POLYNOMIAL CODES AND FINITE GEOMETRIES*

E. F. Assmus, Jr and J. D. Key

October 16, 1996

Contents

*The authors wish to thank Paul Camion, Pascale Charpin and Projet Codes at INRIA for the hospitality and support shown during the preparation of this manuscript. In particular, the first author spent much of 1992-1993 at Projet Codes where the bulk of his work on the chapter was completed.

1 Introduction

The reader familiar with "Designs and their Codes" will soon understand the debt this chapter owes to that book — especially its Chapter 5. We have, however, entirely reworked that material and, more importantly, added a discussion of the group-algebra approach to the Reed-Muller and generalized Reed-Muller codes. This enables us to include a straightforward new proof of Berman's theorem identifying the Reed-Muller codes with the radical powers in the appropriate modular group algebra and to use our treatment of the Mattson-Solomon polynomial to give a proof of the generalization of Berman's theorem to the p-ary case. We have also included Charpin's treatment [16] of the characterization of "affine-invariant" extended cyclic codes due to Kasami, Lin and Peterson.

We have relied heavily on Charpin's doctoral thesis [14, 16] for the new material. The older material relies (as did Chapter 5 of our book) on the treatment of the polynomial codes introduced by Kasami, Lin and Peterson [29] given by Delsarte, Goethals and MacWilliams [18].

Our definition of the generalized Reed-Muller codes is the straightforward generalization of the boolean-function definition of the Reed-Muller codes and, for us, the cyclicity of the punctured variants is simply a consequence of the easily seen fact that their automorphism groups contain the general linear groups.

We are, of course, principally interested in the geometric nature of certain of these codes. Were one interested only in the binary case the development would be very short and our treatment reflects that fact in that we first discuss the Reed-Muller codes giving complete proofs that differ substantially from those given for the general case. In fact, we have here an instance in which the generalization to an arbitrary finite field seems far from trivial, the biggest hurdle being the passage to fields that are not of prime order.

The peculiar nature of the definitions of the geometric codes in the coding-theory literature was due to the interest — at the time of their introduction — in majority-logic decoding of these codes; we therefore also give a short discussion of decoding. On the other hand, we give the natural definitions of the geometric codes (as codes generated by the incidence vectors of the geometric objects at hand) and, hence, our definitions are not the ones found in many engineering texts.

We review the necessary geometry briefly before beginning our discussion of the codes; our treatment is undoubtedly too brief to be useful to a reader with no background whatsoever in finite geometry and such a reader may wish to jump directly to Section 3 — which may even motivate a study of the geometry involved. Much of the material will be understandable even without a firm grip on the geometry and subsequent sections should be of interest to professional coding theorists. We have, at least, tried to make them so.

We assume a knowledge of coding theory and we believe the reader will find in Chapter 1 the coding theory necessary for a study of this chapter.

We have not attempted to discuss open problems or to explore new avenues of research. The reader interested in such matters may wish to consult our book [2] or the articles cited in the bibliography.

2 Projective and affine geometries

Let F be a field and V a vector space over F. We denote by $PG(V)$ the **projective geometry** of V. Its elements are the subspaces of V and its structure is given by set-theoretic inclusion. Similarly, $AG(V)$ denotes the **affine geometry** of V. Its elements are the cosets, $\mathbf{x} + U$, of subspaces U of V, where $\mathbf{x}$ is any vector in V, and again the structure is given by set-theoretic inclusion. The "geometry" of these structures arises by viewing inclusion as an incidence relation.

2.1 Projective geometry

If the vector space V has dimension $n+1$ over F, then $PG(V)$ has **projective dimension** n. We record this with the notation $PG_n(F)$, realizing V

as F^{n+1}. In this case a "point" of the geometry is given in homogeneous coordinates by $(x_0, x_1, \dots, x_n)$ where all x_i are in F and are not all zero; each point has many such coordinate representations[1], in fact $q-1$ when F is F_q, since $(x_0, x_1, \dots, x_n)$ and $(\lambda x_0, \lambda x_1, \dots, \lambda x_n)$ yield the same 1-dimensional subspace of F^{n+1} for any non-zero λ, the 1-dimensional subspaces being the points — or objects of projective dimension 0. Similarly, the **projective dimension** of any subspace is defined to be 1 less than the dimension of the subspace (as a vector space over F).

Thus the **points** of $PG(V)$ are the 1-dimensional subspaces of V, the **lines** are the 2-dimensional subspaces of V, the **planes** the 3-dimensional subspaces of V, and the **hyperplanes** the n-dimensional subspaces of V. Neither $\{0\}$ nor V play a significant role in projective geometry and they are usually ignored. Frequently when working with projective geometry the projective dimension is referred to simply as the *dimension*. The dimension formula for subspaces of V holds for projective dimension as well, provided it is written as follows:

$$\dim(U) + \dim(W) = \dim(U + W) + \dim(U \cap W),$$

where U and W are arbitrary non-zero subspaces of V and $U + W = \langle U \cup W \rangle = \{u + w \mid u \in U, w \in W\}$. Note that we use $\langle S \rangle$ to denote the subspace generated by the set S. The formula has the following important consequence:

Suppose H is a hyperplane of $PG_n(F)$. If U is a subspace of dimension $t > 0$, then $U \cap H$ has dimension t or $t-1$, the former if and only if U is contained in H.

If P and Q are distinct points of $PG(V)$, then $P + Q$ is necessarily a line of $PG(V)$, again by the above formula, and, in fact, it is the unique line through P and Q. Thus every two distinct points lie on a unique line. In projective dimension 2, i.e. in a projective plane, every two distinct lines intersect in a unique point. We will, as here, use geometric terminology whenever convenient.

If $F = F_q$ and V is m-dimensional, one can see by counting bases that the number of subspaces of V of dimension k, where $0 < k \le m$, is

$$\frac{(q^m - 1)(q^m - q)\dots(q^m - q^{k-1})}{(q^k - 1)(q^k - q)\dots(q^k - q^{k-1})}.$$

[1]Except in the binary case; it is this uniqueness that makes the Reed-Muller codes so much easier to analyze than the generalized Reed-Muller codes.

Similarly — or by using the above formula on a quotient space — if V is of dimension m, U a subspace of dimension r, and k an integer with $0 \leq r < k \leq m$, then the number of subspaces of V of dimension k that contain U is

$$\frac{(q^m - q^r)(q^m - q^{r+1}) \dots (q^m - q^{k-1})}{(q^k - q^r)(q^k - q^{r+1}) \dots (q^k - q^{k-1})}.$$

In particular, the number of points of $PG_n(F_q)$ is $\frac{q^{n+1}-1}{q-1} = q^n + q^{n-1} + \cdots + 1$ and the number of lines in the pencil of lines containing a point is $\frac{q^{n+1}-q}{q^2-q} = \frac{q^n-1}{q-1} = q^{n-1} + \cdots + 1$.

Definition 2.1 *If V and W are finite-dimensional vector spaces, then $PG(V)$ and $PG(W)$ are* **isomorphic** *if there is a bijection*

$$\varphi : PG(V) \to PG(W)$$

such that, for $U, U' \in PG(V)$, $U \subseteq U'$ if and only if $U\varphi \subseteq U'\varphi$. If $W = V$, then such a map φ is called an **automorphism** *or* **collineation** *of $PG(V)$.*

Since the projective dimension of $PG(V)$ is equal to the length of the longest chain, $U_1, U_2, \dots, U_k$, of elements of $PG(V)$ satisfying $U_1 \subset U_2 \subset \dots \subset U_k$, it follows that isomorphic geometries have the same projective dimension. That is, V and W must be of the same dimension and, provided they are vector spaces over the same field, they must be isomorphic as vector spaces. Any invertible linear transformation from V to W will induce an isomorphism of the geometries, but something slightly more general will also, a so-called semilinear transformation:

Definition 2.2 *Let F be a field and let V and W be vector spaces over F. A* **semilinear transformation** *of V into W is given by a map*

$$T : V \to W$$

together with an associated automorphism, $\alpha(T)$, of the field F. The map T is additive, i.e. $(\mathbf{v}+\mathbf{u})T = \mathbf{v}T + \mathbf{u}T$ for all $\mathbf{v}, \mathbf{u} \in V$, and $(a\mathbf{v})T = a^{\alpha(T)}(\mathbf{v}T)$ for all $a \in F$ and $\mathbf{v} \in V$.

A semilinear transformation carries subspaces into subspaces, preserving inclusion, and thus induces an incidence-preserving map on the projective geometries. It is an isomorphism of the projective spaces whenever

T is an isomorphism of the additive structures, the inverse being given by T^{-1}, with the associated automorphism of F being $\alpha(T)^{-1}$. Notice that the composition of semilinear transformations is again semilinear and, in fact, $\alpha(ST) = \alpha(S)\alpha(T)$. It follows that when $V = W$ the semilinear isomorphisms form a group and that the map sending T to $\alpha(T)$ defines a homomorphism into the Galois group of F (here the automorphism group of F). The kernel is the group of invertible *linear* transformations of V.

In terms of bases, given ordered bases $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ and $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ of V and W, respectively, then if $(\mathbf{v}_i)T = \sum_{j=1}^{n} a_{ij}\mathbf{w}_j$, $A = (a_{ij})$ and $\alpha = \alpha(T)$, then

$$T : (x_1, x_2, \ldots, x_m) \mapsto (x_1^\alpha, x_2^\alpha, \ldots, x_m^\alpha)A,$$

where, as usual, we have used the bases to identify V with F^m and W with F^n. In matrix form, the composition of two semilinear transformations, (α, A) and (β, B), is $(\alpha\beta, A^\beta B)$, where A^β denotes the matrix (a_{ij}^β). Since a matrix A together with an automorphism α clearly yield, by the above formula, a semilinear transformation, the map sending T to $\alpha(T)$, in the case where $V = W$, is a homomorphism *onto* the Galois group of F.

Thus, for a given vector space V, the group of semilinear isomorphisms of V contains $GL(V)$, the group of invertible linear transformations of V, as a normal subgroup, the quotient being the Galois group of F. The group of semilinear isomorphisms is denoted by $\mathit{\Gamma L}(V)$. Clearly every semilinear isomorphism of V induces an isomorphism of $PG(V)$. The scalar transformations (i.e. those that send $\mathbf{v}$ to $a\mathbf{v}$ for some fixed $a \in F$) induce the identity isomorphism and they are the only semilinear isomorphisms that do. The subgroup of scalar transformations is the centre of $GL(V)$ and a normal subgroup of $\mathit{\Gamma L}(V)$; the quotient groups are denoted, respectively, by $PGL(V)$ — the **projective general linear group** — and $P\mathit{\Gamma L}(V)$ — the **projective semilinear group**. If V is n-dimensional and a basis has been chosen, $PGL(V)$ becomes a matrix group modulo scalar matrices and is denoted by $PGL_n(F)$; similarly in this case we write $P\mathit{\Gamma L}_n(F)$ for $P\mathit{\Gamma L}(V)$. Each of these groups acts as a permutation group on the elements of $PG(V)$, the action on the points of $PG(V)$ being doubly-transitive, which means that given any two pairs of distinct points, (P, Q) and (P', Q'), there is an automorphism in $PGL(V)$ which simultaneously carries P to P' and Q to Q'. In the standard notation, $PGL_n(F)$ acts on $PG_{n-1}(F)$; similarly for the semilinear group.

All the collineations of $PG(V)$ are induced by semilinear transformations; this is the content of the following classical **fundamental theorem**

of projective geometry:

Theorem 2.3 *Let V be a vector space of dimension at least* 3. *Then $P\Gamma L(V)$ is the full automorphism group of $PG(V)$.*

There are well-established proofs of this theorem readily available: see Artin [1, Chapter II], for example, or, for a slightly more modern account, Hahn and O'Meara [24, Chapter 3]. Also note that the theorem starts with planes; the projective line consists merely of points and the lack of any incidences allows an arbitrary permutation to be admitted as an automorphism.

Amongst the automorphisms of $PG_n(F_q)$ there is always one of order $v = (q^{n+1}-1)/(q-1)$ that permutes the points of the geometry in a single cycle of this length, called a **Singer cycle** (after Singer [49]). This automorphism is constructed as follows: view the finite field $K = F_{q^{n+1}}$ as a vector space of dimension $n+1$ over the field $F = F_q$ and let ω be a primitive element of K, that is, a generator of the cyclic group $K^\times = K - \{0\}$. Using the given field structure, it is clear that multiplication by ω induces a linear transformation on the vector space $V = K$. Since the field F has ω^v as primitive element, it is easy to see that this linear transformation induces an automorphism of $PG(V)$ that acts as a cycle of length v on the v points of the geometry. In fact, the 1-dimensional subspaces of $V = K$ given by the non-zero vectors $1, \omega, \omega^2, \ldots, \omega^{v-1}$ represent all the points of $PG_n(F)$, where, of course, ω^v represents the same point as 1, etc.

2.2 Affine geometry

The affine geometry, $AG(V)$, where V is a vector space over a field F, consists of all cosets, $\mathbf{x} + U$, of all subspaces U of V with incidence defined through the natural inclusion relation. Here the dimension is the same as that of the vector space—for obvious geometric reasons. The dimension of a coset is that of the defining subspace U, and if the latter has dimension r, we will also refer to a coset of U as an ***r*-flat**. Thus the **points** are all the vectors, including $\mathbf{0}$, the **lines** are 1-dimensional cosets, or 1-flats, the **planes** are the 2-dimensional cosets, or 2-flats, and so on, with the **hyperplanes** the cosets of dimension $n-1$ — where V is of dimension n over F. We also write $AG_n(F)$ for $AG(V)$, in analogy with the projective case. The affine geometry of these cosets is defined by the inclusion relation which specifies that, if $M = \mathbf{x} + U$ and $N = \mathbf{y} + W$ are cosets in $AG(V)$, then M *contains* N if $M \supseteq N$, from which it follows that W is a subspace of

U. The affine geometry $AG(M)$ is, by definition, the set of cosets of $AG(V)$ that are contained in M together with the induced incidence relation. This is quite clear when M is a subspace but if $M = \mathbf{x} + U$ with $\mathbf{x} \notin U$ it follows also that $AG(M)$ is isomorphic to $AG(U)$ since every element of $AG(M)$ can be written in the form $\mathbf{x} + U'$ for some subspace U' of U. As in the projective case we will use standard geometric terminology — in particular the notion of **parallelism**:

Definition 2.4 *The cosets* $\mathbf{x}+U$ *and* $\mathbf{y}+W$ *in* $AG(V)$ *are* parallel *if* $U \subseteq W$ *or* $W \subseteq U$.

Cosets of the same subspace are thus parallel and cosets of the same dimension are parallel if and only if they are cosets of the same subspace. For a given subspace U of dimension r, its distinct cosets partition V into parallel r-flats and parallelism is an equivalence relation on the set of r-flats of V, the equivalence classes being called parallel classes. Hyperplanes, i.e. $(n-1)$-flats, in $AG_n(F)$ are parallel if and only if they are equal or intersect in the empty set and in $AG_n(F_2)$ a hyperplane and its complement make up a parallel class. In $AG_n(F_q)$ there are q hyperplanes in a parallel class. Here is one more important fact about flats that we will need to properly explain Reed's decoding algorithm for Reed-Muller codes:

If M is an r-flat and N an $(n-r)$-flat in $AG_n(F)$, then either $M \cap N$ is a single point, in which case N meets all the r-flats parallel to M in a single point, or else the intersection of N with an r-flat parallel to M is either a flat of positive dimension or the empty set.

As in the projective case, both $GL(V)$ and $\Gamma L(V)$ act on the geometry, but now we also have V itself acting via translation. The underlying action of the **affine general linear group**, $AGL(V)$, and the **affine semilinear group**, $A\Gamma L(V)$, is given as follow: for $T \in \Gamma L(V)$ and $\mathbf{v} \in V$, the map $(T, \mathbf{v})$ is defined by

$$\mathbf{x}(T, \mathbf{v}) = \mathbf{x}T + \mathbf{v}$$

for each $\mathbf{x} \in V$. Such maps preserve cosets and thus act on $AG(V)$. Composition is given by $(S, \mathbf{v})(T, \mathbf{w}) = (ST, \mathbf{v}T + \mathbf{w})$ and it follows that these affine groups are semi-direct products of the linear and semilinear groups (respectively) with the additive group of V, the action of the linear and semilinear groups on V being the natural one. The permutation action on the points of $AG(V)$, i.e. on the vectors in V, is doubly-transitive and, if $F = F_2$, it is triply-transitive.

Given a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ for V, if $(T, \mathbf{v})$ is an element of $A\Gamma L(V)$, and $\mathbf{v} = \sum_i b_i \mathbf{v}_i$, define the matrix A via $\mathbf{v}_i T = \sum_j a_{ij} \mathbf{v}_j$, and let α be the field automorphism associated with T. Then

$$(T, \mathbf{v}) : (x_1, \ldots, x_n) \mapsto (x_1^\alpha, \ldots, x_n^\alpha)A + (b_1, \ldots, b_n).$$

Moreover, given any triple $(\alpha, A, (b_1, \ldots, b_n))$ where α is an automorphism of the field F, A is an $n \times n$ matrix with entries from F and $(b_1, \ldots, b_n) \in F^n$, the formula above defines an element of $A\Gamma L(V)$ and, in fact, with the obvious multiplication of the triples,

$$(\alpha, A, (b_1, \ldots, b_n))(\beta, B, (c_1, \ldots, c_n)) = (\alpha\beta, A^\beta B, (b_1^\beta + c_1, \ldots, b_n^\beta + c_n)),$$

we have an isomorphism of $A\Gamma L(V)$ with this group, denoted by $A\Gamma L_n(F)$. Similarly we write $AGL_n(F)$ for the affine linear group — when it is given explicitly.

In analogy with the projective case, there is a **fundamental theorem of affine geometry** which states that for $n \geq 2$, $\mathrm{Aut}(AG_n(F)) = A\Gamma L_n(F)$. This is the same theorem, in effect, as the fundamental theorem for projective geometry, if we consider the way in which affine geometries are embedded in projective geometries:

Theorem 2.5 *Let V be a vector space over F, H a hyperplane, and $\mathbf{x}$ a vector in V that is not in H. Set $PG(V)^H = \{U \,|\, U \in PG(V), U \not\subseteq H\}$. Define a map*

$$\varphi : AG(\mathbf{x} + H) \to PG(V)$$

by $M \mapsto \langle M \rangle$ for any coset $M \in AG(\mathbf{x}+H)$. Then φ is an incidence preserving injection with image $PG(V)^H$. Further, the inverse map φ^{-1} satisfies

$$U\varphi^{-1} = U \cap (\mathbf{x} + H),$$

for any $U \in PG(V)^H$.

This is the **fundamental embedding theorem** and the proof is quite direct from the definitions; it can be found in Gruenberg and Weir [23]. Note that the choice of the hyperplane H and vector $\mathbf{x}$ that produce the embedding is not crucial since for another choice, K and $\mathbf{y}$, it is clear that $AG(\mathbf{x} + H)$ is isomorphic to $AG(\mathbf{y} + K)$ and, moreover, H and K are equivalent under the projective group. One generally thinks of the 1-dimensional subspaces of H as the "points at infinity" of the projective space $PG(V)$ and discarding these points leaves the affine geometry of the same dimension. In coordinate terms one can view H as the hyperplane in $F^{n+1} = \{(x_0, x_1, \ldots, x_n) \mid x_i \in F\}$ given by the equation $X_0 = 0$, taking, for convenience, $\mathbf{x} = (1, 0, \ldots, 0)$. Then the embedded affine space is F^n viewed as the last n coordinates, where every projective point *not at infinity* has homogeneous coordinates that can be taken to be $(1, x_1, \ldots, x_n)$. More precisely, the embedded affine geometry of dimension n is obtained from a projective geometry of dimension n by removing a hyperplane and all the subspaces contained in it. The points and subspaces remaining form the affine geometry.

When doing computations one works, normally, in the affine space. In an affine geometry of dimension n, once a basis is chosen for the vector space, any r-flat can be given by a set of $(n - r)$ independent linear equations and solutions are points of the geometry. In the projective case one uses homogeneous equations, of course, and only looks for non-zero solutions — which are not precisely the points but only representatives. So, for example, in $AG_4(F)$ the equations $X_1 + X_2 - X_3 = 0$ and $X_1 + X_4 = 1$ define a 2-flat; it is given by $(0, 0, 0, 1) + U$ where U is the 2-dimensional subspace $\{(x, y, x + y, -x) \mid x, y \in F\}$. In other words the 2-flat consists of all vectors in F^4 of the form $\{(x, y, x + y, 1 - x\}$.

2.3 Designs from geometries

To define incidence structures from $PG(V)$ and $AG(V)$ we need to choose point sets and block sets; the incidence relation will be that of the geometry, namely containment. In every case the point set of our design will be the set of points of the geometry: for projective spaces the 1-dimensional subspaces of V and for affine spaces the vectors of V. For blocks we will take all the subspaces (or cosets in the affine case) of a fixed dimension. In every case the double-transitivity of the group involved will assure us that we are dealing with a 2-design.

Thus, for example, we can consider the design of points and lines, the design of points and planes, or the design of points and hyperplanes of a geometry and be assured of a 2-design. The parameters will depend on both the dimension of the geometry and the cardinality of the finite field. By fixing one of these and letting the other vary we obtain numerous infinite families of designs. Each of these designs will have an automorphism group containing $P\mathit{\Gamma}L(V)$ or $A\mathit{\Gamma}L(V)$ in the projective or affine case, respectively. Except for isolated cases the parameters will admit many designs other than these classical designs, and a large amount of effort has gone into classifying the classical designs amongst those with the same parameters.

Perhaps the most interesting case is that of dimension 2. In the projective case, $PG_2(F_q)$ produces the design of points and lines of a 3-dimensional vector space over a finite field, a classical *projective plane*. It is a design with parameters $2\text{-}(q^2+q+1, q+1, 1)$. For q a proper power of a prime there are many such designs that do not arise from $PG_2(F_q)$, but for q a prime only the classical plane has appeared — and it is possible that there may not be any others. For q *not* a power of a prime *no* designs with these parameters have been discovered. Observe that to recover q from the parameters one must take $\lambda_1 - \lambda_2$ in the notation of Chapter 1; this integer is an important parameter of a design, and is called the **order**.

In the affine case, $AG_2(F_q)$ produces the design of points and lines of a 2-dimensional vector space, i.e. a classical *affine plane*. It is a $2\text{-}(q^2, q, 1)$ design. It also has order q.

Projective planes are **symmetric** designs, i.e. have the same number of points as blocks. For a symmetric $2\text{-}(v, k, \lambda)$ design $\lambda_1 = k$ and the order is given as $k - \lambda$, as it was for projective planes. More generally, the design of points and hyperplanes of a projective geometry produces a symmetric design. If the finite field has q elements and the geometry has projective dimension n, then this design of points and hyperplanes is a symmetric design with parameters

$$2 - \left(\frac{q^{n+1}-1}{q-1}, \frac{q^n-1}{q-1}, \frac{q^{n-1}-1}{q-1} \right)$$

and order q^{n-1}.

2.4 Codes from designs

For any finite incidence structure $\mathcal{D}$ with point set $\mathcal{P}$ and block set $\mathcal{B}$, the **code** $C_p(\mathcal{D})$ of $\mathcal{D}$ over a prime field F_p is the subspace of the space $F_p^{\mathcal{P}}$ of all functions from $\mathcal{P}$ to F_p that is spanned by the incidence vectors of the blocks of $\mathcal{D}$. This code is equivalent to the code given by the row space of any incidence matrix of the incidence stucture — where we use the blocks to index the rows (and the points the columns) of the incidence matrix. Although this is the appropriate way to view the incidence matrix in the context of this chapter, it does sometimes prove useful to examine the code given by the row space of the "point by block" incidence matrix; see, for example, [52].

For any subset $X \subseteq \mathcal{P}$, we denote the characteristic function of X by v^X and refer to v^X as the **incidence vector** of X. Thus

$$v^X(x) = \begin{cases} 1 & \text{if } x \in X \\ 0 & \text{if } x \notin X \end{cases},$$

where $v^X(x)$ denotes the value that the function v^X takes at the point x. Then

$$C_p(\mathcal{D}) = \langle v^B \mid B \in \mathcal{B} \rangle.$$

The dimension of $C_p(\mathcal{D})$ is referred to as the p-rank of $\mathcal{D}$. The rank tends to vary with p in the general case; for so-called 2-designs it is easily determined except for those primes dividing the order of the design.

The minimum weight of the code arising from an incidence structure is clearly at most equal to the cardinality of the smallest block. In general the minimum weight is strictly less than this cardinality, but for the classical geometric designs studied in this chapter there is a distinguished prime one considers, and for these codes we will have equality.

As we will soon see, one of the most widely studied class of binary codes, the Reed-Muller codes, arises precisely as the class of codes given by geometric designs over the binary field — although the original presentation of these codes in 1954 was in the boolean-function context and was given by electrical engineers.

3 The Reed-Muller codes

The Reed-Muller codes have already been defined in Chapter 1 (Section 13). For completeness, and in order to establish our notation for this section and those to follow, we will repeat some of the definitions and results.

3.1 Definitions

Throughout this section F will denote the field F_2. Let V be a vector space of dimension m over F. We let F^V denote the vector space over F of all functions from V to F. As a vector space over F, F^V has dimension 2^m, the cardinality of the set V. Since F^V will be the ambient space for the Reed-Muller codes we must choose a basis for it and we choose the standard basis consisting of the characteristic functions of the elements of the set V. Denoting a typical element of V by $\mathbf{v}$ these basis elements are $\{v^{\mathbf{v}} \mid \mathbf{v} \in V\}$, where we write $v^{\mathbf{v}}$ instead of the more cumbersome $v^{\{\mathbf{v}\}}$. Viewing V as F^m, it too has a standard basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$, where

$$\mathbf{e}_i = (\underbrace{0, 0, \ldots, 1}_{i}, 0, \ldots, 0).$$

Moreover, any function $f \in F^V$ can be given as a function of m variables corresponding to the m coordinate positions: writing the vector $\mathbf{x} \in V$ as

$$\mathbf{x} = (x_1, x_2, \ldots, x_m) = \sum_{i=1}^{m} x_i \mathbf{e}_i,$$

then $f = f(x_1, x_2, \ldots, x_m)$. The "polynomial" x_i is, for example, the linear functional that projects a vector in V onto its i^{th} coordinate in the given basis, its value at $(\sum_{j=1}^{m} x_j \mathbf{e}_j)$ being x_i.

