Much of classical coding theory concentrates on a class of codes called **linear codes**.

Suppose we wish to encode \( k \) bits using \( n \) bits \((n > k)\). The encoding of all \( k \)-bit words to codewords is defined by an \( n \times k \) matrix \( G \):

\[
\begin{pmatrix}
\end{pmatrix}
\]

\[
\begin{pmatrix}
\end{pmatrix}
\]

All arithmetic is done \( \text{mod} \ 2 \) (i.e. over \( \mathbb{Z}_2 \)).

The columns of \( G \) are linearly independent so that every \( n \)-bit word to be encoded is mapped to a unique codeword \( Gv \).

Columns of \( G \) form a basis for the \( k \)-dimensional space of codewords.

**Dual matrix** \( P \) is an \((n-k) \times n\) matrix called the dual matrix. The rows of \( P \) form a basis of the subspace orthogonal to the set of all codewords. The rows of \( P \) along with the columns \( P \) and \( G \) are linearly independent and span the space of all \( n \)-bit strings.

We have \( PG = 0 \). Since any code word \( s \) can be expressed as \( Gv \),

\[
Pv = P G v = 0
\]

In fact, \( P s = 0 \) iff \( s \) is a codeword.

(\( \Rightarrow \)) (comes from the fact that the rows of \( P \) and columns of \( G \) span the whole space of \( n \)-bit strings)
P is called the Parity Check Matrix. It can be used to test whether a word is a valid code word.

The Hamming Distance between two words is the minimum number of bits that must be flipped to turn one into the other. The distance between \( a + b \) is the weight (#1's) in \( (a+b) \).

For a code to correct t errors, the distance between any two codewords must be at least \( 2t+1 \). t errors take a codeword a distance \( t \) from where it started. In order to distinguish between two codewords it is required that the ball of radius \( t \) around each codeword do not intersect.

A code that encodes k bits using n bits with a distance d between codewords is a \([n, k, d]\) code.

We can describe a set of errors with a n-bit vector \( e \) which has a 1 wherever an error occurs and 0's everywhere else. If the original code is \( s \), after the error it becomes \( s' = s + e \). If we apply this to the parity check:

\[
P s' = P (s + e) = P s + P e = P e
\]

\( P s' \) is independent of \( s \) and depends only on \( e \). If \( e \) can be uniquely determined from \( P e \), we can detect and fix the error. \( P e \) is the error syndrome.
\[ P_e = P_f \quad \text{iff} \quad P(e-f) = 0. \quad \Leftrightarrow \quad \text{distance between two codewords can be as small as } \|e-f\|. \]

In order for the distance of the code to be \( d \), we must have for any vector \( \bar{e} \) of unit \( d-1 \) or less, \( P_e \neq 0 \).

\( \Rightarrow \) Any \( d-1 \) columns of \( P \) must be linearly independent.