Week 8

The MacWilliams identity. The Average Weight Equation. Plotkin bound. Simplex codes

Version 2023-11-07. To accessible online version of this chapter

Synopsis. *Remarkably, the weights of codevectors of the dual code $C^\perp$ are completely determined by weights of codevectors of C. This was proved by* **Florence Jessie MacWilliams** *(1917–1990), an English-born American mathematician who spent most of her career at Bell Labs and Harvard in the United States. We state the general case of the MacWilliams identity. We give a proof (not examinable) for codes over $\mathbb{F}_p$ with prime p, and apply the identity to deduce a formula called the Average Weight Equation, as well as the Plotkin bound. We can use the MacWilliams identity to study Hamming codes by analysing their dual codes, called* **simplex codes.**

Theorem 8.1: the MacWilliams identity

If C is a q-ary linear code, $W_{C^\perp}(x,y) = \dfrac{1}{\#C} W_C(x+(q-1)y,\, x-y)$.

Proof for prime $q=p$. This proof is not examinable. Since p is a prime, the field $\mathbb{F}_p$ consists of elements $0, 1, \ldots, p-1$ (residues of integers modulo p). Being able to explicitly list the field elements — not possible for a general prime power q — simplifies the proof.

Let $C \subseteq \mathbb{F}_p^n$ be linear. We fix the complex number $\omega = e^{2\pi i/p}$, a primitive pth root of 1. We have $\omega^p = 1$ and $\omega, \omega^2, \ldots, \omega^{p-1} \neq 1$. We can write ω^a if $a \in \mathbb{F}_p$ — this complex number is well-defined, even though a is only defined modulo p.

Given $\underline{c} \in C$, $\underline{v} \in \mathbb{F}_p^n$, denote

$$\Phi(\underline{c}, \underline{v}) = \omega^{\underline{c}\cdot\underline{v}} x^{n-w(\underline{v})} y^{w(\underline{v})}.$$

We will compute $\sum_{\underline{c}\in C,\ \underline{v}\in\mathbb{F}_p^n} \Phi(\underline{c},\underline{v})$ in two different ways.

Way 1. If $\underline{v}\in C^\perp$, then $\underline{c}\cdot\underline{v}=0$ for all $\underline{c}\in C$, so $\Phi(\underline{c},\underline{v}) = x^{n-w(\underline{v})}y^{w(\underline{v})}$.

If, however, $\underline{v}\notin C^\perp$, there is a codevector $\underline{d}\in C$ such that $\underline{d}\cdot\underline{v}=a\neq 0$ in $\mathbb{F}_p$. Observe that $\Phi(\underline{d}+\underline{c},\underline{v}) = \omega^{\underline{d}\cdot\underline{v}}\Phi(\underline{c},\underline{v}) = \omega^a\Phi(\underline{c},\underline{v})$. We know that $\underline{d}+C=C$, so

$$\sum_{\underline{c}\in C}\Phi(\underline{c},\underline{v}) = \sum_{\underline{c}\in C}\Phi(\underline{d}+\underline{c},\underline{v}) = \omega^a\sum_{\underline{c}\in C}\Phi(\underline{c},\underline{v}) \quad\Longrightarrow\quad (\omega^a-1)\sum_{\underline{c}\in C}\Phi(\underline{c},\underline{v})=0.$$

Since $\omega^a\neq 1$, we have

$$\sum_{\underline{c}\in C}\Phi(\underline{c},\underline{v})=0 \quad \text{for } \underline{v}\notin C^\perp.$$

We conclude that

$$\sum_{\underline{c}\in C,\ \underline{v}\in\mathbb{F}_p^n}\Phi(\underline{c},\underline{v}) = \sum_{\underline{c}\in C,\ \underline{v}\in C^\perp}\Phi(\underline{c},\underline{v}) = \#C\sum_{\underline{v}\in C^\perp} x^{n-w(\underline{v})}y^{w(\underline{v})} = (\#C)W_{C^\perp}(x,y).$$

