DOCUMENTS RESSOURCES

Périphérique "audio" sur PC

Ce périphérique standard permet :

➢ La production et la reproduction de signaux audiofréquences analogiques à partir de :
- logiciels de synthèse sonore (production)
- fichiers musicaux de différents formats (WAV, MP3, …)

➢ La reproduction de signaux audiofréquences au format numérique PDIF

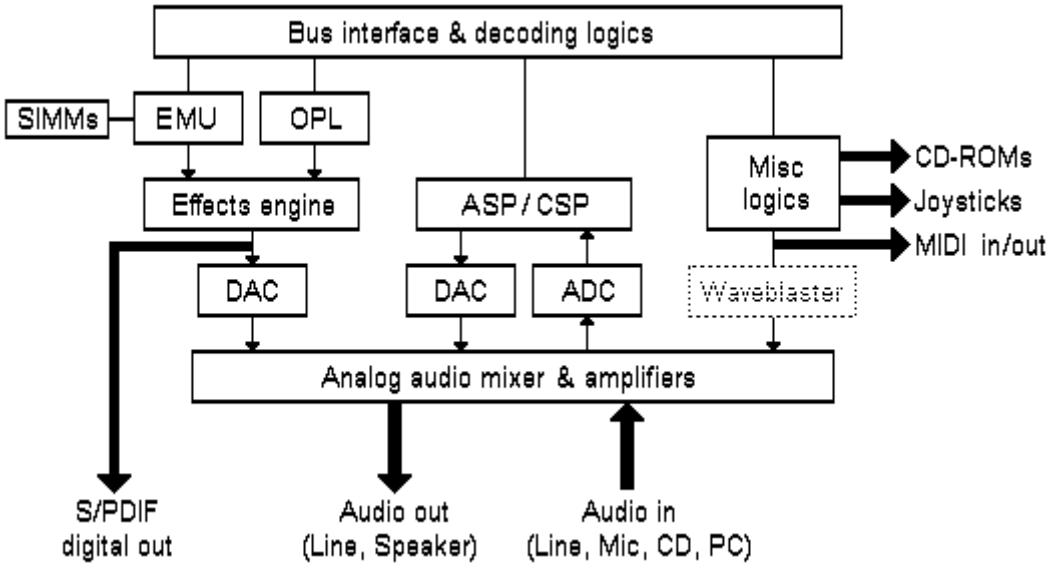
➢ L'acquisition de signaux audiofréquences analogiques de diverses sources: microphone, CD, Line, …

Toutes ces opération sont effectuées en soulageant l'unité centrale le plus possible (on veut écouter ou enregistrer la musique tout en faisant autre chose …).

Ces fonctions nécessitent des opérations de conversions numérique-analogique et analogique-numérique. Les premières cartes Soundblaster avaient une résolution de 8 bits par échantillon (soit 0,4%), tout juste acceptable pour la voix. La "Hi-Fi" (haute fidélité) nécessite une résolution de 16 bits (0,0015%), celle utilisée sur les CD audio.
Aujourd'hui, tous les périphériques audio sur PC ont une résolution de 16 bits, et même 24 bits pour les plus performants.
Pour soulager l'UC, les périphériques audio ont leur propre processeur de signal (DSP) ainsi que de la mémoire locale et un accès direct à la mémoire centrale (DMA). Ainsi, l'UC se "contente" de paramétrer et de donner des ordres au DSP du périphérique qui fonctionne alors de façon autonome.
Toutefois, l'UC traite encore les algorithmes de compression/décompression quand ils nécessitent beaucoup de calculs en temps réel et/ou qu'ils ne sont pas standards (MP3 par exemple).

Les structures associées sont le plus souvent implantées sur la carte mère. Les cartes PCI "audio" proposent aujourd'hui des performances professionnelles pour une utilisation en studio de prise de son.

