TD

Le but de ce TD est d'étudier les codes correcteurs de Reed-Solomon.
Cette grande famille de codes sert notamment en pratique pour les CD et DVD, certaines transmissions type ADSL ou satellites, des sondes spatiales, les QR codes.
Une idée majeure de ces codes est d'exploiter, en plus de la structure linéaire, des résultats sur les polynômes. Ces codes linéaires, sur un corps fini $\mathbb{F}_q$, sont de distance maximale ($d = n - k + 1$) et il existe des algorithmes de décodage assez rapides (en temps polynomial) ne demandant pas de stocker des tables de syndromes.

I. Exemples et premières définitions

1.1 Premiers exemples

1) On définit un code sur $\mathbb{F}_7$ par sa matrice de contrôle :
$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$.
Quels sont les trois paramètres de ce code ?

2) On définit un code $[6, 3, 4]$ sur $\mathbb{F}_7$ par sa matrice de contrôle :
$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1^2 & 2^2 & 3^2 & 4^2 & 5^2 & 6^2 \end{pmatrix}$.
a) Soient α, β et γ 3 éléments 2 à 2 distincts dans un corps. Vérifier que la matrice V suivante (dite de Vandermonde) est inversible.
$V = \begin{pmatrix} 1 & \alpha & \alpha^2 \\ 1 & \beta & \beta^2 \\ 1 & \gamma & \gamma^2 \end{pmatrix}$
b) En déduire que la distance du code défini ci-dessus est bien 4.

1.2 Définition

Soit $\mathbb{F}_q$ un corps fini, $n \in \mathbb{N}^*$ et k un entier compris entre 0 et n.
Soit $\alpha_1, \ldots, \alpha_n$ dans $\mathbb{F}_q^\times$ 2 à 2 distincts.
Soit $v_1, \ldots, v_n$ dans $\mathbb{F}_q^\times$.
Un code de Reed-Solomon généralisé $[n, k, d]$ sur $\mathbb{F}_q$ est un code linéaire dont une matrice de contrôle est $V_{n-k,n}^{(\alpha)} \cdot D_n^{(v)}$ où $V_{n-k,n}^{(\alpha)} = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{n-k-1} & \alpha_2^{n-k-1} & \ldots & \alpha_n^{n-k-1} \end{pmatrix}$ et $D_n^{(v)} = \begin{pmatrix} v_1 & & & \\ & v_2 & & \\ & & \ldots & \\ & & & v_n \end{pmatrix}$.
On appelle les α_i les localisateurs du code et les v_i les multiplicateurs du code.
Si $n = q - 1$, le code est dit primitif. Si $\forall i, v_i = 1$, le code est dit normalisé.
On note GRS (generalised Reed-Solomon) un tel code.

1.3 Définition

Un code RS est un GRS tel que n divise $q-1$ et tel qu'il existe $\alpha \in \mathbb{F}_q$ d'ordre multiplicatif n vérifiant pour tout i entre 1 et n : $\alpha_i = \alpha^{i-1}$ et $v_i = \alpha^{b(i-1)}$ (pour un entier b fixé). On note parfois $\mathrm{RS}(n, k)$ un code RS de paramètres n et k.

1.4 Exemples
- L'ADSL peut utiliser des RS(240, 224) ou RS(255, 239).
- Les CD et les DVD utilisent deux RS différents chacun. Par exemple pour les DVD, un RS(182, 172) et un RS(208, 192). Ceci permet de gérer aussi bien des effacements de données (ou une succession d'erreurs) que les erreurs réparties un peu partout.

1.5 Exercice
- Pour un code GRS, donner le coefficient à la position i, j de $V_{n-k,n}^{(\alpha)} \cdot D_n^{(v)}$.
- Pour un code RS, écrire la matrice de contrôle.

1.6 Exemples
1) Sur $\mathbb{F}_7$, avec $n = 6$ et $k = 4$, on peut prendre $\alpha_i = i$ et $v_i = 1$ pour obtenir un code GRS.
2) Sur $\mathbb{F}_8 = \mathbb{F}_2[X]/(X^3 + X + 1)$, avec $n = 7$ et $k = 4$, on obtient un code RS en posant $\alpha = \overline{X}$ (on rappelle que $\overline{X}$ est d'ordre 7) et $b = 0$.

II. Propriétés des codes GRS et RS

2.1 Propriété
Un code GRS vérifie $d = n - k + 1$. On appelle de tels codes "MDS" (maximable distance separable).

Preuve
- Ecrire la forme d'une matrice de taille $n - k \times n - k$ extraite de la matrice $V_{n-k,n}^{(\alpha)}$.
- Calculer ce déterminant (par récurrence), et montrer qu'il est de la forme $\prod_{i \neq j}(\beta_i - \beta_j)$.
- En déduire le résultat.

