

L' audio numérique

Rappels sur l'audio analogique	3
Représentation spectrale	4
Le signal audio numérique.....	5
1) Les avantages du numérique	5
2) Principe de la conversion analogique/numérique	6
A) Le filtrage.....	6
B) L'échantillonnage.....	6
C) La quantification.....	6
D) Le codage.....	6
E) Transmission / enregistrement.....	6
3) Mise en pratique	8
A) Filtrage	8
B) Échantillonnage du signal audio	8
a) Cas du son pour la vidéo.....	10
b) Cas du téléphone	10
C) La quantification.....	10
a) Principe de base	10
b) Bruit de quantification.....	11
D) Codage des échantillons quantifiés.....	11
E) La transmission et l'enregistrement.....	12
a) Corrections d'erreurs	12
b) Les principales sources d'erreurs	12
c) Bits de parité.....	13
d) Entrelacement des données	14
e) Le codage de canal	14
f) Codage NRZ.....	15
g) Codage NRZ-I.....	15
h) Biphasé mark (Manchester).....	16
i) Miller.....	16
j) Miller ²	16

Rappels sur l'audio analogique

On appelle son tout message naturel ou provoqué perçu par l'intermédiaire du sens de l'ouïe.

Le son s'analyse comme une variation de pression au voisinage de l'oreille, cette onde de pression se propageant de sa source jusqu'à l'oreille avec une célérité de $c=340\text{m/s}$ environ.

Un son est caractérisé par :

- son niveau ou intensité
- sa hauteur liée à la fréquence de son fondamental
- son timbre lié à sa composition spectrale

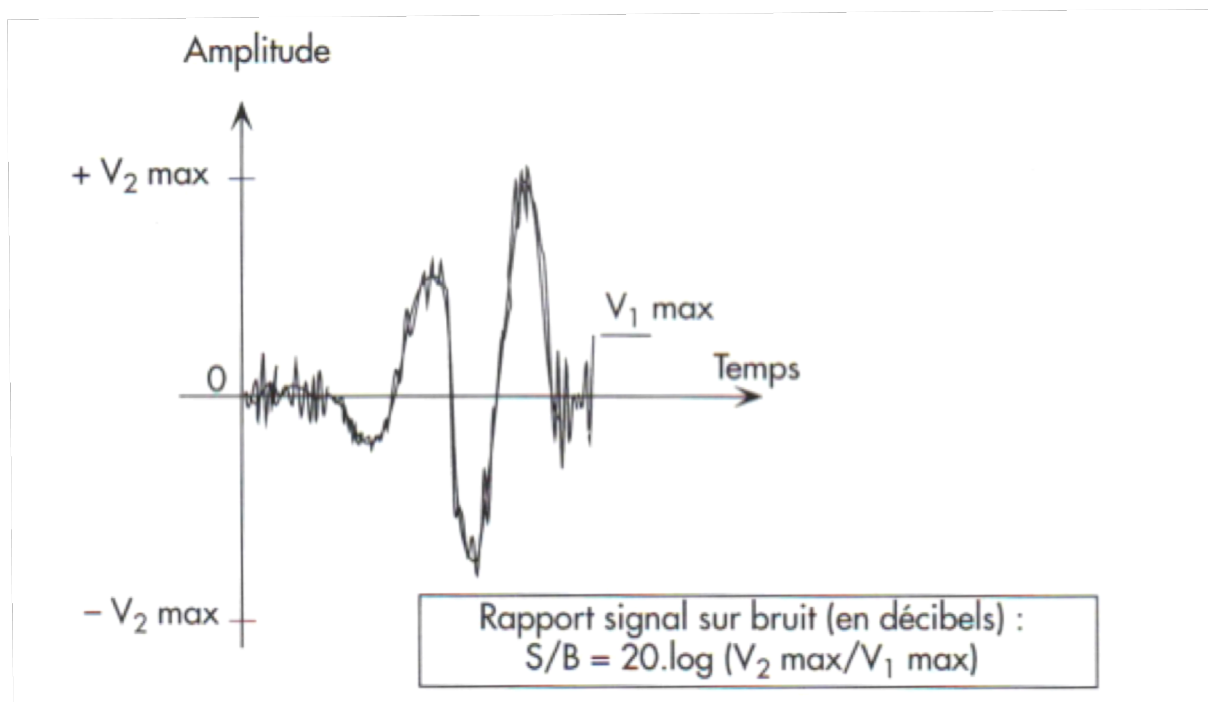
Le signal audio analogique est la traduction d'une grandeur physique en signal analogique (variation d'intensités sonores captées par le micro).

Le signal le plus simple du point de vue contenu fréquentiel est un signal sinusoïdal comme : $x(t) = E\sin(\omega t)$ car il ne contient qu'une seule fréquence $f = \omega/2\pi$

Un signal parlé ou musical est plus complexe puisque son allure varie au cours du temps. Il contient des fréquences graves, moyennes et aiguës. On dit que son spectre s'étend de 20 Hz à 20 kHz et varie en permanence entre ces deux limites.

Une fois capté, le signal va subir de nombreux traitements impliquant des problèmes de bruit. La réponse en fréquence est alors modifiée, le signal distordu. Ces problèmes s'amplifient avec le nombre de copies, en fonction de l'utilisation du matériel, des conditions de stockage...

Le rapport signal sur bruit varie de 40 à 60 dB entre VHS et broadcast.

