Théorie des codes correcteurs d'erreur
Notes de Cours- Prépa Agreg option C*

Thomas Hausberger

14 mars 2025

Table des matières

1. Rudiments de théorie de l'information

1.1. Modèle théorique

Canal bruité : un émetteur envoie un message M, codé en m ; une erreur e de transmission se produit ; au bout du canal bruité, on recueille m', décodé en M', qui est le message reçu. On suppose qu'il n'y a pas d'erreur de codage, ni de décodage, uniquement une erreur de transmission.

But : détecter les erreurs et les corriger, grâce à l'ajout d'une information supplémentaire lors du passage de M à m. Les applications des codes correcteurs d'erreur dans l'industrie concernent le disque compact, le Minitel, la transmission d'images par satellite,...

Représentation du canal (discret, sans mémoire) : c'est la donnée

- d'un alphabet d'entrée $A = \{\alpha_i\}$ et de sortie $B = \{\beta_j\}$
- d'une matrice stochastique de transmission $P = (P_{i,j})$ (avec $\sum_j P_{i,j} = 1\ \forall i$) : lors de la transmission d'un message $m = (m_1, \dots, m_n)$ qui est un mot de longueur n dans l'alphabet A, on reçoit un message $m' = (m'_1, \dots, m'_n)$ de même longueur avec $P(m'_k = \beta_j | m_k = \alpha_i) = P_{i,j}$ pour tout k.

Exemple. Le *canal symétrique binaire* : $A = B = \{0, 1\}$ et $P = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$ où la probabilité d'erreur est $p \ll 1$ pour 0 et 1.

1.2. Codage, décodage, correction des erreurs

Codage : le canal sera supposé symétrique, l'alphabet est $A = B = F$, de cardinal q. Le message M (resp. message codé m) est un mot de longueur k (resp. n) avec $k \leq n$.

- une *application de codage* est une application injective $E : F^k \to F^n$ (E comme encodage) et le code associé son image $C = E(F^k)$. On dit que C est de *dimension* k et de *longueur* n. Si E est l'identité sur les k premières composantes, *i.e.* $E(M) = m$ avec $m_i = M_i$ pour $i \leq k$, on dit que le codage est *systématique*.
- La *distance de Hamming* sur F^n est définie par $d(x, y) = \mathrm{Card}(\{i, x_i \neq y_i\})$. C'est bien une distance sur F^n (le vérifier). On appelle *distance minimale du code* $d(C) = \min_{(x,y)\in C^2, x\neq y} d(x, y)$ et on dit que C est *de type* (n, k, d).
- Le *taux d'information ou rendement* du code est $R = k/n$.

Décodage : on a

$$P(m' = b | m = a) = \prod_i P(m'_i = b_i | m_i = a_i) = p^d (1-p)^{n-d} = (\frac{p}{1-p})^d (1-p)^n$$

où $d = d(a, b)$ et n est la longueur. Cette probabilité décroît lorsque d croît. C'est pourquoi :

- une *application de décodage* est $D : F^n \to F^k$ tel que $D \circ E = \mathrm{Id}_{F^k}$ et $d(E \circ D(m'), m') = \min_{M\in F^k} d(E(M), m')$ pour tout $m' \in F^n$. Si on identifie $E(F^k)$ et F^k, cela signifie que $d(D(m'), m') = \min_{c\in C} d(c, m')$, *i.e.* $D(m')$ se trouve parmi les mots les plus proches de m'. On parle de *décodage selon le principe du maximum de vraissemblance* car les probabilités vont dans ce sens.

— On dit que C est t-correcteur si pour tout $m' \in F^n$ il existe au plus un mot $c \in C$ tel que $d(m', c) \leq t$. C'est équivalent à dire que les boules fermées $B(c, t)$, $c \in C$, sont disjointes. Si la transmission est telle que $d(m, m') \leq t$ pour tout m (*i.e.* il se produit au plus t erreurs), alors le code est t-correcteur si et seulement si on peut corriger t erreurs (*i.e.* retrouver m à partir de m' par une application de décodage qui est alors unique).
— Si les boules $B(c, t)$ forment une partition de F^n on dit que que le code t-correcteur est *parfait*.

