diagonal matrices, and B to be the subgroup of upper triangular matrices. The Weyl group is isomorphic to the symmetric group S_n of permutations of $\{1, \ldots, n\}$. Here is an explanation for the name: let $G = SL(n, k)$, and let V be a vector space of dimension n over k. Then a *flag* in V is a sequence of subspaces $V_1 \subset V_2 \subset \ldots \subset V_n = V$, where $\dim(V_i) = i$. Then G acts transitively on the space of all flags in V with isotropy B.

The representation theory of G is intricate and similar to the representation theory of Lie groups; it was worked out by Chevalley and others ([Ch]). The book [H] is an excellent introduction. It is easiest to describe in terms of the Lie algebra $\underline{\mathrm{g}}$ of G, which is, as a vector space, the tangent space to G at the identity element; the exponential mapping $\exp{:}\underline{\mathrm{g}} \to G$ enables one to go back-and-forth between them. Let $\underline{\mathrm{h}}$ be a

Cartan subalgebra of $\underline{g}$, that is, the Lie algebra of a maximal torus. Let $\rho : \underline{g} \to GL(n, k)$ be a representation on a vector space V of dimension n. A *weight* for ρ is function $\alpha : \underline{h} \to k$ such that $\{v \in V : \rho(h).v = \alpha(h).v$ for all $h \in \underline{h}\}$ is non–empty. It is a theorem that every irreducible representation of $\underline{g}$ (and hence of G) has a unique "highest weight" (highest with respect to a well-defined order on the space of characters). Similarly, given a so–called dominant, integral character of T, its derivative is a weight and is, in fact, the highest weight of some representation of $\underline{g}$. The dimension of such a representation can be determined by the *Weyl character formula* which we will discuss in an *ad hoc* manner only.

Each $\underline{g}$ comes equipped with a representation on $\underline{g}$, called the *adjoint* representation. There is a non–commutative product, the *bracket* $[a, b]$ on $\underline{g}$, and the adjoint representation is defined by $\mathrm{ad}(x)(y) = [x, y]$. A weight of the adjoint representation is called a *root*. The classification of G is based on Dynkin diagrams constructed from the roots: a node of the diagram corresponds to a "simple" root, nodes are connected by edges (with multiplicity) depending on the angle (with respect to an appropriate inner product) between the corresponding subspaces in $\underline{h} \otimes \mathbf{R}$. Any dominant integral weight is a sum of roots with positive integer coefficients.

The most important example, both in the general theory and in this work, is the case of $SL(n)$. Then T and B are defined as above, and G/B is the space of full flags. The Lie algebra $\underline{g}$ is the set of $n \times n$ matrices with the bracket $[x, y] = xy - yx$, and $\underline{h}$ is the subalgebra of diagonal matrices. Define $\epsilon_i : \underline{h} \to k$ as the homomorphism taking an element of $\underline{h}$ to its i^{th} diagonal entry. The roots $\alpha_i = \epsilon_i - \epsilon_{i+1}$ form a basis for the root system. The i^{th} *fundamental representation* has highest weight $\epsilon_1 + \ldots + \epsilon_i$. Its exponential takes a diagonal matrix in T to the product of its first i entries (which is nonzero). Such representations are realized on $k^n \wedge \ldots \wedge k^n$ (i factors); the dimension of the representation space is then $\begin{pmatrix} n \\ i \end{pmatrix}$, as can also be determined from the Weyl Character Formula. A weight (for T) is dominant if it has the form $t_1^{m_1} t_2^{m_2} \ldots t_n^{m_n}$ where t_i is the i^{th} diagonal entry of T and the m_i are integers with $m_1 > m_2 > \ldots > m_n$.