Pour des raisons historiques, tous les périphériques audio sont compatibles "SoundBlaster" (SB) dont on reproduit l'architecture (c'est un exemple) :



Synthétiseur de musique

- **OPL** : Synthétiseur FM Yamaha
- **EMU** : Reproduction multi-voies des échantillons sonores
- **SIMM** + RAM + ROM embarquée : mémoire d'échantillons sonores

- **Effects Engine** : générateur d'effets (réverbération, chorus, QSound, …)
- **S/PDIF** : sortie numérique
- **DAC** : conversion numérique-analogique

Reproduction et enregistrement de messages sonores

- **ASP / CSP** : processeurs numériques de signaux (DSP) permettant
  - Compression/décompression du débit binaire en temps réel pendant l'enregistrement et la reproduction.
  - Filtrages numériques divers : passe-bas, équaliseur, …
  - Synthèse de parole
  - Reproduction quadriphonique
- **DAC** : conversion numérique-analogique stéréo (2 voies) de résolution 8 ou 16 bits (voire 24 bits). La fréquence des échantillons est réglable sur toutes les valeurs classiques (8kHz et 44,1kHz par exemple) et jusqu'à qq centaines de kHz sur certaines cartes.
- **ADC** : conversion analogique-numérique : pour l'enregistrement de sons stéréo. La résolution est de 8 bits ou 16 bits et la fréquence d'échantillonnage est réglable sur toutes les valeurs classiques.

Autres structures (Misc logics)

- **Interface MIDI**
- **Interface Joystick**
- **Interface CD-ROM**

Mélangeur

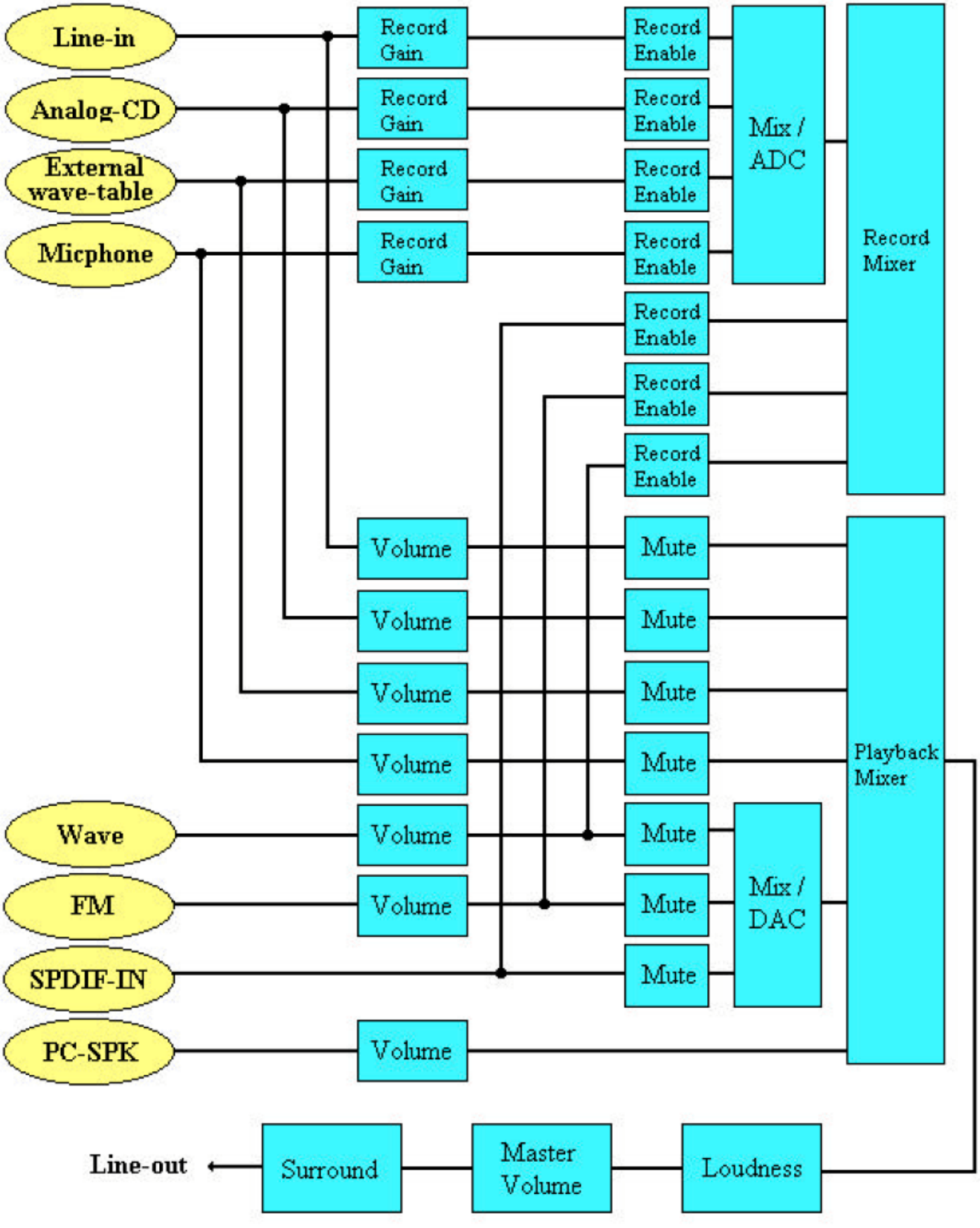
- Mélange analogique et contrôle les niveaux de :
  - Synthétiseurs (EMU + OPL) via un DAC spécifique
  - Sortie "wave" : sortie du DAC de reproduction
  - Entrée "Line" (accès externe)
  - Entrée microphone (accès externe)
  - Entrée CD (connexion interne)
  - Entrée "PC speaker" (connexion interne)
- Contrôle automatique de gain (AGC) de l'entrée microphone
- Sélection des voies d'entrée pour l'enregistrement
- Contrôle du volume, de la tonalité et de l'équilibre G/D du signal de sortie stéréo (Line out, Speaker)

