



Traitement Numérique du signal

Signal audio numérique

- Processus de phonation
- Signal de parole
- Codage de la parole
 - PCM, MIC
 - Format Wave
- Codage audio
 - Format CD
 - Comparaison des formats
- Traitement audio-numérique sur DSP
 - Génération de signaux
- Effets audio-numériques
 - Les effets liés au gain
 - Les effets à retard
 - Filtres numériques
 - Effets de réverbération
 - Effets de distorsion
 - Transformée de Fourier



Bibliographie

- 21061-Data Sheet, Analog Devices
- Applications audio numériques des DSP, Benoît Bouchez, Publitronic, 2003
- Cours de traitement de signal, O. Sentieys



Ondes sonoras

Caractéristiques

- Le signal (ici une onde sonore) sera caractérisée par
 - Une fréquence (hauteur)
 - L'oreille humaine est capable de percevoir des sons dans une plage allant de 20Hz à 20kHz
 - L'amplitude
 - (l'unité dépend du milieu de propagation)
 - Dans l'air elle appelée niveau sonore et se mesure par variation de pression sur une surface. Le seuil de l'oreille humain est estimé à $P_2=2 \cdot 10^{-5} \text{N/m}^2$. C'est le niveau minimum fixant le 0dB SPL (Sound Pressure Level)
 - Niveau SPL(dB) = $20 \log(P_1/P_2)$
 - Le spectre (timbre)
 - Transformée de Fourier

Série de Fourier

- Toute fonction périodique peut se décomposer en une somme de fonctions sinus et cosinus appelées harmoniques, et dont les fréquences sont des multiples de la fréquence la plus basse, appelée fondamentale
- Synthèse :

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi ftn) - \sum_{n=1}^{\infty} b_n \sin(2\pi ftn)$$

Analyse

- Opération inverse de la synthèse
- Les termes a_n et b_n représentent le spectre harmonique du signal

$$a_0 = \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt$$

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos\left(\frac{2\pi ftn}{T}\right) dt$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin\left(\frac{2\pi ftn}{T}\right) dt$$

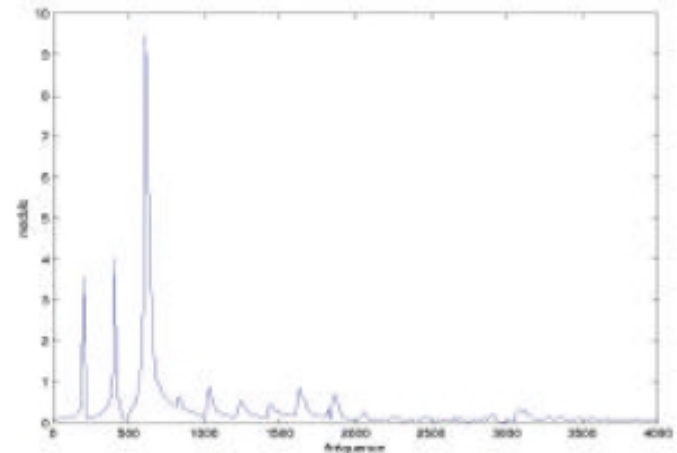
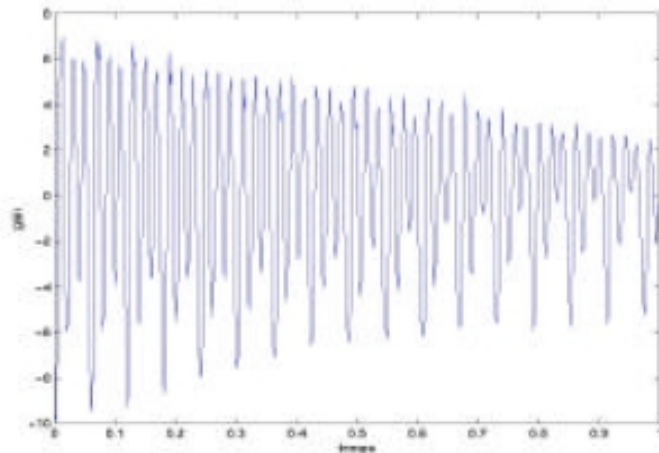
Transformée de Fourier

- La transformée de Fourier s'applique à n'importe quel signal alors que les séries de Fourier ne s'appliquent qu'aux signaux périodiques
 - Cf. chapitre sur la FFT

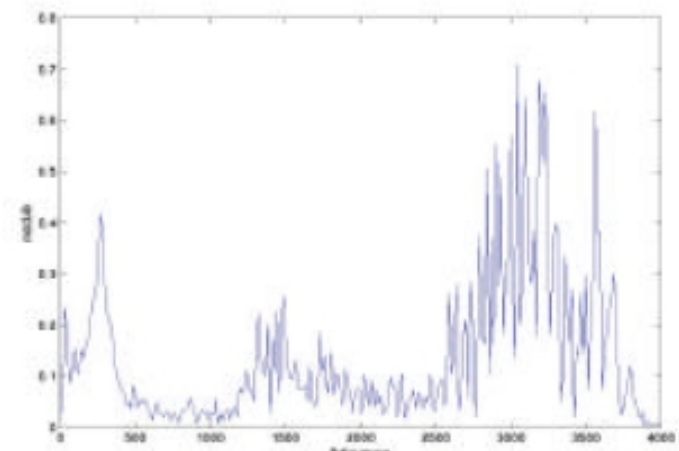
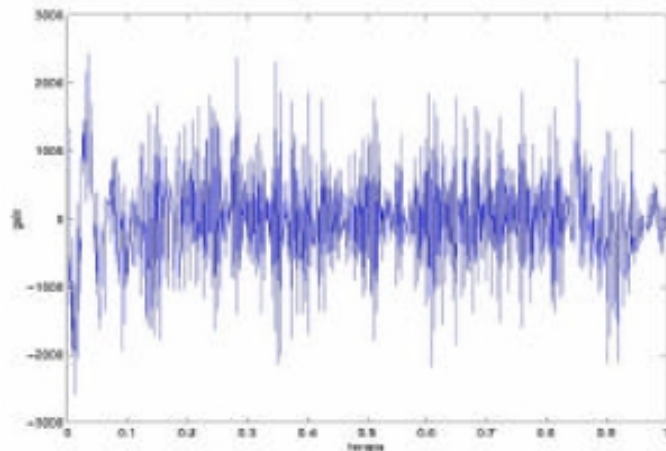
Le son et l'homme

- Le processus de phonation chez l'homme lui permet d'émettre des sons dont le premier harmonique (fréquence fondamentale de vibration des cordes vocales) de 80 à 100Hz chez le male, de 175 à 300Hz chez la femelle et de 200 à 600Hz chez l'enfant.