In the general case, every character χ of T is a dominant weight, and thus corresponds to a unique (up to isomorphism) representation with highest weight χ; it is a fact that this representation extends to B. There are several ways to construct such representations; we will construct them as the vector space of sections of a certain line bundle on G/B. Define an equivalence relation $\equiv$ on the product $G/B \times k$ by $(xbB, k) \equiv (gB, \chi(b)k)$, where $b \in B$. The quotient is a line bundle L_χ over G/B. The cohomology $H^0(G/B, L_\chi)$ is represented by functions $f : G \to k$ satisfying $f(xb) = \chi(b)^{-1} f(x)$. G acts on this cohomology group, thus defining a representation. (Strictly speaking, when using cohomology, one should consider G defined over the algebraic closure of k, and then restrict to the fixed points of the Frobenius mapping.)

The irreducible representations of G are given the structure of a semigroup under the *Young product*; this semigroup is generated by the fundamental representations. The Young product of the irreducible representations with highest weights χ_1 and χ_2 has highest weight $\chi_1\chi_2$, and the representation space is realized by taking the products of sections in $H^0(G/B, L_{\chi_1})$ and $H^0(G/B, L_{\chi_2})$

A basis for $H^0(G/B, L_{\omega_i})$, where ω_i is the character of the i^{th} fundamental representation of $SL(n)$, is constructed as follows. First, there is a "distinguished" section f_0.

PROPOSITION. Define $f_0 : G/B \to k$ by $f_0(g) :=$ upper left $i \times i$ minor of g. Then f_0 is a section of ω_i.

PROOF. A simple computation. **QED**

PROPOSITION. Suppose w is an element of the Weyl group of G and f is in $H^0(G/B, L_\chi)$. Define $f_w(g) := f(wg)$. Then f_w is in $H^0(G/B, L_\chi)$.

PROOF. Another straightforward computation. **QED**

PROPOSITION. The functions f_w constructed from f_0 above span $H^0(G/B, L_{\omega_i})$ when $G = SL(n)$.

PROOF. The Weyl group acts on sections in the indicated manner. The isotropy of f_0 is the direct product of the subgroup that fixes $\{i + 1, \ldots, n\}$ and the subgroup that fixes $\{1, \ldots, i\}$. This has $i!(n - i)!$

elements, so the size of the orbit is $\frac{n!}{i!(n-i)!}$, which is the dimension of $H^0(G/B, L_{\omega_i})$. The distinct f_w are linearly independent by inspection. **QED**

3. Schubert varieties

There is a decomposition of G into disjoint double cosets BwB, where $w \in W$ (this holds under some mild assumptions which apply in the cases of interest here); this is called the Bruhat decomposition, and it naturally leads to a decomposition of G/B into a disjoint union of affine spaces. Let C_w denote the image of BwB in G/B, and let X_w be its closure. X_w is called a *Schubert variety*. It is a fundamental fact that X_w is a disjoint union of certain C_y, $y \in W$. This defines a partial order (the *Bruhat order*) on the set of Schubert varieties in G/B (and hence on the Weyl group) by $X_1 \leq X_2 \iff X_1 \subset \overline{X}_2$. Proctor ([P]) has defined isomorphic partial orders on integer sequences.

In the example of $SL(n)$, the Weyl group is isomorphic to S_n, and the integer sequences are permutations of $\{1, \ldots, n\}$. The order is generated by the "moves" of exchanging s_i and s_j when $s_i < s_j$ and $i < j$. The dimension of a Schubert cell is the length of the associated permutation.

It is easy to count the number of points in a Schubert variety X_w when k is a finite field, because X_w is a disjoint union of affine spaces. Thus, a cell of dimension i has $|k|^i$ points. To find the cardinality of a variety, add up the cardinalities of its constituent cells. In fact, the dimension of the cell C_w is the length w, which is defined because W is a Coxeter group.

A line bundle L over G/B restricts to a line bundle over a Schubert variety $X_w \subset G/B$. This is intuitive when X_w is non–singular, but many Schubert varieties are singular, and the theory becomes less intuitive (recall that a line bundle in this context is really a locally free sheaf of rank 1). The paper [W] discusses a combinatorial algorithm for deciding if X_w is singular when $G = SL(n)$. In what follows we will restrict ourselves to the cell C_w which is an affine space, so there are no singularities. Note that C_w is dense in X_w.

