Introduction à la théorie du codage

Pascal Véron

G.E.C.T.

Mots clés. Codage, Matrice génératrice, Poids minimum, Syndrome

1. Le problème

$$m = (m_1, m_2, \ldots, m_k) \qquad m_i \in \mathcal{A}$$

$$m \longrightarrow \overbrace{\boxed{\text{Canal de transmission}}}^{\text{Perturbation}} \longrightarrow m'$$

Alternative

$$\begin{array}{lllllllll} m & \rightarrow & \boxed{\text{Encodeur}} & \rightarrow & y & \rightarrow & \boxed{\text{Canal de transmission}} & \rightarrow & \\ y' & \rightarrow & \boxed{\text{Correction}} & \rightarrow & y & \rightarrow & \boxed{\text{Décodeur}} & \rightarrow & m \end{array}$$

Fig. 1. Schéma de transmission

- Capacité de correction du code ?
- Nombre de messages que l'on peut coder ?

2. Définitions

Soit $\mathcal{A}$ un ensemble fini non vide, et n un entier naturel non nul. $\mathcal{A}^n$ désigne l'ensemble des $x = (x_1, ..., x_n)$ avec $x_i \in \mathcal{A}$.

DÉFINITION 2.1. *Soient x et y deux éléments de $\mathcal{A}^n$. On appelle* **distance de Hamming** *entre x et y, le nombre de composantes pour lesquelles ces éléments diffèrent. Plus précisément si $x = (x_1, ..., x_n)$ et $y = (y_1, ..., y_n)$, alors*

$$d(x, y) = \text{card}\{i \in \{1, 2, ..., n\} \mid x_i \neq y_i\}$$

Exemple – : Soit $\mathcal{A} = \{+, ?, \neq\}$ et $n = 5$, alors pour $x = (?, ?, +, ?, \neq)$ et $y = (?, \neq, \neq, +, \neq)$ on a $d(x, y) = 3$. ◇

DEFINITION 2.2. *Un code sur $\mathcal{A}$ de longueur n est un sous-ensemble C de $\mathcal{A}^n$. L'ensemble $\mathcal{A}$ est appelé l'alphabet, n la longueur du code C et les éléments de C sont appelés les mots du code.*

3. Codes Correcteurs d'erreurs

DEFINITION 3.3. *Soit $x = (x_1, ..., x_n)$ et r un entier non nul. On appelle boule de centre x et de rayon r, l'ensemble des éléments $y = (y_1, ..., y_n)$ tels que $d(x, y) \leq r$. On note :*

$$B(x, r) = \{y \in \mathcal{A}^n \mid d(x, y) \leq r\}$$

Exemples :

- $\mathcal{A} = \{0, 1\}$
  $n = 2$
  $\mathcal{A}^2 = \{(0,0); (0,1); (1,0); (1,1)\}$
  $B((0,0), 1) = \{(0,0); (0,1); (1,0)\}$

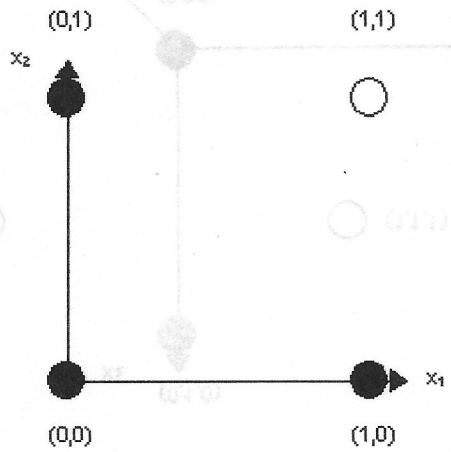

Fig. 2. B((0, 0), 1)

- $\mathcal{A} = \{0, 1\}$
  $n = 3$
  $\mathcal{A}^3 = \{(0,0,0); (0,0,1); (0,1,0); (0,1,1);$
  $(1,0,0); (1,0,1); (1,1,0); (1,1,1)\}$
  $B((0,0,0), 1) = \{(0,0,0); (0,0,1); (0,1,0); (1,0,0)\}$

