THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 4, April, 1972
Printed in U.S.A.

The MacWilliams Identities for Nonlinear Codes

By Mrs. F. J. MACWILLIAMS, N. J. A. SLOANE, and J.-M. GOETHALS

(Manuscript received December 13, 1971)

In recent years a number of nonlinear codes have been discovered which have better error-correcting capabilities than any known linear codes. However, very little is known about the properties of such codes. In this paper we study the most basic property, the weight enumerator. The weight of a codeword is the number of its nonzero components; the weight enumerator gives the number of codewords of each weight, and is fundamental for obtaining the error probability when the code is used for error-correction on a noisy channel. In 1963 one of us showed that the weight enumerator of a linear code is related in a simple way to that of the dual code (Jessie MacWilliams, "A Theorem on the Distribution of Weights in a Systematic Code," Bell System Technical Journal, 42, No. 1 (January 1963), pp. 79–94). In the present paper, which is a sequel, we show that the same relationship holds for the weight enumerator of a nonlinear code. Furthermore, a definition is given for the dual $\mathcal{A}^{\perp}$ of a nonlinear binary code $\mathcal{A}$ which satisfies $(\mathcal{A}^{\perp})^{\perp} = \mathcal{A}$ provided $\mathcal{A}$ contains the zero codeword.

I. INTRODUCTION

In recent years a number of nonlinear codes have been discovered which have better error-correcting capabilities than any known linear codes (e.g., Refs. 1 and 2). However, very little is known about the properties of such codes. In this paper we study the most basic property, the Hamming weight enumerator (defined in Section II), which gives fundamental information about the error probability when the code is used in various error-correction schemes (Ref. 3, Ch. 16). In 1963 one of us showed that the Hamming and the complete weight enumerators of a linear code are related in a simple way to those of the dual code (Ref. 4; Theorems 1 and 3 below). The requirement that the code be linear is unsatisfactory for two reasons: (*i*) Several pairs of nonlinear codes $\mathcal{A}$, $\mathcal{B}$ are known whose weight enumerators satisfy Theorem 3. One example of such a pair is given by the Preparata[2] and Kerdock[1] codes, another by the code shown in Fig. 1. (*ii*) The important theorem of S. P. Lloyd (giving a necessary condition for the existence of a prefect code) may be deduced for linear codes as a corollary to Theorem 3 (Ref. 4, Lemma 2.15), but may be proved directly without assuming linearity (Ref. 5; Ref. 6, p. 111).

It is the purpose of the present paper, therefore, to define the "weight enumerators of the dual code" so as to make Theorems 1 and 3 (and the corresponding theorem for the Lee weight enumerator, Theorem 2) valid even for nonlinear codes.

Furthermore, if $\mathcal{A}$ is a nonlinear binary code which contains the zero codeword, we define the formal dual $\mathcal{A}^{\perp}$ so as to satisfy:

(*i*) $(\mathcal{A}^{\perp})^{\perp} = \mathcal{A}$,
(*ii*) if $\mathcal{A}$ is linear the two definitions of $\mathcal{A}^{\perp}$ agree.

The paper is arranged as follows. Section II states the three MacWilliams identities (Theorems 1, 2, 3). Section III treats the binary case, when the three theorems coincide. The formal dual of a nonlinear binary code is defined in Section 3.5. Section IV treats the general case, first proving Theorem 1 and then deducing Theorems 2, 3 from it. In Section V we discuss properties of the "weights of the dual code" $B(i)$. However, the problem of finding conditions for the $B(i)$ to be positive integers remains unsolved.

0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1
1	0	1	0	0	0	0	0	0	1	0	1	1	1	1	1
1	0	0	1	0	0	0	0	0	1	1	0	1	1	1	1
1	0	0	0	1	0	0	0	0	1	1	1	0	1	1	1
1	0	0	0	0	1	0	0	0	1	1	1	1	0	1	1
1	0	0	0	0	0	1	0	0	1	1	1	1	1	0	1
1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0

Fig. 1—The sixteen rows form a nonlinear code $\mathcal{A}$.

II. WEIGHT ENUMERATORS

Let F be a finite field $GF(q)$, where q is a prime power; and let F^n be a vector space of dimension n over F. A *linear code* $\mathcal{A}$ of length n over $GF(q)$ is a subspace of F^n, and $\mathcal{A}^\perp$ denotes the orthogonal subspace or *dual* code of $\mathcal{A}$. A code is *self-dual* if $\mathcal{A} = \mathcal{A}^\perp$. A *nonlinear* code is any subset of F^n. In this paper a code is linear unless stated otherwise.

We propose to describe the code vectors of a code $\mathcal{A}$ in three ways, giving progressively less information (but becoming progressively easier to handle).

2.1 *The Complete Weight Enumerator*

Let the elements of F be $\omega_0 = 0, \omega_1, \omega_2, \cdots, \omega_{q-1}$, in some fixed order. The *composition* of a vector $\mathbf{v} \in F^n$ is defined to be

$$\text{comp}\,(\mathbf{v}) = \mathbf{s} = (s_0, s_1, \cdots, s_{q-1}), \tag{1}$$

where $s_j = s_j(\mathbf{v})$ is the number of coordinates of $\mathbf{v}$ equal to ω_j. Clearly $\sum_{j=0}^{q-1} s_j = n$.

Let $A(\mathbf{t})$ be the number of vectors $\mathbf{v}$ in $\mathcal{A}$ with comp $(\mathbf{v}) = \mathbf{t}$. The set of integers $\{A(\mathbf{t})\}$ is the *complete weight enumerator* of $\mathcal{A}$.

The first MacWilliams identity relates the complete weight enumerators of $\mathcal{A}$ and $\mathcal{A}^\perp$. (Ref. 4, Lemma 2.7. See also Refs. 7 and 8.)

Theorem 1: If $\mathcal{A}$ is a linear code with complete weight enumerator $\{A(\mathbf{t})\}$, and its dual code $\mathcal{A}^\perp$ has complete weight enumerator $\{B(\mathbf{t})\}$, then

$$\sum_{\mathbf{s}} B(\mathbf{s}) z_0^{s_0} \cdots z_{q-1}^{s_{q-1}} = \frac{1}{|\mathcal{A}|} \sum_{\mathbf{t}} A(\mathbf{t}) \prod_{l=0}^{q-1} \left(\sum_{j=0}^{q-1} \mathcal{X}(\omega_j \omega_l) z_j \right)^{t_l} \tag{2}$$

where the z_i are indeterminates and $\mathcal{X}$ is a character on $GF(q)$ (defined in Section 4.2).