Way 2. If v is a symbol, $v\in\mathbb{F}_p$, we introduce the "weight of v", $w(v)$, as follows: $w(v)=1$ if $v\neq 0$ and $w(v)=0$ if $v=0$. Surely, for a vector $\underline{v}\in\mathbb{F}_p^n$ we have $w(\underline{v}) = w(v_1)+\cdots+w(v_n)$. We then rewrite

$$\begin{aligned}\Phi(\underline{c},\underline{v}) &= \omega^{c_1v_1+\cdots+c_nv_n}x^{1-w(v_1)}y^{w(v_1)}\ldots x^{1-w(v_n)}y^{w(v_n)}\\ &= \omega^{c_1v_1}x^{1-w(v_1)}y^{w(v_1)}\ldots\omega^{c_nv_n}x^{1-w(v_n)}y^{w(v_n)}.\end{aligned}$$

We now sum over $\underline{v}\in\mathbb{F}_p^n$ first: each coordinate of $\underline{v}$ runs over $\mathbb{F}_p=\{0,1,\ldots,p-1\}$. So, for a fixed $\underline{c}\in C$,

$$\begin{aligned}\sum_{\underline{v}\in\mathbb{F}_p^n}\Phi(\underline{c},\underline{v}) &= \sum_{v_1=0}^{p-1}\cdots\sum_{v_n=0}^{p-1}\Phi(\underline{c},\underline{v})\\ &= \sum_{v_1=0}^{p-1}\omega^{c_1v_1}x^{1-w(v_1)}y^{w(v_1)}\cdots\sum_{v_n=0}^{p-1}\omega^{c_nv_n}x^{1-w(v_n)}y^{w(v_n)}. \qquad (*)\end{aligned}$$

Let us analyse the first factor in the product on the right-hand side of (*):

$$\sum_{v_1=0}^{p-1}\omega^{c_1v_1}x^{1-w(v_1)}y^{w(v_1)} = x+\Big(\sum_{v_1=1}^{p-1}\omega^{c_1v_1}\Big)y.$$

If $c_1=0$, the coefficient of y is clearly $1+1+\cdots+1=p-1$, whereas if $c_1\neq 0$, the coefficient of y is the sum of a geometric progression

$$\sum_{v_1=1}^{p-1}\omega^{c_1v_1} = -1+\sum_{v_1=0}^{p-1}\omega^{c_1v_1} = -1+\frac{1-(\omega^{c_1})^p}{1-\omega^{c_1}} = -1+\frac{0}{1-\omega^{c_1}} = -1$$

since $(\omega^{c_1})^p = 1$. Hence the first factor on the right-hand side of (*) is

$$\begin{cases} x + (p-1)y, & \text{if } c_1 = 0, \\ x - y, & \text{if } c_1 \neq 0. \end{cases}$$

The same applies to the second, ..., nth factor in (*), hence (*) has $w(\underline{c})$ factors equal to $x - y$ and $n - w(\underline{c})$ factors equal to $x + (p-1)y$. In other words, (*) evaluates as $(x + (p-1)y)^{n-w(\underline{c})}(x-y)^{w(\underline{c})}$. Therefore,

$$\sum_{\underline{c} \in C} \sum_{\underline{v} \in \mathbb{F}_p^n} \Phi(\underline{c}, \underline{v}) = \sum_{\underline{c} \in C} (x + (p-1)y)^{n-w(\underline{c})}(x-y)^{w(\underline{c})} = W_C(x + (p-1)y,\, x - y).$$

Comparing *Way 2* and *Way 1*, we conclude that $W_C(x+(p-1)y,\, x-y) = (\#C)W_{C^\perp}(x, y)$. This is the MacWilliams identity for $q = p$. □

Simple examples where the MacWilliams identity is used

Let us obtain a short formula for the weight enumerator of the trivial code $\mathbb{F}_q^n$ by writing $\mathbb{F}_q^n$ as the dual code of the **null code** $Null = \{\underline{0}\}$. Of course, every vector in $\mathbb{F}_q^n$ is orthogonal to $\underline{0}$ which explains why $\mathbb{F}_q^n = Null^\perp$.

