

Les entrées/sorties

IX.1 Introduction

L'informatique sert au traitement de l'information. Il est donc indispensable d'être capable de fournir l'information à traiter à un ordinateur et de récupérer les résultats. Le système informatique doit pouvoir communiquer avec l'extérieur. Ces opérations d'échange constituent les Entrées/Sorties (E/S ou I/O : Input/Output). A l'origine l'opérateur entrait les informations à l'aide de clefs ou interrupteurs. On a ensuite utilisé rubans et cartes perforés dont l'invention est largement antérieure à celle de l'informatique. Un ruban perforé, en papier résistant, pouvait contenir des trous sur 5 ou 7 pistes parallèles. D'emploi facile, les rubans étaient cependant fragiles et difficiles à modifier. C'est pourquoi les rubans ont progressivement été remplacés par les cartes perforées. Une carte, en carton et de forme rectangulaire, comptait 80 colonnes et 12 lignes. Elle pouvait donc recevoir jusqu'à 80 mots de 12 bits. Par exemple 80 caractères. Chaque caractère était représenté, selon le code Hollerith, à l'aide de perforations dans la colonne correspondante. Bien qu'ayant disparue, l'influence de la carte perforée s'est longtemps faite sentir dans certains formats, comme celui des écrans alphanumériques à 80 colonnes. Ensuite sont apparus claviers, imprimantes, bandes magnétiques, disques et disquettes, terminaux interactifs, souris, crayons optiques, etc. Les ordinateurs peuvent également recevoir les informations de capteurs et envoyer des signaux électroniques à des périphériques : commande de processus industriels ou acquisition de données. Ils peuvent aussi échanger des informations entre eux par l'intermédiaire des réseaux.

IX.2 Terminal interactif

Un terminal interactif permet un échange bidirectionnel entre l'utilisateur et l'ordinateur. Il comprend au moins un clavier et un écran de visualisation encore appelé moniteur.

IX.2.a Clavier

Le clavier sert à la saisie de texte. Extérieurement un clavier d'ordinateur ressemble à celui d'une machine à écrire, avec éventuellement des touches spéciales supplémentaires et un pavé numérique.

Il existe deux types de claviers : mécanique et à membrane. Un clavier mécanique, le plus fréquent, est une matrice de boutons poussoirs situés au-dessus d'un réseau de fils organisé en lignes et colonnes. Lorsqu'une touche est enfoncée elle réalise un contact à l'intersection d'une ligne et d'une colonne, fermant un circuit électrique. Un circuit de contrôle du clavier détecte

ainsi la pression d'une touche. Le contrôleur est également chargé d'établir et de mémoriser le code correspondant à la touche enfoncée. Il peut alors émettre un signal d'interruption à destination de l'unité centrale.

Un clavier à membrane est constitué de quatre couches. La première est un circuit imprimé supportant un réseau de pistes conductrices. La deuxième couche constitue un isolant percé de trous au niveau des touches. La troisième couche est un circuit imprimé souple supportant un second réseau de pistes avec des bosses au niveau des touches et des trous. Lorsque l'utilisateur enfonce une touche il réalise un contact entre les deux couches conductrices. La quatrième couche, en caoutchouc, absorbe la pression du doigt et donne la sensation d'enfoncement de la touche.

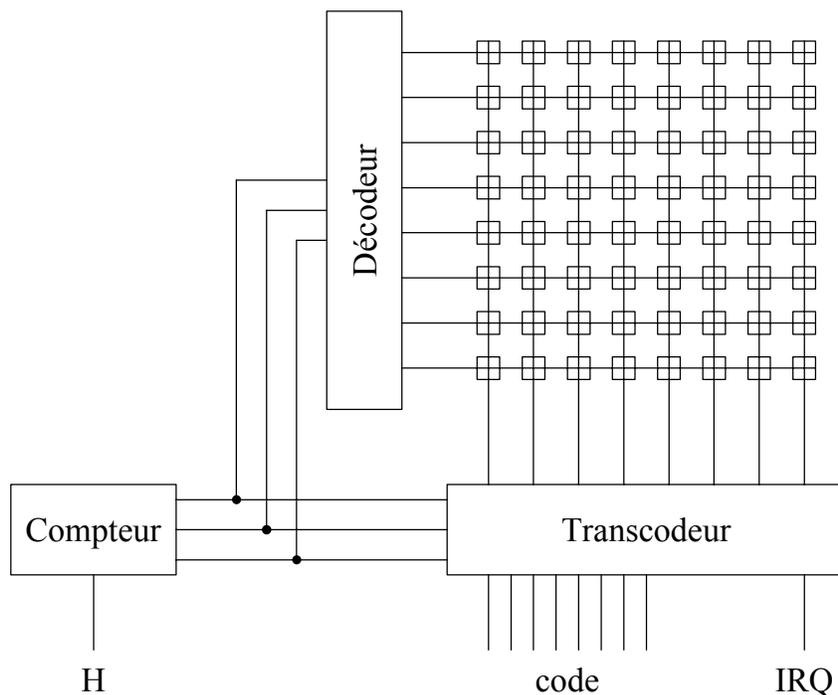


Figure 1

La figure 1 présente un schéma logique simplifié d'un clavier et de son contrôleur. Le clavier est ici constitué d'une matrice 8 x 8. Au rythme du signal d'horloge (H) un compteur modulo 8 envoie un numéro de ligne codé sur 3 bits. Le décodeur met à "1" (par exemple +5 V) la ligne correspondante. Si une touche sur cette ligne est enfoncée la colonne correspondante se trouve également à "1". Chaque touche est identifiée par un numéro de ligne, donné par le compteur, et une colonne. Un transcodeur utilisant les 3 sorties du compteur et les 8 colonnes fournit un code, ici sur 8 bits, et un signal IRQ. De cette façon les lignes sont en permanence explorées en séquence. A part pour certaines touches, comme Majuscule ou Contrôle, l'encodage des colonnes doit avoir une logique prioritaire pour le cas où deux touches d'une même ligne seraient enfoncées simultanément. Il faut également un dispositif anti-rebond. Les fonctionnalités du circuit de la figure 1 sont aujourd'hui intégrées dans un microcontrôleur.

IX.2.b Tube à rayons cathodiques

Tout comme les téléviseurs, la plupart des écrans des terminaux interactifs sont basés sur l'utilisation d'un tube à rayons cathodiques (CRT : Cathodic Ray Tube) dont le principe date de la fin du XIX^{ème} siècle. Un tube cathodique (fig. 2) comprend une source d'électrons, un système d'accélération et de focalisation et un système de déviations verticale et horizontale. Le tout est placé dans une ampoule de verre dans laquelle a été réalisé un vide poussé. Le canon à électrons est constituée d'une cathode en baryum, chauffée par un filament, suivie d'une grille à une tension modulable (wehlnet). On trouve ensuite deux anodes dont les potentiels croissants accélèrent les électrons. Deux jeux d'électrodes planes, permettent de dévier le faisceau d'électrons horizontalement et verticalement. Le système de déflexion peut également être électromagnétique. On obtient ainsi un faisceau d'électrons d'intensité modulable, très fin que l'on peut diriger en un tout point d'un écran. Cet écran est recouvert d'une substance lumineuse (par exemple phosphore) qui, au niveau du point d'impact, transforme l'énergie cinétique des électrons en lumière.

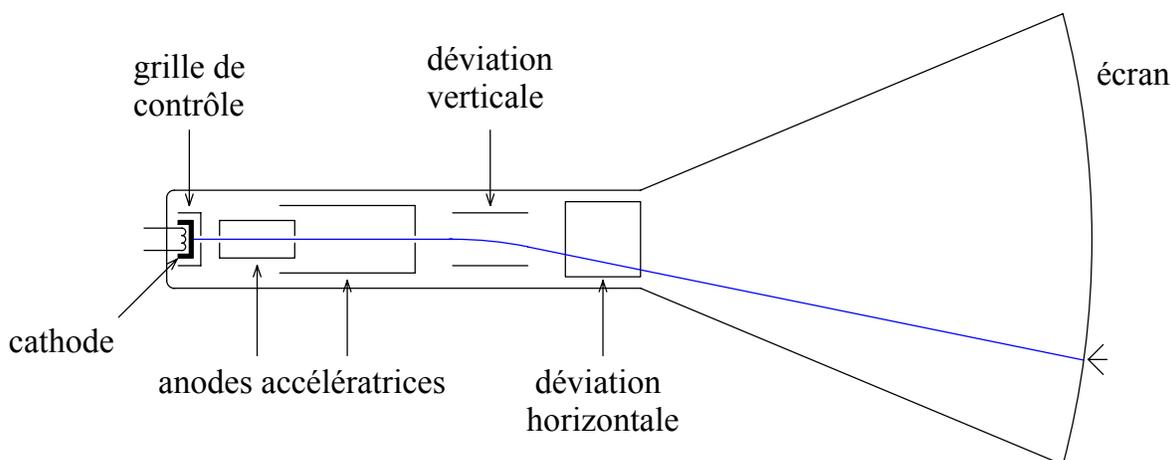


Figure 2

Pour couvrir tout l'écran le faisceau d'électrons balaie celui-ci par lignes, partant du coin en haut à gauche (en regardant l'écran). Deux modes de balayage sont possibles. Le plus simple est visualisé sur la figure 3 : les lignes de balayage sont tracées en noir et les lignes de retour (inactives) en jaune. Il est possible d'obtenir une meilleure résolution en mode entrelacé. Dans ce mode, qui est utilisé en télévision, on rafraîchit alternativement les lignes impaires puis les lignes paires. Le premier balayage part du coin en haut à gauche de l'écran, le second part du centre haut de l'écran (fig. 4). Pour éviter tout effet de clignotement de l'image il faut rafraîchir l'écran au minimum 25 fois par seconde. Il faut également prendre garde à éviter les phénomènes de battement avec la fréquence du secteur. En pratique on travaille à un taux de rafraîchissement d'au moins 60-70 Hz.

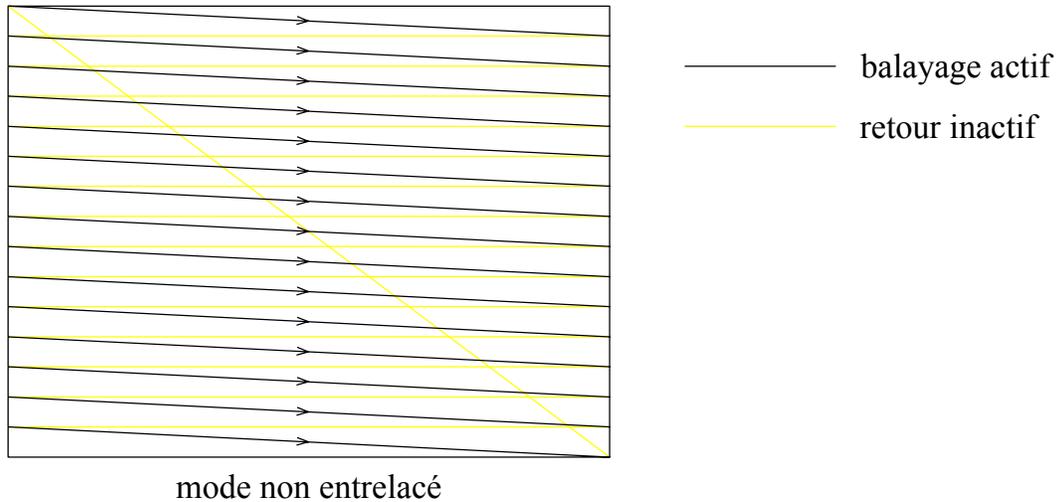


Figure 3

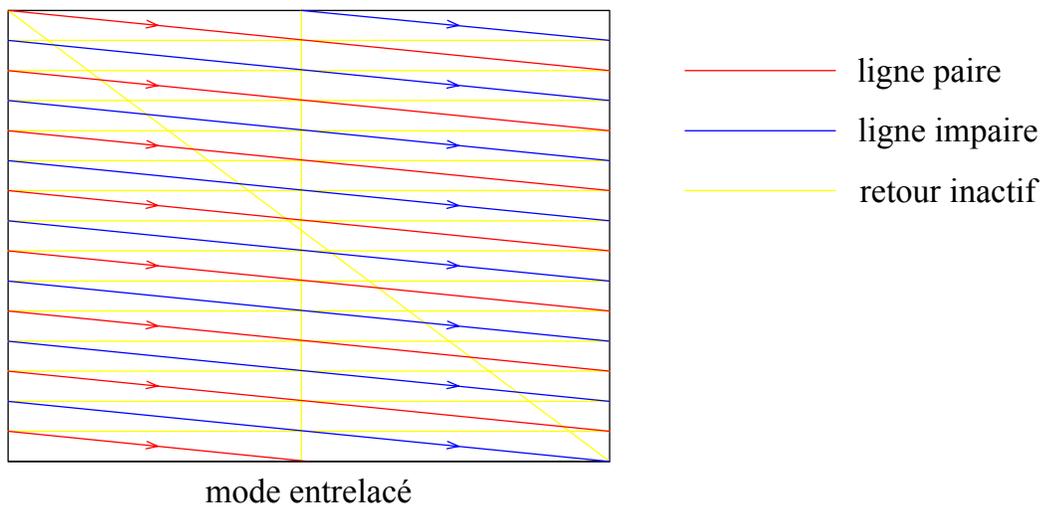


Figure 4

Chaque ligne est divisée en un certain nombre de points ou pixels (contraction de picture element). Le pixel est la plus petite fraction de l'écran qui peut être contrôlée individuellement. Le nombre de ces pixels définit la résolution de l'écran, qui s'exprime sous la forme d'un produit de deux entiers :

