Definition 0.1. The ***inner product*** $\mathbf{u} \cdot \mathbf{v}$ of vectors $\mathbf{u} = u_1 u_2 \cdots u_n$ and $\mathbf{v} = v_1 v_2 \cdots v_n$ in $V(n, q)$ is scalar defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots u_n v_n.$$

If $\mathbf{u} \cdot \mathbf{v} = 0$, then $\mathbf{u}$ and $\mathbf{v}$ are called ***orthogonal***.

Lemma 0.2. For any $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in Vn, q and $\lambda,\ \mu \in GF(q)$,

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\lambda \mathbf{u} + \mu \mathbf{v}) \cdot \mathbf{w} = \lambda(\mathbf{u} \cdot \mathbf{w}) + \mu(\mathbf{v} \cdot \mathbf{w})$

Definition 0.3. Given a linear $[n, k]$-code C, the ***dual code*** of C, denoted by $C^\perp$ is defined as

$$C^\perp = \{\mathbf{v} \in V(n, q) \mid \mathbf{v} \cdot \mathbf{u} = 0 \ \forall\ \mathbf{u} \in C\}.$$

Lemma 0.4. Suppose C is an $[n, k]$-code having a generator matrix G, then $\mathbf{v} \in C^\perp \iff \mathbf{v}G^T = 0$, where G^T is transpose of G.

Proof. $\implies$: This part is obvious as rows of G are codewords.
$\impliedby$: Suppose the rows of G are $\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_k$, and thus $\mathbf{v} \cdot \mathbf{r}_i = 0$ for all i. If $u \in C$, then $u = \sum_{i=1}^{k} \lambda_i \mathbf{r}_i$ for some scalars λ_i, and

$$\begin{aligned} \mathbf{v} \cdot \mathbf{u} &= \sum_{i=1}^{k} \lambda_i (\mathbf{v} \cdot \mathbf{r}_i) \quad \text{(by Lemma 0.2)} \\ &= \sum_{i=1}^{k} \lambda_i 0 = 0. \end{aligned}$$

Hence, $\mathbf{v}$ is orthogonal to all codewords in C. □

Theorem 0.5. Suppose C is an $[n, k]$-code over $GF(q)$. Then the dual code $C^\perp$ is a linear $[n, n - k]$-code.

Proof. Suppose $\mathbf{v}_1, \mathbf{v}_2 \in C^\perp$ and $a \in GF(q)$. Then, for all $\mathbf{u} \in C$,

$$\begin{aligned} (\mathbf{v}_1 + \mathbf{v}_2) \cdot \mathbf{u} &= \mathbf{v}_1 \cdot \mathbf{u} + \mathbf{v}_2 \cdot \mathbf{u} \\ &= 0 \\ (a\mathbf{v}_1) \cdot (u) &= a(\mathbf{v}_1 \cdot \mathbf{u}) \\ &= 0 \end{aligned}$$

Hence, $C^\perp$ is a linear code.
For the dimension part, notice that if two codes C_1 and C_2 are equivalent, then so are $C_1^\perp$ and $C_2^\perp$. Hence it will be enough to show $\dim C^\perp = n - k$ in the case when C has a standard form of generator matrix

$$G = \begin{bmatrix} 1 & \cdots & 0 & a_{11} & \cdots & a_{1,n-k} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & a_{k1} & \cdots & a_{k,n-k} \end{bmatrix}$$

Then

$$C^\perp = \left\{ (v_1, v_2, \ldots, v_n) \in V(n, q) \mid v_i + \sum_{j=1}^{n-k} v_{j+k} a_{ij} = 0,\ \forall\ i \in \{1, 2, \ldots, n-k\} \right\}$$

We have q^{n-k} choices for $(v_{k+1}, \ldots, v_n)$, and for each combination be have unique vector $v_1 v_2 \cdots v_n$ in $C^\perp$. Hence, $|C^\perp| = q^{n-k}$, and $\dim C^\perp = n - k$. □

Theorem 0.6. For any $[n, k]$-code C, $(C^\perp)^\perp = C$.

Proof. $C \subseteq (C^\perp)^\perp$ since every vector in C is orthogonal to every vector in $C^\perp$. But $\dim (C^\perp)^\perp = n - (n - k) = k = \dim C$. Therefore, $C = (C^\perp)^\perp$. □

Definition 0.7. A ***parity-check matrix*** H for an $[n,k]$-code C is a generator matrix of $C^{\perp}$.

