↓ s		c → 000	100	010	110	001	101	011	111
0000	row 1	0000000	0111100	1011010	1100110	1101001	1010101	**0110011**	0001111
1000	row 2	1000000	1111100	0011010	0100110	0101001	0010101	1110011	1001111
0100	row 3	0100000	0011100	1111010	1000110	1001001	1110101	0010011	0101111
0010	row 4	0010000	0101100	1001010	1110110	1111001	1000101	0100011	0011111
0001	row 5	0001000	0110100	1010010	1101110	1100001	1011101	0111011	0000111
0111	row 6	0000100	0111000	1011110	1100010	1101101	1010001	0110111	0001011
1011	row 7	0000010	0111110	1011000	1100100	1101011	1010111	0110001	0001101
1101	row 8	0000001	0111101	1011011	1100111	1101000	1010100	0110010	0001110
1100	row 9	1100000	1011100	0111010	0000110	0001001	0110101	1010011	1101111
1010	row 10	1010000	1101100	0001010	0110110	0111001	0000101	1100011	1011111
0110	row 11	0110000	0001100	1101010	1010110	1011001	1100101	0000011	0111111
1001	row 12	1001000	1110100	0010010	0101110	0100001	0011101	1111011	1000111
0101	row 13	**0101000**	0010100	1110010	1001110	1000001	1111101	**0011011**	0100111
0011	row 14	0011000	0100100	1000010	1111110	1110001	1001101	0101011	0010111
1111	row 15	1000100	1111000	0011110	0100010	0101101	0010001	1110111	1001011
1110	row 16	1110000	1001100	0101010	0010110	0011001	0100101	1000011	1111111

error patterns that this code can correct

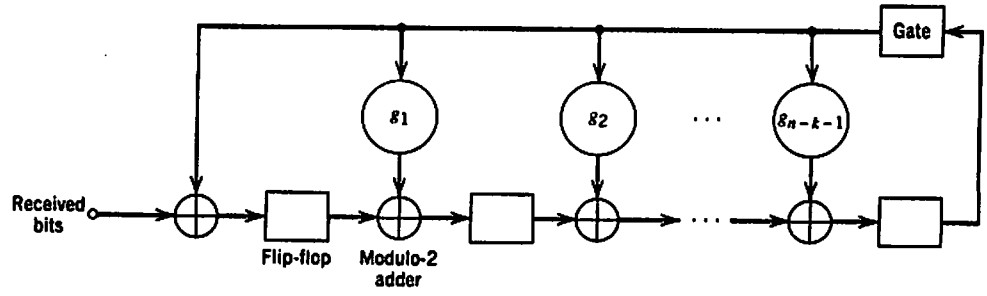
For this code, $n=7$ and $k=3$. Note that the table has $2^{n-k}=16$ rows and $2^k=8$ columns, and therefore has $2^n=128$ entries. These 128 entries are all different and represent all possibilities for the received vector $\underline{r}$. The first column that has the entry 0000000 represents error patterns correctable by the code and the first row that begins with 0000000 represents all the codewords. When one receives a vector $\underline{r}$, one calculates its syndrome $\underline{s}$. There is a table that has the entries of the first column against all values of $\underline{s}$. This yields $\underline{e}$. Then one calculates

$$\underline{c} = \underline{r} + \underline{e}.$$

The last three bits in $\underline{c}$ is the message vector $\underline{m}$. (why?).

As an example, assume one receives 0011011 (bold in the table). The corresponding syndrome is 0101. The corresponding $\underline{e}$ vector is 0101000. This yields $\underline{c}$ as 0101000 + 0011011 0110011. The last three bits are the $\underline{m}$ vector since the code is systematic.

Alternatively to the procedure described above, one can keep



Whenever all the received bits are fed into the shift register, it contains the syndrome $\underline{s}$.

We will show that the syndrome of a received codeword polynomial is the syndrome of the corresponding error polynomial.