Théorème 1.2.1. *Un code de distance minimale d est t-correcteur si et seulement si $d \geq 2t + 1$.*

On dit alors que $\lfloor \frac{d-1}{2} \rfloor$ (partie entière) est la *capacité théorique de correction* du code.

Démonstration. Si C n'est pas t-correcteur, donc s'il existe x dans $B(c, t) \cap B(c', t)$, alors $d(c, c') \leq d(c, x) + d(x, c') \leq 2t$. Réciproquement, s'il existe c et c' avec $d(c, c') \leq 2t$, on peut supposer $c' = (c'_1, \ldots, c'_{2t}, c_{2t+1}, \ldots, c_n)$ (pour faciliter les notations) ; alors $x = (c_1, \ldots, c_t, c'_{t+1}, \ldots, c'_{2t}, c_{2t+1}, \ldots, c_n)$ appartient à $B(c, t) \cap B(c', t)$ donc C n'est pas t-correcteur. □

1.3. Premiers exemples

On prend $F = \{0, 1\}$ (codes binaires).

Exemple (code de répétition pure). $C = \{(0, \ldots, 0), (1, \ldots, 1)\}$, de dimension 1 ; la distance minimale d est égale à la longueur n du code. Il est $\lfloor \frac{d-1}{2} \rfloor$-correcteur (décodage : algorithme majoritaire), mais le rendement $R = 1/n$ est faible.

Exemple (code de parité). A $M = (M_1, \ldots, M_n)$ on rajoute $\sum M_i \mod 2$. Il est de type $(n + 1, n, 2)$. Le rendement est $R = n/(n + 1)$ mais la capacité théorique de correction est nulle (ce code permet de détecter la présence d'une erreur si on sait qu'il y en a au plus une, mais pas de la corriger).

1.4. Théorème de Shannon

On a l'impression que la fiabilité du code (qui se mesure par la probabilité de commettre une erreur en décodant le message transmis à travers le canal probabiliste) ne peut croître qu'au détriment du rendement. Le théorème de Shannon dit que non : quitte à prendre n suffisamment grand, il existe des codes aussi fiables que l'on veut de rendement R arbitrairement bon (mais restant en-dessous de la « capacité » $c(P)$ du canal, quantité qui modélise le débit physique du canal de matrice de probabilités P).

Théorème 1.4.1 (Shannon, 1948)**.** *Les « bons » codes existent quel que soit le canal symétrique binaire.*

C'est un théorème d'existence abstrait (nous n'exposerons pas le formalisme de la théorie de l'information plus précisément ; pour satisfaire la curiosité, $c(p) = 1 + p \ln p + (1-p) \ln(1-p)$ dans le cas du canal symétrique binaire). Encore faut-il savoir construire

de tels codes. Un autre problème, dans la pratique, est de disposer d'un algorithme de décodage efficace.

Exemple. Dans le cas d'une transmission via un fil de cuivre, on mesure une probabilité d'erreur sur 1 bit environ égale à $p = 10^{-4}$. En utilisant un code de type $(15, 8, 5)$ (par exemple, un code BCH binaire de ce type, voir §4), la probabilité de commettre une erreur en décodant par la procédure ci-dessus, lorsque $m = a$ est fixé dans le code, est égale à $\sum_{j=3}^{15} \binom{15}{j} p^j (1-p)^{15-j}$. En effet, avec nos notations, il y a erreur si $m' \notin B(m, t)$, où $t = 2$. Or $\sum_{b \notin B(m,t)} P(m' = b) = \sum_{b \notin B(m,t)} p^{d(m,m')}(1-p)^{15-d(m,m')}$ d'après le calcul de $P(m' = b | m = a)$, indépendemment de a. Or il y a $\binom{15}{j}$ éléments m' tels que $d(m, m') = j$ (connaître le mot m' revient à connaître les j positions où m' diffère de m car le code est binaire). On trouve 5.10^{-10} avec un rendement $8/15$, ce qui est bien mieux que $1 - (1-p)^8 \approx 8.10^{-4}$, probabilité d'erreur sans codage.