2.2 *The Lee Weight Enumerator*

For $q = 2$ this description coincides with the preceding, and for $q = 2^s$, $s > 1$ it is not defined; so in this section q is assumed to be an odd prime power.

For q prime, we wish to classify the coordinates of the code vectors by magnitude. For example, codewords over $GF(5) = \{0, 1, -1, 2, -2\}$ would be classified according to the number of components which are 0, the number which are ± 1, and the number which are ± 2 (but without regard to the actual number which are $1, -1, 2$, or -2).

In general, for q a prime power, let the elements of F be $\omega_0 = 0$, $\omega_1, \cdots, \omega_\delta, \omega_{-\delta}, \omega_{-\delta+1}, \cdots, \omega_{-1}$, where $\omega_{-i} = -\omega_i$ and $\delta = \frac{1}{2}(q-1)$.

Then the *Lee weight* of a vector $\mathbf{v} \varepsilon F^n$ is defined to be

$$\text{Lee}\ (\mathbf{v}) = (l_0\, , l_1\, , \cdots , l_\delta),$$

where $l_i = l_i(\mathbf{v})$ is the number of coordinates of $\mathbf{v}$ equal to either ω_i or $-\omega_i$. In the notation of eq. (1),

$$\begin{aligned} l_0(\mathbf{v}) &= s_0(\mathbf{v}) \\ l_i(\mathbf{v}) &= s_i(\mathbf{v}) + s_{-i}(\mathbf{v}) \quad \text{for} \quad i = 1, \cdots , \delta. \end{aligned} \tag{3}$$

Let $A^L(\mathbf{t})$ be the number of vectors $\mathbf{v}$ in $\mathcal{A}$ with Lee $(\mathbf{v}) = \mathbf{t}$; so that $\{A^L(\mathbf{t})\}$ is the *Lee weight enumerator* of $\mathcal{A}$.

The second MacWilliams identity relates the Lee weight enumerators of $\mathcal{A}$ and $\mathcal{A}^\perp$:

Theorem 2:

$$\sum_{\mathbf{s}} B^L(\mathbf{s}) z_0^{s_0} \cdots z_\delta^{s_\delta}$$

$$= \frac{1}{|\mathcal{A}|} \sum_{t} A^L(\mathbf{t}) \prod_{l=0}^{\delta} \left(z_0 + \sum_{j=1}^{\delta} (\mathfrak{X}(\omega_j \omega_l) + \mathfrak{X}(-\omega_j \omega_l)) z_i \right)^{t_l}, \tag{4}$$

where $\{B^L(\mathbf{s})\}$ *is the Lee weight enumerator for* $\mathcal{A}^\perp$.

(Theorem 2 is believed to be new.) The Lee enumerator is important both because it is an appropriate measure for codes to be used in phase-modulation communication schemes (see Lee, Ref. 9; Berlekamp, Ref. 3, p. 205) and as a compromise in giving much more information than the Hamming enumerator, yet requiring only half as many variables as the complete enumerator.

2.3 *The (Hamming) Weight Enumerator*

For the rest of the paper let q be any prime power.

The (Hamming) *weight* of a vector $\mathbf{v}$, $wt(\mathbf{v})$, is the number of its nonzero coordinates, so that

$$wt(\mathbf{v}) = \sum_{i=1}^{q-1} s_i(\mathbf{v}). \tag{5}$$

Let $\mathcal{A}$ be a linear code of length n over $GF(q)$, and let $A(i)$ be the number of vectors $\mathbf{v}$ in $\mathcal{A}$ with $wt(\mathbf{v}) = i$. Then $\{A(i)\}$ is the (Hamming or ordinary) *weight enumerator* of $\mathcal{A}$. Similarly $\{B(i)\}$ denotes the weight enumerator of the dual code $\mathcal{A}^\perp$. The third MacWilliams identity (Ref. 4, Theorem 1) relates $\{A(i)\}$ and $\{B(i)\}$:

Theorem 3:

$$\sum_{i=0}^{n} B(i)z^{i} = \frac{1}{|\mathcal{A}|} \sum_{i=0}^{n} A(i)(1 + (q - 1)z)^{n-i}(1 - z)^{i}. \tag{6}$$

2.4 *An Example*

Let $\mathcal{A}$ be the self-dual code of length 2 over $GF(5)$ consisting of the code vectors 0 0, 1 2, 2 -1, -2 1, -1 -2.

The complete, Lee, and Hamming weight enumerators are, respectively,

$$\begin{aligned} A(20000) = A(01100) = A(01001) &= A(01010) \\ = A(00011) &= 1, \end{aligned}$$

$$A^{L}(200) = 1, \qquad A^{L}(011) = 4,$$

and

$$A(0) = 1, \qquad A(2) = 4.$$

In this case, $\mathfrak{X}(\omega_j\omega_l) = \alpha^{jl}$ where $\alpha = e^{(2\pi i)/5} = \cos 72° + i \sin 72°$. Theorems 1, 2, 3 assert (correctly) that

$$\begin{aligned} z_0^2 + z_1z_2 + z_2z_{-1} + z_1z_{-2} + z_{-2}z_{-1} = \tfrac{1}{5}[&(z_0 + z_1 + z_2 + z_{-2} + z_{-1})^2 \\ &+ (z_0 + \alpha z_1 + \alpha^2 z_2 + \alpha^3 z_{-2} + \alpha^4 z_{-1})(z_0 + \alpha^2 z_1 + \alpha^4 z_2 + \alpha z_{-2} + \alpha^3 z_{-1}) \\ &+ (z_0 + \alpha^2 z_1 + \alpha^4 z_2 + \alpha z_{-2} + \alpha^3 z_{-1})(z_0 + \alpha^4 z_1 + \alpha^3 z_2 + \alpha^2 z_{-2} + \alpha z_{-1}) \\ &+ (z_0 + \alpha z_1 + \alpha^2 z_2 + \alpha^3 z_{-2} + \alpha^4 z_{-1})(z_0 + \alpha^3 z_1 + \alpha z_2 + \alpha^4 z_{-2} + \alpha^2 z_{-1}) \\ &+ (z_0 + \alpha^3 z_1 + \alpha z_2 + \alpha^4 z_{-2} + \alpha^2 z_{-1})(z_0 + \alpha^4 z_1 + \alpha^3 z_2 + \alpha^2 z_{-2} + \alpha z_{-1})], \end{aligned}$$