Note

$$r(X) = c(X) + e(X)$$

$$e(X) = r(X) + c(X)$$

$$= q(X)g(X) + s(X) + a(X)g(X)$$

$$= v(X)g(X) + s(X)$$

Therefore $s(X)$ is the syndrome of $e(X)$. Note if $e(X)=0$ then $s(X)=0$. When there are errors, in general $s(X) \neq 0$ except for a number of events with small probability of occurrence.

<u>Example</u> There is a family of cyclic codes known as Hamming codes with the following parameters

$$n = 2^m - 1$$

$$k = 2^m - m - 1$$

$$n - k = m$$

where $m \geq 3$. For the $m=3$ Hamming code, the generator matrix $\underline{G}$ is given as

$$\underline{G} = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}}_{\underline{P}} \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\underline{I}_4}$$

The $\underline{H}$ matrix is

$$\underline{H} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\underline{I}_3} \underbrace{\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}}_{\underline{P}^T}$$

The $2^4 = 16$ codewords can be calculated by multiplying $\underline{G}$ on the left by a column vector consisting of four binary values, running from 0000 to 1111.

Message Word	*Code Word*	*Weight of Code Word*	*Message Word*	*Code Word*	*Weight of Code Word*
0000	0000000	0	1000	1101000	3
0001	1010001	3	1001	0111001	4
0010	1110010	4	1010	0011010	3
0011	0100011	3	1011	1001011	4
0100	0110100	3	1100	1011100	4
0101	1100101	4	1101	0001101	3
0110	1000110	3	1110	0101110	4
0111	0010111	4	1111	1111111	7

Note that this code has $d_{min}^{H} = 3$. Therefore this code can correct any 1-bit error among the 7 bits it releases to the channel. The error patterns and the corresponding syndrome values are related through $\underline{s} = \underline{e}\,\underline{H}^T$.

Syndrome	*Error Pattern*
000	0000000
100	1000000
010	0100000
001	0010000
110	0001000
011	0000100
111	0000010
101	0000001

Assume that the codevector 1110010 was sent but it was received as 1010010 (error in the second bit). Let's calculate the syndrome

$$\underline{S} = [1010010] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$= [0\ 1\ 0]$$

Looking at the syndrome table, we see the error pattern 0100000. Adding the pattern to the received codeword we get the transmitted codeword as

$$1\,0\,1\,0\,0\,1\,0 + 0\,1\,0\,0\,0\,0\,0 = 1\,1\,1\,0\,0\,1\,0$$

from which we conclude the message was 0010 because the code is systematic.

Hamming codes are cyclic. The polynomial X^7+1 can be factored as

$$X^7+1 = (1+X)(1+X^2+X^3)(1+X+X^3).$$

The polynomials on the RHS are irreducable, they have no factors with coefficients in the set $\{0,1\}$. By taking

$$g(X) = 1+X+X^3$$

we will get a cyclic code with $n-k=3$ and $n=7$.

Let's encode the message sequence 0010. The message polynomial is $m(X) = X^2$.

$$X^{n-k}\, m(X) = X^5$$

Dividing $X^{n-k} m(X)$ by $g(X)$ and calculating the remainder $b(X)$ by long division, we have

$$
\begin{array}{r}
X^2+1 \\
X^3+X+1 \;\overline{)}\; X^5 \\
\underline{X^5+X^3+X^2} \\
X^3+X^2 \\
\underline{X^3+X+1} \\
X^2+X+1
\end{array}
$$

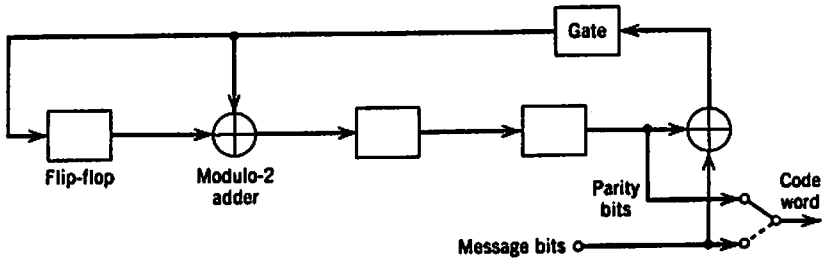
Note the convention of the message sequence

$$
\begin{array}{cccc}
m_0 & m_1 & m_2 & m_3 \\
0 & 0 & 1 & 0
\end{array}
$$

$$
\begin{aligned}
c(X) &= b(X) + X^{n-k} m(X) \\
&= 1 + X + X^2 + X^5
\end{aligned}
$$

The codeword is 1110010 with the four rightmost bits 0010 being the message.