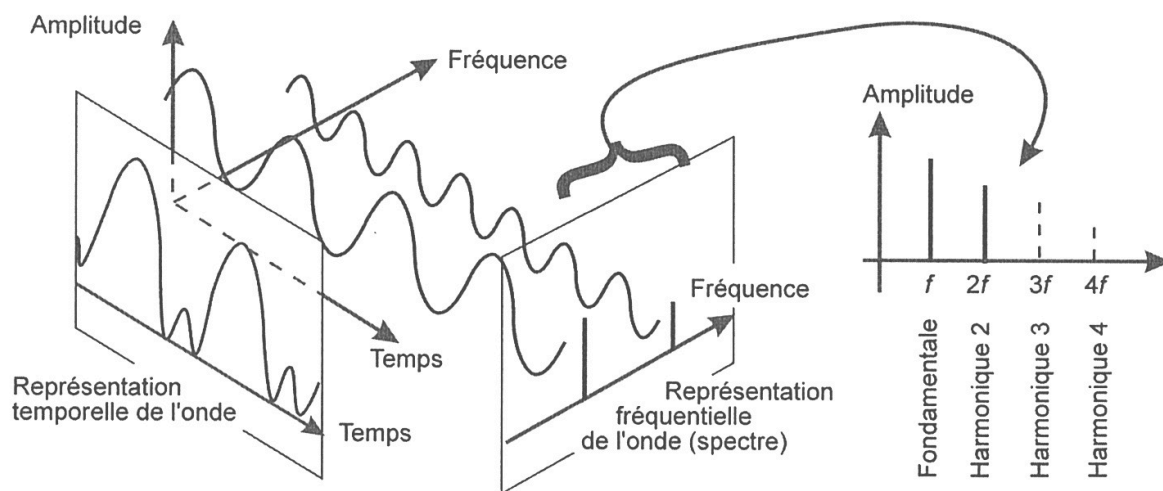


La présence de bruit et de parasite est incontournable dès lors qu'un signal analogique est véhiculée sur une quelconque liaison. De nombreuses techniques permettent de limiter son influence, mais il est presque impossible d'éviter la dégradation du signal original.

Les traitements numériques ne règlent pas complètement ce problème, mais l'influence du bruit sera bien moindre et les techniques de compensation bien plus performantes.

Représentation spectrale

Nous utiliserons régulièrement une approche spectrale du signal conjointement à son analyse temporelle. Le spectre d'un signal permet de représenter son contenu harmonique, c'est à dire une décomposition fréquentielle du signal.

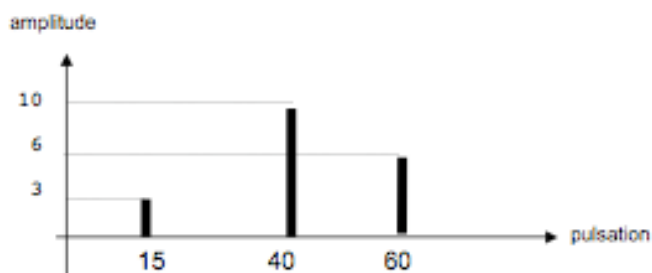


Le spectre d'un signal est la représentation en fonction de la fréquence des amplitudes des différentes composantes présentes dans ce signal.

La transformée de Fourier permet de passer du plan temporel au plan fréquentiel. Cette transformée est réversible. Il est important de comprendre qu'il s'agit d'un seul et même signal représenté de plusieurs manières différentes en fonction des besoins et de la pertinence de la représentation dans un cas précis.

Il est plus commode d'utiliser le signal évoluant dans le domaine temporel pour signifier une modification de sa dynamique (compression, amplification) alors que la représentation spectrale sera plus adaptée pour un traitement lié à la composition spectrale (égalisation, harmonisation)...

Exemple pour : $x(t) = 3\cos(15t) + 10\sin(40t) + 6\cos(60t + \pi/2)$



Le signal audio numérique

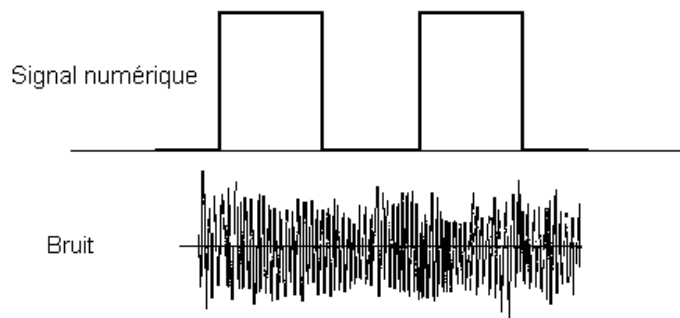
Par définition, il s'agit d'un signal discontinu, basé sur le langage binaire (suite de 0 et de 1), permettant de décrire l'évolution du signal audio d'origine.

- 0 = niveau bas du signal
- 1 = niveau haut du signal

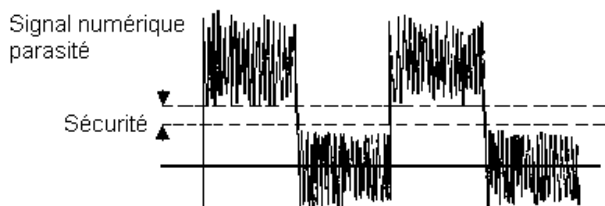
1) Les avantages du numérique

Compte tenu des progrès effectués dans le domaine du traitement numérique, il est possible de réaliser des traitements inimaginables en analogique (par exemple incrustation 3D en temps réel). On peut effectuer de nombreux traitements sans dégradation du signal ce qui est impossible en analogique.

Le bruit et les distorsions du signal n'affectent que les paramètres analogiques du signal numérique et n'altèrent pas l'information si on est capable de reconnaître les 0 et les 1.



Dans un système numérique, le bruit se superpose au signal mais n'atteint pas l'information tant que les 2 niveaux distincts haut et bas sont reconnus par le système.



Dans ces conditions le message reste identique au signal numérique original.