Clearly, $\#Null = 1$ and $W_{Null}(x, y) = x^n$ because N has only one codevector, which is of weight 0. Now use the MacWilliams identity:

Example: the weight enumerator of the trivial code $\mathbb{F}_q^n$

$W_{\mathbb{F}_q^n}(x, y) = \frac{1}{\#Null} W_{Null}(x + (q-1)y, x - y) = (x + (q-1)y)^n.$

We can obtain the same formula for the weight enumerator of the trivial code $\mathbb{F}_q^n$ without the use of MacWilliams identity, see earlier exercises.

The binary ($q = 2$) MacWilliams identity allows us to immediately obtain a short formula for the weight enumerator of the even weight code E_n. Indeed, $E_n = Rep(n, \mathbb{F}_2)^\perp$, and the binary repetition code has weight enumerator $W_{Rep(n,\mathbb{F}_2)}(x, y) = x^n + y^n$ (see example sheets). Also, $\#Rep(n, \mathbb{F}_2) = 2$. Hence

Example: the weight enumerator of E_n

$W_{E_n}(x, y) = \frac{1}{\#Rep(n,\mathbb{F}_2)} W_{Rep(n,\mathbb{F}_2)}(x + y, x - y) = \frac{1}{2}((x+y)^n + (x-y)^n).$

Using the binomial formula, we can expand this sum as $x^n + \binom{n}{2}x^{n-2}y^2 + \binom{n}{4}x^{n-4}y^4 + \ldots$ In particular, this proves that $w(E_n) = d(E_n) = 2$ as the lowest positive power of x in this polynomial is two.

The Average Weight Equation for linear codes

The proof of the following result involves a surprising use of the MacWilliams identity.

Theorem 8.2: the Average Weight Equation

If C is a q-ary linear code of length n, the average of the weights of all the codevectors of C is $(n-z)(1-q^{-1})$, where z is the number of zero columns in a generator matrix of C.

Proof. We count codevectors of weight 1 in the dual code $C^\perp$. By Theorem 5.1, $\underline{v} \in C^\perp$ iff $\underline{v}G^T = \underline{0}$ where G is a generator matrix of C. If $\underline{v}$ is of weight 1 with $v_i \neq 0$, then the ith column of G is zero. The non-zero v_i can be chosen in $q-1$ ways, so each zero column of G gives rise to $q-1$ vectors of weight 1 in $C^\perp$, and there are $z(q-1)$ such vectors in total.

We must get the same number as the coefficient of $x^{n-1}y$ in the weight enumerator $W_{C^\perp}(x,y)$, which by the MacWilliams identity equals

$$\frac{1}{\#C} W_C(x+(q-1)y, x-y) = \frac{1}{\#C} \sum_{\underline{v}\in C} (x+(q-1)y)^{n-w(\underline{v})}(x-y)^{w(\underline{v})}. \qquad (8.1)$$

We put $x=1$ and work out the coefficient of y. By the Binomial Theorem,

$$(1+(q-1)y)^{n-w(\underline{v})} = 1 + (n-w(\underline{v}))(q-1)y \qquad + \text{higher powers of } y,$$
$$(1-y)^{w(\underline{v})} = 1 - w(\underline{v})y \qquad + \text{higher powers of } y,$$

and so the coefficient of y in the product of these two expressions is

$$(n-w(\underline{v}))(q-1) - w(\underline{v}) = n(q-1) - qw(\underline{v}).$$

Summing over $\underline{v} \in C$ then dividing by $\#C$ gives the coefficient of y in (8.1) as $n(q-1) - q\frac{1}{\#C}\sum_{\underline{v}\in C} w(\underline{v})$. We thus get the equation

$$z(q-1) = n(q-1) - q\frac{1}{\#C}\sum_{\underline{v}\in C} w(\underline{v}),$$

hence the average of all weights, $\frac{1}{\#C}\sum_{\underline{v}\in C} w(\underline{v})$, is $(n-z)\frac{q-1}{q}$ as claimed. □