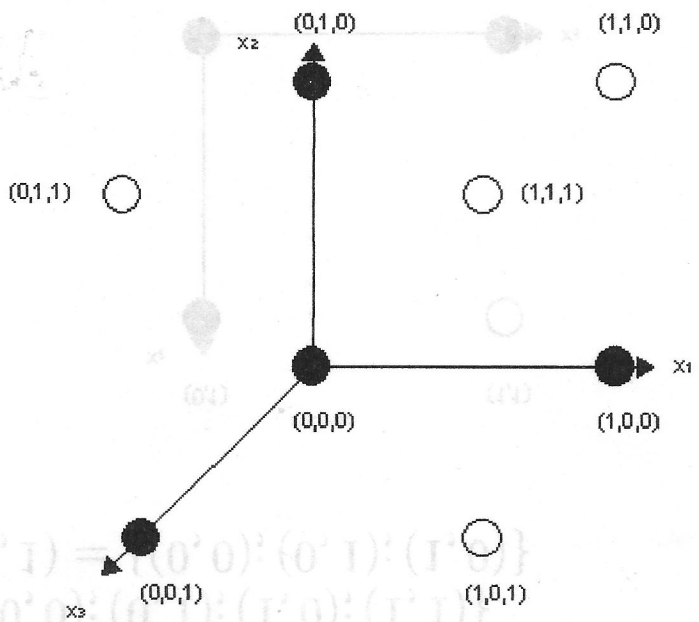

Fig. 3. B((0, 0, 0), 1)

3.1. Condition de décodage d'ordre t

DÉFINITION 3.4. *Un code C de longueur n sur un alphabet $\mathcal{A}$ vérifie la condition de décodage d'ordre t si pour tout x' de $\mathcal{A}^n$, il existe au plus un mot x de C tel que $d(x', x) \leq t$. Cette condition est équivalente à ce que les boules fermées (pour la distance de Hamming) de rayon t, centrées sur les mots de C, soient deux à deux disjointes.*

Exemple : Soient $\mathcal{A} = \{0, 1\}, n = 5, t = 1$, et soit le code :

$$C = \{x = (0,1,1,1,0); y = (1,0,1,0,1); z = (1,1,0,1,1)\}$$

$$\begin{aligned} B(x,1) &= \{(0,1,1,1,0); (1,1,1,1,0); (0,0,1,1,0); \\ &\quad (0,1,0,1,0); (0,1,1,0,0); (0,1,1,1,1)\} \end{aligned}$$

$$\begin{aligned} B(y,1) &= \{(1,0,1,0,1); (0,0,1,0,1); (1,1,1,0,1); \\ &\quad (1,0,0,0,1); (1,0,1,1,1); (1,0,1,0,0)\} \end{aligned}$$

$$\begin{aligned} B(z,1) &= \{(1,1,0,1,1); (0,1,0,1,1); (1,0,0,1,1); \\ &\quad (1,1,1,1,1); (1,1,0,0,1); (1,1,0,1,0)\} \end{aligned}$$

C vérifie la condition de décodage d'ordre 1 : $B(x,1) \cap B(y,1) \cap B(z,1) = \varnothing$.

$$v' = (1,1,1,0,1) \Longrightarrow v = y = (1,0,1,0,1)$$

Contre-exemple : Soient $\mathcal{A} = \{0,1\}, n = 3, t = 1$, et soit le code :

$$C = \{(0,0,0);(1,0,1)\}$$

Notons respectivement x et y les deux mots du code C. On a :

$$\begin{aligned} B(x,1) &= \{(0,0,0);(1,0,0);(0,0,1);\\ &\quad (0,1,0)\} \end{aligned}$$

$$\begin{aligned} B(y,1) &= \{(1,0,1);(0,0,1);(1,1,1);\\ &\quad (1,0,0)\} \end{aligned}$$

Les deux boules ne sont pas disjointes, en effet

$$B(x,1) \cap B(y,1) = \{(1,0,0);(0,0,1)\}.$$

Si le mot reçu est $(1,0,0)$ (ou $(0,0,1)$) il est donc impossible de savoir s'il provient de x ou de y.

3.2. Théorème Fondamental

DÉFINITION 3.5. *La distance minimale d'un code C est la plus petite des distances, non nulle, entre les mots de C. On note :*

$$d_{\min} = \inf\{d(x,y) \mid (x,y) \in C \times C, x \neq y\}$$

THÉORÈME 3.6. *Soit d la distance minimale d'un code C. Si $d \geq 2t+1$,alors C vérifie la condition de décodage d'ordre t.*

COROLLAIRE 3.7. *Un code, dont la distance minimale est d, est capable de corriger au moins t erreurs, où*

$$t = \lfloor \frac{d-1}{2} \rfloor$$

3.3. Borne d'empilement de sphères

Considérons un code C de longueur n sur un alphabet $\mathcal{A}$. Une question alors se pose :

- Quel est le nombre maximum de mots que peut contenir un code C de ca pacité de correction t ?