Thus, H is $(n-k)\times n$ matrix satisfying $GH^T = \mathbf{0}$, where $\mathbf{0}$ is the all-zero matrix. It follows from Lemma 0.4 and Theorem 0.6, that C can be written as

$$C = \{\mathbf{x} \in V(n,q) \mid \mathbf{x}H^T = \mathbf{0}\}.$$

Theorem 0.8. If $G = [I_k \mid A]$ is the standard form of generator matrix of an $[n,k]$-code C, then a parity-check matrix for C is $H = [-A^T \mid I_{n-k}]$.

Proof. Suppose

$$G = \left[\begin{array}{ccc|ccc} 1 & \cdots & 0 & a_{11} & \cdots & a_{1,n-k} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & a_{k1} & \cdots & a_{k,n-k} \end{array}\right]$$

Let

$$H = \left[\begin{array}{ccc|ccc} -a_{11} & \cdots & -a_{k1} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ -a_{1,n-k} & \cdots & -a_{k,n-k} & 0 & \cdots & 1 \end{array}\right]$$

Then H has the required size of a parity-matrix and its rows are linearly independent (as coefficients of 1 in the identitiy part will remain). Also, $\mathbf{g}_i \cdot \mathbf{h}_j$ (where $\mathbf{g}_i$ and $\mathbf{h}_j$ are some rows of G and H respectively) is

$$0 + \cdots + 0 + (-a_{ij}) + 0 + \cdots + 0 + a_{ij} + 0 + \cdots + 0 = 0$$

□

Remark. Minus signs are unnecessary in the binary case.

Definition 0.9. A parity-check matrix is called to be in ***standard form*** if $H = [B \mid I_{n-k}]$.

So from Theorem 0.8, we can get standard form of generator matrix from standard form of parity-check matrix and vice-versa.

Now we will see some more efficient decoding schemes using *parity-check matrices*, but before that some definitions and lemmas.

Definition 0.10. Suppose H is a parity-check matrix of an $[n,k]$-code C. Then for any vector $y \in V(n,q)$, the row vector

$$S(\mathbf{y}) = \mathbf{y}H^T$$

is called the ***syndrome*** of $\mathbf{y}$.

It is quite that $S(\mathbf{y}) = 0$ if and only if $\mathbf{y} \in C$.

Lemma 0.11. Two vectors $\mathbf{u}$ and $\mathbf{v}$ are in the same coset of C if and only if they have the same syndrome.

Proof. $\mathbf{u}$ and $\mathbf{v}$ are in the same coset

$$\begin{aligned} &\iff \mathbf{u} + C = \mathbf{v} + C \\ &\iff \mathbf{u} - \mathbf{v} \in C \\ &\iff (\mathbf{u} - \mathbf{v})H^T = \mathbf{0} \\ &\iff \mathbf{u}H^T = \mathbf{v}H^T \\ &\iff S(\mathbf{u}) = S(\mathbf{v}) \end{aligned}$$

□

Corollary 0.11.1. There is a one-to-one correspondance between cosets and syndromes.

In standard array decoding, if n is large, finding codewords in the array becomes increasingly inefficient. The following *syndrome decoding scheme* is much more efficient.

Syndrome Decoding Scheme: Instead of storing all the cosets in the standard array, if we store the coset leaders along with their corresponding coset syndromes, will be sufficient for achieving the same probability of error detection. When a vector $\mathbf{y}$ is received, we calculate $S(\mathbf{y}) = \mathbf{y}H^T$ and locate $\mathbf{z} = S(\mathbf{y})$ in the syndromes column, find the corresponding coset leader $f(\mathbf{z})$, and decode $\mathbf{y}$ as $\mathbf{x} = \mathbf{y} - f(\mathbf{z})$. This works because of Corollary 0.11.1.

For example, let C be a 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,

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

The syndrome look-up table of C will be

syndrome $\mathbf{z}$	coset leaders f($\mathbf{z}$)
00	0000
11	1000
01	0100
10	0010

Incomplete Decoding Scheme: In this we mix error correction as well detection. If $d(C) = 2t + 1$ or $2t + 1$, we can precisely correct $\leq t$ errors, using syndrome lookup table as all the vectors with weight $\leq t$ will be coset leaders. Otherwise, we will simply seek *re-transmission*. So now, we now need to store even lesser data in the lookup table.