$$\text{Nb pixels/ligne} \times \text{nb lignes}$$

Pour un écran couleur trois faisceaux d'électrons sont nécessaires et chaque pixel est constitué de trois grains de luminophores, groupés en triangle, émettant chacun une couleur différente : rouge, vert et bleu. L'ensemble des couleurs visibles peut s'obtenir en modulant l'intensité de ces trois couleurs de base. La plupart des tubes cathodiques contiennent un masque,

plaqué contre la face interne du revêtement photosensible et percé de trous alignés avec les pixels. Le masque est disposé de façon à ce que chaque canon n'irradie que les luminophores auxquels il est associé. Les luminophores peuvent également être déposés en bandes continues verticales, les trous du masque sont alors remplacés par des fentes.

La dimension physique d'un écran est spécifiée par la longueur de sa diagonale exprimée en pouces. Pour un écran plat l'angle d'incidence du faisceau varie selon la position du point d'impact. Cela a plusieurs conséquences. Tout d'abord si la tache du faisceau au centre de l'écran est circulaire, elle devient elliptique plus on s'approche des bords et surtout des coins. Pour éviter des distorsions et un flou de l'image, la forme du faisceau doit être déformée en conséquence pour compenser cet effet. C'est également une des raisons pour lesquelles la profondeur des tubes est importante : elle ne peut être guère plus courte que la diagonale de l'écran. Si la surface de l'intersection du faisceau avec l'écran varie en fonction du point d'impact, il en est de même pour la densité de puissance reçue par l'écran et surtout le masque. La température de celui-ci n'est pas uniforme. Les déformations thermiques ont alors pour effet de dégrader l'alignement de ses trous ou fentes par rapport aux pixels. Pour éviter ce problème on utilise un matériau à faible coefficient de dilatation. Tout ceci explique pourquoi il a été difficile d'obtenir des écrans plats à coins carrés avec cette technologie.

IX.2.c Ecrans alphanumériques

Bien que ce type d'écran ait presque totalement disparu, nous allons nous intéresser au principe de gestion d'un écran alphanumérique monochrome. Puis nous étudierons quelques évolutions pour parvenir aux écrans graphiques. Afficher une image sur un écran monochrome consiste à allumer ou à éteindre chaque pixel. Mais même pour une image statique, la fréquence du rafraîchissement de l'écran ne permet pas une gestion directe de l'écran par l'unité centrale. Le contrôle de l'écran (balayages horizontal et vertical et intensité du faisceau) est confié à un contrôleur spécialisé. Une mémoire, dite de rafraîchissement, est intercalée entre l'unité centrale et le contrôleur. Les informations à visualiser sont mises à jour dans cette mémoire par l'unité centrale. A chaque cycle de rafraîchissement de l'écran le contrôleur du tube cathodique explore cette mémoire. Les caractères sont généralement codés en ASCII, pour les afficher il faut les transformer en caractères typographiques. Les polices de caractères sont définies par une matrice de $n \times m$ points, à l'intérieur d'une cellule tenant compte de la séparation des caractères. Chaque caractère est défini par les points qui sont à afficher dans la matrice. Il peut donc être représenté par m mots de n bits, les bits à 1 indiquant les points à visualiser :

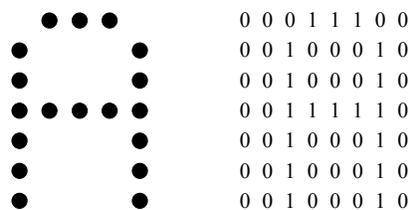


Figure 5

Ces polices de caractères peuvent se trouver en mémoire morte (ROM) pour les polices définies par le constructeur ou être téléchargées par l'utilisateur en mémoire vive (RAM).

En mémoire de rafraîchissement chaque caractère à visualiser est stocké dans une zone pouvant comporter plusieurs champs (fig. 6) : code (par ex. : ASCII), numéro de police et attributs (vidéo inverse, gras, double largeur, souligné, etc...). Cette méthode permet d'utiliser une police et des attributs différents pour chaque caractère à imprimer.

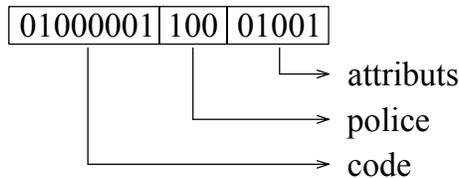


Figure 6

Le contrôleur d'écran doit périodiquement engendrer les signaux de commande du balayage et de l'intensité du faisceau d'électrons. Il boucle en permanence sur les lignes de caractères contenues en mémoire d'écran. Pour chaque ligne il y a m balayages horizontaux à réaliser. Pour chacun de ces balayages le contrôleur d'écran (fig. 7) effectue une boucle sur les colonnes de la ligne et réalise les opérations suivantes :

- Recherche en mémoire de rafraîchissement de l'information (code, police, etc.) sur le caractère à visualiser. Celui-ci est identifié (adressé) par un numéro de ligne et un numéro de colonne.
- Recherche en ROM ou RAM, du mot de la matrice de points correspondant au balayage, au caractère et à la police.
- Emission du signal de commande de l'intensité du faisceau d'électrons en fonction des n bits identifiant les points à allumer.

La figure suivante illustre un concept pouvant s'appliquer à un tel contrôleur. Supposons que l'écran comprend 2^l lignes, comptant chacune 2^c caractères et que tout caractère est défini par une matrice $2^f \times 2^b$ pixels. L'écran est donc constitué de $N = 2^{l+c+r+b}$ pixels au total. Il est donc possible d'utiliser un compteur binaire de $l + c + r + b$ bascules et d'utiliser les sorties de la manière suivante :

- Les l bits de haut poids permettent d'identifier la ligne de caractères à afficher.
- Les r bits suivants (rang sur la figure 7) correspondent aux 2^r balayages horizontaux par ligne de caractères.
- Les c bits suivants permettent d'identifier la colonne contenant le caractère à afficher.
- Les derniers bits comptent les pixels horizontaux dans le caractère à afficher.

Les $l + r$ bits de haut poids identifient la ligne de l'écran en cours de balayage, alors que les $c + b$ bits de bas poids identifient la position horizontale du pixel le long de cette ligne. Ces deux

informations servent à contrôler le mouvement permanent du faisceau d'électrons. On utilise pour cela deux convertisseurs numériques analogiques dont les tensions de sortie commandent les deux jeux de plaques de déflexion.

Un autre convertisseur commande la tension de la grille. Comme il ne reçoit ici en entrée qu'un seul bit provenant d'un registre à décalage il émet un signal de commande en tout ou rien.

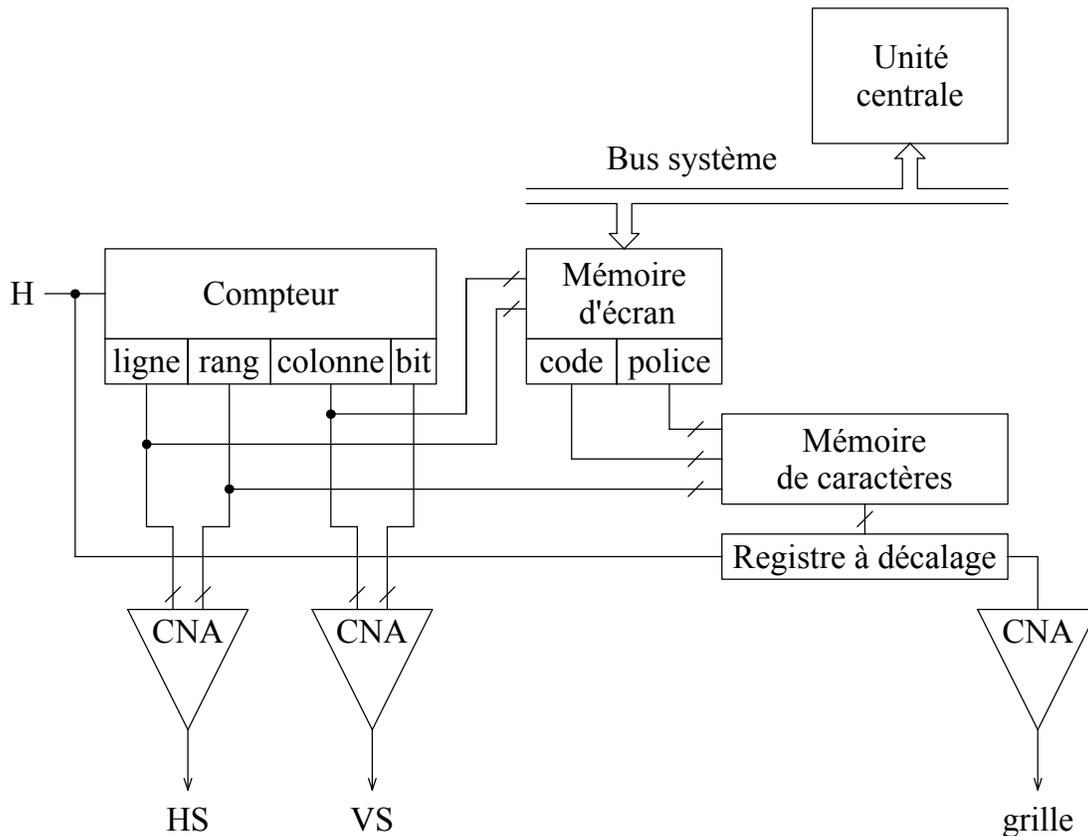


Figure 7

Nous constatons que la mémoire d'écran qui sert d'interface entre l'unité centrale et le contrôleur du tube à rayons cathodiques est une mémoire à double accès. Un accès en écriture depuis le microprocesseur et un accès en lecture pour le contrôleur.

Le système de codage employé en mémoire d'écran dans l'exemple précédent nécessite une étape de décodage avant d'envoyer le signal de commande vidéo (grille) au tube. Une méthode permet d'augmenter le taux de rafraîchissement en évitant cette étape. Il s'agit d'utiliser une mémoire dans laquelle chaque bit correspond à un pixel : "1" ou "0" selon que le pixel associé est allumé ou non. Cette très grande matrice constitue une image binaire de l'écran : mémoire de trame ou bitmap. Le contrôleur peut alors extraire directement le signal de commande de cette mémoire. Un gain de vitesse est obtenu au prix de plus de mémoire. Cette technique est largement employée dans les écrans graphiques.

IX.2.d Ecrans graphiques

Les écrans graphiques permettent l'affichage de caractères et de dessins ou d'images. Il existe deux types d'écrans graphiques : les écrans à mémoire de trame et les écrans à balayage cavalier.

Résumons le principe de la mémoire de trame que nous avons introduit plus haut. L'écran est divisé en petits domaines élémentaires : les pixels. Ces pixels sont répartis en lignes et colonnes. Pour les écrans monochromes chaque pixel possède deux états possibles : éteint ou allumé. Il est donc possible d'avoir une représentation logique d'un tel écran en associant un bit à chaque pixel. On obtient ainsi une image binaire : la mémoire de trame. Le contrôleur du tube cathodique balaie en parallèle ligne par ligne cette mémoire et l'écran, chaque bit à 1 est affiché. Le contrôle de l'intensité se fait en tout ou rien. Initialement cette mémoire de trame contenait l'image des caractères à afficher. Son usage s'est très vite généralisé pour l'affichage de dessins.