that

$$\begin{aligned} z_0^2 + 4z_1z_2 = \tfrac{1}{5}[&(z_0 + 2z_1 + 2z_2)^2 \\ &+ 4(z_0 + (\alpha + \alpha^4)z_1 + (\alpha^2 + \alpha^3)z_2)(z_0 + (\alpha^2 + \alpha^3)z_1 + (\alpha + \alpha^4)z_2)], \end{aligned}$$

and that

$$1 + 4z^2 = \tfrac{1}{5}[(1 + 4z)^2 + 4(1 - z)^2].$$

III. THE BINARY CASE

All the codes in this section are binary, so that Theorems 1 and 2 coincide with Theorem 3.

3.1 *Preliminaries*

Let $F = GF(2)$; let F^n be a vector space of dimension n over F. For purposes of notation we define a group G which is a multiplicative copy of F^n, as follows. Let $x_1, \cdots, x_n$ be indeterminates satisfying $x_i^2 = 1$ and $x_i x_j = x_j x_i$ for $i, j = 1, \cdots, n$. Then G is the multiplicative group consisting of all products $x_1^{v_1} x_2^{v_2} \cdots x_n^{v_n}$ where v_i is 0 or 1. To each vector

$$\mathbf{v} = (v_1, v_2, \cdots, v_n)$$

in F^n we associate the element

$$x^{\mathbf{v}} = x_1^{v_1} x_2^{v_2} \cdots x_n^{v_n}$$

of G. Thus F^n and G are isomorphic, and addition of vectors in F^n corresponds to multiplication in G.

3.2 *Characters*

Let $\mathfrak{X}_{\mathbf{u}}$, $\mathbf{u} \in F^n$, be a character of G given by

$$\mathfrak{X}_{\mathbf{u}}(x^{\mathbf{v}}) = (-1)^a,$$

where $a = \mathbf{u}\mathbf{v}^T$ is the scalar product of $\mathbf{u}$, $\mathbf{v}$ in $GF(2)$.

Let σ_i be the set of vectors of F^n of weight i. Clearly,

$$|\sigma_i| = \binom{n}{i}.$$

Let

$$X_i = \sum_{\mathbf{v} \in \sigma_i} x^{\mathbf{v}}.$$

(For example, $X_1 = x_1 + x_2 + \cdots + x_n$.) X_i is an element of the group algebra QG of G over the field of rational numbers Q.

$\mathfrak{X}_{\mathbf{u}}$ is extended linearly to elements of QG, for example,

$$\mathfrak{X}_{\mathbf{u}}(X_i) = \sum_{\mathbf{v} \in \sigma_i} \mathfrak{X}_{\mathbf{u}}(x^{\mathbf{v}}).$$

Note that $\mathfrak{X}_{\mathbf{u}}(X_i)$ is a rational integer, not an element of $GF(2)$.

Let S_n be the group of all permutations of n symbols, i.e., the group of all $n \times n$ permutation matrices. $\mathbf{v}\pi$ is the vector obtained from $\mathbf{v}$ by multiplying by the permutation matrix π.

Lemma 3.1:

$$\mathfrak{X}_{\mathbf{u}\pi}(x^{\mathbf{v}}) = \mathfrak{X}_{\mathbf{u}}(x^{\mathbf{v}\pi^T}) \quad \textit{for any } \pi \textit{ in } S_n.$$

Proof:

$$\mathfrak{X}_{\mathbf{u}\pi}(x^{\mathbf{v}}) = (-1)^a,$$

$$a = \mathbf{u}\pi\mathbf{v}^T = \mathbf{u}(\mathbf{v}\pi^T)^T. \qquad \text{Q.E.D.}$$

3.3 *Krawtchouk Polynomials*

The *Krawtchouk polynomial* $P_s(i)$ (a polynomial in s) is defined by

$$(1 + z)^{n-s}(1 - z)^s = \sum_{i=0}^{n} P_s(i)z^i, \tag{7}$$

so that

$$P_s(i) = \sum_{r=0}^{\min(i,s)} (-1)^r \binom{s}{r}\binom{n-s}{i-r} \qquad i = 0, \cdots, n. \tag{8}$$

It follows from the definition that

$$\sum_{i=0}^{n} P_s(i) = 2^n\, \delta_{s,0}\,. \tag{9}$$

Other properties may be found in Refs. 10 and 11.

Let J_s be the vector with $v_1 = v_2 = \cdots = v_s = 1$ and $v_{s+1} = \cdots = v_n = 0$.

Lemma 3.2: *If* $\mathbf{u}$ *has weight* s,

$$\mathfrak{X}_{\mathbf{u}}(X_i) = P_s(i).$$

Proof: Since X_i is clearly invariant under any permutation in S_n we may suppose, by (3.1), that $\mathbf{u} = J_s$.

Consider the formal sum

$$\sum_{i=0}^{n} \mathfrak{X}_{J_s}(X_i)z^i = \mathfrak{X}_{J_s}\left(\sum_{i=0}^{n} X_i z^i\right).$$

Now

$$\sum_{i=0}^{n} X_i z^i = \prod_{j=1}^{n} (1 + x_j z), \tag{10}$$

and

$$\mathfrak{X}_{J_s}(1 + x_j z) = \begin{cases} 1 - z & \text{if} \quad j = 1, \cdots, s, \\ 1 + z & \text{if} \quad j = s + 1, \cdots, n. \end{cases}$$

Thus

$$\sum_{i=0}^{n} \mathfrak{X}_{J_s}(X_i)z^i = (1 + z)^{n-s}(1 - z)^s. \qquad \text{Q.E.D.}$$

Lemma 3.3:

$$\binom{n}{i}P_i(s) = \binom{n}{s}P_s(i).$$

Proof: By rearranging the binomial coefficients in eq. (8).