Son voisé



un son voisé et son spectre (son " eu ")



un son non voisé et son spectre (son " ch ")



Acquisition du signal

Echantillonnage du son

- Passage d'un monde continu à un monde discret
 - Échantillonnage
- Théorème de Nyquist :
 - Si la fréquence d'échantillonnage d'un signal continu est au moins égale au double de la composante de fréquence la plus élevée de ce signal, on retrouvera après échantillonnage l'intégralité des composantes (spectrales) du signal d'origine.

Quantification

- En sortie d'un échantillonneur bloqueur, le signal est toujours analogique,
- L'étape préalable au traitement est donc la numérisation du signal
- Réalisée par un convertisseur analogique/numérique (CAN/ADC)
- La largeur du mot de quantification (n bits) est appelée résolution
 - Exemple sur 8 bits, la résolution est de 256, soit pour un signal d'entrée entre 0 et 1V, $1/256 = 0,00390625V$
 - Cette dernière valeur représente donc l'erreur de quantification maximale
 - L'erreur de quantification est calculée par $1/\sqrt{12}$ LSB, soit 0,29 LSB
 - Cette erreur se traduit par l'ajout d'un bruit aléatoire appelé bruit de quantification

Comparaison des supports

Support	GSM	Mini DV	CD Audio	DVD vidéo	DVD HD
Fréquence d'échantillonnage(Hz)	8000	32,000 / 48,000	44,100	48,000	96,000
Quantification (bits)	13	12 / 16	16	16	24
Canaux	1	2	1 ou 2	1 à 6 (5.1)	1 à 8

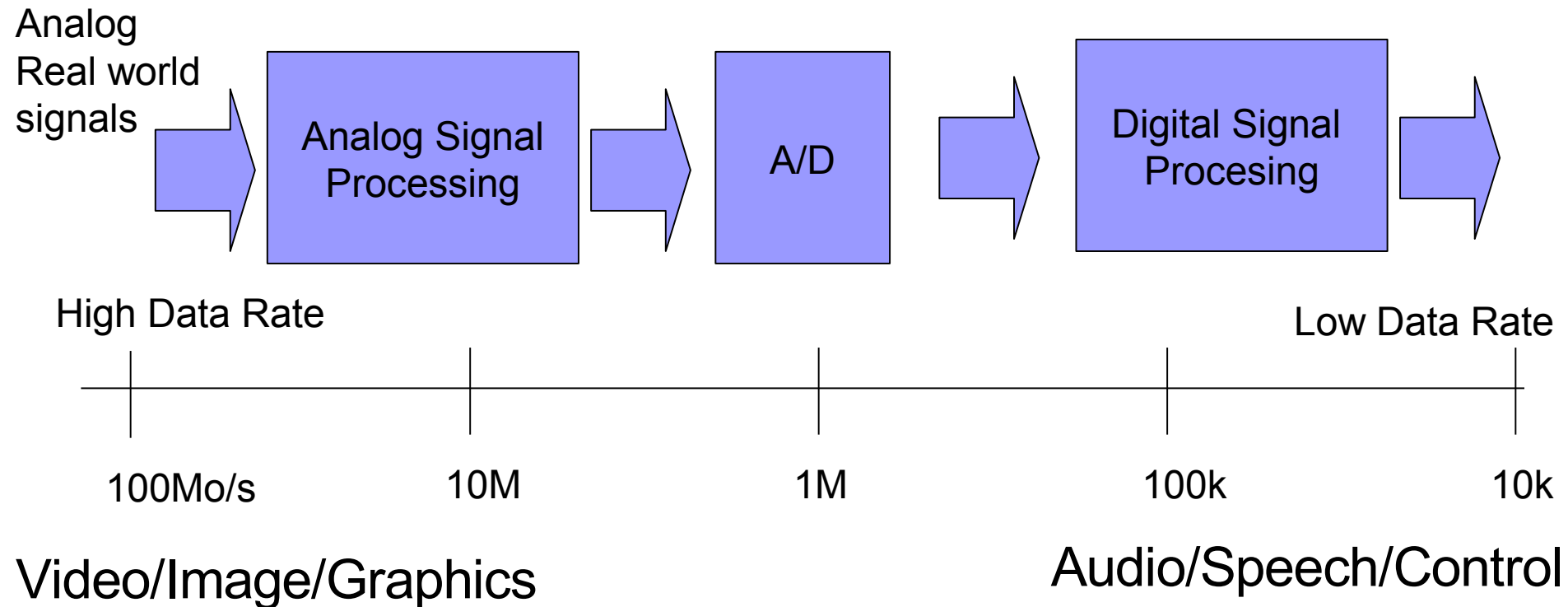
Précision

- On appelle précision d'un convertisseur le rapport entre une valeur binaire obtenue en sortie du convertisseur et la valeur binaire qu'il aurait fallu obtenir.
- La précision s'exprime en bits utiles car les circuits électroniques ajoutent des bruits au signal d'entrée

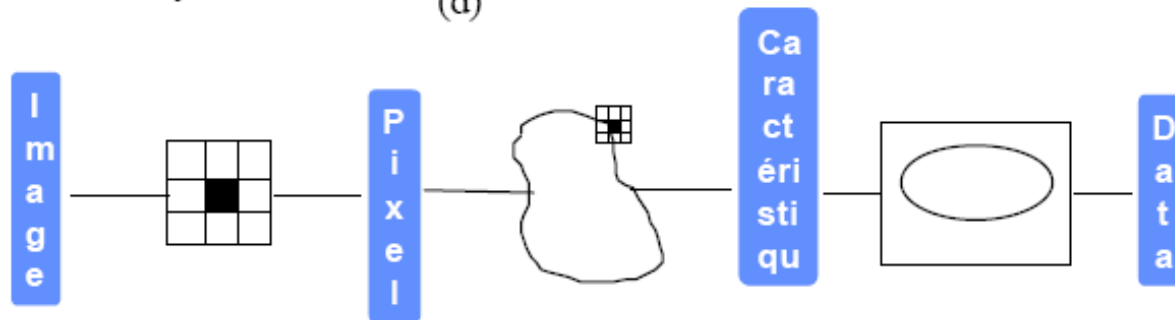
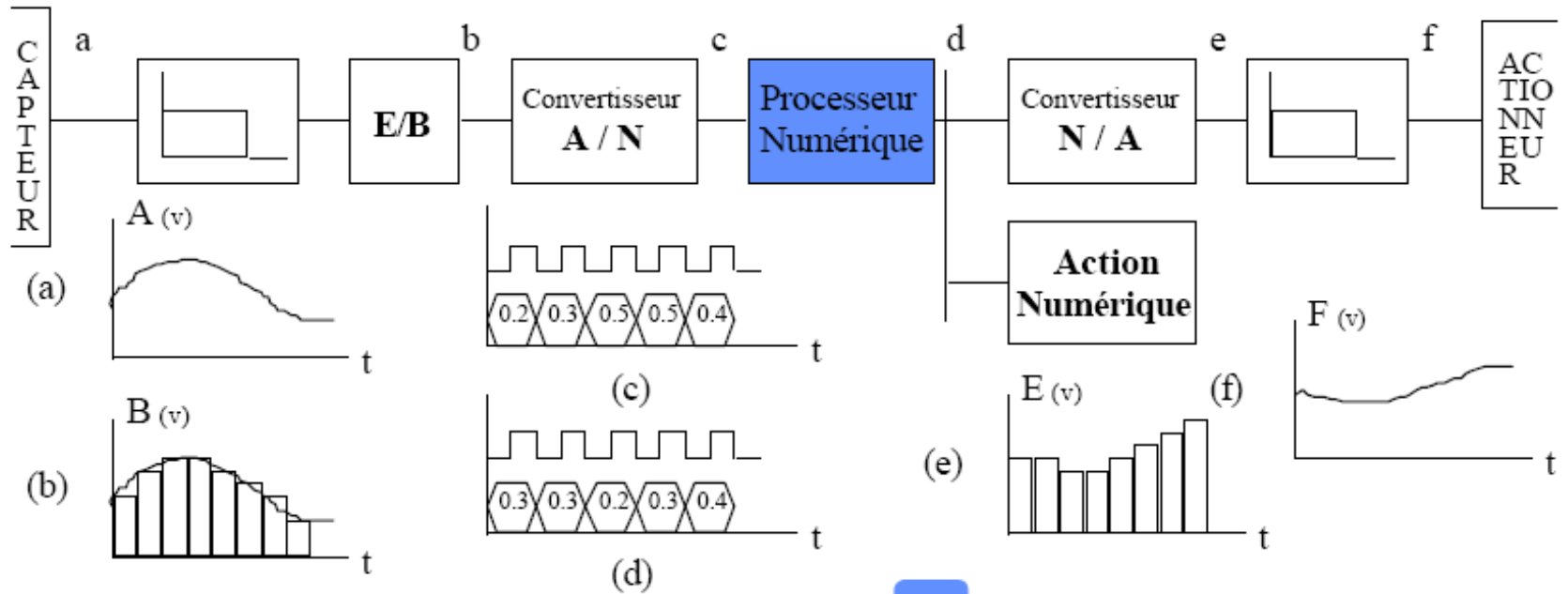
Dynamique