La figure suivante détaille la fonction "mélangeur" associée auc DAC et DAC.

- **Entrées analogiques** : elles sont regroupées dans le haut de la figure ("Line in", "Analog CD", "External wave table" et "Microphone").
  - Chacune est dosée indépendemment (avec des potentiomètres analogiques), puis est sélectionnée avant d'être mélangée de façon **analogique** avec les autres voies.
  - Le résultat analogique est converti en numérique avec l'ADC
  - Le résultat est mélangé de façon **numérique** avec les autres sources numériques (Wave, FM, SPDIF) pour traitements par le PC.
- **Entrées numériques** : elles sont regroupées dans le bas de la figure ("Wave", "FM" et "SPDIF-in")
  - L'entrée "Wave" est le signal numérique produit par les logiciels de lecture des fichiers sonores. Le niveau est dosable.
  - L'entrée "FM" est le signal numérique produit par le synthétiseur. Le niveau est dosable.
  - L'entrée "SPDIF-in" provient d'une source numérique externe.
  - Chacune de ces entrées peut être sélectionnée avant d'être mélangée puis appliquée au DAC de reproduction

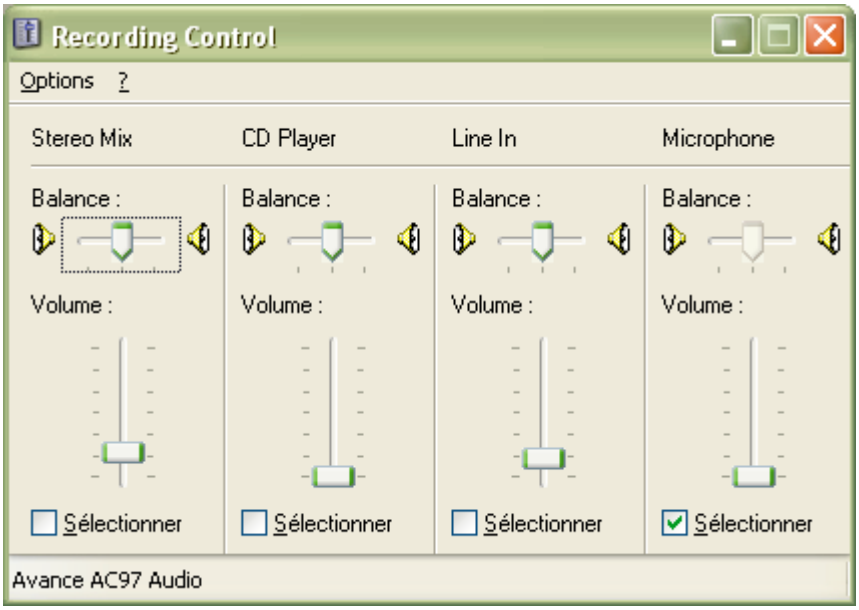
- Le résultat analogique est mélangé avec les autres entrées analogiques (dosage indépendant) et la sortie "Speaker" (PC-SPK) du PC
- Le résultat final stéréo est appliqué sur les sorties analogiques "Line out" et "Speaker out"

Note : les mélangeurs "reproduction" (playback) et "enregistrement" sont pilotés par le "**Contrôle de volume**" de Windows.