3.4 *Definition of B(i) and Proof of Theorem 3*

Let $\mathcal{A}$ be an arbitrary (linear or nonlinear) code, i.e., any subset of F^n; let $A(i)$ be the number of vectors in $\mathcal{A}$ of weight i. Define

$$\boldsymbol{\mathcal{A}} = \sum_{\mathbf{v} \varepsilon \mathcal{A}} x^{\mathbf{v}};$$

$\boldsymbol{\mathcal{A}}$ is an element of QG. Corresponding to $\mathcal{A}$ we define numbers $B(i)$, $i = 0, 1, \cdots, n$, by

$$B(i) = \frac{1}{|\mathcal{A}|} \sum_{\mathbf{u} \varepsilon \sigma_i} \mathfrak{X}_{\mathbf{u}}(\boldsymbol{\mathcal{A}}). \tag{11}$$

Note that $B(i)$ is a rational number, perhaps negative.

With this definition of $B(i)$ we can now prove the binary version of Theorem 3, as follows. Define

$$\boldsymbol{\mathcal{A}}^{\pi} = \sum_{\mathbf{v} \varepsilon \mathcal{A}} x^{\mathbf{v}\pi}.$$

We average $\boldsymbol{\mathcal{A}}$ over all equivalent codes $\boldsymbol{\mathcal{A}}^{\pi}$:

Lemma 3.4:

$$\sum_{\pi \varepsilon S_n} \boldsymbol{\mathcal{A}}^{\pi} = \sum_{i=0}^{n} A(i) i!\,(n-i)!\,X_i\,.$$

Proof: Let $\mathbf{v}$ be a vector of weight i in $\mathcal{A}$. The $i!$ permutations of the nonzero symbols of $\mathbf{v}$ leave $\mathbf{v}$ unchanged, as do the $(n-i)!$ permutations of the places in which $\mathbf{v}$ contains zero. Thus

$$\sum_{\pi \varepsilon S_n} x^{\mathbf{v}\pi} = i!\,(n-i)!\,X_i\,. \qquad \text{Q.E.D.}$$

Lemma 3.5:

$$B(j) = \frac{1}{|\mathcal{A}|}\,\frac{1}{j!\,(n-j)!} \sum_{\pi \varepsilon S_n} \mathfrak{X}_{J_j\pi}(\mathcal{A}).$$

Proof: As π runs through S_n, $J_j\pi$ runs through $j!(n-j)!$ copies of σ_j.

Q.E.D.

Proof of Theorem 3: By (3.5), (3.1):

$$\begin{aligned} B(j) &= \frac{1}{|\mathcal{A}|}\,\frac{1}{j!\,(n-j)!}\,\mathfrak{X}_{J_j}\Big(\sum_{\pi\,\varepsilon\, S_n}\mathcal{A}^{\pi}\Big) \\ &= \frac{1}{|\mathcal{A}|}\,\frac{1}{j!\,(n-j)!}\,\mathfrak{X}_{J_j}\Big(\sum_{i=0}^{n} A(i)i!\,(n-i)!\,X_i\Big) \quad \text{by (3.4)}, \\ &= \frac{1}{|\mathcal{A}|}\sum_{i=0}^{n} A(i)\,\frac{i!\,(n-i)!}{j!\,(n-j)!}\,P_j(i) \quad \text{by (3.2)}, \\ &= \frac{1}{|\mathcal{A}|}\sum_{i=0}^{n} A(i)P_i(j) \quad \text{by (3.3)}. \end{aligned}$$

Multiply both sides by z^j and sum on j:

$$\begin{aligned} \sum_{j=0}^{n} B(j)z^j &= \frac{1}{|\mathcal{A}|}\sum_{i=0}^{n} A(i)\sum_{j=0}^{n} P_i(j)z^j \\ &= \frac{1}{|\mathcal{A}|}\sum_{i=0}^{n} A(i)(1+z)^{n-i}(1-z)^{i}. \qquad \text{Q.E.D.} \end{aligned}$$

In the next section we show that in the case $\mathcal{A}$ is linear, $B(i)$ is the usual weight distribution of the dual code.

3.5 *The Dual Code*

If $\boldsymbol{\mathcal{A}} = \sum_{\mathbf{v}\varepsilon F^n} \alpha_{\mathbf{v}} x^{\mathbf{v}}$, $\alpha_{\mathbf{v}}\ \varepsilon\ Q$, is any element of QG for which $A(0) = 1$, we define its formal weight distribution to be $\{A(i)\}$, where

$$A(i) = \sum_{\mathbf{v}\varepsilon\sigma_i} \alpha_{\mathbf{v}}\ , \tag{12}$$

$$|\boldsymbol{\mathcal{A}}| = \sum_{i=0}^{n} A(i), \tag{13}$$

and its formal dual to be

$$\boldsymbol{\mathcal{A}}^{\perp} = \frac{1}{|\boldsymbol{\mathcal{A}}|}\sum_{\mathbf{u}\varepsilon F^n} \mathfrak{X}_{\mathbf{u}}(\boldsymbol{\mathcal{A}})x^{\mathbf{u}}. \tag{14}$$

It follows from (12) that the formal weight distribution of $\boldsymbol{\mathcal{A}}^{\perp}$ is $\{B(i)\}$, where

$$B(i) = \frac{1}{|\boldsymbol{\mathcal{A}}|}\sum_{\mathbf{u}\varepsilon\sigma_i} \mathfrak{X}_{\mathbf{u}}(\boldsymbol{\mathcal{A}}). \tag{11$'$}$$

If $\boldsymbol{\mathcal{A}}$ is a linear or nonlinear code, then clearly (12), (13) give the usual weight distribution and total number of codewords, and eq. (11′) for $B(i)$ coincides with eq. (11) of Section 3.4.

