Weight Distributions of Codes (Peterson + Weldon, pp. 64-; Berlekamp, pp. 397 -)

Let A_i = number of codewords of weight i in an (N, K) block code. The weight enumerator for the code is

$$A(z) = \sum_{i=0}^{N} A_i z^i$$

The weight distribution can be used to calculate the undetected error probability. Any error pattern equal to a nonzero codeword causes an undetected error. For a code over GF(2) used on a BSC with crossover probability p

$$P(\text{undetected error}) = \sum_{i=1}^{N} A_i \, p^i (1-p)^{N-i}$$

$$= (1-p)^N A[p/(1-p)] - (1-p)^N$$

The weight distribution $\{A_i\}$ has been found analytically for only a few codes, for example, the maximal length codes, Hamming codes, and Reed-Solomon codes. If K is not too large, $A(z)$ can be determined by examining all codewords using a computer. If K is large, $N-K$ may be small. We will derive a formula relating the weight distributions of an (N, K) code and its $(N, N-K)$ dual code in this section.

Definitions

Let U and V be subspaces of W.

Let $U \cap V$ = set of vectors in both U and V. Then $U \cap V$ is also a subspace of W.

Let $U \oplus V$ be the set of vectors of the form $u+v$ with $u \in U$ and $v \in V$. $U \oplus V$ is a subspace of W

Lemma 1

The sum of the dimensions of $U \cap V$ and $U \oplus V$ is the sum of the dimensions of U and V.

Proof:

Let $k_1 = \dim U$, $k_2 = \dim V$, and $k_0 = \dim U \cap V$. There is a basis of k_0 vectors for $U \cap V$. It is possible to find a basis for U consisting of these k_0 vectors and $k_1 - k_0$ others not in $U \cap V$, and a basis for V consisting of these k_0 and $k_2 - k_0$ others not in $U \cap V$. Then the k_0 vectors in the basis for $U \cap V$, the $k_1 - k_0$ additional vectors in the basis for U, and the $k_2 - k_0$ additional vectors in the basis for V taken together form a basis for $U \oplus V$. Thus $\dim U \oplus V = k_0 + (k_1 - k_0) + (k_2 - k_0)$
$= k_1 + k_2 - k_0$

Q.E.D.

Lemma 2

Let $U^\perp$ be the null space of U and $V^\perp$ be the null space of V. Then $(U \oplus V)^\perp = U^\perp \cap V^\perp$

Proof:

Since $U \subset U \oplus V$, every vector in $(U \oplus V)^\perp$

must be in $U^\perp$, i.e. $(U \oplus V)^\perp \subset U^\perp$. Similarly, $V \subset U \oplus V$, so $(U \oplus V)^\perp \subset V^\perp$. Therefore $(U \oplus V)^\perp \subset U^\perp \cap V^\perp$. Every vector in $U \oplus V$ has the form $u + v$ with $u \in U$ and $v \in V$. If w is in $U^\perp \cap V^\perp$, then $u \cdot w = v \cdot w = 0$ so $(u+v) \cdot w = 0$. Therefore $U^\perp \cap V^\perp \subset (U \oplus V)^\perp$.

Q.E.D.

<u>Theorem</u> (MacWilliams, 1963)

Let V be an (N, K) block code and $V^\perp$ be the $(N, N-K)$ dual code. Let A_i and B_i denote the number of vectors of weight i in V and $V^\perp$, respectively. Then, for codes over $GF(q)$,

$$\sum_{i=0}^{N} B_i z^i = q^{-K} \sum_{i=0}^{N} A_i (1-z)^i [1 + (q-1) z]^{N-i} \qquad (1)$$

or

$$B(z) = q^{-K} [1 + (q-1) z]^N A[(1-z)/(1+(q-1)z)] \qquad (1')$$

Proof:

Let $s = (s_1, s_2, \ldots, s_m)$ be a set of m distinct integers from the set $n = (1, \ldots, N)$ and let $t = (t_1, \ldots, t_{N-m}) = \bar{s}$. Let F_s be the subspace of all vectors whose components in positions $s_1, \ldots, s_m$ may be nonzero but must be zero in $t_1, \ldots, t_{N-m}$. Define F_t similarly. Then $F_t = F_s^\perp$.

Now consider the subspace of all vectors in V which are zero in positions $t_1, \ldots, t_{N-m}$. This subspace

is $V \cap F_s$. Similarly $V^\perp \cap F_t$ is the set of all vectors in $V^\perp$ with zeros in positions $s_1, \ldots, s_m$. By lemma 2, $(V \cap F_s)^\perp = V^\perp \oplus F_s^\perp = V^\perp \oplus F_t$. Let $d_s = \dim V \cap F_s$ and $d_t = \dim V^\perp \cap F_t$, then $\dim V^\perp \oplus F_t = N - d_s$ because it is the null space of $V \cap F_s$. Also, according to Lemma 1, $\dim V^\perp \oplus F_t = (N-K) + (N-m) - d_t$. Thus

$$N - d_s = (N-K) + (N-m) - d_t$$

or

$$d_t = d_s + N - K - m$$

Now consider pairs consisting of a set s of m integers and a vector $v \in V \cap F_s$. For each s there are q^{d_s} such pairs since $\dim V \cap F_s = d_s$. Considering all choices of s, the total number of such pairs is

$$\sum_{\text{all } s} q^{d_s}$$

On the other hand, each vector of weight j in V has $N-j$ zero components, and any set t which is a subset of $N-m$ of the indices of these positions defines a set s which can be paired with this vector. There are $\binom{N-j}{N-m}$ choices with $N-m$ integers in t or m in s. There are A_j vectors of weight j in V. Therefore, the number of pairs (v, s) is also