L'étape suivante a consisté à passer du "noir et blanc" à des niveaux de gris. Pour ce type d'écran il faut moduler l'intensité du faisceau d'électrons pour chaque pixel. Il faut alors plusieurs bits (par exemple 8 bits pour représenter 256 niveaux de gris) par pixel. Un convertisseur numérique analogique est alors utilisé pour commander l'intensité du canon à électrons. Le principe est toujours le même pour les écrans en couleur : il faut moduler les trois couleurs de base (bleu, rouge et vert). Leur mélange donne naissance aux autres couleurs. On peut utiliser trois mémoires de trame. Avec 1 bit par couleur fondamentale on peut obtenir huit couleurs, avec 8 bits par couleur on dispose d'une palette de seize millions de couleurs ($2^{24} = 16 \cdot 10^6$). Une autre solution consiste à avoir une seule mémoire, avec n bits par pixel. Entre la mémoire de trame et les trois convertisseurs numériques analogiques il faut intercaler une table de correspondance (look-up table) en mémoire morte ou vive. Cette table de couleurs permet d'associer à chacune des 2^n combinaisons une couleur particulière, définie par trois intensités. Lorsqu'elle est rangée en mémoire vive cette table peut être redéfinie par l'utilisateur.

Le rafraîchissement de l'écran est géré par le contrôleur du tube cathodique qui explore périodiquement la mémoire de trame. Lorsque l'image doit être modifiée l'unité centrale ne doit intervenir qu'au niveau de cette mémoire. Cependant compte tenu de l'importance de plus en plus grande du graphisme dans les applications et en particulier du 3D dans les jeux et les applications multimédia, la mise à jour de la mémoire de trame mobilise de plus en plus l'unité centrale et engorge le bus système. Une solution à ce problème consiste à insérer entre l'unité centrale et la mémoire de trame un processeur spécialisé (fig. 8). L'unité centrale transmet les actions à réaliser à ce processeur graphique grâce à des macro-instructions (tracer une droite, un cercle, déplacer l'image, etc.). Le processeur graphique se charge de la mise à jour de la mémoire de trame. L'unité centrale est alors libérée pour ses autres tâches et le trafic sur le bus système nettement réduit. L'architecture des processeurs graphiques s'est très vite spécialisée. Cette interface graphique est reliée à la mémoire de trame, encore appelée mémoire vidéo (VRAM), par un bus réservé. Une caractéristique importante de cette mémoire vidéo est son double accès en écriture et en lecture.

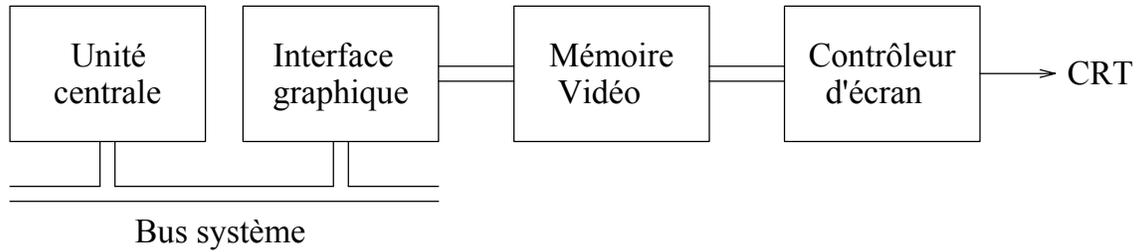


Figure 8

Dans un écran à balayage cavalier on utilise le faisceau d'électrons comme un crayon pour dessiner. Il est programmé pour dessiner une image par segments. Cette technique, adaptée aux structures filiformes, est utilisée dans certaines stations de conception assistée par ordinateur (CAO) car elle permet une très grande finesse du trait.

IX.2.e Ecrans à cristaux liquides

L'apparition des écrans plats à cristaux liquides (LCD : Liquid Crystal Display) a permis, grâce à leur encombrement réduit et leur faible consommation, l'éclosion des "portables". Les cristaux liquides sont des liquides dont les molécules longues s'organisent selon un réseau comme les atomes d'un cristal. Elles se disposent en couches dans lesquelles elles ont toutes la même orientation.

Il est possible de contrôler l'orientation de ces molécules en les plaçant au contact avec une surface très finement gravée. Dans un écran à cristaux liquides ceux-ci sont pris en sandwich entre deux plaques de ce type dont les rainures sont perpendiculaires. Les molécules dans les deux couches au contact avec ces plaques ont donc des orientations orthogonales. L'orientation dans les couches intermédiaires varie continûment avec la profondeur de 0° à 90° .

Les longues molécules des cristaux liquides se polarisent facilement. Cette caractéristique a deux conséquences qui sont mises à profit dans les écrans à cristaux liquides. Si on place ces molécules dans un champ électrique elles s'alignent parallèlement à celui-ci. Par ailleurs l'interaction collective de ces molécules avec le champ électrique de la lumière a pour effet de faire tourner le plan de polarisation de la lumière pour le rendre parallèle à l'orientation des molécules.

Dans un écran à cristaux liquides ceux-ci sont placés entre deux filtres polarisants croisés. La polarisation de chaque filtre est parallèle à celle des rainures de la face en regard. Ainsi la lumière qui pénètre dans le liquide est-elle polarisée parallèlement avec la première couche de molécules. Au fur et à mesure que la lumière traverse le liquide son plan de polarisation tourne avec les molécules. A la sortie le plan de polarisation a donc pivoté de 90° pour se trouver parallèle au second filtre polarisant. Celui-ci laisse donc passer la lumière : l'écran est transparent.

On place deux électrodes planes de part et d'autre des cristaux liquides. Lorsque ces électrodes sont soumises à une tension elles créent un champ électrique perpendiculaire à l'écran. Les molécules s'alignent avec ce champ et n'interagissent alors plus avec la lumière dont le plan de polarisation n'est pas modifié par la traversée du liquide. La lumière est alors absorbée par le second filtre : l'écran est opaque. Cependant l'alignement des molécules avec le champ électrique n'est pas instantané. La vitesse de ce réarrangement dépend de l'intensité du champ électrique et de la viscosité du liquide. Il est ainsi possible de moduler la quantité de lumière transmise.

La figure suivante illustre le principe d'un écran à cristaux liquides dit à matrice passive. Il est constitué de plusieurs couches de matériaux. On trouve dans le sens du parcours de la lumière :

- Un filtre polarisant.
- Une plaque de verre.
- Une série d'électrodes en oxyde métallique transparent en forme de longues bandes parallèles.
- Une plaque de polymère gravée avec de très fines rainures parallèles à la polarisation du premier filtre.
- Les cristaux liquides.
- Une seconde plaque d'alignement aux rainures orthogonales à celles de la première.
- Une seconde série d'électrodes en bandes orthogonales aux premières électrodes.
- Une couche de protection.
- Un filtre coloré constitué de bandes rouges, vertes et bleues parallèles et alternées.
- Une plaque de verre.
- Un second filtre polarisant croisé par rapport au premier.

Tous les matériaux utilisés sont transparents. Les deux plaques de verre servent de supports mécaniques aux autres couches. L'épaisseur du liquide est contrôlée par de fines entretoises.

Les deux couches d'électrodes correspondent aux lignes et aux colonnes. Un pixel est défini par l'intersection de deux électrodes. Comme dans un tube à rayons cathodiques les pixels sont balayés périodiquement. Une différence de potentielle est appliquée sur les deux électrodes (ligne et colonne) correspondantes pendant un petit intervalle de temps. L'amplitude de cette tension permet de moduler d'intensité lumineuse du pixel correspondant. Sa couleur est définie par le filtre coloré. La juxtaposition de pixels avec les trois couleurs de base permet d'obtenir les autres couleurs.

La face arrière de l'écran est éclairée par des tubes fluorescents basse tension dont la lumière blanche est répartie uniformément sur toute la surface par un diffuseur. Cette technique permet la réalisation d'écrans plats, de faibles épaisseurs et à consommations réduites. Contrairement aux tubes cathodiques pour lesquels la lumière est émise au niveau de l'écran dans un grand angle solide, il faut regarder un écran à cristaux liquides de face.

Très bon marché cette technologie à matrice passive présente quelques inconvénients. Elle est assez lente ce qui réduit le taux de rafraîchissement. Lorsqu'une tension est appliquée entre une ligne et une colonne le champ électrique influence les pixels voisins. Une autre technique, dite à matrice active, permet de remédier à ces difficultés. Les électrodes sont segmentées petits

carrés correspondants aux pixels. Ce découpage permet d'éviter d'influencer les pixels voisins. Chaque pixel est commandé par un transistor, ce qui permet un gain important en vitesse. Ces transistors sont fabriqués sur des films de silicium très fins déposés sur un substrat de quartz de grandes dimensions. Ces films contiennent plusieurs millions de transistors qui ne doivent pas présenter de défaut. Les films de transistors peuvent maintenant être réalisés avec du silicium polycristallin déposé sur un substrat en verre (le dépôt du silicium se fait à plus basse température que pour obtenir un monocristal). On cherche à utiliser du silicium amorphe.

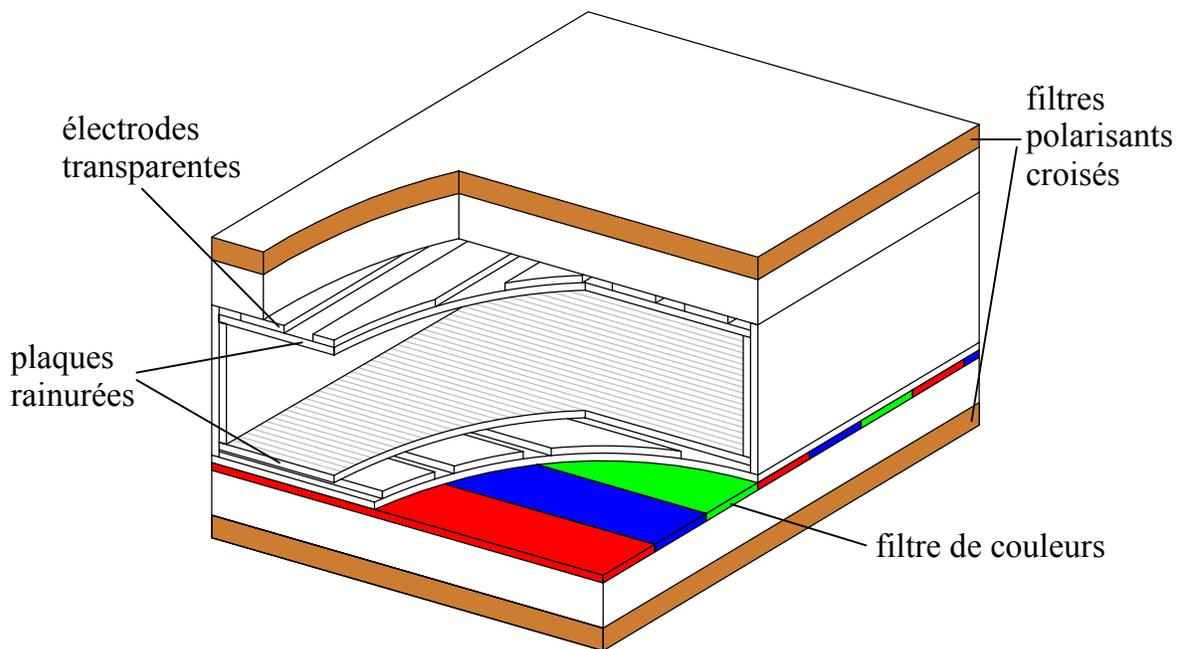


Figure 9

IX.2.f Souris

Outre le clavier l'utilisateur communique avec un ordinateur par l'intermédiaire d'une interface graphique et d'une souris. La plupart des souris actuelles sont optoélectroniques. Une boule en acier recouverte de caoutchouc entraîne deux roues dont les axes sont perpendiculaires. Une troisième roue équipée d'un ressort maintient la boule en contact avec les deux autres roues. Sur l'axe de chacune de ces deux roues se trouve un disque percé de fentes radiales. Chaque disque tourne entre deux cellules décalées. Chaque cellule contient une diode électroluminescente et une photodiode. Le comptage des impulsions lumineuses permet de mesurer le déplacement de la souris dans une direction. Le décalage entre les deux cellules permet de déterminer le sens du mouvement.