Theorem 4: *If* $\mathbf{\mathcal{A}}$ *is a linear code, then the expressions (14), (11′) for its dual code and weight distribution of dual code, coincide with the usual definitions.*

Proof: If $\mathbf{u}$ is in the dual subspace to $\mathcal{A}$, then $\mathfrak{X}_{\mathbf{u}}(x^{\mathbf{v}}) = 1$ for all $\mathbf{v} \in \mathcal{A}$, so $\mathfrak{X}_{\mathbf{u}}(\mathbf{\mathcal{A}}) = |\mathcal{A}|$. If $\mathbf{u} \notin \mathcal{A}^{\perp}$, then $\mathbf{uv}^T \equiv 1$ (modulo 2) for exactly half the vectors $\mathbf{v} \in \mathcal{A}$, so

$$\mathfrak{X}_{\mathbf{u}}(\mathbf{\mathcal{A}}) = 0 \quad \text{for} \quad \mathbf{u} \notin \mathcal{A}^{\perp}.$$

Therefore from (14),

$$\mathbf{\mathcal{A}}^{\perp} = \frac{1}{|\mathcal{A}|} \sum_{\mathbf{u} \in \mathcal{A}\perp} x^{\mathbf{u}}. \qquad \text{Q.E.D.}$$

Combining Theorem 4 with the results of the last section, we have completed the proof of Theorem 3 for binary linear codes.

Theorem 5: *Let* $\mathbf{\mathcal{A}} = \sum_{\mathbf{v} \in F^n} \alpha_{\mathbf{v}} x^{\mathbf{v}}$, $\alpha_{\mathbf{v}} \in Q$, *be any element of* QG *for which* $A(0) = 1$, *with formal dual* $\mathbf{\mathcal{A}}^{\perp}$ *given by eq. (14). Then*

(*i*) $|\mathbf{\mathcal{A}}|\,|\mathbf{\mathcal{A}}^{\perp}| = 2^n$,
(*ii*) $(\mathbf{\mathcal{A}}^{\perp})^{\perp} = \mathbf{\mathcal{A}}$.

(Note that by the earlier remarks this theorem includes linear and nonlinear binary codes as a special case.)

Proof: (*i*) Set $z = 1$ in Theorem 3.
(*ii*) From (14), $(\mathbf{\mathcal{A}}^{\perp})^{\perp} = \sum_{\mathbf{u} \in F^n} \beta_{\mathbf{u}} x^{\mathbf{u}}$, where

$$\begin{aligned}\beta_{\mathbf{u}} &= \frac{1}{|\mathbf{\mathcal{A}}^{\perp}|} \mathfrak{X}_{\mathbf{u}}(\mathbf{\mathcal{A}}^{\perp}),\\ &= \frac{1}{2^n} \sum_{\mathbf{v} \in F^n} \mathfrak{X}_{\mathbf{v}}(\mathbf{\mathcal{A}}) \mathfrak{X}_{\mathbf{u}}(x^{\mathbf{v}}) \quad \text{by } (i), (14),\\ &= \frac{1}{2^n} \sum_{\mathbf{v} \in F^n} \mathfrak{X}_{\mathbf{v}}\Big(\sum_{\mathbf{w} \in F^n} \alpha_{\mathbf{w}} x^{\mathbf{w}} \Big) \mathfrak{X}_{\mathbf{u}}(x^{\mathbf{v}}),\\ &= \frac{1}{2^n} \sum_{\mathbf{w} \in F^n} \alpha_{\mathbf{w}} \sum_{\mathbf{v} \in F^n} (-1)^{\mathbf{v}(\mathbf{u}+\mathbf{w})^T},\\ &= \frac{1}{2^n} (2^n \alpha_{\mathbf{u}}),\end{aligned}$$

since the innermost sum is zero unless $\mathbf{u} = \mathbf{w}$. Q.E.D.

Remarks: In spite of Theorem 5, eq. (14) is not always a satisfactory definition of the dual of a nonlinear code, even in the binary case.

For example, Fig. 1 shows a nonlinear code with weight distribution $A(0) = A(8) = 1$, $A(2) = A(6) = 7$, and

$$\mathfrak{A} = 1 + x_1(x_2 + x_3 + \cdots + x_8) + x_1 \cdots x_8\left(1 + \frac{1}{x_1}\left(\frac{1}{x_2} + \cdots + \frac{1}{x_8}\right)\right).$$

When the weight distribution is substituted in the right-hand side of the MacWilliams identity (6), $B(i)$ is found to be the same as $A(i)$ (Ref. 4, bottom of p. 82) so that this code is in some sense self-dual. However, although eq. (11) correctly gives the weight distribution $B(0) = B(8) = 1$, $B(2) = B(6) = 7$, eq. (14) gives

$$\mathfrak{A}^{\perp} = 1 - \tfrac{1}{2}x_1(x_2 + x_3 + \cdots + x_8) + \tfrac{1}{2} \sum_{2 \leqq i < j \leqq 8} x_i x_j + \cdots$$

which seems unsatisfactory. A better definition of the dual of a nonlinear code has recently been given by P. Delsarte and J. -M. Goethals (private communication).

IV. THE GENERAL CASE

4.1 *Preliminaries*

Let $q = p^f$, $f \geqq 1$, where p is prime; and let $F = GF(q) = \{\omega_o = 0, \omega_1, \cdots, \omega_{q-1}\}$. Let $x_i^{(\omega_j)}$ be commuting indeterminates satisfying

$$x_i^{(\omega_j)} x_i^{(\omega_k)} = x_i^{(\omega_j + \omega_k)};$$

and let G be the multiplicative group consisting of all products $x_1^{(v_1)} x_2^{(v_2)} \cdots x_n^{(v_n)}$, $v_i \in F$. To each vector $\mathbf{v} = (v_1, \cdots, v_n)$ in F^n we associate the element $x^{(\mathbf{v})} = x_1^{(v_1)} \cdots x_n^{(v_n)}$ of G; as in Section 3.1, G is a multiplicative copy of F^n. Let $\mathcal{C}G$ be the group algebra of G over the complex numbers.

