Definition 1. Linear Code
An (n, k) *linear code* over a finite field F is a k-dimensional subspace V of the vector space

$$F^n = \underbrace{F \oplus F \oplus \cdots \oplus F}_{n \text{ copies}}$$

over F. The members of V are called the *code words.* The ratio k/n is called the *information rate* of the code. When F is $\mathbb{Z}_2$, the code is called binary.

Example 1. The Hamming (7,4) Code
Assuming that our message consists of all possible 4-tuples of 0's and 1's (i.e., we wish to send a sequence of 0's and 1's of length 4). Encoding will be done by viewing these messages as four-dimensional vectors over the field $\mathbb{Z}_2$ and multiplying each of the 16 possible messages on the right by the matrix

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

The resulting seven-dimensional vectors are called *code words.* See Table 1.

Message	Encoder G	Code Word
0000	$\rightarrow$	0000000
0001	$\rightarrow$	0001111
0010	$\rightarrow$	0010110
0100	$\rightarrow$	0100101
1000	$\rightarrow$	1000011
1100	$\rightarrow$	1100110
1010	$\rightarrow$	1010101
1001	$\rightarrow$	1001100
0110	$\rightarrow$	0110011
0101	$\rightarrow$	0101010
0011	$\rightarrow$	0011001
1110	$\rightarrow$	1110000
1101	$\rightarrow$	1101001
1011	$\rightarrow$	1011010
0111	$\rightarrow$	0111100
1111	$\rightarrow$	1111111

Table 1

Definition 2. Hamming Distance, Hamming Weight
The *Hamming distance* between two vectors of a vector space is the number of components in which they differ. The *Hamming weight* of a vector is the number of nonzero components of the vector.
We will use $d(u, v)$ to denote the Hamming distance between the vectors u and v and $\text{wt}(u)$ for the Hamming weight of the vector u.

Definition 3. Nearest-Neighbor Decoding
For any received vector v, the corresponding code word sent is a code word v' such that the Hamming distance $d(v, v')$ is a minimum. If there is more than one such v', we decide arbitrarily.

Theorem 1. *Properties of Hamming Distance and Hamming Weight*
For any vectors u, v and w of a linear code, $d(u, v) \leq d(u, w) + d(w, v)$ and $d(u, v) = wt(u - v)$.

Theorem 2. *Correcting Capability of a Linear Code*
If the Hamming weight of every nonzero code word in a linear code is at least $2t+1$, then the code can correct for any t or fewer errors. Furthermore, the same code can detect any $2t$ or fewer errors..

Proof. Using the nearest-neighbor decoding, suppose a transmitted code word u is received as the vector v and at most t errors were made in transmission. Then, by definition, $d(v, u) \leq t$. If w is any code word other than u, then $w - u$ is a nonzero code word. Thus, by assumption,

$$2t + 1 \leq \mathrm{wt}(w - u) = d(w, u) \leq d(w, v) + d(v, u) \leq d(w, v) + t$$

and it follows that $t + 1 \leq d(w, v)$. So the code word closest to the received vector v is u and, therefore v is correctly decoded as u.
To show that the code can detect $2t$ errors, we suppose a transmitted code word u is received as the vector v and at least one error, but no more than $2t$ errors, was made in transmission. Because only code words are transmitted, an error will be detected whenever a received word is not a code word. But, u cannot be a code word, since $d(v, u) \leq 2t$, while we know that the minimum distance between distinct code words is at least $2t + 1$. $\square$

Theorem 3. *Singleton Bound*
Let C be a code of length n over and alphabet of size q with minimum Hamming distance d. Then $\log_q(|C|) \leq n - d + 1$.

Definition 4. Maximum Distance Separable code
A code of length n over an alphabet of size q with $|C| = q^k$ and minimum Hamming distance d satisfying $k = n - d + 1$ is said to be a Maximal Distance Separable (MDS) code.

Theorem 4. *Sphere Packing Bound*
Let C be a code of length n over alphabet of size q with minimum Hamming distance $2t + 1$. Then

$$|C| \left(\sum_{s=0}^{t} \binom{n}{s} (q - 1)^s \right) \leq q^n$$

Definition 5. Perfect Code
A perfect code is a code C of length n where every vector in $\mathbb{F}_q^n$ is contained in precisely one sphere of radius t centered about a codeword.

Example 2. Perfect Codes
The following are all prefect codes:

- the codes $C = \mathbb{F}_q^n$
- the codes consisting fo exactly one codeword (the zero vector in the case of linear codes)
- the binary repetition codes of odd length (i.e., $\mathbf{1} = 11111$, length $n = 5$)
- the binary codes of odd length consisting of a vector c and the complementary vector $\bar{c}$ with 0's and 1's interchanged.

Theorem 5. *If C is a linear code over a ring R, the th minimum Hamming distance and the minimum Hamming weight are equal.*

Definition 6. Generator Matrix
A Generator Matrix is any $k \times n$ matrix G whose rows forma basis for C.

Definition 7. Parity-check Matrix
A parity-check matrix H is a $(n-k) \times n$ matrix, for a (n,k) code C, defined by

$$C = \Big\{ x \in \mathbb{F}_q^n \mid Hx^T = 0 \Big\}.$$

Note that C is the kernel of the linear transformation H, because a linear code is a subspace of a vector space.

Theorem 6. *If $G = [I_k \mid A]$ is a generator matrix for the (n,k) code C in standard form, then $H = [-A^T \mid I_{n-k}]$ is a parity check matrix for C.*

Definition 8. Let R be a finite ring. A linear code C over the alphabet R of length n is a submodule of R^n.

Note: If R is a field then the linear codes are vector spaces and we have the full force of linear algebra at our disposal!