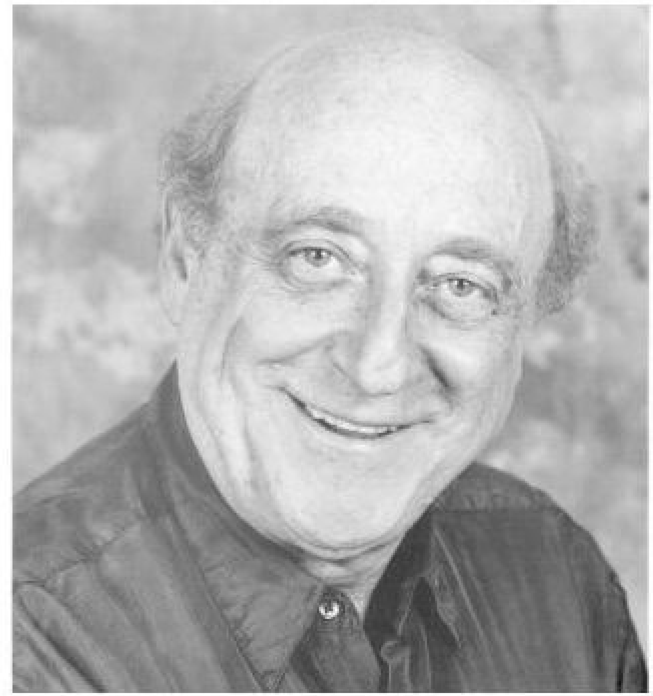
ECE 405/511
Error Control Coding

Reed-Solomon Codes

Irving Reed (1923-2012) Gus Solomon (1930-1996)



Polynomial Codes Over Certain Finite Fields, 1960

Reed-Solomon Codes

- Nonbinary BCH codes
- Consider GF(q)   (q=p^r, p prime)
- To construct a t error correcting nonbinary BCH code with symbols from GF(q), use the same technique as for binary BCH codes.
- Roots of $g(x)$ are in GF(q^m), $n|q^m$-1

  n-$k \leq 2mt$ product of at most 2t minimal polynomials of degree m

  $d \geq 2t$+1

- Choose $2t$ consecutive powers of α, an element of order n in GF(q^m), as roots of $g(x)$.
- For RS codes, m=1 and α is a primitive element in GF(q), then

$$n = q-1$$

$$n-k \leq 2t \longrightarrow n-k = 2t$$

$$d \geq 2t+1 \longrightarrow d \geq n-k+1$$

Singleton Bound

- The minimum distance for an (n,k) linear code is bounded by

$$d \leq n\text{-}k+1$$

- For an RS code $d \geq n\text{-}k+1$, so $d = n\text{-}k+1$ and all RS codes meet the Singleton bound with equality
  - they are optimal $(n,k,n\text{-}k+1)$ codes, $n = q\text{-}1$
- Codes that meet the Singleton bound are called Maximum Distance Separable (MDS)

Reed-Solomon Codes – Minimal Polynomials

- Coefficients of $g(x)$ are in GF(q), roots of $g(x)$ are also in GF(q).
- Minimal polynomial of α is x-α. There are no conjugates since $\alpha^q = \alpha^{q-1}\alpha = \alpha$.
- BCH: $M_1(x) = (x-\alpha)(x-\alpha^q)(x-\alpha^{q^2})\cdots$

  RS: $M_1(x) = (x-\alpha)$
- RS codes are a subclass of BCH codes with m = 1.

Example t=2 GF(8)

- n = 8-1 = 7 Form GF(8) from x^3+x+1

α^0	1
α^1	α
α^2	α^2
α^3	$\alpha+1$
α^4	$\alpha^2+\alpha$
α^5	$\alpha^2+\alpha+1$
α^6	α^2+1

$$g(x)=(x-\alpha)(x-\alpha^2)(x-\alpha^3)(x-\alpha^4)$$
$$=x^4+\alpha^3x^3+x^2+\alpha x+\alpha^3$$

$$\mathbf{H}=\begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha & \alpha^3 & \alpha^5 \\ 1 & \alpha^3 & \alpha^6 & \alpha^2 & \alpha^5 & \alpha & \alpha^4 \\ 1 & \alpha^4 & \alpha^8 & \alpha^5 & \alpha^2 & \alpha^6 & \alpha^3 \end{bmatrix}$$

- (7,3,5) RS code

Comparison: RS vs Binary BCH

- RS: $n = q^m\text{-}1 \quad q = 8, m = 1$ (7,3,5)

$$g(x) = (x-\alpha)(x-\alpha^2)(x-\alpha^3)(x-\alpha^4)$$

- Binary BCH: $n = q^m\text{-}1 \quad q = 2, m = 3$ (7,1,7)

$$g(x) = (x-\alpha)(x-\alpha^2)(x-\alpha^3)(x-\alpha^4)\boxed{(x-\alpha^6)(x-\alpha^5)}$$

- RS code: $q^k = 8^3 = 512$ codewords
- Binary BCH code: $q^k = 2^1 = 2$ codewords

Comparison: RS vs Binary BCH

- Each symbol can be represented as 3 bits,
  a codeword has n = 7 symbols = 21 bits and
  k = 3 data symbols = 9 bits.
- The (7,3,5) RS code can be considered as a (21,9) binary code.
- t = 2 symbol error correction
  - since 5 bit errors may cover 3 symbols, corrects any burst error of 4 bits or less.