Ceci nous donne le nombre de messages que l'on pourra coder.

PROPOSITION 3.8. *Soit C un code de longueur n sur un alphabet $\mathcal{A}$. Pour tout entier $r \leq n$, on a :*

$$\forall x \in \mathcal{A}^n \qquad \mid B(x,r) \mid = \sum_{i=0}^{r} (\mid \mathcal{A} \mid -1)^i \binom{n}{i}$$

On a alors :

PROPOSITION 3.9. *Soit C un code de longueur n sur un alphabet $\mathcal{A}$ de capacité de correction t, alors*

$$\mid C \mid \leq \frac{\mid \mathcal{A} \mid^n}{\sum_{i=0}^{t} (\mid A \mid -1)^i \binom{n}{i}}$$

Preuve – En effet :

$$\bigcup_{x \in C} B(x,t) \subset \mathcal{A}^n$$

$$\sum_{x \in C} \mid B(x,t) \mid \leq \mid \mathcal{A}^n \mid$$

or d'après la proposition précédente on a :

$$\sum_{x \in C} \mid B(x,t) \mid = \mid C \mid \sum_{i=0}^{t} (\mid \mathcal{A} \mid -1)^i \binom{n}{i}$$

ce qui prouve le résultat. □

3.4. Codes équivalents

Soit C un code de longueur n et de distance minimale d . Soit σ une permutation de l'ensemble $\{1, \ldots, n\}$. Considérons le code C' dont chaque mot x' provient d'un mot x de C auquel on a appliqué la permutation σ. Alors la distance minimale du code C' est encore d.

DÉFINITION 3.10. *Deux codes C et C' sont dits équivalents si l'un peut être obtenu en permutant les symboles de l'autre.*

Exemple – :

$$\begin{array}{lcll} C & = & \{(1,0,0,1);(1,1,1,0);(0,0,1,1)\} & d=2 \\ \sigma & = & (4,2,3,1) & \\ C' & = & \{(1,0,0,1);(0,1,1,1);(1,0,1,0)\} & d'=2 \end{array}$$

Problème :

Trouver des codes sur $\{0,1\}$:

- Possèdant de bonnes structures algébriques.
- dont il est inutile de donner la liste des mots.

$\longrightarrow$ Les codes linéaires

4. Les codes linéaires

L'alphabet $\mathcal{A}$ sera l'ensemble $\{0,1\}$ doté des opération suivantes :

$$\forall x \in \{0,1\} \quad x+0=x, x+x=0, x\times 0=0, x\times x=x$$

Soient $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_n)$, et $\lambda \in \{0,1\}$, on a :

$$x+y=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$$

et :

$$\lambda x=(\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

Exemple : $n=4, x=(0,1,1,0), y=(1,1,0,1), \lambda=0,$

$$\begin{aligned} x+y &= (1,0,1,1) \\ 0x &= (0,0,0,0) \end{aligned}$$

DEFINITION 4.11. *Soit C un code de longueur n sur $\{0,1\}$. C est un code linéaire si :*

$$\forall (x,y) \in C \times C, \qquad x+y \in C$$
$$\underbrace{(0,\ldots,0)}_{n} \in C$$

Exemple : $C = \{(0,0,0,0,0);(1,0,1,1,0);(0,1,1,1,1);(1,1,0,0,1)\}$.

DÉFINITION 4.12. *Un code linéaire de longueur n sur $\{0,1\}$ et de dimension k est un sous-espace vectoriel de dimension k de $\{0,1\}^n$. On le note $C(n,k)$.*

C sous-espace vectoriel de $\{0,1\}^n \rightarrow C$ vérifie la propriété précédente.