- La gamme de valeurs que peut prendre le signal d'entrée est appelé dynamique
- La dynamique d'un système audionumérique sur n bits est calculé par
 - $\text{Dynamique (dB)} = 6,02n + 1,76$
 - Soit à peu près 6dB par bit de résolution

Traitement numérique du signal (Digital Signal Processing)



Chaîne de TNS





Format WAV

Format WAVE

- Le format WAV est un fichier suivant la norme R.I.F.F. (Resource Interchange File Format). Un fichier de format R.I.F.F. est un fichier formé de différentes parties appelées CHUNK. Un CHUNK est en fait un descripteur de fiches ayant la structure suivante :
- **Structure d'un CHUNK**
 - char FOURCC[4] Identificateur du CHUNK
 - DWORD size Taille du buffet suivant en octets
 - BYTE data[size] Les données de la fiche.

Chunk RIFF

- Un fichier RIFF commence toujours par le premier CHUNK appelé CHUNK RIFF.
- Il commence à l'adresse 0 du fichier et a une structure de taille fixe.
- Le CHUNK RIFF est l'exception qui confirme la règle car le champ « size » donne la taille du fichier entier en octets et non pas la taille de la fiche comme c'est le cas d'habitude

Format Wave

- 4 octets : Chunk 'RIFF'
- 4 octets : Taille du chunk (taille du fichier – 8)
- 4 octets : Chunk 'WAVE'
- 4 octets : Chunk 'fmt '
- 4 octets : Taille du chunk fmt
- 2 octets : Encodage audio (1 = PCM)
- 2 octets : Nombre de canaux
- 4 octets : Fréquence d'échantillonnage
- 4 octets : Nb d'octets par seconde
- 2 octets : Nb d'octets par bloc
- 2 octets : Nb de bits par échantillon
- 4 octets : Chunk 'DATA'
- 4 octets : Taille du chunk DATA
- Les échantillons, exemple sur 16 bits, 2 canaux
 - 2 octets : Sample0Channel0
 - 2 octets : Sample0Channel1
 - 2 octets : Sample1Channel0
 - 2 octets : Sample1Channel1
 - ...



Génération de signaux

Génération de signaux par table d'échantillons

- Utilisés dans de nombreux synthétiseurs, l'onde est lue de manière cyclique dans une table d'échantillons,
- Cette technologie porte le nom de PCM (Pulse Code Modulation) (ou AWM, DCO...)
- La table est manipulée par le DSP et les DAG par un tampon circulaire (ring buffer)
 - Pour les signaux périodiques, on ne mémorise qu'une période
 - Pour les autres toute la durée => stockage mémoire !

Table d'échantillons à fréquence variable

■ Exemple :

- Une table d'onde de 256 échantillons représentant une sinusoïde lue à 44,1KHz génère un signal de fréquence de $44100/256=172,265625\text{Hz}$
- Pour changer dynamiquement la fréquence du signal
 - Soit changer le nombre d'échantillons (impossible)
 - Soit changer la fréquence d'échantillonnage (lire la table plus ou moins vite)

Fréquence de lecture variable

■ Lecture rapide

- Incrémenter l'index de 2, 3, 4 cases par cycle revient à produire des signaux avec une fréquence de $2f$, $3f$, $4f$... sans perte de qualité

■ Lecture lente

- Pour réaliser l'opération inverse, et générer des signaux de $f/2$, $f/3$...
- Il faudra incrémenter l'index de 0.5, 0.33 ... par cycle
- Possible si le compteur est codé en virgule fixe et si le même échantillon est lu tant que le compteur n'avance pas d'une unité!
- La précision sur 16 bits (format Q8.8) à 44,1kHz est de 0,6728Hz :
 - Avec un incrément de 0x0100 (1,0), $f=172,26\text{Hz}$
 - Avec un incrément de 0x0101 (1,0038), $f=172,83\text{Hz}$
 - ...

