D. Particular Cyclic Codes

1. Maximal Length Codes

Def: Let $h(D) = h_0 + \cdots + D^m$ be the minimal polynomial of a primitive element $\alpha \in GF(p^m)$. Then $h(D)$ is called a <u>primitive polynomial</u>.

Note: See Peterson for lists of primitive polynomials of various degrees

<u>Theorem 16</u>

a cyclic block code of length $N = p^m - 1$ with a check polynomial $h(D)$ which is a primitive polynomial of degree m over $GF(p)$ is a maximal length code.

<u>Proof:</u>

The number of information symbols is $K = \deg h(D) = m$. Therefore there must be p^m codewords including the all 0 word. Therefore there are $p^m - 1$ nonzero codewords. One codeword is $g(D) = (D^{p^m-1} - 1) / h(D)$. Each cyclic shift of $g(D)$ must be a codeword. Thus the codewords corresponding to $g(D)$ and its cyclic shifts are a set of $N = p^m - 1$ words. It will now be shown that each of these cyclic shifts is distinct. Suppose that the ith and jth shifts are identical for $0 \le i < j < p^m - 1$. Then

$$[D^i g(D)] \bmod (D^{p^m-1} - 1) = [D^j g(D)] \bmod (D^{p^m-1} - 1)$$

Thus for some $a(D)$ and $b(D)$

$$D^i g(D) - a(D)[D^{p^m-1} - 1] = D^j g(D) - b(D)[D^{p^m-1} - 1]$$

or $$D^i - a(D)h(D) = D^j - b(D)h(D)$$

and $[b(D) - a(D)]\,h(D) = D^j - D^i = D^i[D^{j-i} - 1]$. Since $j - i < p^m - 1$, α cannot be a root of $D^{j-i} - 1$ and $h(D)$ cannot divide the right hand side by Theorem 10. Therefore each of the $p^m - 1$ cyclic shifts of the codeword corresponding to $g(D)$ is a unique code word. A generator matrix for the code is

$$G = \begin{array}{c} \\ \\ \\ \\ \\ \end{array}\begin{bmatrix} g_0 & g_1 & \cdots & g_{N-m-1} & 1 & \overbrace{0 \cdots 0}^{m-1} \\ 0 & g_0 & \cdots & g_{N-m-1} & 1 & 0 \cdots 0 \\ \vdots & & & & & \\ \vdots & & \cdots & & 1 & 0 \\ 0 \cdots & 0 & g_0 & \cdots & g_{N-m-1} & 1 \end{bmatrix} \Big\updownarrow m$$

$$\longleftarrow N = p^m - 1 \longrightarrow$$

The codeword corresponding to $g(D)$ is

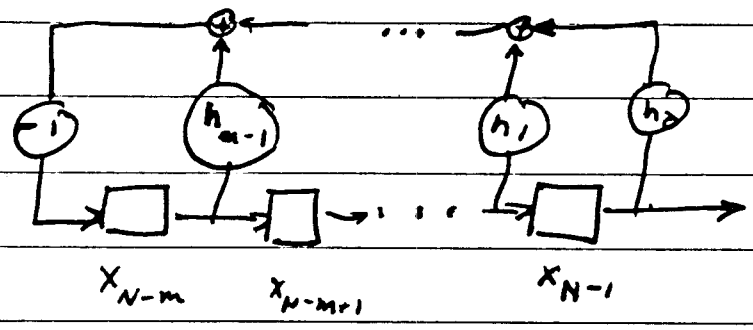
$$[g_0 \; g_1 \; \cdots \; g_{N-m-1} \; \underbrace{1 \; 0 \cdots 0}_{m}]$$

If $x_{N-1}, \cdots x_{N-m}$ are taken as the information digits, this codeword corresponds to the information sequence $\vec{u} = [x_{N-m}, \cdots x_{N-1}] = [1\,0 \cdots 0]$ which corresponds to the last column of G. The cyclic shift $T\vec{g} = [0 \; g_0 \; g_1 \; \cdots \; g_{N-m-1} \; 1 \; 0 \cdots 0]$ is also a codeword and corresponds to $\vec{u} = [g_{N-m-1} \; 1 \; 0 \cdots 0]$ and the next to last column of G. Similarly each column of G is the set of information digits for one of the $p^m - 1$ cyclic shifts of $\vec{g}$. Since these cyclic

shifts are all distinct, the information sequences must be distinct and hence the columns of G are all the $p^m - 1$ nonzero m-tuples and G is the generator for a maximal length code.