1.5. Borne de Singleton

Proposition 1.5.1 (borne de Singleton)**.** *Soit C un (n, k, d)-code. On a $d \leq n - k + 1$.*

Démonstration. Soit $E' : F^k \xrightarrow{E} F^n \xrightarrow{p} F^{n-d+1}$ où p désigne la projection sur les $n-d+1$ premières composantes. Alors E' est injective car la distance minimale de C est d. Donc $k \leq n - d + 1$. □

C'est cohérent avec l'intuition que lorsque l'information supplémentaire diminue la capacité de correction diminue. Nous verrons que les codes de Reed-Solomon, par exemple, sont optimaux pour la borne de Singleton (l'inégalité est une égalité : on dit alors que le code est MDS, de l'anglais « maximum distance separable »). Il en est de même des codes cités en exemple plus haut (mais ces derniers sont peu utiles en pratique !)

2. Codes linéaires

L'idée est qu'en rajoutant de la structure on peut obtenir des algorithmes de codage et de décodage plus efficaces.

2.1. Définitions

L'alphabet $F = \mathbb{F}_q$ est un corps fini ($q = p^r$, p premier). Alors $m' = m + e$, où $m', m, e \in \mathbb{F}_q^n$.

- On définit $\omega(e) = d(e, 0) = d(m, m')$ le *poids de Hamming* de l'erreur.
- On demande que l'application de codage $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ soit linéaire ; noter que messages émis et mots du code sont des *vecteurs lignes*. Ainsi E est définie par sa matrice $G \in \mathrm{M}_{k,n}(\mathbb{F}_q)$ (dans les bases canoniques, les vecteurs étant rangés en ligne) appelée *matrice génératrice* du code. On a $C = \{uG, u \in \mathbb{F}_q^k\}$; noter le gain en terme d'information stoquée pour définir C. Si le codage est systématique, alors G est de la forme $(I_k | P)$.

— Comme E est injective, C s'identifie à un sous-espace vectoriel de dimension k de $\mathbb{F}_q^n$.
— La distance minimale du code est

$$d(C) = \min_{(x,y)\in C^2, x\neq y} d(x,y) = \min_{0\neq z\in C} \omega(z)$$

(car C est un espace vectoriel).

2.2. Matrice de contrôle, syndrôme

Le *code dual* $C^\perp$ du code linéaire C est par définition l'orthogonal de C dans l'espace vectoriel F^n pour le produit scalaire usuel $x \cdot y = \sum_{i=1}^n x_i y_i$. Il est donc de dimension $n-k$. On a

$$C^\perp = \{u \in F^n,\ u \cdot v = 0\ \forall v \in C\}.$$

Une *matrice de contrôle* (ou de parité) H du code C est par définition une matrice génératrice du code dual $C^\perp$, c'est-à-dire une matrice $H \in \mathrm{M}_{n-k,n}(F)$ telle que $C^\perp = \{wH, w \in F^{n-k}\}$.

On a de plus

$$C = (C^\perp)^\perp = \{v \in F^n,\ u \cdot v = 0\ \forall u \in C^\perp\}.$$

Comme les lignes L_i de H engendrent $C^\perp$, ceci est équivalent à $v\ {}^tL_i = 0$ pour $1 \leq i \leq n$, ou encore à $v\ {}^tH = 0$, donc à $H\ {}^tv = 0$. On prend parfois cette condition comme définition d'une matrice de contrôle (ce qui évite de parler de code dual, mais rend la définition moins naturelle).

La matrice ${}^tH$ définit une application linéaire $S : F^n \to F^{n-k}$ (par $S(x) = x\ {}^tH$) dont on a vu que C est le noyau. Elle est donc en particulier surjective (par le théorème du rang). On dit que S est l'*application syndrôme* du code C associée à H et que $S(x)$ est le *syndrôme* de x (associé à la matrice de contrôle H).