Des systèmes de correction d'erreurs sophistiqués permettent de prévenir les désagréments engendrés par des perturbations lors du stockage ou de transmission de l'information (parasites magnétiques influant sur un câble, rayures sur un support optique...). La copie est parfaite et la dégradation de génération n'existe pas. La copie est un clone de la source, on ne parle plus de générations.

Il ne faut pas perdre de vue que la capture et la restitution seront toujours analogiques. Il y a aura toujours des micros/des enceintes et des objectifs/écrans... Seul le traitement de l'information s'effectue de manière digitale.

2) Principe de la conversion analogique/numérique

A) Le filtrage

Avant de passer à l'étape de l'échantillonnage, il est crucial de s'assurer que le signal est bien limité à une bande passante compatible avec la fréquence d'échantillonnage choisie. C'est une opération qui prépare le signal pour que la conversion s'effectue correctement.

On utilise un filtre passe-bas dont la réponse est linéaire avant de chuter brutalement.

B) L'échantillonnage

On prélève ponctuellement des échantillons du signal analogique à intervalles réguliers et suffisamment rapprochés conformément au théorème de Shannon. Les échantillons prélevés décrivent alors la forme du signal point par point.

Il est nécessaire de maintenir la valeur de l'échantillon le temps de procéder à son échantillonnage. Le dispositif prenant en charges ces opérations est appelé échantillonneur-bloqueur.

C) La quantification

La quantification consiste à donner la valeur la plus proche à l'échantillon par rapport au nombre de valeurs disponibles dans l'échelle de quantification.

Cette étape provoque forcément des erreurs d'approximations. C'est l'erreur de quantification, qui a pour répercussion l'apparition d'un bruit de quantification.

D) Le codage

Une fois que les échantillons sont quantifiés, qu'on leur a attribué une valeur, il reste à coder cette valeur afin de pouvoir l'exploiter facilement lors des étapes de traitement numérique.

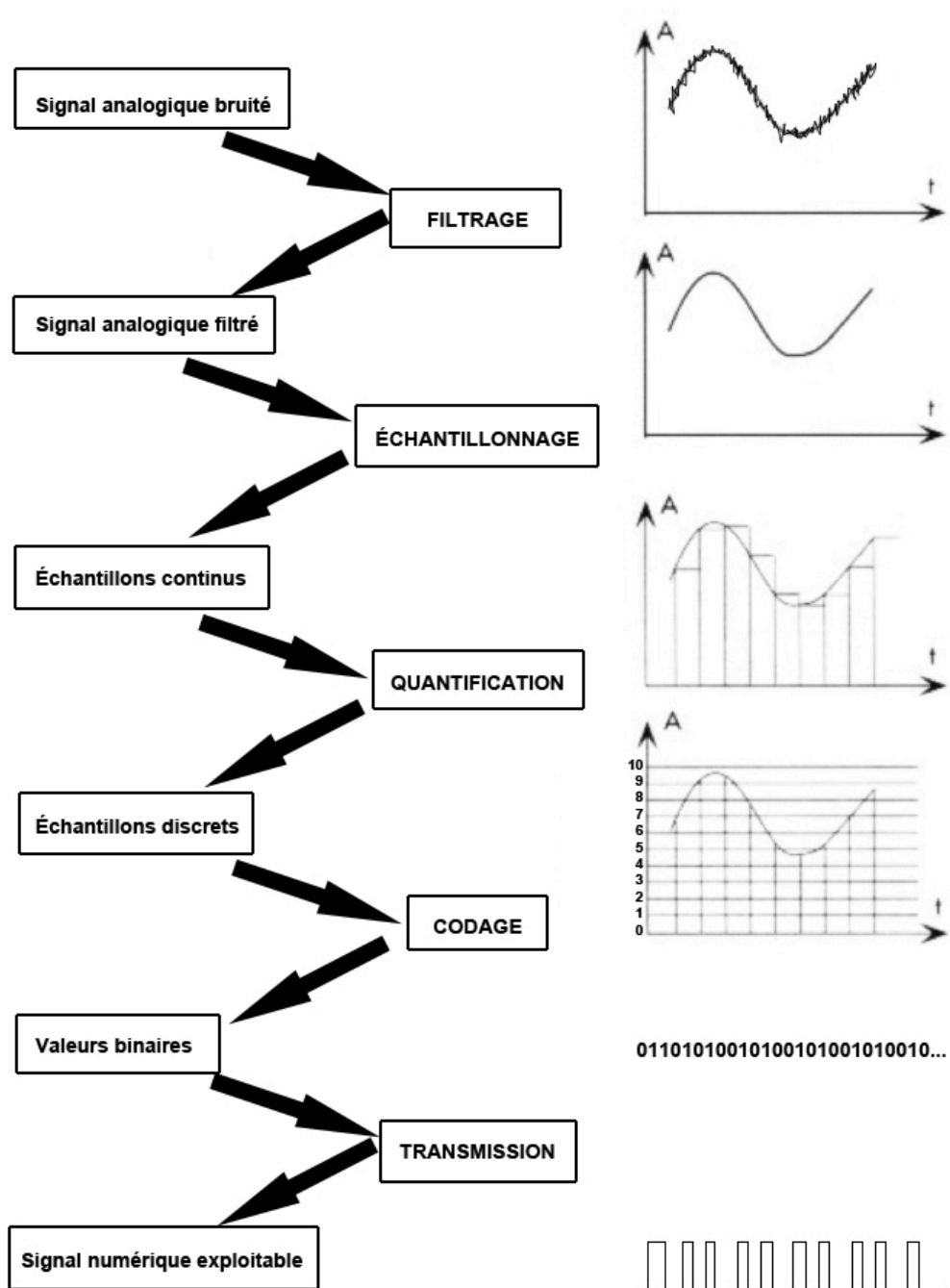
Le plus souvent le codage s'effectue sous forme binaire, mais il existe des variantes dans des cas particuliers.