A simple example where we verify the Average Weight Equation

The easiest case where we can explicitly verify the Average Weight Equation is $C = Rep(n, \mathbb{F}_q)$, the q-ary repetition code of length n. The code consists of the zero vector and $q-1$ vectors of the form $aa\ldots a$ where $a \in \mathbb{F}_q \setminus \{0\}$, of weight n. The total number of codevectors is q. The one-row generator matrix $\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}$ of the code does not contain a zero column, so $z=0$. We arrive at the following

Example: average weight of a codevector of $Rep(n, \mathbb{F}_q)$

The average weight of a codevector of $Rep(n, \mathbb{F}_q)$ is

$$\frac{1 \times 0 + (q-1) \times n}{q} = n(1 - q^{-1}),$$

which agrees with the Average Weight Equation.

Exercise. Verify the Average Weight Equation by explicit calculation for the trivial code $\mathbb{F}_q^n$.

Simplex codes

What is the weight enumerator of $\mathrm{Ham}(r, q)$? This question can be answered using the MacWilliams identity. In the particular case $q = 2$, the answer can be explored further to give the probability P_{undetect} for the binary Hamming code (we do not pursue this here).

Recall from the previous chapter that the Hamming codes are defined via an interesting check matrix whose columns form *a maximal set of columns where no two columns are proportional*. What is the code *generated* by this matrix? We analyse these codes in the rest of this chapter.

Definition: simplex code

A **simplex code** $\Sigma(r, q)$ is defined as $\mathrm{Ham}(r, q)^{\perp}$.

Remark: recall that a *regular simplex* in an n-dimensional euclidean space $\mathbb{R}^n$ is a convex polytope whose vertices are $n+1$ points with the same distance between each pair of points. Thus, a 2-dimensional regular simplex is an equilateral triangle, and a 3-dimensional regular simplex is a regular tetrahedron. The following result motivates our terminology.

Theorem 8.3: properties of a simplex code

The simplex code $\Sigma(r, q)$ has length $n = (q^r - 1)/(q - 1)$ and dimension r. The Hamming distance between each pair of codevectors is q^{r-1}.

Proof. The length and dimension of $\Sigma(r, q) = \mathrm{Ham}(r, q)^{\perp}$ are dictated by the parameters of the Hamming code, see Theorem 7.3. It remains to calculate the distances.

Since $\Sigma(r, q)$ is linear, it suffices to show that every non-zero $\underline{v} \in \Sigma(r, q)$ has weight q^{r-1}.

By linear algebra, there is a basis of $\Sigma(r, q)$ which contains $\underline{v}$, hence $\underline{v}$ is the first row of some generator matrix H' of $\Sigma(r, q)$.

Since H' is a check matrix for $\mathrm{Ham}(r,q)$ and $d(\mathrm{Ham}(r,q)) = 3$, by Distance Theorem 7.2 no two columns of H' are proportional, hence the columns of H' represent distinct lines in $\mathbb{F}_q^r$. Therefore, the weight of $\underline{v}$ (the first row of H') is the number of lines where the *first* entry of a representative vector is not zero.

The total number of possible columns of size r with non-zero top entry is $(q-1)$ (choices for the top entry) $\times q^{r-1}$ (choices for the other entries which are unrestricted). But $(q-1)$ non-zero columns form a line, hence the number of required lines is $(q-1)q^{r-1}/(q-1) = q^{r-1}$. Hence $w(\underline{v}) = q^{r-1}$ as claimed. □

The weight enumerator of a binary Hamming code

By Theorem 8.3, the weight enumerator of the simplex code $\Sigma(r,q)$ is

$$W_{\Sigma(r,q)}(x,y) = x^n + (q^r-1)x^{n-q^{r-1}}y^{q^{r-1}}$$

where $n = \dfrac{q^r-1}{q-1}$. This formula reflects the fact that there is one codevector of weight 0 and q^r-1 codevectors of weight q^{r-1} in $\Sigma(r,q)$.