Example t=3 GF(64)

- n = 64-1 = 63
- α a root of the primitive polynomial x^6+x+1

$$g(x) = (x-\alpha)(x-\alpha^2)(x-\alpha^3)(x-\alpha^4)(x-\alpha^5)(x-\alpha^6)$$
$$= x^6 + \alpha^{59}x^5 + \alpha^{48}x^4 + \alpha^{43}x^3 + \alpha^{55}x^2 + \alpha^{10}x + \alpha^{21}$$

- (63,57,7) RS code
- $64^{57} = 8.96\text{x}10^{102}$ codewords
- $64^{63} = 6.16\text{x}10^{113}$ vectors
- sphere volume is $9.94\text{x}10^9$ so the spheres fill about 14.5% of the vector space

GF(7) Example

- RS codes can be constructed over any finite field
- Consider q = 7 so that n = q-1 = 6, and t = 2
- First find a primitive element in GF(7)

ø(6) = 2 so two primitive elements

$3^1=3 \quad 3^2=2 \quad 3^3=6 \quad 3^4=4 \quad 3^5=5 \quad 3^6=1 \longrightarrow$ 3 is primitive

$b=1 \qquad g(x) = (x-3^1)(x-3^2)(x-3^3)(x-3^4)$

$\qquad\qquad = (x-3)(x-2)(x-6)(x-4)$ (6,2,5) RS code

$b=2 \qquad g(x) = (x-3^2)(x-3^3)(x-3^4)(x-3^5)$

$\qquad\qquad = (x-2)(x-6)(x-4)(x-5)$ (6,2,5) RS code

- One can pick any group of consecutive roots

$$g(x) = (x-3^1)(x-3^2)(x-3^3)$$
$$= (x-3)(x-2)(x-6) \qquad \text{(6,3,4) RS code}$$
$$= x^3+3x^2+x+6$$

$$g(x) = (x-3^2)(x-3^3)(x-3^4)$$
$$= (x-2)(x-6)(x-4) \qquad \text{(6,3,4) RS code}$$
$$= x^3+2x^2+2x+1 = g^*(x) \quad \text{self reciprocal}$$

$$g(x) = (x-3^1)(x-3^2)(x-3^3)(x-3^4)(x-3^5)$$
$$= (x-3)(x-2)(x-6)(x-4)(x-5) \qquad \text{(6,1,6) RS code}$$
$$= x^5+x^4+x^3+x^2+x+1 = g^*(x) \quad \text{self reciprocal}$$