As a function on V, $x_i^k = x_i$ whenever $k > 0$, so we obtain all the monomial functions via the 2^m monomial functions:

$$\mathcal{M} = \{x_1^{i_1} x_2^{i_2} \ldots x_m^{i_m} \mid i_k = 0 \text{ or } 1; k = 1, 2, \ldots, m\},$$

where we write 1 for the constant function $x_1^0 x_2^0 \ldots x_m^0$ with value 1 at all points of V; as a code vector it is the all-one vector $\mathbf{1}$. The linear combinations over F of these 2^m monomials give all the polynomial functions, since, once again, we can reduce any polynomial in the x_i modulo $x_i^2 - x_i$, for $i = 1, 2, \ldots, m$. The set $\mathcal{M}$ of 2^m monomials is another basis for the vector space F^V; the following lemma indicates how each of our given basis elements of characteristic functions of the vectors in V is given as a polynomial, i.e. as a sum of elements of $\mathcal{M}$. This not only proves the assertion but also shows that the set $\mathcal{M}$ is a linearly independent set of vectors in F^V.

Lemma 3.1 *Set $K = \{1, 2, \ldots, m\}$ and, for $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in V$, let $I_{\mathbf{w}} = \{i \in K \mid w_i = 1\}$. Then*

$$v^{\mathbf{w}} = \prod_{k=1}^{m} (x_k + 1 + w_k) = \sum_{K \supseteq J \supseteq I_{\mathbf{w}}} \prod_{j \in J} x_j.$$

Proof: The proof is simple: the first polynomial is easily seen to define the characteristic function of the vector $\mathbf{w}$; and the expansion of this product is clearly the sum on the right. $\square$

We repeat the definition of the Reed-Muller codes:

Definition 3.2 *Let V denote the vector space of dimension m over $F = F_2$ and let r satisfy $0 \leq r \leq m$. The* **Reed-Muller code of order r**, *denoted by $\mathcal{R}(r,m)$, is the subspace of F^V (with basis the characteristic functions of the vectors of V) that consists of all polynomial functions in the x_i of degree at most r, i.e.*

$$\mathcal{R}(r,m) = \left\langle \prod_{i \in I} x_i | I \subseteq \{1,2,\ldots,m\}, 0 \leq |I| \leq r \right\rangle.$$

Example 3.3 The first-order Reed-Muller code $\mathcal{R}(1,m)$ consists of all linear combinations of the monomials x_i and 1 and hence each codeword, apart from 0 and $\mathbf{1}$, is given either by a non-zero linear functional on V or by 1 plus such a functional. Since any non-zero linear functional has 2^{m-1} zeros, every vector of $\mathcal{R}(1,m)$, apart from 0 and $\mathbf{1}$, has weight 2^{m-1}. A generator matrix for $\mathcal{R}(1,m)$ using the basis $x_1, x_2, \ldots, x_m, 1$ can be written so that the first $2^m - 1$ columns and m rows are the binary representations of the numbers between 1 and $2^m - 1$, whereas the last column is all 0, apart from a final row where all entries are equal to 1. This is clearly a generator matrix for the orthogonal of the extended Hamming code, i.e. $\mathcal{R}(1,m) = (\widehat{\mathcal{H}_m})^{\perp}$, where $\widehat{C}$ denotes the code obtained from C by adding an overall parity check.

As an immediate consequence of the definition and the linear independence of the functions in $\mathcal{M}$, we have that

$$\dim(\mathcal{R}(r,m)) = \binom{m}{0} + \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{r}.$$

In particular, $\dim(\mathcal{R}(1,m)) = 1 + m$.

The trivial cases include the repetition code, $\mathcal{R}(0,m) = F\mathbf{1}$, $\mathcal{R}(m,m) = F^V$ and the code $\mathcal{R}(m-1,m)$, which is of codimension 1 in F^V and equal to $(F\mathbf{1})^{\perp}$. The Reed-Muller codes are a nested sequence of codes. That is,

$$\mathcal{R}(r,m) \subseteq \mathcal{R}(s,m)$$

whenever $0 \leq r \leq s \leq m$.

We mentioned above that the orthogonal of $\mathcal{R}(0,m) = F\mathbf{1}$ is $\mathcal{R}(m-1,m)$. This is a special case of the following result, which was proved in Chapter 1:

Theorem 3.4 *For any $m \geq 1$ and any r such that $0 \leq r < m$,*

$$\mathcal{R}(r,m)^{\perp} = \mathcal{R}(m-r-1,m).$$

We will, in fact, reprove this result in Section 5 when we give its straightforward generalization to generalized Reed-Muller codes, Theorem 5.8.

Example 3.5 From Theorem 3.4 we get immediately that

$$\mathcal{R}(1,m)^{\perp} = \widehat{\mathcal{H}_m} = \mathcal{R}(m-2,m).$$

Thus, extended Hamming codes are Reed-Muller codes.

In the next subsection we will see the connection between the Reed-Muller codes and the codes of the designs of points and flats in affine space over F_2. The codes of the analogous designs from projective spaces over F_2 arise as *punctured* Reed-Muller codes:

Definition 3.6 *For* $0 \leq r < m$ *the* **punctured Reed-Muller code of order** $\boldsymbol{r}$, $\mathcal{R}(r,m)^*$, *is the code obtained from* $\mathcal{R}(r,m)$ *by puncturing at the vector* $\mathbf{0} \in V$.

One could puncture at *any* vector of V and get an isomorphic code since the set of polynomial functions is invariant under translation in V; i.e. if f is a polynomial in the x_i's of degree s then so is g where $g = f(x_1 + a_1, \ldots, x_m + a_m)$ for any vector $\mathbf{a} = (a_1, \ldots, a_m) \in V$, which means that the automorphism group of any Reed-Muller code acts transitively on the coordinates.

Example 3.7 If $m = 3$, $r = 1$, $\mathcal{R}(1,3)$ is a self-dual $[8,4,4]$ binary code, and $\mathcal{R}(1,3)^*$ is a $[7,4,3]$ code, *viz.* the Hamming code $\mathcal{H}_3$. Example 3.5 gives the reason for this and shows that Hamming codes are punctured Reed-Muller codes.

Proposition 3.8 *For* $r < m$ *the punctured Reed-Muller code* $\mathcal{R}(r,m)^*$ *is a*

$$[2^m - 1, \tbinom{m}{0} + \tbinom{m}{1} + \cdots + \tbinom{m}{r}]$$

binary code.

Proof: This follows easily: the dimension must be that of $\mathcal{R}(r,m)$ since all the vectors in this code are of even weight and the projection cannot, therefore, have a nontrivial kernel. $\square$

Finally, note that it follows from Theorem 3.4 that

$$(\mathcal{R}(r,m)^*)^{\perp} = \mathcal{R}(m-r-1,m)^* \cap (F_2\mathbf{1})^{\perp}$$

provided $r < m$. That is, $(\mathcal{R}(r,m)^*)^{\perp}$ consists of the vectors of $(\mathcal{R}(m-r-1,m))$ with a zero at $\mathbf{0}$, that coordinate being discarded.

3.2 Geometries and Reed-Muller codes

The set of vectors V is the point set for any design defined from an affine geometry $AG_m(F_2)$; moreover the binary codes of all the associated designs of points and flats are subspaces of F^V. Similarly, the designs from the projective geometry $PG_{m-1}(F_2)$ all have point set $V^* = V - \{\mathbf{0}\}$ and F^{V^*} is the ambient space of their binary codes. In this section we indicate how to associate these design codes with the Reed-Muller and punctured Reed-Muller codes of the last section.

Consider the generating elements of $\mathcal{R}(r, m)$: the polynomial x_i as a codeword has value 1 at a point $\mathbf{x}$ in V if the vector $\mathbf{x}$ has a 1 in the coordinate position i and value 0 otherwise. Thus $1 + x_i = v^H$, where H is the hyperplane with the equation $X_i = 0$. Also, x_i is the characteristic function of the complement of this hyperplane, i.e. the $(m-1)$-flat with equation $X_i = 1$. Similarly, $(1 + x_i)(1 + x_j)$, for $i \neq j$, is the characteristic function of the intersection of two hyperplanes, a subspace of dimension $m - 2$. In general, all the elements of $\mathcal{M}$ are the incidence vectors of flats in the affine geometry and $\mathcal{R}(r, m)$ is spanned by the incidence vectors of these $(m - s)$-flats, for $0 \leq s \leq r$. In order to show that $\mathcal{R}(r, m)$ is the binary code of the design of points and $(m - r)$-flats of $AG_m(F_2)$, which is our aim, we need to show that the vectors given by the $(m - r)$-flats span $\mathcal{R}(r, m)$. Notice that we already have this result for the first-order Reed-Muller codes, since the linear equations certainly define $(m - 1)$-flats and, furthermore, $\mathcal{R}(1, m)$ has precisely $2(2^m - 1)$ such vectors, the number of $(m - 1)$-flats in $AG_m(F_2)$. Thus, if $\mathcal{A}$ is the affine design of points and $(m - 1)$-flats, we have that

$$\mathcal{R}(1, m) = C_2(\mathcal{A}).$$

The general case is almost as easy. First of all we have that the flats are in the Reed-Muller code:

Proposition 3.9 *The incidence vectors of the $(m - r)$-flats of $AG_m(F_2)$ are all in $\mathcal{R}(r, m)$.*

Proof: Any $(m-r)$-flat T in $AG_m(F_2)$ consists of all the vectors (points of the affine space) $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ that satisfy r linear equations,

$$\sum_{j=1}^{m} a_{ij} X_j = b_i, \text{ for } i = 1, 2, \ldots, r \ ,$$

where all the a_{ij} and b_j are in F_2. The polynomial,

$$\prod_{i=1}^{r}\left(b_i+1+\sum_{j=1}^{m}a_{ij}x_j\right),$$

has degree at most r and thus is in $\mathcal{R}(r,m)$. Moreover it is clearly the characteristic function v^T of T. $\square$

In fact the degree of the polynomial is exactly r when the equations are independent and the proof actually shows that all the $(m-s)$-flats are in $\mathcal{R}(r,m)$ provided $s \le r$.

Theorem 3.10 *Let $\mathcal{A}$ be the design of points and r-flats of the affine geometry $AG_m(F_2)$, where $0 \le r \le m$. Then the binary code $C_2(\mathcal{A})$ is the Reed-Muller code $\mathcal{R}(m-r,m)$. Its dimension is*

$$\binom{m}{0}+\binom{m}{1}+\cdots+\binom{m}{m-r}.$$

Let $\mathcal{P}$ be the design of points and r-dimensional subspaces of the projective geometry $PG_{m-1}(F_2)$ where $1 \le r \le m-1$. Then the binary code $C_2(\mathcal{P})$ is the punctured Reed-Muller code $\mathcal{R}(m-r-1,m)^*$. Its dimension is

$$\binom{m}{0}+\binom{m}{1}+\cdots+\binom{m}{m-r-1}.$$

Proof: The characteristic function of any $(t+1)$-flat is the sum of the characteristic functions of two t-flats contained in it and thus the binary code of the design of points and $(m-r)$-flats contains, by a trivial induction, the characteristic function of every $(m-s)$-flat for $0 \le s \le r$ and hence the code of this design *is* $\mathcal{R}(r,m)$. Reversing the roles of r and $m-r$ gives the first part of the theorem.

For the second part of the theorem, notice first that the code of the design is contained in the punctured Reed-Muller code. Extend the code of the design by an overall parity check and note that this extended code is a subcode of $\mathcal{R}(m-r-1,m)$ and that incidence vectors of the $(r+1)$-dimensional *subspaces* of V generate this extended code. Now every $(r+2)$-dimensional subspace of V has an incidence vector that is the sum of all the incidence vectors of the $(r+1)$-dimensional subspaces it contains. But, over F_2, an $(r+1)$-flat that is not a subspace consists of an $(r+2)$-dimensional subspace from which the points of an $(r+1)$-dimensional subspace, the subspace of which it is a coset, have been removed. Thus, in the code of the design it is the sum of the incidence vectors of an $(r+2)$-dimensional subspace and an $(r+1)$-dimensional subspace. It follows that all the $(r+1)$-flats of V are in the extended code of the design and it is, therefore, $\mathcal{R}(m-r-1,m)$. $\square$

In the proof the essential new point is that subspaces alone generate the Reed-Muller codes because flats can be obtained from subspaces, a fact which makes the discussion of the binary case very easy. We record this as

Corollary 3.11 *The Reed-Muller code $\mathcal{R}(m-r,m)$ is generated by the characteristic functions of the r-dimensional subspaces of F_2^m or, indeed, by the r-flats containing any fixed point of F_2^m.*

The characteristic functions of the r-flats are vectors of weight 2^r and are precisely the minimum-weight vectors of $\mathcal{R}(m-r,m)$, as we shall soon prove. Before doing so, we introduce two exact sequences that arise naturally from the geometric nature of the Reed-Muller and punctured Reed-Muller codes.

Lemma 3.12 *Any embedding of $PG_{m-1}(F_2)$ into $PG_m(F_2)$ gives rise to the following two short exact sequences whenever $0 \leq r < m$:*

(i) $0 \to \mathcal{R}(m-r-1,m)^* \to \mathcal{R}(m-r,m+1)^* \to \mathcal{R}(m-r,m) \to 0$;

(ii) $0 \to \mathcal{R}(m-r-1,m) \to \mathcal{R}(m-r,m+1)^* \to \mathcal{R}(m-r,m)^* \to 0$.

Proof:

Let W be the $(m+1)$-dimensional vector space defining $PG_m(F_2)$. Then an embedding of $PG_{m-1}(F_2)$ in $PG_m(F_2)$ is given by a hyperplane H of W and, moreover, the complement of H in W, $\overline{H} = W - H$, is a copy of $AG_m(F_2)$, as we explained in Section 2.2.

Let $\mathcal{D}$ be the design of points and r-dimensional subspaces of $PG_m(F_2)$. Using $PG(H)$ we form the design $\mathcal{D}_1$ of r-dimensional subspaces in $PG_{m-1}(F_2)$, and from $AG(\overline{H})$ we form the design $\mathcal{D}_2$ of r-flats in $AG_m(F_2)$. By Theorem 3.10, $C_2(\mathcal{D}) = \mathcal{R}(m-r,m+1)^*$, $C_2(\mathcal{D}_1) = \mathcal{R}(m-r-1,m)^*$ and $C_2(\mathcal{D}_2) = \mathcal{R}(m-r,m)$.

Any block of the design $\mathcal{D}$ is either in H or meets it in an $(r-1)$-dimensional (projective) subspace. The intersection with $\overline{H}$ is thus empty or an r-flat; clearly every r-flat of $AG(\overline{H})$ arises in this way. Thus $C = C_2(\mathcal{D})$ projects onto $C_2(\mathcal{D}_2)$, and $C_2(\mathcal{D}_1)$ is in the kernel. Thus $\dim(C_2(\mathcal{D}_1)) \leq$

$\dim(C_2(\mathcal{D})) - \dim(C_2(\mathcal{D}_2))$, and using the formula for the dimension of the Reed-Muller codes, we have

$$\sum_{i=0}^{m-r-1} \binom{m}{i} \le \sum_{i=0}^{m-r} \binom{m+1}{i} - \sum_{i=0}^{m-r} \binom{m}{i}.$$

Using the identity $\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k-1}$ repeatedly, shows that this is actually an equality, and hence that $C_2(\mathcal{D}_1)$ is the whole kernel. This yields the short exact sequence *(i)*.

To obtain the second sequence we use the same embedding but now project C onto the coordinate positions corresponding to the points of $PG(H)$. Let $\mathcal{E}_2$ be the design of points and $(r-1)$-dimensional subspaces of $PG_{m-1}(\mathbf{F}_2)$, and $\mathcal{E}_1$ the design of points and $(r+1)$-flats of $AG_m(\mathbf{F}_2)$. Then $C_2(\mathcal{E}_2) = \mathcal{R}((m-1)-(r-1),m)^*$, and $C_2(\mathcal{E}_1) = \mathcal{R}(m-r-1,m)$. Certainly C projects onto $C_2(\mathcal{E}_2)$ since every r-dimensional projective subspace of $PG(W)$ meets H in an $(r-1)$-dimensional subspace — or is contained in H — and every $(r-1)$-dimensional subspace arises in this way. Two r-dimensional subspaces of $PG(W)$ that meet $PG(H)$ in the same $(r-1)$-dimensional subspace have disjoint intersections in $\overline{H}$ and thus form two cosets of the same r-dimensional subspace of $AG(\overline{H})$. Together they form an $(r+1)$-dimensional space. It follows immediately that the kernel of the projection is $C_2(\mathcal{E}_1)$ and thus yields the sequence *(ii)*. □

We next draw out the consequences of Lemma 3.12 and in so doing prove that the minimum weights of the Reed-Muller codes are as we have indicated and, more importantly, determine the nature of the minimum-weight vectors.

Theorem 3.13 *For $0 \le r \le m$ the minimum weight of $\mathcal{R}(m-r,m)$ is 2^r and the vectors of minimum weight are the incidence vectors of the r-flats of $AG_m(F_2)$. For $1 \le r \le m$ the minimum weight of $\mathcal{R}(m-r,m)^*$ is $2^r - 1$ and the vectors of minimum weight are the incidence vectors of the $(r-1)$-dimensional subspaces of $PG_{m-1}(F_2)$.*

Proof: Clearly the minimum weights are at most 2^r and 2^r-1 by Theorem 3.10. Now we use the short exact sequences and induction on m, the result being trivial for $m=1$. Assume the result true for m and all $r<m$. Thus we assume that $\mathcal{R}(m-s,m)$ has minimum weight 2^s for $m-s \le m$ and $\mathcal{R}(m-s,m)^*$ has minimum weight 2^s-1 for $m-s<m$ and that the minimum-weight vectors are as announced.

Since, for $r = 0$, the result is trivial for any m we may assume $r > 0$ and consider dimension $m+1$. If $r = m$ the results are easy, for then we have $\mathcal{R}(1, m+1)$, a case we have already discussed. We suppose $0 < r < m$ and use the notation of Lemma 3.12. Thus $\mathcal{D}$ is the design of points and r-dimensional subspaces of $PG_m(\mathbf{F}_2)$. Let v be a minimum-weight vector of $C = C_2(\mathcal{D})$, so that $\mathrm{wt}(v) \le 2^{r+1} - 1$. If v is zero at the coordinates corresponding to $\overline{H} = W - H$, then v can be viewed in $C_2(\mathcal{D}_1)$, from the short exact sequence (i), and hence v has weight $2^{r+1}-1$ and is the incidence vector of an r-dimensional subspace of H (and hence of W), by the induction hypothesis. If v is zero at the coordinates corresponding to H, then v can be viewed in $C_2(\mathcal{E}_1) = \mathcal{R}(m-(r+1), m)$, from the short exact sequence (ii), and thus has weight at least 2^{r+1}, which is not possible. Thus v can be taken to have support meeting both H and $\overline{H}$. Again by the induction hypothesis, the weight is at least $2^r + 2^r - 1 = 2^{r+1} - 1$, using the last non-zero terms of the short exact sequences, and hence has exactly this weight. Furthermore, restricted to $PG(H)$, v is the incidence vector of an $(r-1)$-dimensional subspace. To show that v is the incidence vector of an r-dimensional subspace of $PG_m(\mathbf{F}_2)$, construct an r-dimensional subspace of $PG_m(\mathbf{F}_2)$ whose incidence vector w coincides with v on $PG(H)$ and that contains at least one point in $\overline{H}$ in common with the support of v. Then the weight of $v - w$ is easily seen to be less than $2^{r+1} - 1$ and hence $v = w$. This gives the projective result for projective dimension m from which the affine result for dimension $m+1$ follows since the Reed-Muller codes are invariant under translation in V — as we remarked in the last section — which means it is sufficient to consider only those minimum-weight vectors of the Reed-Muller code with a 1 at $\mathbf{0}$. $\square$

It should be noted that the code of any projective-geometry design is cyclic due to the existence of Singer cycles (as already mentioned in Section 2.1) and hence the punctured Reed-Muller codes are cyclic.

We summarize the results obtained on the properties of the Reed-Muller codes and finite geometries over the field F_2:

Theorem 3.14 *Let m be any positive integer.*

(1) If $\mathcal{A}$ is the design of points and r-flats of the affine geometry $AG_m(F_2)$, where $0 \le r \le m$, then the binary code $C = C_2(\mathcal{A})$ is the Reed-Muller code $\mathcal{R}(m-r, m)$. It is a $[2^m, \binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{m-r}, 2^r]$ binary code and the minimum-weight vectors are the incidence vectors of the

r-flats. Further, C contains the incidence vectors of all t-flats for $r \leq t \leq m$.

For $r > 0$ the orthogonal, $C^\perp$, is the Reed-Muller code $\mathcal{R}(r-1, m)$, which is the binary code of the design of points and $(m-r+1)$-flats of the affine geometry $AG_m(F_2)$.

(2) *If $\mathcal{D}$ is the design of points and r-dimensional subspaces of the projective geometry $PG_m(F_2)$, where $0 \leq r \leq m$, then the binary code $C = C_2(\mathcal{D})$ is the punctured Reed-Muller code $\mathcal{R}(m-r, m+1)^*$. It is a $[2^{m+1}-1, \binom{m}{0} + \binom{m+1}{1} + \cdots + \binom{m+1}{m-r}, 2^{r+1}-1]$ binary cyclic code and the minimum-weight vectors are the incidence vectors of the r-dimensional subspaces. Further, C contains the incidence vectors of all t-dimensional subspaces for $r \leq t \leq m$.*

The code orthogonal to $\mathcal{R}(m-r, m+1)^*$ is $\mathcal{R}(r, m+1)^* \cap (F_2\mathbf{1})^\perp$, which is the even-weight subcode of the binary code of the design of points and $(m-r)$-dimensional subspaces of $PG_m(F_2)$.

Example 3.15 (1) The code of the design of points and lines in $AG_4(F_2)$ is $\mathcal{R}(3,4)$, which is the even-weight subcode of F^V. Its orthogonal is $F\mathbf{1} = \mathcal{R}(0,4)$. The code of the design of points and planes is $\mathcal{R}(2,4)$, of dimension 11, with orthogonal the code from the design of points and hyperplanes, of dimension 5, i.e. $\mathcal{R}(1,4)$.

(2) The code of the design of points and lines in $PG_3(F_2)$ is $\mathcal{R}(2,4)^*$, of dimension 11 and minimum weight 3; it is, of course, a binary Hamming code.

(3) A basis consisting of the incidence vectors of lines in $PG_m(F_2)$ for the code $\mathcal{R}(m-2, m)^* = \mathcal{H}_m$ can be found as follows (as described in Key and Sullivan [32]): take any line and include its incidence vector; take any point off the line, and include the three incidence vectors of the three lines joining the new point to the points on the first line. Continue in this way: at each stage, if there is a point not yet incident with a chosen line then simply take all the incidence vectors of the lines joining that point to the points already obtained. These incidence vectors are clearly linearly independent and, as is easily seen, are equal in number to the dimension; hence they yield a basis. The successive dimensions are

$$1,\ 1+3=4,\ 4+7=11,\ 11+15=26,\ 26+31=57,\ \ldots$$

Moreover, the incidence vectors chosen that have a point in common with a fixed point of the first line form a collection of parity checks (of the simplex code that is dual to the Hamming code) that are "focused on" that fixed point — see Section 3.3 — and can be used for majority-logic decoding.

The group $AGL_m(F_2)$, in its natural action on $V = F_2^m$, yields a group of automorphisms of every Reed-Muller code $\mathcal{R}(r,m)$ and $PGL_m(F_2)$, in its natural action on $V^* = V - \{\mathbf{0}\}$, yields a group of automorphisms of every punctured Reed-Muller code $\mathcal{R}(r,m)^*$. That $PGL_m(F_2)$ is the full group of automorphisms of $\mathcal{R}(r,m)^*$ whenever $0 < r < m-1$ follows from Theorem 3.13 and the fundamental theorem of projective geometry, Theorem 2.3. ¿From this it follows that $AGL_m(F_2)$ is the full group of automorphisms of $\mathcal{R}(r,m)$ whenever $0 < r < m-1$. One must be careful here and note that it is not the entire projective space that must be preserved, but only part of it, to ensure that the automorphism comes from the general linear group.

3.3 Decoding

One of the attractions of Reed-Muller codes is the simple and easily implemented decoding scheme that is available, with decoding decisions made by majority vote, just as with the repetition code — which is, of course, the simplest Reed-Muller code, $\mathcal{R}(0,m)$. Since the scheme is related to the geometric nature of these codes we describe it here. The scheme dates from the very beginning of coding theory and is due to Reed [46]. It was Reed's algorithm that prompted the investigation of majority-logic decoding and the rather peculiar definition of so-called Euclidean-geometry codes as maximal cyclic subspaces of duals of the codes generated by certain flats in $AG_m(F_{2^s})$. We begin by describing majority-logic decoding.

Let C be an arbitrary linear code contained in the ambient space F_q^n. Recall that a parity check is simply a code vector in the orthogonal code, $C^\perp$, and that the support of a vector in F_q^n is the set of coordinate positions in which it has non-zero entries. Suppose we are given J parity checks and a coordinate position, i say, such that the intersection of the supports of any two of the given parity checks is precisely the singleton set $\{i\}$. If a received vector has been perturbed by t or fewer errors during transmission, where $2t \leq J$, then, clearly, at least half of the parity checks will give zero (i.e. check) when applied to the received vector *unless* the symbol at coordinate i

is in error. Moreover, had we normalized the given parity checks so that each had a 1 at coordinate i, then, in the event that one of the t or fewer errors had occured at coordinate i, at least half of the parity checks would record that error. Thus a majority "vote" of the values of the parity checks corrects the entry at coordinate i. This is the essence of majority-logic decoding.

