

# Disque compact et code correcteur

**Niveau :** terminale S spécialité.

**Lien avec le programme :** problème de codage, (matrices et calculs matriciels), matrice carrée, matrice inverse d'une telle. Mise en œuvre d'un algorithme.

**Lien avec Les maths au quotidien :** Codage, Musique.

Les informations contenues sur un disque compact, comme de la musique, sont numérisées par une piste de plusieurs kilomètres de long, succession de trames d'octets.

Un octet est un mot de 8 bits. Par exemple :

00101110, qui a pour valeur décimale  $0 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 46$

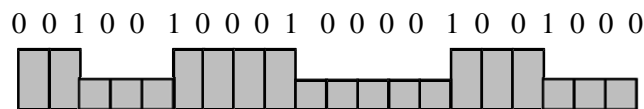
01101011 qui a pour valeur décimale  $0 \times 2^7 + 1 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 107$

Ces octets sont gravés sur le disque à l'aide de creux et de plats.

Les informations gravées sur la surface sont lues à l'aide d'une diode laser.

L'information « 0 » ou « 1 » est alors obtenue de la façon suivante :

- On a un « 1 » logique lors du passage d'un creux à un plat ou d'un plat à un creux (appelé front).
- On a un « 0 » logique lorsqu'il n'y a pas de front.



Lors de la lecture d'un disque compact par un lecteur, il peut y avoir des erreurs :

- des erreurs isolés - un bit ou une suite de quelques bits, dues à un défaut de pressage ou une erreur de lecture.
- une salve d'erreurs (plusieurs octets), causées par une rayure, une empreinte de doigt ou un grain de poussière.

Les erreurs vont cependant en général pouvoir être corrigées : c'est la fonction de l'algorithme de détection et de correction d'erreurs « CIRC ».

Si les erreurs sont sur des octets isolés, le lecteur peut détecter les octets corrompus et corriger les erreurs à l'aide de ce qu'on appelle les données de parité : il s'agit pour chaque trame de 24 octets de deux fois deux octets supplémentaires contenant respectivement la somme des valeurs décimales des octets de la trame (*Checksum*), ainsi que la somme pondérée de ces valeurs décimales, les poids étant les numéros des emplacements des octets sur la trame (poids de 1 à 24).

S'il y a une salve d'erreurs, le lecteur est capable de séparer les octets corrompus, par une méthode appelée désentrelacement, puis ensuite de corriger les erreurs sur les octets corrompus isolés avec les données de parité.

Il aura fallu d'abord entrelacer, c'est-à-dire mélanger les octets de la piste du CD, avant la gravure du disque.

Exemple sur 24 octets :

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

10	7	19	3	24	14	4	20	2	15	6	21	8	12	16	23	1	13	21	5	9	17	22	11
----	---	----	---	----	----	---	----	---	----	---	----	---	----	----	----	---	----	----	---	---	----	----	----

Entrelacement des octets lors de la fabrication du disque

10	7	19	3	24	14	4	20	2	15	6	21	8	12	16	23	1	13	21	5	9	17	22	11
----	---	----	---	----	----	---	----	---	----	---	----	---	----	----	----	---	----	----	---	---	----	----	----

Lecture de la piste contenant 4 octets erronés

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Désentrelacement avec séparation des octets erronés

## Partie A- Entrelacement-Désentrelacement

Pour illustrer le principe d'entrelacement-désentrelacement des octets sur le disque, on va prendre un exemple avec une piste de 5 octets seulement.

Soit  $A = (a_1 \ a_2 \ a_3 \ a_4 \ a_5)$  dont les coefficients sont les valeurs décimales des 5 octets (comprises entre 0 et 255).

$$\text{Soient } M = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \text{ et } N = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}. M \text{ est la matrice d'entrelacement.}$$

- Calculer  $B = AM$ . ( $B$  est la matrice dont les coefficients sont les valeurs décimales des octets qui vont être gravés dans cet ordre sur le disque).
- Calculer  $MN$ . Qu'en déduit-on pour  $M$  ?
  - Déterminer la matrice de désentrelacement : celle qui permet de retrouver  $A$  à partir de  $B$ .
- Soit  $B = (a_2 \ a_4 \ a_1 \ a_5 \ a_3)$ . Déterminer la matrice d'entrelacement  $M$ , puis la matrice de désentrelacement.