Si le codage est systématique et $G = (I_k|P)$, alors on peut montrer que $H = (-{}^tP|I_{n-k})$ est une matrice de contrôle.

Proposition 2.2.1. *Soit C un code linéaire de type (n,k,d) et H une matrice de contrôle. Alors s colonnes quelconques de H sont linéairement indépendantes si et seulement si $d \geq s+1$.*

Démonstration. ($\Leftarrow$) : nous démontrons la contraposée, *i.e.* s'il existe s colonnes C_i linéairement dépendantes, alors $d \leq s$. On supppose donc $\sum_{i\in I} \lambda_i C_i = 0$ avec les λ_i non tous nuls. Comme ${}^tC_i = S(e_i)$, où e_i désigne le vecteur ligne ayant pour unique coefficient non nul un 1 en position i. On a $x = \sum_{i\in I} \lambda_i e_i \in C$ car $S(x) = 0$ et $\omega(x) \leq \mathrm{Card}\, I = s$, donc $d \leq s$.

($\Rightarrow$) : on suppose que s colonnes sont toujours linéairement indépendantes et il s'agit de montrer qu'un élément $0 \neq x = (x_i) \in C$ est toujours de poids $\omega(x) \geq s+1$. Comme $S(x) = 0$, on a $\sum_{i=1}^n x_i C_i = 0$, donc $\mathrm{Card}(\{i, x_i \neq 0\}) \geq s+1$. □

Corollaire. *Soit H une matrice de contrôle du code C. Alors $d = d(C)$ est caractérisé par :*

(i) $d-1$ colonnes sont toujours linéairement indépendantes ;

(ii) il existe d colonnes liées.

Décodage par tableau de syndrôme : si l'on reçoit m' tel que $S(m') \neq 0$ alors $m' \notin C$. On cherche m tel que $\omega(e)$ soit minimal, où $e = m' - m$. On a $S(e) = S(m')$ donc e est le vecteur de syndrôme $S(m')$ de poids minimal. On trace donc le tableau de syndrôme, qui associe à chaque syndrôme possible un mot de poids minimal ayant ce syndrôme. Pour cela, on parcourt les mots par poids croissant. Si le tableau de syndrôme est établi une fois pour toutes, cette méthode est par contre impratiquable pour des codes de très grande dimension.

2.3. Cas particulier des codes de Hamming binaires

Ils sont définis par la matrice de contrôle H de taille $m \times (2^m - 1)$, où la j-ième colonne est l'écriture de j en base 2.

La matrice H est de rang m, donc le noyau de dimension $k = 2^m - m - 1$. Deux colonnes sont toujours linéairement indépendantes mais les trois premières colonnes sont liées, donc $d = 3$. Ce code est donc de type $(2^m - 1, 2^m - m - 1, 3)$. Il est 1-correcteur. C'est de plus un code parfait : les boules de rayon 1 centrées en les q^k éléments sont de cardinal $(q-1)n+1$ et sont disjointes. Or $q^k[(q-1)(2^m-1)+1] = q^n$ car $n - k = m$ et $q = 2$.

On décode comme suit : partant de $m' - m = e = \sum \lambda_i e_i$, avec $\lambda_i \in \{0, 1\}$, on a $S(m') = S(e) = \lambda_i {}^tC_i$. S'il y a une seule erreur, on a $e = e_i$ et $S(m') = {}^tC_i$. Si de plus toutes les colonnes C_j sont distinctes, on peut retrouver i. C'est le cas ici, où C_j est l'écriture de j en base 2. Pour résumer, $S(m')$ donne i directement, à travers son écriture en base 2 ! Enfin, comme le code est binaire, connaître la position de l'erreur suffit à la corriger.