Such a collection of parity checks is said to be **focused on $\boldsymbol{i}$**[2]. Note that if each of the coordinates of C has a collection of J parity checks focused on it, then the code will necessarily correct t or fewer errors — where again $2t \leq J$ — and therefore C must have minimum weight at least $2t+1$. Indeed, it is very easy to see that if there is a set of J parity checks focused on a coordinate i, then any code vector with a non-zero entry at i must have weight at least $J+1$ in order to satisfy the J parity checks. Note also that any code with a transitive automorphism group (and, in particular, a cyclic code) will have minimum weight at least $J+1$ provided that one, and hence all coordinates, has a collection of J parity checks focused on it. In the cyclic case majority-logic decoding, when it is available, is particularly simple.

An instructive example is the dual C of a binary Hamming code, frequently referred to as a simplex code. It has the classical Steiner triple system, namely the lines of the projective geometry, among its parity checks and the pencil of lines through a point of the geometry fulfills the requirements for a collection of parity checks focused on the given point; if the Hamming code is of block length 2^m-1, then $J=2^{m-1}-1$. Indeed, C has minimum weight 2^{m-1} and $t=2^{m-2}-1$ errors can be corrected. A simpler, but still important, case is the following:

Example 3.16 The repetition code of length n over F_q has, clearly, $n-1$ parity checks of weight 2 focused on any given coordinate and, for odd n, one simply uses a majority vote to determine the symbol sent, obtaining the correct symbol provided at most $\frac{n-1}{2}$ errors occurred during transmission.

To use majority logic to correct many errors one must have many parity checks focused on each coordinate, which entails that the minimum weight $d^\perp$ of $C^\perp$ be small; in fact, in order to have J parity checks focused on a given coordinate we must have $(d^\perp-1)J \leq n-1$. The examples above have the

[2]Unfortunately the term "orthogonal on i" is the terminology of most of the coding literature. Blahut, recognizing the problem with that terminology, used "concurrent on i" in [7], but that does not seem to have been adopted. We here make another attempt at change.

smallest possible minimum weights for their duals and allow error-correction via majority logic to correct up to the full error-correcting capacity of the code. In the coding literature such codes are said to be "completely orthogonalizable".

Consider next the Reed-Muller code $C = \mathcal{R}(r, m)$ where $r < m$. A basis for C is the set of monomials of degree r or less. The idea of Reed's decoding scheme is to determine first the "information bits" corresponding to monomials of degree r, thus reducing the problem to decoding in the Reed-Muller code $\mathcal{R}(r-1, m)$. Let K be a subset of $\{1, 2, \ldots, m\}$ of cardinality r and let L be the complement of K. Now the monomial of degree r,

$$\prod_{k \in K} x_k,$$

as an element of $C = \mathcal{R}(r, m)$, is the characteristic function of the $(m-r)$-flat S given by the equations

$$X_k = 1, k \in K.$$

Also

$$\prod_{l \in L} x_l$$

is the characteristic function of the r-flat T given by

$$X_l = 1, l \in L$$

and, moreover, is an element of $\mathcal{R}(r-1, m)^{\perp}$. Each of the 2^{m-r} translates of T meets the flat S precisely once, but any other $(m-r)$-flat given by a different monomial of degree r evenly. (To see this the reader may wish to think of *subspaces* of the relevant dimensions; in one case the intersection is the zero vector and S is a transversal to the 2^{m-r} translates of T; in the other the intersection is a subspace of positive dimension and S meets a translate of T either in a flat of that dimension or not at all.) Thus, a majority vote of the parity checks corresponding to these 2^{m-r} translates will record only the information bit corresponding to $\prod_{k \in K} x_k$ *provided* fewer than $2^{m-r-1} - 1$ errors have been made in transmission. Note that one retrieves the information bit directly by majority vote and that, after determining those information bits corresponding to the $\binom{m}{r}$ monomials of degree r, the received vector is adjusted and decoding proceeds in $\mathcal{R}(r-1, m)$ via precisely the same method.

Finally, we make contact, briefly, with so-called *L-step majority-logic decoding*. In our discussion of Reed's algorithm we used parity checks which were *not* in the dual of the code in question: the r-flat T above has a characteristic function in $\mathcal{R}(r-1,m)^{\perp}$ but not in $\mathcal{R}(r,m)^{\perp}$. However, if T' is any tranlate of T then $T \cup T'$ is an $(r+1)$-flat whose characteristic function *is* in $\mathcal{R}(r,m)^{\perp}$. Moreover, the $2^{m-r}-1$ flats of this form (namely, $T \cup T'$, where T' is a distinct translate of T) are **focused on $\boldsymbol{T}$** in the sense that the intersection of the supports of any two of them is precisely the set T. Thus, provided sufficiently few errors were made in transmission, a majority vote using these parity checks will give the sum of the error bits contained in the coordinate positions corresponding to the flat T. Such a "divide and conquer" technique using majority or threshold circuitry was thoroughly investigated early in the history of coding theory and was the subject of Massey's thesis, [41]. For a fuller discussion of the decoding of Reed-Muller and Generalized Reed-Muller codes and L-step majority-logic decoding the reader may wish to consult [40] or a textbook on error control, for example, [7] or [36].

4 The group-algebra approach

We have not, so far, taken full advantage of the fact that the coordinate set of the codes in question is, itself, endowed with structure. We did, of course, use that structure in defining the Singer cycles — where the coordinate set was given the structure of a field — and, moreover, we used the additive structure to discuss the minimum-weight vectors and the automorphism groups of the Reed-Muller codes. But, for example, we have not yet explicitly shown how to get the generator polynomials of the punctured Reed-Muller codes although we know they are cyclic. In fact, of course, there are many Singer cycles and the codes are therefore cyclic in many ways — which is another way of saying that one must specify explicitly how a code is cyclic before one can compute the generator polynomial. The group-algebra approach allows one to naturally specify the roots of the generator polynomial without actually choosing the Singer cycle and it is this intrinsic nature of the approach which makes the group-algebra setting so attractive.

We follow Charpin [14] but Landrock and Manz [34] have also given an expository account; the original source of this approach was Berman's seminal paper, [6]. One exploits the modular group algebra of an elementary abelian[3] group, the additive group of the field that labels the coordinates,

[3]The group, in fact, need not be elementary abelian nor even abelian and the general case has been treated; see, for example, the chapter by Ward in this Handbook. But, if one

as the ambient space for the codes.

4.1 Elementary results and Berman's theorem

We proceed in full generality but the reader interested only in Berman's result and the Reed-Muller codes can take G below to be F_2^m and F to be F_2.

Let $q = p^m$ and set G equal to the additive subgroup of the field F_q. We will very soon regard G as F_q, but for the moment it is merely an elementary abelian p-group[4] of order p^m. Let F be any subfield of F_q and set $\boldsymbol{R} = F[G]$, the group algebra of G over the field F. Recall that the elements of $\boldsymbol{R}$ are simply functions from G to F and therefore, when G and F are taken as suggested above, $\boldsymbol{R}$ is the ambient space of the Reed-Muller codes. We choose, however, to formulate things a bit differently and view the group algebra in a polynomial way — as one frequently does with group algebras given by abelian monoids[5]. Thus a typical element of $\boldsymbol{R}$ is a formal sum $\sum_{g\in G} x_g X^g$ where the x_g are elements of F and, as a function, it is simply the one that assigns x_g to the element g of G. Addition and scalar multiplication are component-wise and the multiplication is given by the addition in G. Thus,

$$\sum_{g\in G} x_g X^g + \sum_{g\in G} y_g X^g = \sum_{g\in G} (x_g + y_g) X^g$$

and, for $c \in F$,

$$c(\sum_{g\in G} x_g X^g) = \sum_{g\in G} (cx_g) X^g;$$

using the "polynomial" multiplication $X^g X^h = X^{g+h}$ gives the usual multiplication formula in the group algebra:

$$(\sum_{g\in G} x_g X^g)(\sum_{h\in G} y_h X^h) = \sum_{g,h\in G} x_g y_h X^{g+h} = \sum_{k\in G}(\sum_{h\in G} x_{k-h} y_h) X^k.$$

Notice that X^0 is the unit element of the commutative ring $\boldsymbol{R}$; i.e. $X^0 a = a$ for every $a \in \boldsymbol{R}$. The augmentation map $\boldsymbol{R} \to F$ given by $\sum_{g\in G} x_g X^g \mapsto$

restricts oneself to p-groups, a result of Faldum's [20] shows that one might as well restrict oneself to the elementary abelian case — as far as producing "good" codes is concerned.

[4] That is, G is an abelian group all of whose non-identity elements have order p or, in other words, a vector space over the field F_p. Since the group operation is being written additively "order p" means simply that $pg = 0$ for every $g \in G$.

[5] The paradigm is the ordinary polynomial ring where the monoid in question is the set of non-negative integers under addition.

$\sum_{g\in G} x_g$ is clearly a linear transformation of the vector space structure of $\boldsymbol{R}$ onto F; moreover, it is an algebra homomorphism — as one can easily check from the multiplication formula. We denote the kernel of this augmentation map by $\boldsymbol{M}$; it is, of course, an ideal of $\boldsymbol{R}$, but much more is true: since we are in characteristic p we have the Frobenius homomorphism, $a \mapsto a^p$, at our disposal and the fact that G is an elementary abelian p-group gives

$$(\sum_{g\in G} x_g X^g)^p = \sum_{g\in G} x_g^p X^0 = (\sum_{g\in G} x_g^p) X^0 = (\sum_{g\in G} x_g)^p X^0,$$

which shows that every element not in $\boldsymbol{M}$ is invertible in $\boldsymbol{R}$ and hence that $\boldsymbol{M}$ is the unique maximal ideal of $\boldsymbol{R}$.

In the binary case, with the interpretation suggested above, $\boldsymbol{M}$ is the Reed-Muller code $\mathcal{R}(m-1, m)$; we shall shortly see that the powers of the ideal $\boldsymbol{M}$ give precisely the Reed-Muller codes.

Observe that in our present notation the characteristic function of a subset S of G is given by the element $\sum_{g\in S} X^g$ of the group algebra. Consider next the element $X^g - X^0 = X^g - 1$ of the ideal $\boldsymbol{M}$, where we have set $X^0 = 1$ since it is the unit element of $\boldsymbol{R}$. Provided $g \neq 0$,

$$(X^g - 1)^{p-1} = \sum_{i=0}^{p-1} (-1)^{p-1-i} \binom{p-1}{i} X^{ig} = (-1)^{p-1} \sum_{i=0}^{p-1} X^{ig} = \sum_{h\in U} X^h$$

where $U = \langle g \rangle = \{ig \mid 0 \leq i < p\}$ is the subspace over F_p spanned by g. We have here used the fact that $(-1)^{p-1} = 1$, even when $p = 2$, and the fact that $\binom{p-1}{i} = (-1)^i$ since we are working in a field of characteristic p. Moreover, if we are given a set of linearly independent elements of G, $g_1, g_2, \ldots, g_r$ say, then one checks easily that $\prod_{i=1}^r (X^{g_i} - 1)^{p-1} = \sum_{g\in U} X^g$, where now U is the subspace spanned by $\{g_1, g_2, \ldots, g_r\}$. In fact, we have the following more precise statement:

Lemma 4.1 *Let S be a non-empty subset of G. Then*

$$\prod_{g\in S} (X^g - 1)^{p-1} = \begin{cases} \sum_{g\in\langle S\rangle} X^g & \textit{if } S \textit{ is a linearly independent set} \\ 0 & \textit{otherwise} \end{cases}$$

Proof: We have already remarked on the case of a linearly independent set S, so suppose S is linearly dependent. We wish to show that the product is zero. If $0 \in S$, that result is immediate; otherwise, let S' be a linearly independent subset of S with the property that there is a $g_0 \in S - S'$

contained in $\langle S'\rangle$. Then the product in question is $(\prod_{g\in S'}(X^g-1)^{p-1})(X^{g_0}-1)^{p-1}a$ where a is an element of $\boldsymbol{R}$.

By the first part of the lemma,

$$(\prod_{g\in S'}(X^g-1)^{p-1})(X^{g_0}-1)^{p-1}=(\sum_{g\in\langle S'\rangle}X^g)(\sum_{h\in\langle g_0\rangle}X^h)=\sum_{h\in\langle g_0\rangle}\sum_{g\in\langle S'\rangle}X^{g+h}.$$

Since $g+h$ runs through $\langle S'\rangle$ as g does for every $h\in\langle g_0\rangle$, this latter sum is $\sum_{h\in\langle g_0\rangle}\sum_{g\in\langle S'\rangle}X^g=p\sum_{g\in\langle S'\rangle}X^g=0$ and we have the result. $\square$

Now since the ideal $\boldsymbol{M}$ is generated *linearly* over the field F by the elements X^g-1, the ideal $\boldsymbol{M}^r$ is generated linearly by elements of the form $\prod_{g\in S}(X^g-1)$ where S is a subset of G of cardinality r. Moreover, in characteristic 2, the subsets S can be taken to be linearly independent subsets of the vector space G over F_2, by the above Lemma. Hence in this binary case $\boldsymbol{M}^r$ is generated linearly by the characteristic functions of the r-dimensional subspaces of the vector space G over F_2. Because of the simple result (Corollary 3.11) that $\mathcal{R}(m-r,m)$ is generated by the characteristic functions of *subspaces* of dimension r, we have proved Berman's theorem:

Theorem 4.2 *In the group algebra $F_2[G]$, where G is an m-dimensional vector space over F_2, the Reed-Muller code $\mathcal{R}(m-r,m)=\boldsymbol{M}^r$, where $\boldsymbol{M}$ is the unique maximal ideal of $F_2[G]$.*

Remark: The theorem is even true for $r=0$ provided we define $\boldsymbol{M}^0=\boldsymbol{R}$, as is customary, it being the ideal generated by 1. Observe that for $r=m$ we have the repetition code and, indeed, $\prod_{g\in B}(X^g-1)=\mathbf{1}$ for every basis B of G.

4.2 Isometries of the group algebra

If $R=F[G]$ is the group algebra of any group G, abelian or not, then any automorphism σ of the group G induces an automorphism of R, which we also denote by σ, via

$$\sigma(\sum_{g\in G}x_gX^g)=\sum_{g\in G}x_gX^{\sigma(g)},$$

as one can easily check. Moreover, in the basis given by X^g, the coding-theory basis we have chosen, such an automorphism is weight preserving — i.e. it is also an **isometry** preserving the Hamming metric. If σ is *any* automorphism of the algebra R that is weight preserving, then it must be given monomially; i.e. $\sigma(X^g)$ must be of the form aX^h for some $h \in G$ and $a \in F^\times$; for G an elementary abelian p-group we have — since automorphisms preserve the unit element — $\sigma(X^0) = \sigma((X^g)^p) = a^p X^0 = X^0$, which implies that $a = 1$, since F is of characteristic p, and that the automorphism is given by a coordinate permutation. But now one easily checks that setting $h = \sigma(g)$ defines an automorphism of G that induces the given isometry. In the case of an elementary abelian p-group G, the automorphism group is simply $GL(G)$ where G is viewed as a vector space over the field F_p. In the case at hand we can choose a basis for G and then $GL_m(F_p)$, where $|G| = p^m$, is *precisely* the group of isometric automorphisms of our group algebra. We record this fact with

Proposition 4.3 *The group of isometric automorphisms of the group algebra $F[G]$, where G is an elementary abelian p-group and F a field of characteristic p, is $GL(G) = \mathrm{Aut}(G)$ in its natural action on the coordinate set G.*

A group algebra, $F[G]$, comes canonically equipped with an involutory anti-automorphism induced by the map $G \to G$ which sends a group element to its inverse. When G is abelian this canonical map is an involutory *automorphism* and clearly isometric. We denote this canonical involution by $x \mapsto \overline{x}$; in $GL(G)$ it is represented by the map $g \mapsto -g$. As we shall see this canonical automorphism plays an important role when discussing the orthogonal of a code viewed in $\boldsymbol{R}$. It is the analogue of taking the "reverse" when computing orthogonals to cyclic codes.

The ideal $\boldsymbol{M}$ of $\boldsymbol{R}$ is intrinsically defined since it consists of the nilpotent[6] elements of $\boldsymbol{R}$. It follows that every automorphism fixes $\boldsymbol{M}$ and hence all powers of $\boldsymbol{M}$; in particular, isometric automorphisms fix $\boldsymbol{M}^r$ for all r. There are isometries not given by automorphisms, of course. For example, multiplication by X^g yields an isometry of $\boldsymbol{R}$; such an isometry is clearly given by a translation in the vector space G. Since any ideal of $\boldsymbol{R}$ is fixed by such a multiplication, all the powers of $\boldsymbol{M}$ are fixed. We thus have $AGL(G)$ acting as a group of isometries of $\boldsymbol{R}$ and fixing the ideals that are here of interest. We have now explained in our new language — but in a more general setting — what we already know about the Reed-Muller codes.

[6] An element of a ring is nilpotent if some power of it is 0; in our case the generators of M, *viz.* $X^g - 1$, are 0 when raised to the p^{th} power and it follows easily that all elements of M are nilpotent. M is the *Jacobson radical* of the ring R.

In the early history of coding theory there was great interest in deciding which extended cyclic codes were "affine invariant" and Kasami, Lin and Peterson [29] settled the question. Because of the historic interest and the motivation it will provide for the rest of this chapter, we discuss and prove their result. We are here, as in all of this section, following Charpin [14].

First of all it must be emphasized that "affine invariant" refers not to the group of isometries discussed above but to a smaller group; a more precise name would be "translation-invariant extended cyclic codes". The point is that one does not demand invariance under the group $AGL(G)$, but only under the subgroup $AGL_1(F_q)$, where now we are viewing G as the field F_q. There are many more codes invariant under this smaller group, even if one insists that the codes be self-dual: see, for example, [17] where all binary, affine-invariant self-dual codes of block length at most 512 have been found and where evidence is presented to suggest that the number goes to infinity with the admissible block length. On the other hand, the only binary codes invariant under the larger group are the Reed-Muller codes (see Theorem 4.17 below).

We shall see in a moment how to extend the cyclic codes in question so that they will lie in the ideal $\boldsymbol{M}$, but let us note first that a linear subspace of $\boldsymbol{R}$ invariant under translation is simply an ideal of $\boldsymbol{R}$. Our aim, therefore, is to characterize those ideals invariant under the isometric automorphisms given by $X^g \mapsto X^{ug}$ where u is a non-zero field element and where we have identified G with F_q.

4.3 Translation-invariant extended cyclic codes

We prove here the theorem of Kasami, Lin and Peterson characterizing "affine-invariant" cyclic codes.

Set $n = p^m - 1$ and suppose $C \subset F^n$ is a cyclic code. Now C is specified completely by the n^{th} roots of unity that are roots of its generator polynomial. We shall assume that 1 is not a root for we wish to extend C by an overall parity check and we wish to avoid trivial cases. Let α be a primitive n^{th} root of unity, i.e. a primitive element of F_q, where $q = p^m$. Then the set of roots of the generator polynomial are specified by that subset T of $\{1, 2, \dots, n-1\}$ where α^i is a root if and only if $i \in T$. We shall refer to T as the **defining set** of the cyclic code C. We embed C in $\boldsymbol{R}$ as follows:

$$(c_0, c_1, \dots, c_{n-1}) \mapsto (-\sum_{i=0}^{n-1} c_i)X^0 + \sum_{i=0}^{n-1} c_i X^{\alpha^i}.$$

Clearly the image, which we denote by $\widehat{C}$, is invariant under the map

$$\sum_{g\in G} x_g X^g \mapsto \sum_{g\in G} x_g X^{\alpha g}$$

and, indeed, any linear subspace of $\boldsymbol{R}$ invariant under this map comes from a cyclic code. Since α is a generator of $F_q^\times$ the image is invariant under the maps given by $X^g \mapsto X^{ug}$ for *all* non-zero $u \in F_q$. We have, by our choice of α, embedded all cyclic codes over F in $\boldsymbol{M}$.

Consider next the following F-linear maps ϕ_s of $\boldsymbol{R}$ into the space $G = F_q$:

$$\phi_s(\sum_{g\in G} x_g X^g) = \sum_{g\in G} x_g g^s,$$

where $0 \leq s < n$. With the proviso that $0^0 = 1$, the map ϕ_0 is simply the augmentation map with kernel $\boldsymbol{M}$. Observe that if i is in the defining set of a cyclic code C, then the image, $\widehat{C}$, of C has the property that $\phi_i(c) = 0$ for all $c \in \widehat{C}$. Moreover, for any embedded cyclic code, $\phi_0(c) = 0$ for all $c \in \widehat{C}$. Thus, we "extend" T to $\widehat{T} = T \cup \{0\}$ and note that the image of the cyclic code is defined by $\widehat{T}$ in the sense that $c \in \widehat{C}$ if and only if $\phi_i(c) = 0$ for all $i \in \widehat{T}$. Hence we abuse the terminology and refer to $\widehat{T}$ as the **defining set** of $\widehat{C}$.

Unlike ϕ_0, ϕ_s is *not* an algebra homomorphism for $s > 0$. It does, however, have an important multiplicative property which we now explain. Let $N = \{0, 1, \ldots, n\}$ and define a partial order on N by $k \preceq l$ if and only if $k_\nu \leq l_\nu$ for all ν, where $k = \sum_{\nu=0}^{m-1} k_\nu p^\nu$ and $l = \sum_{\nu=0}^{m-1} l_\nu p^\nu$ are the p-ary expansions of k and l. We give $k \prec l$ the obvious meaning: $k \preceq l$ but $k \neq l$. Then

Proposition 4.4 *For all $x, y \in \boldsymbol{R}$,*

$$\phi_s(xy) = \sum_{i \preceq s} \binom{s}{i} \phi_i(x)\phi_{s-i}(y).$$

Proof: Setting $x = \sum_{g\in G} x_g X^g$ and $y = \sum_{h\in G} y_h X^h$ and writing out the definition of $\phi_s(xy)$ yields

$$\phi_s(xy) = \sum_{i=0}^{s} \binom{s}{i} \phi_i(x)\phi_{s-i}(y)$$

and an application of Lucas's theorem[7] gives the result. □

We have immediately the following

Corollary 4.5 *If I is an ideal of $\boldsymbol{R}$ and $\phi_s(x) = 0$ for all $x \in I$, then $\phi_i(x) = 0$ for all $x \in I$ and $i \preceq s$.*

Proof: Since $xX^g \in I$ for all $x \in I$ the formula above yields

$$\phi_s(xX^g) = \sum_{i \preceq s} \binom{s}{i} \phi_i(x) g^{s-i} = 0$$

for all $g \in G^\times$ and unless the $\phi_i(x) = 0$ for all $i \preceq s$ we would have a non-zero polynomial, $\sum_{i \preceq s} \binom{s}{i} \phi_i(x) Z^{s-i}$, of degree less than n with n roots, namely the elements of $G^\times$. □

The discussion above and the corollary yield the theorem of Kasami, Lin and Peterson:

Theorem 4.6 *A cyclic code of block length $p^m - 1$ has an extension which is translation invariant if and only if its defining set T does not contain 0 and has the property that $s \in T$ implies $i \in T$ for all $i \preceq s$.*

Proof: Clearly an extended cyclic code that is translation invariant is an ideal and hence its defining set has the required property. On the other hand if an extended cyclic code has a defining set with the required property the formula shows that $\phi_s(cX^g) = 0$ for all $c \in \widehat{C}$ and all $s \in T$ or, in other words, that $c \in \widehat{C}$ implies $cX^g \in \widehat{C}$ for all $g \in G$ or that $\widehat{C}$ is translation invariant. □

The powers of $\boldsymbol{M}$ are, of course, extended cyclic codes that are translation invariant since they are invariant under the larger group $AGL(G)$. Thus we should be able to determine their generator polynomials. The reader should observe that these polynomials will depend on the choice of α, but the defining sets are intrinsic to $\boldsymbol{R}$ since they are given by the appropriate ϕ_i's. This intrinsic nature of the group-algebra approach has been exploited in diverse directions by Charpin and her students. The interested reader may wish to consult [16, 17]. In the following section we give the promised defining sets for the Reed-Muller codes and look briefly at their p-ary analogues.

[7] Lucas's theorem states that $\binom{s}{i} \equiv \prod_\nu \binom{s_\nu}{i_\nu}$ modulo p. Hence $\binom{s}{i}$ is non-zero if and only if $i \preceq s$.

4.4 The generator polynomials of punctured Reed-Muller codes and their p-ary analogues

We will in fact determine the defining set of $\mathcal{R}(r,m)$. If that set is $\widehat{T} = T \cup \{0\}$ and α is the chosen n^{th}-root of unity then the generator polynomial of $\mathcal{R}(r,m)^*$ is simply $\prod_{i \in T}(Z - \alpha^i)$. Now $\mathcal{R}(r,m) = \boldsymbol{M}^{m-r}$ and we are interested in the case where $r < m$. For $r = m-1$ we know that $T = \emptyset$ since $\boldsymbol{M}$ is, clearly, annihilated only by ϕ_0 and, of course, $\mathcal{R}(m-1,m)^* = F_2^{2^m-1}$, the whole ambient space. Since the dimension of $\mathcal{R}(r,m)$ is $k = \sum_{i=0}^{r}\binom{m}{i}$ we know that $|T| = 2^m - 1 - k = \sum_{i=0}^{m-r-1}\binom{m}{i} - 1$ and that therefore $|\widehat{T}| = \sum_{i=0}^{m-r-1}\binom{m}{i}$. This is the cardinality of the set of integers less than n whose binary expansions have fewer than $m-r$ entries equal to 1. As the next proposition will show, $\widehat{T}$ is precisely this defining set and therefore the generator polynomial of $\mathcal{R}(r,m)^*$ is

$$\prod_{0<\mathrm{wt}_2(i)<m-r} (Z - \alpha^i)$$

where wt_2 is the function given by the following more general

Definition 4.7 *For any integers, $k \geq 0$ and $q > 1$, the q-**weight** of k, written $\mathrm{wt}_q(k)$, is*

$$\mathrm{wt}_q(k) = \sum_{\nu=0}^{\infty} k_\nu,$$

where $k = \sum_{\nu=0}^{\infty} k_\nu q^\nu$ is the q-ary expansion of k.

