# Feuille d'exercices 2 bis

*Remarque:* Tous les exercices ne seront pas traités en séance de TD, j'indiquerai au fur et à mesure sur la page du forum (http://cours-jussieu-nombres.monforum.com/cours-et-td-2007-vf6.html) les exercices que nous aurons abordés.

## 1 Codes correcteurs

Pour transmettre une information on utilise l'alphabet $\mathbb{F}_q$; on envoie des messages de $n$ lettres. Le principe des codes correcteurs d'erreurs est de pouvoir corriger des erreurs de transmission (cf. les CD, les transmissions par satellite...). L'ensemble des mots $\mathbb{F}_q^n$ peut être muni de la distance de Hamming définie comme suit: pour $(x_1, \cdots, x_n)$ et $(x'_1, \cdots, x'_n)$ dans $\mathbb{F}_q^n$ alors

$$d(x, x') := \text{card}\{i \in [1, n] \ / \ x_i \neq x'_i\}$$

On vérifie aisément qu'il s'agit bien d'une distance.

Un code est un sous-ensemble $\mathcal{C} \subset \mathbb{F}_q^n$ comportant au moins deux éléments de $\mathbb{F}_q^n$; on définit la distance d'un code comme

$$d(\mathcal{C}) := \min_{x \neq x' \in \mathcal{C}} d(x, x').$$

Le principe consiste, une fois choisi un code $\mathcal{C}$, à n'envoyer que des messages avec des mots appartenant à $\mathcal{C}$; on peut alors repérer jusqu'à $d(\mathcal{C}) - 1$ erreurs de transmission sur un mot en outre si le nombre d'erreurs commises $t$ est tel que $2t + 1 \leq d(\mathcal{C})$, on voit qu'il existe un seul mot de $\mathcal{C}$ situé à une distance $\leq t$ du mot reçu. Le code permet donc de corriger $t := \lfloor \frac{d(\mathcal{C})-1}{2} \rfloor$ erreurs. On introduit le taux de corrections $\frac{t}{n}$ et le taux d'information $\log \text{card}(\mathcal{C})/n \log q$. La théorie de l'information développée par Shannon, indique que si l'on accepte d'envoyer des messages de plus en plus long, il existe des codes aussi sûrs que l'on veut avec un taux d'information proche de 1: cependant le théorème de Shannon est un théorème d'existence, il ne dit pas comment construire les codes en question.

**Exercice 1.** *On considère des codes linéaires, i.e. $\mathcal{C} \subset \mathbb{F}_q^n$ est un sous-espace vectoriel. Pour tout $x \in \mathcal{C}$, on définit son poids $\omega(x)$ comme le nombre de composantes non nulles.*

*(1) Montrer que $d(\mathcal{C}) = \min_{0 \neq x \in \mathcal{C}} \min \omega(x)$.*

*(2) **exemple du bit de parité:** pour transmettre $(x_1, \cdots, x_{n-1}) \in \mathbb{F}_2^{n-1}$ on envoie $x = (x_1, \cdots, x_{n-1}, x_1 + \cdots + x_{n-1}) \in \mathbb{F}_2^n$. Montrer qu'il s'agit d'un code cyclique qui permet de repérer une erreur mais pas de la corriger.*

*(3) **Code de Hamming:** prenons l'ensemble des mots de sept chiffres binaires, $q = 2$, $n = 7$ et $\mathcal{C}$ le code ayant pour base*

$$e_0 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix},$$

*Pour transmettre un message $m = (m_0, m_1, m_2, m_3)$, on transmet $x = m_0e_0 + m_1e_1 + m_2e_2 + m_3e_3$. Expliquez le décodage dans le cas où une erreur au plus est commise.*

*(4) Une matrice génératrice $A$ d'un code $\mathcal{C}$ est une matrice dont les lignes forment une base. Une matrice vérificatrice $B$ d'un code $\mathcal{C}$ est une matrice dont les lignes forment une base des formes linéaires s'annulant sur $\mathcal{C}$. Montrer que $A^tB = 0$ et que la distance du code $\mathcal{C}$ est le plus petit nombre $d$ tel qu'il existe $d$ vecteurs colonnes de $B$ distincts et liés.*

(5) *Supposons un code $\mathcal{C}$ donné avec une matrice vérificatrice $B$ et supposons que le code est 1-correcteur. Soit alors un message $x'$ reçu différant du message envoyé $x$ en au plus une coordonnée: on note $\epsilon = x' - x$ l'erreur commise. Montrer comment calculer $\epsilon$ à l'aide de $B$.*

(6) *Soit $\mathcal{C}$ un code de longueur $n$ sur $\mathbb{F}_q$. Donnez la distance et des matrices génératrices et vérificatrices des codes suivants:*

