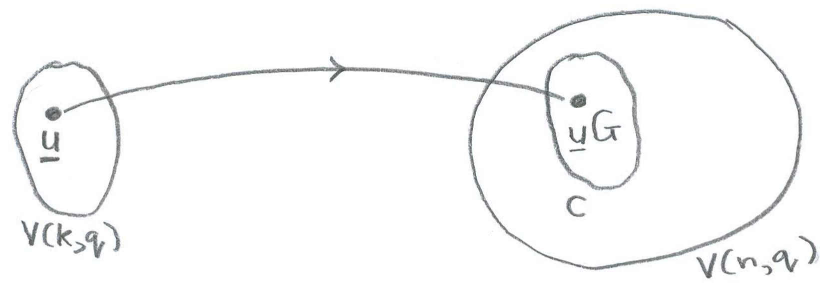
Encoding with a linear code

Let C be a q-ary $[n,k]$-code.

Since $M = q^k$, we can communicate q^k messages — identify these with the q^k vectors of $V(k,q)$. If C has generator matrix G, then we encode these q^k message vectors with the map

$$V(k,q) \to V(n,q) : \quad \underline{u} \mapsto \underline{u}G$$



Example Consider the binary $[7,4]$-code C

with generator matrix $G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$

The message vectors are the $2^4 = 16$ vectors of $V(4,2)$.

A message vector $u_1 u_2 u_3 u_4$ is encoded by

$$[u_1\ u_2\ u_3\ u_4]\begin{bmatrix} 1&0&0&0&1&0&1\\ 0&1&0&0&1&1&1\\ 0&0&1&0&1&1&0\\ 0&0&0&1&0&1&1 \end{bmatrix}$$

$$= [\underbrace{u_1\ \ u_2\ \ u_3\ \ u_4}_{\text{message digits}}\ \ \underbrace{u_1+u_2+u_3\ \ \ u_2+u_3+u_4\ \ \ u_1+u_2+u_4}_{\text{check digits}}]$$

(added redundancy for protection against noise)

The $2^4 = 16$ codewords are obtained by encoding the $2^4 = 16$ message vectors; for example, the message vector 1011 is encoded to the codeword 1011000.

Cosets

Let C be a q-ary $[n,k]$-code.

For any $\underline{a} \in V(n,q)$, the set

$$\underline{a} + C = \{\underline{a} + \underline{x} : \underline{x} \in C\}$$

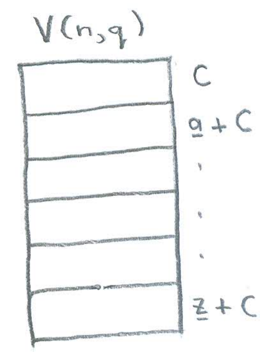
is called a coset of C.

Remember: $V(n,q)$ is a group under vector addition and C is a subgroup of $V(n,q)$.

Lemma 3.3 If $\underline{a}+C$ is a coset of C and $\underline{b} \in \underline{a}+C$, then $\underline{b}+C = \underline{a}+C$.

Theorem 3.4 Lagrange's Theorem

Suppose C is a q-ary $[n,k]$-code. Then



1. Every vector of $V(n,q)$ belongs to some coset of C.
2. Every coset contains q^k vectors.
3. Any two cosets are disjoint.

Example: Let C be the binary $[4,2]$-code with generator matrix $G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$.

Then $C = \{0000, 1011, 0101, 1110\}$ and the cosets of C are

$$\begin{aligned}
0000 + C &= C \\
1000 + C &= \{1000, 0011, 1101, 0110\} \\
0100 + C &= \{0100, 1111, 0001, 1010\} \\
0010 + C &= \{0010, 1001, 0111, 1100\}
\end{aligned}$$

Note that, as representative for each coset, we have chosen a vector of minimum weight in the coset.