Many methods for describing Schubert cells have evolved; for our purposes we use the following. Suppose $G = SL_n$. Pick $d = d_1, d_2, \ldots, d_n$, a permutation of the list $1, 2, \ldots, n$. Arrange these in a matrix

$$\begin{array}{cccc} d_{1,1} & & & \\ d_{2,1} & d_{2,2} & & \\ & \cdots & & \\ d_{n,1} & d_{n,2} & \cdots & d_{n,n} \end{array} .$$

Here, $d = \{d_1, \ldots, d_i\} = \{d_{i,1}, \ldots, d_{i,i}\}$, $d_{i,j} < d_{i,j+1}$, and $d_{i,j}$ is the smallest integer such that $\dim(V_i \cap F_{d_{i,j}}) = j$. The dimension of such a variety is

$$\sum_i \#\{j > i : d_j < d_i\}.$$

Some examples of this appear in a later section.

4. Codes from Schubert cells

Assume from now on that k is a finite field. Choose an algebraic group G over k, a Borel subgroup B, and character χ of a maximal torus $T \subset B$. The character χ extends to a character of B. Then we have a line bundle L_χ over G/B. Let $m = \dim\ (H^0(G/B, L_\chi))$, and let $f_1, \ldots, f_m$ be a basis for $H^0(G/B, L_\chi)$. Choose a Schubert cell C_w and let $N = |C_w| = 2^{\mathrm{len}(w)}$. Order the points of X_w in some arbitrary manner $x_1, \ldots, x_N$, and choose arbitrary representatives for the x_i (recall that the x_i are cosets in G/B). Then define the linear code $C(k, G, B, \chi, w)$ as follows.

DEFINITION. Let V be a k–vector space with N points, and define a linear transformation $C : H^0(G/B, L_\chi) \to V$ by $C(f) = (f(x_1), \ldots, f(x_N))$. The code is the image of C.

This is an $[N, \leq m]$ code over k. If the map C is *injective* then the code has parameters $[N, m]$. By [Ch, Exposé 15], the i^{th} fundamental representation has a section whose zero–locus is a certain codimension one Schubert variety $X_{w'}$. The map $C(k, G, B, \chi, w)$ then fails to be injective. Also, if $y < w'$ in the Bruhat order, the map $C(k, G, B, \chi, y)$ fails to be injective. However, this particular failure of injectivity has no effect on the asymptotic results below. Examples presented later show how to pick w so that C is injective.

Recall that the *rate* of an $[N, m, d]$ code is m/N.

CONJECTURE. There exists a family of codes $C(k, G, B, \chi, w)$ whose rate is asymptotically 1.

For example, let $G = SL(n, k)$, let B be the subgroup of upper triangular matrices, and let χ be the character of the i^{th} fundamental representation. Then $m = \binom{n}{i} = O(n^i)$. Perhaps one can choose a w with length $O(i)$ such that C is injective. The examples later illustrate some of the difficulties involved.

Estimating d, the minimum weight of the code, appears more complicated in general. By definition, the weight of the codeword corresponding to the image of a section f is $N - |\{x_i : f(x_i) = 0\}|$.

PROPOSITION. Let ω_i be the minimal weight of the code $C(k, G, B, \chi_i, w)$, $i = 1, 2$, and let ω_{12} be the minimal weight of the code constructed from the Young product of χ_1 and χ_2. Then $\omega_{12} \leq \min(\omega_1, \omega_2)$.

PROOF. The zero-locus of a product of two sections contains the zero-loci of the factors. **QED**

Thus, we focus on the fundamental representations. First, consider the i^{th} fundamental representation of $SL(n)$. Recall that a basis for $H^0(G/B, \chi)$ is given by the determinant of the upper-left $i \times i$ minor of row permutations of g, the matrix representing a flag. In this case, the functions we are evaluating are polynomials of degree i in in variables, namely the entries in the leftmost i columns of the matrix. The minimum distance, then, is bounded below by the minimum distance of the Reed-Müller code $\mathcal{R}(ni, i)$; this is known to be 2^{ni-1}.