Q.E.D.

One circuit for generating these codes is shown below.



The maximal length codes are also known as <u>pseudo-noise</u> (P-N) sequences. It was shown earlier that each nonzero $\alpha \in GF(p)$ appears p^{m-1} times in each codeword. Consider the binary codes, i.e., $p=2$. Then each codeword has 2^{m-1} ones and $2^m - 1 - 2^{m-1} = 2^{m-1} - 1$ zeros so that $P(1) \cong P(0) \cong 1/2$ for moderate m. If the sequences were truly binary random sequences one would expect a run of k 1's or k 0's to have probability 2^{-k}. A run of k 1's is a sequence of k consecutive 1's preceded and followed by a 0. A run of k 0's is defined similarly. In other words one would expect to find runs of length 1 about 1/2 the time, runs of

length 2 about 1/4 of the time, etc. Consider the encoder shown on p. 140. As it is shifted $2^m - 1$ times each of the possible $2^m - 1$ non zero sequences of m binary digits appears once in the register. The sequence of m 1's occurs once. It must be preceded and followed by 0 or else there would be more than one sequence of m ones in the codeword.

A zero followed by $m-1$ ones occurs exactly once. This is accounted for by the sequence of m 1's preceeded by a 0. Similarly $m-1$ ones followed by a 0 is accounted for. Thus there can be no run of $m-1$ ones

Let $0 < k < m-1$. To find the number of runs of 1's of length k, consider m consecutive digits beginning with 0, then k 1's, then a 0 and the remaining $m-2-k$ terms arbitrary. This can occur 2^{m-k-2} ways since each combination of the last $m-2-k$ digits can occur. Analogous reasoning holds for runs of k 0's for $0 < k < m-1$.

No run of m 0's can occur but a 1 followed by $m-1$ 0's is valid and so is $m-1$ 0's followed by a 1. Thus a run of $m-1$ 0's occurs.

The total number of runs is $2 \times \sum_{k=1}^{m-2} 2^{m-k-2} + 1 + 1$

$= 2^{m-1} = (N+1)/2$. Half are runs of 1's and half are runs of 0's, i.e., 2^{m-2} runs of each. Of the runs of 1's half have length 1, 1/4 have length 2, ..., $2^{-(m-2)}$ have length $m-2$ and similarly for runs of k 0's with $1 \le k \le m-2$. Then there is a run of $m-1$ 0's and a run of m 1's.

Correlation Properties

The binary maximal length sequences have an interesting correlation property that makes them ideal for synchronization and ranging applications. Let $\vec{A} = \{a_0, \ldots, a_{2^m-2}\}$ be a binary maximal length sequence. Define $\vec{B} = \{b_0, \ldots b_{2^m-2}\}$ to be an analog sequence determined from $\vec{A}$ by the transformation $b_i = \begin{cases} +1 & \text{for } a_i = 0 \\ -1 & \text{for } a_i = 1 \end{cases}$.

Mod 2 Addition Table

	0	1
0	0	1
1	1	0

Analog Multiplication Table

	+1	-1
+1	1	-1
-1	-1	1

From the two tables above it is clear that mod 2 addition is equivalent to analog multiplication of the b's. If $\vec{A}_1$ and $\vec{A}_2$ are maximal length sequences so is $\vec{A}_3 = \vec{A}_1 + \vec{A}_2$ so that $\vec{A}_3$ also contains 2^{m-1} 1's and $2^{m-1}-1$ 0's unless ~~Equivalently~~