E) Transmission / enregistrement

Le signal analogique étant numérisé, il reste encore à le mettre en forme pour pouvoir utiliser l'information dans de bonnes conditions.

On ajoute un certain nombre d'informations supplémentaires qui permettront de garder les données intactes lors du décodage et de corriger les éventuelles erreurs de transmission. Ces données sont redondantes et accroissent la quantité d'informations.

Une autre étape consiste à modifier la manière dont les données sont codées pour s'adapter un mode de fonctionnement particulier (électrique, optique...).

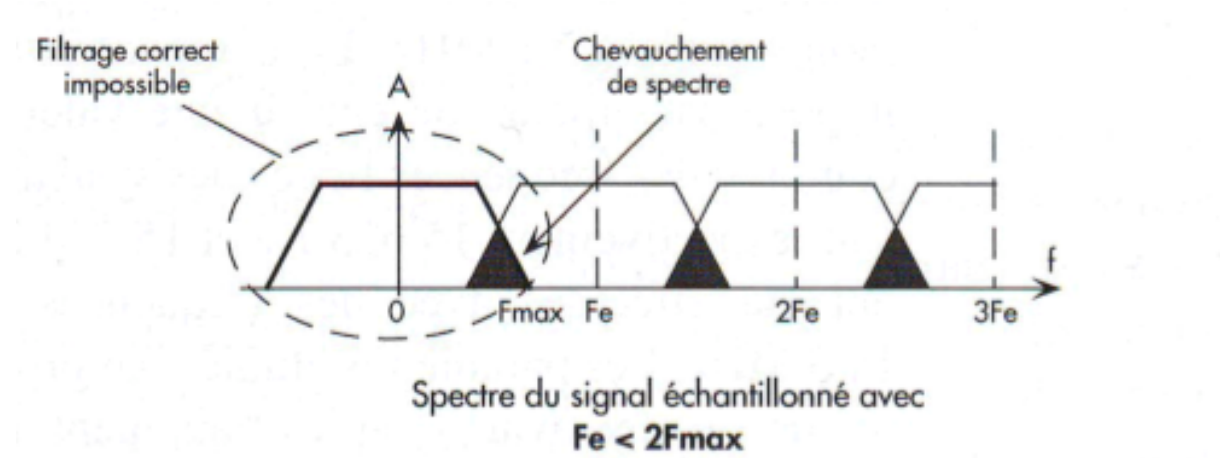


3) Mise en pratique

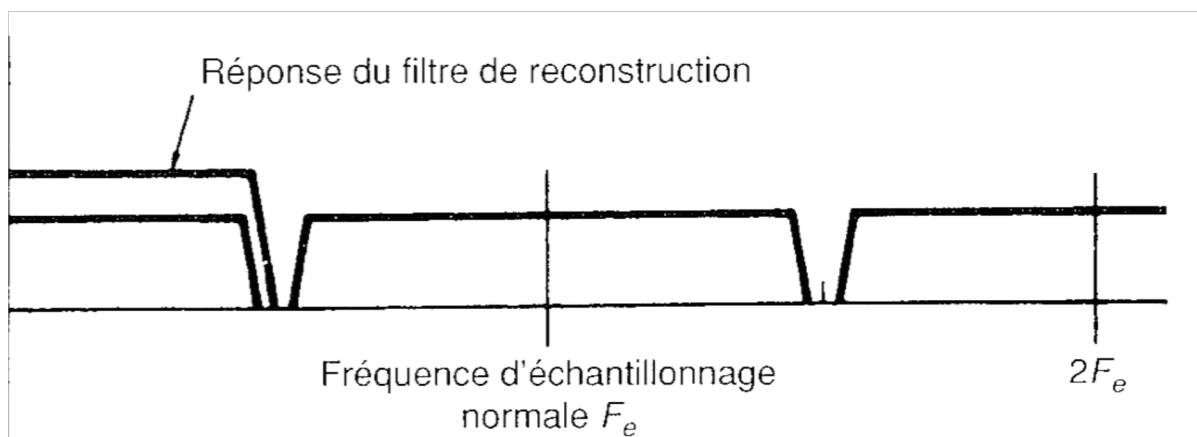
A) Filtrage

La loi de shannon nyquist indique $F_e \geq 2 \times F_{\max}$ du signal à numériser

On admet que pour l'audio $F_{\max} = 20 \text{ KHz}$ (pour la vidéo $F_{\max} = 6\text{MHz}$)
Si cette condition n'est pas respectée, il y a apparition de fréquences aberrantes dans le signal du fait du repliement de spectre.