Recall that the *relative minimum distance* of an $[N, m, d]$ code is d/N. The argument above (with the example of $G = SL(n, k)$) shows:

PROPOSITION. There exists a family of codes $C(k, G, B, \chi, w)$ whose relative minimum distance is asymptotically greater than or equal to 1/2. **QED**

5. The parameters of $SL(n)$ codes

In this section we consider the problem of finding more precisely the parameters of such a code in the case of a line bundle corresponding to the i^{th} fundamental representation of SL_n. There are some useful simplifications. First, following a suggestion from R. Pellikaan, we restrict attention to an *affine* piece of the Schubert variety, which is isomorphic to an affine space over k. Next, we take advantage of the special data in these cases to reindex the sections.

First, the i^{th} fundamental representation has $\binom{n}{i}$ independent sections, which are *a priori* indexed by a certain set of cosets in the symmetric group. These are evaluated by taking the determinant of a certain $i \times i$ minor of a matrix representing a flag. By examining the specified action of the Weyl group, it is easy to see that this can be simplified to the following procedure: take the first i columns of the matrix. Then each section consists of taking the determinant of the $i \times i$ matrix obtained by selecting i rows $r_1, \ldots, r_i$. Thus, we can index the sections as

$$f_{r_1,\ldots,r_i},$$

with $r_1 < \cdots < r_i$, of which there are obviously $\binom{n}{i}$.

To determine the dimension of the image, we calculate the kernel of the evaluation map

$$f \mapsto (f(p_1), \ldots, f(p_N)).$$

The general case follows from the case $i = 2$ easily, so we will illustrate that case. The key parameter is the second row of the matrix describing the Schubert cell, namely $[d_{21}d_{22}]$. This means that the flags in the cell satisfy $V_1 \subset F_{d_{21}}$ and $V_2 \subset F_{d_{22}}$. Consider the section f_{ij}; if $j > d_{22}$, then the matrix must have a row of zeroes, so f_{ij} is identically zero in this cell, *ie*, f_{ij} is in the kernel. Similarly, if $i > d_{21}$, then the matrix whose determinant we are taking has a column of zeroes. It is also possible to exhibit matrices representing flag for whcih this function does not vanish: namely, put 1 in places $(1, i)$ and $(2, j)$. This proves:

THEOREM. The condition that $f_{r_1,\ldots,r_i}$ *not* be in the kernel is $r_j < d_{ij}$.

COROLLARY. The dimension of the code is $O(d_{i1}d_{i2}\cdots d_{ii})$.

COROLLARY. If $d_{ii} = n$ and $d = [n-i+1,\ldots,n]$ then the evaluation map is injective.

PROOF. Examine the matrix for such a cell: the i^{th} column can have nonzero entries for its whole length, the previous column can have nonzero entries for its whole length except for the last, etc; thus, one of the matrices representing a flag in this cell can have the $i \times i$ identity matrix as its lower left-hand corner. Similarly, the identity can appear as any $i \times i$ minor constructed from the leftmost i rows.

The exact dimension is easy to work out in examples, but a closed form expression for it seems unnecessarily difficult.

Since the sections are polynomials of bounded degree (namely, i), the code constructed here is, in some sense, a subcode of a Reed–Müller code. The Reed–Müller code $\mathcal{R}(m, r)$ is constructed from the polynomials of degree at most r in m variables, using an evaluation map similar to the one used here. The minimum distance of such a code is 2^{m-r}; see [MS, Chapter 13]. Since the sections $f_{r_1,\ldots,r_i}$ operate on the first i columns of the representative matrix, the number of variables available is ni; thus

THEOREM. The minimum distance of the code from the i^{th} fundamental representation of SL_n is at least 2^{ni-i}.

6. Example

Consider the case of the second fundamental representation of SL_6 and the case of a Schubert cell represented by an array whose second row is [56]. The number of non-vanishing sections f_{ij} on this cell is $\binom{6}{2} = 15$. The minimum distance of a code so constructed is $2^{6\cdot 2-2} = 2^{10}$.