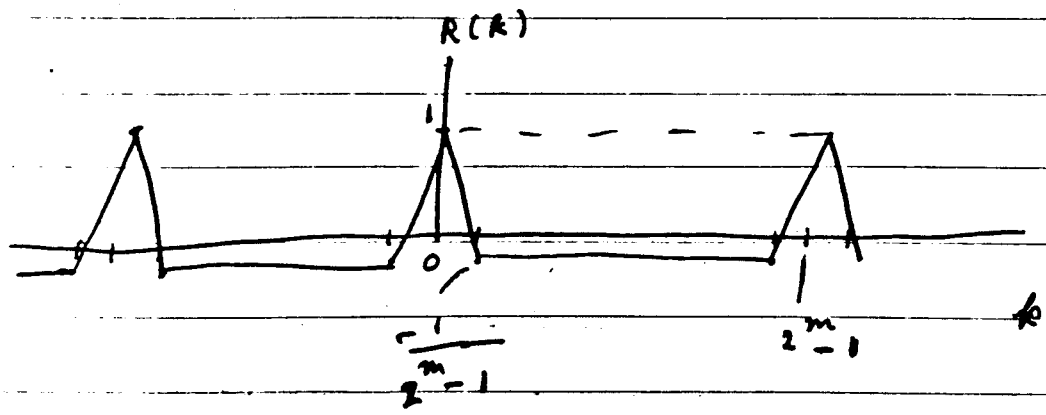
$\vec{A}_1 = \vec{A}_2$ so that $\vec{A}_3 = \vec{0}$. Equivalently $\vec{B}_3 = \vec{B}_1 \vec{B}_2$ $= \{ b_{11} b_{21}, b_{12} b_{22}, \ldots, b_{1N} b_{2N} \}$ corresponds to a maximal length sequence so that $\vec{B}_3$ has 2^{m-1} "-1's" and $2^{m-1}-1$ "+1's" unless $\vec{B}_1 = \vec{B}_2$ and then $\vec{B}_3$ has $2^m - 1$ "+1's".

The cyclic or periodic correlation function for a sequence $\vec{B}$ is defined as

$$R(k) = \frac{1}{2^m - 1} \sum_{l=0}^{2^m-2} b_l \, b_{(l+k) \bmod (2^m-1)}$$

If $\vec{B}$ corresponds to a maximal length sequence then so does the cyclic shift $T^{-k}\vec{B}$ and also $\vec{B}(T^{-k}\vec{B})$. Therefore

$$R(k) = \begin{cases} 1 & \text{for } k = 0, \pm(2^m-1), \ldots \\ -\dfrac{1}{2^m-1} & \text{otherwise} \end{cases}$$



Syndrome Decoding of Cyclic Codes

Let $X(D)$ be a transmitted codeword in the cyclic code generated by $g(D) = g_0 + \cdots + D^{N-K}$

and let the channel error pattern be $e(D) = e_0 + e_1 D + \cdots + e_{N-1} D^{N-1}$. The received word is $y(D) = y_0 + \cdots + y_{N-1} D^{N-1} = X(D) + e(D)$. Let the code be a systematic code, encoded by a type A or B encoder. The check symbols are then $x_0 + x_1 D \cdots + x_{N-K-1} D^{N-K-1}$ $= -[x_{N-K} D^{N-K} + \cdots + x_{N-1} D^{N-1}] \bmod g(D)$. The syndrome

$S(D)$ is the received check symbols minus the check symbols estimated from the received data digits, i.e.