Proposition 4.8 *The defining set of the ideal $\boldsymbol{M}^t$ in $F_2[G]$, where G is the elementary abelian 2-group of order 2^m, is that subset of $\{0,1,\ldots,2^m-2\}$ whose elements have binary expansions containing fewer than t entries equal to 1. That is, the defining set is $\{i \mid 0 \leq i < 2^m - 1$ and $\mathrm{wt}_2(i) < t\}$.*

Proof: We use induction on t. We have the result for $t = 1$ since 0 is the only integer k with $\mathrm{wt}_2(k) = 0$. Suppose the result true for t and consider $t+1$. Now, a typical generating element of $\boldsymbol{M}^{t+1}$ is of the form $x(X^g - 1)$ where $x \in \boldsymbol{M}^t$. By the nested nature of the ideals we know, of course, that $\boldsymbol{M}^{t+1}$ is annihilated by all ϕ_s with $\mathrm{wt}_2(s) < t$ and we need only show that it is annihilated by those ϕ_s with $\mathrm{wt}_2(s) = t$. For such an s we have that

$$\phi_s(x(X^g - 1)) \quad = \quad \sum_{i \preceq s} \binom{s}{i} \phi_i(x)\phi_{s-i}(X^g - 1)$$

$$= \quad \phi_s(x)\phi_0(X^g - 1) + \sum_{i \prec s} \binom{s}{i} \phi_i(x)\phi_{s-i}(X^g - 1).$$

But the first summand on the right side is 0 since $X^g - 1 \in \boldsymbol{M}$ and the second summand is 0 since $i \prec s$ implies $\mathrm{wt}_2(i) < \mathrm{wt}_2(s) = t$. Since we know the dimension of the ideal, we must have precisely the defining set. □

The above proof is due to Charpin [13]; observe that it does not depend on the fact that we are in characteristic 2 and, therefore, proves more. We have, in fact, proved the following

Proposition 4.9 *Let $\boldsymbol{R} = F[G]$ where G is an elementary abelian p-group and F a field of characteristic p. Then the ideal $\boldsymbol{M}^t$, where $\boldsymbol{M}$ is the ideal of nilpotent elements of $\boldsymbol{R}$, is annihilated by all ϕ_s with $\mathrm{wt}_p(s) < t$.*

In the event that the field F is *not* a subfield of G one must take an overfield of both in order to have a target for the functions ϕ_s, but this does not effect the proof.

Of course $\boldsymbol{M}^t$ is an extended cyclic code invariant under translation, but since we have not yet computed its dimension, we cannot assert that we have its defining set — as we did in the binary case. The proposition does, however, show that the dimension is at most equal to $|\{i \,|\, 0 \le i < p^m - 1, \mathrm{wt}_p(i) \ge t\}|$ since we are in the presence of an extended cyclic code — which means that $\dim_F(\boldsymbol{M}^t) = p^m - |\widehat{T}|$, where $\widehat{T}$ is the defining set of $\boldsymbol{M}^t$. We will soon exhibit linearly independent elements that will give us not only this dimension but also the so-called "Jennings[8] Basis" of the algebra $F[G]$.

Let $\{g_0, g_1, \ldots, g_{m-1}\}$ be a basis of the F_p-space G. For any $k = \sum_{\nu=0}^{m-1} k_\nu p^\nu$, where $0 \le k_\nu < p$ for all ν, set $J_k = \prod_{\nu=0}^{m-1}(X^{g_\nu} - 1)^{k_\nu}$. Clearly $J_k \in \boldsymbol{M}^t$ whenever $\mathrm{wt}_p(k) \ge t$. Moreover, these elements are linearly independent over F, where F is any field of characteristic p. For suppose $\sum_{\mathrm{wt}_p(k) \ge t} a_k J_k = 0$, where all $a_k \in F$. Choose j such that $\mathrm{wt}_p(j)$ is a minimum with $a_j \ne 0$ and set $j = \sum_{\nu=0}^{m-1} j_\nu p^\nu$. Multiplying the linear relation by $\prod_{\nu=0}^{m-1}(X^{g_\nu} - 1)^{p-1-j_\nu}$, bearing in mind that $(X^g - 1)^p = 0$ for any g, yields $a_j \prod_{\nu=0}^{m-1}(X^{g_\nu} - 1)^{p-1} = a_j \mathbf{1} = 0$ by Lemma 4.1, and hence that $a_j = 0$.

We have thus proved the following

[8] In fact, this basis first appeared in a paper by Lombardo-Radice, [38]. Lombardo-Radice goes over the same ground as Jennings did ([28]) but only for abelian groups; Jennings was aware of the work of Lombardo-Radice and extended that work to arbitrary p-groups.

Theorem 4.10 *Let G be an elementary abelian p-group of order p^m and F a field of characteristic p. For any basis $\{g_0, \dots, g_{m-1}\}$ of G, the p^m elements*

$$\prod_{\nu=0}^{m-1} (X^{g_\nu} - 1)^{e_\nu}$$

where $0 \le e_\nu < p$ form a linear basis for $\boldsymbol{R} = F[G]$. Moreover,

$$\{\prod_{\nu=0}^{m-1} (X^{g_\nu} - 1)^{e_\nu} \mid \sum_{\nu=0}^{m-1} e_\nu \ge t, 0 \le e_\nu < p\}$$

form a basis of $\boldsymbol{M}^t$, where $\boldsymbol{M}$ is the radical of $\boldsymbol{R}$.

Such a basis for $F[G]$ was exploited by Jennings [28] and is called a **Jennings basis** of the group algebra. It is, as the construction shows, independent of the coefficient field of the modular algebra and simultaneously exhibits bases for all powers of the radical.

We note here that the index of nilpotency of the radical is $1 + m(p-1)$; i.e.

$$\boldsymbol{M}^{1+m(p-1)} = 0$$

but $\boldsymbol{M}^k \neq 0$ for any smaller k. Just as in the binary case, $\boldsymbol{M}^{m(p-1)}$ is the repetition code generated by $\prod_{\nu=0}^{m-1}(X^{g_\nu} - 1)^{p-1} = \sum_{g \in G} X^g = \mathbf{1}$; it is *the* minimal ideal of $\boldsymbol{R}$, which means that it is contained in every non-zero ideal of $\boldsymbol{R}$, a fact that is easily seen using the Jennings basis.

Corollary 4.11 *The code $\boldsymbol{M}^t$ is a code of block length p^m, dimension $|\{k|0 \le k < p^m, \mathrm{wt}_p(k) \le m(p-1) - t\}|$ and minimum weight $(b+1)p^a$, where $t = a(p-1) + b$ with $0 \le b < p-1$. As an extended cyclic code its defining set is $\{i \,|\, 0 \le i < p^m - 1,\ \mathrm{wt}_p(i) < t\}$.*

Proof: The dimension is $|\{k|0 \le k < p^m, \mathrm{wt}_p(k) \ge t\}|$, of course, but taking the set of complements, $(p^m - 1) - k$, gives the above description — which is sometimes more useful. As for the minimum weight, the BCH bound implies that the minimum weight is at least as announced, since $k = \sum_{\nu=0}^{a-1}(p-1)p^\nu + bp^a = (b+1)p^a - 1$ is the smallest integer with $\mathrm{wt}_p(k) = t$. On the other hand, $(X^{g_0} - 1)^b \prod_{\nu=1}^{a}(X^{g_\nu} - 1)^{p-1}$ yields a vector of the given weight since the product is the characteristic function of the a-dimensional subspace generated by $\{g_1, \dots, g_a\}$, and multiplying by $(X^{g_0} - 1)^b$ merely takes the sum of $b+1$ distinct, weighted translates.□

Observe that the minimum-weight vectors we have exhibited have their supports lying in an $(a+1)$-dimensional subspace of G, namely the subspace generated by $\{g_0, g_1, \ldots, g_a\}$. When t is divisible by $p-1$ these minimum-weight vectors *are* characteristic functions of subspaces — just as in the binary case.

Corollary 4.12 *If $t = a(p-1)$ then $\boldsymbol{M}^t$ contains the incidence vector of every a-flat of the affine geometry $AG_m(F_p)$ and hence the code over F_p of the design of points and a-flats of $AG_m(F_p)$.*

Example 4.13 There is a simple formula, easily derived, for the dimension of $\boldsymbol{M}^{(m-1)(p-1)}$, since it is the number of ways of selecting at most $p-1$ objects — repetitions allowed — from a set of m objects. One has then (cf. Example 5.6) that

$$\dim(\boldsymbol{M}^{(m-1)(p-1)}) = \binom{m+p-1}{m}$$

and that among the minimum-weight vectors one finds the characteristic functions of flats of codimension 1. As we shall see it *is* the code over F_p of this affine design.

4.5 Orthogonals and annihilators

We have already seen that for the Reed-Muller codes

$$\mathcal{R}(r,m)^{\perp} = \mathcal{R}(m-r-1,m)$$

for $0 \le r < m$, or — in the current language — that the Reed-Muller code $(\boldsymbol{M}^{m-r})^{\perp}$ is precisely $\boldsymbol{M}^{r+1}$. Moreover, the same equality is true if we replace the orthogonal by the annihilator in the group algebra.

More precisely, if S is any subset of $\boldsymbol{R}$ we set

$$\mathrm{Ann}(S) = \{x \in \boldsymbol{R} \,|\, xs = 0 \text{ for all } s \in S\}.$$

Then, since $\boldsymbol{M}^{m(p-1)+1} = \{0\}$, $\mathrm{Ann}(\boldsymbol{M}^t) \supseteq \boldsymbol{M}^{m(p-1)+1-t}$ and a dimension argument yields the equality. We explain the entire matter using the canonical automorphism

$$x = \sum_{g \in G} x_g X^g \mapsto \sum_{g \in G} x_g X^{-g} = \sum_{g \in G} x_{-g} X^g = \overline{x}$$

introduced in Section 4.2.

In order to eliminate visual confusion we will use the following notation for the usual inner product:

$$[x, y] = \left[\sum x_g X^g, \sum y_g X^g\right] = \sum x_g y_g.$$

Of course, for a code $C \subseteq \boldsymbol{R}$,

$$C^{\perp} = \{x \in \boldsymbol{R} \,|\, [x, c] = 0 \text{ for all } c \in C\}.$$

We have immediately the following "adjoint" relationship:

$$[\overline{x}, y] = [x, \overline{y}].$$

This shows that $\overline{S}^{\perp} = \overline{S^{\perp}}$, where we have set, for $S \subseteq \boldsymbol{R}$, $\overline{S} = \{\overline{s} \,|\, s \in S\}$. Now, since $xy = 0$ if and only if $\sum_{h \in G} x_{k-h} y_h = \sum_{h \in G} x_{k+h} y_{-h} = 0$ for all $k \in G$, $xy = 0$ if and only if $[X^k x, \overline{y}] = 0$ for all $k \in G$, and we get the following result — which is the analogue of Theorem 5.23 of Chapter 1 giving the orthogonals to cyclic codes and, moreover, admits a generalization to a more general case: Proposition 1.2 of Chapter (Ward).

Proposition 4.14 *For an ideal I of $\boldsymbol{R}$*

$$\mathrm{Ann}(I) = \overline{I^{\perp}} = \overline{I}^{\perp}.$$

The ideals we are concerned with are invariant under all the isometric automorphisms and, in particular, under the canonical automorphism. Hence we have

Corollary 4.15 *If an ideal I is invariant under the canonical automorphism, i.e. if $\overline{I} = I$, then $\mathrm{Ann}(I) = I^{\perp}$. In particular, in the group algebra $\boldsymbol{R}$ we have that*

$$\mathrm{Ann}(\boldsymbol{M}^t) = (\boldsymbol{M}^t)^{\perp} = \boldsymbol{M}^{m(p-1)+1-t}.$$

Example 4.16 Taking $t = m(p-1)$, we have that $\boldsymbol{M} = (F_p\mathbf{1})^{\perp}$ and, in particular, is of codimension 1 in $\boldsymbol{R}$, a fact that has emerged in various ways during our discussion of the group-algebra approach.

Remark: Observe that even if an ideal is not invariant under the canonical automorphism, the proposition shows that its annihilator and its orthogonal are equivalent codes.

4.6 The codes of the designs from $AG_m(F_p)$

We next prove that $\boldsymbol{M}^{r(p-1)}$ is the code, when p is prime, of the design of points and r-flats of the affine geometry $AG_m(F_p)$, generalizing what we already know when the prime is 2 (the Reed-Muller case). This more general result follows easily from a corollary of a theorem of Delsarte [19] which characterizes the subspaces left invariant by $AGL_m(F_q)$ acting naturally on the group ring $F_q[G]$ where G is the additive group of the field F_{q^m}. The corollary characterizes the codes in the group ring $\boldsymbol{R}$ invariant under $AGL_m(F_p)$ acting naturally on $\boldsymbol{R}$ as precisely the powers of $\boldsymbol{M}$. A recent elegant proof of this result by Weidner [54], using the Jennings basis, will be sketched now, but we will return to this matter in the next section.

Theorem 4.17 *Let $\boldsymbol{R}$ be the group ring over F_p of an elementary abelian p-group, G, of order p^m with $AGL(G) = AGL_m(F_p)$ — the automorphism group of the group G — acting naturally on $\boldsymbol{R}$. Let $\boldsymbol{M}$ be the radical of $\boldsymbol{R}$. Then the only subspaces of $\boldsymbol{R}$ invariant under $AGL_m(F_p)$ are the powers of $\boldsymbol{M}$. In group-theoretical terms $\boldsymbol{R}$, viewed as a module over $AGL_m(F_p)$, is uniserial[9] and $\boldsymbol{M}^t/\boldsymbol{M}^{t+1}$ is an irreducible $GL_m(F_p)$-module for $0 \leq t \leq m(p-1)$.*

Proof: We know, of course, that the powers of $\boldsymbol{M}$ are invariant under $AGL_m(F_p)$. Any subspace invariant under $AGL_m(F_p)$ is necessarily an ideal of $\boldsymbol{R}$ and, since $\boldsymbol{M}^{m(p-1)+1} = 0$, given any ideal I of $\boldsymbol{R}$ there is a smallest t with $\boldsymbol{M}^t \subseteq I$ and $\boldsymbol{M}^{t-1} \not\subseteq I$ — unless, of course, $I = \boldsymbol{R}$ in which case we have our assertion since, by convention, $\boldsymbol{R} = \boldsymbol{M}^0$. If $I \neq \boldsymbol{M}^t$, then there is an $x \in I$ which is not in $\boldsymbol{M}^t$ and because I is an ideal multiplying by a suitable element of $\boldsymbol{R}$ insures that $x \in \boldsymbol{M}^{t-1} \cap I$ but $x \notin \boldsymbol{M}^t$. But then $\boldsymbol{M}^{t-1} \cap I$ would be a proper $AGL_m(F_p)$-submodule of $\boldsymbol{M}^{t-1}$ strictly containing $\boldsymbol{M}^t$, an impossibility whenever $\boldsymbol{M}^{t-1}/\boldsymbol{M}^t$ is irreducible. Thus we are reduced to showing that $\boldsymbol{M}^t/\boldsymbol{M}^{t+1}$ is irreducible as an $AGL_m(F_p)$-module for all t. In fact we will show that it is an irreducible $GL_m(F_p)$-module for all t. (A slight change at the end of the argument would, in fact, show that it is an irreducible $SL_m(F_p)$-module but we do not need this generality for the purpose at issue.)

Of course, the action of $GL_m(F_p)$ on $\boldsymbol{R}$ is given once a basis of the elementary abelian group G is given and then that action is given by the

[9] A module is "uniserial" if it has only one composition series and "irreducible" if it has no proper submodules.

realization of the group via non-singular $m \times m$ matrices over F_p. We slightly change our notation letting the basis of the elementary abelian group be $g_1, g_2, \ldots, g_m$. Then, for an element $\sigma \in GL_m(F_p)$ represented by the matrix (a_{ij}), we have

$$\sigma g_i = \sum_{j=1}^{m} a_{ij} g_j$$

and hence, by definition,

$$\sigma X^{g_i} = X^{\sigma g_i} = X^{\sum_{j=1}^{m} a_{ij} g_j}.$$

Now set $x_i = X^{g_i} - 1$. A basis for $\boldsymbol{M}/\boldsymbol{M}^2$ is given by the images of $x_1, x_2, \ldots, x_m$ and a basis for the quotient $\boldsymbol{M}^t/\boldsymbol{M}^{t+1}$ is given by the images of elements of the form $\prod x_i^{k_i}$ where $0 \leq k_i \leq p-1$ and $\sum k_i = t$. These elements are part of the Jennings basis of $\boldsymbol{R}$ given by our choice of the basis of G. Moreover, the image of

$$\{\prod_{i=1}^{m} x_i^{k_i} | 0 \leq k_i \leq p-1, \sum k_i = t\}$$

in $\boldsymbol{M}^t/\boldsymbol{M}^{t+1}$ is a basis (over F_p) of that $GL_m(F_p)$-module. We will systematically throughout the proof work with these elements and ignore any elements of $\boldsymbol{M}^{t+1}$ that arise during calculations; this is tantamount to working in the quotient space. As an example of this caveat we note — since

$$X^{g+h} - 1 = (X^g - 1) + (X^h - 1) + (X^g - 1)(X^h - 1)$$

and *since we are working over a prime field* — that, modulo $\boldsymbol{M}^2$,

$$\sigma(x_i) = \sigma(X^{g_i} - 1) = X^{\sum a_{ij} g_j} - 1 \equiv \sum_{j=1}^{m} a_{ij} x_j$$

when σ is given by the matrix (a_{ij}). Of course, since the σ are algebra homomorphisms, we have that

$$\sigma(\prod_{i=1}^{m} x_i^{k_i}) = \prod_{i=1}^{m} (\sigma x_i)^{k_i}.$$

It is well-known — and easy to prove — that a p-Sylow subgroup of $GL_m(F_p)$ is given by the lower triangular matrices with 1's on the diagonal and that any non-zero S-module (when the field is F_p) contains a non-zero element left fixed by all elements of S.[10] Letting S be this p-Sylow subgroup we investigate the action of S on $\boldsymbol{M}^t/\boldsymbol{M}^{t+1}$, which we also denote by $\boldsymbol{V}_t$.

One first shows by induction on m that every non-zero S-submodule, W, of $\boldsymbol{V}_t$ contains the image of the element $w_t = x_1^{p-1} \cdots x_a^{p-1} x_{a+1}^b$ where $t = a(p-1)+b$ with $0 \leq b < p-1$. For $m=1$ this is obvious since, in this case, modulo $\boldsymbol{M}^{t+1}$, $\boldsymbol{M}^t$ is generated over F_p by x_1^t. Let $m > 1$ and let w be an element of $\boldsymbol{M}^t$ not in $\boldsymbol{M}^{t+1}$ that represents an element of the submodule W fixed by all elements of S; write

$$w = \sum_{i=0}^{p-1} w_i x_m^i$$

where $w_i \in \boldsymbol{M}^{t-i} \cap \langle x_1, \ldots, x_{m-1} \rangle$. Set $k = \max\{i | w_i \neq 0\}$. For $k=0$ the induction on m gives the result. Suppose $k \neq 0$. For $1 \leq i < m$ define $\sigma_i \in S$ by:

$$\sigma_i(x_j) = \begin{cases} x_j & \text{if } j \neq m\ ; \\ x_m + x_i & \text{if } j = m\ . \end{cases}$$

Then

$$\sigma_i w = \sum_{j=0}^{k} w_j (x_m + x_i)^j = \sum_{j=0}^{k} v_j x_m^j$$

for some $v_j \in \boldsymbol{M}^{t-j} \cap \langle x_1, \ldots, x_{m-1} \rangle$ with $v_{k-1} = w_{k-1} + k x_i w_k$. On the other hand $\sigma_i w = w$ and hence $v_{k-1} = w_{k-1}$ yielding $k x_i w_k = 0$. But k is a positive integer less that p and hence non-zero in F_p. So we have that $x_i w_k = 0$ for $1 \leq i < m$ and it follows that w_k is a scalar multiple of $\prod_1^{m-1} x_i^{p-1}$ which entails $a = m-1$ and $b = k$. But then, w_i, for $i < k$, is in $\boldsymbol{M}^{(m-1)(p-1)+k-i} \cap \langle x_1, \ldots, x_{m-1} \rangle = 0$ and w is a scalar multiple of the sought w_t. Thus, we have the assertion.

Next, setting $t' = m(p-1) - t$, consider the bilinear map

$$\phi : \boldsymbol{V}_t \times \boldsymbol{V}_{t'} \to \boldsymbol{V}_{m(p-1)} \approx F_p$$

given by $\phi(x,y) = xy$. It is invariant under $GL_m(F_p)$, i.e. $\phi(\sigma x, \sigma y) = \phi(x,y)$ for all σ. Moreover, the form is non-degenerate. Thus there is a $v_t \in \boldsymbol{V}_t$ with $\phi(v_t, w_{t'}) = 1$. It follows that

$$v_t = x_a^b \prod_{i=m+1-a}^{m} x_i^{p-1}.$$

[10] In other words, over F_p a p-group has only the trivial irreducible representation.

We claim that v_t generates $\boldsymbol{V}_t$ as an S-module: let W be the S-submodule generated by v_t and set

$$W^{\perp} = \{y \in \boldsymbol{V}_{t'} \mid \phi(w, y) = 0 \text{ for all } w \in W\}.$$

$W^{\perp}$ is also an S-submodule because of the invariance and, because $w_{t'}$ is not in $W^{\perp}$, $W^{\perp} = 0$ — which gives that $W = \boldsymbol{V}_t$.

Finally, consider the the element of $GL_m(F_p)$ that sends x_i to x_{m+1-i} for $1 \le i \le m$. It sends w_t to v_t which shows that any non-trivial $GL_m(F_p)$-submodule of $\boldsymbol{V}_t$ must, in fact, be $\boldsymbol{V}_t$. Thus these modules are irreducible. □

Note: Another proof — somewhat more robust since it has something to say about non-prime fields — of this result due to Mortimer [44] will be given in Section 5.5.

Now the code generated by the r-flats of the affine geometry $AG_m(F_p)$ is clearly invariant under the group $AGL_m(F_p)$ and is contained in $\boldsymbol{M}^{r(p-1)}$ but not in $\boldsymbol{M}^{r(p-1)+1}$. This yields

Theorem 4.18 *For any prime p, the code of the design of points and r-flats of the affine geometry $AG_m(F_p)$ is $\boldsymbol{M}^{r(p-1)}$.*

Since the dimension of $\boldsymbol{M}^{(m-1)(p-1)}$ is easy to compute we have a proof of the following important fact:

Corollary 4.19 *The dimension of the code over F_p of the design of points and $(m-1)$-flats of $AG_m(F_p)$ is*

$$\binom{m+p-1}{m}.$$

It is equally easy to compute the dimension of $\boldsymbol{M}^{p-1}$, which gives us the following

Corollary 4.20 *The dimension of the code over F_p of the design of points and lines of $AG_m(F_p)$ is*

$$p^m - \binom{m+p-2}{m}.$$

Proof: The proof consists of observing that, by complementation, one need only count the number of ways of choosing $p-2$ objects from m objects — repetitions allowed. □

Remark: In both cases the formulas are simple binomial coefficients since one does not have to worry about the constraint $p-1$ on the number of repetitions allowed.

Although we now know the dimensions and minimum weights of the codes coming from $AG_m(F_p)$ we have not yet determined the nature of the minimum-weight vectors nor have we discussed the codes coming from $PG_m(F_p)$. We postpone that discussion until Section 5.7 where all the relevant facts are established. See, for example, Theorem 5.42 and Theorem 5.44.

5 Generalized Reed-Muller codes

5.1 Introduction

Our description of generalized Reed-Muller codes is based, primarily, on the now classic paper of Delsarte, Goethals and MacWilliams [18]. We have, however, reworked the material in several important respects and have introduced a different notation. The definitions are based on the polynomial codes introduced by Kasami, Lin and Peterson [30, 31]; these authors introduced the primitive generalized Reed-Muller codes and Weldon [55] introduced the non-primitive generalized Reed-Muller codes and the single-variable approach using the Mattson-Solomon polynomial. Our treatment of that polynomial appears to be new in that we view it in a quotient ring that is slightly different from the one traditionally used.

Were it not for the complication introduced by moving from a prime field to F_q, where q is a proper prime power, much of the material in this section could be avoided. The reader interested *only* in the prime case will, however, need to read Section 5.7 and the previous material necessary for its understanding.

5.2 Definitions

First we describe the so-called *m-variable* approach. This is entirely analogous to our approach to the Reed-Muller codes (which are, simply, the generalized Reed-Muller codes for $q = 2$) and the generalization is straightforward (the functions involved being F_q-valued — rather than boolean — and having F_q-valued variables).

Let $q = p^t$, where p is a prime. Set $E = F_q$ and let V be a vector space of dimension m over E. Again we will denote a general vector in V by $\mathbf{v}$, and we will take V to be the space E^m of m-tuples, with standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$, where

$$\mathbf{e}_i = (\underbrace{0, 0, \ldots, 1}_{i}, 0, \ldots, 0).$$

Our codes will be q-ary codes, i.e. codes over E, and the ambient space will be the function space E^V, with the usual basis of characteristic functions of the vectors of V. As in Section 3, we can denote the members $f \in E^V$ by functions of the m-variables denoting the coordinates of a variable vector in V, i.e. if

$$\mathbf{x} = (x_1, x_2, \ldots, x_m) \in V,$$

then $f \in E^V$ is given by

$$f = f(x_1, x_2, \ldots, x_m)$$

and the x_i take values in E. Since every element in E satisfies $a^q = a$, the polynomial functions in the m variables can be reduced modulo $x_i^q - x_i$ (as was done in Section 3 for $q = 2$) and we can again form the set $\mathcal{M}$ of q^m monomial functions

$$\mathcal{M} = \{x_1^{i_1} x_2^{i_2} \ldots x_m^{i_m} \mid 0 \le i_k \le q-1, k = 1, 2, \ldots, m\}. \tag{1}$$

For a monomial in $\mathcal{M}$ the degree ρ is the total degree, i.e. $\rho = \sum_{k=1}^m i_k$ and we have that $0 \le \rho \le m(q-1)$. We will shortly show that $\mathcal{M}$ forms another basis (that we will not use for the codes) of E^V — as was done for $q = 2$; we do this in an entirely analogous way by expressing each characteristic function of a vector as a polynomial — in fact, a linear combination, with coefficients in F_q, of members of $\mathcal{M}$:

Lemma 5.1 *For* $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in V$,

$$v^{\mathbf{w}} = \prod_{i=1}^{m} \left(1 - (x_i - w_i)^{q-1}\right).$$

Proof: Since $a^{q-1} = 1$ for any non-zero $a \in E$, $1 - (x_i - w_i)^{q-1} = 0$ whenever $x_i \ne w_i$; thus the polynomial function on the right is clearly the same as the characteristic function on the left. $\square$

Example 5.2 The polynomial $1-(x_i-a)^{q-1}$ is the characteristic function of the $(m-1)$-flat in E^m given by the equation $X_i = a$. This polynomial is *not* linear unless $q = 2$.