Le code est de dimension k s'il suffit de connaitre k éléments du code $(x_1,\ldots,x_k)$ pour connaitre tous les autres mots du code. Les éléments $x_1,\ldots,x_k$ sont appelés la **base** du code. Chaque mot x du code est alors de la forme :

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k, \qquad \lambda_i \in \{0,1\}$$

Exemple : le code suivant est un code de dimension 2 et de longueur 5

$$C = \{(0,0,0,0,0);(1,0,1,1,0);(0,1,0,1,0);(1,1,1,0,0)\}$$

une base du code est

$$\begin{aligned}(1,0,1,1,0) &= x_1\\ (0,1,0,1,0) &= x_2\end{aligned}$$

En effet les autres mots du code sont obtenus de la façon suivante :

$$\begin{aligned}(0,0,0,0,0) &= 0\times x_1 + 0 \times x_2\\ (0,1,0,1,0) &= 1\times x_1 + 1 \times x_2\end{aligned}$$

PROPOSITION 4.13. *Si C est un code linéaire de longueur n et de dimension k sur $\{0,1\}$ alors le code comporte 2^k mots.*

DÉFINITION 4.14. *Le poids d'un élément $x = (x_1, \ldots, x_n)$ est le nombre de ses composantes non nulles. On note*

$$\omega(x) = \text{card}\{i \in \{1, \ldots, n\} \mid \omega_i \neq 0\}$$

PROPOSITION 4.15. *Soient d la distance de Hamming, et ω la fonction poids, alors pour x et y dans $\{0,1\}^n$ on a :*

- $d(x,y) = \omega(x+y)$,
- $\omega(x) = d(x,0)$,
- $\omega(x) = 0 \Leftrightarrow x = 0$,
- $\omega(x+y) \leq \omega(x) + \omega(y)$.

DÉFINITION 4.16. *Soit C un code linéaire de longueur n et de dimension k, on appelle poids du code, le plus petit poids, non nul, des élements du code. On note :*

$$\omega(C) = \min\{\omega(x) \mid x \in C, x \neq (0, \ldots, 0)\}$$

PROPOSITION 4.17. *Le poids minimum d'un code linéaire est égal à la distance minimum du code.*

PROPOSITION 4.18. *Un code linéaire $C(n,k)$ est t-correcteur si son poids minimal est $2t+1$ ou $2t+2$.*

4.1. Matrice génératrice

DÉFINITION 4.19. *On appelle matrice génératrice, d'un code linéaire $C(n,k)$, le tableau dont chaque ligne est composé des éléments de la base du code. On note :*

$$G = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$$

Exemple : Pour le code précedent on a

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

On montre :
- qu'un code linéaire possède plusieurs matrices génératrices,
- que tout code linéaire $C(n,k)$ est équivalent à un code $C'(n,k)$ de matrice génératrice

$$G' = k \left(\overbrace{\begin{matrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{matrix}}^{k} \quad \overbrace{\begin{matrix} \\ \\ M \\ \\ \\ \end{matrix}}^{n-k} \right)$$

M étant une matrice comportant k lignes et $n-k$ colonnes dont les éléments font partie de l'ensemble $\{0,1\}$. La matrice, à k lignes et k colonnes, comportant uniquement des 1 sur la diagonale est notée I_k.

Construire un code linéaire $C(n,k) \rightarrow$ exhiber sa matrice génératrice :

$$G = (I_k \mid M)$$

Chaque mot du code est obtenu en calculant :

$$\lambda_1 G_1 + \cdots + \lambda_k G_k \quad \text{où} \lambda_i \in \{0,1\}$$

G_i représentant la ligne n° i de G .

Exemple : La matrice suivante est la matrice génératrice d'un code de longueur 5 et de dimension 3 :

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \begin{matrix} G_1 \\ G_2 \\ G_3 \end{matrix}$$

Les mots du code sont :

$$\begin{aligned}
(1,0,0,1,0) &= 1 \times G_1 + 0 \times G_2 + 0 \times G_3 \\
(0,1,0,1,1) &= 0 \times G_1 + 1 \times G_2 + 0 \times G_3 \\
(0,0,1,0,1) &= 0 \times G_1 + 0 \times G_2 + 1 \times G_3 \\
(1,1,0,0,1) &= 1 \times G_1 + 1 \times G_2 + 0 \times G_3 \\
(1,0,1,1,1) &= 1 \times G_1 + 0 \times G_2 + 1 \times G_3 \\
(0,1,1,1,0) &= 0 \times G_1 + 1 \times G_2 + 1 \times G_3 \\
(1,1,1,0,0) &= 1 \times G_1 + 1 \times G_2 + 1 \times G_3 \\
(0,0,0,0,0) &= 0 \times G_1 + 0 \times G_2 + 0 \times G_3
\end{aligned}$$