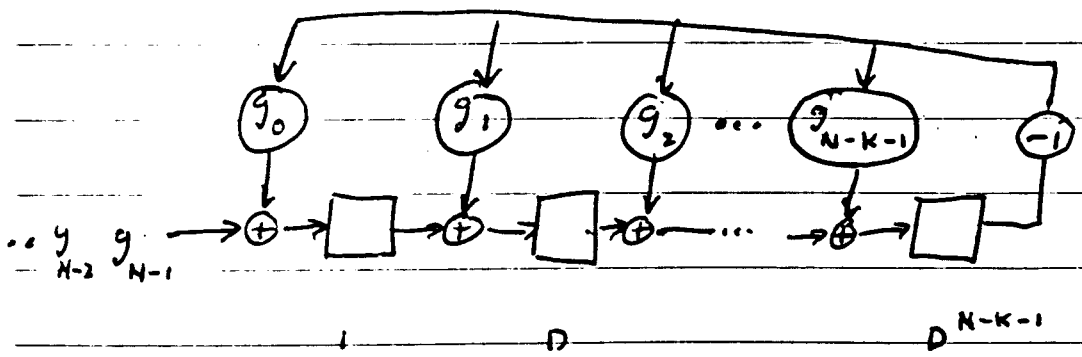
$$S(D) = y_0 + \cdots + y_{N-K-1} D^{N-K-1} + [y_{N-K} D^{N-K} + \cdots + y_{N-1} D^{N-1}] \bmod g$$

Since $\deg g(D) = N-K$,

$$S(D) = [y_0 + \cdots y_{N-K-1} D^{N-K-1} + \cdots y_{N-1} D^{N-1}] \bmod g(D)$$

$$= y(D) \bmod g(D) = [X(D) + e(D)] \bmod g(D) = e(D) \bmod g$$

The circuit shown below calculates $S(D)$ if the register is initially set to $\bar{0}$ and the

