THEOREM OF THE DAY

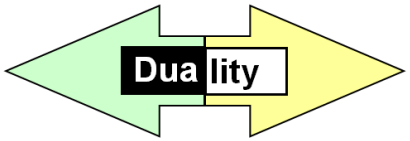
MacWilliams' Identity *Let C be a linear code of length n over the finite field of q elements and let $C^\perp$ be the dual code of C. With $A_w(X)$ denoting the number of words of weight w in a code X, let $W_C(x,y) = \sum_{w=0}^{n} A_w(C)x^w y^{n-w}$ and $W_{C^\perp}(x,y) = \sum_{w=0}^{n} A_w(C^\perp)x^w y^{n-w}$ be the homogeneous weight enumerators of C and $C^\perp$, respectively. Then*

$$W_{C^\perp}(x,y) = |C|^{-1} W_C(y-x, y+(q-1)x)$$

C

All weighted sums, modulo 3, of *a* and *b*.

Name	codeword					weight
zero	0	0	0	0	0	0
a	1	0	1	1	0	3
b	0	1	2	0	2	3
2a	2	0	2	2	0	3
2b	0	2	1	0	1	3
a+b	1	1	0	1	2	4
2a+b	2	1	1	2	2	5
a+2b	1	2	2	1	1	5
2a+2b	2	2	0	2	1	4



***C*$^\perp$, *on* the right, consists of all (0,1,2)-vectors of length 5 having zero product, modulo 3, with everything in *C*, on the left. For example:**

$b \times h = 0\,1\,2\,0\,2 \times 0\,0\,1\,2\,2$
$= 0.0+1.0+2.1+0.2+2.2$
$= 0 + 0 + 2 + 0 + 1$
$= 2 + 1 = 0 \pmod 3$

Name	codeword					weight
zero	0	0	0	0	0	0
f	1	0	0	2	0	2
g	0	1	0	0	1	2
h	0	0	1	2	2	3
2f	2	0	0	1	0	2
2g	0	2	0	0	2	2
2h	0	0	2	1	1	3
f+g	1	1	0	2	1	4
f+h	1	0	1	1	2	4
:						:
f+2g+h	1	2	1	1	1	5
:						:
2f+2g+2h	2	2	2	2	0	4

$C^\perp$

All weighted sums, modulo 3, of *f*, *g* and *h*.

The weight of a codeword is the number of places in which it differs from zero. The distribution of weights for a linear error-correcting code determines the probability of incorrectly decoding a transmitted codeword. For suppose the probability of a single symbol change, during transmission, is p. Without loss of generality, consider incorrect decoding for the zero codeword: due to errors we receive a non-zero codeword c of weight w. For this to occur, w symbols must change, each one choosing, from $q-1$ nonzero symbols, the appropriate entry in c. Meanwhile $n-w$ symbols are unchanged. The resulting probability is $p^w \times (1/(q-1))^w \times (1-p)^{n-w}$. Over all possible c, the probability of incorrect decoding enumerates to $W_C(p/(q-1), 1-p) - (1-p)^n$ (the last term accounting for c actually being the zero codeword).

For the code C above, we have $W(C) = y^5 + 4x^3y^2 + 2x^4y + 2x^5$. Suppose that p is, say, 0.01. Substituting $x = p/2$ and $y = 1-p$, we find that the probability of incorrect decoding, as specified above, is about one in two million. Now $W(C^\perp) = (1/9)W_C(y-x, y+2x) = 1/9\,(y+2x)^5 + 4/9\,(y+2x)^2\,(y-x)^3 + 2/9\,(y+2x)\,(y-x)^4 + 2/9\,(y-x)^5 = y^5 + 4x^2y^3 + 8x^3y^2 + 12x^4y + 2x^5$. And we learn, without knowing a single nonzero codeword of $C^\perp$, that $C^\perp$ has $2 + 12 + 8 + 4 + 1 = 27$ codewords and that the probability of incorrect decoding compares very unfavourably at about one in ten thousand.

In 1961, Jessie MacWilliams took leave from her job as a programmer at Bell Labs to complete (in one year!) a PhD at Harvard. This remarkable theorem, indispensable in the study of error-correcting codes, appeared in her dissertation.

Web link: www.math.msu.edu/~jhall/classes/codenotes/coding-notes.html (Chapter 9).

Further reading: *The Theory of Error-Correcting Codes* by F.J. MacWilliams and N.J.A. Sloane, North Holland, 1983.