■ Ces opérations d'allongement/rétrécissement du signal ne sont pas sans effet sur le spectre et provoquent une transposition

Code SHARC du générateur

// Bloc de code principal du générateur

Generation:

```
    bit clr mode1 ALUSAT;           // Permet le rebouclage du compteur
    r0=dm(Compteur_Wavetable);
    r1=dm(Increment);
    r0=r0+r1;                       // Calcul adresse échantillon suivant
    dm(Compteur_Wavetable)=r0;

    r0=lshift r0 by -25;           // Recupération des 7 bits d'adresse (table d'onde 128
    échantillons)
    i0=Wavetable;                 // Pointe sur la table d'onde
    l0=0;                          // Buffer non circulaire

    m0=r0;
    nop;
    r0=dm(m0, i0);                // lecture échantillon pointé

    dm(Sortie_Gauche)=r0;
    dm(Sortie_Droite)=r0;
```

Interpolation

- On ne peut déterminer la valeur de l'échantillon pour des index fractionnaires,
- Il s'en suit un défaut dans le signal généré appelé artefact de numérisation se traduisant par des 'marches d'escalier' dans le signal faisant apparaître des harmoniques indésirables détectables à l'oreille.
- Pour corriger :
 - Soit filtre à fréquence de coupure variable
 - Soit interpolation des échantillons :

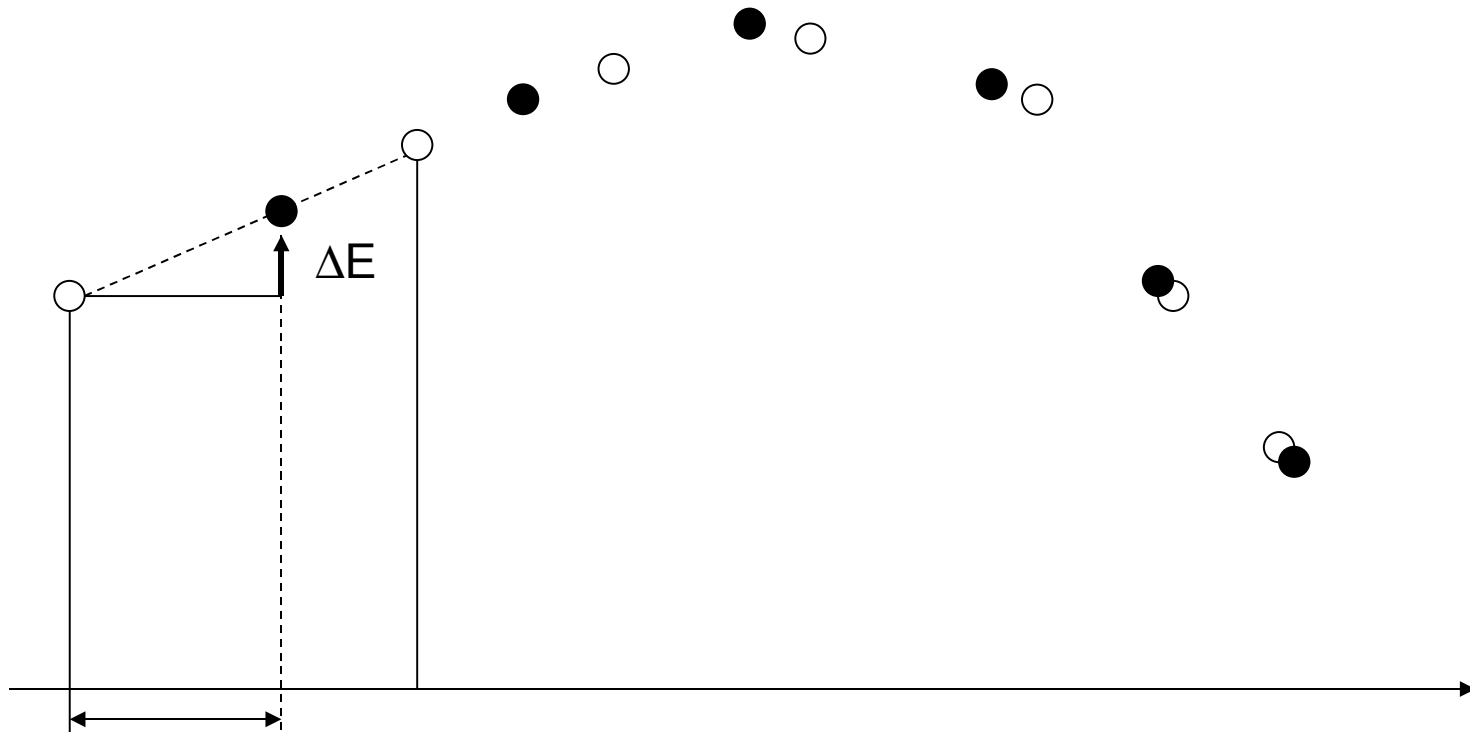
$$y_t = x_t + ((x_t - x_{t-1}) \times \textit{frac}(\textit{index}))$$

Interpolation

○ Echantillons mémorisés en table d'onde

● Echantillons interpolés

$$\Delta E = (x(n) - x(n-1)) \times F(n)$$



$F(n)$ = Partie fractionnaire de l'incrément de lecture

Code SHARC de l'interpolation

Generation:

```
bit clr mode1 ALUSAT; // Permet le rebouclage du compteur
r0=dm(Compteur_Wavetable);
r1=dm(Increment);
r0=r0+r1; // Calcul adresse échantillon suivant
dm(Compteur_Wavetable)=r0;

// Récupération échantillon adresse N
r0=lshift r0 by -25; // Recupération des 7 bits d'adresse (table d'onde 128 échantillons)
i0=Wavetable; // Pointe sur la table d'onde
l0=0; // Buffer non circulaire
m0=r0;
nop;
r7=dm(m0, i0); // R7 = Echantillon N. Adressage premodifié (sans modification de i0)

// Récupération échantillon adresse N+1
r1=1;
r0=r0+1; // Incrémentation adresse sur 7 bits
r2=0x7F;
r0=r0 and r2;
m0=r0;
nop;
r6=dm(m0, i0); // R6 = Echantillon N+1

// Interpolation linéaire sur partie fractionnaire de l'adresse
r0=r6-r7; // r0=[N+1]-[N]
r1=dm(Compteur_Wavetable);
r1=lshift r1 by 6; // Justification adresse Q7.25 en Q1.31
r2=0x7FFFFFFF;
r1=r1 and r2; // Masque bit de signe
r2=r1*r0(SSF); // r2=frac(adresse)*(Echantillon[N+1]-Echantillon[N])
r0=r7+r2; // Sortie=Echantillon[N]+(frac(adresse)*(Echantillon[N+1]-Echantillon[N]))

dm(Sortie_Gauche)=r0;
dm(Sortie_Droite)=r0;
```

Générateur de bruit

- Un bruit est un signal composé de toute les fréquences
- Bruit brownien
 - Produit par le mouvement des atomes autour du noyau des atomes
- Bruit blanc
 - Lorsque toutes l'amplitude est la même sur toute l'amplitude du spectre
- Bruit rose
 - Atténuation du contenu spectral de -3dB/octave



