The Topology of Hamada's Formula ($p = 2$)

James S. Wolper
Department of Mathematics
Idaho State University
Pocatello, ID 83209–8085 USA
wolperj@pequod.isu.edu

12 January 1996

Research partially supported by the Isaac Newton Institute for Mathematical Science, University of Cambridge, UK. Preprint number NI96002.

1. Introduction

Hamada's Formula gives the dimension of the projective geometry code $C^*(r, n, q)$, and several combinatoric and geometric proofs have been given; see *eg* [CH] and the references there. In this note, a new proof is given for the case $C^*(r, n, 2)$, using a method with a topological inspiration. While this case is perhaps not of the greatest interest (since such codes are Reed–Müller codes), the method appears to be very amenable to generalization. And while these dimensions were known, the recurrence relation presented here appears to be completely unrelated to other expressions for these quantities. It is not clear at this writing whether the technique used here works over other fields.

It is also hoped that other "continuous" methods could be similarly adapted to "discrete" problems.

I would like to thank J.W.P. Hirschfeld for his help with this project. Some of this work was performed while the author was a participant in the Isaac Newton Institute for Mathematical Sciences programme on *Computer Security, Cryptology, and Coding Theory*; their support is gratefully acknowledged.

2. Preliminaries

We find the dimension of $C(r, n, 2)$, the dual to the projective geometry code. It is easy to see that that

$$\dim C(r, n, q) + \dim C^*(r, n, q) = \frac{q^{n+1} - 1}{q - 1},$$

where the right hand side is the number of points in $\mathbf{P}^n(\mathbf{F}_q)$, the projective space over the field with q elements. The result is:

THEOREM.

$$\dim C(r, n, 2) = \sum_{i=0}^{r-1} \dim C(r - i, n - i - 1, 2) + \sum_{i=0}^{n-r} 2^i. \tag{1}$$

The code is constructed by considering all r-dimensional projective subspaces of $\mathbf{P}^n$, and taking rows of the incidence matrix whose rows represent the subspaces and whose columns represent the points. Over any groundfield k, this is equivalent to looking in the *grassmannian* of $r + 1$-dimensional vector subspaces of

k^{n+1}, or $Gr(r+1, n+1)$. This latter space is an algebraic variety over k, wtih dimension $(r+1)(n-r)$; in fact, when $k = \mathbf{C}$ or $k = \mathbf{R}$, the grassmannian is a smooth manifold of the indicated dimension.

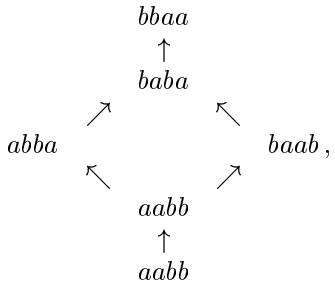
In the continuous cases, the grassmannian $Gr(r, n)$ (note change of parameters) has a cell decomposition, that is, it can be written as a disjoint union of proper submanifolds, each of which is isomorphic to the "cell" k^r. This is defined as follows: fix a *flag* of subspaces of k^n, that is, a sequence of subspaces $F_1 \subset F_2 \subset \cdots \subset F_n$, where $\dim F_i = i$. Then for each sequence $i_1, \ldots, i_r$ of distinct integers with $i_j \leq n$, there is a cell given by

$$\{V : \dim (V \cap F_{i_j}) = j, \dim (V \cap F_{i_j}) \neq j+1\}.$$

Such a cell is conveniently represented by a word of length n consisting of r "a"s and $n-r$ "b"s, where the "a"s go in positions i_j. Thus, in $Gr(2,4)$, the six cells are *bbaa*, *baba*, *baab*, *abba*, *abab*, and *aabb*.

These cells form a poset in two (or more) distinct ways. Topologically, $w < v$ if the cell represented by w is in the closure of the cell represented by v. This is expressed combinatorially by noting that the poset is generated by the "moves": $wabv$ is in the closure of $wbav$, where w and v are arbitrary subwords. This is variously called the *Bruhat order* or the *adherence order*. Both of these techniques can be generalized to more general *flag varieties*. A good reference for grassmannians and the cell decomposition is [GH].

The cells with this partial order define a lattice, called the *Hasse diagram*. For $Gr(2,4)$ it is



where $w \to w'$ indicates that $w \subset \overline{w'}$. A good reference for grassmannians and the cell decomposition is [GH].

For our case, the cell decomposition sets up an inductive method for counting the number of independent rows in the incidence matrix describing the code. The calculation is arranged as follows: The first part is induction from $C(1, n, q)$ to $C(1, n+1, q)$. The second is induction from each cell in the difference $C(1, n+1, q) - C(1, n, q)$ to the one above; that is, the difference $C(r, n+1, q) - C(r, n, q)$ is a union of cells, and we must count the contribution of each cell. The final part tackles the general case of $C(r, n, 2)$.

