The Hamming codes, defined in this section, are an importtant family of single-error-correcting codes, which are easier to to encode and decode than other similar codes.
These codes are best defined using their parity-check matrix:

Definition 0.1. Let r be a positive integer and let H be an $r \times (2^r - 1)$ matrix whose columns are the distinct non-zero vectors of $V(r, 2)$. Then the code having H as its parity-check matrix is called a ***binary Hamming code*** and is denoted by $\mathrm{Ham}(r, 2)$.

Remark. (i) $\mathrm{Ham}(r, 2)$ has length $n = 2^r - 1$ and dimension $k = n - r$. r is also called the ***redundancy*** of the code.
(ii) $\mathrm{Ham}(r, 2)$ is the code generated by any of the matrix H, since its colums can be taken in any order.

Theorem 0.2. The binary Hamming code $\mathrm{Ham}(r, 2)$, for $r \geq 2$,

(i) is a $[2^r - 1, 2^r - 1 - r]$-code;
(ii) has a minimum distance 3 (therefore, single-error correcting);
(iii) is a perfect code.

Proof.

(i) By definition, $\mathrm{Ham}(r, 2)^{\perp}$ is a $[2^r - 1, r]$-code ans so $\mathrm{Ham}(r, 2)$ is a $[2^r - 1, 2^r - 1 - r]$-code.
(ii) Since, $\mathrm{Ham}(r, 2)$ is a linear code, it is enough, by Theoem **??**, to show that every non-zero codeword has weight≥ 3.
Firstly, suppose that the code has codeword $\mathbf{x}$ of weight 1, with non-zero element at ith position, but as it is orthogonal to every row in H, this implies that the ith column of H is the all-zero vector, contradicting the definition of H.
Similarly, if codeword $\mathbf{x} \in \mathrm{Ham}(r, 2)$, where

$$\mathbf{x} = 0 \cdots 010 \cdots 010 \cdots 0$$

with 1s in the ith and jth places. Let $H = \{h_{pq}\}$ then since $\mathbf{x}$ is orthogonal to each row of H, this implies

$$h_{si} = h_{sj} \pmod 2 \quad \forall\, s \in \{1, 2, \ldots, r\}$$

Hence the ith and jth columns of H are identical, again contradicting the definition. Thus $d(\mathrm{Ham}(r, 2)) \geq 3$, and it is easy to see that the equality exists, for example vector $11100 \cdots 0$ can exist in $\mathrm{Ham}(r, 2)$, by a suitable rearrangement of columns in H.
(iii) The left-hand side of shere packing bound is

$$2^{n-r}\left(1 + \binom{n}{1}\right) = 2^{n-r}(1 + n) = 2^n$$

Therefore, the bound is acheived and the code is a perfect code.

□

As a binary Hamming code is a perfect code, the coset leaders are precisely the 2^r vectors of $V(n, 2)$ of weight ≤ 1. Now if we arrange the columns of H in order of increasing binary numbers, we can have the following nice **decoding algorithm**. We calculate the syndrome $S(\mathbf{y})$ of the received vector $\mathbf{y}$ as usual, if it is $\mathbf{0}$, then assume it was the codeword sent, otherwise assume single error and the value of $S(\mathbf{y})$ gives the position of the error, as the syndromes are the columns of H itself.

Definition 0.3. The ***extended binary Hamming code*** $\hat{\mathrm{H}}\mathrm{am}(r, 2)$ is the code obtained from $\mathrm{Ham}(r, 2)$ by adding an overall parity-check.

It can be proven by above results that $\hat{\mathrm{H}}\mathrm{am}(r, 2)$ is a $[2^r, 2^r - 1 - r, 4]$-code. It is of no better than $\mathrm{Ham}(r, 2)$ when used for complete decoding, but can be used for incomplete decoding, for it can simultaneously correct any single error and detect any double error.

Now we shall define Hamming codes over arbitrary field $GF(q)$, before that we need to understand the following relationship between minimum distance of a linear code and linear inidependence property between columns of parity-check matrix.