Un crayon lumineux est un détecteur optique, qui en contact avec l'écran, permet de repérer le passage du faisceau d'électrons du tube cathodique. Il envoie un signal au contrôleur vidéo qui peut ainsi déterminer sa position par rapport à l'écran.

IX.3 Imprimantes

Dès les débuts de l'informatique, l'imprimante est apparue comme un outil indispensable pour fournir à l'homme les résultats des calculs. Il existe une grande variété d'imprimantes : à impact, thermiques, à jet d'encre, à laser, etc.

IX.3.a Imprimantes à impact de marteaux

Parmi ces imprimantes citons les imprimantes à marteaux. Il y a un marteau par colonne. Les caractères sont gravés en relief sur un ruban métallique qui tourne en permanence devant les marteaux. Un ruban encreur est placé entre le ruban métallique et le papier. Le papier avance ligne par ligne. L'électronique de l'imprimante déclenche la frappe de chaque marteau lorsque le bon caractère gravé passe devant : frappe à la volée. Cette technologie permet de réaliser des imprimantes très rapides (jusqu'à 3000 lignes par minute). Dans certains modèles le ruban métallique peut être remplacé par un tambour horizontal ou une "marguerite". Ces premiers types d'imprimantes, parmi lesquelles nous pouvons encore mentionner l'imprimante à boule d'IBM, sont les seules à disposer de lettres préformées, assurant une excellente qualité d'impression (comparable à celle des machines à écrire). Par contre les polices de caractères sont limitées en nombre et elles sont incapables d'imprimer des graphiques.

IX.3.b Imprimantes à aiguilles

Dans ces imprimantes les caractères sont représentés par des points dans une grille ou matrice. La tête d'impression comporte autant d'aiguilles que de points par colonne. La tête se déplace horizontalement devant la ligne à imprimer. Pour chaque caractère les aiguilles correspondant aux points à marquer sont lancées par de petits électro-aimants. Elle frappent ainsi le ruban encreur situé entre la tête et le papier. Compte tenu de l'inertie des aiguilles il est difficile de réaliser des imprimantes rapides. Les polices de caractères sont stockées en mémoire. Dans certains cas elles peuvent être téléchargées par l'ordinateur. De plus elles peuvent imprimer point par point des graphiques et des dessins. Simple et fiable, cette technologie permet de produire des imprimantes bon marché. Permettant de faire des copies carbone elles sont très utilisées dans les administrations.

La qualité des imprimantes est déterminée par la densité de points exprimée en dpi (dots per inch) ou ppp (points par pouce).

IX.3.c Imprimantes thermiques

Les imprimantes thermiques ressemblent aux imprimantes matricielles à aiguilles. Au lieu de frapper un ruban les aiguilles chauffent un papier spécial sensible à la chaleur. Peu utilisées pour l'impression monochrome, elles permettent l'impression couleur.

On utilise un rouleau de cellophane recouvert d'une encre solide, composé d'une succession de régions jaune, cyan, magenta (les trois couleurs primaires) et noire éventuellement. Une série de têtes thermiques est disposée perpendiculairement au sens de défilement du papier. L'impression de chacune des couleurs primaires nécessite un passage du papier devant les têtes. L'impression se fait donc en trois ou quatre passes. Il existe deux technologies. Dans le cas du transfert thermique on utilise une encre à base de cire qui fond lorsqu'une tête thermique est activée. Par pression mécanique on obtient un point. Les couleurs de base (rouge, vert, bleu et noir) sont obtenues par superposition des couleurs primaires. Les autres couleurs sont obtenues par juxtaposition de points. La résolution est d'environ 300 ppp. Pour la technologie de la sublimation thermique le colorant passe de l'état solide à l'état gazeux lorsqu'une tête thermique est activée. Il diffuse alors sur le papier où il revient à l'état solide. La quantité de colorant vaporisée peut être contrôlée par la température de la tête thermique. Il est donc possible de définir l'intensité des couleurs primaires. Par superposition on peut disposer d'une très large palette de couleurs. L'impression sans trame est d'excellente qualité.

IX.3.d Imprimantes à jet d'encre

Dans une imprimante à jet d'encre la tête de lecture est composée de petits tubes qui projettent de minuscules gouttelettes d'encre. Ces gouttelettes, chargées électrostatiquement, sont dirigées vers le papier avec une grande précision par un système de plaques de déviation. L'absence d'impact mécanique rend son fonctionnement très silencieux. D'autre part, en augmentant le nombre des tubes il est possible imprimer simultanément plusieurs couleurs. L'impression est de très bonne qualité. La résolution varie entre 200 ppp et 400 ppp.

IX.3.e Imprimantes à laser

Le fonctionnement des imprimantes à laser s'inspire de celui des photocopieuses. L'image de la page à imprimer est construite sur un tambour photoconducteur. Cette image est obtenue avec un faisceau laser réfléchi sur un miroir mobile, piloté par l'électronique de l'imprimante. Le tambour passe devant une station d'encrage où l'encre sèche, chargée électrostatiquement, est attirée aux endroits chargés du tambour. L'image est ensuite transférée sur le papier par contact. Le papier est chauffé pour fixer définitivement les particules d'encre. Cette technique permet l'impression sur papier ordinaire. On atteint des résolutions variant entre 300 et 2000 ppp. Dans certains modèles les polices de caractères sont définies géométriquement. Il est alors possible d'obtenir une qualité proche de celle d'un imprimeur.

IX.4 Supports magnétiques

IX.4.a Enregistrement magnétique

Le principe de l'enregistrement magnétique est identique à celui mis en œuvre dans un magnétophone. On utilise les propriétés magnétiques de certains matériaux ferromagnétiques comme l'oxyde de fer ou certains alliages (phosphore-nickel ou phosphore-cobalt). Une couche magnétique, d'une épaisseur de quelques microns ou même inférieure au micron aujourd'hui, est déposée sur un support souple ou rigide. Elle est constituée de domaines de Weiss microscopiques. Initialement l'orientation de l'aimantation de chacun de ces domaines est aléatoire et l'aimantation résultante est nulle. En présence d'un champ magnétique, suffisamment intense, l'aimantation de chaque domaine bascule pour être parallèle au champ inducteur. Lorsque celui-ci disparaît l'aimantation de chaque domaine persiste (rémanence) dans cette position, induisant une aimantation globale. Il est donc possible d'aimanter certaines zones ou cellules de la couche magnétique.

Ce principe est mis en œuvre dans une tête d'écriture (fig. 10). Celle-ci est constituée par l'entrefer d'un électro-aimant. En alimentant la bobine on crée un champ magnétique au voisinage de l'entrefer, son orientation dépend du sens du courant. Ce champ traverse la couche magnétique. Les domaines de Weiss en regard de la tête prennent une aimantation parallèle au champ magnétique. Ils constituent une cellule dont l'aimantation, parallèle aux lignes de champ dans l'entrefer, a deux orientations possibles. Cette technique est donc adaptée à un codage binaire.

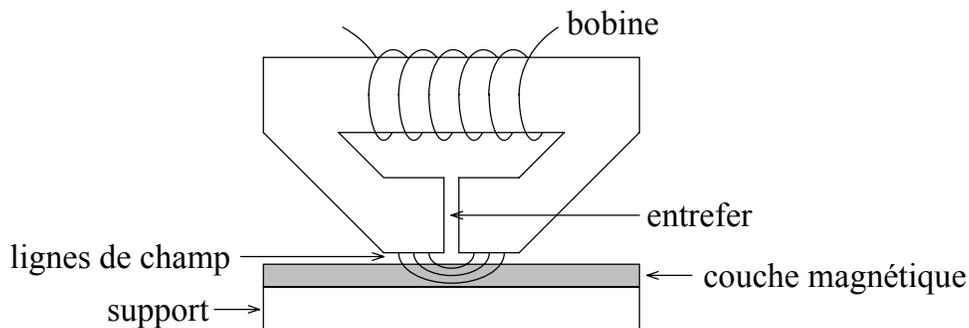


Figure 10

Pour la lecture on peut utiliser la même tête non alimentée. Chaque cellule constitue un aimant permanent. Lorsqu'une cellule passe devant l'entrefer, par induction elle modifie le flux magnétique traversant la bobine, ce qui crée un courant. Le sens de ce courant dépend du sens de variation de l'aimantation, donc de l'information binaire stockée dans la cellule.

La densité linéaire d'enregistrement se mesure en bpi (bits per inch). Cette densité dépend de la nature de la couche magnétique, de la distance entre cette couche et la tête de lecture/écriture et de l'ouverture de l'entrefer.

Une tête de lecture est sensible aux variations du flux magnétique, donc aux transitions, ou changements de sens d'aimantation. Différentes techniques d'enregistrement sont utilisées pour le codage des disques et des bandes magnétiques. Nous allons en étudier quelques-unes, illustrées par la figure 11.

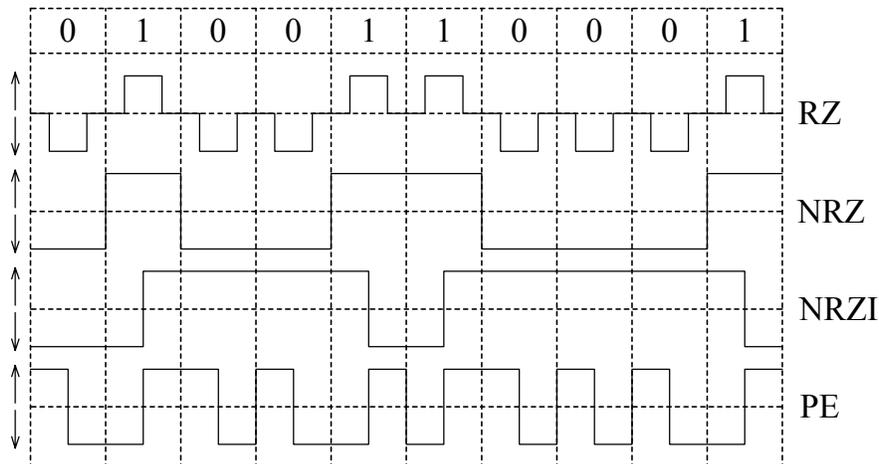


Figure 11

RZ (Return to Zero) On revient à zéro entre chaque bit. On sépare les cellules magnétisées par une zone non aimantée. La valeur 1 correspond à une impulsion positive et la valeur 0 à une impulsion négative au niveau de la bobine. L'enregistrement est autosynchronisé : il n'y a pas besoin de base de temps externe pour compter tous les bits. Par contre ce codage n'est pas favorable à la densité.

NRZ (Non Return to Zero) On supprime les zones non aimantées. Il n'y a une transition, donc une impulsion, uniquement lorsque les informations passent de 0 à 1 et de 1 à 0. On peut ainsi augmenter la densité. Cependant, comme une séquence de bits égaux engendre une impulsion seulement au début, il faut une base de temps pour compter les bits.

NRZI (NRZ Inverted) Le sens de l'aimantation ne porte plus directement l'information. Seuls les bits à 1 provoquent une inversion de magnétisation. Il est nécessaire d'avoir une base de temps.

PE (Phase Encoding) Chaque cellule est divisée en deux parties magnétisées en sens opposé. Il y a donc une transition au milieu de chaque cellule, même dans une séquence de bits égaux. Le sens de cette transition indique la valeur binaire de l'information. Ce codage est autosynchronisateur. Il est facile de conserver la base de temps, repérant les milieux des

cellules, en phase avec la vitesse de défilement. Cette technique est cependant limitée par la fréquence des transitions.

GCR (Group Code Recording) Pour atteindre des densités supérieures, on reprend le principe du NRZI en supprimant les séquences de zéros pour faciliter la synchronisation de la base de temps. Pour cela on utilise 5 bits pour coder des groupes de 4 bits selon la table 1. Chaque groupe de 4 bits est codé avant d'être enregistré, en utilisant la méthode NRZI, et décodé à la lecture. Avec ce code il est impossible de trouver plus de deux zéros consécutifs. Dans ces conditions il est plus facile de recadrer en permanence la base de temps sur la fréquence de défilement des informations.

| Donnée (4 bits) | Code (5 bits) | Donnée (4 bits) | Code (5 bits) |
|-----------------|---------------|-----------------|---------------|
| 0000 | 11001 | 1000 | 11010 |
| 0001 | 11011 | 1001 | 01001 |
| 0010 | 10010 | 1010 | 01010 |
| 0011 | 10011 | 1011 | 01011 |
| 0100 | 11101 | 1100 | 11110 |
| 0101 | 10101 | 1101 | 01101 |
| 0110 | 10110 | 1110 | 01110 |
| 0111 | 10111 | 1111 | 01111 |

Table 1

IX.4.b Disques durs

Un disque dur constitue la mémoire de masse non volatile d'un ordinateur. Le premier disque a été développé en 1954 par IBM.