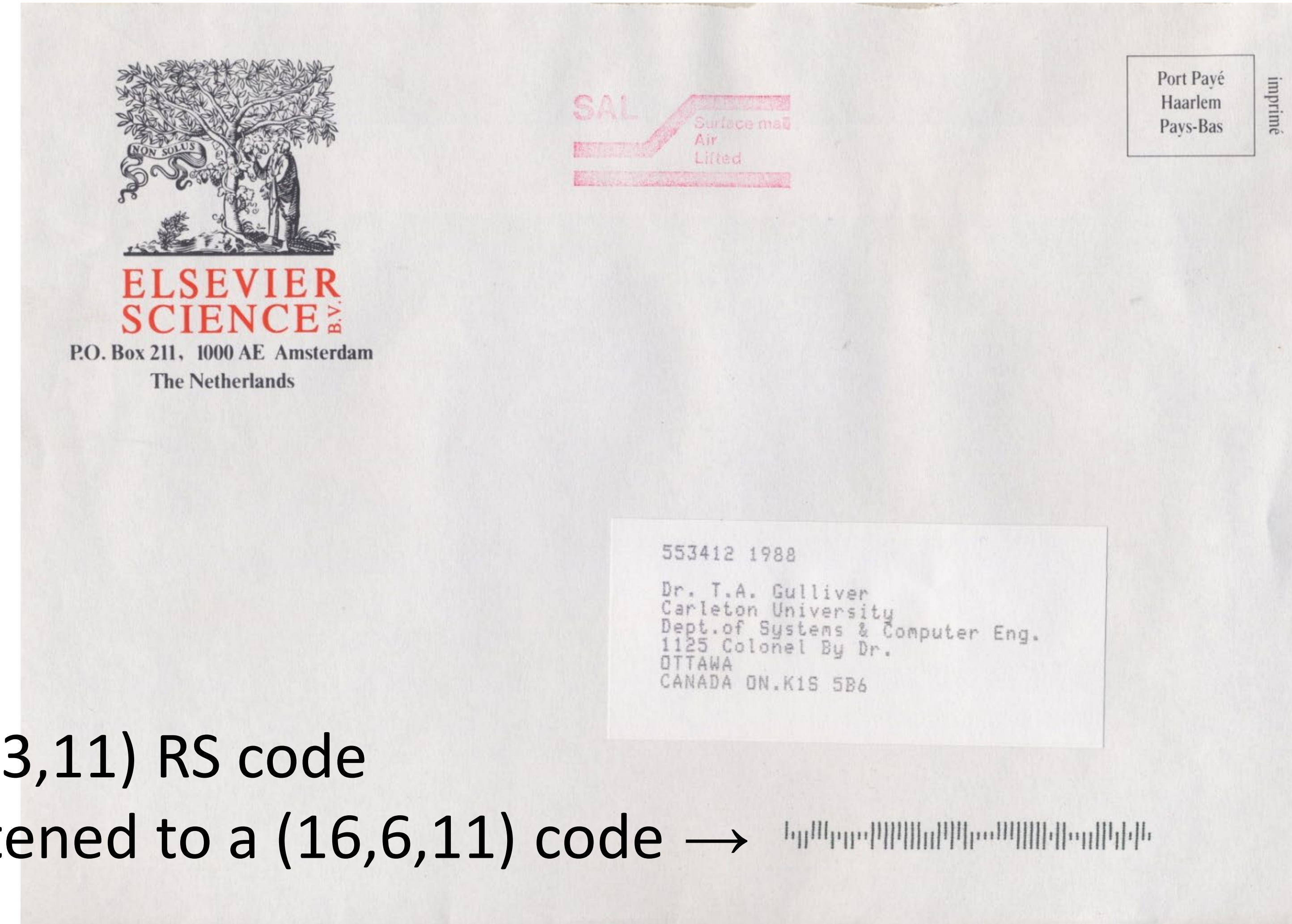
Properties of RS Codes

- The dual code of an RS code is also MDS
  - C (6,2,5) code over GF(7)
  - $C^{\perp}$ (6,4,3) code over GF(7)
- Since RS codes are cyclic codes, they can always be put in systematic form $x^{n-k}m(x)+d(x)$
- A shortened RS codes is MDS

  $(n,k,n\text{-}k+1) \longrightarrow (n\text{-}u,k\text{-}u,n\text{-}k+1)$ $(6,4,3) \longrightarrow (5,3,3)$
- A punctured RS code is MDS

  $(n,k,n\text{-}k+1) \longrightarrow (n\text{-}u,k,n\text{-}k\text{-}u+1)$ $(6,4,3) \longrightarrow (5,4,2)$

Example: Bar Codes over GF(64)



(63,53,11) RS code
shortened to a (16,6,11) code ⟶

Extended RS Codes

- An (n,k) RS code over GF(q) with $n = q$-1 can be extended twice to a $(q+1,k)$ MDS code
- There is a technique for constructing such codes which are cyclic
- A very few RS codes can be triply extended to obtain an MDS code with $n = q+2$
  - $k = 3$ or $n\text{-}k = 3$ and $q = 2^m$

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 & 0 \\ 1 & \alpha & \cdots & \alpha^{q-2} & 0 & 1 & 0 \\ 1 & \alpha^2 & \cdots & \alpha^{2(q-2)} & 0 & 0 & 1 \end{bmatrix}$$

Extended RS Codes

- The (6,3,4) RS code over GF(4) has generator matrix

$$\mathrm{G} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & \alpha & \alpha^2 & 0 & 1 & 0 \\ 1 & \alpha^2 & \alpha & 0 & 0 & 1 \end{bmatrix}$$

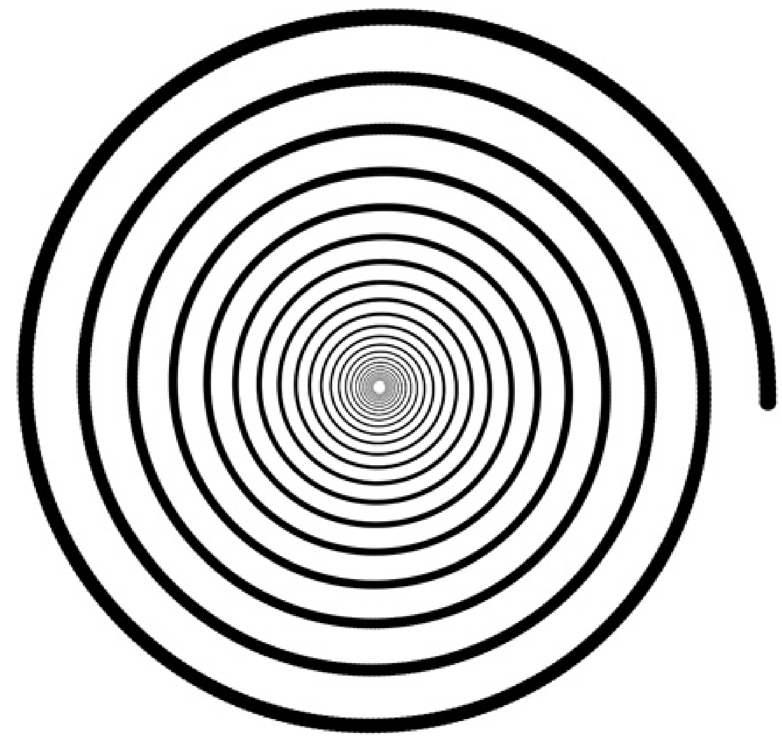
Example: NASA/JPL Code

- q = 256, n = q-1 = 255
- (255,223,33) RS code over GF(2^8)