Réalisation du schéma :

$$m \longrightarrow \boxed{\text{encodeur}} \longrightarrow y$$

On dispose d'un code $C(n,k)$ sur $\{0,1\}$. Le message à transmettre est un message binaire de longueur k, i.e.

$$m = (m_1, \ldots, m_k) \qquad m_i \in \{0,1\}$$

Soit G la matrice génératrice du code C, on associe alors à m le mot suivant :

$$y = m_1 G_1 + m_2 G_2 + \cdots + m_k G_k$$

Exemple : Le code C définie ci-dessus permet de coder des messages de longueur 3. Au message $m = (1,0,1)$, on associe le mot

$$\begin{array}{rcl} y & = & 1 \times (1,0,0,1,0) + 0 \times (0,1,0,1,1) + 1 \times (0,0,1,0,1) \\ & = & (1,0,1,1,1) \end{array}$$

L'étape de décodage

$$y \longrightarrow \boxed{\text{décodeur}} \longrightarrow m$$

est donc triviale. Elle consiste à simplement récupérer les k premiers éléments du mot y, car $G = (I_k \mid M)$.

DÉFINITION 4.20. *Les k premières composantes de y sont appelées les* **symboles d'information**, *les $n-k$ suivantes sont appelées les* **symboles de contrôle**.

Si cette technique de construction du code ne nous permet pas à l'avance de déterminer sa capacité de correction, on a tout de même le résultat suivant :

PROPOSITION(BORNE DE SINGLETON) 4.21. *La distance minimale d d'un code linéaire $C(n,k)$ vérifie :*

$$d \leq n - k + 1$$

Preuve – Ceci se déduit à partir de la matrice génératrice du code. On a $G = (I_k \mid M)$, et $d \leq \omega(G_i)$, pour $i = 1, \ldots, k$. Or, au pire, M contient une ligne dont tous les éléments valent 1, d'où $d \leq n - k + 1$. □

4.2. Matrice de contrôle

DÉFINITION 4.22. *Soient x et y deux mots de longueur n sur $\{0,1\}$, on appelle* **produit scalaire** *de x par y la quantité suivante :*

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Exemple : n=4, $x = (0,1,1,0), y = (1,1,0,0), \langle x, y \rangle = 1$.

DÉFINITION 4.23. *Soit C un code linéaire de longueur n.On appelle code orthogonal de C, et on note $C^{\perp}$, le code de longueur n défini par :*

$$C^{\perp} = \{y \in \{0,1\}^n \mid \langle x, y \rangle = 0 \quad \forall x \in C\}$$

PROPOSITION 4.24. *Soit C un code linéaire de longueur n sur $\{0,1\}$ et de dimension k, alors $C^{\perp}$est un code linéaire de longueur n et de dimension $n - k$.*

DÉFINITION 4.25. *On appelle matrice de contrôle d'un code C, toute matrice génératrice de son orthogonal.*

DÉFINITION 4.26. *Soit M une matrice comportant n colonnes et k lignes. On note m_{ij} l'élément situé à la ligne i et à la colonne j. On appelle transposé de M et on note ${}^tM$, la matrice comportant k colonnes et n lignes avec ${}^tm_{ij} = m_{ji}$.*

Exemple :

$$M = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \quad {}^tM = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

• Relation liant la matrice génératrice et la matrice de contrôle d'un code.

PROPOSITION 4.27. *Soit C un code de longueur n et de dimension k . Soit G sa matrice génératrice :*

$$G = (I_k \mid M)$$

la matrice H de contrôle du code est alors

$$H = ({}^tM \mid I_{n-k})$$

De même si $G = (M \mid I_k)$ alors $H = (I_{n-k} \mid {}^t M)$

Exemple :

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \quad H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

PROPOSITION(PROPRIÉTÉ FONDAMENTALE) 4.28. *Soit H une matrice de contrôle d'un code C et $x = (x_1, \ldots, x_n)$. Le mot x appartient au code C si et seulement si :*

$$x_1{}^t(H^1) + x_2{}^t(H^2) + \cdots + x_n{}^t(H^n) = \overbrace{(0 \cdots 0)}^{n-k}$$

où H^i représente la colonne i de la matrice H .