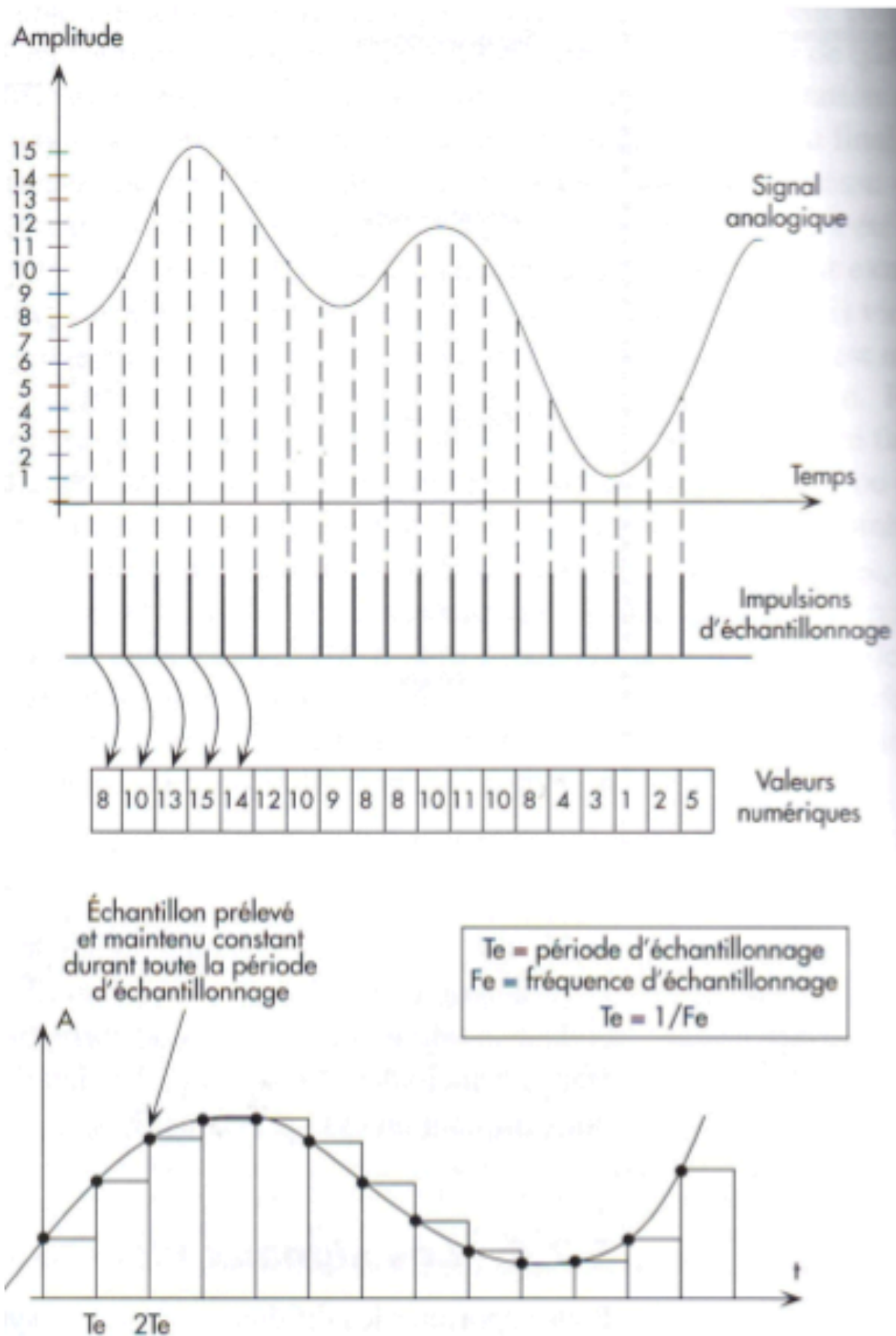


Pour éviter ce problème, on utilise en amont un filtre anti-aliasing (anti-repliement ou de reconstruction) pour supprimer les fréquences indésirables supérieures à la moitié de la fréquence d'échantillonnage (bruits, parasites etc...)



B) Échantillonnage du signal audio

On effectue ensuite un découpage du signal analogique à intervalles réguliers.
La valeur obtenue est alors bloquée le temps de la quantifier.



Le critère de Shanon-Nyquist ($F_e \geq 2 \times F_{max}$) est un point de départ permettant de déterminer la fréquence d'échantillonnage. Mais il est judicieux de tenir compte d'autres paramètres liés à l'utilisation de ce signal en cours de numérisation.

a) Cas du son pour la vidéo

L'utilisation du son pour la vidéo nous amène à considérer les systèmes à 625 lignes/50Hz et 525 lignes/60Hz.

Dans le premier cas 37 lignes sont nécessairement noires et donc 588 lignes sont utilisables soit 294 pour chaque trame.

$$50 \times 294 = 14700 \text{ Hz}$$

De même pour le système à 525 lignes, 35 sont noires, 490 lignes disponibles soit 245 ligne pour chaque trame

$$60 \times 245 = 14700 \text{ Hz}$$

Ces fréquences ne sont pas utiles à proprement parler, mais il est aisé d'isoler une horloge battant à 14700 Hz.

Nous cherchons une fréquence F_e supérieure ou égale à 40 000 Hz en ayant pour repère une horloge dont la fréquence est de 14 700 Hz. Il suffit de multiplier par 3 cette fréquence : $14\,700 \times 3 = 44\,100 \text{ Hz}$. Cette valeur répond à nos conditions de départ et son utilisation sera facile.

b) Cas du téléphone

On cherche avant tout à véhiculer le son d'une voix humaine dont le spectre ne couvre pas la totalité du spectre audible (100 Hz – 5000 Hz environ). La qualité n'est pas un critère essentiel, peu importe si la voix est déformée, du moment qu'elle reste intelligible.

Les compagnies téléphoniques filtrent le spectre pour conserver les fréquences entre 300 et 3400 Hz. La fréquence d'échantillonnage peut être réduite à 7 KHz, ce qui limite la quantité de données à gérer.