2.2 Propriété
Le code dual d'un code GRS est encore GRS, pour les mêmes localisateurs.

Preuve
L'idée est de chercher une matrice G (génératrice de C, donc de contrôle de $C^\perp$), pour les localisateurs α_i de C mais avec des multiplicateurs v'_i.
- Ecrire la relation vérifiée entre G et H, et écrire la relation obtenue entre la $i$ème ligne de G et la $j$ème ligne de H.
- Interpréter la famille de relations ci-dessus comme étant $V_{n-1,n}^{(\alpha)} \cdot D_n^{(v)} \cdot {}^t(v'_1, \ldots, v'_n) = 0$.
- En utilisant la propriété précédente, en déduire qu'il existe des multiplicateurs v'_i tous non nuls.

Remarque
En fait, on pourrait de plus prouver qu'un choix possible de v'_i vérifie $v_i v'_i \prod_{j \neq i}(\alpha_i - \alpha_j) = 1$, via l'interprétation polynomiale ci-dessous et des polynômes d'interpolation de Lagrange.

2.3 Interprétation polynomiale
Soit C un code GRS de paramètres $[n, k, n - k + 1]$, de localisateurs α_i et de multiplicateurs v_i.
En utilisant une matrice génératrice de C, montrer que :

$$C = \{(v'_1 P(\alpha_1), \ldots, v'_n P(\alpha_n)) \ ; \ P \in \mathbb{F}_q[X] \text{ où } degP \leq k - 1\}.$$

Grâce à cette écriture du code, retrouver la propriété qu'il est MDS.

2.4 Exemple

Soit C le code $[6,2,5]$ sur $\mathbb{F}_7$ défini par $G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$

et ayant comme matrice de contrôle $H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1^2 & 2^2 & 3^2 & 4^2 & 5^2 & 6^2 \\ 1^3 & 2^3 & 3^3 & 4^3 & 5^2 & 6^3 \end{pmatrix} \cdot \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \dots & \\ & & & 6 \end{pmatrix}$.

Expliciter l'interprétation polynomiale de ce code : Pour un signal initial $m = (m_0, m_1)$, quel est le message envoyé ?

III. Décodage

3.1 Algorithme de Berlekamp-Welch

Soit C un code GRS de paramètres $[n, k, n-k+1]$, de localisateurs α_i, et dont le code dual a des multiplicateurs $v_i' = 1$.

$$C = \{(P(\alpha_1), \dots, P(\alpha_n)) \ ; \ P \in \mathbb{F}_q[X] \text{ où } degP \leq k-1\}.$$

On suppose d impair et $t = \frac{d-1}{2}$. Soit $c = (P(\alpha_1), \dots, P(\alpha_n))$ et $r \in \mathbb{F}_q^n$ tels que $d_H(c, r) \leq t$.
On voit r comme un message reçu, avec un nombre d'erreurs tel qu'il soit possible de retrouver c. En pratique, on cherche P.
Notons $I = \{i \in \{1, \dots, n\} \ ; \ c_i \neq r_i\}$. On a donc $|I| \leq t$ et $\forall i \in \{1, \dots, n\} \setminus I, P(\alpha_i) = r_i$.
Les r_i et α_i sont les valeurs connues, l'inconnue est P, de degré $\leq k-1$.
Pour trouver P, on va construire deux polynômes Q_0 et Q_1 (de degrés respectifs $< n-t$ et $< t+1$) tels que

$$\forall i \in \{1, \dots, n\}, Q_0(\alpha_i) = r_i Q_1(\alpha_i)$$

puis ces deux polynômes nous donneront P.

a) En considérant les équations satisfaites par Q_0 et Q_1, montrer qu'il existe deux tels polynômes non nuls.
b) On considère maintenant le polynôme $\phi \in \mathbb{F}_q[X]$ défini par $\phi(X) = Q_0(X) - P(X)Q_1(X)$. Montrer que ϕ est le polynôme nul, et déduire P.

3.2 Exemple de décodage

On reprend le code C du 2.4. Le message reçu est $y = (1, 2, 6, 1, 4, 3)$. Expliciter le système à résoudre pour décoder ce message.
On admet que pour ces valeurs r_i, on trouve $Q_0(X) = 5 + 6X + 6X^2 + X^3$ et $Q_1(X) = 6 + 6X + 6X^2$.
En déduire P puis le message initialement envoyé.

3.3 Algorithme de Gao

On considère le polynôme $F(X) = \prod_{i=1}^{n}(X - \alpha_i)$ et R le polynôme de degré $< n$ vérifiant $R(\alpha_i) = r_i$ (interpolation de Lagrange).
Montrer que l'ensemble des équations $Q_0(\alpha_i) = r_i Q_1(\alpha_i)$ est équivalent à

$$Q_0 \equiv RQ_1 \mod F.$$

Il est ensuite possible de montrer que les polynômes Q_0 et Q_1 peuvent se calculer par un algorithme d'Euclide étendu (ce n'est pas du tout immédiat à voir).