Construction d'un code binaire RS

- *Canal binaire symétrique* (BSC) avec un taux τ d'erreurs de bits. Exemple: τ=1%

  On suppose m fixé: V=GF(2^m); un chiffre de V = m bits. (parfois m=8, 16, 32 etc)
  p=Prob (erreur transmission d'un chiffre de m bits) = $1-(1-\tau)^m$.

- Un code RS(n,k) sur V est de rendement n/k et corrige t=(n-k)/2 erreurs de chiffres. La probabilité ε d'une erreur non corrigée (***erreur résiduelle***) est :

$$\varepsilon \leq \sum_{i=t+1}^{n} C_n^i p^i (1-p)^{n-i} = \sum_{j=0}^{n-t-1} C_n^j p^{n-j} (1-p)^j$$

  - *Définition*: C = **Capacité** BSC(τ) = $1+\tau.\log_2 \tau + (1-\tau).\log_2(1-\tau)$
    Théorème de Shannon :
    pour tout R < C, il existe un code de rendement R et d'erreur résiduelle arbitrairement petite.

  - Exemple: τ=1% => capacité = 0,91… => on cherche un code correcteur de rendement proche de 91% et d'erreur résiduelle faible.

Le choix de m et t définit le code RS … et impose R et ε

- Pour **τ=5.10^{-4} = 0,0005 :** rendement < capacité = 0,993

m	t	r	n	R	ε
8	2	4	255	0,984	0,083314104
8	4	8	255	0,969	0,003850848
8	8	16	255	0,937	1,16402E-06
8	12	24	255	0,906	6,0687E-11
8	16	32	255	0,875	9,02514E-16
8	32	64	255	0,749	1,01325E-38
16	500	1000	65535	0,985	0,830965135
16	600	1200	65535	0,982	0,00038756
16	700	1400	65535	0,979	4,80635E-14
16	800	1600	65535	0,976	4,27197E-30
16	1000	2000	65535	0,969	5,75845E-78

=> Code(255, 223) sur GF(256)
de rendement 87% et 16-correcteur
avec erreur résiduelle = 10^{-15}

NB Chiffres de 8 bits:
bloc de 1784 bits codés par 2040 bits

- Pour **τ=1% = 0,01 :** rendement < capacité = 0,91

m	t	r	n	R	ε
8	8	16	255	0,9373	0,998133479
8	16	32	255	0,8745	0,769179312
8	32	64	255	0,749	0,002590694
8	48	96	255	0,6235	3,08393E-09
8	64	128	255	0,498	6,64789E-18
8	80	160	255	0,3725	7,67393E-29
16	9000	18000	65535	0,7253	1
16	10000	20000	65535	0,6948	0,001808289
16	10500	21000	65535	0,6796	4,22996E-17
16	11000	22000	65535	0,6643	7,73789E-43

Compléments

- Codes raccourcis de Reed-Solomon sur GF(q)
  ∀n' < q : construction d'un code linéaire sur GF(q) de distance arbitraire d=n-k+1 (et maximale !)
- Correction de paquet d'erreurs :
  un code t-correcteur sur $GF(2^m)$ permet de corriger tout paquet de longueur ≤ m.(t-1)+1
- Correction d'effacements :
  Si toute sous-matrice de G formée de k colonnes est de rank k, il suffit de résoudre un système linéaire pour corriger
- Entrelacement :
  - Avec profondeur
  - Avec retard

IV Exemples de codes utilisés dans des matériels

Code du minitel

- Faible taux d'erreur, mais paquets longs :
  - Code de Hamming + detection paquets
- Source = suite de paquets de 15 octets = 120 bits
- Correction d'1 erreur : Code Hamming(127,120,3) :
  - 7 bits + 1 bit parité pour les 7 bits contrôle = 1 octet
- Détection de paquets: 1octet avec que des 0
  - Si erreur détectée : ARQ
- Total : 17 octets, rendement = 15/17 = 88%

Codes réseaux informatiques

- CRC : cyclic redundancy check avec r bits de redondance
  Polynome générateur $g= \Sigma_i g_i.x^i$ de degré r sur F_2
- Source: x=k bits ⇔ m(x) polynome de degré k-1
  - Exemple: m=01101 $m(x) = x+x^2+x^4$
- Codage: $c = m.x^r - (m.x^r \bmod g) = m.x^r + (m.x^r \bmod g)$
  - On a c multiple de g
  - Redondance r bits ($g_r = 1$)
- Propriétés: CRC détecte :
  - erreurs de poids 1 si : $g_0 = 1$
  - erreurs de poids 2 si g(x) a un facteur avec au moins 3 termes
  - nombre impair d'erreurs si g a (x+1) comme facteur

CRC standards

- CRC-8 : x^8+x^2+x+1
- CRC-10 : $x^{10}+x^9+ x^5+x^4 x+1$
- CRC-12 : $x^{12}+x^{11}+ x^3+x^2+1$
- CRC-16 : $x^{16}+x^{15}+ x^2+1$
- CRC-CCITT : $x^{16}+x^{12}+ x^5+1$
- CRC-32 : $x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+ x^8+x^7+x^5+x^4+ x^2+x+1$

Satellites : ex. Voyager

- Photos de Saturne et Jupiter (1977)
  - Données peu critiques (images 800x800 - 8 bits)
  - Données critiques: GSE (General Science&Engineering)
    - Mesures
    - Contrôle
- Données GSE codées par Golay(24,12) sur F_2 (6-correcteur)
- Autres données: code convolutif

CDROM / DVDROM (1/3)

- Codes par bloc cycliques
- Exemple : CD Audio :
  - Données = octets
  - K=trame = 24 octets codés sur 32 octets sur le CD
  - CIRC = Code de Reed-Solomon entrelacé croisé
    - Code (32, 24) = entrelacement de 2 codes cycliques
  - Base:
    - code de Reed-Solomon (255, 251, 5)
    - Code C1(28,24, 5)
    - Code C2(32,28,5)
  - Codes en détail
  - Description de l'entrelacement

CDROM : code CIRC (2/3)

Le code C_1

Il s'agit d'un code (28,24,5) sur le corps à 256 éléments.