4.2 *Characters*

Let $p(x)$ be a primitive irreducible polynomial of degree f over $GF(p)$, and let α be a root of $p(x)$. Then any element $\lambda \in GF(q)$ has the canonical representation

$$\lambda = \lambda_0 + \lambda_1\alpha + \lambda_2\alpha^2 + \cdots + \lambda_{f-1}\alpha^{f-1}, \qquad \lambda_i \in GF(p).$$

If $GF(q)$ is considered as an additive group, it forms an abelian group, denoted by $(GF(q), +)$, which is isomorphic to the direct product of f copies of $GF(p)$; the isomorphism being given for example by

$$\lambda \leftrightarrow (\lambda_0, \lambda_1, \cdots, \lambda_{f-1}).$$

A character $\mathcal{X}$ on $GF(q)$ is a homomorphism from $(GF(q), +)$ to the multiplicative group of the complex numbers. Define a fixed character on $GF(q)$ by

$$\mathcal{X}(\lambda) = \xi^{\lambda_0}$$

where $\xi = e^{(2\pi i)/p}$, and

$$\mathcal{X}(\lambda + \mu) = \xi^{\lambda_0 + \mu_0}.$$

All characters on $GF(q)$ are now given by

$$\mathcal{X}_\nu(\lambda) = \mathcal{X}(\lambda\nu), \qquad \text{all } \nu \in GF(q).$$

All of the following depends on the choices of $p(x)$, α, and $\mathcal{X}$; this dependence on coordinatization seems inevitable in studying codes over $GF(q)$.

Define a character $\mathcal{X}_{\mathbf{u}}$ on G by

$$\mathcal{X}_{\mathbf{u}}(x^{\mathbf{v}}) = \mathcal{X}(\mathbf{u}\mathbf{v}^T) = \mathcal{X}\left(\sum_{i=1}^{n} u_i v_i\right) \tag{15}$$

where $\sum_{i=1}^{n} u_i v_i \in GF(q)$. These characters form a group isomorphic to G (and to F^n): $\mathcal{X}_{\mathbf{u}} \leftrightarrow x^{\mathbf{u}}$. We extend $\mathcal{X}_{\mathbf{u}}$ to $\mathcal{C}G$ by linearity.

Lemma 4.1:

$$\mathcal{X}_{\mathbf{u}\pi}(x^{(\mathbf{v})}) = \mathcal{X}_{\mathbf{u}}(x^{(\mathbf{v}\pi^T)}) \quad \text{for any } \pi \in S_n .$$

The proof is straightforward and is omitted.

4.3 *Generalized Krawtchouk Polynomials.*

Let $\mathbf{s} = (s_0, s_1, \cdots, s_{q-1})$, $\mathbf{t} = (t_0, t_1, \cdots, t_{q-1})$ be compositions as defined in Section 2.1. The *generalized Krawtchouk polynomial* $P_{\mathbf{s}}(\mathbf{t})$ is defined by

$$\prod_{l=0}^{q-1} \left(\sum_{i=0}^{q-1} \mathcal{X}(\omega_i \omega_l) z_i \right)^{s_l} = \sum_{\mathbf{t}} P_{\mathbf{s}}(\mathbf{t}) z_0^{t_0} z_1^{t_1} \cdots z_{q-1}^{t_{q-1}}. \tag{16}$$

$$\text{Let} \quad X_{\mathbf{t}} = \sum_{\substack{\mathbf{v} \in F^n \\ \text{comp}(\mathbf{v}) = \mathbf{t}}} x^{\mathbf{v}}.$$

Lemma 4.2:

$$\prod_{k=1}^{n} \sum_{i=0}^{q-1} x_k^{(\omega_i)} z_i = \sum_{\mathbf{t}} X_{\mathbf{t}} z_0^{t_0} z_1^{t_1} \cdots z_{q-1}^{t_{q-1}}.$$

This is a straightforward generalization of eq. (10). For example, expand the product ($n = 3$, $q = 4$)

$$(x_1^{(\omega_0)}z_0 + x_1^{(\omega_1)}z_1 + x_1^{(\omega_2)}z_2 + x_1^{(\omega_3)}z_3)$$
$$\cdot(x_2^{(\omega_0)}z_0 + x_2^{(\omega_1)}z_1 + x_2^{(\omega_2)}z_2 + x_2^{(\omega_3)}z_3)$$
$$\cdot(x_3^{(\omega_0)}z_0 + x_3^{(\omega_1)}z_1 + x_3^{(\omega_2)}z_2 + x_3^{(\omega_3)}z_3).$$

Lemma 4.3: For any composition **s** *let*

$$\mathbf{u} = (\overset{\longleftarrow s_0 \longrightarrow}{\omega_0\omega_0 \cdots \omega_0}\overset{\longleftarrow s_1 \longrightarrow}{\omega_1\omega_1 \cdots \omega_1} \cdots \overset{\longleftarrow s_{q-1} \longrightarrow}{\omega_{q-1}\omega_{q-1} \cdots \omega_{q-1}})$$

so that comp (**u**) = **s**. *Then*

$$\mathfrak{X}_{\mathbf{u}}(X_{\mathbf{t}}) = P_{\mathbf{s}}(\mathbf{t}).$$

Proof: Consider the formal sum

$$\sum_{\mathbf{t}} \mathfrak{X}_{\mathbf{u}}(X_{\mathbf{t}}) z_0^{t_0} \cdots z_{q-1}^{t_{q-1}} = \mathfrak{X}_{\mathbf{u}}\left(\prod_{k=0}^{n} \sum_{i=0}^{q-1} x_k^{(\omega_i)} z_i\right) \quad \text{by (4.2)},$$
$$= \prod_{k=1}^{n} \sum_{i=0}^{q-1} \mathfrak{X}_{\mathbf{u}}(x_k^{(\omega_i)}) z_i$$
$$= \prod_{k=1}^{n} \sum_{i=0}^{q-1} \mathfrak{X}(u_k\omega_i) z_i \quad \text{by eq. (15)},$$
$$= \prod_{l=0}^{q-1} \left(\sum_{i=0}^{q-1} \mathfrak{X}(\omega_i\omega_l) z_i\right)^{s_l} \text{by the form of } \mathbf{u},$$
$$= \sum_{\mathbf{t}} P_{\mathbf{s}}(\mathbf{t}) z_0^{t_0} \cdots z_{q-1}^{t_{q-1}}$$

by eq. (16). Q.E.D.

For a composition **s**, let $\binom{n}{\mathbf{s}}$ denote the multinomial coefficient $n!/(s_0!s_1! \cdots s_{q-1}!)$.