$$\frac{\text{\# of codewords} \times \text{volume}}{\text{size of vector space}} = 2.78 \times 10^{-14}$$

Example: Compact Discs

- 44.1 kHz sample rate
- 16 bit stereo samples
- 2×16×44100=1.41 Mbps
- Original CD capacity: 74 minutes of audio or 650 MB of data
- Data stored on a spiral, not concentric circles
  - length 5.38 km
  - velocity 1.2-1.4 m/s



Kees Schouhamer Immink (1946-)

Sources of Error

1) Defects caused during disc production
   - inferior disc pits and bubbles during disc formation
   - defects in the aluminum film and a poor reflective index
2) Defects caused in handling
   - fingerprints and scratches
   - dust
3) Variations and disturbances during playback
   - disturbance of the servo mechanism
4) Jitter - time variation of the signal
5) Interference

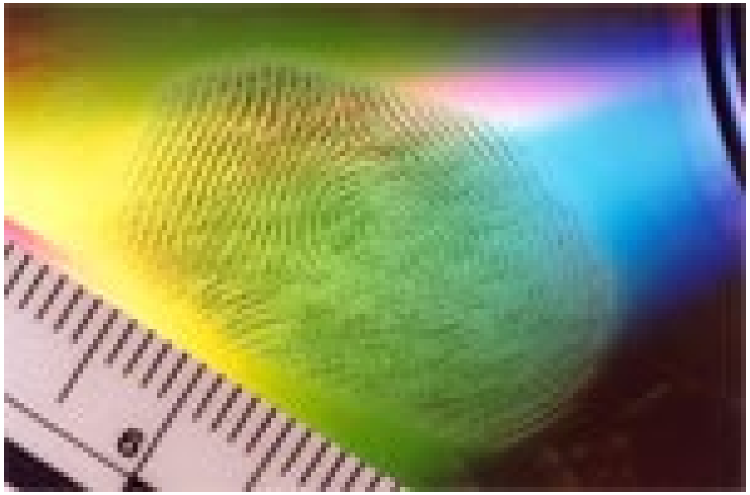
(1)-(3) cause burst errors

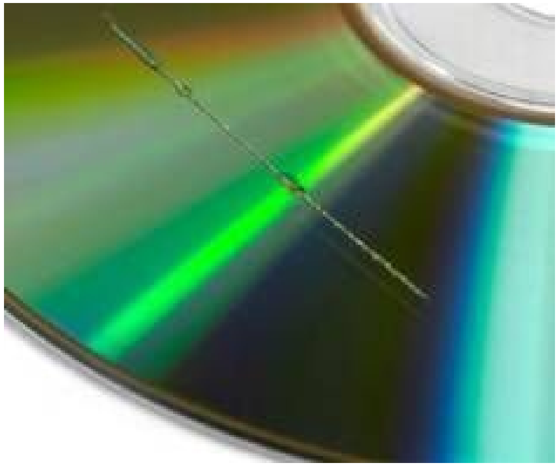
(4) and (5) cause random errors

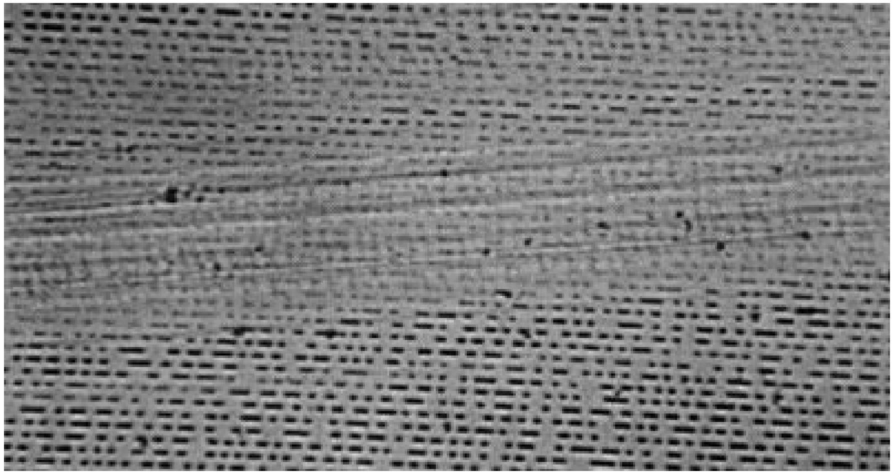
Causes of Disc Errors

- Fingerprints cause 43% of errors
- General wear and tear causes 25% of errors
- Player-related issues cause 15% of errors
- User-related issues cause 12% of errors
- Manufacturing defects cause 2% of errors