Si x est un mot de 24 octets le mot de code de C_1 qui lui correspond est le mot de 28 octets égal à $(x, x R_1^t)$, où la matrice de dimension (4,24) R_1 est définie par

$$R_1 = \begin{bmatrix} a^{6} & a^{192} & a^{142} & a^{159} & a^{99} & a^{88} & a^{104} & a^{144} & a^{55} & a^{180} & a^{174} & a^{101} & a^{111} & a^{118} & a^{169} & a^{107} & a^{132} & a^{25} & a^{167} & a^{239} & a^{168} & a^{188} & a^{1} \\ a^{45} & a^{108} & a^{248} & a^{131} & a^{64} & a^{221} & a^{100} & a^{235} & a^{147} & a^{45} & a^{198} & a^{21} & a^{228} & a^{186} & a^{231} & a^{56} & a^{68} & a^{81} & a^{46} & a^{32} & a^{60} & a^{225} & a^{13} \\ a^{50} & a^{52} & a^{59} & a^{132} & a^{186} & a^{81} & a^{128} & a^{126} & a^{133} & a^{32} & a^{213} & a^{195} & a^{43} & a^{198} & a^{194} & a^{13} & a^{167} & a^{167} & a^{252} & a^{61} & a^{3} & a^{12} & a^{6} \\ a^{42} & a^{136} & a^{153} & a^{93} & a^{82} & a^{98} & a^{138} & a^{49} & a^{174} & a^{168} & a^{95} & a^{105} & a^{112} & a^{163} & a^{101} & a^{126} & a^{19} & a^{161} & a^{233} & a^{162} & a^{182} & a^{105} & a^{3} \end{bmatrix}$$

Le code C_2

Il s'agit d'un code (32,28,5) sur le corps à 256 éléments.

Si x est un mot de 28 octets le mot de code de C_2 qui lui correspond est le mot de 32 octets égal à $(x, x R_2^{tr})$, où la matrice $(4,28) R_2$ est :

$$R_2 = (R_1 \mid R') \text{ où } R' = \begin{bmatrix} a^{232} & a^{98} & a^{54} & a^{174} \\ a^{167} & a^{211} & a^{180} & a^{143} \\ a^{24} & a^{41} & a^{188} & a^{164} \\ a^{92} & a^{48} & a^{168} & a^{67} \end{bmatrix}$$

CDROM : code CIRC (3/3)

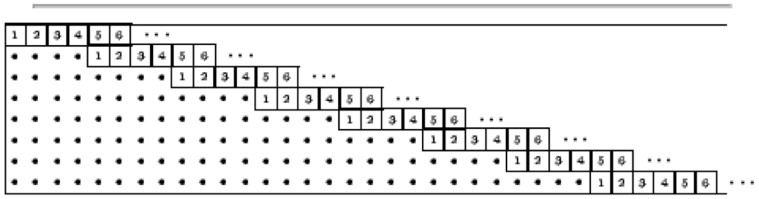

Figure 1: Table d'entrelacement à retard 4 de profondeur 8

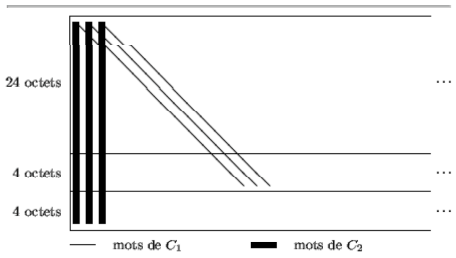

Figure 2: Schéma de codage

GSM

- Signal parole: par tranche de 20 ms
- Codec: Numérisation : 260 bits
  = 50 très critiques + 132 critiques + 78 complémentaires
- Codage de canal : 456 bits
  - » 50 bits : CRC (X3+x+1)
  - » 182+3 : codage convolutif : x2 : 378 bits
  - » +78= 4 bits controles :
- Entrelacement (diagonal : sur plusieurs trames de 456 bits)
- Chiffrement / modulation

Autres codes et applications

Codes Goppa(n,k,d) : définis par une courbe f(x,y) sur un GF

- n=nombre de points de f et d ≥ n-deg(f); ex: $x^3y+y^3+x=0$ ➔ Goppa(24,3,20)
- Pour l'instant difficiles à rendre efficaces mais très grande distance minimale

Turbo codes [C. Berrou, A. Glavieux] : codes de convolution

- fonctionnent seulement sur un petit alphabet (n et k ≤ 8)
- Décodage est complexe mais la correction est très grande

Cryptographie : système McEliece : code linéaire 2^{500} mots

- Problème à sens unique : trouver un mot de poids minimal
  - Facile si la matrice génératrice normalisée est connue
  - Difficile (énumérer tous les mots !) si la matrice est mélangée
  - Clef publique : matrice génératrice permutée
  - Clef privée : matrice génératrice normalisée