3. $\dim (C(1, n, 2))$, Part One

For the rest of this note we will identify a cell with the word describing it.

It is trivial to check that the dimension of $C(1, 2, 2)$ is 4. Now, note that all cells in $Gr(r+1, n+1)$ whose word ends in b are really cells in $Gr(r+1, n)$. Further, every point in every subspace in this portion is *not* in F_{n+1}. Thus, the incidence matrix splits into two rows of blocks, where the first row consists of subspaces in cells whose word ends in a, and the second row consists of subspaces in cells whose word ends in b. The first column block consists of points not in F_{n+1} (*ie*, last coordinate is zero), and the second column block consists of points whose last coordinate is one:

$$\begin{pmatrix} * & 0 \\ * & * \end{pmatrix}$$

Clearly, all of the vectors in the lower block row are independent of those in the upper, and the number of independent vectors in the upper row is $\dim (C(1, n-1, 2)$.

Notice that this step generalizes immediately to any $C(r, n, 2)$.

To continue the proof we need to show that dim $(C(1, n, 2))$ − dim $(C(1, n-1, 2)) = 2^n - 1$. This is part two.

4. dim $(C(1, n, 2))$, Part Two

Now we count the "new" independent vectors added by each cell in the difference $C(r, n+1, q) - C(r, n, q)$. To do this, we need a list of the cells whose last letter is a. This is $ab^{n-1}a$, $bab^{n-2}a$, ..., $b^{n-1}aa$. Note that the dimension of $b^r ab^s a$ is $2(n-1) - s$.

The bottom cell is $ab^{n-1}a$, which consists of the subspaces V which contain F_1 and are not contained in F_n. A basis for such a subspace is $\{(1, 0, \ldots, 0), (0, *, \ldots, *, 1)\}$, where $*$ is any element of the groundfield. Different choices for $*$ lead to disjoint subspaces, so each of the 2^{n-2} choices leads to an incidence vector which has a 1 in a position where no other incidence vector has a 1. Thus, all 2^{n-2} vectors are independent.

Now consider the transition to $b^r ab^s a$ from $b^{r-1}ab^{s+1}a$. The subspaces in the bigger cell satisfy the condition that $V \cap F_{r+1}$ has dimension 1, while $V \cap F_r$ has dimension zero. Thus, each such subspace has a basis of the form $\{(*, \ldots, *, 1, 0, \ldots, 0), (*, \ldots, *, 0, *, \ldots, *, 1)\}$, where the number of leading $*$s in the first and second vectors is r. Notice that the sum of any two incidence vectors constructed from any two such subspaces has the form $(*, \ldots, *, 0, *, \ldots, *, 1)$, that is, is in a subspace in the smaller cell. Thus, each choice of the r $*$s leads to only one independent vector (although each choice does yield one new vector).

Now add up the numbers of independent vectors in each cell. Starting at the top, the sum is $1 + 2 + 4 + \cdots + 2^{n-1} = 2^n - 1$.

5. General r

The calculation of dim $C(r, n, 2)$ for general values of r is organized as follows: the Hasse diagram for $Gr(r+1, n+1)$ is divided into regions according to the word's ending. The lowest region consists of those words ending in b; thus, it is isomorphic to the Hasse diagram for $C(r-1, n, 2)$. As above, the points in subspaces in cells in this region all have $x_{n+1} = 0$, and so this creates a block of zeroes in the incidence matrix defining the code.

The next region consists of cells whose words end in ba. Since the final two letters are already determined, the cells in this region have the form wba, where w is a word representing a cell in $Gr(r, n)$. The next region consists of words ending in aa, which puts its contents poset isomorphic with the Hasse diagram for $C(r-2, n-1)$. The next region's words end in baa, etc. The final region consists of the words ending with r consecutive as.

So the calculation is clear: the incidence matrix is partitioned into blocks, one block for each region. All but the final region contribute dim $C(r-i, n-i-1, 2)$ new independent rows, while the final region contributes $1 + 2 + \cdots + 2^{n-r}$ new independent rows, using the same argument as above. The exclusion of points with $x_{n+1-i} = 1$ in the i^{th} region means that no new dependency relations occur.

It should be noted that Hirschfeld has verified formula (1) using binomial coefficient identities.

6. Conclusions

The strategy of this note has shown that a topological perspective "reorganizes" the calculation of dim $(C(r, n, 2))$; this part of the argument is independent of the groundfield. The tactics, however, take advantage of some of the special relationships that occur when the ground field has 2 elements. Future work will address these deficiencies.

REFERENCES

[CH] Ceccherini, P. V. and J. W. P. Hirschfeld, "The dimension of projective geometry codes", *Discrete Mathematics* **106/107** (1992), 117 – 126.

[GH] Griffiths, P., and J. Harris, *Prinicples of Algebraic Geometry*, Wiley–Interscience (1978).