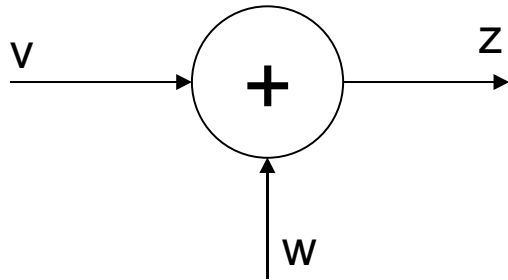
Production de bruits blancs

- Corps de Gallois

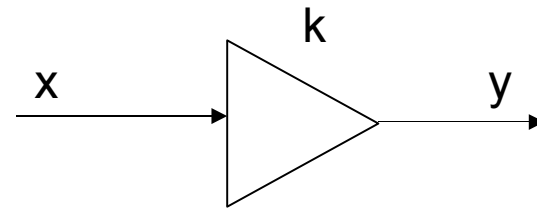


Effets liés au gain

symboles

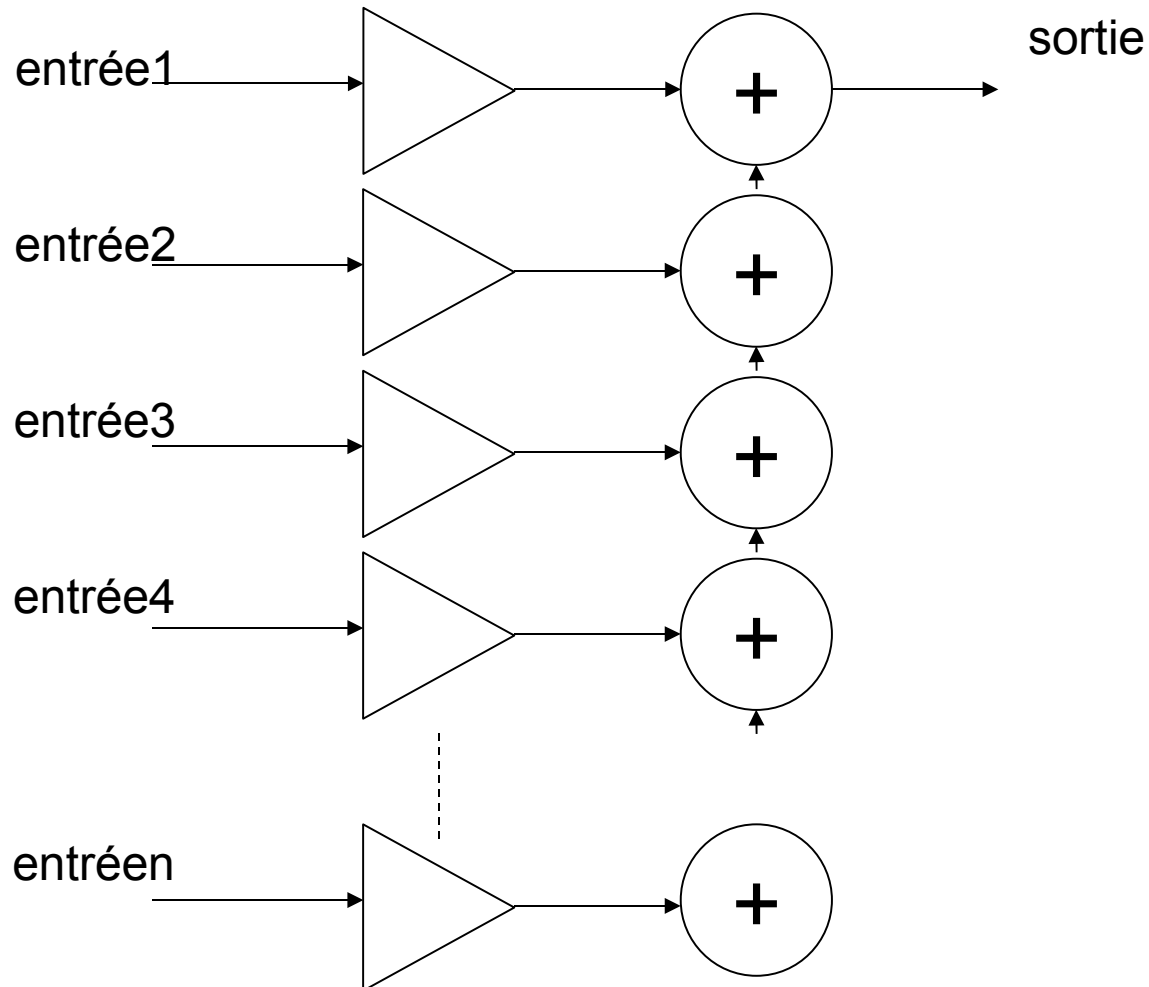


$$z(n) = v(n) + w(n)$$

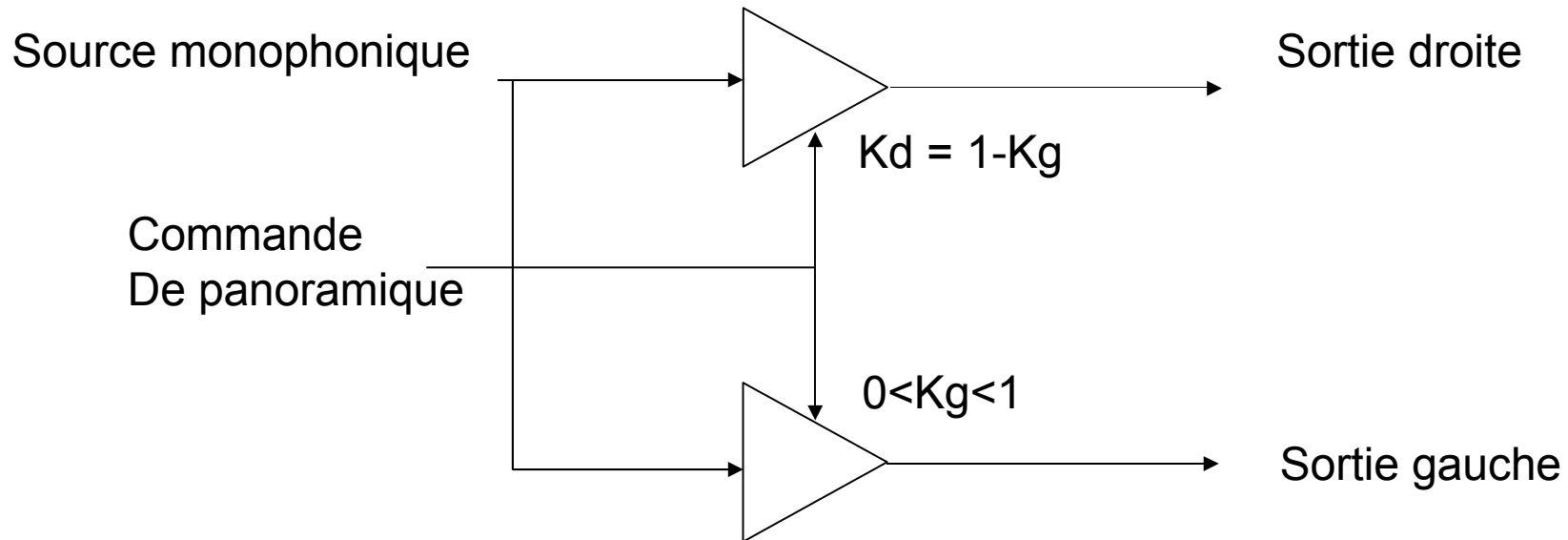


$$y(n) = x(n).k$$

Mélangeur numérique

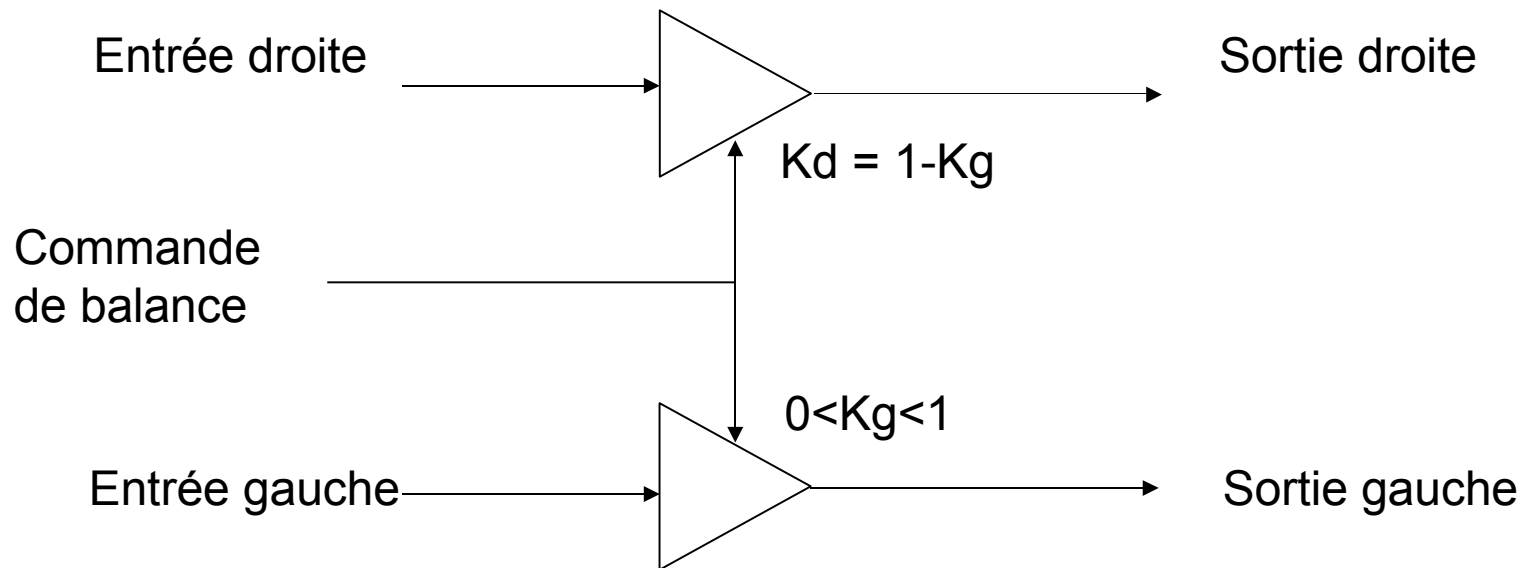


Contrôle d'image panoramique mono-> stéréo



Utilisé pour le mixage d'instruments en enregistrement

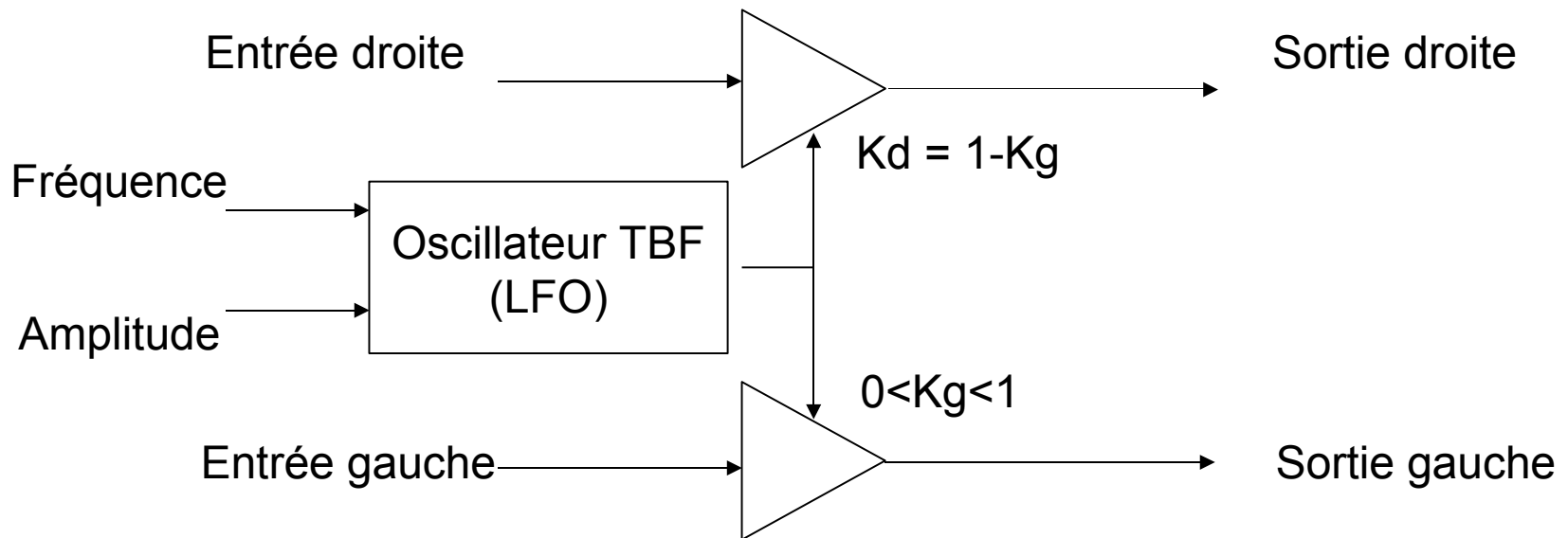
Balance stéréo



Trémolo

- Le trémolo est une modulation d'amplitude sinusoidale ou triangulaire à une fréquence comprise entre 0,5 et 5Hz.
- Sa réalisation se fait sur DSP par l'intermédiaire de LFO (Low Frequency Oscillator)
 - Ce n'est rien d'autre qu'une table d'onde dont on contrôle la fréquence par le pas de lecture
 - Et l'amplitude par un coefficient de réglage de la sortie de la table

Trémolo





Effets de retard

Retard spatial

- Le cerveau humain est capable de détecter des décalages temporels infimes dans la perception d'un son par les 2 oreilles. Si un même son parvient à l'oreille la plus éloignée de la source avec un retard de qqs ms par rapport à l'autre, il replacera la source du signal dans l'espace.
- Par ce principe, il est possible de placer les sources de son dans l'espace.
- (vitesse de propagation dans l'espace : 340m/s)

Retard, écho et réverbération

- Le retard est le temps mis par un signal pour se propager
- L'écho consiste à simuler la réflexion d'un son dans un espace
- La réverbération est un phénomène acoustique plus complexe que l'écho.