2.4. Equivalence de codes

Définition 2.4.1. *Deux codes C_1 et C_2 sont dits* équivalents *s'ils se déduisent l'un de l'autre par une permutation des coordonnées. Précisément, il existe $\sigma \in S_n$ tel que $C_2 = C_1^\sigma$, où $C_1^\sigma = \{(m_{\sigma(1)}, \ldots, m_{\sigma(n)}),\ m = (m_1, \ldots, m_n) \in C_1\}$.*

Deux codes équivalents ont même longeur, dimension, distance minimale. En théorie de l'information, on étudie donc les codes à équivalence près (en exercice : l'équivalence de code est bien une relation d'équivalence).

De façon équivalente (en exercice), deux codes C_1 et C_2 donnés par des matrices génératrices respectives G_1 et G_2 sont équivalent si $G_2 = MG_1P_\sigma$, où $M \in GL_k(F)$ et P_σ est la matrice de permutation associée à σ (telle que $(m_{\sigma(1)}, \ldots, m_{\sigma(n)}) = (m_1, \ldots, m_n)P_\sigma$).

Proposition 2.4.1. *Tout code linéaire est équivalent à un code donné par un codage systématique.*

Démonstration. Soit G une matrice génératrice. On peut obtenir une décomposition $QGP = (I_k|A)$ où Q est inversible et P est une matrice de pemutation en appliquant l'algorithme de Gauss-Jordan sur les lignes puis en effectuant une permutation des colonnes afin de regrouper les pivots. □

3. Codes cycliques

3.1. Définition

Définition 3.1.1. *Un code linéaire $C \subset \mathbb{F}_q^n$ est dit* cyclique *si*

$$(c_0, \ldots, c_{n-1}) \in C \Longleftrightarrow (c_{n-1}, c_0, \ldots, c_{n-2}) \in C.$$

Cette condition d'invariance par permutation circulaire s'exprime de façon naturelle dans la description plus algébrique suivante.

On considère le quotient de la $\mathbb{F}_q$-algèbre $\mathbb{F}_q[x]$ par l'idéal $(x^n - 1)$: tout élément de $\mathbb{F}_q[x]/(x^n - 1)$ peut être représenté par un polynôme de degré strictement inférieur à n (on prend le reste de la division euclidienne par $x^n - 1$ comme représentant de la classe d'un élément de $\mathbb{F}_q[x]$). On obtient ainsi un isomorphisme de $\mathbb{F}_q$-espaces vectoriels $\mathbb{F}_q^n \xrightarrow{\sim} \mathbb{F}_q[x]/(x^n-1)$ donné par $(c_0, \ldots, c_{n-1}) \mapsto \sum_{i=0}^{n-1} c_i \bar{x}^i$ dont on se sert pour identifier un élément de $\mathbb{F}_q^n$ avec un polynôme modulo $x^n - 1$.

Théorème 3.1.1. *Le code linéaire C est cyclique si et seulement si C est un idéal de $\mathbb{F}_q[x]/(x^n - 1)$.*

Démonstration. Si C est un idéal de $\mathbb{F}_q[x]/(x^n - 1)$ et $c = (c_0, \ldots, c_{n-1}) \in C$ alors

$$x \cdot c = (c_{n-1}, c_0, \ldots, c_{n-2}) \in C.$$

Réciproquement, si C est cyclique alors pour tout $c \in C$ on a $\bar{x}c \in C$, puis $\bar{x}^2 c \in C$, *etc.* Donc un produit $\bar{b}c$ est encore un mot du code pour tout polynôme $b \in \mathbb{F}_q[x]$, si bien que C est un idéal de $\mathbb{F}_q[x]/(x^n - 1)$. □

3.2. Construction par polynôme générateur

L'anneau $\mathbb{F}_q[x]$ est principal, donc tout idéal du quotient $\mathbb{F}_q[x]/(x^n-1)$ est principal (bien que l'anneau quotient $\mathbb{F}_q[x]/(x^n - 1)$ ne soit pas principal puisqu'il n'est pas intègre !) En effet, l'application canonique $\pi : \mathbb{F}_q[x] \to \mathbb{F}_q[x]/(x^n - 1)$ de passage au quotient induit une bijection entre les idéaux de $\mathbb{F}_q[x]$ contenant $x^n - 1$ et les idéaux du quotient. Or les idéaux de $\mathbb{F}_q[x]$ sont principaux et l'image d'un idéal principal par un morphisme d'anneaux est encore un idéal principal.