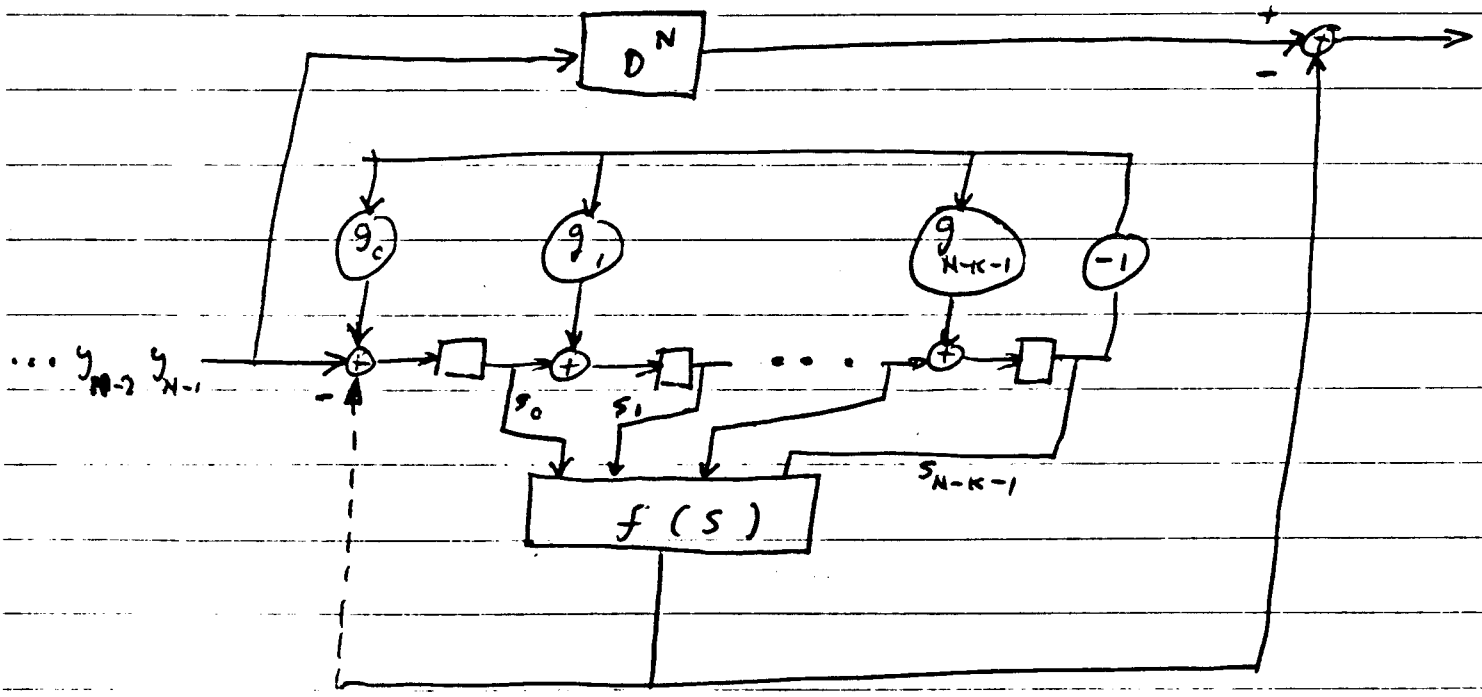


received codeword $y_{N-1}, y_{N-2} \cdots y_0$ is shifted into the register starting with the highest degree digits. If the register contents are

$c(D) = c_0 + \cdots + c_{N-K-1} D^{N-K-1}$ at a given time, on the next shift the new contents are $\{D\, c(D)\} \bmod g(D) + y_j = \{D\, c(D) + y_j\} \bmod g(D)$ where y_j is the input. Thus as $y(D)$ is shifted into the register the contents become

$$y_{N-1},\quad y_{N-2} + D\, y_{N-1},\ \ldots,\quad y(D) \bmod g(D)$$

This suggests a decoder of the following form:



The received word is shifted into the decoder and $s(D)$ appears in the register after all N coefficients are shifted in. At this time y_{N-1} appears at the output of the D^N delay. The function $f(s)$ forms an estimate of the error e_{N-1} and subtracts it from the received data resulting in the decoded digit $x^*_{N-1} = y_{N-1} - e^*_{N-1}$.

Because of the cyclic structure of the code the same logic function can be used to decode each digit. After the first N shifts when all the received digits have entered the register the input is set to 0. On the N+1st shift the new syndrome register contents are

$$s'(D) = [D\,y(D)] \bmod g(D)$$

$$= [D\,y(D) - y_{N-1}(D^N - 1)] \bmod g(D) \quad \text{since}$$

$g(D)$ divides $D^N - 1$. But $D\,y(D) - y_{N-1}(D^N-1)$ corresponds to a cyclic shift of $y(D) = X(D) + e(D)$ so

$$s'(D) = \{D\,e(D) - e_{N-1}(D^N - 1)\} \bmod g(D)$$

$$= e_{N-1} + e_0 D + \cdots + e_{N-K-2} D^{N-K-1} + [e_{N-K-1} D^{N-K} + \cdots + e_{N-2} D^{N-1}] \bmod g(D)$$

This is just the syndrome corresponding to the received data $T\vec{y} = T\vec{X} + T\vec{e}$ and X_{N-2} is now the coefficient of D^{N-1} in the codeword. Thus $f(s)$ should be used to decode X_{N-2} using the new syndrome $s'(D)$. The register is shifted a total of 2N times decode the received block.

Notice that after the N+1st shift, e_{N-1} only affects s_0, the constant term of $s'(D)$. This effect can be subtracted from the syndrome by closing the dotted path after the Nth shift. After 2N shifts the syndrome register contents would then be $[e(D) - e^*(D)] \bmod g(D)$. If the error pattern is correctable then $e = e^*$ and the register contents would be $\vec{0}$.

Therefore with the dotted connection, uncorrectable errors are detected.

The practicality of a syndrome decoder depends on the complexity of the logic function $f(S)$. For maximal length codes it will be shown below that $f(S)$ is a threshold element:

Threshold Decoding of Maximal Length Codes (Binary)