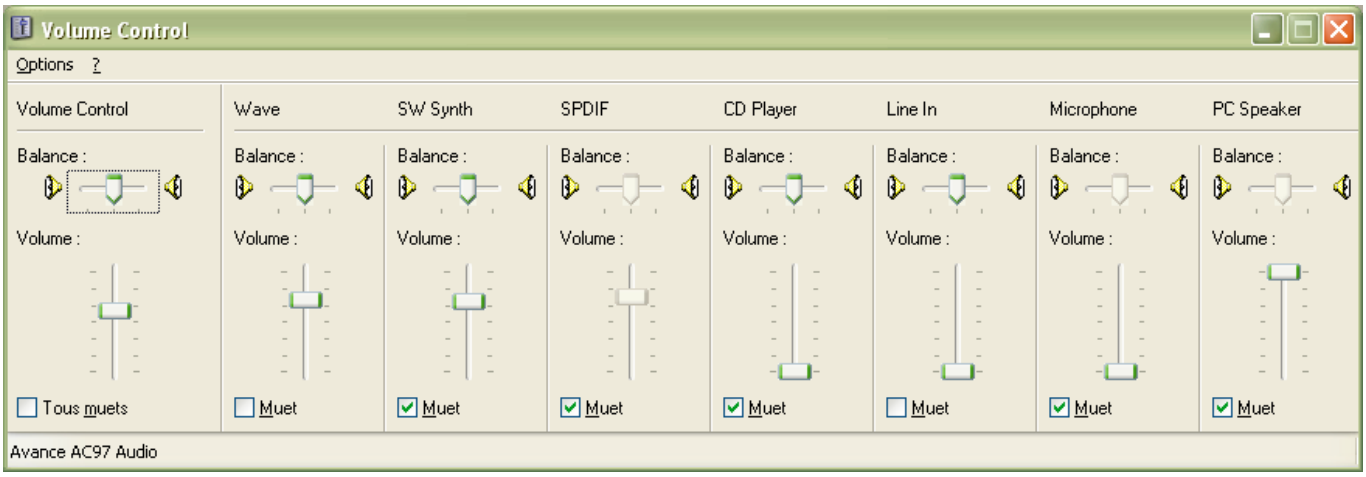
CMI8330 Mixer Block Diagram

"Contrôle du volume" en mode enregistrement



Les sources "CD Player", "Line In" et "Microphone" se retrouvent dans la partie supérieure du schéma bloc précédent. L'entrée "External wave-table" n'existe pas sur l'ordinateur utilisé.
Les curseurs "Volume" et "Balance" correspondent aux blocs "Record Gain" du schéma bloc. Les voies G et D sont dosées simultanément avec "Volume" et l'équilibre peut être ajusté avec "Balance".
Pour la majorité des interfaces "audio", le niveau ne peut être réglé que sur 32 crans, ce qui n'est pas très fin. Cette limitation est liée aux spécifications des circuits intégrés utilisés et devrait s'améliorer dans l'avenir.
Les cases à cocher "Sélectionner" correspondent aux blocs "Record Enable". On ne peut sélectionner qu'une seule source à la fois car le CAN n'a que 2 entrées : les voies G et D. Pour mélanger les diverses sources, il faut choisir "Stereo Mix".
A noter que les sources "Wave", "SW Synth" (FM) et "SPDIF" sont dosées et/ou sélectionnées dans la fenêtre "reproduction" du "Contrôle de volume".

"Contrôle du volume" en mode reproduction (ou lecture)



Il s'agit d'une véritable table de mixage identifiée par le bloc "Playback Mixer" dans le schéma bloc précédent. On peut donc sélectionner autant de sources que nécessaire.
Les sources externes analogiques "CD Player", "Line In" et "Microphone" sont dosées et sélectionnées indépendemment des réglages "enregistrement" (voir schéma bloc). Ce n'est pas le cas des sources numériques internes ("Wave", "SW Synth" et "SPDIF").
Le curseur "Volume Control" correspond au bloc "Master Volume" du schéma précédent.