Un code linéaire cyclique C de dimension k et de longueur n est donc défini par un polynôme g de $\mathbb{F}_q[x]$ tel que $\bar{g}$ engendre C vu comme idéal de $\mathbb{F}_q[x]/(x^n - 1)$. Le polynôme g est alors de degré $n - k$ et divise $x^n - 1$. On a $C = \langle \bar{g}, \bar{x}\bar{g}, \ldots, \bar{x}^{k-1}\bar{g} \rangle$.

Soit $x^n - 1 = f_1 \ldots f_m$ la décomposition de $x^n - 1$ en facteurs irréductibles sur $\mathbb{F}_q$. Nous supposerons désormais que $(n, q) = 1$ ce qui élimine les facteurs multiples. Comme f_i est irréductible, l'idéal $(\bar{f}_i)$ est maximal ; il définit un code cyclique maximal C_i. On trouve tous les codes cycliques de longueur n sur $\mathbb{F}_q$ en choisissant un diviseur quelconque parmi les 2^m diviseurs de $x^n - 1$.

Si $g = \sum_{i=0}^{n-k} g_i x^i$, une matrice génératrice de C est

$$G = \begin{pmatrix} g_0 & g_1 & \cdots & g_{n-k} & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_{n-k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 \cdots & \cdots & 0 & g_0 & g_1 & \cdots & g_{n-k} & \end{pmatrix}$$

Remarque. Définissant $h = (x^n - 1)/g(x) = \sum_{i=0}^{k} h_i x^i$ (appelé *polynôme correcteur* de C), on peut démontrer qu'une matrice de contrôle est

$$H = \begin{pmatrix} 0 & 0 & \cdots & 0 & h_k & h_{k-1} & \cdots & h_0 \\ 0 & 0 & \cdots & h_k & h_{k-1} & \cdots & h_0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ h_k & h_{k-1} & \cdots h_0 & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

3.3. Construction à partir des racines

Lorsque l'on a défini C à partir d'un polynôme générateur g, les éléments de C sont les multiples de g, donc, puisque les racines de g dans un corps de décomposition sont simples (on rappelle que l'on a supposé $(n, q) = 1$), les éléments qui s'annulent en les $\deg(g)$ racines de g. Cela amène à définir g par l'ensemble de ses racines.

Soient donc $\alpha_1, \ldots, \alpha_s$ des éléments d'une extension finie $\mathbb{F}_{q^m}$ de $\mathbb{F}_q$. Nous notons p_i les polynômes minimaux de α_i sur $\mathbb{F}_q$. Un polynôme de $\mathbb{F}_q[x]$ qui s'annule en α_i est un multiple de p_i. Soit $n \in \mathbb{N}$ tel que $\alpha_i^n = 1$ pour tout i (par exemple $n = q^m - 1$) et $g = \mathrm{ppcm}(p_1, \ldots, p_s)$. Alors g divise $x^n - 1$ (car p_i divise $x^n - 1$ pour tout i). Il définit donc un code linéaire cyclique de longueur n sur $\mathbb{F}_q$.

Théorème 3.3.1. *Soit $C \subset \mathbb{F}_q[x]/(x^n - 1)$ un code cyclique de polynôme générateur g dont les racines sont $\alpha_1, \ldots, \alpha_{n-k}$. Alors f est un mot du code si et seulement si le vecteur ${}^t(f_0, \ldots, f_{n-1})$ des coefficients de f est dans le noyau de*

$$H = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \alpha_{n-k} & \alpha_{n-k}^2 & \cdots & \alpha_{n-k}^{n-1} \end{pmatrix}$$

Démonstration. $f \in C$ si et seulement si $f(\alpha_i) = 0$ pour $1 \leq i \leq n - k$. □

La matrice H joue donc le rôle de matrice de contrôle (bien qu'elle ne soit pas à coefficients dans $\mathbb{F}_q$).