For the binary systematic maximal length codes generated by a Type A or B encoder, the generator matrix has the form

$$G = \left[P^T \,\vdots\, I_m \right]_{m \times (2^m - 1)}$$

where m is the degree of the primitive check polynomial $h(D)$. The columns of P^T are all the nonzero binary m-tuples of weight 2 or more. Therefore the check matrix has the form

$$H = \left[I_{2^m-1-m} \,\vdots\, P \right]_{(2^m-1-m)\times(2^m-1)}$$

Thus $\vec{S} = \vec{y} H^T = [s_0, \dots, s_{2^m-m-2}] = \vec{e} H^T$

$$= [e_0, e_1, \dots e_{N-1}] [I_{N-m} \,\vdots\, P]^T$$

corresponds to the syndrome polynomial

$$S(D) = e(D) \bmod g(D) = s_0 + s_1 D + \dots + s_{N-K-1} D^{N-K-1}$$

As an example consider the (15, 4) code with the primitive check polynomial $h(D) = D^4 + D + 1$. Then $H = [I_{11} \mid P]$ where P is an 11×4 matrix. The rows of H form a basis of the null space of the set of codewords. Since $h(D)$ divides $D^N - 1$ the null space must also be a cyclic code. Therefore any cyclic shift of a row of H must be in the null space. To find H for this code in the proper canonic form, first observe that the N-tuple $[0 \cdots 0\ 1\ 0\ 0\ 1\ 1]$ corresponding to $h(D)$ can be used as the last row according to p. 120. Similarly $(0 \cdots 0\ 1\ 0\ 0\ 1\ 1\ 0)$ and $[0 \cdots 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0]$ can be used as the next two rows. However, the next shift $[0 \cdots 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 0]$ cannot be used for the canonic form. However, any linear combination of rows can be used so that this shift plus the last row $= [0 \cdots 0\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1]$ can be used. Similarly the shift of this row + last row can be used as the next row. Thus at each step the next row from the bottom is taken as the left shift of the previous row if no 1 appears in the 5th position from the right after the shift or else it is taken as the shift + last row.

For this example the check matrix becomes

$$H = \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & & 1 & 0 & 0 & 1 \\ & 1 & \cdots & & 1 & 1 & 0 & 1 \\ & & & & 1 & 1 & 1 & 1 \\ & & \ddots & & 1 & 1 & 1 & 0 \\ & & & & 0 & 1 & 1 & 1 \\ & & & & 1 & 0 & 1 & 0 \\ & & & & 0 & 1 & 0 & 1 \\ & & & & 1 & 0 & 1 & 1 \\ & & & & 1 & 1 & 0 & 0 \\ & & & & 0 & 1 & 1 & 0 \\ & & & 1 & 0 & 0 & 1 & 1 \end{array}\right] = [I_{11} \mid P]$$

$\leftarrow m \rightarrow$

For maximal length codes $d_{avg} = d_{min} = (q-1)q^{k-1}$. So that for binary codes $d_{min} = 2^{m-1}$. For the above example $d_{min} = 2^3 = 8$. The following are a set of $J = d_{min} - 1 = 7$ checks orthogonal on $e_{N-1} = e_{14}$:

$$A_1 = s_0 = e_0 + e_{11} + e_{14}$$
$$A_2 = s_8 + s_1 = e_1 + e_8 + e_{14}$$
$$A_3 = s_9 + s_2 = e_2 + e_3 + e_{14}$$
$$A_4 = s_9 + s_4 = e_4 + e_9 + e_{14}$$
$$A_5 = s_6 = e_6 + e_{12} + e_{14}$$
$$A_6 = s_5 + s_7 = e_5 + e_7 + e_{14}$$
$$A_7 = s_{10} = e_{10} + e_{13} + e_{14}$$