Notes sur la résolution des convertisseurs et l'échantillonnage

On ne présente pas ici en détails les techniques de conversions A/N et N/A, sans intérêt si on est qu'utilisateur.
Par contre, même en tant qu'utilisateur, il faut savoir choisir la **résolution des convertisseurs** et la **fréquence d'échantillonnage** en fonction de l'application.

Dans les applications "audio", les signaux en bouts de chaînes de traitement sont toujours de nature analogique :

- le **microphone** converti la pression acoustique en signal électrique **analogique**
- le **haut-parleur** converti un signal électrique **analogique** en pression acoustique

Par contre, les traitements appliqués aux signaux audio (hormis l'amplification, bien entendu) sont aujourd'hui essentiellement **numériques** :

- mixage, filtrage, traitements spéciaux (simulation de son "cathédrale" par ex.)
- compression/décompression et mémorisation : sur disque ou silicium (mémoires flash)
- surround et dérivés

Mais, malgré ce fait, de nombreux audiophiles refusent encore aujourd'hui tout appareil numérique et préfèrent le disque vynil au CD !

Les signaux audio passent donc nécessairement par des fonctions de conversion analogique-numérique et numérique-analogique.
L'objectif est de ne **pas altérer l'information** transportée par le signal.
Dans le cas d'une prise de son, il s'agit de la pression acoustique prélevée en un point de l'auditorium. A priori, le signal le plus fidèle est celui produit par le microphone, à condition qu'il soit de qualité, bien placé et bien connecté au préamplificateur !
Les premiers traitements analogiques (amplification par ex.) altèrent déjà l'information : distorsion et bruit. Mais on sait faire aujourd'hui d'excellents préamplificateurs.
En sortie du préamplificateur, on dispose d'un signal électrique analogique très fidèle. Les caractéristiques électriques de ce signal sont les suivantes (en "Hi-Fi") :

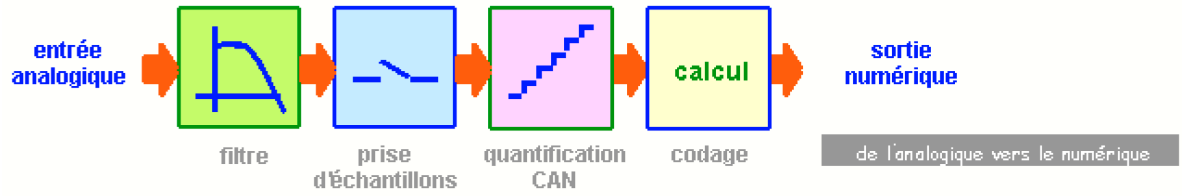
- Valeur efficace nominale : 0,775V, soit une amplitude de 1,1V environ
- Valeurs crêtes instantanées : +10V et –10V (soit 20 dB environ au delà du niveau nominal)
- Bruit : inférieur à 0,775mV eff entre 20Hz et 20kHz (soit 60dB de moins que le niveau nominal)
- Gamme de fréquences des composantes : 20Hz à 20kHz

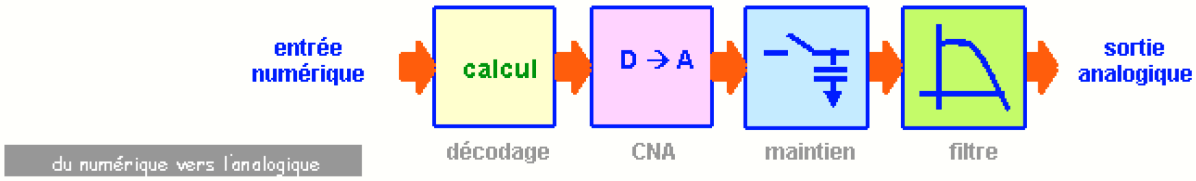
Ce signal doit être converti en numérique sans altérer ses caractéristiques. Les problèmes se situent dans cette fonction ! En effet, il faut passer d'un signal analogique **continu** (sens mathématique) à un signal numérique **discret** : c'est à dire une suite de nombres. Or :

