Math 75 notes, Lecture 16 outline

P. Pollack and C. Pomerance

References below are to Pretzel's *Error-correcting codes and finite fields*:

- We reviewed the connection between a generator matrix for a code and a check matrix. In particular, we did this for the standard generator and check matrices for the $(6, 3)$ triple check code over $\mathbb{F}_2$.

- We multiplied this check matrix by the 0-vector and the 6 possible weight 1 vectors, getting 7 of the 8 possible vectors of length 3. We found an eighth vector giving rise to the 8th length-3 vector, namely $(1, 0, 0, 0, 0, 1)$ checks to $(1, 1, 1)$.

- These different vectors of length 3 are called syndromes, and we saw that if the word w has syndrome s, then the set of words having the exact same syndrome s is $C+w$, namely the equivalence class (coset) containing w.

- If we take as coset representatives (called leaders) words of minimal weight, we thus have a mechanism for error correction. For example, if $(1, 1, 0, 0, 0, 0)$ is the received word, we can multiply it by H to see if it is a code word. Well no, it isn't, the product is the syndrome $(0, 1, 1)$, which is not the 0-vector, so w is not a code word. But the weight 1 vector $(0, 0, 1, 0, 0, 0)$ has the same syndrome, so it is reasonable to suspect that this is the error pattern. That is, we should subtract (same as add in characteristic 2) $(0, 0, 1, 0, 0, 0)$ from the received word to get $(1, 1, 1, 0, 0, 0)$ to get the likely code word that was sent (which then decodes to real word $(1, 1, 1)$, since we are dealing with standard matrices).

- We noticed that if e_i is the ith standard basis vector in F^n and c_i is a scalar (element of F), then $H(c_i e_i)^T$ is just c_i times the ith column of the check matrix H. And so if $w = \sum c_i e_i$ is a linear combination of the standard basis vectors, then Hw^T is exactly $\sum c_i H_i$, where H_i is the ith column of H. We used this to prove the following theorem, which is stated a little differently in the book (see p. 59).

  Theorem 1. *For a check matrix H of the linear code C, let d_H be the minimal size of a set of linearly dependent columns of H. Then $d_H = d(C)$.*

  This has the corollary that if over $\mathbb{F}_2$ the matrix H has no zero column and the columns are all different, then $d_H \geq 3$, so therefore $d(C) \geq 3$, and therefore the code can correct at least 1 error.

- We introduced Ham_k, the binary Hamming code with parameter k. The check matrix H_k is just a listing of all the nonzero vectors of length k, so is a $k \times (2^k - 1)$ matrix. The corresponding code has length $2^k - 1$ and dimension $2^k - k - 1$. For example, when $k = 3$, we get a $(7, 4)$ code. It has minimal distance 3, so can correct 1 error. Note that it is denser (more efficient) than the $(6, 3)$ triple check code, since its density is $4/7$ in comparison to $3/6 = 1/2$ for the triple check code.