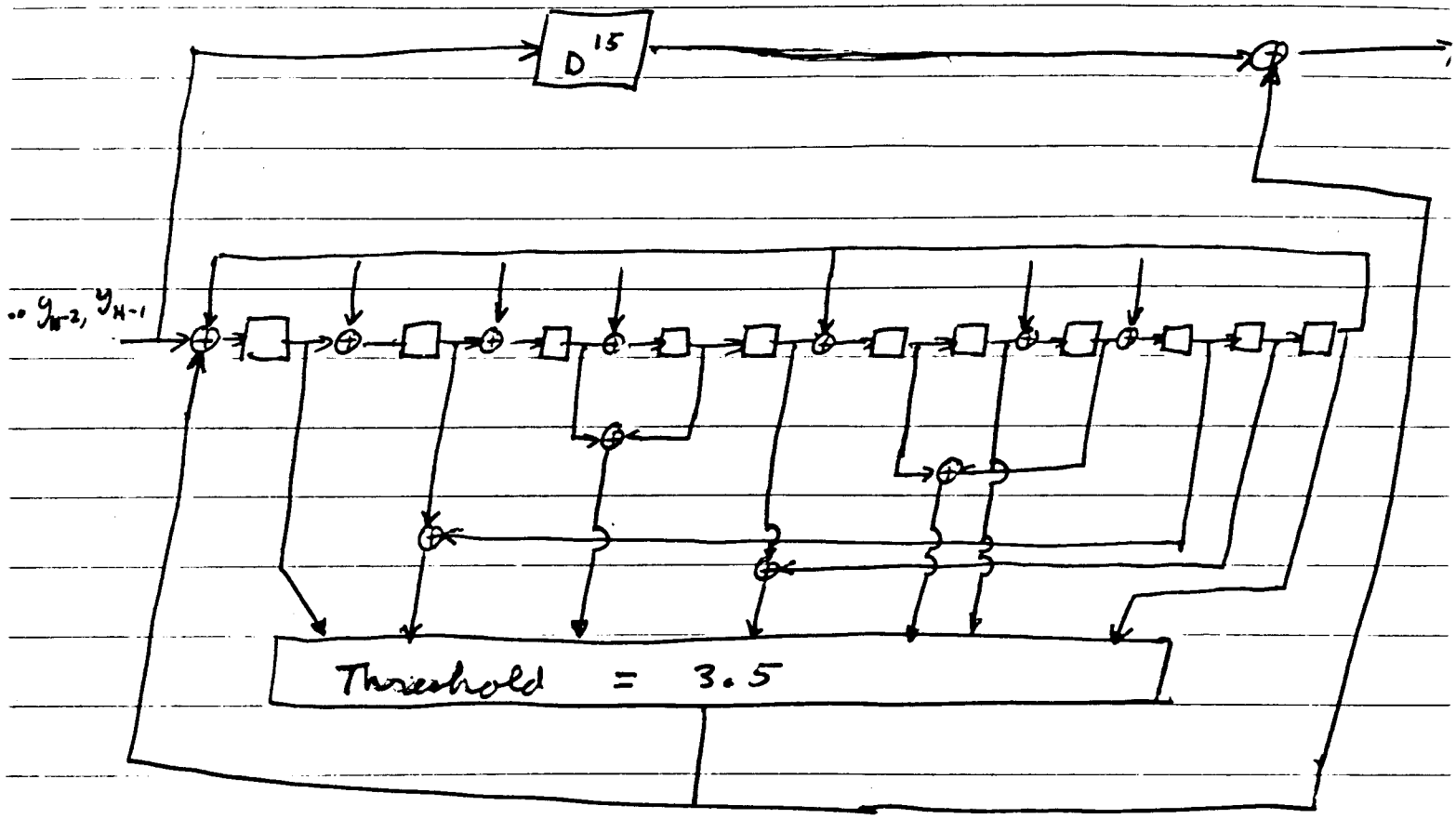
For any maximal length code over GF(2) $J = d_{min} - 1$ checks orthogonal on e_{N-1} can be found. Any row of H with a 1 in the last position checks e_{N-1}. There are m-1 rows that check e_{N-1}, only one other information noise bit and a check noise bit. Each check noise bit is different so that these rows form a set of m-1 checks $\perp$ on e_{N-1}. If more than one other information noise bit is checked by a row that checks e_{N-1}, there is a unique row that checks the same information noise bits but not e_{N-1}. The sum of these two rows gives a check on e_{N-1} and two check noise bits. A total of

$$\sum_{i=2}^{m-1} \binom{m-1}{i} = 2^{m-1} - (m-1) - 1$$ such pairs exist.

Thus there are $J = 2^{m-1} - 1 = d_{min} - 1$ checks $\perp$ on e_{N-1}.

Note: Since the dual code to a binary maximal length code is a Hamming single error correcting code, all the rows of H must have weight greater than or equal to 3.