The circuit to implement this encoder is



Let's calculate the contents of the shift register as the message is entered (note the convention 0010 ← first to enter, last to enter)

Shift	Input	Register Contents
		000 (Initial state)
1	0	000
2	1	110
3	0	011
4	0	111 ← Corresponds to $1+X+X^2$

Note the convention that the leftmost register contains b_0 and the rightmost b_{n-k-1}.

We transmit $c(X)$ on the channel (first c_{n-1}, last $c_0 = b_0$). Assume c_1 is errored. Therefore

$$c(X) = 1 + X + X^2 + X^5$$

is received as

$$r(X) = 1 + X^2 + X^5.$$

Let's calculate the syndrome $s(X)$ as the remainder of the division of $r(X)$ by $g(X)$

$$\begin{array}{r} X^2 + 1 \\ X^3 + X + 1 \;\overline{)\; X^5 + X^2 + 1} \\ \underline{X^5 + X^3 + X^2} \\ X^3 + 1 \\ \underline{X^3 + X + 1} \\ X \end{array}$$

$$s(X) = X$$

One can calculate a $\underline{G}$ matrix for this code. Recall that

$$g(X) = 1 + X + X^3$$

then

$$X\,g(X) = X + X^2 + X^4$$
$$X^2 g(X) = X^2 + X^3 + X^5$$
$$X^3 g(X) = X^3 + X^4 + X^6$$

By using the coeffients of these polynomials as the rows of a 4x7 matrix, we get the following matrix

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

However, this matrix is not in the systematic form. By modulo-2 addition of the first row to the third and fourth rows and by the modulo-2 addition of the second row to the fourth, it can be made systematic.

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

From this generator matrix, one can generate a table of syndromes and 1-bit error patterns and based on $s(X)$ calculate $e(X)$ and therefore $c(X)$ and $m(X)$. (In this case, recognize this $\underline{G}$ matrix as the one belonging to the Hamming code studied earlier. Therefore, we have

$$e(X) = X$$

and

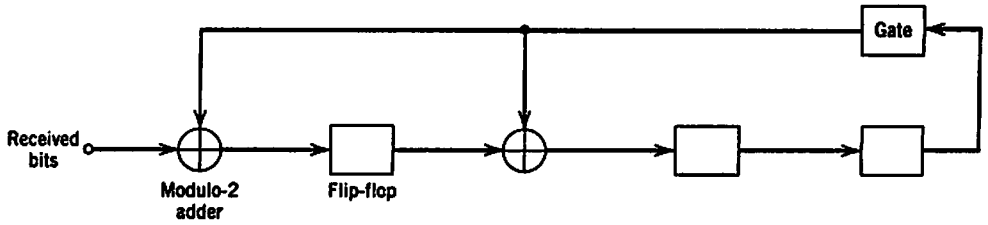
$$c(X) = 1 + X + X^2 + X^5$$

$$\underline{c} = 1110010$$

and

$$\underline{m} = 0010.$$

The circuit to calculate the syndrome is given on the next page.



Let's calculate the contents of the shift register as the received codeword 1010010 is entered into the shift register.

Shift	Input	Contents of Shift Register
		000 (Initial state)
1	0	000
2	1	100
3	0	010
4	0	001
5	1	010
6	0	001
7	1	010

And therefore $s(X) = X$, as calculated previously.

Cyclic codes are used in various communication systems. One group that is used in many communications protocols is known as Cyclic Redundancy Check (CRC) codes. With CRC codes, the purpose is to detect the presence of errors. If the syndrome is not zero, there are errors. In data communications, when there are errors a request for retransmission is made. Such protocols are called Automatic Repeat Request (ARQ) protocols.

Some cyclic codes used for error detection and correction are BCH (Bose-Chaudhuri-Hocquenghem) codes and Reed-Solomon codes.