Une unité de disques magnétiques est généralement constituée de plusieurs disques ou plateaux tournant à la même vitesse (plusieurs milliers de tours par minute). Un plateau est constitué d'un substrat non magnétique, usiné et poli. On utilise soit l'aluminium, soit le verre ou la céramique plus résistants et plus fins. Les deux faces d'un plateau sont recouvertes d'une couche magnétisable (oxyde métallique ou métal). Entre ces plateaux se trouvent les têtes de lecture/écriture montées sur un même bras mobile (fig. 12).

Les disques durs utilisent des têtes ultra-légères, qui grâce à leur profil aérodynamique flottent sur un coussin d'air. L'altitude de vol varie entre 0.25 et 2.5 μm . Comme le montre la figure 13, où nous n'avons pas pu visualiser un cheveu de diamètre 50 à 100 μm , cela impose de grandes contraintes sur la planéité des surfaces et leur propreté. Pour cela les disques durs doivent être placés dans des boîtiers hermétiques, exempt de poussières (technologie Winchester, car mise au point pour l'IBM 3030, calibre de la carabine Winchester). A l'arrêt les têtes se posent sur une partie du plateau préparée à cet effet ne contenant pas de données. Le ralentissement de la rotation du disque, avec diminution progressive du film d'air entraîné, permet un atterrissage en douceur.

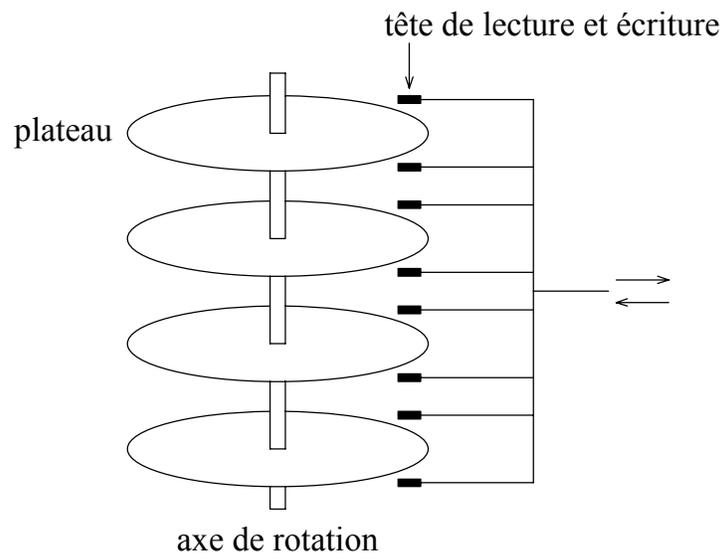


Figure 12

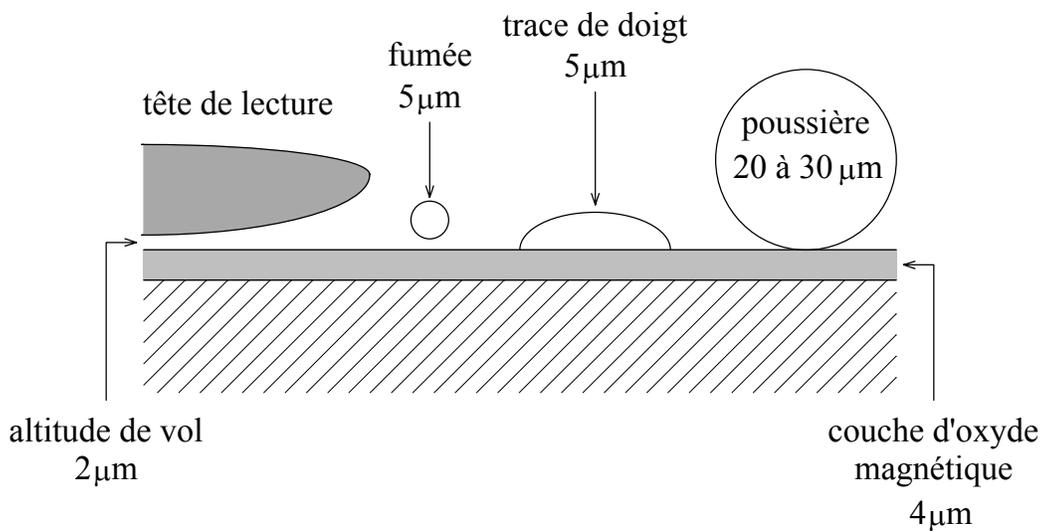


Figure 13

Les surfaces de chaque disque sont structurées (fig. 14) en pistes (track) : cercles concentriques centrés sur l'axe de rotation. Un disque est également divisé en segments, ou "parts de gâteau" d'angle fixe. L'intersection d'une piste et d'un segment est appelée secteur. La taille des secteurs est généralement de 512 octets. Le secteur représente l'unité élémentaire d'accès d'un

disque. Chaque secteur est séparé du suivant par une petite zone non utilisée qui permet de repérer les secteurs et de synchroniser les signaux de commande. L'ensemble des pistes de tous les plateaux à un rayon donné forment un cylindre. Comme les têtes de lecture/écriture sont solidaires, les secteurs d'un cylindre peuvent être accédés simultanément. Le bras supportant ces têtes pivote autour d'un axe et sa position est définie par une bobine électromagnétique. Ce système permet un positionnement rapide et précis de la tête sur la piste. Ce positionnement est par ailleurs asservi sur des repères placés sur le disque.

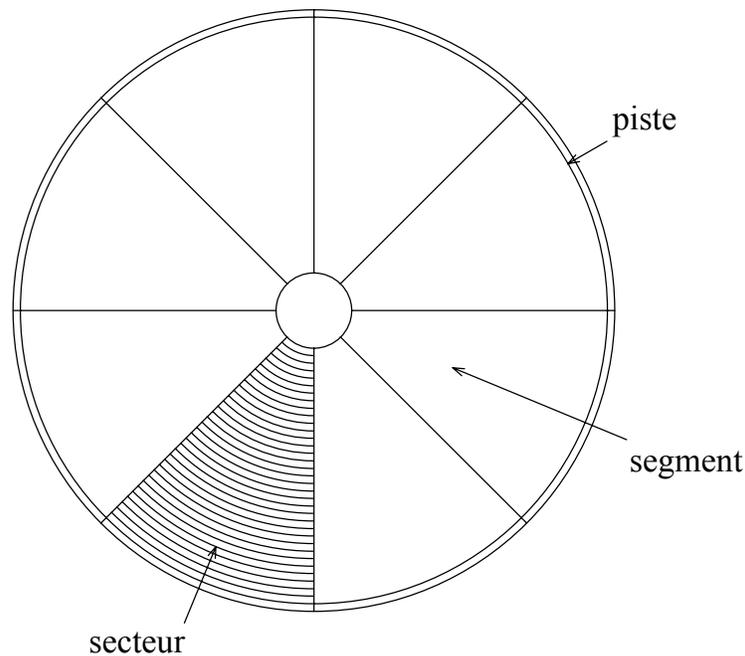


Figure 14

La surface d'un disque vierge est uniforme, aucun repérage absolu n'est possible. Avant de pouvoir l'utiliser il faut réaliser un premier formatage dit de bas niveau, c'est-à-dire mettre en place sur la surface des repères permettant de retrouver pistes et secteurs. Chaque secteur est généralement structuré en deux zones : un entête permettant son repérage et les données. Cette structure dépend de l'architecture, on peut trouver :

- un intervalle de secteur : 10-20 octets à 0 ;
- un label précédé d'un marqueur : adresse du secteur (numéros de disque, de face, de cylindre et de secteur) ;
- un code détecteur d'erreur ;
- une zone de synchronisation : permet de passer de la lecture d'adresse à l'écriture des données, sert également au recadrage des signaux d'horloge ;
- un marqueur de données ;
- la zone de données.

Les données sont écrites et lues par secteurs. Pour identifier un bloc de données il faut spécifier le numéro de l'unité de disques, la tête (donc le plateau et la face), la piste (ou le cylindre) et le secteur. En fait la capacité des disques durs et la taille moyenne des fichiers ont augmenté de façon telle qu'il n'est plus possible d'adresser individuellement chaque secteur. Un disque de 1 Go contient 2^{21} secteurs. Ceux-ci sont groupés en grappes ou clusters. Cette organisation, la taille des grappes et l'adressage logique des secteurs et des clusters dépendent du système d'exploitation. Cela requiert un formatage logique, avec préparation de vecteurs d'état, de tables d'allocation des fichiers, etc. Un vecteur d'état permet de savoir si un secteur ou cluster est utilisable, utilisé ou libre.

La capacité d'un disque est exprimée en multiples de 1024 octets. La table suivante résume la valeur des préfixes utilisés pour exprimer les capacités des disques. J'ai quelque peu anticipé certains multiples ne servent pas encore même pour les sauvegardes.

| Symbole | Préfixe | Capacité | |
|---------|---------|----------|--------------|
| k | kilo | 2^{10} | = 1024 |
| M | méga | 2^{20} | = $(1024)^2$ |
| G | giga | 2^{30} | = $(1024)^3$ |
| T | téra | 2^{40} | = $(1024)^4$ |
| P | péta | 2^{50} | = $(1024)^5$ |
| E | exa | 2^{60} | = $(1024)^6$ |
| Z | zetta | 2^{70} | = $(1024)^7$ |
| Y | yotta | 2^{80} | = $(1024)^8$ |

Table 2

L'accès aux données sur un disque dur peut se décomposer en trois étapes :

- positionnement des têtes au niveau du cylindre;
- passage du secteur devant la tête de lecture et repérage;
- transfert des données.

Un disque est dit à accès semi-séquentiel : accès direct à la piste avec une constante de temps liée à la mécanique, puis accès séquentiel au secteur dans la piste. Le temps de déplacement de la tête est appelé temps de recherche. Il dépend de l'amplitude du mouvement du bras (passage à un cylindre voisin ou traversée complète du disque) et de l'inertie de celui-ci. Une fois la tête positionnée, le temps d'accès au secteur sur la piste est appelé temps de latence. Celui-ci dépend des positions initiales respectives de la tête et du secteur, en moyenne il faut parcourir la moitié de la piste. Il dépend également de la vitesse de rotation des plateaux. Ensuite le taux de transfert des données est caractérisé par débit.

Il existe deux configurations possibles des pistes. Elles peuvent contenir toutes le même nombre de secteurs (fig. 14). Dans ce cas, comme la vitesse de rotation est constante, la densité d'information varie en fonction du rayon de la piste, mais le débit est constant. Une autre organisation consiste à garder la densité d'information à peu près constante. Dans ce cas le

nombre de secteurs par piste varie en fonction du rayon. Le disque est divisé en zones, le nombre de secteurs par piste étant identique à l'intérieur d'une zone. La capacité du disque est alors accrue et le débit dépend également du rayon. Il est plus important sur le bord des plateaux. Le système doit donc chercher à installer les fichiers d'utilisation fréquente à cet endroit.

IX.4.c Disquettes

Dans leur principe les disquettes fonctionnent comme les disques durs. Amovibles les disquettes sont principalement utilisées dans les petits systèmes. Les premières disquettes avaient un diamètre de 8 pouces, puis sont apparues les disquettes 5"1/4 et enfin 3"1/2. Les disquettes sont enfermées dans une enveloppe souple ou rigide dont les faces internes sont couvertes d'un revêtement limitant le frottement.

Outre qu'une disquette est souple, la différence principale avec un disque dur réside dans la tête de lecture/écriture. Tout d'abord cette tête est en contact avec la surface magnétique, ce qui limite la vitesse de rotation : 300 tours/minute pour les disquettes 3" 1/2. De plus le positionnement de la tête est obtenu par une vis actionnée par un moteur pas à pas, sans asservissement. Ce système nécessite une marge de sécurité entre les pistes, ce qui limite la densité : 135 tpi (tracks per inch) pour les classiques disquettes de 1.44 Mo.