Lemma 4.4:

$$\binom{n}{\mathbf{s}} P_{\mathbf{s}}(\mathbf{t}) = \binom{n}{\mathbf{t}} P_{\mathbf{t}}(\mathbf{s}).$$

Proof: Set $\alpha_l = \sum_{i=0}^{q-1} \mathfrak{X}(\omega_i\omega_l) z_i$, so (16) becomes

$$\prod_{l=0}^{q-1} \alpha_l^{s_l} = \sum_{\mathbf{t}} P_{\mathbf{s}}(\mathbf{t}) \prod_i z_i^{t_i}.$$

Multiply by $\prod_{l=0}^{q-1} \binom{n}{\mathbf{s}} y_l^{s_l}$ and sum on **s**:

$$\sum_{\mathbf{s}} \binom{n}{\mathbf{s}} \prod_{l=0}^{q-1} (\alpha_l y_l)^{s_l} = \sum_{\mathbf{s},\mathbf{t}} \binom{n}{\mathbf{s}} P_{\mathbf{s}}(\mathbf{t}) \prod_i z_i^{t_i} \prod_l y_l^{s_l}. \tag{17}$$

The left-hand side is

$$(\alpha_0 y_0 + \alpha_1 y_1 + \cdots + \alpha_{q-1} y_{q-1})^n$$

which rearranged becomes

$$(\beta_0 z_0 + \beta_1 z_1 + \cdots + \beta_{q-1} y_{q-1})^n, \tag{18}$$

where

$$\beta_i = \sum_{l=0}^{q-1} \mathfrak{X}(\omega_i \omega_l) y_l .$$

Expanding (18) we get

$$\sum_{\mathbf{t}} \binom{n}{\mathbf{t}} \prod_{i=0}^{q-1} \left(\sum_{l=0}^{q-1} \mathfrak{X}(\omega_i \omega_l) y_l \right)^{t_i} z_i^{t_i} = \sum_{\mathbf{s},\mathbf{t}} \binom{n}{\mathbf{t}} \sum_{\mathbf{s}} P_{\mathbf{t}}(\mathbf{s}) \prod_l y_l^{s_l} \prod_i z_i^{t_i}. \tag{19}$$

Equating coefficients in (17), (19) gives the result. Q.E.D.

4.4 *Definition of $B(\mathbf{s})$ and Proof of Theorem 1*

As in Section 2.1, let $\mathcal{A}$ be any code in F^n, with complete weight enumerator $\{A(\mathbf{t})\}$; and let

$$\boldsymbol{\mathcal{A}} = \sum_{\mathbf{v} \in \mathcal{A}} x^{\mathbf{v}}$$

be the corresponding element of $\mathcal{C}G$. For each composition $\mathbf{s}$ define

$$B(\mathbf{s}) = \frac{1}{|\mathcal{A}|} \sum_{\substack{\mathbf{u} \in F^n \\ \text{comp}(\mathbf{u}) = \mathbf{s}}} \mathfrak{X}_{\mathbf{u}}(\mathcal{A}). \tag{20}$$

In general $B(\mathbf{s})$ is a complex number. With this definition of $B(\mathbf{s})$ we can now prove Theorem 1.

Remark: If $\mathcal{A}$ is a linear code it follows immediately (as in the proof of Theorem 4) that $\{B(\mathbf{s})\}$ is the composition of the dual code to $\mathcal{A}$.

We first average $\boldsymbol{\mathcal{A}}$ over all equivalent codes. For a vector $\mathbf{u}$ of composition t,

$$\sum_{\pi \in S_n} x^{\mathbf{u}\pi} = \prod_{i=0}^{q-1} (t_i!) \sum_{\substack{\mathbf{v} \in F^n \\ \text{comp}(\mathbf{v}) = \mathbf{t}}} x^{\mathbf{v}}.$$

Set $d(\mathbf{t}) = \prod_{i=1}^{q-1} (t_i!)$. Then

$$\sum_{\pi \in S_n} \boldsymbol{\mathcal{A}}^{\pi} = \sum_{\mathbf{t}} d(\mathbf{t}) A(\mathbf{t}) X_{\mathbf{t}} . \tag{21}$$

Proof of Theorem 1:

From eq. (20),

$$|\mathcal{A}|\, B(\mathbf{s}) = \sum_{\substack{\mathbf{u} \in F^n \\ \text{comp}(\mathbf{u}) = \mathbf{s}}} \mathfrak{X}_{\mathbf{u}}(\boldsymbol{\mathcal{A}})$$

$$= \frac{1}{d(\mathbf{s})} \sum_{\pi \in S_n} \mathfrak{X}_{\mathbf{u}\pi}(\boldsymbol{\mathcal{A}})$$

$$= \frac{1}{d(\mathbf{s})} \mathfrak{X}_{\mathbf{u}}\Big(\sum_{\pi \in S_n} \boldsymbol{\mathcal{A}}^{\pi}\Big) \quad \text{by (4.1),}$$

[**u** is now the vector defined in Lemma (4.3)],

$$= \frac{1}{d(\mathbf{s})} \sum_{\mathbf{t}} d(\mathbf{t}) A(\mathbf{t}) \mathfrak{X}_{\mathbf{u}}(X_t) \qquad \text{by (21),}$$

$$= \sum_{t} \frac{d(\mathbf{t})}{d(\mathbf{s})} A(\mathbf{t}) P_{\mathbf{s}}(\mathbf{t}) \qquad \text{by (4.3),}$$

$$= \sum_{\mathbf{t}} P_{\mathbf{t}}(\mathbf{s}) A(\mathbf{t}) \qquad \text{by (4.4).}$$

Multiply both sides by $z_0^{s_0} \cdots z_{q-1}^{s_{q-1}}$ and sum over all compositions **s**.

Q.E.D.

4.5 *Proofs of Theorems 2 and 3.*

We use the notation of Sections 2.2 and 2.3.

Proof of Theorem 2:

In eq. (2) replace z_i by z_i for $1 \leqq i \leqq \delta$. Then using eq. (3), we see that eq. (2) collapses into eq. (4). Q.E.D.