Théorème 3.3.2. *Le code cyclique binaire de longueur $n = 2^m - 1$ dont un polynôme générateur est le polynôme minimal sur $\mathbb{F}_2$ d'un élément primitif de $\mathbb{F}_{2^m}$ (i.e. un générateur du groupe cyclique $\mathbb{F}_{2^m}^\times$) est équivalent au code de Hamming binaire de longueur n.*

Démonstration. Soit α un élément primitif de $\mathbb{F}_{2^m}$, *i.e.* $\mathbb{F}_{2^m}^\times = \{1, \alpha, \ldots, \alpha^{2^m-2}\}$. Alors $\mathbb{F}_{2^m} = \mathbb{F}_2(\alpha)$ (c'est donc un élément primitif de l'extension de corps $\mathbb{F}_{2^m}$ sur $\mathbb{F}_2$; la réciproque n'est par contre pas vraie en général : prendre $m = 4$ et considérer α^3) et $\beta = \{1, \alpha, \ldots, \alpha^{m-1}\}$ est une base de $\mathbb{F}_{2^m}$ sur $\mathbb{F}_2$. Soit H la matrice dont la j-ième colonne est le vecteur colonne des coefficients de α^j dans cette base. Alors, pour $a = \sum_{i=0}^{n-1} a_i x^i$ dans $\mathbb{F}_2[x]$, on a $H^t a = a(\alpha)$ exprimé dans la base β. Donc H est une matrice de contrôle du code engendré par g.

Or l'ensemble des colonnes de H coïncide, à permutation près, avec l'écriture des $2^m - 1$ premiers entiers en base 2. Le code cyclique engendré par g est donc équivalent au code de Hamming binaire de longueur $n = 2^m - 1$. □

3.4. Codage et décodage des codes linéaires cycliques

Au mot $M = (M_0, \ldots, M_{k-1})$ on fait correspondre le polynôme $\sum_{i=0}^{k-1} M_i x^i$ noté encore M (en fait sa classe, mais on identifie classe et représentant de degré plus petit que k dans l'implémentation informatique). Le mot codé $m = MG$, où l'on a noté G la matrice génératice correspondant au polynôme générateur g (les lignes de G sont les $x^i g$, pour $0 \leq i \leq k-1$, exprimés dans la base $1, x, \ldots, x^{n-1}$), correspond alors au produit de polynômes $m = Mg$. Le morphisme de codage, dans le formalisme polynômial, correspond donc à une simple multiplication de polynômes dans $\mathbb{F}_q[x]$.

Nous allons voir comment décoder efficacement les codes BCH, qui sont des codes linéaires cycliques particuliers définis à partir de racines.

4. Codes BCH

4.1. Définition

Les codes BCH constituent une classe populaire de codes introduite par Bose, Ray-Chaudhuri et Hocquenghem.

Définition 4.1.1. *Soit β une racine primitive n-ième de l'unité dans $\overline{\mathbb{F}}_q$, pour n un entier premier avec q et g le* ppcm *(unitaire) des polynômes minimaux de $\beta, \ldots, \beta^{\delta-1}$. Le code $BCH(q, n, \delta)$ est le code cyclique sur $\mathbb{F}_q$ défini par les racines $\beta, \ldots, \beta^{\delta-1}$. Il est de longueur n et dimension $k = n - \deg g$, le polynôme g étant un polynôme générateur. On dit que δ est la distance assignée du code.*

On rappelle que l'ensemble des racines n-ièmes de l'unité sur $\mathbb{F}_q$ (*i.e.* les racines du polynôme séparable $x^n - 1 \in \mathbb{F}_q[x]$ dans un corps de décomposition) forme un groupe cyclique isomorphe à $\mathbb{Z}/n\mathbb{Z}$. Les générateurs sont les racines primitives n-ièmes de l'unité ; ce sont les racines dans $\overline{\mathbb{F}}_q$ du polynôme cyclotomique Φ_n sur $\mathbb{F}_q$ (qui est obtenu en réduisant modulo p le n-ième polynôme cyclotomique de $\mathbb{Z}[x]$). L'extension cyclotomique engendrée est de degré m lorsque $n = q^m - 1$ et alors, $\mathbb{F}_q(\beta)^\times = \{1, \beta, \ldots, \beta^{n-1}\}$. On parle de code BCH *primitif*. Dans le cas général, si l'on pose $m = [\mathbb{F}_q(\beta) : \mathbb{F}_q]$, on sait juste que l'ordre n de β divise le cardinal $q^m - 1$ de $\mathbb{F}_q(\beta)^\times$.