Exemple : Considérons le code défini ci-dessus. Le mot $x = (1,1,0,0,0,0)$ appartient au code. En effet :

$$1\times(1,0)+1\times(1,0)+0\times(0,1)+0\times(1,1)+0\times(1,0)+0\times(0,1) = (0,0)$$

Par contre le mot $x = (1,1,0,0,0,1)$ n'appartient pas au code car :

$$1\times(1,0)+1\times(1,0)+0\times(0,1)+0\times(1,1)+0\times(1,0)+1\times(0,1) = (0,1)$$

Remarque Importante : D'après la proposition précédente, il existe dans le code C un mot de poids r si et seulement si il existe r colonnes de H de somme nulle.

PROPOSITION 4.29. *Le poids minimum d'un code linéaire C sur $\{0,1\}$ est le plus petit entier r tel qu'il existe r colonnes de H de somme nulle. Un tel code corrigera alors au moins $\lfloor \frac{r-1}{2} \rfloor$ erreurs.*

Nous sommes donc en mesure de construire un code dont on connaitra à l'avance la capacité de correction. En particulier,

Mot de poids 1 dans C	$\Leftrightarrow$	Au moins une colonne nulle dans$H$
Mot de poids 2 dans C	$\Leftrightarrow$	Au moins deux colonnes égales dans H
Mot de poids au moins 3 dans C	$\Leftrightarrow$	Toutes les colonnes de H sont distinctes et aucune n'est nulle.

Ainsi pour construire un code 1-correcteur (poids minimal du code ≥ 3), il suffit que H vérifie la troisième condition.

Exemple : Construction d'un code C, 1-correcteur, de longueur 7. (Taux d'erreur du canal : 1 erreur pour 7 bits transmis).

Quelle est la longueur maximale des messages que l'on pourra coder, i.e. quelle est la dimension maximale que peut avoir le code ?

D'après la borne de Singleton $\dim(C)=k \leq n-d+1$, d'où $k \leq 5$.

- $k = 5$. $\dim(C^{\perp}) = n - k = 7 - 5 = 2$. D'où

$$H = (I_2 \mid N)$$

Code 1-correcteur $\Rightarrow$ les 7 colonnes de cette matrice sont distinctes et aucune n'est nulle.

Or

$$H = \begin{pmatrix} 1 & 0 & \\ 0 & 1 & N \end{pmatrix}$$

Il n'existe qu'un seul élément non nul distinct de $(1,0)$ et de $(0,1)$ dans $\{0,1\}^2$, l'élément $(1,1)$.

THÉORÈME 4.30. *Il existe un code linéaire 1-correcteur sur* $\{0,1\}$, *de longueur* n *et de dimension* k, *si et seulement si* $2^{n-k} \geq n+1$

.

- $k = 4$. Le code $C^{\perp}$ est alors de dimension 3, donc :

$$H = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & 1 & 0 & N \\ 0 & 0 & 1 & \end{pmatrix}$$

où N est une matrice comportant 3 lignes et 4 colonnes. Toutes les colonnes de H doivent être distinctes et aucune ne doit être nulle, on peut donc prendre par exemple :

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

On a alors :

$$G = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

le code C comporte 16 mots et son poids minimum est 3.

4.3. Correction d'erreurs à l'aide de la matrice de contrôle.

C un code t-correcteur sur $\{0,1\}$, de longueur n et de dimension k.
H sa matrice de contrôle.
Soit $x \in C$, $y = x + e$ le mot reçu, e un mot erreur de longueur n et $\omega(e) \leq t$

Soit h l'application définie par :

$$\begin{array}{rccc} h: & \{0,1\}^n & \longrightarrow & \{0,1\}^{n-k} \\ & z & \longmapsto & z_1{}^t(H^1) + z_2{}^t(H^2) + \cdots + z_n{}^t(H^n) = h(z) \end{array}$$

DÉFINITION 4.31. *Le syndrome d'un élément z de $\{0,1\}^n$ (par rapport à H) est le vecteur $h(z)$ de $\{0,1\}^{n-k}$.*

On a donc :

$$\begin{aligned} h(y) &= h(x) + h(e) \\ &= h(e) \end{aligned}$$

En effet $h(x) = 0$ puisque $x \in C$.