A decoder for the (15, 4) code is shown on the following page.



2. Binary Single Error Correcting Hamming Codes

It was shown earlier that the binary maximal length and Hamming single error correcting codes are duals. If the generator polynomial for a code is a primitive polynomial of degree $N-K=m$, then the check matrix has all the 2^m-1 nonzero binary m-tuples as it's columns. This follows from the results for maximal length codes. Thus a primitive polynomial of degree m generates an $(2^m-1, 2^m-1-m)$ binary Hamming single error correcting code.

Recoding

assuming that only single errors are introduced by the channel

$$e(D) = D^{\ell} \qquad 0 \le \ell \le 2^m - 2$$

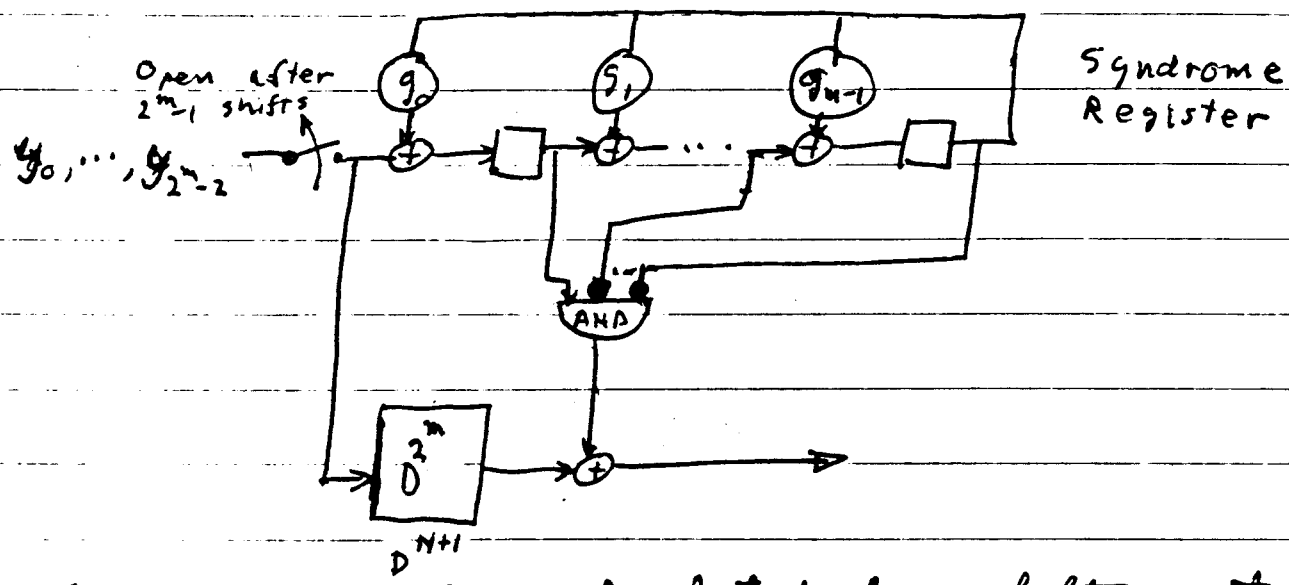
and $y(D) = x(D) + D^{\ell}$ so that

$$s(D) = y(D) \bmod g(D) = D^{\ell} \bmod g(D)$$

Since $s(D) \in GF(2^m)$ its multiplicative inverse is

$$s^{-1}(D) = D^{2^m-1-\ell} \bmod g(D)$$

a decoder is shown below that uses this property.



The syndrome is calculated by shifting the received word into the register. The input is then set to 0. after $2^m-1-\ell + (2^m-1) = N-\ell+N$ shifts the register contents are $[D^{\ell} \cdot D^{2^m-1-\ell}] \bmod g(D)$ $= D^{2^m-1} \bmod g(D) = 1$, The register cannot become 1 before this because $g(D)$ is a primitive polynomial at this time the symbol that appears at the output of the delay is $y_{N-(2^m-1-\ell)} = y_{\ell}$ which is