Cette capacité est de moins en moins adaptée à la taille des fichiers. Citons deux techniques employées pour augmenter la capacité et le débit dans d'autres types de supports magnétiques amovibles. La première consiste à faire flotter le disque souple entre les deux têtes de lecture/écriture (effet Bernoulli). Cela permet une plus grande vitesse de rotation, donc un débit plus élevé. Une autre consiste à utiliser un faisceau laser pour asservir le déplacement de la tête. Le disque dispose des pistes optiques de référence. Ce système permet un positionnement très précis de la tête, donc une plus grande densité de pistes.

IX.4.d Bandes et cartouches magnétiques

Dans les bandes ou cartouches magnétiques le support de la couche magnétisable est constitué par un ruban souple comparable à ce qui est utilisé pour les enregistrements audio ou vidéo. Une bande est logiquement divisée en pistes parallèles, par exemple neuf pistes (fig. 15). Les octets sont enregistrés transversalement : 8 bits sur huit pistes. La neuvième piste est utilisée pour le bit de parité. Les données sont enregistrées par blocs. A la fin d'un bloc on peut placer 9 bits supplémentaires, chacun représentant la parité de sa piste dans le bloc. Si une erreur se produit pour un seul bit, celle-ci induit deux erreurs de parité : dans la piste et l'octet. Cette méthode permet de repérer et d'identifier le bit en erreur. Une correction est alors possible.

D'accès séquentiel, ce sont des supports de données relativement lents. Par contre ils sont très économiques. Sous forme de cassettes ou cartouches elles sont d'un usage plus aisé. Il est en particulier possible dans les grands systèmes d'utiliser des robots pour installer les cartouches dans les lecteurs.

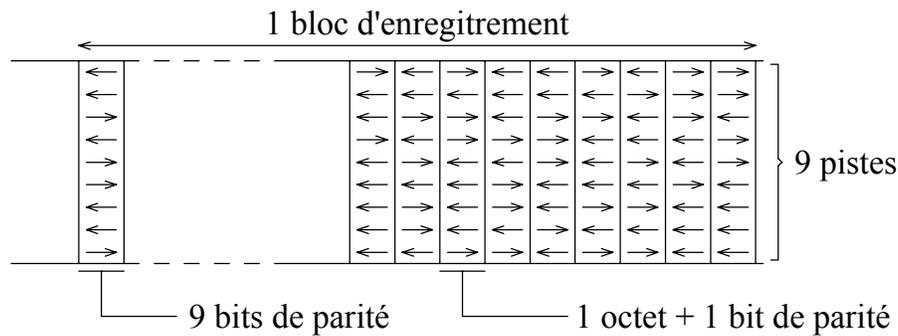


Figure 15

Cependant les enregistrements magnétiques s'altèrent avec le temps. Il faut les réécrire tous les trois ou cinq ans. La durée de vie des bandes est également limitée. Il faut les remplacer tous les dix ou quinze ans. Dans un grand centre de calcul cela représente beaucoup de travail et d'argent.

IX.5 Disques optiques numériques

Les disques compacts sont des objets standardisés tant au niveau de leurs caractéristiques physiques que de leur format logique. Cette standardisation qui assure une compatibilité dans un très vaste domaine est certainement une des raisons du succès de ce support. La normalisation initiale, correspondant au disque compact audio (CD-A), est décrite dans le "Livre Rouge" rédigé en 1982 par Sony et Philips. Nous commençons donc notre étude par ce type de CD.

IX.5.a Format physique

Un disque compact a un diamètre externe de 120 mm pour une épaisseur de 1.2 mm. Il est percé d'un trou central de 15 mm. Il est constitué de trois couches. Sur un support en polycarbonate (plastique transparent résistant ayant un indice de réfraction de 1.55) est déposée une fine couche réfléchissante d'aluminium. Cette couche est protégée des rayures et de l'oxydation par une laque protectrice, sur laquelle peut être imprimée une étiquette.

Les informations sont imprimées dans la couche métallique sous forme de cuvettes (pits), d'une profondeur de 0.12 μm , gravées le long d'une spirale, située dans une zone de diamètre compris entre 46 et 117 mm. Cette zone est divisée en trois régions concentriques : un en-tête (lead-in) de 2 mm de large, suivi des données sur une largeur de 33 mm au maximum puis d'une plage de fin (lead-out). Les informations sont arrangées du centre vers l'extérieur. Le pas entre spires est de 1.6 μm , ce qui correspond à une densité de 16000 tpi. La longueur totale de la spirale (22188 spires) est d'environ 5.6 km, pour une longueur utile (données) de 5.38 km. La

largeur des creux est de $0.6 \mu\text{m}$. La surface vierge qui sépare deux cuvettes consécutives est appelée méplat (land). La spirale apparaît donc comme une alternance de créniaux. La longueur de chaque créneau (cuvette ou méplat) est comprise entre $0.833 \mu\text{m}$ et $3.054 \mu\text{m}$.

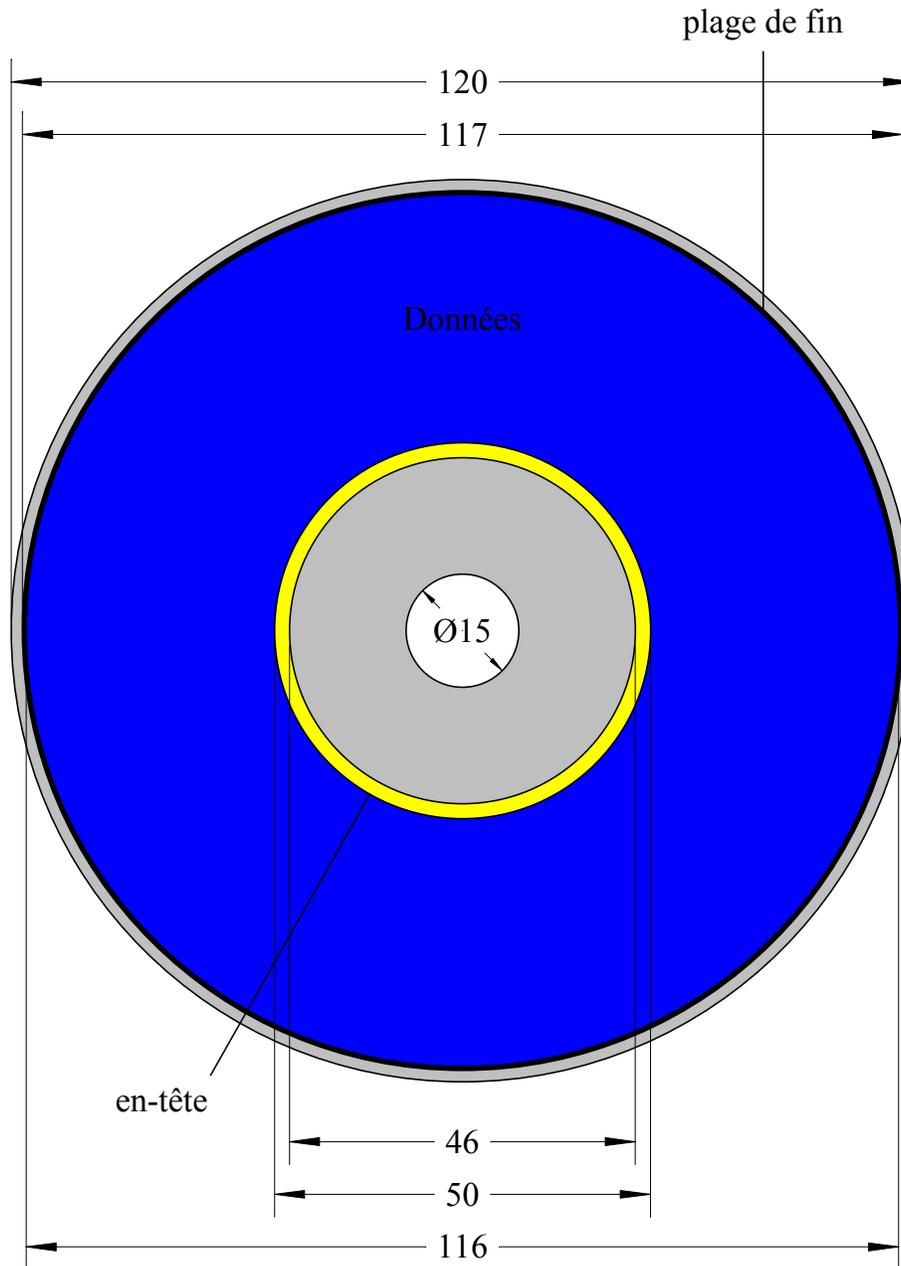


Figure 16

IX.5.b Dispositif de lecture

La lecture des données se fait au moyen d'un laser à semi-conducteur en arséniure de gallium (AsGa) d'une longueur d'onde de 780 nm (infrarouge). La lumière émise par ce laser traverse le substrat en polycarbonate puis se réfléchit sur la couche d'aluminium. Le rayon réfléchi est ensuite guidé par un miroir semi-transparent vers une photodiode. Le faisceau laser est focalisé en une tache d'environ $1.1 \mu\text{m}$ de diamètre. Sur une surface vierge (méplat) la lumière est totalement réfléchie et la photodiode reçoit alors au moins 70 % de la lumière émise. Par contre lorsque l'image du laser chevauche une cuvette, la lumière réfléchie au fond de la cuvette parcourt une distance légèrement plus longue que la lumière réfléchie sur la surface extérieure. La profondeur des cuvettes est telle que la différence de chemin est égale à une demie longueur d'onde. Comme la lumière émise par un laser est cohérente, il y a alors interférence destructive. En pratique il n'y a pas extinction totale, mais la spécification impose une quantité de lumière réfléchie dans ces zones inférieure à 28 % de la lumière émise.

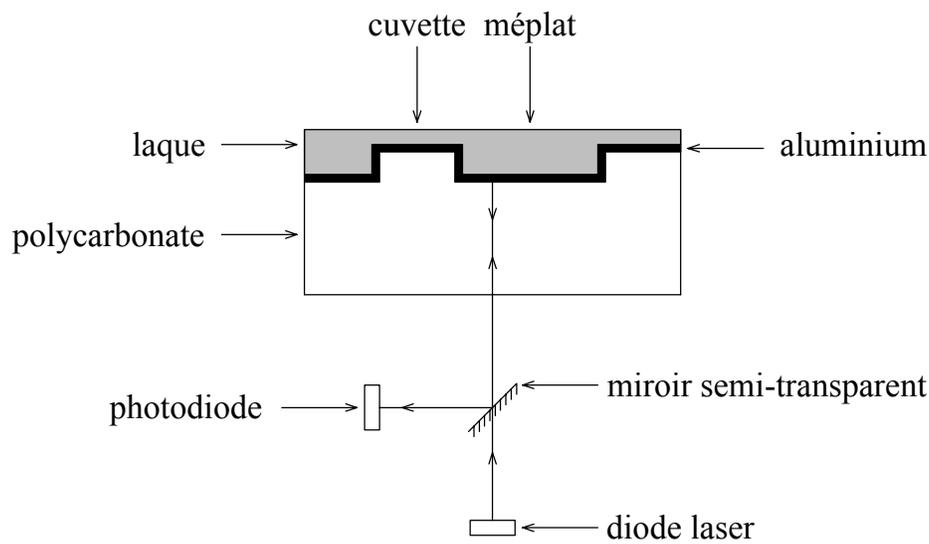


Figure 17

Le principe de la lecture est donc assez simple, mais il faut ajouter des systèmes d'asservissement pour conserver l'alignement de la tête de lecture le long de la spirale et la focalisation au niveau de la couche réfléchissante ($\pm 2 \mu\text{m}$). La tête n'étant pas en contact avec le disque il n'y a peu d'usure.

Contrairement à un disque dur, la vitesse de rotation d'un disque compact est variable de façon à ce que la vitesse linéaire de défilement de la spirale devant la tête de lecture soit constante (CLV : Constant Linear Velocity). La vitesse de lecture pour un CD audio est de 1.2 m s^{-1} . La vitesse angulaire varie donc de 200 à 460 tours par minute selon la position de la tête le long de la spirale. La durée d'enregistrement maximum est d'environ 74 minutes. Le temps d'accès est plus long que pour un disque dur, car il faut positionner la tête de lecture d'inertie

mécanique plus grande et il faut accélérer ou ralentir la rotation du disque pour ajuster la vitesse linéaire de défilement.