- ces nombres sont codés en binaire et il faut obligatoirement limiter le nombre de bits
- le rythme de ces nombres devrait être infini pour se rapprocher au mieux d'un signal continu. Mais dans ce cas, le nombre de bits à traiter par seconde est infini aussi, ce qui est inenvisageable ! Par ailleurs, les CAN de précision ne réalisent pas instantanément leur conversion, ce qui limitera aussi le rythme du signal numérique.
- la technologie actuelle permet de réaliser des conversions jusqu'à 24 bits par échantillon

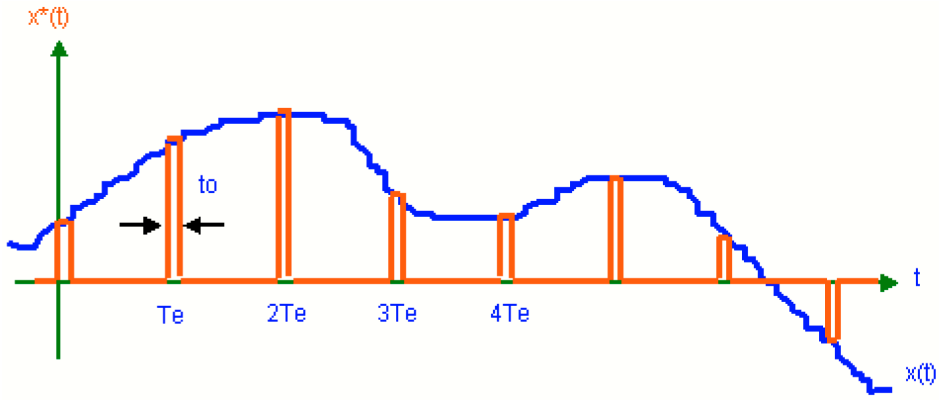
Des problèmes identiques se posent à l'autre bout de la chaîne de traitement, mais ils sont généralement plus facile à résoudre.

Une chaîne de traitement ressemble à celle-ci :





Echantillonnage

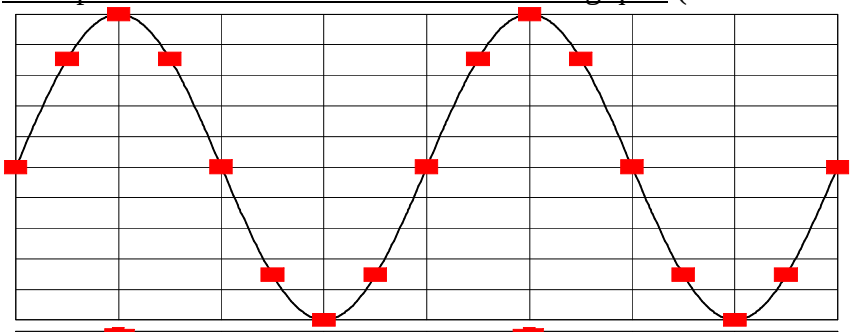


La fonction "échantillonnage" prélève des échantillons analogiques du signal d"entrée au ryhme de la fréquence FE (instants Te, 2Te, 3Te … sur la figure ci-contre).
Chacun de ces échantillons est converti en numérique pour produire une séquence de nombres : le signal numérique.

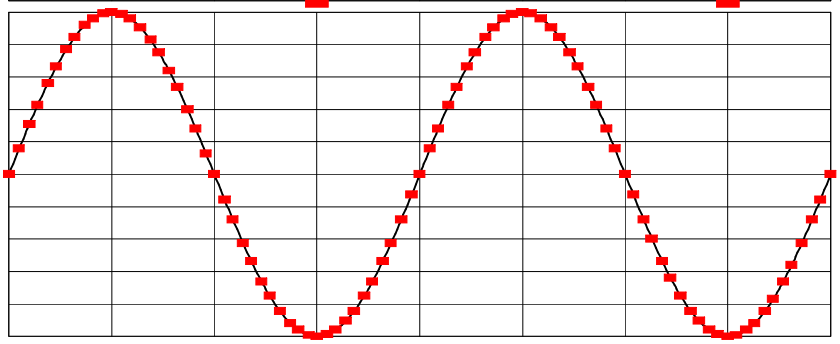
Le choix de FE dépend de la composition spectale du signal d'entrée :

Règle de Shannon : **la fréquence d'échantillonnage FE doit être supérieure à deux fois la fréquence de la composante la plus aigue du signal d'entrée**.
Dans notre cas il s'agit de 20kHz, ce qui justifie FE = 44,1kHz utilisées sur les CD audio.
En téléphonie, le spectre audio est volontairement limité à 3,5kHz. Les centraux téléphoniques numérisent alors le signal reçu avec une FE de 8 kHz.