C'est le résultat de la réflexion multiple d'une onde sonore diffusée dans toutes les directions

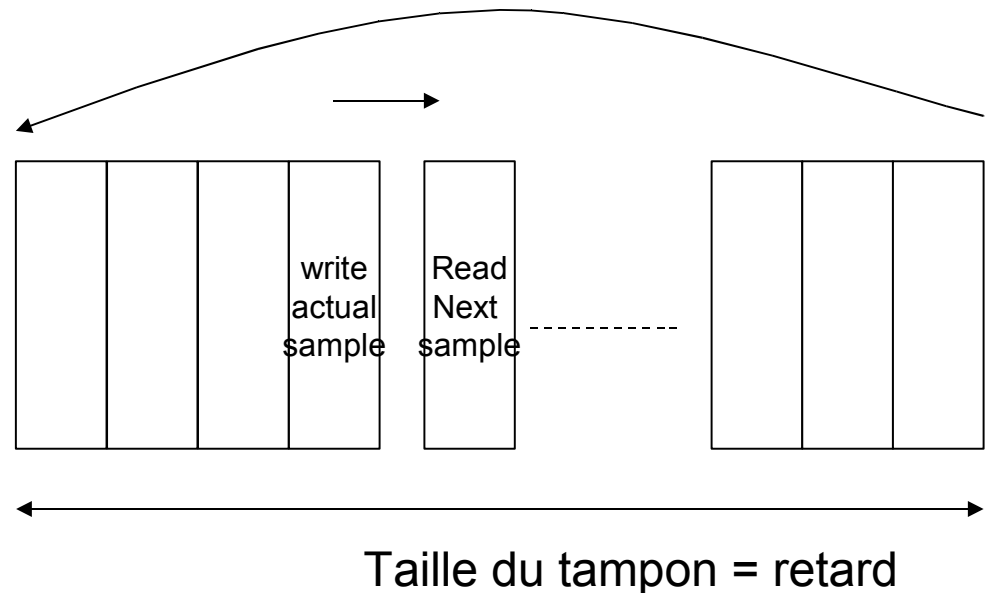
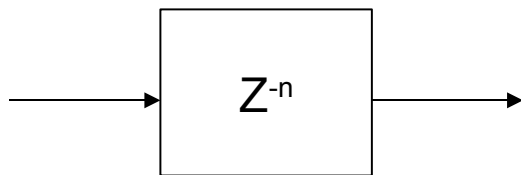
Ligne à retard avec les DSP

- On utilise un tampon circulaire (DAG) dont la lecture et l'écriture sont retardées de la longueur du buffer

$$\text{Retard(s)} = \text{tailleBuffer}/F_s$$

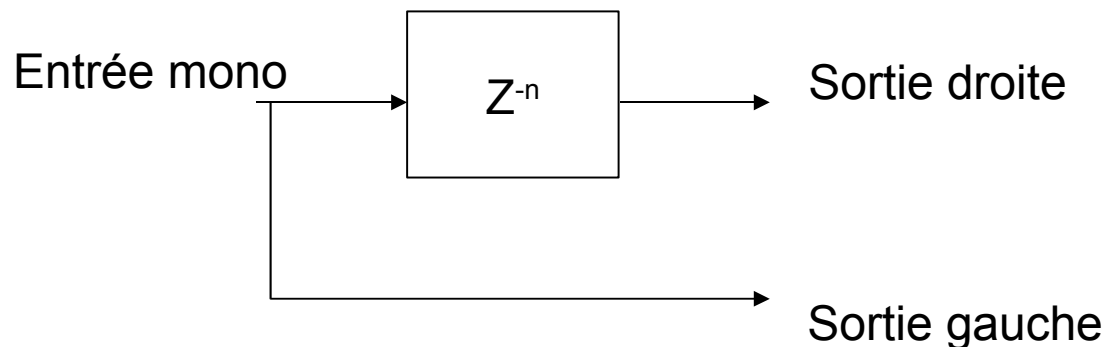
Exemple pour $\text{tailleBuffer} = 2048$,

$\text{Retard} = 2048/44100 = 46,4\text{ms}$



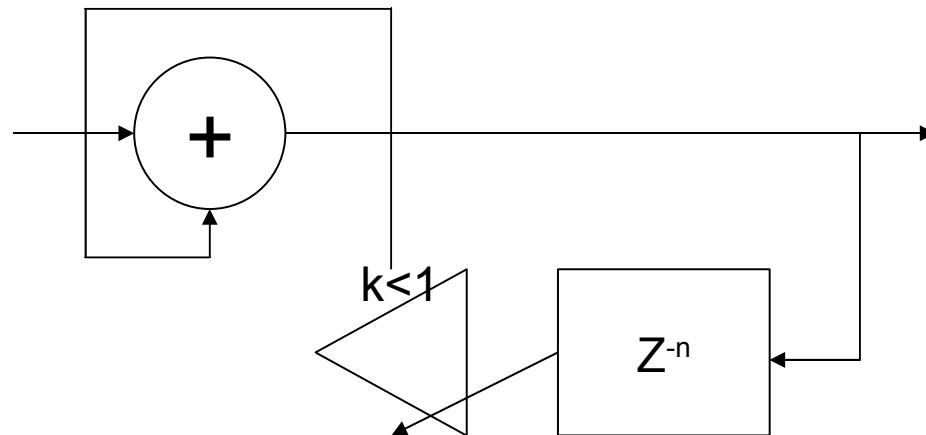
Doubleur

- En mélangeant le signal retardé et le signal original on obtient l'effet de doubleur (Slapback) : dédoublement du signal
- Le retard utilisé est entre 10 et 30ms
- Si ce mélange se fait de manière stéréophonique, on provoque un déplacement du son dans l'image stéréophonique