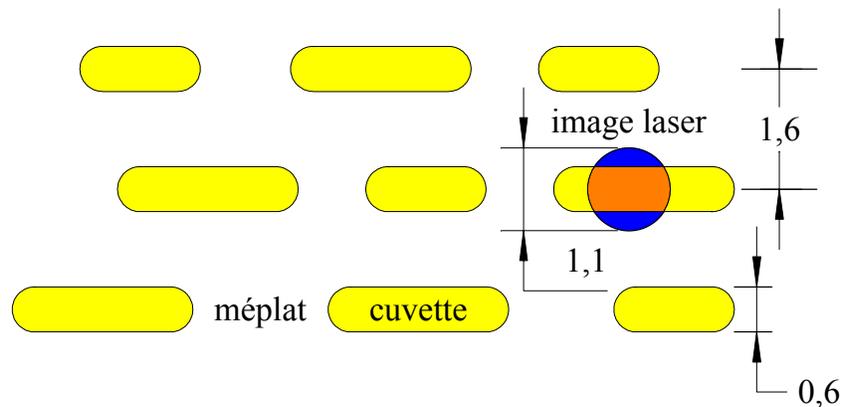


Figure 18

IX.5.c Codage de l'information

Contrairement à ce qu'on pourrait penser les cuvettes et les méplats ne sont pas directement associés aux deux valeurs logiques "0" et "1". En fait la valeur logique "1" est définie par une transition de l'intensité lumineuse détectée. Cela correspond donc au passage d'un méplat à une cuvette et vice-versa. La longueur des cuvettes et des zones vierges permet ensuite de compter le nombre de zéros. Cependant une détection efficace de deux transitions consécutives nécessite un créneau d'une longueur minimale. Il est donc impossible de coder deux "1" consécutifs. En pratique il faut au moins deux "0" entre deux "1". L'horloge est synchronisée à chaque transition. Il faut éviter d'avoir des séries de "0" trop longues. Il y aurait alors un risque de perdre la base de temps, et donc de faire une erreur dans le comptage des "0". La limitation adoptée est d'au plus dix "0" entre deux "1". On associe donc à chaque créneau (pit ou land) une séquence binaire débutant par un "1" et comptant entre 3 et 11 bits. La longueur d'un créneau est d'environ $0.278 \mu\text{m}$ par bit.

Pour tenir compte de ces contraintes on est amené à utiliser 14 bits pour coder un octet. On dénombre en effet 267 combinaisons de 14 bits pour lesquelles on trouve au moins deux et au plus dix "0" entre deux "1". Ce codage est appelé E.F.M. pour Eight to Fourteen Modulation. La table de correspondance a été choisie de manière à minimiser le nombre de portes logiques nécessaires à la réalisation du transcodeur qui doit être intégré dans tout lecteur de disque compact. Lors de la concaténation de deux codes de 14 bits, il est encore possible de trouver deux "1" consécutifs ou une série de plus de dix "0". Pour éviter cette difficulté, on intercale 3 bits de fusion tous les 14 bits. Il faut donc au total 17 bits pour représenter un octet. Pour améliorer le rapport signal/bruit on cherche à faire en sorte que les longueurs cumulées des creux et des bosses soient en permanence presque égales. C'est un des rôles des bits de fusion.

Le système de codage E.F.M. est destiné à rendre le système de lecture le plus tolérant possible à divers défauts techniques : trou excentré, disque voilé, laser mal focalisé, profondeur des creux, épaisseur du disque, etc. Il subsiste cependant d'autres sources d'erreurs. Ce peut être de microscopiques bulles d'air ou poussières piégées dans le polycarbonate pendant la fabrication du disque, ou des rayures, salissures et autres empreintes de doigts sur le disque. La spécification exige qu'il y ait moins de 250 erreurs de lecture par seconde. Ce taux d'erreur est trop important pour assurer une bonne qualité de son. Pour éliminer ces erreurs on fait appel à un code de détection et de correction d'erreurs : le code Reed Solomon entrelacé croisé (CIRC : Cross Interleaved Reed Solomon Code). Cette technique permet d'atteindre un taux d'erreurs de 10^{-8} sur les données utiles.

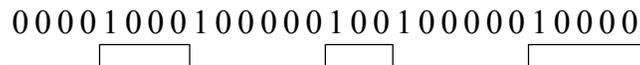


Figure 19

Sur un CD audio les données sont groupées en trames. Une trame comprend 24 octets utiles auxquels il faut ajouter 8 octets pour le code de détection et de correction des erreurs. On ajoute également un octet de contrôle. Chacun de ces 33 octets est codé avec 14 bits. Une trame débute par deux séquences de synchronisation identiques de 11 bits ("1" suivi de dix "0"), suivies par 5 bits de fusion. Cette configuration ne peut se trouver dans les données car elle correspond à une des onze combinaisons de 14 bits non utilisées par le codage E.F.M.. Une trame comprend donc 588 bits.

| | | |
|---------|---|--------------|
| 1 trame | Synchronisation | 27 bits |
| | 1 octet de contrôle | 1 x 17 bits |
| | 24 octets de données | 24 x 17 bits |
| | 8 octets de détection et correction d'erreurs | 8 x 17 bits |
| | Total | 588 bits |

Table 3

L'unité d'enregistrement sur un CD est le secteur qui regroupe 98 trames, soient 57624 bits. Une seconde d'enregistrement correspond à 75 secteurs.

| | | |
|-----------|-------------|--|
| 1 secteur | 98 trames | 57624 bits 2352 octets utiles |
| 1 seconde | 75 secteurs | $4.3218 \cdot 10^6$ bits 176400 octets utiles |

Table 4

Le son sur un CD est échantillonné sur 16 bits et sur deux canaux (stéréo). Une seconde d'enregistrement correspond donc à 44100 échantillons. La fréquence d'échantillonnage est donc

de 44.1 kHz. Le débit global de lecture est de 4.32 Gb.s^{-1} , pour un débit utile de 172 ko.s^{-1} . La durée maximale d'enregistrement de 74 minutes correspond à 747 Mo utiles. L'adressage d'un secteur sur un CD se fait sous la forme : minute:seconde:secteur. Un CD audio peut être divisé en pistes (jusqu'à 99). L'adressage peut être absolu (depuis le début de la spirale) ou relatif (depuis le début de la piste).

IX.5.d Disque compact en informatique

Le disque compact est sorti peu de temps après l'IBM PC-XT qui était alors équipé d'un disque dur de 10 Mo. Il est vite apparu que les disques compacts pouvaient servir de support à des données informatiques. Cela nécessitait cependant une réduction importante du taux d'erreurs et un repérage plus précis des secteurs. C'est pourquoi le format des secteurs a été revu et précisé dans le Livre Jaune rédigé par Philips et Sony en 1984.

Dans un CD-ROM on distingue deux formats pour les secteurs. La structure des secteurs décrite pour les disques audio, à savoir 98 trames de 588 bits, est conservée. Par contre l'utilisation des 2352 octets disponibles est modifiée. Le premier format, nommé CD-ROM Mode 1, est prévu pour les données informatiques. Celles-ci représentent 2048 octets, quantité adaptée à l'adressage binaire. Ce mode contient un niveau supplémentaire de détection et de correction d'erreurs en couche qui nécessite 4 octets pour la détection (EDC : Error Detection Code) et 276 octets pour la correction (ECC : Error Correction Code). Cette technique permet d'obtenir un taux d'erreur de 10^{-12} . Huit octets sont laissés libres entre les octets de détection et les octets de correction. Pour faciliter le repérage et l'adressage des secteurs ceux-ci débutent par 12 octets de synchronisation suivis de 4 octets d'en-tête (trois octets pour l'adressage minute:seconde:secteur et un octet pour le mode).

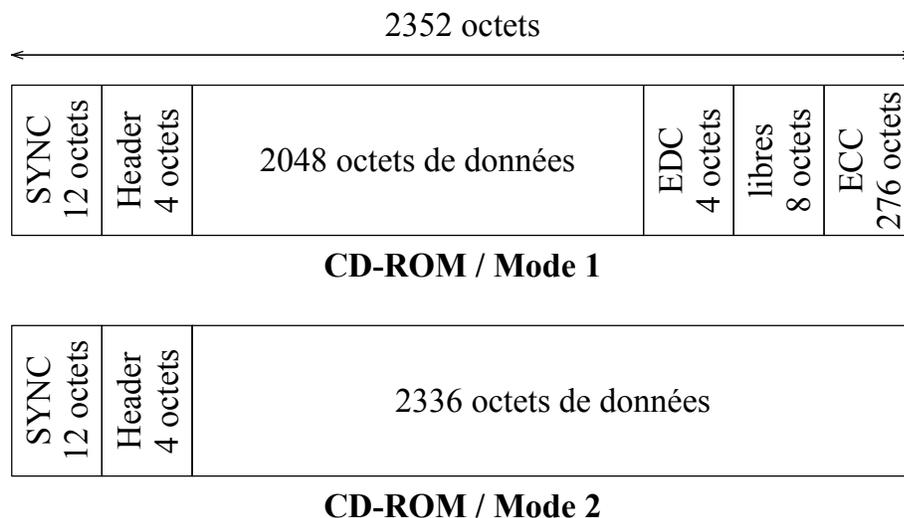


Figure 20

Ce format correspond à un débit pour les données utiles de 150 ko.s^{-1} . et une capacité utile de 650 Mo. Il existe un second format (CD-ROM Mode 2) qui n'utilise pas de détection d'erreurs. Il permet d'utiliser 2336 octets par secteur, pour un débit de 172 ko.s^{-1} . Ce format convient aux données compressées audio ou vidéo.

Si des données sonores de haute qualité sont nécessaires les formats CD-A et CD-ROM peuvent être combinées. Sur un CD en mode mixte, la première piste contient les données informatiques. Elle est suivie par une ou plusieurs pistes audio.

Les livres rouge et jaune n'autorisent pas le mélange de secteurs de types différents sur une même piste. Cela s'avère une limitation majeure pour des applications multimédia qui manipulent simultanément du texte, du son et de la vidéo. La tête de lecture est alors obligée de se déplacer en permanence pour alimenter les divers flux. Le format CD-ROM XA a été introduit en 1989 par Philips, Sony et Microsoft pour palier cette limitation. Il permet l'entrelacement de secteurs de types différents sur une même piste. A côté de cet apport, le format des secteurs est légèrement modifié. Il utilise les huit octets laissés libres par le format CD-ROM. Les deux formats CD-ROM XA sont dénommés Forme 1 et Forme 2. Ils sont résumés dans la figure suivante.

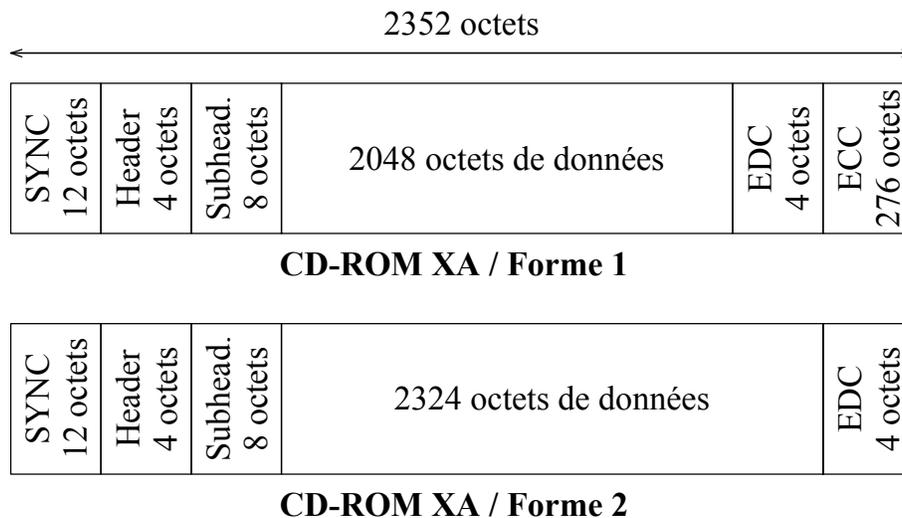


Figure 21

Si la vitesse de lecture des disques audio est fixée par la fréquence d'échantillonnage, il n'y a aucune contrainte pour les données informatiques. Cela permet d'accélérer le débit de transfert, qui est généralement exprimé comme un multiple de 150 ko.s^{-1} (nX). Il est également possible de fonctionner à vitesse de rotation constante (CAV : Constant Angular Velocity).