Since E^V has dimension q^m, $\mathcal{M}$, being of cardinality q^m, is another basis for E^V — by the above lemma. The space E^V can then be viewed as the space of all polynomials (reduced modulo $x_i^q - x_i$) in the m variables; i.e. all linear combinations with coefficients in E of the monomials in $\mathcal{M}$. We use this interpretation to define the generalized Reed-Muller codes:

Definition 5.3 *Let $E = F_q$, where $q = p^t$ and p is a prime, and set $V = E^m$. Then for any ρ such that $0 \le \rho \le m(q-1)$, the* **$\boldsymbol{\rho}^{\text{th}}$ order generalized Reed-Muller code** *$\mathcal{R}_E(\rho, m)$ over E is the subspace of E^V (with basis the characteristic functions of the vectors in V) of all reduced m-variable polynomial functions of degree at most ρ. Thus*

$$\mathcal{R}_E(\rho,m) = \mathcal{R}_{F_q}(\rho,m) = \left\langle x_1^{i_1} x_2^{i_2} \ldots x_m^{i_m} \in \mathcal{M} \mid \sum_{k=1}^{m} i_k \le \rho \right\rangle .$$

Example 5.4 (1) For $q = 2$ and $0 \le \rho \le m$, $\mathcal{R}_{F_2}(\rho, m) = \mathcal{R}(\rho, m)$.

(2) For any q, $\mathcal{R}_{F_q}(0,m) = F_q\boldsymbol{\jmath} = \langle \boldsymbol{\jmath} \rangle$ and $\mathcal{R}(m(q-1),m) = F_q^{q^m}$, the entire ambient space.

The dimension of a generalized Reed-Muller code can be obtained by simply counting the number of elements in its obvious monomial basis:

Theorem 5.5 *For any ρ such that $0 \le \rho \le m(q-1)$,*

$$\dim(\mathcal{R}_{F_q}(\rho,m)) = \sum_{i=0}^{\rho} \sum_{k=0}^{m} (-1)^k \binom{m}{k} \binom{i-kq+m-1}{i-kq}.$$

Proof: We use the fact that the number of ways of picking j objects from a set of m objects — with repetitions allowed — is $\binom{j+m-1}{m-1} = \binom{j+m-1}{j}$. An inclusion-exclusion argument shows that the inner sum is the number of ways of picking i objects from a set of m objects, when no object can be chosen more than $q-1$ times. Summing on i yields the result. □

Example 5.6 (1) By a direct count we have $\dim(\mathcal{R}_{F_q}(1,m)) = 1+m$, just as in the binary case — but note that the function x_i is zero on the hyperplane $X_i = 0$ but takes on many different values off the hyperplane unless $q = 2$. Since every non-constant linear polynomial in the m variables has q^{m-1} zeros, these codes have minimum weight $q^m - q^{m-1} = q^{m-1}(q-1)$ and all code vectors except the zero vector and non-zero multiples of $\mathbf{1}$ have this weight.

(2) By a direct count we have, for $0 \le \rho \le q-1$, $\dim(\mathcal{R}_{F_q}(\rho,1)) = 1+\rho$, and, since a polynomial in one variable of degree at most ρ can have at most ρ distinct roots, the minimum weight in this code is $q - \rho$, there being a polynomial of degree ρ with exactly ρ distinct roots since $\rho < q$. These codes are of genus zero in the sense of Tsfasman[11] and Vlăduţ, [53].

(3) More generally, if $\rho \le q-1$ no choice of $i_1, i_2, \ldots, i_m$ with $\sum i_k = \rho$ will ever have an $i_k > q-1$ and inclusion-exclusion is unnecessary in the proof. Moreover, by introducing a "dummy" object the number of ways of choosing *at most* ρ things — repetitions allowed — from a set of m objects is easily seen to be $\binom{\rho+m}{m}$ and hence one sees directly that

$$\dim(\mathcal{R}_{F_q}(\rho,m)) = \binom{\rho+m}{m}, \text{ for } 0 \le \rho \le q-1.$$

Just as in the binary case, the codes orthogonal to generalized Reed-Muller codes are again generalized Reed-Muller codes. Again the proof is entirely analogous to the binary case, but we need a lemma which generalizes the result that $\mathcal{R}(m-1,m)$ consists of even-weight vectors. The required result is an easy consequence of the orthogonality relations; thus the proof depends simply on the fact that $\sum_{i=0}^{n-1} \omega^i = 0$ whenever ω is an n^{th} root of unity different from 1.

Lemma 5.7 *If $f \in E^V$ has degree $\rho < m(q-1)$ as a polynomial in $x_1, x_2, \ldots, x_m$ then*

$$\sum_{\mathbf{w} \in V} f(\mathbf{w}) = 0.$$

In fact, if f is any linear combination of the elements of $\mathcal{M}$ with coefficients in any overfield of E, the same result holds.

[11] These codes are usually referred to as MDS codes, or sometimes optimal codes, in the coding literature; we have chosen to adopt the new terminology introduced in [53].

Proof: The result is clearly true for any constant function (since the block length of the code is a multiple of, in fact a power of, p, the characteristic of F_q) so we need to prove the assertion only for monomial functions, i.e. elements of $\mathcal{M}$, of positive degree less than $m(q-1)$. Moreover, if in such a monomial any $i_k = 0$, the sum is again a multiple of q and hence 0. We thus restrict ourselves to those monomials in which every x_i appears. The orthogonality relations for the group $E^\times \times \ldots \times E^\times$, using E itself as the field where the characters take their values, yields immediately, taking η as the principal character and χ the character sending $\mathbf{a} = (a_1, \ldots, a_m)$ to $a_1^{i_1} \ldots a_m^{i_m}$, that

$$\sum_{\mathbf{a}} a_1^{i_1} \ldots a_m^{i_m} = 0$$

since there is some k for which $i_k < q-1$, and since the sum only need be taken over those vectors all of whose entries are non-zero. □

Theorem 5.8 *For $\rho < m(q-1)$*

$$\mathcal{R}_{F_q}(\rho, m)^\perp = \mathcal{R}_{F_q}(m(q-1) - 1 - \rho, m).$$

Proof: If f has degree at most ρ and g has degree at most $m(q-1)-1-\rho$, then the product fg has degree less than $m(q-1)$. Thus Lemma 5.7 implies that

$$\sum_{\mathbf{w} \in V} f(\mathbf{w}) g(\mathbf{w}) = 0$$

and the corresponding codewords are orthogonal. Hence

$$\mathcal{R}_{F_q}(\rho, m)^\perp \supseteq \mathcal{R}_{F_q}(m(q-1) - 1 - \rho, m)$$

and now we need only check the dimensions: the involution of $\mathcal{M}$ that sends $x_1^{i_1} \ldots x_m^{i_m}$ to $x_1^{q-1-i_1} \ldots x_m^{q-1-i_m}$ yields the fact that the number of monomials of degree greater than $m(q-1)-1-\rho$ is equal to the number of degree less than or equal to ρ and hence $\dim(\mathcal{R}_{F_q}(m(q-1)-1-\rho, m)) = q^m - \dim(\mathcal{R}_{F_q}(\rho, m))$, as required. □

The generalized Reed-Muller codes are codes over possibly non-prime fields and thus could not be the codes of designs coming from affine geometries — unless q happens to be a prime. They do contain the incidence vectors of flats in the geometry, as we will shortly see. In order to demonstrate this it is convenient, notationally, to make an observation — important in itself — about the automorphism group of $\mathcal{R}_{F_q}(\rho, m)$:

Theorem 5.9 *For $0 \le \rho \le m(q-1)$, the automorphism group of $\mathcal{R}_{F_q}(\rho, m)$ contains the affine group $AGL_m(F_q)$ in its natural action on $V = F_q^m$.*

Proof: Recall that for any code $C \subseteq F^{\mathcal{P}}$, an automorphism of C is a permutation σ of $\mathcal{P}$ that preserves C, i.e. for which, if $c \in C$, $c^\sigma \in C$, where c^σ is defined by $c^\sigma(P) = c(\sigma(P))$ for $P \in \mathcal{P}$.

Now $\gamma \in AGL_m(F_q)$ is given by

$$\gamma : \mathbf{v} \mapsto A\mathbf{v} + \mathbf{a},$$

where $\mathbf{v}, \mathbf{a} \in V = E^m$ — viewed as column vectors — and A is a non-singular $m \times m$ matrix over E. Thus, for $f \in \mathcal{R}_{F_q}(\rho, m)$, f^γ is defined by

$$f^\gamma(\mathbf{x}) = f(A\mathbf{x} + \mathbf{a}),$$

and so, clearly, $f^\gamma \in \mathcal{R}_{F_q}(\rho, m)$. $\square$

Note: Berger and Charpin [5] have shown that $AGL_m(F_q)$ is the *full* group of permutation automorphisms of these generalized Reed-Muller codes — when, of course, $0 < \rho < m(q-1)$. See Chapter (Huffman) for the details.

Theorem 5.10 *For $0 \le r \le m$ and $\rho \ge r(q-1)$, the generalized Reed-Muller code $\mathcal{R}_{F_q}(\rho, m)$ contains the incidence vector of any $(m-r)$-flat of $AG_m(F_q)$.*

Proof: Any $(m-r)$-flat in $AG_m(F_q)$ consists of all the points $\mathbf{x}$ of V satisfying r independent equations

$$\sum_{j=1}^{m} a_{ij} X_j = w_i, \text{ for } i = 1, 2, \ldots, r$$

where all a_{ij} and w_i are in F_q. If the code $\mathcal{R}_{F_q}(\rho, m)$ contains the incidence vector of some t-flat, then it will contain the incidence vector of every t-flat, since the affine group $AGL_m(F_q)$ acts transitively on t-flats and, as we have just seen, preserves the code. So we need only construct one $(m-r)$-flat that is in $\mathcal{R}_{F_q}(\rho, m)$.

Consider the polynomial

$$p(x_1, \ldots, x_m) = \prod_{i=1}^{r} \left(1 - x_i^{q-1}\right),$$

of degree $r(q-1)$. Then $p(x_1,\ldots,x_m)=0$ on V unless $x_i=0$ for $i=1,2,\ldots,r$. Thus the codeword corresponding to $p(\mathbf{x})$ has the entry 1 at points on the $(m-r)$-flat defined by the r equations

$$X_1=0, X_2=0,\ldots,X_r=0$$

and the entry 0 at points off the flat. Hence it is the incidence vector of this $(m-r)$-flat. Since $p(x_1,\ldots,x_m)\in\mathcal{R}_{F_q}(\rho,m)$ for $\rho\geq r(q-1)$, we have the result. □

As in the binary case the proof shows more, namely that the generalized Reed-Muller code contains all $(m-s)$-flats for $0\leq s\leq r$. Moreover, the subcode generated by the $(m-r)$-flats contains, by the same induction argument used in the binary case, all $(m-s)$-flats for $0\leq s\leq r$. Note, however, that when using characteristic functions of t-flats to obtain characteristic functions of $(t+1)$-flats one could use coefficients other than 1 provided $q>2$ and hence obtain vectors that are supported on the $(t+1)$-flat but are not characteristic functions.

Example 5.11 Take $q=3$ and $m=2$. The geometry is then $AG_2(F_3)$, the affine plane of order 3. Let $C=C_3(AG_2(F_3))$ be the code over F_3 associated with this plane, i.e. the code generated by the incidence matrix of the plane. The incidence vectors of the lines (1-flats) will be in $\mathcal{R}_{F_3}(\rho,2)$ for $\rho=2,3$ and 4. In fact $C=\mathcal{R}_{F_3}(2,2)$, while, as we know, $\mathcal{R}_{F_3}(3,2)=(F_3\boldsymbol{\jmath})^{\perp}$ and $\mathcal{R}_{F_3}(4,2)=F_3^9$, the entire ambient space.

This example is indicative of what happens in the case of planes over prime fields. A rather easy argument using elementary divisors (see [2, Chapter 6]) shows that the dimension of the code of any affine plane of prime order p is $\binom{p+1}{2}$ and since the computation of the dimension of $\mathcal{R}_{F_p}(p-1,2)$ is also easy (see Example 5.6 above) and also yields $\binom{p+1}{2}$ we have a proof, which avoids the use of Delsarte's theorem, of the following

Proposition 5.12 *The code over F_p of the desarguesian affine plane of prime order p is the generalized Reed-Muller code $\mathcal{R}_{F_p}(p-1,2)$.*

Since it is easy to see ([2, Corollary 6.4.1]) that the code over F_p of any affine plane of order p has as minimum-weight vectors only the scalar multiples of the characteristic functions of lines of the plane, the above proposition yields an elementary proof of the following

Corollary 5.13 *The generalized Reed-Muller code* $\mathcal{R}_{F_p}(p-1,2)$ *is a*

$$[p^2, \frac{p(p+1)}{2}, p]$$

code over F_p *all of whose minimum-weight vectors are scalar multiples of characteristic functions of 1-flats of* F_p^2.

In terms of the modular algebra $\boldsymbol{R} = F_p[G]$, where G is the elementary abelian p-group of order p^2, the result above is expressed as

$$\boldsymbol{M}^{p-1} = \mathcal{R}_{F_p}(p-1,2) = C_p(AG_2(F_p)).$$

These equalities should, strictly speaking, be isomorphisms but if we take the point of view that the set on which the various functions involved are defined is a fixed copy of F_p^2, we actually can assert equality. We shall see later on (Theorem 5.19) that, more generally, in the prime case $\boldsymbol{M}^{m(p-1)-\rho} = \mathcal{R}_{F_p}(\rho, m)$, which is a generalization of Berman's theorem.

Just as for the Reed-Muller codes, we can remove a coordinate position to obtain a code of length $q^m - 1$, which turns out to be cyclic:

Definition 5.14 *The* $\boldsymbol{\rho}^{\text{th}}$ **order punctured generalized Reed-Muller code**, *where* $0 \le \rho < m(q-1)$, *denoted by* $\mathcal{R}_{F_q}(\rho, m)^*$, *is the code of length* $q^m - 1$ *obtained by deleting the coordinate position* $\mathbf{0}$ *from* $\mathcal{R}_{F_q}(\rho, m)$.

For $q = 2$, $\mathcal{R}_{F_2}(\rho, m)^* = \mathcal{R}(\rho, m)^*$, the punctured Reed-Muller code. These are also called *shortened* generalized Reed-Muller codes (see van Lint [37]) or *cyclic* generalized Reed-Muller codes (see Blake and Mullin [8]) since our next result will show they are cyclic (which we already know for $q = 2$). Observe also that any coordinate position can be deleted in place of $\mathbf{0}$, since $AGL_m(F_q)$ acts transitively on the vectors of $V = E^m$.

Theorem 5.15 *For any* ρ *such that* $0 \le \rho < m(q-1)$, *the automorphism group of* $\mathcal{R}_{F_q}(\rho, m)^*$ *contains the general linear group* $GL_m(F_q)$. *In particular,* $\mathcal{R}_{F_q}(\rho, m)^*$ *is a cyclic code.*

Proof: The group $GL_m(F_q)$ is the stabilizer of $\mathbf{0}$ in $AGL_m(F_q)$, so it obviously acts on $\mathcal{R}_{F_q}(\rho, m)^*$. We can obtain a cyclic group of order $q^m - 1$ acting on it in the usual way: consider the field $K = F_{q^m}$ and let ω be a primitive element in K; now K, as a vector space over $F_q = E$, is isomorphic to V and multiplication by ω simply cycles the elements of $K^\times = V - \{\mathbf{0}\}$. (This map also yields a Singer cycle on the projective points — as discussed earlier in Section 2.) $\square$

Corollary 5.16 *Provided that $\rho < m(q-1)$, the generalized Reed-Muller codes $\mathcal{R}_{F_q}(\rho, m)$ are extended cyclic codes and of the same dimension as the corresponding cyclic codes $\mathcal{R}_{F_q}(\rho, m)^*$.*

Proof: By Lemma 5.7, $f(\mathbf{0}) = -\sum_{\mathbf{w}\neq\mathbf{0}} f(\mathbf{w})$ provided that the degree of f is less than $m(q-1)$. □

5.3 The single-variable approach

We introduce now the single-variable approach to the generalized Reed-Muller codes utilizing the Mattson-Solomon polynomial [42].

Taking ω to be a primitive element of $K = F_{q^m}$ and, using the same notation as above, set $v = q^m - 1$ and consider the vector space of polynomials in Z with coefficients in K and of degree less than v. Then Lagrange interpolation (see, for example, [48]) shows that any function from $K^\times$ to K is given uniquely by such a polynomial, viewed as a polynomial function in the single variable Z. In terms of the characteristic functions of the points ω^i of $V^* = V - \{\mathbf{0}\} \approx K^\times$, where $V = F_q^m \approx K$, such a polynomial function can be written as

$$P(Z) = \sum_{i=0}^{v-1} P(\omega^i) g_i(Z), \tag{2}$$

where $g_i(Z)$ denotes the characteristic function of $\{\omega^i\}$, i.e.

$$g_i(Z) = -\omega^i \frac{(Z^v - 1)}{(Z - \omega^i)}. \tag{3}$$

Clearly the polynomials Z^i, for $i = 0, 1, \ldots, v-1$, form an alternative basis, and the correspondence is given as follows: if

$$P(Z) = \sum_{j=0}^{v-1} c_j Z^j, \text{ where } c_j \in K,$$

then, using the discrete Fourier transform and noting that $1/v$ is -1 when viewed in K,

$$c_j = -\sum_{i=0}^{v-1} P(\omega^i)\omega^{-ji} = -\phi_{v-j}(P), \tag{4}$$

where ϕ_s is the function defined in Section 4.3. Then

$$P(Z) = \sum_{j=0}^{v-1} c_j Z^j = -\sum_{j=0}^{v-1} \phi_{v-j}(P) Z^j,$$

is the **Mattson-Solomon polynomial** of the function P.

We now restrict to those functions taking values in $E = F_q \subseteq K$, i.e. we require that $P(\omega^i) \in E$ for all i. This is equivalent, from Equation (4), to requiring that $(c_j)^q = c_{qj}$, where the subscripts are taken modulo $v = q^m - 1$. These functions form an E-subspace of K^{V^*}. Denoting this subspace by L and writing its vectors in terms of the basis of characteristic functions, we have

$$L = \left\{ (P(1), \dots, P(\omega^{v-1})) \mid P(Z) = \sum_{j=0}^{v-1} c_j Z^j, c_j \in K, (c_j)^q = c_{qj} \right\}.$$

The vector space L over the field E corresponds with the vector-space structure of the polynomial ring $E[Y]/(Y^v - 1)$ via

$$(P(1), \dots, P(\omega^{v-1})) \mapsto \sum_{i=0}^{v-1} P(\omega^i) Y^i.$$

In fact a polynomial $f(Y) = \sum_{i=0}^{v-1} a_i Y^i$ corresponds to the function $P(Z)$ defined by $P(Z) = \sum_{j=0}^{v-1} c_j Z^j$, where $c_j = -f(\omega^{-j})$. If the polynomial $g(Y)$ divides $Y^v - 1$, then the cyclic code generated by $g(Y)$ contains $f(Y)$ if and only if $f(\omega^{-j}) = 0$ for all zeros ω^{-j} of $g(Y)$. The corresponding $P(Z) \in L$ has the property that $c_j = 0$ if ω^{-j} is a root of $g(Y)$. Thus the cyclic code with zeros $\{\omega^{-j} \mid j \in T\}$, where $T \subseteq \{0, 1, \dots, v-1\}$, can be characterized as

$$\left\{ (P(1), \dots, P(\omega^{v-1})) \mid P(Z) = \sum_{j=0}^{v-1} c_j Z^j \in L, c_j = 0 \text{ if } j \in T \right\}.$$

Note that if a positive integer u has an orbit of length i under the map $j \mapsto jq$ modulo v, i.e. if $uq^i \equiv u \pmod{v}$ and i is the smallest integer satisfying the congruence, then the choice of the coefficient of Z^u must be in a field of degree i over E; this agrees, of course, with the dimensional requirements.

For the extended codes we adjoin the extra coordinate position corresponding to $0 \in K$, where the entry is $-\sum_{i=0}^{v-1} P(\omega^i)$. Since $P(0) = c_0$, Equation (4) implies that all extended cyclic codes are contained in

$$M = \{(P(0), P(1), P(\omega), \dots, P(\omega^{v-1})) \mid P(Z) \in L\}, \tag{5}$$

the subspace of E^{q^m} consisting of those vectors the sum of whose coordinates is zero. (We have deliberately called this subspace M to suggest to the reader the correspondence with $\boldsymbol{M}$.) It follows from the above that every polynomial $P(Z)$ in $K[Z]$ of degree less than v has the property that $\sum_{z\in K} P(z) = 0$; thus a polynomial $r(Z) = \sum_{j=0}^{v} c_j Z^j$ satisfies $\sum_{z\in K} r(z) = 0$ if *and only if* it is of degree less than v — i.e. has $c_v = 0$ — since such polynomials form a subspace of codimension 1 (as a vector space over K).

We next establish the correspondence between reduced polynomials in the variables $x_1, \ldots, x_m$ of degree less than $m(q-1)$ and polynomials $P(Z) \in L$. In fact, we do more and establish an algebra isomorphism between $E[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$ and the E-subalgebra of $K[Z]/(Z^{q^m} - Z)$ given by the set of fixed points of the Frobenius map $x \mapsto x^q$. The difference between what follows and the development given in [18] is that here the Mattson-Solomon polynomials live in $K[Z]/(Z^{q^m} - Z)$ rather than $K[Z]/(Z^{q^m-1} - 1)$.

In order to explain the correspondence we first introduce the trace: let $\mathrm{Tr}_{K/E}$ denote the trace from K to E, i.e. for $z \in K$,

$$\mathrm{Tr}_{K/E}(z) = z + z^q + z^{q^2} + \cdots + z^{q^{m-1}}.$$

Since the trace is a linear transformation from the vector space K over E onto E, given any basis $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ for K over E there is a unique complementary basis $\{\beta_1, \beta_2, \ldots, \beta_m\}$ for K over E such that

$$\mathrm{Tr}_{K/E}(\alpha_i \beta_j) = \delta_{ij},$$

where δ_{ij} denotes the Kronecker delta function. (See, for example, Lidl and Niederreiter [35, Theorem 2.24].)

Using the basis $\{1, \omega, \omega^2, \ldots, \omega^{m-1}\}$, where ω is a primitive element for K, let the complementary basis be $\{\beta_1, \beta_2, \ldots, \beta_m\}$. Then $z \in K$ satisfies $z = \sum_{i=1}^{m} a_i \omega^{i-1}$, where the a_i are in E, if and only if $a_i = \mathrm{Tr}_{K/E}(\beta_i z)$.

Since $E \subseteq K$, we can define a ring homomorphism θ by

$$\theta : \begin{cases} E[x_1, \ldots, x_m] & \to \quad K[Z] \\ \qquad\qquad x_i & \mapsto \quad \beta_i Z + (\beta_i Z)^q + \ldots + (\beta_i Z)^{q^{m-1}} = \mathrm{Tr}_{K/E}(\beta_i Z), \end{cases}$$

where we are utilizing the Frobenius map of $K[Z]$ into itself and slightly abusing the trace notation. Following θ by the natural map

$$K[Z] \to K[Z]/(Z^{q^m} - Z),$$

using the standard representatives — namely polynomials in Z of degree less than or equal to v — and viewing Z as $Z + (Z^{q^m} - Z)$, we see that

$(\mathrm{Tr}_{K/E}(\beta_i Z))^q = \mathrm{Tr}_{K/E}(\beta_i Z)$; hence we get the induced ring homomorphism,

$$\bar{\theta} : E[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m) \to K[Z]/(Z^{q^m} - Z).$$

We can thus convert any "reduced" polynomial in the variables x_i with coefficients in E into a polynomial in Z, of degree less than or equal to $v = q^m - 1$, with coefficients in K. It follows from Lemma 5.7 and the fact that the vector

$$(\mathrm{Tr}_{K/E}(\beta_1 z), \ldots, \mathrm{Tr}_{K/E}(\beta_m z))$$

takes on every value in E^m as z varies over K, that the image of $p(x_1, \ldots, x_m)$ is of degree *less than* v provided $p(x_1, \ldots, x_m)$ is of degree less than $m(q-1)$. Moreover, the polynomial $P(Z) = \sum_{j=0}^{v-1} c_j Z^j$ has $c_j^q = c_{jq}$ (with subscripts computed modulo v) if and only if $P(Z)^q = P(Z)$ in the ring $K[Z]/(Z^{q^m} - Z)$ — since $qj \neq v$ for $j < v$ and therefore computing subscripts modulo v is the same as viewing the polynomial in $K[Z]/(Z^{q^m} - Z)$. Since $(p(x_1, \ldots, x_m))^q = p(x_1, \ldots, x_m)$ in the ring $E[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$, the reduced polynomials of degree less than $m(q-1)$ have images, under $\bar{\theta}$, in L.