C) La quantification

a) Principe de base

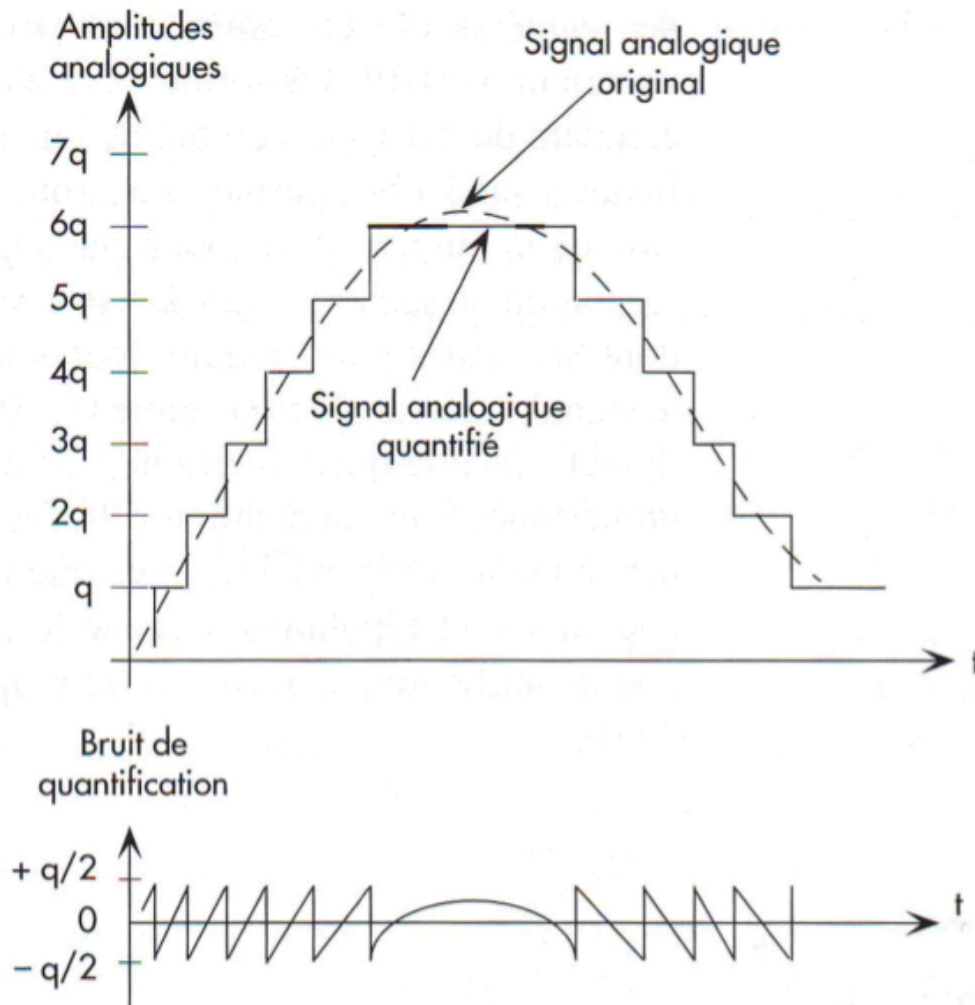
On établit une correspondance entre l'amplitude discrète mesurée lors de l'échantillonnage et la plus proche des valeurs codée sur n bits.

Rappel : en binaire on a 2^n combinaisons possibles pour un codage sur n bits.

Cela implique inévitablement des imprécisions. La précision du signal dépendra donc du nombre de bits alloués à la quantification.

Si le pas de quantification vaut q , alors l'erreur de quantification est au maximum de $q/2$ (approximation de la valeur discrète).

Ces imprécisions se traduisent par l'apparition d'un bruit de quantification.



b) Bruit de quantification

Le rapport signal sur bruit en télévision peut-être calculé de manière approximative par la relation suivante : $S/N = 6 \cdot n + 2$ (n représente la nombre de bits alloués pour le codage de l'information.)

On en déduit que chaque bit supplémentaire améliore le rapport signal/bruit de 6 dB.

- Pour 8 bits = $6 \times 8 + 2 = 50$ dB
- Pour 10 bits = $6 \times 10 + 2 = 62$ dB
- Pour 16 bits = $6 \times 16 + 2 = 98$ dB

On utilise habituellement une quantification de 16 bits soit 65536 niveaux disponibles. La dynamique de l'oreille est de 140 dB, ce qui correspond à une quantification sur 23 bits.

D) Codage des échantillons quantifiés

On utilise généralement le codage binaire pour transformer les valeurs issues de la quantification. Ce codage s'appuie sur une représentation à 2 symboles, **0** ou **1** dans laquelle on associe les symboles en leur attribuant un "poids" en fonction des puissances de 2.

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

Dans un nombre binaire, la valeur d'un bit, appelée poids, dépend de la position du bit en partant de la droite. A la manière des dizaines, des centaines et des milliers pour un nombre décimal, le poids d'un bit croît d'une puissance de deux en allant de la droite vers la gauche comme le montre le tableau suivant :

Nombre binaire	1	1	1	1	1	1	1	1
Poids	$2^7 = 128$	$2^6 = 64$	$2^5 = 32$	$2^4 = 16$	$2^3 = 8$	$2^2 = 4$	$2^1 = 2$	$2^0 = 1$

Bit de poids fort (MSB)

Bit de poids faible (LSB)

Pour convertir un mot binaire en nombre décimal, il suffit de multiplier la valeur de chaque bit par son poids, puis d'additionner chaque résultat. Ainsi, le mot binaire 0101 vaut en décimal :

$$\begin{aligned} 0101_{(2)} &= (2^3 \times 0) + (2^2 \times 1) + (2^1 \times 0) + (2^0 \times 1) \\ &= 8 \times 0 + 4 \times 1 + 2 \times 0 + 1 \times 1 \\ &= 5_{(10)} \end{aligned}$$

Il suffit de faire le cheminement inverse pour coder un entier naturel en binaire :

$$\begin{aligned} 13_{(10)} &= 1 \times 8 + 1 \times 4 + 0 \times 2 + 1 \times 1 \\ &= (1 \times 2^3) + (1 \times 2^2) + (0 \times 2^1) + (1 \times 2^0) \\ &= 1101_{(2)} \end{aligned}$$

E) La transmission et l'enregistrement

a) Corrections d'erreurs

En théorie on ne devrait pas avoir besoin de correcteur d'erreurs, si tout se passait toujours dans conditions optimales. En pratique des erreurs apparaissent quand les supports et les voies de transmission présentent des défauts qui modifient le message.