IX.5.e Fabrication d'un CD

La fabrication d'un CD-ROM commence par la mise en forme des données : découpage en secteurs, addition des octets de synchronisation, contrôle, détection et correction d'erreurs, codage E.F.M., etc. Ensuite on prépare une première matrice en verre. Une très fine couche de résine photosensible est déposée sur plaque de verre parfaitement polie. Cette couche est impressionnée par un laser qui crée une image des cuvettes. Après développement, la résine non polymérisée est retirée. Le négatif des cuvettes apparaît alors sur la plaque de verre sous forme de bosses. Une très fine couche d'argent est alors vaporisée sur la face photosensible. Du nickel est déposé par galvanisation au-dessus de cette couche argentée. Après séparation de la plaque de verre on obtient une première copie du disque appelé père. Ce père peut alors être employé pour estamper en série les disques en polycarbonate. Cependant pour éviter toute détérioration de l'original on effectue une empreinte dite mère, qui sert à la fabrication de plusieurs fils. Père ou fils servent au moulage du polycarbonate par injection ou pression. Les disques de polycarbonate sont ensuite recouverts d'une fine couche d'aluminium, puis d'une laque.

Ce procédé permet une production rapide et à faible coût unitaire pour des grandes séries. La fabrication du père et des fils est en effet longue, délicate et coûteuse. Ce n'est donc rentable que si ces frais sont répartis sur un grand nombre de disques produits. Cette solution n'est pas adaptée aux petites séries ou au développement de prototypes. On fait alors appel aux disques compacts enregistrables.

IX.5.f Disque compact enregistrable

Il existe deux types de disques optiques enregistrables. Commençons par le disque optique enregistrable une seule fois : CD-R (Recordable). Le contenu de ce type de disque peut être enregistré une fois, il fonctionne ensuite comme CD-ROM classique en lecture uniquement. On parle en anglais de disque WORM : Write Once/Read Many.

Un disque compact enregistrable se compose de quatre couches. Entre le substrat en polycarbonate et la couche réfléchissante, qui est ici en argent ou en or, se trouve un colorant organique photosensible (cyanine, phtalocyanine, AZO, Formazan). La couche réfléchissante est toujours protégée par une laque.

L'écriture des données est réalisée par un laser puissant, qui permet de chauffer localement le colorant. Au-dessus d'une température critique, il se produit une réaction chimique qui rend le colorant opaque. Le codage des données est identique à celui d'un CD-ROM. A la lecture les zones "brûlées" absorbent la lumière, alors qu'ailleurs la lumière peut traverser le colorant pour être réfléchi. Le signal lumineux renvoyé vers la photodiode est donc comparable à celui d'un disque compact. Un CD-R peut être utilisé dans un lecteur de CD classique. La puissance du laser à l'écriture dépend de nombreux paramètres : type de colorant, température, hygrométrie, vitesse de défilement, etc. Au centre du disque, entre le trou et la plage de début, est

définie une zone (PCA : Power Calibration Area) qui permet le calibrage du faisceau laser. Cette zone contient un compteur du nombre de calibrages réalisés. Ce compteur est limité à 99.

Sans repère il serait très difficile, dans la phase d'écriture, de suivre une spirale au pas de 1.6 μm . C'est pourquoi un marquage de la spirale est réalisé à la fabrication du disque au moyen de cuvettes gravées dans la couche réfléchissante. Ces cuvettes sont réparties de manière à permettre également un asservissement précis de la vitesse de rotation du disque et de la puissance du laser, la fréquence de défilement devant être de 22.05 kHz.

La stabilité des colorants, qui doivent par exemple résister aux ultraviolets de la lumière solaire, garantit une durée de vie d'au moins 30 ans des disques enregistrables. Par rapport aux bandes magnétiques, ils autorisent un accès aléatoire aux données. Ces deux arguments font des disques compacts enregistrables un bon support pour l'archivage.

Le marquage du colorant est irréversible, il n'est donc pas possible de modifier le contenu du disque. Ne pouvoir écrire les données qu'une fois sur un disque peut être une limitation pour certaines applications. C'est ce qui a conduit au disque réinscriptible : CD-RW (ReWritable).

IX.5.g Disque compact réinscriptible

La structure d'un disque compact réinscriptible comporte six couches. Le colorant est remplacé par un alliage d'argent, indium, antimoine et tellure (Ag-In-Sb-Te), pris en sandwich entre deux couches de diélectrique. Dans un disque vierge, l'alliage est dans une forme cristalline. Il réfléchit alors en partie la lumière. L'enregistrement se fait au moyen d'un laser puissant qui chauffe localement l'alliage pour le porter au-delà de sa température de fusion (entre 500 et 700°C). Si le refroidissement est rapide les atomes n'ont pas le temps de s'organiser et l'alliage revient à l'état solide dans un état amorphe. Il absorbe et diffuse alors la lumière. La lecture d'un CD-RW se fait ensuite avec un laser de faible puissance comme pour un CD classique. Cependant la faible quantité de lumière réfléchie (15-25 % pour les lands) nécessite une photodiode plus sensible.

Si on chauffe suffisamment longtemps une zone amorphe à une température un peu inférieure à la température de fusion, les atomes ont la possibilité de se réarranger ce qui permet de retrouver une structure cristalline après refroidissement. Il est donc possible d'effacer un enregistrement. Cette technique d'enregistrement est dit à changement de phase.

Ces disques, comme les disques enregistrables une fois, sont fabriqués avec une spirale de guidage pour la phase d'écriture. La couche d'enregistrement fabriquée par pression est initialement amorphe. Le disque subit ensuite un traitement thermique pour la faire passer en phase cristalline. Ainsi chaque disque CD-RW est-il vendu vierge.

IX.5.h Disque optique numérique haute densité

Alors que le disque compact a été initié comme support audio, le disque haute densité avait pour objectif initial la vidéo. Les études lancées par les industriels avaient conduit à deux formats concurrents : le MMCD (Multi Media Compact Disc) et le SDD (Super Density Disc). Lorsqu'elle s'est intéressée à ce type de support, l'industrie informatique a imposé le choix d'un seul format. C'est ainsi qu'en 1995 le Digital Versatile Disc (DVD) a remplacé le Digital Video Disc.

Les dimensions d'un DVD sont identiques à celles d'un CD : 120 mm de diamètre externe, 1.2 mm d'épaisseur et percé d'un trou de 15 mm de diamètre. La capacité a été accrue en augmentant la densité et le nombre de couches. Le principe de lecture optique reste le même avec des améliorations. L'utilisation d'un laser avec une plus petite longueur d'onde (635 ou 650 nm, rouge) permet de réduire la taille des motifs à détecter. En outre, les améliorations de l'optique autorisent une plus grande ouverture numérique (0.6 contre 0.45 pour le CD) ce qui permet de réduire la taille l'image du faisceau laser dans le plan de focalisation. Les créneaux peuvent donc être réduits en largeur et longueur et les spires peuvent être plus serrées. L'expérience acquise avec les disques compacts, en particulier au niveau de la focalisation et de l'asservissement du suivi de piste, a également permis de réduire les marges de sécurité définies au début de leur développement. La lecture d'un DVD se fait avec une vitesse linéaire constante 3.49 m.s^{-1} . Ces améliorations techniques permettent une densité d'information brute environ quatre fois et demie plus grande, avec un pas des spires de $0.74 \mu\text{m}$ et une longueur minimale de créneau de $0.4 \mu\text{m}$:

$$G_1 = \frac{1.6}{0.74} \times \frac{0.833}{0.400} = 4.5$$

D'autre part, le DVD utilise la capacité d'un faisceau lumineux de traverser plusieurs surfaces semi-transparentes et d'être peu influencé par des perturbations hors du point de focalisation. Ces propriétés, associées à l'utilisation de photodiodes plus sensibles, permettent l'utilisation de deux couches de réflexion superposées. La première couche est semi-réfléchissante, la seconde l'est totalement. Le faisceau laser est lui focalisé sur l'une ou l'autre couche. Cela nécessite une épaisseur maximum du substrat traversé par le faisceau inférieure à 0.6 mm . L'épaisseur de 1.2 mm est obtenue en collant dos à dos deux disques de 0.6 mm . Les deux faces d'un DVD peuvent porter de l'information, éventuellement en deux couches. Les DVD existent en quatre versions : mono-face / mono-couche, mono-face / bi-couche, bi-face / mono-couche ou bi-face / bi-couche.

La densité de la seconde couche, plus profonde, est légèrement inférieure. Le pas des spires est de $0.80 \mu\text{m}$ et la longueur minimale des créneaux est de $0.44 \mu\text{m}$. La lecture d'un DVD se fait avec une vitesse linéaire constante 3.84 m.s^{-1} , de manière à conserver le même débit de lecture. Dans un DVD vidéo la lecture de la première couche s'effectue de l'intérieur vers l'extérieur du disque, et en sens inverse pour la seconde couche. Cela évite une interruption du flux de données qui serait visible au cours de la lecture d'un film. Ce n'est pas nécessairement le cas pour les DVD-ROM.

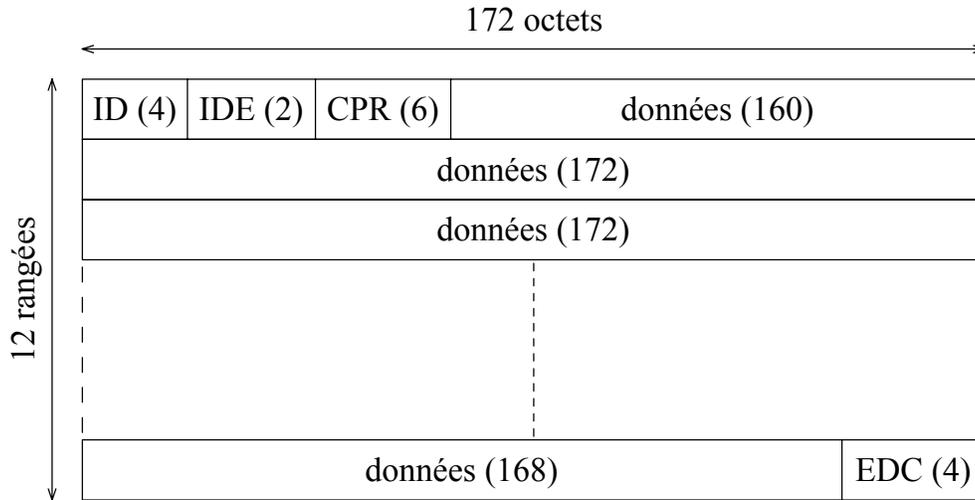


Figure 22

Un autre gain a été obtenu au niveau du codage de l'information. Le codage d'un octet sur un CD nécessite 17 bits (14 bits E.F.M. + 3 bits de fusion). Sur un DVD on n'utilise que 16 bits par octet. Par ailleurs les codes de détection et de correction d'erreurs se sont développés depuis la conception du CD-A. De même la puissance des contrôleurs embarqués chargés de la mise en œuvre de ces codes s'est nettement accrue. Il a donc été possible d'augmenter la fiabilité des données d'un ordre de grandeur en réduisant le nombre d'octets nécessaires à la synchronisation, au repérage, ainsi qu' à la détection et à la correction des erreurs.

Dans une trame de DVD, illustrée par la figure 22, les 2048 octets utiles sont accompagnés de 10 octets de gestion et 6 octets de détection d'erreurs. Un secteur DVD regroupe 16 trames avec 2080 et 2752 octets de détection et correction d'erreurs. Un secteur, qui correspond donc à 32 ko utiles, nécessite donc 605696 bits sur le DVD. Nous avons vu que sur un CD un secteur correspond à 2048 octets utiles et occupe 57624 bits sur le disque. Cette nouvelle organisation logique permet donc un gain de :

$$G_2 = \frac{57624/2}{605696/32} = 1.5$$

Ces gains permet d'atteindre les capacités suivantes pour les diverses déclinaisons du DVD :

| DVD-ROM | Capacité |
|--------------------------|----------|
| Mono-face et mono-couche | 4.7 Go |
| Mono-face et bi-couche | 8.5 Go |
| Bi-face et mono-couche | 9.4 Go |
| Bi-face et bi-couche | 17 Go |

Table 5

Certains des développements en cours pour augmenter ces capacités portent sur une nouvelle réduction de la longueur d'onde du laser, avec par exemple un laser bleu (405 nm). On parle de disques HD-ROM à 600 Go.