Remarque. La définition précédente ne reflète pas le fait qu'un code $BCH(q, n, \delta)$ dépend du choix de β. Cependant, les propriétés du code sont essentiellement indépendantes de β, et en particulier la distance minimale.

Théorème 4.1.1. *La distance minimale de $C = BCH(q, n, \delta)$ est au moins égale à la distance assignée δ. On pourra donc corriger au moins $\lfloor \frac{\delta-1}{2} \rfloor$ erreurs.*

Démonstration. Un élément $f = \sum_{i=0}^{n-1} f_i x^i$ appartient au code si et seulement si

$$\begin{pmatrix} 1 & \beta & \ldots & \beta^{n-1} \\ 1 & \beta^2 & \ldots & \beta^{2(n-1)} \\ \vdots & \vdots & & \vdots \\ 1 & \beta^{\delta-1} & \ldots & \beta^{(\delta-1)(n-1)} \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{n-1} \end{pmatrix} = 0.$$

Parce que tous les déterminants de taille $\delta - 1$ extraits de la matrice ci-dessus sont, à une constante non nulle près, des déterminants de Vandermonde dont les coefficients parmi les β^i sont deux à deux distincts, on voit que ce système d'équations n'admet pas de solution $f \neq 0$ tel que $\omega(f) \leq \delta - 1$. Tout élément non nul c de C vérifie donc $\omega(c) \geq \delta$. □

4.2. Construction pratique et décodage

Voir énoncé du TD-TP correspondant : racines de l'unité dans un corps fini et codes BCH

4.3. Codes de Reed-Solomon

Les codes de Reed-Solomon sont des codes $BCH(q, n, \delta)$ primitifs avec $n = q - 1$ (*i.e.* $m = 1$). On écrira $RS(q, \delta)$.

On peut déterminer la distance minimale de façon exacte :

Proposition 4.3.1. *Un code $RS(q, \delta)$ est un $(q - 1, q - \delta, \delta)$-code.*

Démonstration. En effet, la racine primitive n-ième de l'unité β est alors un générateur de $\mathbb{F}_q^\times$ et $g = \prod_{i=1}^{\delta-1}(x - \beta^i)$. La dimension du code est $k = n - \deg g = q - \delta$. En tant que cas particulier de code BCH, la distance minimale d vérifie $d \geq \delta$. Par ailleurs, la borne de Singleton donne $q - \delta = k \leq n - d + 1 = q - d$, d'où $\delta \geq d$. Finalement $d = \delta$. □

Un aspect important de ces codes réside dans leur capacié de corection lorsque l'on a une série contigüe d'erreurs (effet « burst ») ou bien des effacements. En pratique (pour le CD et le DVD par exemple), on utilise une méthode « d'entrelacement croisé » (on combine deux codes de Reed-Solomon). Nous n'en dirons pas plus et renvoyons le lecteur intéressé à la littérature...

Références

[D] M. Demazure, *« Cours d'algèbre : primalité, divisibilité, codes »*, Nouvelle Bibliothèque Mathématique, Cassini, Paris, 1997.

[L] J.H. van Lint, *"Introduction to coding theory"*, troisième édition, Springer-Verlag, 1999.

[PW] O. Papini et J. Wolfman, *« Algèbre discrète et codes correcteurs »*, collection Math. et Applications, Springer-Verlag, 1995.

[VZGG] J. Von Zur Gathen, J. Gerhard, *"Modern Computer Algebra"*, Cambridge University Press, second ed., 2003.