Causes of Disc Errors







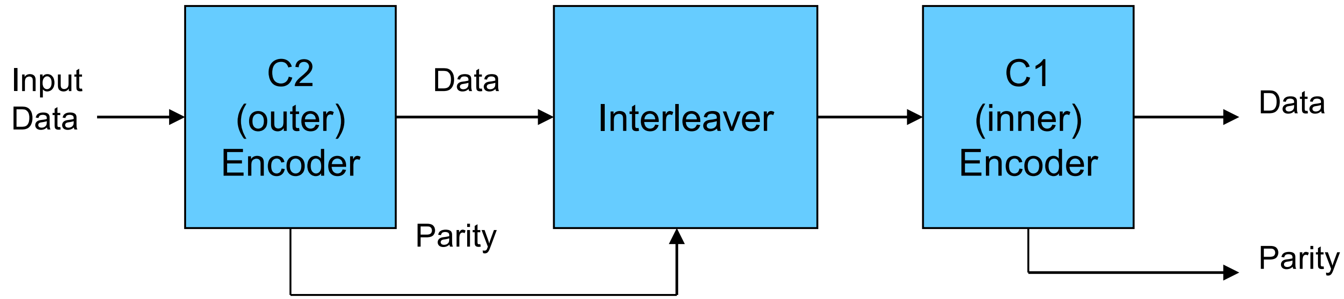
Error Correction

- Reed-Solomon code
  - (255,251,5) code over GF(2^8)
- Shortened to a (28,24,5) outer code
- These codewords are interleaved to reduce the effects of burst errors
- (32,28,5) inner code
- Overall code rate is

$$\frac{24}{28} \times \frac{28}{32} = 0.75$$

CIRC Encoder

- CIRC – Cross Interleaved Reed-Solomon Code



- Interleaving disperses the codewords so they are not contiguous on the disc
- mitigates long burst errors associated with scratches and fingerprints
  - Maximum correctable burst error length
    - 3874 bits ≈ 2.5 mm

Encoding Algorithm

- Samples are split into two 8 bit symbols
- Six samples from each channel are grouped to obtain 24 symbols
- Four outer RS code parity symbols are generated to give a frame of 28 symbols
- Symbols are interleaved over 109 frames
- Four inner RS parity symbols are generated to give 32 symbols
- These frames are also interleaved

Control and Error Correction

- Skips are caused by physical disturbances
  - Wait for disturbance to subside
  - Retry
- Read errors caused by disc/servo problems
  - Detect error
  - Choose location for retry
  - Retry, if it fails interpolate if applicable

Interpolation

- Used when decoding fails
- Fill missing audio data using adjacent data
  - time or channel
- Only valid for audio CDs

Decoding Reed-Solomon Codes

- $c(x)$ is the transmitted codeword
- $2t$ consecutive powers of α are roots

$$c(\alpha^{b}) = c(\alpha^{b+1}) = \cdots = c(\alpha^{b+2t-1}) = 0$$

- The received word is $r(x) = c(x)+e(x)$
- The error polynomial is

$$e(x) = e_0+e_1x+ \ldots +e_{n-1}x^{n-1}$$

- The syndromes are

$$S_j = r(\alpha^j) = e(\alpha^j) = \sum_{k=0}^{n-1} e_k(\alpha^j)^k, \qquad j = 1,\cdots,2t$$

Decoding Reed-Solomon Codes

- Suppose there are v errors in locations

$$i_1, i_2, \cdots, i_\nu$$

- The syndromes can be expressed in terms of these error locations

$$S_j = \sum_{l=1}^{v} e_{i_l} (\alpha^j)^{i_l} = \sum_{l=1}^{v} e_{i_l} (\alpha^{i_l})^j = \sum_{l=1}^{v} e_{i_l} X_l^j, \qquad j = 1, \cdots, 2t$$

- The X_l are the error locators
- The $2t$ syndrome equations can be expanded in terms of the v unknown error locations

GF(8) formed from x^3+x^2+1

Power of α	Polynomial in α	Vector
$-\infty$	0	000
0	1	100
1	α	010
2	α^2	001
3	α^2+1	101
4	$\alpha^2+\alpha+1$	111
5	$\alpha+1$	110
6	$\alpha^2+\alpha$	011