Conversely, we define a ring homomorphism,

$$K[Z] \to K[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m),$$

by

$$Z \mapsto \sum_{i=1}^{m} x_i \omega^{i-1}.$$

Since $(\sum_{i=1}^m x_i \omega^{i-1})^{q^m} = \sum_{i=1}^m x_i \omega^{i-1}$, we obtain a ring homomorphism

$$K[Z]/(Z^{q^m} - Z) \to K[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m).$$

If $P(Z)^q = P(Z)$, then the image of $P(Z)$ must lie in $E[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$, since this is the subring of $K[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$ left fixed by the Frobenius map, $x \mapsto x^q$. Let R denote the subring of $K[Z]/(Z^{q^m} - Z)$ left pointwise fixed by the Frobenius map; then we have a ring homomorphism

$$\psi : R \to E[x_1, \ldots x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$$

and using the fact that $\mathrm{Tr}_{K/E}(x_i \beta) = x_i \mathrm{Tr}_{K/E}(\beta)$ it follows easily that $\psi \circ \bar{\theta}$ is the identity map. Moreover, since in $K[Z]/(Z^{q^m} - Z)$ we have $(Z^v)^k = Z^v$

for any positive integer k, $P(Z)^q = P(Z) = \sum_{j=0}^{v} c_j Z^j$ if and only if $c_v \in E$ and $\sum_{j=0}^{v-1} c_j Z^j \in L$. Both rings have dimension q^m as E-algebras and hence $\bar{\theta}$ is an isomorphism of rings, with ψ the inverse of $\bar{\theta}$. In addition, under this ring isomorphism

$$M \approx \mathcal{R}_{F_q}(m(q-1)-1, m).$$

In fact, much more is true: a simple calculation shows that Z^j corresponds to a polynomial in the x_i's of degree[12] $\mathrm{wt}_q(j)$; thus, if $P(Z) \in L$ is such that $c_j = 0$ for $\mathrm{wt}_q(j) > \rho$, its image in $E[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$ has degree less than or equal to ρ. Hence the isomorphism above carries, for $0 \leq \rho < m(q-1)$, the extended cyclic code with defining set $\widehat{T} = \{j \mid \mathrm{wt}_q(j) < m(q-1)-\rho\}$ onto the generalized Reed-Muller code $\mathcal{R}_{F_q}(\rho, m)$. Hence we have proved the following

Theorem 5.17 *Let R be the subring of $F_{q^m}[Z]/(Z^{q^m} - Z)$ left pointwise fixed by the Frobenius homomorphism, $x \mapsto x^q$. Then R is isomorphic (as an algebra over the field F_q) to $F_q[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$. Moreover, the isomorphism can be chosen in such a way that the extended cyclic code over F_q with defining set $\{j \mid \mathrm{wt}_q(j) < m(q-1) - \rho\}$ is carried onto the generalized Reed-Muller code $\mathcal{R}_{F_q}(\rho, m)$.*

5.4 Roots, dimensions and minimum weights

In this section we draw out the consequences of what we have just proved and discuss the minimum weights of the relevant codes.

Theorem 5.18 *If $K = F_{q^m}$, $E = F_q$, and ω is a primitive element of K, then, for $0 \leq u \leq q^m - 2$, ω^u is a root of the generator polynomial of the code $\mathcal{R}_{F_q}(\rho, m)^*$ if and only if $0 < \mathrm{wt}_q(u) \leq m(q-1) - 1 - \rho$.*

This is a consequence of Theorem 5.17 and the discussion at the begining of Section 4.3, where α^i is a root if and only if i is in the defining set T.

We also obtain the promised generalization (due to Charpin, [14, 15]) of Berman's theorem:

Theorem 5.19 *For any prime p, and any ρ such that $0 \leq \rho < m(p-1)$, if $\boldsymbol{M}$ is the radical of $F_p[G]$, the group algebra over F_p of the elementary abelian group G of order p^m, then the code given by $\boldsymbol{M}^{m(p-1)-\rho}$ is the generalized Reed-Muller code $\mathcal{R}_{F_p}(\rho, m)$.*

[12] The reduction modulo $x_i^q - x_i$ can only reduce the degree of a given monomial and, for future reference, we note that the reduction is by a multiple of $q - 1$.

This is a consequence of Theorem 5.17 and Corollary 4.11.

We draw out the consequences of Theorem 5.18 below.

Corollary 5.20 *For $0 \leq \rho < m(q-1)$ the code $\mathcal{R}_{F_q}(\rho, m)^*$ is the cyclic code with generator polynomial*

$$g(Y) = \prod_{\substack{0<u<q^m-1 \\ \mathrm{wt}_q(u) \leq m(q-1)-1-\rho}} (Y - \omega^u),$$

where ω is a primitive element of F_{q^m}.

Corollary 5.21 *For $0 \leq \rho < m(q-1)$ the code $(\mathcal{R}_{F_q}(\rho, m)^*)^\perp$ is the cyclic code with generator polynomial*

$$g(Y) = \prod_{\substack{0 \leq u<q^m-1 \\ \mathrm{wt}_q(u) \leq \rho}} (Y - \omega^u),$$

where ω is a primitive element of F_{q^m}. Moreover,

$$(\mathcal{R}_{F_q}(\rho, m)^*)^\perp = (F_q\boldsymbol{\jmath})^\perp \cap \mathcal{R}_{F_q}(m(q-1)-1-\rho, m)^*.$$

Proof: That the generating polynomial is as asserted follows from results in Chapter 1 on cyclic codes. The second statement then follows from Corollary 5.20, with the extra factor $(Y-1)$ placing the code inside $(F_q\boldsymbol{\jmath})^\perp$. □

Corollary 5.22 *For $0 \leq \rho < m(q-1)$, the dimensions of both $\mathcal{R}_{F_q}(\rho, m)^*$ and $\mathcal{R}_{F_q}(\rho, m)$ are given by*

$$|\{u | 0 \leq u \leq q^m - 1 \textit{ and } \mathrm{wt}_q(u) \leq \rho\}|.$$

Proof: This is simply a restatement of the value of the dimension in terms of the q-weight. □

Example 5.23 For $m = 2$ and $q = 3$, $E = F_3$ and $K = F_9$. The quadratic $f(Z) = Z^2 + Z - 1$ is a primitive polynomial for K, with primitive root ω. Since $m(q-1) = 4$, $\mathcal{R}_E(1, 2)^*$ and $\mathcal{R}_E(2, 2)^*$ are the only interesting cases. The generator polynomial for $\mathcal{R}_E(2,2)^*$ is $g(Y) = (Y-\omega)(Y-\omega^3) = Y^2+Y-1$ and the code has dimension 6. A generator matrix is

$$G = \begin{pmatrix} -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 \end{pmatrix}.$$

The extended code, $\mathcal{R}_E(2,2)$, has a generator matrix that is G augmented by an extra column whose entries are -1's: this is a generator matrix for the code of the affine plane $AG_2(F_3)$. If the extra column, corresponding to $\mathbf{0}$, is labelled 0, and added as the first column, and the columns of G then labelled 1 to 8, then the plane can be pictured as in Figure 1, with incidence matrix as given in Figure 2, where the rows are arranged in parallel classes. The columns then correspond to the points

$$(0,0),(1,0),(0,1),(1,-1),(-1,-1),(-1,0),(0,-1),(-1,1),(1,1);$$

equivalently, they correspond to the elements of F_9 in the order

$$0,1,\omega,\omega^2,\omega^3,\omega^4,\omega^5,\omega^6,\omega^7.$$

The matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$$

cycles the last eight of these points and correspondes to multiplication by ω. The line $\{0,1,5\}$, for example, has the equation $X_2 = 0$ and the line $\{6,5,8\}$ has the equation $X_1+X_2+1=0$.

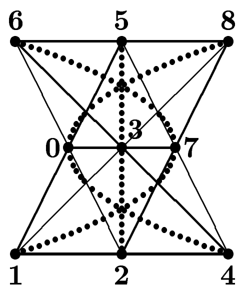

Figure 1: The affine plane $AG_2(F_3)$

$$
\begin{array}{c}
\begin{array}{ccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\
\left[\begin{array}{ccccccccc}
 & 1 & 1 & & 1 & & & & \\
1 & & & 1 & & & & 1 & \\
 & & & & & 1 & 1 & & 1 \\ \hline
 & 1 & & & & & 1 & 1 & \\
1 & & & & 1 & & & & 1 \\
 & & 1 & 1 & & 1 & & & \\ \hline
 & 1 & & 1 & & & & & 1 \\
 & & & & 1 & 1 & & 1 & \\
1 & & 1 & & & & 1 & & \\ \hline
1 & 1 & & & & 1 & & & \\
 & & 1 & & & & & 1 & 1 \\
 & & & 1 & 1 & & 1 & &
\end{array}\right]
\end{array}
$$

Figure 2: Incidence matrix for $AG_2(F_3)$

The generator polynomial for $\mathcal{R}_E(1,2)^*$ is

$$f(Y) = (Y-\omega)(Y-\omega^2)(Y-\omega^3)(Y-\omega^4)(Y-\omega^6)$$

and that for $(\mathcal{R}_E(2,2)^*)^\perp$ is

$$(Y-1)f(Y) = Y^6 + Y^5 - Y^4 - Y^2 - Y + 1,$$

so a check[13] matrix for $\mathcal{R}_E(2,2)^*$ is

$$H = \begin{pmatrix} 1 & -1 & -1 & 0 & -1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & -1 & 1 & 1 \end{pmatrix}.$$

Theorem 5.24 *If $\rho = r(q-1)+s$, where $0 \le s < q-1$, then the code $\mathcal{R}_{F_q}(\rho,m)^*$ is a subcode of a BCH code of length q^m-1 over F_q with designed distance*

$$(q-s)q^{m-r-1} - 1.$$

Proof: From Theorem 5.18, for a primitive element ω of F_{q^m}, ω^u is a root of the generator polynomial of $\mathcal{R}_{F_q}(\rho,m)^*$ if and only if $0 < \mathrm{wt}_q(u) < m(q-1)-\rho$. Now

$$m(q-1)-\rho = (m-r)(q-1)-s = (m-r-1)(q-1)+(q-1-s),$$

[13]This matrix is frequently called a *parity-check matrix* in the literature.

so if we let h be the smallest integer with $\mathrm{wt}_q(h) = m(q-1) - \rho$, then

$$h = (q-s-1)q^{m-r-1} + \sum_{i=0}^{m-r-2} (q-1)q^i = (q-s)q^{m-r-1} - 1.$$

It follows that every integer u with $0 \leq u < h$ satisfies $\mathrm{wt}_q(u) < m(q-1) - \rho$, and thus the elements $\omega^1, \omega^2, \ldots, \omega^{h-1}$ are all roots of the generator polynomial of the code. Thus $\mathcal{R}_{F_q}(\rho, m)^*$ is a subcode of a BCH code of designed distance $(q-s)q^{m-r-1} - 1$ as stated. □

The designed distance is the true minimum distance, as the following theorem shows by an explicit construction of codewords of this weight.

Theorem 5.25 *For any ρ such that $0 \leq \rho < m(q-1)$, where $\rho = r(q-1)+s$ with $0 \leq s < q-1$, $\mathcal{R}_{F_q}(\rho, m)$ has vectors of weight $(q-s)q^{m-r-1}$ that consist of the sum of multiples of the incidence vectors of $(q-s)$ parallel $(m-r-1)$-flats, all contained in an $(m-r)$-flat.*

Proof: Given arbitrary elements $w_i \in E$, for $i = 1, \ldots, r$, and s *distinct* elements w'_j in E, let

$$p(x_1, \ldots, x_m) = \prod_{i=1}^{r} \left(1 - (x_i - w_i)^{q-1}\right) \prod_{j=1}^{s} (x_{r+1} - w'_j).$$

Then $p(x_1, \ldots, x_m)$ has degree $r(q-1) + s = \rho$ and is zero in $E^m = V$ unless

$$\begin{aligned} x_i &= w_i, \text{ for } i = 1, \ldots, r, & (6) \\ x_{r+1} &\neq w'_j \text{ for } j = 1, \ldots, s. & (7) \end{aligned}$$

There are $(q-s)q^{m-r-1}$ vectors in E^m satisfying both equations and the codeword corresponding to $p(x_1, \ldots, x_m)$ has this weight.

To establish the geometric nature of the codewords defined by such polynomials, consider the q^{m-r-1} points of E^m satisfying (6) and the additional equation $x_{r+1} = c$, where c is an element of E that is not amongst the w'_j. Then these points all belong to an $(m-r-1)$-flat and the corresponding coordinate positions in the codeword of $p(\mathbf{x})$ have the constant value

$$\prod_{j=1}^{s} (c - w'_j)$$

on these points. The same is true of each of the $(q-s)$ elements of E that are not amongst the w'_j, and hence we get a vector of the stated form. □

Remark: If $s=0$, then the polynomial $p(x_1,\ldots,x_m)$ is the incidence vector of an $(m-r)$-flat.

Corollary 5.26 *If $\rho = r(q-1)+s < m(q-1)$ with $0 \le s < q-1$, then $\mathcal{R}_{F_q}(\rho,m)$ has minimum weight $(q-s)q^{m-r-1}$ and $\mathcal{R}_{F_q}(\rho,m)^*$ has minimum weight $(q-s)q^{m-r-1}-1$.*

Proof: By taking the flat with $w_i = 0$ for $i = 1,\ldots,r$, and $w'_j \neq 0$ for $j = 1,\ldots,s$, it follows that the coordinate at the point $\mathbf{0}$ of the corresponding polynomial is non-zero, so that the corresponding codeword in $\mathcal{R}_{F_q}(\rho,m)^*$ has weight $(q-s)q^{m-r-1}-1$. This is the minimum weight by Theorem 5.24. By translation invariance the minimum weight of $\mathcal{R}_{F_q}(\rho,m)$ must be $(q-s)q^{m-r-1}$ since it has vectors of that weight and, as we have just seen, the minimum weight of $\mathcal{R}_{F_q}(\rho,m)^*$ is $(q-s)q^{m-r-1}-1$. □

Corollary 5.27 *Let p be a prime. The code over F_p of the design of points and r-flats of the affine geometry $AG_m(F_{p^t})$ has minimum weight p^{tr}.*

Proof: Apply Theorem 5.25 with $s=0$. Since the code of the design is a subset of the generalized Reed-Muller code, it must have at least this minimum weight, and since it has vectors of this weight, this must be the minimum weight. □

Corollary 5.28 *Let p be a prime. The code over F_p generated by the differences of the incidence vectors of two parallel r-flats of the affine geometry $AG_m(F_{p^t})$ has minimum weight $2p^{tr}$.*

Proof: Set $q=p^t$ and take $\rho = (m-r-1)(q-1)+(q-2)$. Each of the generating vectors of the code in question is in $\mathcal{R}_{F_q}(\rho,m)$ which, even as a code over F_q, has minimum weight $2q^r$. Thus the code of the design, being a subset and having vectors of this weight, must also have minimum weight $2q^r$. □

5.5 Codes invariant under the full affine group

Delsarte [19], generalizing the ideas of Kasami, Lin and Peterson and Theorem 4.6, characterized those codes invariant under $GL_m(F_q)$ in its natural action on the non-zero vectors of an m-dimensional vector space over F_q; he also discussed, in a very general way, the "projective" case. In our report [3] we described in full Delsarte's proof for the affine case when q is a prime. Here we give an alternative approach to this case due to Mortimer [44]; it is more direct and more suitable for our purposes.

Mortimer's results culminate in the proof that the only codes in E^{E^m}, where $E = F_p$ and p is a prime, invariant under G, where $ASL_m(F_p) \subseteq G \subseteq AGL_m(F_p)$, are the generalized Reed-Muller codes, $\mathcal{R}_E(\rho, m)$. We will prove his more general results leading to this, all of which can be found in [44, Chapter 5]. Note that we have already sketched in Section 4.6, Theorem 4.17, a new proof, due to Weidner, of the main result. Weidner's proof is in the single-variable context and utilizes the Jennings basis.

As usual, let $E = F_q$, where $q = p^t$ and p is a prime, and set $V = E^m$. We will be slightly more general than in Definition 5.3 and let K denote any extension field of E, and consider the vector space K^V — viewing that space as the space of linear combinations of $\mathcal{M}$ *with coefficients in the field* K, where $\mathcal{M}$ denotes the set of monomials in m variables, as in Section 5.2, Equation (1), page 43. For any ρ with $0 \leq \rho \leq m(q-1)$, we write

$$K_\rho = \{f \mid f \in K^V, \text{ and } \deg(f) \leq \rho\}, \tag{8}$$

where the degree of f is the total degree, i.e. the maximum value of $\sum a_i$ for the monomials $x_1^{a_1} x_2^{a_2} \ldots x_m^{a_m}$ that actually occur in the expression for f. For $K = E$ we have $K_\rho = \mathcal{R}_E(\rho, m)$.

For any integers $b \geq 0$ and i and j such that $1 \leq i, j \leq m$, define linear transformations, δ_i^b and $\epsilon_{i,j}^b$ from K^V to itself by giving them as follows on our chosen basis, the monomials in $\mathcal{M}$:

$$(x_1^{a_1} x_2^{a_2} \ldots x_m^{a_m})\delta_i^b = \binom{a_i}{b} x_1^{a_1} x_2^{a_2} \ldots x_i^{a_i - b} \ldots x_m^{a_m} \tag{9}$$

$$(x_1^{a_1} x_2^{a_2} \ldots x_m^{a_m})\epsilon_{i,j}^b = \binom{a_i}{b} x_1^{a_1} x_2^{a_2} \ldots x_i^{a_i - b} \ldots x_j^{a_j + b} \ldots x_m^{a_m}. \tag{10}$$

Since $\binom{a_i}{b} = 0$ for $a_i < b$, δ_i^b annihilates the monomial $x_1^{a_1} x_2^{a_2} \ldots x_m^{a_m}$ unless $b \leq a_i$; similarly $\epsilon_{i,j}^b$ annihilates $x_1^{a_1} x_2^{a_2} \ldots x_m^{a_m}$ unless $b \leq a_i$. Both δ_i^0 and $\epsilon_{i,j}^0$ are the identity on K^V.

Theorem 5.29 *Let T be the translation subgroup of $AGL_m(F_q)$. Then a subspace C of K^V is a T-module if and only if it is invariant under δ_i^b for all i and b such that $1 \leq i \leq m$ and $0 \leq b \leq q-1$.*

Proof: For $u \in E$ and $1 \leq i \leq m$ let τ_i^u denote the translation of V such that

$$\tau_i^u : (x_1, x_2, \ldots, x_m) \mapsto (x_1, \ldots, x_i - u, \ldots, x_m). \tag{11}$$

It is clearly sufficient to show that each τ_i^u in its action on K^V is a linear combination of the δ_i^b and conversely.

Let $f = \sum_j p_j x_i^j$ be any function in K^V where the p_j are polynomials independent of x_i. Then

$$\begin{aligned}
(f)\tau_i^u &= \sum_j p_j (x_i + u)^j \\
&= \sum_j p_j \sum_b \binom{j}{b} x_i^{j-b} u^b \\
&= \sum_b u^b \sum_j \binom{j}{b} p_j x_i^{j-b} \\
&= \sum_b u^b (f)\delta_i^b,
\end{aligned}$$

and thus $\tau_i^u = \sum_b u^b \delta_i^b$. On the other hand,

$$\begin{aligned}
(f) \sum_{u \in E^\times} u^{-b} \tau_i^u &= \sum_{u \in E^\times} u^{-b} \sum_j p_j (x_i + u)^j \\
&= \sum_{u \in E^\times} u^{-b} \sum_j p_j \sum_k \binom{j}{k} x_i^{j-k} u^k \\
&= \sum_k \sum_j p_j \binom{j}{k} x_i^{j-k} \sum_{u \in E^\times} u^{k-b} \\
&= \begin{cases} -(f)\delta_i^b & \text{if } b \neq 0, q-1 \\ -(f)\delta_i^0 - (f)\delta_i^{q-1} & \text{if } b = 0 \text{ or } q-1 \end{cases},
\end{aligned}$$

so that $\delta_i^b = -\sum_{u \in E^\times} u^{-b} \tau_i^u$ for $b \neq 0, q-1$, and $\delta_i^{q-1} = -\delta_i^0 - \sum_{u \in E^\times} \tau_i^u$. Thus each translation is a linear combination of the δ_i^b over K, and conversely, giving the theorem. □

Now we show that invariance under transvections is equivalent to invariance under the $\epsilon_{i,j}^b$. In fact, we need only the following transvections: for $u \in E$ and $i, j = 1, 2, \ldots m$ and $i \neq j$, define

$$\gamma_{i,j}^u : (x_1, x_2, \ldots, x_m) \mapsto (x_1, \ldots, x_i - ux_j, \ldots, x_m). \tag{12}$$

Then $\gamma_{i,j}^u$ is a transvection with axis given by $x_j = 0$. Recall that the *special* affine group, $ASL_m(F_q)$, is generated by the translations and transvections: see, for example, [23].

Theorem 5.30 *A subspace C of K^V is invariant under $ASL_m(F_q)$ if and only if it is invariant under all the transformations δ_i^b and $\epsilon_{i,j}^b$ with $0 \leq b \leq q-1$ and $i \neq j$ satisfying $1 \leq i, j \leq m$.*

Proof: In view of Theorem 5.29, since $SL_m(F_q)$ is spanned by the transvections $\gamma_{i,j}^u$, we need only show that each of these is a linear combination over E of the $\epsilon_{i,j}^b$, and conversely.

Any $f \in K^V$ can be written in the form

$$f = \sum_{r,s} p_{r,s} x_i^r x_j^s$$

where $p_{r,s}$ is a polynomial which is independent of x_i and x_j. Then

$$\begin{aligned}
(f)\gamma_{i,j}^u &= \sum_{r,s} p_{r,s}(x_i + ux_j)^r x_j^s \\
&= \sum_{r,s} p_{r,s} \sum_b \binom{r}{b} x_i^{r-b} x_j^{s+b} u^b \\
&= \sum_b u^b \sum_{r,s} \binom{r}{b} p_{r,s} x_i^{r-b} x_j^{s+b} u^b \\
&= \sum_b u^b (f)\epsilon_{i,j}^b.
\end{aligned}$$

Thus $\gamma_{i,j}^u = \sum_b u^b \epsilon_{i,j}^b$, and we can invert this formula to obtain the converse exactly as in the proof of Theorem 5.29. □

Theorem 5.31 *Let C be a subspace of K^V. Then C is invariant under $AGL_m(F_q)$ if and only if*

1. *C is invariant under the transformations δ_i^b and $\epsilon_{i,j}^b$ for $i \neq j$ and $1 \leq i, j \leq m$ and $0 \leq b \leq q-1$, and*

2. C is spanned by monomials.

Proof: By the previous theorems, the first condition characterizes subspaces invariant under $ASL_m(F_q)$, so it will suffice to show that an $ASL_m(F_q)$-invariant subspace is also invariant under $AGL_m(F_q)$ if and only if it is spanned by monomials. This is equivalent to showing that if a monomial appears with a non-zero coefficient in a function in C, then the monomial itself is in C.

The group $AGL_m(F_q)$ is generated by $ASL_m(F_q)$ and the dilations η_i^u defined by

$$\eta_i^u : (x_1, x_2, \dots, x_m) \mapsto (x_1, \dots, ux_i, \dots, x_m) \tag{13}$$

for $i = 1, 2, \dots, m$ and $u \in E^\times$. Suppose C is an $ASL_m(F_q)$-module spanned by monomials. Each η_i^u maps each monomial to a scalar multiple of itself, so C is invariant under η_i^u. Thus C is an $AGL_m(F_q)$ module.

Conversely suppose that C is an $AGL_m(F_q)$-invariant subspace that is not spanned by monomials. Thus C is invariant under the transformations

$$\lambda_i^k = -\sum_{u \in E^\times} u^k \eta_i^u, \tag{14}$$

for $0 \le k \le q-1$, and $1 \le i \le m$. Then

$$\begin{aligned}(x_1^{a_1} x_2^{a_2} \dots x_m^{a_m})\lambda_i^k &= (-\sum_{u \in E^\times} u^{k-a_i}) x_1^{a_1} x_2^{a_2} \dots x_m^{a_m} \\ &= \begin{cases} x_1^{a_1} x_2^{a_2} \dots x_m^{a_m} & \text{if } k \equiv a_i \pmod{(q-1)} \\ 0 & \text{otherwise} \end{cases} .\end{aligned}$$

Thus if $f \in C$ then $(f)\lambda_i^k \in C$ and consists of the terms of f that contain x_i^k if $k \neq 0, q-1$ and consists of those containing x_i^{q-1} or independent of x_i if $k = 0$ or $q-1$.

Choose $f \in C$ such that none of the monomial terms in it are in C. Subject to this condition choose f to have a minimal number of terms. Within f choose a term g with a maximal number of exponents which are neither 0 nor $q-1$. Subject to this condition choose g with a maximal number of exponents $q-1$. Relabelling subscripts we have

$$g = x_1^{a_1} x_2^{a_2} \dots x_r^{a_r} x_{r+1}^{q-1} \dots x_{r+s}^{q-1},$$

where $0 < a_i < q-1$.

The function $(f)\lambda_1^{a_1}\dots\lambda_r^{a_r}$ of C contains the terms of f that contain x_i raised to the exponent a_i for $i=1,\dots,r$. By the minimality of the number of terms in f, we have $f=(f)\lambda_1^{a_1}\dots\lambda_r^{a_r}$. Thus every monomial in f begins with $x_1^{a_1}x_2^{a_2}\dots x_r^{a_r}\dots$ and by the choice of g, the remaining exponents are 0 or $q-1$.

If $r+s<m$ then the monomial g is fixed by $\kappa_i=\epsilon_{i,m}^{q-1}\epsilon_{m,i}^{q-1}$ for $i=r+1,\dots,r+s$. If κ_i annihilates any monomial of f then $(f)\kappa_i$ contains fewer terms than f, contradicting the choice of f. Thus each term of f begins

$$x_1^{a_1}x_2^{a_2}\dots x_r^{a_r}x_{r+1}^{q-1}\dots x_{r+s}^{q-1}\dots.$$

Since g has the maximal number of exponents $q-1$ we have $f=ag$ for some $a\in E$, contradicting our hypothesis.

Thus $r+s=m$ and $g=x_1^{a_1}x_2^{a_2}\dots x_r^{a_r}x_{r+1}^{q-1}\dots x_m^{q-1}$. There must be another term h of f and we can take this to be

$$h=x_1^{a_1}x_2^{a_2}\dots x_r^{a_r}x_{r+1}^{q-1}\dots x_t^{q-1},$$

where $r<t<m$, by changing the last variables if necessary (leaving g fixed). Now $(h)\delta_m^{q-1}=0$, and

$$(f)\delta_m^{q-1}\delta_{m-1}^{q-1}\dots\delta_{t+1}^{q-1}=h+\dots$$

contains fewer terms than f and is still in C. This contradicts the choice of f as a function none of whose terms lies in C with a minimal number of terms. This contradiction gives the theorem. □

Now take $q=p$ a prime, so that K is any field of characteristic p.