## Partie B- Détection et correction d'une erreur

Dans la suite, on considère qu'après le désentrelacement, les octets erronés sur le disque sont isolés :

On considère que pour toute suite de 3 octets, il y a au plus une erreur sur un octet et que l'on connaît la somme et la somme pondérée des 3 octets (sur un vrai disque compact, c'est 24 octets). Ces deux informations sont présentes sur le disque et on suppose qu'elles ne sont pas altérées.

Soit  $A$  la matrice ligne  $(a_1 \ a_2 \ a_3)$  dont les coefficients sont des nombres entiers compris entre 0 et 255. Soit la matrice  $A' = (a_1 \ a_2 \ a_3 \ s_1 \ s_2)$  avec  $s_1$  et  $s_2$  définis (sur 2 octets chacun) par :

$$s_1 = a_1 + a_2 + a_3 \text{ et } s_2 = a_1 + 2a_2 + 3a_3.$$

- Écrire  $A'$  lorsque  $A = (67 \ 4 \ 192)$ .

- On définit la matrice  $G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 3 \end{pmatrix}$ . Montrer que  $AG = A'$ .

La matrice ligne  $A = (a_1 \ a_2 \ a_3)$  contient au plus un octet erroné. Le lecteur lit à la place  $B = (b_1 \ b_2 \ b_3)$ .

Soit  $B' = (b_1 \ b_2 \ b_3 \ s_1 \ s_2)$ , qui diffère de la matrice  $A'$  par au plus l'un de ses trois premiers coefficients :

$$B' = A' + E \text{ avec } E = (x \ 0 \ 0 \ 0 \ 0) \text{ ou } E = (0 \ x \ 0 \ 0 \ 0) \text{ ou } E = (0 \ 0 \ x \ 0 \ 0) \text{ avec } x \text{ entier relatif.}$$

- Détection d'erreur** À partir de  $B'$ , le CIRC calcule la matrice  $S = (b_1 + b_2 + b_3 - s_1 \ b_1 + 2b_2 + 3b_3 - s_2)$ . La matrice  $S$  s'appelle le *syndrome* de  $B$ .

  - Dans cette question,  $A = B$ . Interpréter la situation. Calculer le syndrome de  $B$ .
  - Que vaut le syndrome de  $B$  lorsque  $A = (3 \ 43 \ 12)$  et que le second octet est lu comme étant 40 ?
  - Montrer qu'il n'y a pas eu d'erreur dans la transmission si et seulement si  $S = (0 \ 0)$ .

- Correction d'erreur** On définit la matrice  $C = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$

  - Montrer que  $S = B'C$ .
  - Montrer que  $A'C = (0 \ 0)$ . En déduire que  $S = EC$ .
  - En déduire que  $S$  est de l'une des formes suivantes :  $S = (x \ x)$  ou  $S = (x \ 2x)$  ou  $S = (x \ 3x)$  avec  $x$  entier.
  - Quelle est la matrice  $E$  si  $S = (3 \ 3)$  ? si  $S = (7 \ 21)$  ?
  - On suppose ici que  $B' = (5 \ 17 \ 8 \ 33 \ 72)$ . Corriger l'éventuelle erreur.
  - En une phrase, expliquer comment corriger l'éventuelle erreur en « observant » les coefficients de  $S$ .
  - Écrire alors un algorithme qui prend en entrée les deux coefficients du syndrome  $S$  et qui indique en sortie s'il y a une erreur, et si oui, indique la position de l'octet corrompu et la valeur décimale de l'erreur. On pourra programmer l'algorithme en Python et fournir le script.