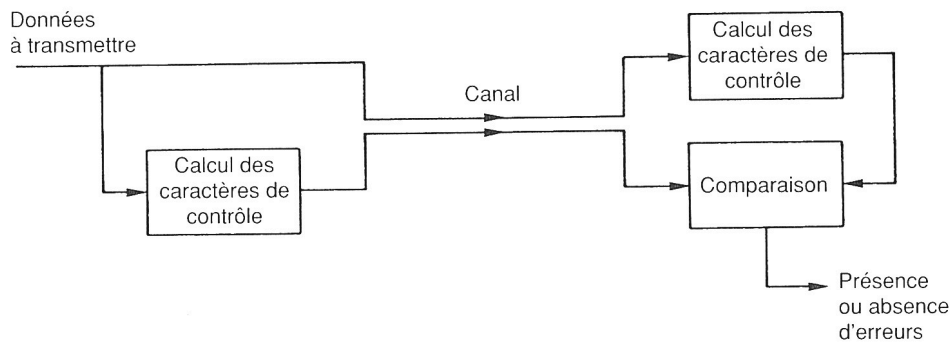
b) Les principales sources d'erreurs

- Problèmes de liés à la lecture (mécanique, vibration...)
- Fluctuation du niveau du signal
- Détériorations du support (magnétique, optique...)
- Bruits et distorsions liés au câblage, à la connectique.

Une erreur en numérique se caractérise par une inversion entre un 0 et un 1 ou une perte de valeurs. Les données peuvent alors conserver une mise en forme acceptable mais leur interprétation est faussée.

c) Bits de parité

Des systèmes redondants permettent de corriger les inversions de valeurs.



Voici un exemple simplifié en illustrant le fonctionnement :

Données de base

0	0	1
1	1	0
0	1	1

On ajoute des codes de corrections d'erreurs avant de transmettre ces données. Ce sont les bits de parités. On additionne les valeurs sur une ligne et sur une colonne, sans tenir compte des retenues.

Données avec le contrôle d'erreur

0	0	1	1
1	1	0	0
0	1	1	0
1	0	0	

Lors de la réception des données, une étape supplémentaire de décodage est nécessaire avant de pouvoir utiliser les informations reçues. Cette étape permettra de s'assurer que les données sont conformes aux données initiales et d'effectuer les corrections nécessaires si des erreurs sont détectées.

Données erronées

0	0	1	1
1	1	0	0
0	0	1	0
1	0	0	

Un 1 a été transformé en 0 lors de la transmission. L'analyse des bits de parités permet de s'apercevoir qu'une erreur s'est glissée au croisement de la troisième ligne et de la deuxième colonne. Il est facile de modifier la valeur du bit en question pour reconstituer les données de départ.

Données corrigées

0	0	1
1	1	0
0	1	1

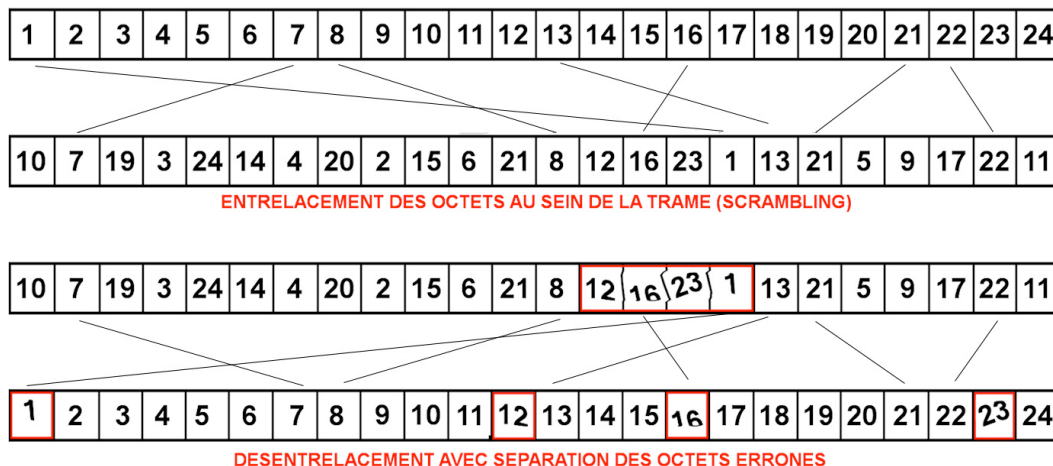
Ce système a néanmoins des inconvénients :

- Il est capable de passer au travers d'erreurs (2 erreurs sur une même ligne ou même colonne se compensent). Il n'est donc pas infallible.
- L'utilisation des bits de parités augmente le poids des données.

d) Entrelacement des données

Les perturbations du signal numérique affectent souvent plusieurs dizaines de bits consécutifs. On procède alors à un brassage ou entrelacement des données. On mélange les données dans un ordre connu par le codeur et le décodeur.

Les erreurs sont dispersées lorsqu'on reforme le signal et ne sont donc plus consécutives. Elles sont plus facilement corrigées.