Lemma 5.32 *The collection of transformations $\epsilon_{i,j}^k$ act transitively on the set of all monomials of fixed degree (ignoring scalar multiples) when $q=p$ is a prime.*

Proof: We prove this recursively. Let $g=x_1^{a_1}\dots x_m^{a_m}$ and $h=x_1^{b_1}\dots x_m^{b_m}$ be two monomials with $a_1+\dots+a_m=b_1+\dots+b_m$. Suppose that after a change of variables (if necessary) we have

$$a_1=b_1,\dots,a_{r-1}=b_{r-1},a_r>b_r,\dots,a_{s-1}>b_{s-1},a_s<b_s,\dots,a_m<b_m.$$

Clearly

$$a_r-b_r\le(b_s-a_s)+\dots+(b_m-a_m)$$

and thus there are integers c_j for $j = s, \dots, m$ with

$$a_r - b_r = c_s + \cdots + c_m$$

and $0 \leq c_j \leq b_j - a_j$. Thus

$$\begin{aligned} e &= (g)\epsilon_{r,s}^{c_s} \dots \epsilon_{r,m}^{c_m} \\ &= ux_1^{b_1} \dots x_{r-1}^{b_{r-1}} x_r^{a_r - c_s - \dots - c_m} x_{r+1}^{a_{r+1}} \dots x_{s-1}^{a_{s-1}} x_s^{a_s + c_s} \dots x_m^{a_m + c_m} \\ &= ux_1^{b_1} \dots x_r^{b_r} x_{r+1}^{a_{r+1}} \dots x_{s-1}^{a_{s-1}} x_s^{a_s + c_s} \dots x_m^{a_m + c_m} \end{aligned}$$

for some non-zero $u \in E$ is a monomial with one more exponent in common with h than g has. The lemma now follows by induction. □

Theorem 5.33 *Let*

$$ASL_m(F_p) \subseteq G \subseteq AGL_m(F_p)$$

where p is a prime. If C is a non-trivial G-invariant subspace of K^V, where $V = F_p^m$, then $C = K_k$ for some k such that $0 \leq k \leq m(p-1)$.

Proof: The proof is by induction on m. For $m = 1$ let $f \in C$ be of maximal degree k, say. Then $(f)\delta_1^i$ has degree $k - i$, and so C contains functions of each degree degree less than k. It follows that $C = K_k$.

Now suppose $m \geq 2$ and let $f \in C$ be of maximal degree k. If $k = 0$ then $C = K_0 = \langle \mathbf{1} \rangle$. If $k = 1$ then C contains a linear function, and hence all such, since G is transitive on linear functions; thus $C = K_1$. Now use induction on k. Then for some i $(f)\delta_i^1 \in C$ has degree $k - 1$ and hence $K_{k-1} \cap C$ is not trivial and thus equal to K_{k-1} by the induction hypothesis. Thus $K_{k-1} \subset C$. The function f can then be taken to be homogeneous of degree k.

Suppose $k \leq (m-1)(p-1)$. Choose a monomial $g = x_1^{a_1} \dots x_m^{a_m}$ amongst the terms of f with a_1 the maximal exponent of x_1 in f. From the lemma above we have a product σ of transformations $\epsilon_{i,j}^k$ such that $(g)\sigma$ is a monomial of degree k that is independent of x_1. Thus $h = (f)\sigma \in C$ is, by the maximality of a_1, independent of x_1. The subspace C' of C consisting of the functions in C that are independent of x_1 is invariant under $ASL_{m-1}(F_p)$. By the induction hypothesis then C contains every function of degree k that is independent of x_1. In particular, C contains a monomial of degree k and since G acts transitively on the monomials of degree k, it follows that C contains all the monomials of degree k. Since $C \supset K_{k-1}$, we have $C = K_k$.

Now suppose that $(m-1)(p-1) < k < m(p-1)$. Since $K_k^\perp = K_{m(p-1)-k-1}$ and

$$K_{k-1} \subset C \subseteq K_k,$$

we have

$$K_{m(p-1)-k-1} \subseteq C^\perp \subset K_{m(p-1)-k}.$$

Since from the above inequality we have that

$$m(p-1)-k \leq m(p-1)-(m-1)(p-1)-1 \leq (m-1)(p-1),$$

we get $C^\perp = K_{m(p-1)-k-1}$ by the argument above, and thus $C = K_k$. □

Corollary 5.34 *With the natural action of $AGL_m(F_p)$ on a vector space V of dimension m over F_p, where p is a prime, the only subspaces of F_p^V left invariant by $AGL_m(F_p)$ are the generalized Reed-Muller codes $\mathcal{R}_{F_p}(\rho, m)$.*

5.6 The geometric codes

We are now ready to consider the so-called non-primitive codes. We restrict ourselves to the case of geometric interest; the reader interested in the general case may wish to consult [2, Section 5.6].

Set $n = (q^m-1)/(q-1) = q^{m-1}+\ldots+1$ and observe that n is the number of points of $PG_{m-1}(F_q)$. We wish to look at those polynomials $P(Z) = \sum_{j=0}^{v-1} c_j Z^j \in L$ for which $P(\omega^i) = P(\omega^{n+i})$ for all i, in order that they should define functions on the projective points, ω^n being a primitive element of the field F_q. It follows from Equation (4) that, for such a polynomial, $c_j = \omega^{jn} c_j$ for each j, so that $c_j = 0$ unless $j \equiv 0 \pmod{q-1}$. Hence the polynomial has the form

$$P(Z) = \sum_{i=0}^{n-1} c_{i(q-1)} Z^{i(q-1)}. \tag{15}$$

We set L_{proj} equal to the subspace of all such polynomials of L.

Now since $P(\omega^i) = P(\omega^{i+n})$, the usual vector of length q^m-1 will consist of $q-1$ repetitions of the vector

$$(P(1), P(\omega), P(\omega^2), \ldots, P(\omega^{n-1})),$$

and we will take n to be the length of the geometric codes we will consider.

Since, for any integer j, $j \equiv \mathrm{wt}_q(j) \pmod{q-1}$, in the isomorphism we have given between R and $E[x_1, \ldots, x_m]/(x_1^q - x_1, \ldots, x_m^q - x_m)$, the polynomials in L_{proj} will correspond to reduced polynomials all of whose monomials will have degree divisible by $q-1$. It follows that we should look only at those generalized Reed-Muller codes of order divisible by $q-1$. Thus, we shall change notation and use r rather than $\rho = r(q-1)$ in the following

Definition 5.35 *The* r^{th} **order projective generalized Reed-Muller code**

$$\mathcal{P}_{F_q}(r,m)$$

where $0 \le r < m$ *is the code of length* $n = (q^m-1)/(q-1)$ *given as the set of vectors*

$$\{(P(1), P(\omega), \ldots, P(\omega^{n-1})) \mid P(X) \in L_{proj}, c_j = 0 \text{ for } \mathrm{wt}_q(j) > r(q-1)\}.$$

Here $P(Z)$ is defined in Equation (15) and L_{proj} is defined above. The r^{th} order projective generalized Reed-Muller code is also given by

$$\left\langle x_1^{i_1} x_2^{i_2} \ldots x_m^{i_m} \mid \sum_{k=1}^{m} i_k \equiv 0 \pmod{q-1}, \sum_{k=1}^{m} i_k \le r(q-1) \right\rangle,$$

where these polynomials are only evaluated on a set of representatives in F_q^m of the projective points. Observe that these codes are still cyclic since they are invariant under a Singer cycle. This is clear, of course, from the definition — since it is phrased in the single-variable language. The following proposition is also clear.

Proposition 5.36 *The dimension of* $\mathcal{P}_{F_q}(r,m)$ *is*

$$|\{j|0 \le j \le q^m - 1, q-1 \text{ divides } j, \mathrm{wt}_q(j) \le r(q-1)\}|.$$

Here the weight $\mathrm{wt}_q(j)$ is defined in Definition 4.7, on page 33.

Example 5.37 To construct the code $\mathcal{P}_{F_3}(1,3)$, of length 13 and dimension 7, the multi-variable formulation is the easiest to give, since the generating monomials are readily seen to be

$$\{1, x_1x_2, x_1x_3, x_2x_3, x_1^2, x_2^2, x_3^2\}.$$

If the irreducible cubic $X^3 - X^2 + 1$, with root ω, is used to obtain F_{27}, then the matrix

$$\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

can be used to generate representatives of the projective points, starting, say with $(1,0,0)^t$. The seven generating monomials yield the generator matrix

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & -1 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 1 & 1 & -1 & 0 & -1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

This code is the code over F_3 of the projective plane of order 3. The code vectors in G can be described geometrically: for example, labelling the columns 1 to 13, to represent the points, and using sets of these numbers to represent lines, then the penultimate row (corresponding to the monomial x_2^2) represents the complement of the line $\{1,3,4,8\}$, i.e. the vector $\mathbf{1}-v^{\{1,3,4,8\}}$ in our usual notation for characteristic functions. The second row (corresponding to the monomial x_1x_2) is the vector $v^{\{3,5,6,10\}} - v^{\{3,9,11,12\}}$. In terms of homogeneous coordinates for the projective geometry, the line $\{1,3,4,8\}$ represents the point set $\{(1,0,0),(0,0,1),(-1,0,1),(1,0,1)\}$, which is $(0,1,0)^t$ in homogeneous coordinates.

Since our projective codes are cyclic we can use the roots to obtain the orthogonal in the usual way. The code orthogonal to $\mathcal{P}_{F_q}(r,m)$ is obtained as follows:

Theorem 5.38 *If* $0 \le r < m$ *then*

$$(\mathcal{P}_{F_q}(r,m))^\perp = \mathcal{P}_{F_q}(m-r-1,m) \cap (F_q\mathbf{1})^\perp.$$

Remark: In Example 5.37, $r=1$ and the orthogonal is $\mathcal{P}_{F_3}(1,3) \cap (F_3\mathbf{1})^\perp$, as expected.

Theorem 5.39 *The minimum weight of* $\mathcal{P}_{F_q}(m-r,m)$ *is*

$$(q^r-1)/(q-1) = q^{r-1} + \ldots + 1.$$

Proof: Since, by Corollary 5.26, the minimum weight of $\mathcal{R}_{F_q}((m-r)(q-1),m)^*$ is q^r-1, the minimum weight is at least $(q^r-1)/(q-1)$. But the polynomial

$$p(x_1,\ldots,x_m) = \prod_{i=1}^{m-r}\left(1-(x_i)^{q-1}\right)$$

is such that each of its monomials has degree divisible by $q-1$ and yields a code vector. Since it takes the value 1 at $\mathbf{0}$, it obviously yields a vector of weight $(q^r-1)/(q-1)$ in $\mathcal{P}_{F_q}(m-r,m)$. $\Box$

The polynomial above that yields a minimum-weight vector is, in fact, the incidence vector of an r-dimensional subspace of F_q^m. Thus projectively it is an $(r-1)$-dimensional subspace of $PG_{m-1}(F_q)$.

5.7 The codes of the designs from $PG_m(F_p)$

We have already discussed in Section 4.6 the codes coming from $AG_m(F_p)$, showing that the code of the design of points and r-flats is precisely $\boldsymbol{M}^{r(p-1)}$, which, by Theorem 5.19, is $\mathcal{R}_{F_p}((m-r)(p-1),m)$. We are now in a position to consider the projective case — when $q=p$ is a prime[14] — and we wish to show that the code of the design of points and projective r-dimensional subspaces of $PG_m(F_p)$ is precisely $\mathcal{P}_{F_p}(m-r,m+1)$. If $\mathcal{P}$ is this design we already know from above, since $\mathcal{P}$ has an automorphism group that acts transitively on the set of r-dimensional subspaces, that $C_p(\mathcal{P}) \subseteq \mathcal{P}_{F_p}(m-r,m+1)$, since the characteristic functions of the r-dimensional subspaces are in $\mathcal{P}_{F_p}(m-r,m+1)$. Clearly, we can use a dimension argument to get the equality. The following lemma gives a recursion for the dimension of the projective generalized Reed-Muller codes at hand and will allow us to use an induction argument. In view of our treatment of the Reed-Muller codes, this recursion can be viewed as a substitute for the Pascal-triangle property enjoyed by the binomial coefficients.

Lemma 5.40 *For p a prime, the dimensions of the generalized Reed-Muller codes and the projective generalized Reed-Muller codes are related by the following recursion:*

$$\dim(\mathcal{P}_{F_p}(r-1,m)) + \dim(\mathcal{R}_{F_p}(r(p-1),m)) = \dim(\mathcal{P}_{F_p}(r,m+1)).$$

Proof: Let Q, A and P be the sets of integers whose cardinalities give the dimensions of $\mathcal{P}_{F_p}(r-1,m)$, $\mathcal{R}_{F_p}(r(p-1),m)$ and $\mathcal{P}_{F_p}(r,m+1)$, respectively. We must show that $|Q|+|A|=|P|$. Now Q is the set of integers satisfying $0 \le u \le p^m-1$, where $(p-1)$ divides u, and $\mathrm{wt}_p(u) \le (r-1)(p-1)$. Similarly, A is the set of integers satisfying $0 \le u \le p^m-1$ and $\mathrm{wt}_p(u) \le r(p-1)$ while P the set of integers satisfying $0 \le u \le p^{m+1}-1$, $p-1$ divides u, and $\mathrm{wt}_p(u) \le r(p-1)$.

[14]The general case will be treated in the following section.

Divide P into the following two disjoint sets: Q', the set of those integers in P whose p-ary expansion has $u_m = p-1$, and A', those integers in P whose p-ary expansion has $u_m < p-1$. For $u \in P$ where $u = u_0 + \cdots + u_m p^m$, set $f(u) = u - u_m p^m$. The reader will have no difficulty in seeing that f yields a one-to-one correspondence between Q' and Q and between A' and A. $\square$

We now use an embedding of $PG_{m-1}(F_p)$ in $PG_m(F_p)$ just as we did in the Reed-Muller case; here we know that the code of the projective design of r-dimensional subspaces projects *onto* $\mathcal{R}_{F_p}((m-r)(p-1), m)$, which is the code of the design of r-dimensional flats of $AG_m(F_p)$, and that the kernel contains the code of r-dimensional subspaces of $PG_{m-1}(F_p)$. An induction on m now yields the dimensional equality we seek and hence the following

Theorem 5.41 *For p a prime, the code over F_p of the design of points and r-dimensional subspaces of $PG_m(F_p)$ is the projective generalized Reed-Muller code $\mathcal{P}_{F_p}(m-r, m+1)$. Moreover, we have the following exact sequence for these codes:*

$$0 \to \mathcal{P}_{F_p}(m-r-1, m) \to \mathcal{P}_{F_p}(m-r, m+1) \to \mathcal{R}_{F_p}((m-r)(p-1), m) \to 0.$$

Remark: The sequence above can also be read as an exact sequence of the geometric codes and, as such, is

$$0 \to C_p(\mathcal{Q}) \to C_p(\mathcal{P}) \to C_p(\mathcal{A}) \to 0,$$

where $\mathcal{Q}$ is the design of points and r-dimensional subspaces of $PG_{m-1}(F_p)$, $\mathcal{P}$ the design of points and r-dimensional subspaces of $PG_m(F_p)$ and $\mathcal{A}$ the design of points and r-flats of $AG_m(F_p)$.

We now want to identity the minimum-weight vectors in these geometric codes. We already know the minimum weights and that among the minimum-weight vectors one finds the relevant geometric objects: the characteristic functions of the r-flats in the affine case and the characteristic functions of the r-dimensional projective subspaces in the projective case. We must, therefore, prove that only these vectors and their scalar multiples are minimum-weight vectors. We begin with the affine case:

Theorem 5.42 *For p a prime, the minimum-weight vectors of $\mathcal{R}_{F_p}((m-r)(p-1), m)$ are the scalar multiples of the characteristic function of the r-flats of $AG_m(F_p)$.*

Proof: We know that the scalar multiples of the r-flats are minimum-weight vectors. Moreover, if any minimum-weight vector has as its support an r-flat, then it clearly must be a scalar multiple of the characteristic function of that flat. Suppose, therefore, that we have a minimum-weight vector v whose support, X say, is *not* an r-flat. Of course, $|X| = p^r$. Without loss of generality we may assume that X contains the zero vector. Now, since we are over a prime field, the set X cannot be closed under addition (for then it would be a subspace). Let $\mathbf{x} \in X$ be such that $\mathbf{x} + X \neq X$. Now the vector v is a linear combination of characteristic functions of r-flats, i.e.

$$v = \sum_{S \in \mathcal{S}} a_S v^S,$$

where $\mathcal{S}$ is a collection of r-flats and the a_S are in F_p. Let w be the translate of v by $\mathbf{x}$. Then the support of w is $\mathbf{x} + X$, $v \neq w$, and

$$w = \sum_{S \in \mathcal{S}} a_S v^{\mathbf{x}+S}.$$

Since $v - w$ is in the code generated by the differences of the incidence vectors of parallel r-flats, it has, by Corollary 5.28, weight at least $2p^r$. But this is impossible since its support is a subset of $X \cup (\mathbf{x}+X)$, which is of cardinality less than $2p^r$ since $x \in X \cup (\mathbf{x} + X)$. Thus every minimum-weight vector is supported on an r-flat and hence is a scalar multiple of the characteristic function of an r-flat. □

We complete our discussion of the geometric codes in the case in which $q = p$ is a prime by showing that we have the analogous result in the projective case. In order to do so we first proceed more generally with q arbitrary and introduce some temporary notation.

Let $\mathcal{A}_{r,m}$ denote the design of points and r-flats of $AG_m(\mathbf{F}_q)$ and $\mathcal{P}_{r,m}$ denote the design of points and r-dimensional subspaces of $PG_m(\mathbf{F}_q)$.

Consider next the subcode $E_{r,m}$ of $C_p(\mathcal{A}_{r,m})$ generated by the differences of incidence vectors of parallel r-flats. Just as in the binary case (see Section 3.2) $E_{r,m}$ is in the kernel of the projection of $C_p(\mathcal{P}_{r,m})$ onto the coordinates corresponding to the embedded $(m-1)$-dimensional projective space, the image of the projection being $C_p(\mathcal{P}_{r-1,m-1})$. Observe that by using the $(q-1)$-to-1 map of $V - \{\mathbf{0}\}$ onto $PG_{m-1}(\mathbf{F}_q)$, where V is the m-dimensional vector space over $\mathbf{F}_q$ defining the projective space, we can pull the code $C_p(\mathcal{P}_{r-1,m-1})$ back to $\mathbf{F}_p^{V^*}$, where we are writing V^* for $V - \{\mathbf{0}\}$; this simply amounts to repeating each column $(q-1)$ times. By adjoining an overall parity check to this pull-back we get the code in $\mathbf{F}_p^V$ that is generated by the incidence vectors of the r-dimensional *subspaces* of V. Call this code $P_{r,m}$. Viewing $C_p(\mathcal{A}_{r,m})$ and $E_{r,m}$ in this same ambient space we have, clearly, that

$$E_{r,m} + P_{r,m} = C_p(\mathcal{A}_{r,m}).$$

This equation points to the reason why the binary case is so easy: when $q = 2$, $E_{r,m} \subseteq P_{r,m}$ and thus we need analyse only the projective geometry codes.

For the same reason as in the binary case, $P_{r+1,m} \subseteq P_{r,m}$ and, furthermore, $P_{r+1,m} \subseteq E_{r,m}$ since, if T is any $(r+1)$-dimensional subspace, S any r-dimensional subspace contained in it, and $\mathbf{v}$ is in T but not in S, then

$$-v^T = \sum_{a \in F_q, a \neq 0} (v^S - v^{a\mathbf{v}+S}).$$

Letting $a_{r,m}$ be the p-rank of $\mathcal{A}_{r,m}$, $p_{r,m}$ the p-rank of $\mathcal{P}_{r,m}$ and setting $e_{r,m} = \dim(E_{r,m})$, we have that $\dim(P_{r,m}) = p_{r-1,m-1}$ and that

$$p_{r-1,m-1} + e_{r,m} \geq a_{r,m} + p_{r,m-1}, \tag{16}$$

since the intersection, $P_{r,m} \cap E_{r,m}$, contains $P_{r+1,m}$. Further, we have the following:

Lemma 5.43 *Given an embedding of $PG_{m-1}(\mathbf{F}_q)$ in $PG_m(\mathbf{F}_q)$, if the first of the following sequences,*

$$0 \to C_p(\mathcal{P}_{r,m-1}) \to C_p(\mathcal{P}_{r,m}) \to C_p(\mathcal{A}_{r,m}) \to 0$$

and

$$0 \to E_{r,m} \to C_p(\mathcal{P}_{r,m}) \to C_p(\mathcal{P}_{r-1,m-1}) \to 0,$$

that arise from the embedding is exact, then so is the second and, moreover, in that case we have

$$p_{r,m} = a_{r,m} + p_{r,m-1} \text{ and } e_{r,m} + p_{r-1,m-1} = p_{r,m}.$$

Proof: Clearly, just as in the binary case, the sequences follow easily from the embedding, and we need only check that the kernels are as described. That the codes are contained in the kernels is obvious; thus in order to prove that they are the kernels we must check the dimensions. From the discussion preceding the lemma, in particular (16), and the second sequence, we have that

$$p_{r,m-1} + a_{r,m} \leq p_{r-1,m-1} + e_{r,m} \leq p_{r,m}$$

and the result follows since, if the first sequence is exact, $p_{r,m-1} + a_{r,m} = p_{r,m}$. $\square$

With this machinery in place one can now, for any prime p, imitate the proof for the binary case: see Theorem 3.13. Note that here we need only identify the vectors in the projective codes since we have already determined the minimum-weight vectors in the affine case. We leave to the reader the proof of the following

Theorem 5.44 *For p a prime, the minimum-weight vectors of the code of the design of points and r-dimensional subspaces of $PG_m(F_p)$ are the scalar multiples of the incidence vectors of these subspaces.*

Remark: We note that we have thus determined the minimum-weight vectors of the codes $\mathcal{P}_{F_p}(r,m)$ in the prime case, since these codes are codes of designs arising from $PG_{m-1}(F_p)$.

5.8 The subfield subcodes

To obtain the codes of the designs coming from affine and projective spaces over F_q, in the case in which q is a proper prime power, we need to restrict the codes, $\mathcal{R}_{F_q}((m-r)(q-1),m)$ and $\mathcal{P}_{F_q}((m-r),m+1)$, to **subfield subcodes** — also defined in Chapter 1, Section 5.

Definition 5.45 *Let C be a linear code over a field E and let F be a subfield of E. The set C' of vectors in C, all of whose coordinates lie in F, is called the* **subfield subcode** *of C over F.*

It is easy to verify that C' is a linear code over F and that any permutation of the coordinate positions preserving C also preserves C'. We are interested here **only** in the case where $F = F_p$, the **prime** subfield of $E = F_q$. In what follows $q = p^t$.

Definition 5.46 *Denote by $\mathcal{A}_{F_q/F_p}(\rho,m)$ the subfield subcode of the generalized Reed-Muller code $\mathcal{R}_{F_q}(\rho,m)$ and by $\mathcal{P}_{F_q/F_p}(r,m)$ the subfield subcode of the projective generalized Reed-Muller code $\mathcal{P}_{F_q}(r,m)$.*

Taking first the single-variable approach, $P(Z) = \sum_{j=0}^{v-1} c_j Z^j$ yields a vector in $\mathcal{A}_{F_q/F_p}(\rho, m)$ if $P(\omega^j) \in F$ for all j, and $P(0) \in F$, which is equivalent to the condition that $c_{jp} = c_j^p$ for all j, the subscripts being read modulo $v = q^m - 1$ as usual. Writing

$$V_\rho = \{u \mid 0 \leq u \leq q^m - 1, \mathrm{wt}_q(up^j) \leq \rho \text{ for } j = 0, 1, \ldots, t-1\}$$

(where up^j is taken reduced modulo $q^m - 1$, for the same reasons as before), we have that $\mathcal{A}_{F_q/F_p}(\rho, m)$ is given by

$$\left\{ (P(0), \ldots, P(\omega^{v-1})) \mid P(Z) = \sum_{u \in V_\rho} c_u Z^u, c_u \in F_{q^m}, c_{up} = (c_u)^p \right\}$$

and that

$$\dim(\mathcal{A}_{F_q/F_p}(\rho, m)) = |V_\rho|.$$

Clearly, by Theorem 5.10, $\mathcal{A}_{F_q/F_p}(\rho, m)$ will contain the incidence vector of any $(m-r)$-flat when $\rho \geq r(q-1)$. Its minimum weight d_ρ is bounded by

$$(q-s)q^{m-r-1} \leq d_\rho \leq q^{m-r},$$

where $\rho = r(q-1) + s$ and $0 \leq s < q-1$, and, from Theorem 5.25, attains the lower bound if there are s distinct w'_j in E such that the c_k, as defined there, are in F. In particular, if $s = 0$, then this holds and $d_\rho = q^{m-r}$. Further, in the case $s = q-2$ this is also the case: the vector obtained is the difference of the incidence vectors of two parallel r-flats, which is clearly a vector of the subfield subcode.