$2t$ Equations in $2v$ Unknowns

$$S_1 = e_{i_1} X_1 + e_{i_2} X_2 + \cdots + e_{i_v} X_v$$

$$S_2 = e_{i_1} X_1^2 + e_{i_2} X_2^2 + \cdots + e_{i_v} X_v^2$$

$$S_3 = e_{i_1} X_1^3 + e_{i_2} X_2^3 + \cdots + e_{i_v} X_v^3$$

$$\vdots$$

$$S_{2t} = e_{i_1} X_1^{2t} + e_{i_2} X_2^{2t} + \cdots + e_{i_v} X_v^{2t}$$

The Error Locator Polynomial

- The error locator polynomial $\Lambda(x)$ has as its roots the inverses of the v error locators $\{X_l\}$

$$\Lambda(x) = \prod_{l=1}^{v}(1 - X_l x) = \Lambda_v x^v + \ldots + \Lambda_1 x + \Lambda_0$$

- The roots of $\Lambda(x)$ are then $X_1^{-1}, X_2^{-1}, \ldots, X_v^{-1}$

$$\Lambda(X_l^{-1}) = \Lambda_v X_l^{-v} + \ldots + \Lambda_1 X_l^{-1} + \Lambda_0 = 0$$

$$e_{i_l} X_l^j \left(\Lambda_v X_l^{-v} + \ldots + \Lambda_1 X_l^{-1} + \Lambda_0\right) = 0$$

$$\Lambda(X_l^{-1}) = \Lambda_v X_l^{-v} + \ldots + \Lambda_1 X_l^{-1} + \Lambda_0 = 0$$

$$e_{i_l} X_l^j \left(\Lambda_v X_l^{-v} + \ldots + \Lambda_1 X_l^{-1} + \Lambda_0 \right) = 0$$

$$e_{i_l} \left(\Lambda_v X_l^{j-v} + \ldots + \Lambda_1 X_l^{j-1} + \Lambda_0 X_l^j \right) = 0$$

$$\sum_{l=1}^{v} e_{i_l} \left(\Lambda_v X_l^{j-v} + \ldots + \Lambda_1 X_l^{j-1} + \Lambda_0 X_l^j \right)$$

$$= \Lambda_v \sum_{l=1}^{v} e_{i_l} X_l^{j-v} + \ldots + \Lambda_1 \sum_{l=1}^{v} e_{i_l} X_l^{j-1} + \Lambda_0 \sum_{l=1}^{v} e_{i_l} X_l^j$$

$$= \Lambda_v S_{j-v} + \ldots + \Lambda_1 S_{j-1} + \Lambda_0 S_j = 0$$

$$\Lambda_v S_{j-v} + \ldots + \Lambda_1 S_{j-1} = -S_j$$

$$\begin{bmatrix} S_1 & S_2 & S_3 & S_4 & \cdots & S_{t-1} & S_t \\ S_2 & S_3 & S_4 & S_5 & \cdots & S_t & S_{t+1} \\ S_3 & S_4 & S_5 & S_6 & \cdots & S_{t+1} & S_{t+2} \\ S_4 & S_5 & S_6 & S_7 & \cdots & S_{t+2} & S_{t+3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ S_{t-1} & S_t & S_{t+1} & S_{t+2} & \cdots & S_{2t-3} & S_{2t-2} \\ S_t & S_{t+1} & S_{t+2} & S_{t+3} & \cdots & S_{2t-2} & S_{2t-1} \end{bmatrix} \begin{bmatrix} \Lambda_t \\ \Lambda_{t-1} \\ \Lambda_{t-2} \\ \Lambda_{t-3} \\ \vdots \\ \Lambda_2 \\ \Lambda_1 \end{bmatrix} = \begin{bmatrix} -S_{t+1} \\ -S_{t+2} \\ -S_{t+3} \\ -S_{t+4} \\ \vdots \\ -S_{2t-1} \\ -S_{2t} \end{bmatrix}$$