Exemples de successions d'échantillons analogiques (avant conversion)



Echantillonage d'un signal 1 kHz à FE=8kHz



Echantillonage d'un signal 1 kHz à FE=44,1kHz

Note : le filtre passe-bas d'entrée est adaptée à la fréquence d'échantillonnage pour atténuer toutes les composantes du signal d'entrée ne respectant pas la règle de Shannon.

Fréquences d'échantillonnage standards : 8kHz – 11025Hz – 22050Hz – 32kHz – 44,1kHz – 48kHz
Certaines cartes audio vont jusqu'à quelques centaines de kHz

Conversion numérique-analogique
Le rythme du signal numérique est normalement égal à la fréquence d'échantillonnage FE (sauf cas particuliers) et le CNA associé à la fonction "maintien" reproduit la succession des échantillons pour produire un signal analogique en "marche d'escalier".
Le filtre passe-bas en sortie atténue les composantes indésirables (fréquences FE et multiples). Ses caractérisques sont adaptées à la fréquence d'échantillonnage.

Résolution

La résolution est liée au nombre de bits de chaque échantillon numérique.

Si n est la taille des échantillons numériques, la résolution vaut : $R = \frac{1}{2^n}$

Soit R=0,39%=3900ppm si n=8 et R=15ppm si n=16
Quelle résolution permet de respecter les caractéristiques du signal analogique données précédemment ?
On peut comparer la résolution exprimée en dB à la dynamique du signal analogique. Cette dernière vaut:

$$\text{Dynamique}_{\text{analogique}} = 20 \times \text{Log}_{10}\left(\frac{\text{Valeur_crête_max}}{\text{Valeur_crête_bruit}}\right) = 20\text{x}\text{Log}_{10}\left(\frac{10}{1{,}1\text{mV}}\right) = 80\text{dB}$$

La résolution exprimée en dB (en fait l'inverse) doit être meilleure que la dynamique :

- R_{dB} = 48dB si n = 8
- R_{dB} = 96dB si n =16

On constate que la résolution de 16 bits convient, ce qui justifie son choix pour les CD audio.
Bien entendu, cela ne suffit pas pour assurer l'intégrité de l'information sonore, mais c'est une condition nécessaire.
A l'époque de la définition des normes des CD audio la technologie ne permettait pas de dépasser la résolution de 16 bits. Aujourd'hui on atteint 24 bits à un coût raisonnable.

Codage

Le codage définit les lois de correspondance entre les échantillons analogiques et leurs équivalents numériques des CAN et CNA.
En notant A la valeur de l'échantillon analogique (en Volt) et N son équivalent numérique, les lois de correspondance sont :

➢ Pour le CNA : $\boxed{A_{CNA} = N_{CNA} \times \frac{V_{REF}}{2^{n-1}}}$ Avec $-2^{n-1} \le N_{CNA} \le 2^{n-1} - 1$

Si n = 8 : $-128 \le N_{CNA} \le 127$ (il y a 256 valeurs différentes)
$-V_{REF} \le A_{CNA} \le 0{,}992 \text{ x } V_{REF}$ par pas de $7{,}8125E^{-3} \text{ x } V_{REF}$ (quantum)
Si n = 16 : $-32768 \le N_{CNA} \le 32767$ (il y a 2^{16} valeurs différentes)
$-V_{REF} \le A_{CNA} \le 0{,}99997 \text{ x } V_{REF}$ par pas de $30{,}5E^{-6} \text{ x } V_{REF}$ (quantum)
V_{REF} est une tension de référence interne au CNA.