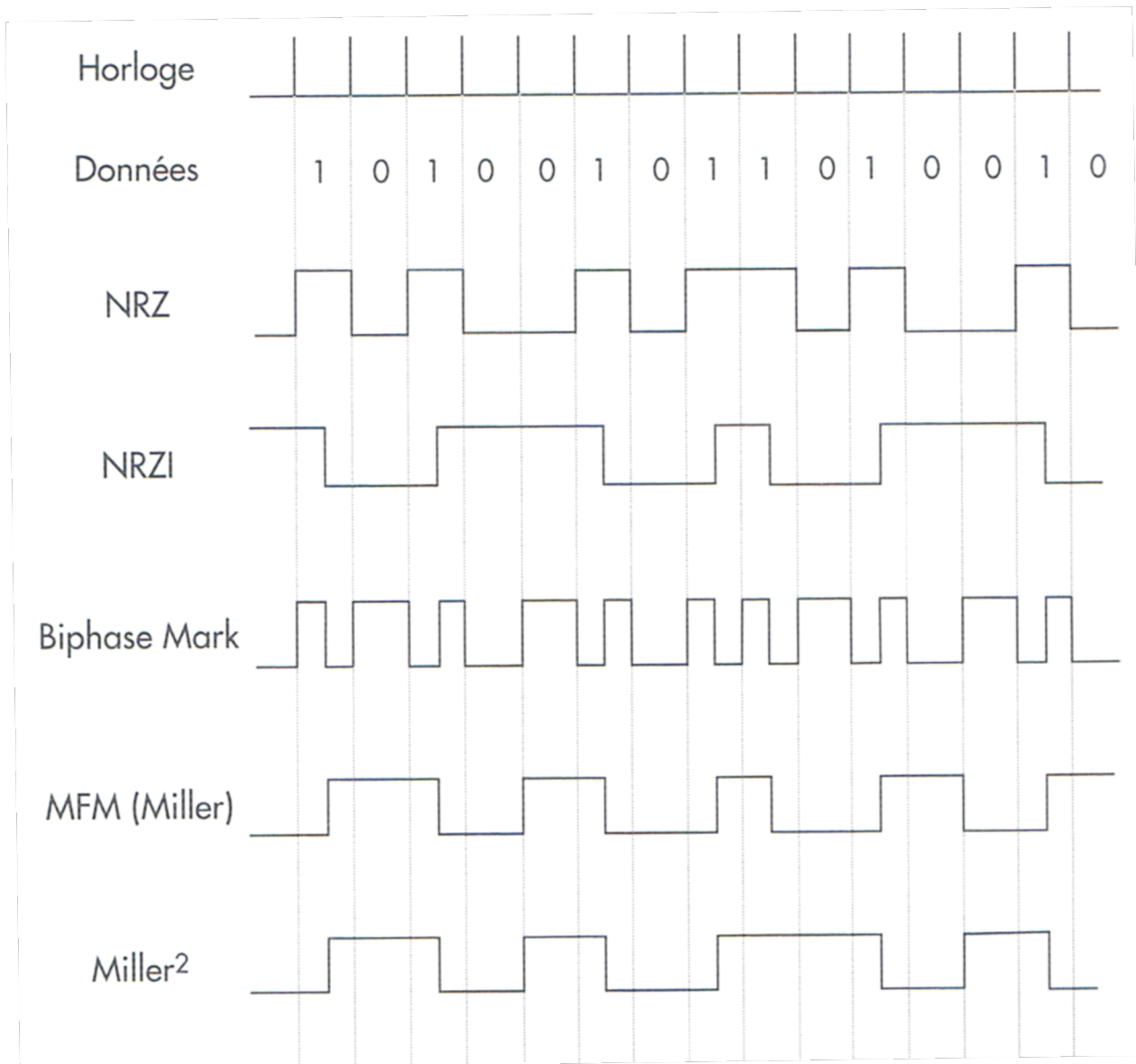
e) Le codage de canal

Il s'agit de mettre en forme le signal d'une manière particulière pour l'adapter au mieux au canal de transmission ou au support sur lequel on souhaite l'enregistrer.

Si on a de longues suites de 1 ou de 0, on ne peut pas facilement relire le signal. On parle alors d'une composante continue. Comment faire alors pour "découper" ces plages et retranscrire correctement l'information ?

Comment faire la différence entre l'absence de signal et un signal de valeur nulle ?

La solution consiste à utiliser un codage particulier qui modifie juste l'agencement du signal mais pas son contenu.



f) Codage NRZ

Le codage NRZ (signifiant No Return to Zero, soit Non Retour à Zéro) est le premier système de codage, car le plus simple. Il consiste tout simplement à transformer les 0 en -X et les 1 en +X, de cette façon on a un codage bipolaire dans lequel le signal n'est jamais nul. Par conséquent, le récepteur peut déterminer la présence ou non d'un signal. Il est difficile de reconstituer le signal d'horloge du fait des plages continues.

g) Codage NRZ-I

Le codage NRZI est sensiblement différent du codage NRZ. Avec ce codage, lorsque le bit est à 1, le signal change d'état après le top de l'horloge. Lorsque le bit est à 0, le signal ne subit aucun changement d'état.

Il possède un défaut: la présence d'une composante continue lors d'une suite de zéro, gênant la synchronisation entre émetteur et récepteur.

On l'utilise dans les liaisons série 4:2:2 en vidéo.

h) Biphase mark (Manchester)

Le codage Manchester est également appelé codage biphase ou PE (pour Phase Encode).

Chaque bit est délimité à gauche par une inversion de phase. Si le bit transmis est '1', une inversion de phase supplémentaire est générée au milieu de la période d'horloge.

- Le 1 produit une transition pendant la demi période d'horloge.
- Le 0 produit une transition et un maintien pendant la durée de la période d'horloge

La composante continue est nulle, le code contient tous les fronts d'horloges (transition à chaque période) et est dit auto synchroniseur.

Ce code est utilisé par les magnétoscopes pour le LTC et dans le cas de liaison SPDIF.

i) Miller

- Un 1 donne une transition pendant la demi période d'horloge.
- Un 0 isolé ne fait rien, mais deux 0 de suite provoque à la période d'horloge.

Ce code comporte une composante continue faible mais il est assez aisé d'extraire le signal d'horloge.

j) Miller²

Ce code reprend les principes du code Miller avec une règle supplémentaire.

Une suite de deux 1 ne provoque pas de transition à la demi période.

Il existe d'autres types de codage : HDM1, 3PM, EIAT, DASH, PRODIGI...