$$\mathbf{A}'\boldsymbol{\Lambda} = \mathbf{S}$$

Berlekamp-Massey Algorithm

$$\Lambda_v S_{j-v} + \ldots + \Lambda_1 S_{j-1} = -S_j$$

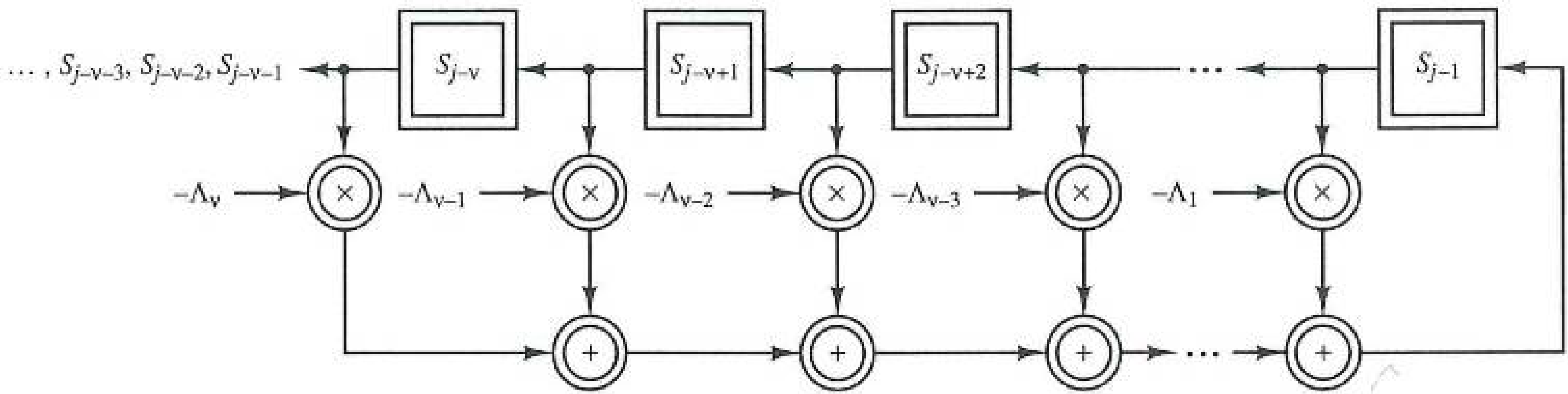


Figure 9-3. LFSR Interpretation of Eq. (9-27)

Berlekamp-Massey Algorithm

1. Compute the syndromes $S_1, S_2, \ldots, S_{2t}$ for the received word.
2. Set $k = 0$, $\Lambda^{(0)}(x) = 1$, $L = 0$, $T(x) = x$
3. Set $k = k + 1$. Compute the discrepancy $\Delta^{(k)} = S_k - \sum_{i=1}^{L} \Lambda_i^{(k-1)} S_{k-i}$
4. If $\Delta^{(k)} = 0$, go to step 8.
5. Modify the connection polynomial $\Lambda^{(k)} = \Lambda^{(k-1)}(x) - \Delta^{(k)} T(x)$
6. If $2L \geq k$, go to step 8.
7. Set $L = k - L$ and $T(x) = \Lambda^{(k-1)}(x)/\Delta^{(k)}$
8. Set $T(x) = xT(x)$
9. If $k < 2t$, go to step 3.
10. Determine the roots of $\Lambda(x) = \Lambda^{(2t)}(x)$. If the roots are distinct and lie in the right field, then determine the error magnitudes, correct the corresponding locations in the received word, and STOP.
11. Declare a decoding failure and STOP.

(7,3,5) Reed-Solomon Code

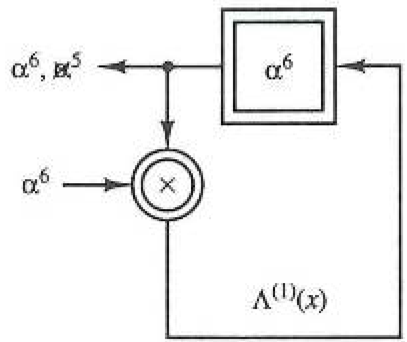
α^0	1
α^1	α
α^2	α^2
α^3	$\alpha+1$
α^4	$\alpha^2+\alpha$
α^5	$\alpha^2+\alpha+1$
α^6	α^2+1

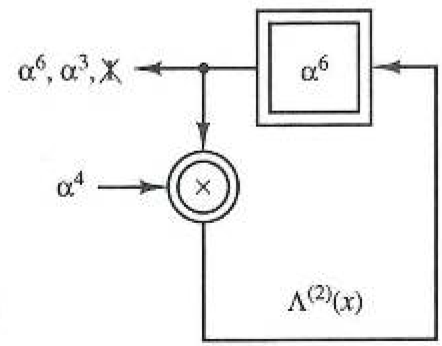
$$g(x)=(x-\alpha)(x-\alpha^2)(x-\alpha^3)(x-\alpha^4)$$