Ainsi **le mot reçu y et le mot erreur e ont même syndrome.**

PROPOSITION 4.32. *Soit C un code linéaire t-correcteur, alors tous les mots erreur de poids inférieurs ou égaux à t ont des syndromes distincts.*

Preuve – Soient e_1 et e_2 tels que $\omega(e_1) \leq t$ et $\omega(e_2) \leq t$ alors :

$$\begin{array}{rccc} & h(e_1) & = & h(e_2) \\ \Leftrightarrow & h(e_1 + e_2) & = & 0 \\ \Leftrightarrow & e_1 + e_2 & \in & C \end{array}$$

or

$$\omega(e_1 + e_2) \leq \omega(e_1) + \omega(e_2) \leq 2t$$

et tous les mots de C sont de poids $p \geq 2t + 1$ (puisque le code est t-correcteur). D'où $h(e_1) \neq h(e_2)$. □

Principe du décodage : (pour un code t-correcteur, de longueur n, sur $\{0,1\}$)

- Construire un tableau de déchiffrement composé de 2 colonnes. La première colonne comporte tous les mots de poids inférieur ou égal à t (mots erreurs). Dans la deuxième colonne se trouvent les syndromes correspondants.

- Soit y un mot reçu ($y = x + e$).

- Calculer $h(y)$.

- Chercher dans la deuxième colonne du tableau de déchiffrement la quantité $h(y)$, on trouve alors sur la même ligne, dans la première colonne, le mot erreur e correspondant.

- On a alors $x = y + e$.

La méthode est fiable si le mot y reçu provient effectivement d'un mot du code entaché d'au plus t erreurs.

Remarque : Cette méthode s'avère très vite irréalisable dès que n et t croient, puisque le tableau de déchiffrement comporte $\binom{n}{t}$ lignes.

5. Exemple d'application

• $C(7,4)$ 1-correcteur

$$G = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Mot erreur	Syndrome
$(1,0,0,0,0,0,0)$	$(1,0,0)$
$(0,1,0,0,0,0,0)$	$(0,1,0)$
$(0,0,1,0,0,0,0)$	$(0,0,1)$
$(0,0,0,1,0,0,0)$	$(0,1,1)$
$(0,0,0,0,1,0,0)$	$(1,0,1)$
$(0,0,0,0,0,1,0)$	$(1,1,0)$
$(0,0,0,0,0,0,1)$	$(1,1,1)$

Le code étant de dimension 4, les messages à envoyer seront de longueur 4. Soit $m = (1,0,1,1)$.

• Codage du message : Au message m on associe le mot de code

$$\begin{aligned} x &= 1 \times (0,1,1,1,0,0,0) + 0 \times (1,0,1,0,1,0,0) + \\ &\quad 1 \times (1,1,0,0,0,1,0) + 1 \times (1,1,1,0,0,0,1) \\ &= (0,1,0,1,0,1,1) \end{aligned}$$

Ce mot est transmis via le canal défectueux. Supposons que le mot reçu soit

$$y = (0,1,1,1,0,1,1)$$

(une erreur a été commise en troisième position).

• Correction de l'erreur :

$$\begin{aligned} h(y) &= 0 \times (1,0,0) + 1 \times (0,1,0) + 1 \times (0,0,1) + \\ &\quad 0 \times (1,0,1) + 1 \times (1,1,0) + 1 \times (1,1,1) \\ &= (0,0,1) \end{aligned}$$

Dans le tableau de déchiffrement, on voit que l'erreur correspondant au syndrome $(0,0,1)$ est $(0,0,1,0,0,0,0)$. Le mot envoyé est donc :

$$\begin{aligned} x &= (0,1,1,1,0,1,1) + (0,0,1,0,0,0,0) \\ &= (0,1,0,1,0,1,1) \end{aligned}$$

• Décodage du message : La récupération des 4 dernières composantes de x donne

$$m = (1,0,1,1)$$

La méthode est efficace puisque une seule erreur a été commise.

Si le mot reçu est :

$$\begin{aligned} y &= (0,0,1,1,0,1,1) \\ &= x + (0,1,1,0,0,0,0) \end{aligned}$$

alors $h(y) = (0,1,1)$ et le mot erreur associé est $(0,0,0,1,0,0,0)$. L'étape de correction donne

$$x = (0,0,1,0,0,1,1)$$

on en déduit que $m = (0,0,1,1)$. Et donc le message initial n'est pas retrouvé puisque plus d'une erreur a été commise lors de la transmission.