Proof of Theorem 3:

In eq. (2) set $z_0 = 1$, $z_i = z$ for $i \neq 0$, and use eq. (5) to obtain (6).

Q.E.D.

V. DISCUSSION

We return to the binary case, which is easier to visualize.

The Hamming distance between vectors **u**, **v** is the weight of **u** + **v** (the weight of **u** − **v** if not binary). Coding theorists are interested in the distance structure of a code, not just in its weight structure. For linear codes, these are the same; they may also be the same for nonlinear codes, as in the example in Fig. 1. The following lemma is obvious.

Lemma 5.1: The distance and weight structure of a code $\mathcal{A}$ are the same if and only if the weight structure of $\mathcal{A} + \mathbf{v}$ is the same as that of $\mathcal{A}$ for all $\mathbf{v} \varepsilon \mathcal{A}$.

A code of this type will be said to have property 5.1. From now on we restrict ourselves to such codes.

A code with property 5.1 clearly contains the vector $\mathbf{0}$. The element of QG corresponding to $\mathcal{A} + \mathbf{v}$ is $\boldsymbol{\mathcal{A}}x^{\mathbf{v}}$.

Property 5.1 implies that

$$|\mathcal{A}|\, B(s) = \sum_{\mathbf{u}\, \varepsilon\, \sigma_s} \mathcal{X}_{\mathbf{u}}(\boldsymbol{\mathcal{A}}) = \sum_{\mathbf{u}\, \varepsilon\, \sigma_s} \mathcal{X}_{\mathbf{u}}(\boldsymbol{\mathcal{A}}x^{\mathbf{v}}) \quad \text{for} \quad \mathbf{v}\, \varepsilon\, \mathcal{A}.$$

Lemma 5.2: Property 5.1 implies that $B(s) \geqq 0$.

Proof: Take the sum over all $\mathbf{v}\, \varepsilon\, \mathcal{A}$ of the equation

$$\begin{aligned} |\mathcal{A}|\, B(s) &= \sum_{\mathbf{u}\, \varepsilon\, \sigma_s} \mathcal{X}_{\mathbf{u}}(\boldsymbol{\mathcal{A}}x^{\mathbf{v}}). \\ |\mathcal{A}|^2\, B(s) &= \sum_{\mathbf{u}\, \varepsilon\, \sigma_s} \mathcal{X}_{\mathbf{u}} \sum_{\mathbf{v}\, \varepsilon\, \mathcal{A}} (\boldsymbol{\mathcal{A}}x^{\mathbf{v}}) \\ &= \sum_{\mathbf{u}\, \varepsilon\, \sigma_s} \mathcal{X}_{\mathbf{u}}(\mathcal{A}) \sum_{\mathbf{v}\, \varepsilon\, \mathcal{A}} \mathcal{X}_{\mathbf{u}}(x^{\mathbf{v}}) \\ &= \sum_{\mathbf{u}\, \varepsilon\, \sigma_s} (\mathcal{X}_{\mathbf{u}}(\mathcal{A}))^2. \end{aligned}$$

Q.E.D.

Corollary 5.3: If $B(s) = 0$ then $\mathcal{X}_{\mathbf{u}}(\mathcal{A}) = 0$ for each $\mathbf{u}\, \varepsilon\, \sigma_s$.

Property 5.1 does not imply that $B(s)$ is an integer. Since by Theorem 5, $\sum_s B(s) = 2^n/|\mathcal{A}|$, $B(s)$ cannot all be integers unless $|\mathcal{A}| = 2^k$. For example, the code $\begin{pmatrix} 000 \\ 110 \\ 011 \end{pmatrix}$ has property 5.1, but the $B(s)$ are not all integers.

At present we have a satisfactory interpretation for $A(s)$, $B(s)$ if $\sum_{\pi\, \varepsilon\, S_n} \mathcal{A}^{\pi}$ can be generated by a linear code. ($\mathcal{A}$ need not be linear; any collection of vectors with the same weights as the vectors of a linear code will give the same average.) It would be very desirable to find an explanation for the cases in which $A(s)$, $B(s)$ can be thought of as the weight distribution of nonlinear codes.

VI. ACKNOWLEDGMENT

The authors would like to state that they found the basic idea of this paper, and much more, in J. H. van Lint's book *Coding Theory.*[6]

Added to galley proof:

Since this paper was written, it has come to our attention that Neal Zierler (unpublished) discovered the nonlinear MacWilliams identity for Hamming weight enumerators in 1966.

REFERENCES

1. Kerdock, A. M., "A Class of Low-Rate Nonlinear Codes," to appear in Info. and Control.
2. Preparata, F. P., "A Class of Optimum Nonlinear Double-Error Correcting Codes," Info. and Control, *13*, No. 4 (October 1968), pp. 378–400.
3. Berlekamp, E. R., *Algebraic Coding Theory*, New York: McGraw-Hill, 1968.
4. MacWilliams, F. J., "A Theorem on the Distribution of Weights in a Systematic Code," B.S.T.J., *42*, No. 1 (Janaury 1963), pp. 79–94.
5. Lloyd, S. P., "Binary Block Coding," B.S.T.J., *36*, No. 2 (March 1956), pp. 517–535.
6. van Lint, J. H., *Coding Theory*, New York: Springer-Verlag, 1971.
7. Assmus, E. F., Jr., "Research to Develop the Algrbraic Theory of Codes," Sylvania Electronic Systems, Waltham, Mass., Report AFCRL-67-0365, June 1967; especially Part V.
8. Gleason, A. M., "Weight Polynomials of Self-Dual Codes and the MacWilliams Identities," Actes, Congrès intern. Math., 1970, Vol. 3, pp. 211–215; Paris: Gauthier-Villars, 1971.
9. Lee, C. Y., "Some Properties of Non-binary Error-Correcting Codes," IEEE Trans. Info. Theory, *IT-4*, No. 2 (June 1958), pp. 77–82.
10. Krawtchouk, M., "Sur une généralisation des polynomes d'Hermite," Comptes Rendus, *189*, 1929, pp. 620–622.
11. Szegö, G., *Orthogonal Polynomials*, Colloquium Publications, Vol. 23, New York: American Mathematical Society, revised edition, 1959, pp. 35–37.