*(i)* ***Code raccourci****: soit $d(\mathcal{C}) \leq l \leq n$, on pose $\mathcal{C}^{(l)} := \{x \in \mathbb{F}_q^l \ / \ (x; 0, \cdots, 0) \in \mathcal{C}\}$.*

*(ii)* ***Code étendu****: $\bar{\mathcal{C}} := \{(x_1, \cdots, x_{n+1}) \in \mathbb{F}_q^{n+1} \ / \ (x_1, \cdots, x_n) \in \mathcal{C} \text{ et } x_1 + \cdots + x_{n+1} = 0\}$.*

*(iii)* ***Code dual****: $\mathcal{C}^* := \{x' \in \mathbb{F}_q^n \ / \ \forall x \in \mathcal{C}, \ < x, x' >= 0\}$ où $<,>$ est le produit scalaire canonique.*

(7) *Soit $\mathcal{C}$ un code de dimension $k$ et de longueur $n$ sur $\mathbb{F}_q$, montrer que $d(\mathcal{C}) \leq n + 1 - k$ et que si $\mathcal{C}$ est $t$-correcteur alors*

$$1 + C_n^1(q-1) + C_n^2(q-1)^2 + \cdots + C_n^t(q-1)^t \leq q^{n-k}$$

*Un code tel que $d(\mathcal{C}) = n + 1 - k$ sera dit MDS maximal distance separable. Un code $t$-correcteur tel que $\mathcal{C} = \bigcup_{x \in \mathcal{C}} B(x, t)$ est dit $t$-correcteur parfait.*

*Montrer que le code de Hamming de longueur 7 est 1-correcteur parfait mais qu'il n'est pas MDS.*

**Exercice 2.** ***Codes linéaires cycliques*** *Un code linéaire cyclique est un code $\mathcal{C}$ linéaire de longueur $n$, stable par la permutation $T(a_1, \cdots, a_n) = (a_n, a_1, \cdots, a_{n-1})$.*

(1) *En utilisant l'isomorphisme naturel d'espace vectoriel $\psi : \mathbb{F}_q^n \simeq \mathbb{F}_q[X]/Q(X)$ où $Q = X^n - 1$, montrer que $\mathcal{C} \subset \mathbb{F}_q^n$ est stable par $T$ si et seulement si son image par $\psi$ est un idéal. En déduire alors qu'il existe une bijection entre les codes cycliques de longueur $n$ et les polynômes unitaires divisant $X^n - 1$.*

(2) *Rappeler la factorisation en irréductibles des polynômes cyclotomiques $\Phi_n$ dans $\mathbb{F}_q$, et en déduire une bijection entre les codes cycliques de longueur $n$ et les parties $I \subset (\mathbb{Z}/n\mathbb{Z})^\times$ stables par la multiplication par $q$.*

(3) *Soit $C$ un code linéaire cyclique de longueur $n$ sur $\mathbb{F}_q$ associé à $I \subset (\mathbb{Z}/n\mathbb{Z})^\times$ et supposons qu'il existe $i$ et $s$ tels que $\{i+1, i+2, \cdots, i+s\} \subset I$. Montrer alors que $d(\mathcal{C}) \geq s + 1$.*

(4) ***Codes de Hamming****: soit $n = \frac{q^r - 1}{q-1}$ et $I := \{1, q, q^2, \cdots, q^{r-1}\}$. Montrer que $d(\mathcal{C}) = 3$ ou $4$ et qu'il est parfait 1-correcteur.*

*Remarque: Pour $r = 3$, $q = 2$ et $n = 7$ on retrouve le code étudié précédemment.*

*En construisant une matrice vérificatrice montrer qu'en fait on a $d(\mathcal{C}) = 3$.*

(5) ***Codes de Reed-Solomon****: ce code est utilisé dans les CD. Soit $n = q - 1$ et soit $\alpha$ un générateur de $\mathbb{F}_q^\times$. Pour $k$ fixé on pose*

$$g(X) := \prod_{i=1}^{q-1-k} (X - \alpha^i)$$

*Montrer que le code linéaire cyclique correspondant est MDS et que pour $q = 2^f$, on a $2t + 1 \leq d(\mathcal{C}) = q - k$.*

(6) ***Code ternaire de Golay****: on a $3^5 - 1 = 11.23$; on choisit $q = 3$, $n = 11$ et la partie de$(\mathbb{Z}/11\mathbb{Z})^\times$ engendrée par 3, i.e. $i = \{1, 3, 4, 5, 9\}$. On note $\mathcal{G}_{11}$ le code linéaire cyclique correspondant. Montrer que $d(\mathcal{G}_{11}) = 4, 5$ puis que $\mathcal{G}_{11}$ est 2-correcteur parfait (il n'est pas MDS).*