The weight enumerator of $\mathrm{Ham}(r,q) = \Sigma(r,q)^\perp$ can then be obtained using the MacWilliams identity. We do this for a binary Hamming code.

Proposition 8.4: the weight enumerator of $\mathrm{Ham}(r,2)$

$W_{\mathrm{Ham}(r,2)}(x,y) = \frac{1}{n+1}\left((x+y)^n + n(x+y)^{\frac{n-1}{2}}(x-y)^{\frac{n+1}{2}}\right)$ where $n = 2^r-1$.

Proof. The MacWilliams identity, Theorem 8.1, in the case of binary codes gives $W_{C^\perp}(x,y) = \frac{1}{\#C}W_C(x+y, x-y)$. We put $C = \Sigma(r,2)$ so that $C^\perp = \mathrm{Ham}(r,2)$. By Theorem 7.3, $n = 2^r-1$ so that $\#C = 2^r = n+1$ and the weight of each non-zero codevector in $\Sigma(r,2)$ is $q^{r-1} = 2^{r-1} = \frac{n+1}{2}$. We also have $n - q^{r-1} = n - \frac{n+1}{2} = \frac{n-1}{2}$.

Substituting these in the MacWilliams identity, we obtain $W_{\mathrm{Ham}(r,2)}$ as stated. □

Example: weight enumerator of the "original" Hamming code

$$W_{\mathrm{Ham}(3,2)} = \frac{1}{8}\left((x+y)^7 + 7(x+y)^3(x-y)^4\right) = x^7 + 7x^4y^3 + 7x^3y^4 + y^7.$$

Exercise: explicitly expand the left-hand side in the formula for $W_{\mathrm{Ham}(3,2)}$.

Exercise: Use Proposition 8.4 to show that every binary Hamming code contains the vector $111\ldots1$ (all bits equal to 1).

The Plotkin Bound

The Plotkin bound was obtained by Morris Plotkin in 1960 for arbitrary (not necessarily linear) binary codes. It applies to codes with very large minimum distance: $d > n/2$ where n is the length of the code. A proof of the general case of the bound by a direct counting argument can be found in the literature. We will only prove the statement for linear codes, which will serve as an example of the power of the MacWilliams identity and its corollary, the Average Weight Equation. (*Historical note*: the MacWilliams identity was proved in 1961, i.e., after the Plotkin bound.)

Proposition 8.5: The Plotkin bound for binary linear codes

If $C \subseteq \mathbb{F}_2^n$ is a linear code such that $d = d(C) > n/2$, then $\#C \leq \dfrac{d}{d - n/2}$.

Proof. Let $M = \#C$. The code C contains the zero vector, $\underline{0}$, and $M-1$ vectors of weight *at least* d. Then the average weight of a codevector of C is *at least*

$$\frac{1 \times 0 + (M-1) \times d}{M} = \big(1 - \frac{1}{M}\big)d.$$

So from the Average Weight Equation (where z is the number of zero columns in a generator matrix of C) we obtain

$$(n-z)\big(1 - \frac{1}{2}\big) \geq \big(1 - \frac{1}{M}\big)d \quad \Longrightarrow \quad \frac{n}{2} \geq \big(1 - \frac{1}{M}\big)d \quad \Longleftrightarrow \quad \frac{n}{2d} \geq 1 - \frac{1}{M}$$

so that $1/M \geq 1 - n/(2d) = (2d-n)/(2d)$ and $M \leq 2d/(2d-n)$, as claimed. □