$$\sum_{\text{all } s} q^{d_s} = \sum_{j=0}^{N} A_j \binom{N-j}{N-m}$$

Similarly, considering $V^\perp$ and sets t of $N-m$ integers yields

$$\sum_{\text{all } t} q^{d_t} = \sum_{i=0}^{N} B_i \binom{N-i}{m}$$

Since $d_t = d_s + N - K - m$ and each t determines a unique s

$$\sum_t q^{d_t} = \sum_t q^{d_s + N - K - m} = \sum_s q^{d_s - N - K - m}$$

$$= q^{N-K-m} \sum_s q^{d_s}$$

or

$$\sum_{i=0}^{N} B_i \binom{N-i}{m} = q^{N-K-m} \sum_{j=0}^{N} A_j \binom{N-j}{N-m} \qquad (2)$$

Define $\binom{n}{k} = 0$ for $k > n$. Multiplying (2) by y^m and summing yields

$$\sum_{m=0}^{N} \sum_{i=0}^{N} B_i \binom{N-i}{m} y^m = q^{-K} \sum_{m=0}^{N} q^{N-m} \sum_{j=0}^{N} A_j \binom{N-j}{N-m} y^m$$

$$\sum_{i=0}^{N} B_i \sum_{m=0}^{N} y^m \binom{N-i}{m} = q^{-K} \sum_{j=0}^{N} A_j \sum_{m=0}^{N} \binom{N-j}{N-m} q^{N-m} y^m$$

(let $N-m \to m$)

$$\sum_{i=0}^{N} B_i (1+y)^{N-i} = q^{-K} \sum_{j=0}^{N} A_j \sum_{m=0}^{N} \binom{N-j}{m} q^m y^{N-m-j} y^j$$

$$\sum_{i=0}^{N} B_i (1+y)^{N-i} = q^{-K} \sum_{j=0}^{N} A_j y^j (q+y)^{N-j}$$

Now let $z = (1+y)^{-1}$, then $y = (1-z)/z$

$$\sum_{i=0}^{N} B_i z^i = q^{-K} \sum_{j=0}^{N} A_j z^N \frac{(1-z)^j}{z^j} \left(q + \frac{1-z}{z}\right)^{N-j}$$

$$= q^{-K} \sum_{j=0}^{N} A_j (1-z)^j [1 + (q-1)z]^{N-j} \qquad (3)$$

or

$$B(z) = q^{-K}\,[1+(q-1)z]^N\, A\left[\frac{1-z}{1+(q-1)z}\right]$$

Q. E. D.

An explicit formula for B_i can be determined from (3) as follows:

$$\sum_{i=0}^{N} B_i z^i = q^{-K} \sum_{j=0}^{N} A_j \left(\sum_{s=0}^{j} \binom{j}{s} (-1)^s z^s\right)\left(\sum_{t=0}^{N-j} \binom{N-j}{t} (q-1)^t z^t\right).$$

Letting $s+t=i$, and eliminating t gives

$$B(z) = q^{-K} \sum_{j=0}^{N} \sum_{s=0}^{N} \sum_{i=0}^{N} A_j \binom{j}{s}\binom{N-j}{i-s}(-1)^s (q-1)^{i-s} z^i$$

$$= q^{-K} \sum_i z^i \sum_j A_j \sum_s \binom{j}{s}\binom{N-j}{i-s}(-1)^s (q-1)^{i-s}$$

Equating coefficients of z^i yields

$$B_i = q^{-K} \sum_{j=0}^{N} A_j \sum_{s=0}^{N} \binom{j}{s}\binom{N-j}{i-s}(-1)^s (q-1)^{i-s} \qquad (4)$$

The sum of the kth powers of the weights can be found by using the identity:

$$\left(z\frac{d}{dz}\right)^k \sum_{i=0}^{N} B_i z^i = \sum_{i=0}^{N} i^k B_i z^i$$

and letting ~~with~~ $z=1$. Due to the factor $(1-z)$ on the right of (1), it follows that all the terms involving A_i for $i>k$ disappear. The results for $k=0,1,2$ are

$$\sum_{i=0}^{N} B_i = q^{N-K}$$

$$\sum_{i=1}^{N} i B_i = q^{N-K-1} [N(q-1) - A_1]$$

$$\sum_{i=1}^{N} i^2 B_i = q^{N-K-2} \{ N(q-1)(Nq-N+1) - A_1 [q + 2(N-1)(q-1)] + 2A_2 \}$$

<u>Example</u>

Consider the Hamming code with $N = 2^m - 1$ and $N - K = m$. Its dual is the $(2^m - 1, m)$ maximal length code over $GF(2)$. The weight distribution for the maximal length code is $A_0 = 1$, $A_{2^{m-1}} = 2^m - 1$, $A_i = 0$ otherwise. Thus $A(z) = 1 + (2^m - 1) z^{2^{m-1}}$ and

$$B(z) = 2^{-3} (1+z)^{2^m - 1} \left[1 + (2^m - 1) \left(\frac{1-z}{1+z} \right)^{2^{m-1}} \right]$$

Suppose $m = 3$ to give the $(7,4)$ Hamming code. Then

$$B(z) = 2^{-3} (1+z)^7 \left[1 + 7 \left(\frac{1-z}{1+z} \right)^4 \right]$$

$$= 2^{-3} \left[(1+z)^7 + 7 (1+z)^3 (1-z)^4 \right]$$

$$= 1 + 0z + 0z^2 + 7z^3 + 7z^4 + 0z^5 + 0z^6 + z^7$$