$$=x^4+\alpha^3x^3+x^2+\alpha x+\alpha^3$$

$$r(x)=\alpha^2x^6+\alpha^2x^4+x^3+\alpha^5x^2$$

Berlekamp-Massey Algorithm

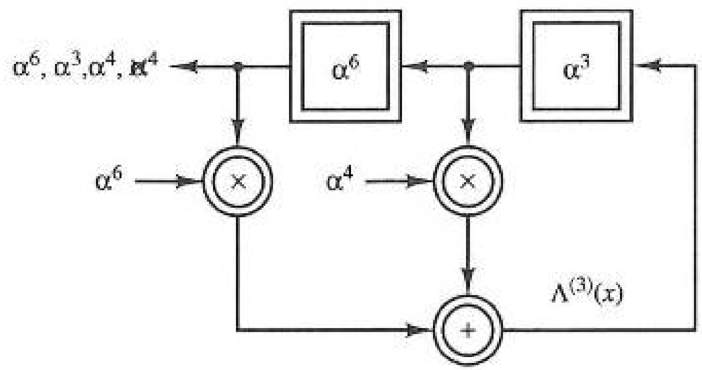


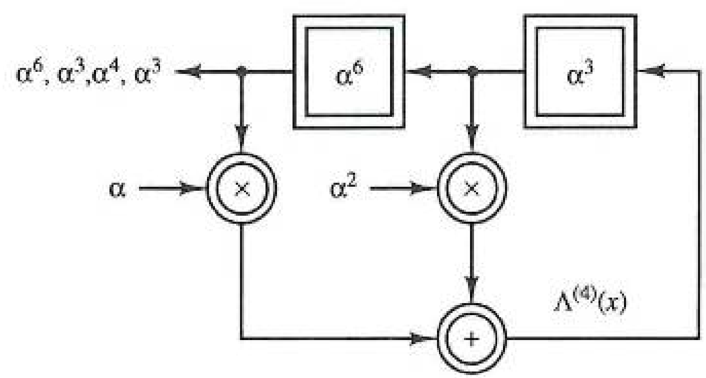


Berlekamp-Massey Algorithm

k	S_k	$\Lambda^{(k)}(x)$	$\Delta^{(k)}(x)$	L	$T(x)$
0	–	1	–	0	x
1	α^6	$1+\alpha^6 x$	S_1-0 = α^6	1	αx
2	α^3	$1+\alpha^4 x$	S_2-α^5 = α^2	1	αx^2
3	α^4	$1+\alpha^4 x+\alpha^6 x^2$	S_3-1 = α^5	2	$\alpha^2 x+\alpha^6 x^2$
4	α^3	$1+\alpha^2 x+\alpha x^2$	S_4-α^4 = α^6	–	–

Berlekamp-Massey Algorithm





Decoding Reed-Solomon Codes

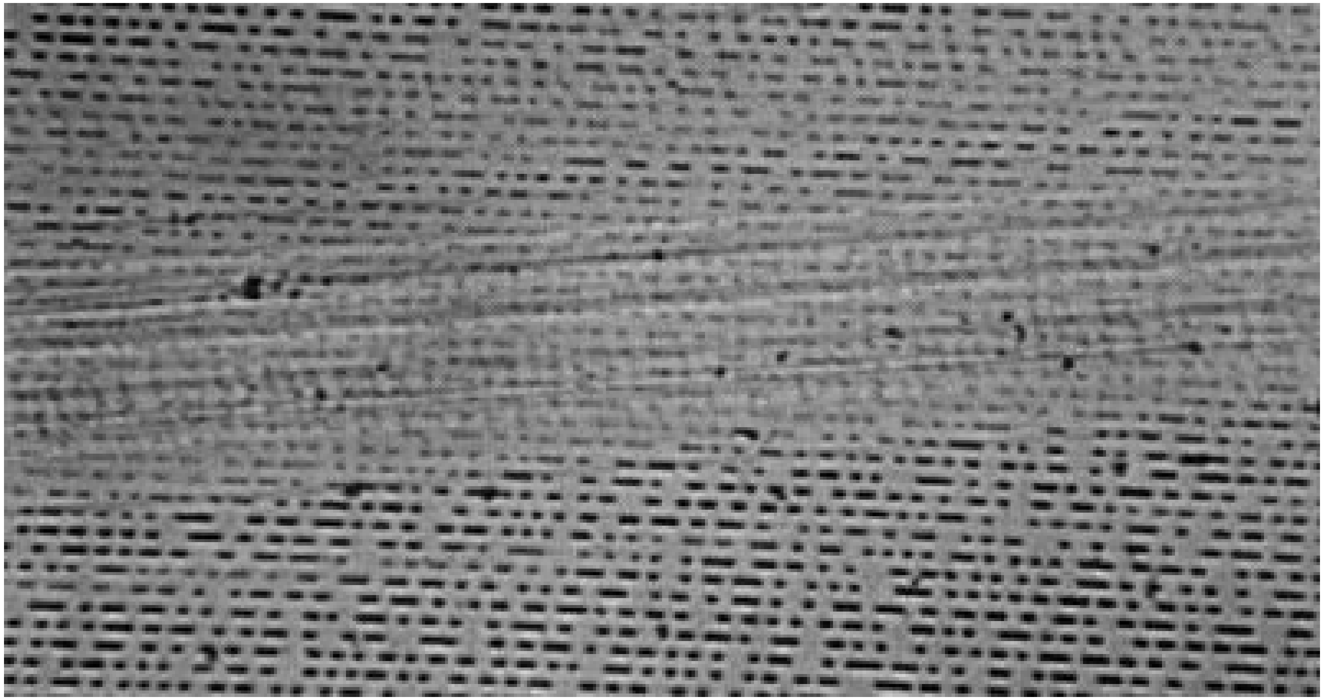
1. Compute the syndromes
2. Determine the error locator polynomial $\Lambda(x)$
3. Determine the error magnitudes from $\Lambda'(x)$ and $\Omega(x)$

$$\Omega(x) = [1 + S(x)]\Lambda(x) \bmod x^{2t+1}$$

$$e_{i_k} = \frac{-X_k \Omega(X_k^{-1})}{\Lambda'(X_k^{-1})}$$

4. Evaluate the error locations and the error values at those locations

CD Errors due to a Ball Point Pen



A Highly Corroded Disc

- Two minutes can still be played.

Audio Data Format