➢ Pour le CAN : le résultat numérique N_{CAN} est tel que :

$$\boxed{\frac{V_{REF}}{2^{n-1}}\left(N_{CAN} - \frac{1}{2}\right) \le A_{CAN} < \frac{V_{REF}}{2^{n-1}}\left(N_{CAN} + \frac{1}{2}\right)} \quad \text{Avec } -2^{n-1} \le N_{CAN} \le 2^{n-1} - 1$$

Ou plus simplement : $\boxed{N_{CAN} = 2^{n-1} \times \frac{A_{CAN}}{V_{REF}}}$ arrondi à l'unité la plus proche.

Pour des raisons inconnues, le **codage dépend de la résoltion** :

➢ Le **binaire décalé** (Bd) pour une résolution de 8 bits
➢ Le **complément à 2** (C2) pour une résolution de 16 bits

Définitions des codes Bd et C2 :

Notations : N : nombre entier à représenter (positif et négatif)
Rep(N) : sa représentation en binaire sur 8 ou 16 bits
$pos_{2,n}$ (X) : représentation en binaire natuel sur n bits du nombre entier positif X

1. Binaire décalé

La représentation Bd est adaptée aux convertisseurs A/N et N/A bipolaires.

$$Rep(N)_{Bd} = pos_{2,n}\left(N + 2^{n-1}\right)$$

Exemples sur 8 bits :

N		$Rep(N)_{Bd}$
0	→	10000000 (ou 80h)
+1	→	10000001 (ou 81h)
-1	→	01111111 (ou 7Fh)
+127	→	11111111 (ou FFh)
-128	→	00000000 (ou 00h)

2. Complément à 2

La représentation C2 est adaptée aux calculs.

$$\text{Si } N \geq 0 : \; Rep(N)_{Bd} = pos_{2,n}(N)$$
$$\text{Si } N < 0 : \; Rep(N)_{Bd} = pos_{2,n}\left(N + 2^{n}\right)$$

Exemples sur 16 bits:

N		$Rep(N)_{C2}$
0	→	00000000 00000000 (ou 0000h)
+1	→	00000000 00000001 (ou 0001h)
-1	→	11111111 11111111 (ou FFFFh)
+32767	→	01111111 11111111 (ou 7FFFh)
-32768	→	10000000 00000000 (ou 8000h)

Note : le bit MSB peut être interprété comme bit de signe : "0" ⇒ positif et "1" ⇒ négatif

3. Transcodage Bd ↔ C2

Les représentations doivent avoir la même taille.
Dans ce cas, **il suffit d'inverser l'état du bit MSB** quel que soit le sens du transcodage.

Exemples sur 8 bits :

N		$Rep(N)_{Bd}$	$Rep(N)_{C2}$
0	→	10000000	00000000
+1	→	10000001	00000001
-1	→	01111111	11111111
+127	→	11111111	01111111
-128	→	00000000	10000000

Questions d'étude

1. Que signifient les sigles et/ou acronymes : WAV, MP3, DSP, DMA, Hi-Fi, PDIF, CD, UC, Bd, C2, ADC, CDA, FE
2. Peut-on travailler avec l'ordinateur tout en écoutant et/ou enregistrant de la musique sans créer de perturbations ?
   Dans l'affirmative, quel procédé technique permet cette performance ?
3. Proposer une configuration des 2 fenêtres du "Contrôle de volume" (modes "enregistrement" et "reproduction") pour :
   - Enregistrer un mélange des sources "Line In" et "Wave".
   - Ecouter un mélange des sources "Wave", "CD Player" et "SPDIF".
4. La source "wave" est-elle numérique ou analogique ?
   Elle est interne au PC, mais qui ou quoi la produit ?
5. La source "Line In" provient d'un lecteur externe MP3 et comporte des composantes de fréquences 50Hz à 10kHz. La source "Wave" interne comporte des composantes de fréquences 100Hz à 3kHz.
   → Choisir la résolution du CAN et la fréquence d'échantillonnage les plus adaptées.
6. Codage : donner les représentations binaires des valeurs 0, max et min dans les codes Bd et C2 sur 8 bits et 16 bits.
7. Un fichier sonore de type "WAV" de durée une minute a été créé avec une fréquence d'échantillonnage de 44,1kHz. Cette information est mémorisée dans l'entête de ce fichier.
   On modifie ce paramètre avec 22050Hz sans toucher aux autres. On écoute alors ce fichier sonore avec un outil de lecture.
   → Quel sera le résultat ?