The *largest* cell with this second row corresponds to the array

$$\begin{array}{cccccc} 6 & & & & & \\ 5 & 6 & & & & \\ 4 & 5 & 6 & & & \\ 3 & 4 & 5 & 6 & & \\ 2 & 3 & 4 & 5 & 6 & \\ 1 & 2 & 3 & 4 & 5 & 6 \end{array}.$$

This cell thus corresponds to the integer sequence 6,5,4,3,2,1; the dimension is 15. This is quite large compared to the minimum distance, but the corresponding Schubert *variety* has subcells in each dimension. For example, the subcell from the sequence 5,6,1,2,4,3 has dimension 11, so the code has

length	2^{11}
dimension	15
minimum distance	2^{10}.

7. Comparison with Reed–Müller Codes

It is important to compare these codes with Reed-Müller codes, since the minimum distance estimate was based on the minimum distance for such a code. In fact, the codes constructed here can be better in the sense of having the same error correcting capability of a Reed-Müller code but a higher rate. For example, the $[2^{11}, 15, 2^{10}]$ code from the previous section has a higher rate than the Reed-Müller code $\mathcal{R}(11, 1)$ which has parameters $[2^{11}, 12, 2^{10}]$.

7. Further questions

The example in the previous section and its generalizations are amenable to computer analysis.

Since the codes constructed here are linear codes, standard decoding algorithms apply. In particular, Reed–Müller codes are easy to decode in hardware; see [MS]. It would be nice to have a decoding algorithm specifically suited to these codes.

The construction here can be generalized as follows: let P be a parabolic subgroup, that is, a subgroup containing a Borel subgroup. Then a representation of G with character χ determines a vector bundle over G/P. By the Bott–Borel–Weil theorem, exactly one of the cohomology groups $H^i(G/P, \chi)$ is non-zero, and G acts on this group. It is not clear whether better codes can be constructed in this manner. Nor is it clear whether better codes can be constructed from other groups G.

Finally, we have not attempted the more general case of codes over a field with $q \neq 2$ elements. One of our steps here requires the $p = 2$ hypothesis, namely, the claim that the sections $f_{r_1,\ldots,r_i}$ are well-defined, although it is possible to work around this restriction.

7. References

[Bo] Borel, Armand, *Linear Algebraic Groups*. Graduate Texts in Mathematics, Number 126. NY: Springer–Verlag.

[C] Chakravarti, M., "The generalized Goppa codes and related designs from Hermitian surfaces ...", in *Coding Theory and Applications*, Lecture Notes in Computer Science, volume 311. NY: Springer–Verlag.

[Ch] Chevalley, Claude, *Seminaire sur la classification des groupes de Lie algébriques.* (Mimeographed notes) Paris (1956 – 8).

[D] Demazure, Michel, "Une démonstration algébrique d'un théorème de Bott", *inventiones mathematicae* **5**(1968), 349 – 356.

[H] Humphreys, James E., *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics, Number 9. NY: Springer–Verlag.

[HLP] Høholdt, Tom, Jacobus van Lint, and Ruud Pellikaan, *Algebraic Geomtric Codes*, preprint (1996).

[MS] MacWilliams, F. J., and N. J. A. Sloane, *The theory of error–correcting codes*, North–Holland Mathematical Library, v. 16 (1978).

[P] Proctor, R., "Classical Bruhat orders and lexicographic shellability", *Journal of Algebra* **77** (1982), 104 – 126.

[RR] Ryan, Charles T. and Kevin M. Ryan, "An application of geometry to the calculation of weight enumerators", *Congr. Numer.* **67**(1988), 77–89.

[TV] Tsfasman, M. A., and S. G. Vlăduţ. *Algebraic–Geometric codes.* Mathematics and its Applications, Soviet Series, 58. Kluwer Academic Publishers, Dordrecht, 1991.

[W] Wolper, J., "A combinatorial approach to the singularities of Schubert varieties", *Advances in Math.* **76** (1989) 184 – 193.