Theorem 0.4. Suppose C is a linear $[n, k]$-code over $GF(q)$ with parity-check matrix H. Then the minimum distance of C is d if and only if any $d-1$ columns of H are linearly independent but some d columns are linearly dependent.

Proof. By Theoem **??**, the minimum distance of C is equal to the smallest of weights of the non-zero codewords. Let $\mathbf{x} = x_1x_2\cdots x_n$ be a vector in $V(n, q)$. Then

$$\begin{aligned}\mathbf{x} \in C &\iff \mathbf{x}H^T = 0\\ &\iff x_1\mathbf{H}_1 + x_2\mathbf{H}_2 + \cdots + x_n\mathbf{H}_n = \mathbf{0}\end{aligned}$$

where $\mathbf{H}_i$ denote the columns of H.
Therefore, for each codeword $\mathbf{x}$ of weight d, there is a set of d linearly dependent columns of H. On the other hand, if there existed some $d-1$ dependent columns of H, then we would have a codeword, defined by coefficients (not all zero) of those $d-1$ columns, with coeffients being the value of codeword's ith position (where i belongs ti $d-1$ column positions in H), and 0 at other places. Thus, this codeword will have weight $< d$, hence contradiction. □

Now, any non-zero vector $\mathbf{v}$ in $V(r, q)$ has exactly $q-1$ non-zero scalar multiples, forming set $\{\lambda\mathbf{v}|\lambda \in GF(q), \lambda \neq 0\}$. Hence, q^r-1 non-zero vectors of $V(r, q)$ may be partitioned into $(q^r-1)/(q-1)$ vectors, where each set consists of elements that are scalar multiples of other elements of that set. Now by choosing one vector from each set of $(q^r-1)/(q-1)$ vectors, no two of which are linearly dependent. Therfore, by theorem 0.4, taking these as the columns of H gives a parity-check matrix of a $\left[\frac{(q^r-1)}{q-1}, \frac{q^r-1}{q-1} - r, 3\right]$-code. This is called ***q-ary Hamming code*** and is denoted by $\mathrm{Ham}(r, q)$. Here also, $\mathrm{Ham}(r, q)$ is unique upto equivalence.

Example: A parity-check matrix for $\mathrm{Ham}(2, 11)$ is

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{bmatrix}$$

Theorem 0.5. $\mathrm{Ham}(r, q)$ is a perfect single-error-correcting code.

Proof. $\mathrm{Ham}(r, q)$ was constructed to be an $(n, M, 3)$-code with $n = (q^r-1)/(q-1)$ and $M = q^{n-r}$. The left-hand side of the sphere-packing bound becomes

$$\begin{aligned} q^{n-r}(1 + n(q-1)) &= q^{n-r}(1 + q^r - 1) &(1)\\ &= q^n &(2)\end{aligned}$$

which is the right-hand side, so $\mathrm{Ham}(r, q)$ is a perfect code. □

Corollary 0.5.1. If q is a prime power and if $n = (q^r-1)/(q-1)$, for some $r \geq 2$, then

$$A_q(n, 3) = q^{n-r}$$

Decoding with a q-ary Hamming code:
Since a Hamming code is a perfect single-error correcting code, the codet leaders, other than $\mathbf{0}$, are precisely the vectors of weight 1. The syndrome of such a coset leader $codex = 0\cdots 0b0\cdots 0$, with non-zero element at the jth place, is

$$S(\mathbf{x}) = b\mathbf{H}_j^T,$$

where $\mathbf{H}_j$ denotes the jth column of H.
So the decoding scheme is as follows. If the received vector is $\mathbf{y}$, then first calculate the syndrome $S(\mathbf{y})$, if $S(\mathbf{y}) = \mathbf{0}$, assume no errors, otherwise $S(\mathbf{y}) = b\mathbf{H}_j^T$ for some b and j and the assumed single error is corrected by subtracting b from the jth entry of $\mathbf{y}$.