(7) ***Code binaire de Golay****: on a $2^{11} - 1 = 23.89$, on choisit $q = 2$, $n = 23$ et $I = (2) \subset (\mathbb{Z}/23\mathbb{Z})^\times$. On note $\mathcal{G}_{23}$ le code linéaire cyclique correspondant. Montrer que $d(\mathcal{G}_{2^n}) = 5, 6, 7$ puis que $\mathcal{G}_{23}$ est "-correcteur parfait.*

**Exercice 3.** ***Code du minitel***

(a) *Montrez que le polynôme $X^7 + X^3 + 1$ est irréductible sur $\mathbb{F}_2$ et en déduire que $\mathbb{F}_{128} \simeq \mathbb{F}_2[X]/(X^7 + X^3 + 1)$ et montrez que $X$ est un générateur du groupe multiplicatif.*

(b) *Pour envoyer un message de* 15 *octets (soit 120 bits) de la forme* $M = a_0 a_1 \cdots a_{119}$ *où les* $a_i$ *sont des éléments de* $\mathbb{F}_2$ *(des bits), on considère l'élément suivant de* $\mathbb{F}_{128}$

$$\beta = a_0\alpha^{126} + \cdots + a_{119}\alpha^7 = a_{120}\alpha^6 + \cdots + a_{125}\alpha + a_{126}$$

*On envoie alors le message* $a_0 a_1 \cdots a_{126} a_{127}$ *où* $a_{127}$ *est un bit de parité, soit* 16 *octets. Le message reçu est* $a'_0 \cdots a'_{127}$ *où certains* $a'_i$ *sont distincts de* $a_i$ *à cause d'une erreur de transmission.*

(i) *Expliquez pour quoi si* $a'_0\alpha^{126} + \cdots + a'_{125}\alpha + a'_{126} = 0$ *alors il est raisonnable de penser qu'il n'y a pas eu d'erreurs de transmission. Quel est alors le message?*

(ii) *On suppose que les erreurs de transmissions sont suffisamment rares pour qu'au plus une erreur se soit produite, par exemple au bit* $k$*, i.e.* $a_i = a'_i$ *pour* $i \neq k$ *et* $a'_k = a_k + 1$*. Expliquez comment trouver* $k$ *et donc le message initial.*

(iii) *Commentez le choix de* 128*.*

**Exercice 4.** ***Les disques compacts***

(a) *Montrez que* $P(X) = X^8 + X^7 + X^6 + X^5 + X^4 + X^2 + 1$ *est irréductible sur* $\mathbb{F}_2$ *et en déduire que* $\mathbb{F}_{256} \simeq \mathbb{F}_2[X]/(P(X))$*. Montrez que* $\alpha$*, l'image de* $X$*, est un générateur du groupe multiplicatif.*

(b) *On représente un octet par un élément de* $\mathbb{F}_{256}$*. Considérons un mot* $M = a_0 \cdots a_{250}$ *constitué de* 251 *octets, i.e.* $a_i \in \mathbb{F}_{256}$*. On considère*

$$(\sum_{i=0}^{250} a_i X^i)(X+\alpha)(X+\alpha^2)(X+\alpha^3)(X+\alpha^4) = \sum_{i=0}^{254} b_i X^i$$

*et on transmet le message* $M = b_0 \cdots b_{255} b_{256}$ *où* $b_{256}$ *est un bit de parité.*

(i) *Supposons que deux erreurs au plus se produisent dans la lecture de* $M$*. Comment savoir s'il y a eu zéro, une ou deux erreurs et expliquez comment les corriger.*

(ii) *On suppose désormais que quatre octets quelconques de* $M$ *sont illisibles. Expliquez comment retrouver les bonnes valeurs.*

(iii) *Dans un CD, on code les informations musicales par paquets de* 24 *octets auquels on adjoint* 4 *octets comme précédemment afin de pouvoir corriger deux erreurs ou 4 effacements. On obtient ainsi des mots de* 28 *octets, dont le* $i$*ème mot est noté* $M_i$ *de* $k$*-ième octet est* $M_i(k)$*. Les mots sont alors entrelacés comme suit: chaque sillon est constitué de* 28 *octets, le* $i$*-ème sillon contient alors les octets suivants*

$$M_i(1)\ M_{i-4}(2)\ M_{i-8}(3) \cdots M_{i-108}(28)$$

*ou de manière équivalente* $M_i$ *est constitué de* $S_i(1)S_{i+4}(2)\cdots S_{i+108}(28)$*. Chaque sillon de* 28 *octets est complété de* 4 *octets comme précédemment. Expliquez comment nos lecteurs de CD se jouent des rayures (de 2mm de large).*