For the orthogonal code, $\mathcal{A}_{F_q/F_p}(\rho, m)^\perp$, clearly we have

$$\mathcal{A}_{F_q/F_p}(\mu, m) \subseteq \mathcal{A}_{F_q/F_p}(\rho, m)^\perp,$$

where $\mu = m(q-1) - \rho - 1 = (m-r-1)(q-1) + (q-2-s)$ (and $\rho = r(q-1) + s$, and $0 \leq s < q-1$, as above). Its minimum weight $d_\rho^\perp$ thus certainly satisfies

$$d_\rho^\perp \leq d_\mu \leq q^{r+1}, \qquad (17)$$

from the above discussion. A lower bound for $d_\rho^\perp$ follows from the BCH bound, and some evaluations of these are quoted in Delsarte et al. [18, Theorem 4.3.1]. In particular, for $\rho = r(q-1)$ this gives

$$d_{r(q-1)}^\perp \geq (p+q)q^{r-1},$$

and for $\rho = r(q-1) + (q-2) = (r+1)(q-1) - 1$, it gives

$$d^{\perp}_{r(q-1)+(q-2)} \geq q^{r+1},$$

which, from (17), yields

$$d^{\perp}_{r(q-1)+(q-2)} = q^{r+1}.$$

For the codes of designs arising from projective geometries, we must take the subfield subcodes of the codes $\mathcal{P}_{F_q}(m-r, m+1)$. As we have already indicated the minimum weight of this code is $q^r + q^{r-1} + \cdots + 1$ and the incidence vectors of the projective subspaces of dimension r are minimum-weight vectors.

Our interest is in the codes given by the designs of r-flats of the affine spaces and r-dimensional subspaces of the projective spaces. Just as in the binary case we must first analyse the codimension 1 case — in the projective case the design of points and hyperplanes of a projective space. This case was, historically, the one given the most attention and was introduced for projective *planes* by Prange with Rudolph considerably enriching the subject and making serious conjectures. The first systematic treatment in the case of planes was given by Graham and MacWilliams [22]. These results were generalized to higher dimensions by Goethals and Delsarte [21] and MacWilliams and Mann [39]; in particular, these authors computed the dimension of the code of the design of points and hyperplanes of an arbitrary projective space. The results are highly diverse and some of the proofs very technical: thus we only state what we need and show the reader how to construct these codes, using results of Delsarte et al. [18]. More recently, Rose [47] has given elegant new proofs of some of the results and Brouwer and Wilbrink [10, Theorem 4.8] have given a simple method to compute the p-ranks of the codes in question. We refer the reader also to a fuller account in Assmus and Key [2].

We thus now simply state the general theorem. Notice that everything stated is true also for $p = 2$, but that Theorem 3.14 gives more precise information in that case.

Theorem 5.47 *Let m be any positive integer, $q = p^t$ where p is a prime, and let $0 \leq r \leq m$.*

(1) The code over F_p of the design of points and r-flats in the affine geometry $AG_m(F_q)$, is $\mathcal{A}_{F_q/F_p}((m-r)(q-1), m)$. It has minimum weight

q^r and the minimum-weight vectors are the multiples of the incidence vectors of the r-flats. The p-rank is given by the cardinality of the set of integers u satisfying

- $0 \leq u \leq q^m - 1$
- $\mathrm{wt}_q(up^j) \leq (m-r)(q-1), j = 0, 1, \ldots, t-1$

where up^j is reduced modulo $q^m - 1$. The orthogonal code satisfies

$$\mathcal{A}_{F_q/F_p}((m-r)(q-1), m)^{\perp} \supseteq \mathcal{A}_{F_q/F_p}(r(q-1)-1, m)$$

and

$$\mathcal{A}_{F_q/F_p}(r(q-1)-1, m) = \langle v^M - v^N \mid M, N \textit{ parallel } (m-r)\textit{-flats in } V \rangle.$$

This latter code has minimum weight $2q^{m-r}$ with minimum-weight vectors multiples of the difference of the incidence vectors of two parallel $(m-r)$-flats . The minimum weight, $d^{\perp}_{(m-r)(q-1)}$, of the orthogonal code satisfies

$$(q+p)q^{m-r-1} \leq d^{\perp}_{(m-r)(q-1)} \leq 2q^{m-r}.$$

When $q = p$ the subfield codes are the generalized Reed-Muller codes, i.e.

$$\mathcal{A}_{F_p/F_p}((m-r)(p-1), m) = \mathcal{R}_{F_p}((m-r)(p-1), m)$$

and

$$\mathcal{R}_{F_p}((m-r)(p-1), m)^{\perp} = \mathcal{R}_{F_p}(r(p-1)-1, m).$$

(2) *The code over F_p of the design of points and r-dimensional subspaces of the projective geometry $PG_m(F_q)$ is $\mathcal{P}_{F_q/F_p}(m-r, m+1)$. It has minimum weight $(q^{r+1}-1)/(q-1)$ and the minimum-weight vectors are the multiples of the incidence vectors of the blocks. The p-rank is given by the cardinality of the set of integers u satisfying*

- $0 \leq u \leq q^{m+1} - 1$
- $(q-1)$ *divides* u
- $\mathrm{wt}_q(up^j) \leq (m-r)(q-1), j = 0, 1, \ldots, t-1$

where up^j is reduced modulo $q^{m+1} - 1$. The orthogonal code satisfies

$$\mathcal{P}_{F_q/F_p}(m-r, m+1)^{\perp} \supseteq \mathcal{P}_{F_q/F_p}(r, m+1) \cap \langle \mathbf{1} \rangle^{\perp}$$

and has minimum weight at least $(q^{m-r+1}-1)/(q-1)+1$*; if* $q=p$ *the subfield codes are the non-primitive generalized Reed-Muller codes and this becomes an equality.*

In particular, the code over F_p of the design of points and hyperplanes in the affine geometry $AG_m(F_q)$ is $\mathcal{A}_{F_q/F_p}(q-1,m)$ and the code over F_p of the design of points and hyperplanes of the projective geometry $PG_m(F_q)$ is $\mathcal{P}_{F_q/F_p}(1,m+1)$.

In order to actually construct the subfield subcodes, the m-variable approach is once again the most straightforward; it is described fully in [18]. Before describing the construction, we need some notation: if k satisfies $0\le k\le q-1$, and $k=\sum_{i=0}^{t-1}k_ip^i$, where $0\le k_i\le p-1$, then write

$$[pk]=k_{t-1}+k_0p+\cdots+k_{t-2}p^{t-1}=pk-k_{t-1}(q-1), \tag{18}$$

i.e.

$$[pk]=\begin{cases} pk \bmod (q-1) & \text{if } k<q-1\\ q-1 & \text{if } k=q-1.\end{cases}$$

Further, write $[k]=k$.

Theorem 5.48 *For any* ρ *such that* $0\le\rho\le m(q-1)$*, the code*

$$\mathcal{A}_{F_q/F_p}(\rho,m)$$

consists of the following polynomial functions, in terms of the usual basis of characteristic functions on F_q^m*:*

$$p(x_1,\ldots,x_m)=\sum_{l_1,\ldots,l_m}d(l_1,l_2,\ldots,l_m)x_1^{l_1}x_2^{l_2}\ldots x_m^{l_m},$$

where $0\le l_i\le q-1$*,* $d(l_1,l_2,\ldots,l_m)\in F_q$*, and*

(1) $\sum_{i=1}^m[p^jl_i]\le\rho$*, for* $j=0,1,\ldots,t-1$*;*

(2) $d([p^jl_1],\ldots,[p^jl_m])=(d(l_1,\ldots,l_m))^{p^j}$*, for* $j=0,1,\ldots,t-1$*.*

Example 5.49 Take $m=2$ and $q=4=2^2$. Thus $t=2$ and $0\le\rho\le 6$. Taking $\rho=3$ will give $\mathcal{A}_{F_4/F_2}(3,2)=C_2(AG_2(F_4))$. Then $V_3=\{0,1,2,3,4,6,8,9,12\}$ and so $\dim(\mathcal{A}_{F_4/F_2}(3,2))=9$. If ω is a primitive element for $E=F_4$, a root of $X^2+X+1=0$, then polynomials that generate the code, according to Theorem 5.48, are $\{1,x_1^3,x_2^3,x_1+x_1^2,x_2+x_2^2,\omega x_1+$

$\omega^2 x_1^2, \omega x_2 + \omega^2 x_2^2, x_1 x_2^2 + x_1^2 x_2, \omega x_1 x_2^2 + \omega^2 x_1^2 x_2\}$. A generator matrix from these polynomials can be constructed, and the entries are all, of course, in F_2. For example, if $K = F_{16}$ is constructed from E using the primitive polynomial $X^2 + \omega X + \omega$ and a is a root of this, then ordering the vectors of E^2 in the usual way, i.e. $\mathbf{0}, 1, a, a^2, \ldots, a^{14}$. Then the codeword obtained from the polynomial $\omega x_2 + \omega^2 x_2^2$ is

$$(0,0,1,1,1,1,0,1,0,1,1,0,0,1,0,0).$$

The codes $\mathcal{P}_{F_q/F_p}(r,m)$ can be constructed in a manner analogous to the primitive case as in Theorem 5.48. With the added condition that $(q-1)$ divides $\sum_i l_i$, the codewords are given by the first $n = (q^m - 1)/(q-1)$ coordinates, as in Definition 5.35.

Example 5.50 For $m = 3$, $q = 4$, $n = (4^3 - 1)/(4-1) = 21$, and taking $r = 1$, will produce $\mathcal{P}_{F_4/F_2}(1,3)$ as the binary code of the projective plane of order 4, $PG_2(F_4)$. If ω is a primitive element for $E = F_4$, then the polynomials that generate (over F_2) $\mathcal{P}_{F_4/F_2}(1,3)$ are $\{1, x_1^3, x_2^3, x_3^3, x_1 x_2^2 + x_1^2 x_2, \omega x_1 x_2^2 + \omega^2 x_1^2 x_2, x_2 x_3^2 + x_2^2 x_3, \omega x_2 x_3^2 + \omega^2 x_2^2 x_3, x_3 x_1^2 + x_3^2 x_1, \omega x_3 x_1^2 + \omega^2 x_3^2 x_1\}$. The dimension is thus 10 and a generator matrix is given by the codewords corresponding to each of these ten polynomials $p(x_1, \ldots, x_m)$, appropriately evaluated. Taking $X^3 + \omega^2 X^2 + \omega X + \omega$ for the generating polynomial of F_{64} over F_4. For example, the codeword corresponding to $p(x_1, \ldots, x_m) = \omega x_1 x_2^2 + \omega^2 x_1^2 x_2$ is

$$(0,0,0,1,1,0,1,0,0,1,1,0,0,1,0,0,1,0,1,0,0).$$

This is the vector $v^L - v^M$ where L and M are lines, $L = \{3,4,7,17,19\} = (1,1,0)^t$, and $M = \{3,5,10,11,14\} = (1,\omega,0)^t$, where the points are labelled $1, 2, \ldots, 21$ in the order given[15].

Note: The designs formed from affine or projective geometries may happen to have orders divisible by primes other than the characteristic prime for the geometry. The codes for such primes will not be of any interest — a result that follows from work of Mortimer [45] on the modular representations of doubly-transitive groups.

[15]Computations here and elsewhere were with Cayley[9] and Magma[11].

5.9 Formulas for p-ranks

The dimensions of the codes arising from finite geometries are given in the preceding section in terms of the number of integers with q-weight satisfying certain properties and this may very well be the most efficient way to calculate the dimension in the general case, since there seems to be no simple formula that will cover all possibilities — except in the case of the Reed-Muller codes, where $q = p = 2$. There are, however, simplifications in certain cases, and we give some of these below.

But we first give Hamada's [25, 26] rather complicated general formula:

Result 5.51 (Hamada) *Let $q = p^t$ and let $\mathcal{D}$ denote the design of points and r-dimensional subspaces of the projective geometry $PG_m(F_q)$, where $0 \leq r \leq m$. Then the p-rank of $\mathcal{D}$ is given by*

$$\sum_{s_0} \cdots \sum_{s_{t-1}} \prod_{j=0}^{t-1} \sum_{i=0}^{L(s_{j+1},s_j)} (-1)^i \binom{m+1}{i} \binom{m+s_{j+1}p-s_j-ip}{m},$$

where $s_t = s_0$ and summations are taken over all integers s_j (for $j = 0, 1, \ldots, t-1$) such that

$$r+1 \leq s_j \leq m+1, \text{ and } 0 \leq s_{j+1}p - s_j \leq (m+1)(p-1),$$

and

$$L(s_{j+1}, s_j) = \lfloor \frac{s_{j+1}p - s_j}{p} \rfloor,$$

i.e. the greatest integer not exceeding $(s_{j+1}p - s_j)/p$.

This formula is deduced in [25, 26]. It simplifies in certain cases, in particular in the case of designs of points and hyperplanes, when the formula becomes that found earlier by Graham and MacWilliams [22] for planes, and by Goethals and Delsarte [21], MacWilliams and Mann [39], and Smith [50], for general m. It becomes in that case:

Result 5.52 *If $q = p^t$, the p-rank of the design of points and hyperplanes of $PG_m(F_q)$ is*

$$\binom{m+p-1}{m}^t + 1.$$

If $q = p^t$, the p-rank of the design of points and $(m-1)$-flats of the affine geometry $AG_m(F_q)$ is

$$\binom{m+p-1}{m}^t.$$

Observe that the passage from the projective dimension to the affine dimension is, in the codimension-1 case, quite easy: since the minimum-weight is given by the number of points of a hyperplane and since the incidence vector of a hyperplane is a minimum-weight vector one simply projects onto the affine points off one such hyperplane — losing one dimension.

In the case $q = p = 2$, the codes are Reed-Muller, and the 2-ranks are explicitly given in Theorem 3.14. When $q = p$ in general, Theorem 5.5 gives the p-rank; we restate it here for the affine geometry designs.

Result 5.53 *For any r such that $0 \leq r \leq m$,*

$$\dim(\mathcal{R}_{F_p}((m-r)(p-1), m)) = \sum_{i=0}^{(m-r)(p-1)} \sum_{k=0}^{m} (-1)^k \binom{m}{k} \binom{i-kp+m-1}{i-kp}.$$

A simplification of this for $q = p$ prime and $r = 1$, when we have the design of points and lines, was obtained by Ceccherini and Hirschfeld [12]; a general summation formula in the prime case but for any r has been established by Hirschfeld and Shaw [27] and is as follows:

Result 5.54 *For p a prime, the p-rank of the design of points and r-dimensional projective subspaces of $PG_m(F_p)$ is*

$$p^m + p^{m-1} + \ldots + 1 - \sum_{i=0}^{r-1} (-1)^i \binom{(r-i)(p-1)-1}{i} \binom{m+(r-i)p-r}{m-i}.$$

For $r = 1$ Result 5.54 has a simple form which follows from Theorem 5.38 and the simple form for the dimension of points and hyperplanes; note that the derivation of the formula below is particularly easy since one can utilize Corollary 4.19 to get the dimension of the projective code of points and hyperplanes — as explained above — and not even invoke Theorem 5.38, but instead note that the complements of the hyperplanes give the necessary orthogonal (with a compensating loss of the gained dimension). Thus the code over F_p of the design of points and lines of the projective geometry $PG_m(F_p)$ where p is a prime has dimension

$$p^m + p^{m-1} + \ldots + 1 - \binom{m+p-1}{m}.$$

These simple derivations for the case $q = p$ a prime highlight the difficulties of the prime-power case.

As we have seen in Corollary 4.20:

Result 5.55 *The p-rank of the design of points and lines of the affine geometry $AG_m(F_p)$, where p is a prime, is*

$$p^m - \binom{m+p-2}{m}.$$

In the case of $q = p = 3$ and $r = 1$, we have a Steiner triple system. Hence we have

Result 5.56 *The 3-rank of the Steiner triple system of points and lines of $AG_m(F_3)$ is $3^m - 1 - m$.*

Another particular case that gave a bound for the p-rank of a translation plane is the following from Key and Mackenzie [33]:

Result 5.57 *If $\mathcal{D}$ is the design of points and m-flats in $AG_{2m}(F_p)$, where p is a prime, then the p-rank of $\mathcal{D}$ is given by*

$$\dim\left(\mathcal{R}_{F_p}(m(p-1), 2m)\right) = \mathrm{rank}_p(\mathcal{D}) = \sum_{i=0}^{m-1} (-1)^i \binom{2m}{i} \binom{m+(m-i)p}{2m}.$$

There are also other cases where simpler arguments give the p-rank and even a basis in terms of incidence vectors of the geometric objects involved. For example, Bagchi and Sastry [4] have produced a simple derivation of the dimension of the binary code of the design of points and planes in $PG_3(F_{2^t})$ by finding a set of planes whose incidence vectors form a basis:

Result 5.58 (Bagchi and Sastry) *Let $\mathcal{D}$ be the design of points and planes in $PG_3(F_{2^t})$ and let $\mathcal{O}$ be an ovoid in $PG_3(F_{2^t})$. Then the incidence vectors of the tangent planes to the ovoid form a basis for $C_2(\mathcal{D})$. It follows that* $\dim(C_2(\mathcal{D})) = 2^{2t} + 1$.

Another result which describes an explicit basis of incidence vectors — in this case lines — of the affine plane $AG_2(F_p)$, where p is a prime, has been obtained by Moorhouse [43]. In this case the dimension of the code over F_p is $\binom{p+1}{2} = \sum_{i=1}^{p} i$, and a basis can be had by ordering, arbitrarily, any p of the $p+1$ parallel classes and taking one line from the first, two from the second, etc., even the choices of the lines being made arbitrarily; the basis consists of the incidence vectors of the selected lines.

A conjecture of Hamada (see [25, 26]) that the p-rank of the design of points and r-dimensional flats of a finite-geometry design over a field of characteristic p is always the smallest for designs with the same parameters and also *characterizes* such designs, is false in general, the counter-examples first occuring for 2-(31,7,7) designs: see Tonchev [51] and Delsarte and Goethals [21]. However, the minimality of the p-rank still appears to be true, and the conjecture still stands for designs of points and hyperplanes and also for designs of points and lines; moreover, when $p = q = 2$, this limited conjecture is valid.

References

[1] E. Artin. *Geometric Algebra.* New York: Wiley Interscience, 1957.

[2] E. F. Assmus, Jr. and J. D. Key. *Designs and their Codes.* Cambridge University Press, 1992. Cambridge Tracts in Mathematics, Vol. 103 (Second printing with corrections, 1993).

[3] Edward F. Assmus, Jr. and Jennifer D. Key. Codes and finite geometries. Technical report, INRIA, 1993. Report No. 2027.

[4] Bhaskar Bagchi and N. S. Narasimha Sastry. Even order inversive planes, generalized quadrangles and codes. *Geom. Dedicata*, 22:137–147, 1987.

[5] Thierry Berger and Pascale Charpin. The automorphism group of Generalized Reed-Muller codes. *Discrete Math.*, 117:1–17, 1993.

[6] S. D. Berman. On the theory of group codes. *Kibernetika*, 3(1):31–39, 1967.

[7] Richard E. Blahut. *Theory and Practice of Error Control Codes.* Addison-Wesley, 1983.

[8] Ian F. Blake and Ronald C. Mullin. *The Mathematical Theory of Coding.* New York: Academic Press, 1975.

[9] W. Bosma and J. Cannon. *Handbook of Cayley Functions.* Department of Mathematics, University of Sydney, January 1993.

[10] Andries E. Brouwer and Henny A. Wilbrink. Block designs. In F. Buekenhout, editor, *Handbook of Incidence Geometry*, pages 349–382. Elsevier, 1995. Chapter 8.

[11] John Cannon and Catherine Playoust. *An Introduction to Magma.* School of Mathematics and Statistics, University of Sydney, 1994.

[12] P. V. Ceccherini and J. W. P. Hirschfeld. The dimension of projective geometry codes. *Discrete Math.*, 106/107:117–126, 1992.

[13] P. Charpin. Représentation des codes cycliques primitifs dans une algèbre modulaire. Technical report, 1991. Rapport n^o 82, Département de mathématiques et d'informatique, Faculté des Sciences, Université de Sherbrooke, Edité par I. Assem et B. Courteau.

[14] Pascale Charpin. *Codes cycliques étendus invariants sous le groupe affine.* Thèse de Doctorat d'État, Université Paris VII, 1987.

[15] Pascale Charpin. Une generalisation de la construction de Berman des codes de Reed et Muller p-aires. *Communications in Algebra*, 16:2231–2246, 1988.

[16] Pascale Charpin. Codes cycliques étendus affines-invariants et antichaines d'un esemble partiellement ordonne. *Discrete Math.*, 80:229–247, 1990.

[17] Pascale Charpin and Françoise Levy-dit-Vehel. On self-dual affine-invariant codes. *J. Combin. Theory, Ser. A.* To appear.

[18] P. Delsarte, J. M. Goethals, and F. J. MacWilliams. On generalized Reed-Muller codes and their relatives. *Inform. and Control*, 16:403–442, 1970.

[19] Philippe Delsarte. On cyclic codes that are invariant under the general linear group. *IEEE Trans. Inform. Theory*, 16:760–769, 1970.

[20] Andreas Faldum. Reed-Muller codes are optimal in the class of the radical codes. *Arch. Math.* To appear.

[21] Jean-Marie Goethals and Philippe Delsarte. On a class of majority-logic decodable cyclic codes. *IEEE Trans. Inform. Theory*, 14:182–188, 1968.

[22] R. L. Graham and Jessie MacWilliams. On the number of information symbols in difference-set cyclic codes. *Bell System Tech. J.*, 45:1057–1070, 1966.

[23] K. W. Gruenberg and A. J. Weir. *Linear Geometry.* Graduate Texts in Mathematics: 49. New York: Springer Verlag, 2nd edition, 1977.

[24] A. J. Hahn and O. T. O'Meara. *The Classical Groups and K-Theory.* Grundlehren der mathematischen Wissenschaften 291. New York: Springer-Verlag, 1989.

[25] N. Hamada. The rank of the incidence matrix of points and d-flats of in finite geometries. *J. Sci. Hiroshima Univ. Ser. A-I*, 32:381–396, 1968.

[26] N. Hamada. On the p-rank of the incidence matrix of a balanced or partially balanced incomplete block design and its applications to error correcting codes. *Hiroshima Math. J.*, 3:153–226, 1973.

[27] J. W. P. Hirschfeld and R. Shaw. Projective geometry codes over prime fields. In *Finite Fields: Theory, Application and Algorithms.* American Mathematical Society. Contemporary Mathematics Series.

[28] S. A. Jennings. The structure of the group ring of a p-group over a modular field. *Trans. Amer. Math. Soc.*, 50:175–185, 1941.

[29] T. Kasami, S. Lin, and W. W. Peterson. Some results on cyclic codes which are invariant under the affine group and their applications. *Inform. and Control*, 11:475–496, 1967.

[30] Tadeo Kasami, Shu Lin, and W. Wesley Peterson. New generalizations of the Reed-Muller codes. Part I: Primitive codes. *IEEE Trans. Inform. Theory*, 14:189–199, 1968.

[31] Tadeo Kasami, Shu Lin, and W. Wesley Peterson. Polynomial codes. *IEEE Trans. Inform. Theory*, 14:807–814, 1968.

[32] J. D. Key and F. E. Sullivan. Codes of Steiner triple and quadruple systems. *Des. Codes Cryptogr.*, 3:117–125, 1993.

[33] Jennifer D. Key and Kirsten Mackenzie. An upper bound for the p-rank of a translation plane. *J. Combin. Theory, Ser. A*, 56:297–302, 1991.

[34] Peter Landrock and Olaf Manz. Classical codes as ideals in group algebras. *Des. Codes Cryptogr.*, 2:273–285, 1992.

[35] Rudolf Lidl and Harald Niederreiter. *Introduction to Finite Fields and their Applications.* Cambridge: Cambridge University Press, 1986.

[36] Shu Lin and Daniel J. Costello, Jr. *Error Control Coding: Fundamentals and Applications.* Prentice-Hall, 1983. Englewood Clifts, NJ.

[37] J. H. van Lint. *Introduction to Coding Theory.* Graduate Texts in Mathematics 86. New York: Springer-Verlag, 1982.

[38] Lucio Lombardo-Radice. Intorno alle algebre legate ai gruppi di ordine finito. *Rend. Sem. Mat. Fac. Sci. R. Univ. Roma (4)*, 2:312–322, 1938.

[39] F. J. MacWilliams and H. B. Mann. On the p-rank of the design matrix of a difference set. *Inform. and Control*, 12:474–489, 1968.

[40] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes.* Amsterdam: North-Holland, 1983.

[41] James L. Massey. *Threshold Decoding.* The M.I.T. Press, 1963.

[42] H. F. Mattson and G. Solomon. A new treatment of Bose-Chaudhuri codes. *J. Soc. Indust. Appl. Math.*, 9:654–669, 1961.

[43] G. Eric Moorhouse. Bruck nets, codes, and characters of loops. *Des. Codes Cryptogr.*, 1:7–29, 1991.

[44] Brian Mortimer. *Some problems on permutation groups: affine groups and modular permutation representations.* PhD thesis, Westfield College, University of London, 1977.

[45] Brian Mortimer. The modular permutation representations of the known doubly transitive groups. *Proc. London Math. Soc. (3)*, 41:1–20, 1980.

[46] I. S. Reed. A class of multiple-error-correcting codes and the decoding scheme. *IRE Trans.*, IT-4:38–49, 1954.

[47] Kimberly Jan Rose. *Generalized Reed-Muller codes and finite geometries.* PhD thesis, Lehigh University, 1993.

[48] Joseph J. Rushanan. On the Vandermonde matrix. *Amer. Math. Monthly*, 96:921–924, 1989.

[49] James Singer. A theorem in finite projective geometry and some applications to number theory. *Trans. Amer. Math. Soc.*, 43:377–385, 1938.

[50] K. J. C. Smith. On the p-rank of the incidence matrix of points and hyperplanes in a finite projective geometry. *J. Combin. Theory*, 7:122–129, 1969.

[51] Vladimir D. Tonchev. Quasi-symmetric 2-(31,7,7) designs and a revision of Hamada's conjecture. *J. Combin. Theory, Ser. A*, 42:104–110, 1986.

[52] Vladimir D. Tonchev. Quasi-symmetric designs, codes, quadrics, and hyperplane sections. *Geom. Dedicata*, 48:295–308, 1993.

[53] M. A. Tsfasman and S. G. Vlăduţ. *Algebraic-Geometric Codes.* Kluwer Academic Publishers, 1991. Mathematics and Its Applications (Soviet Series).

[54] Michael Weidner. The action of $GL(m, F_p)$ on $F_pE(m,p)$. Private communication, 1993.

[55] Edward J. Weldon, Jr. New generalizations of the Reed-Muller codes. Part II: Nonprimitive codes. *IEEE Trans. Inform. Theory*, 14:199–205, 1968.

Index

Glossary