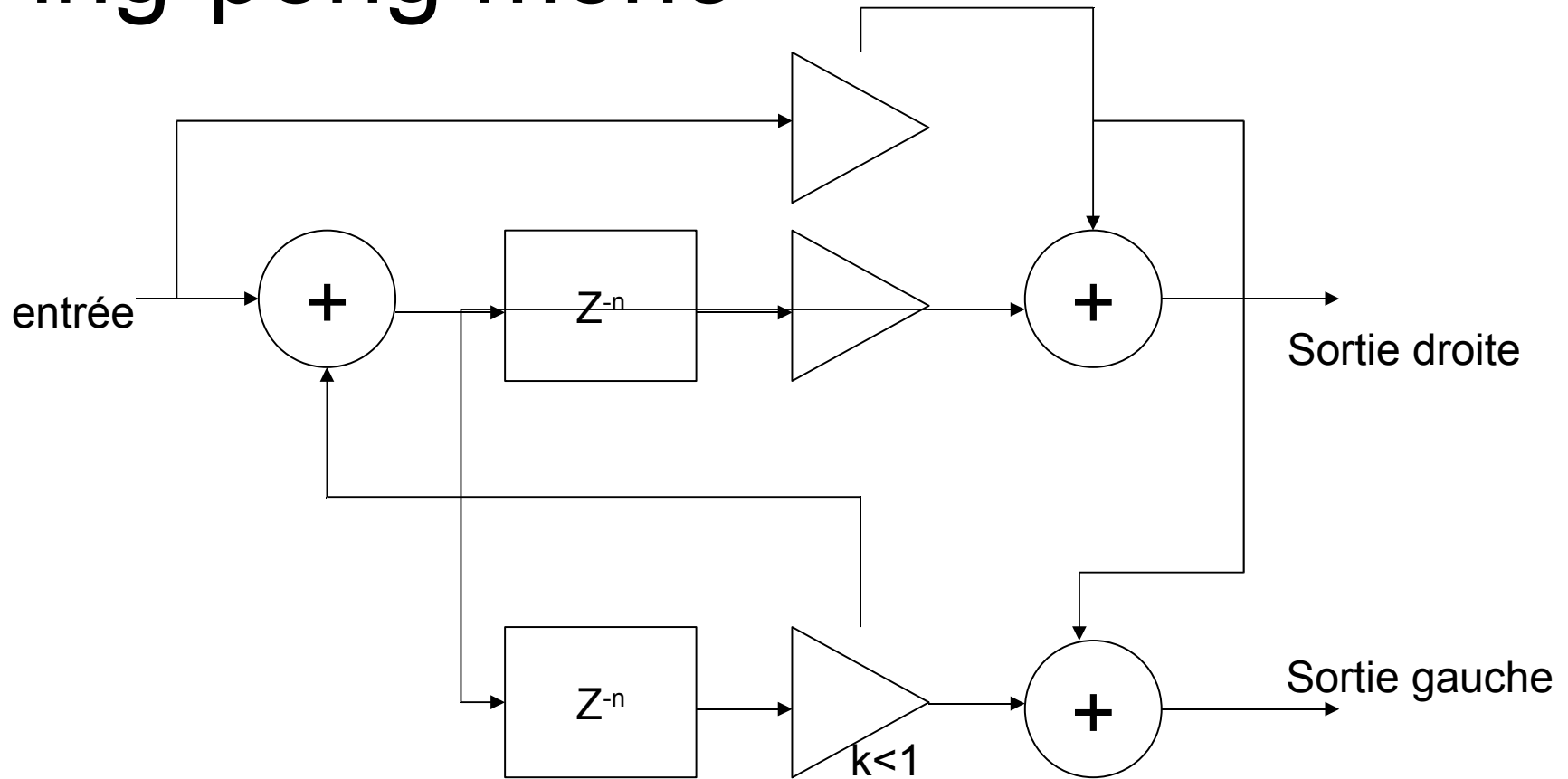


L'écho

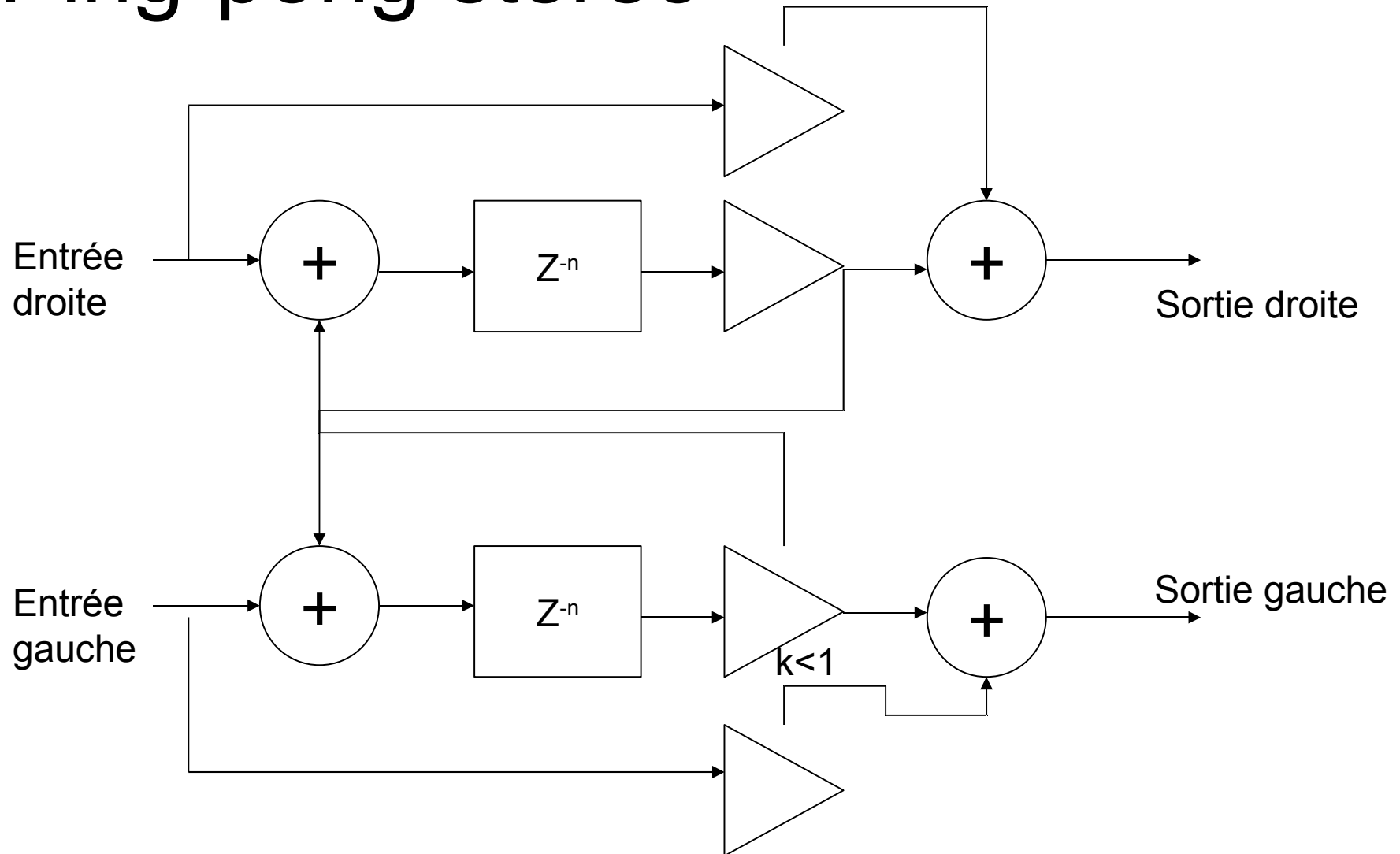
- Plus le coefficient de réinjection k est proche de 1, plus l'atténuation de l'écho est progressive et plus l'effet dure longtemps




Ping-pong mono



Ping-pong stéréo





Filtre à peigne



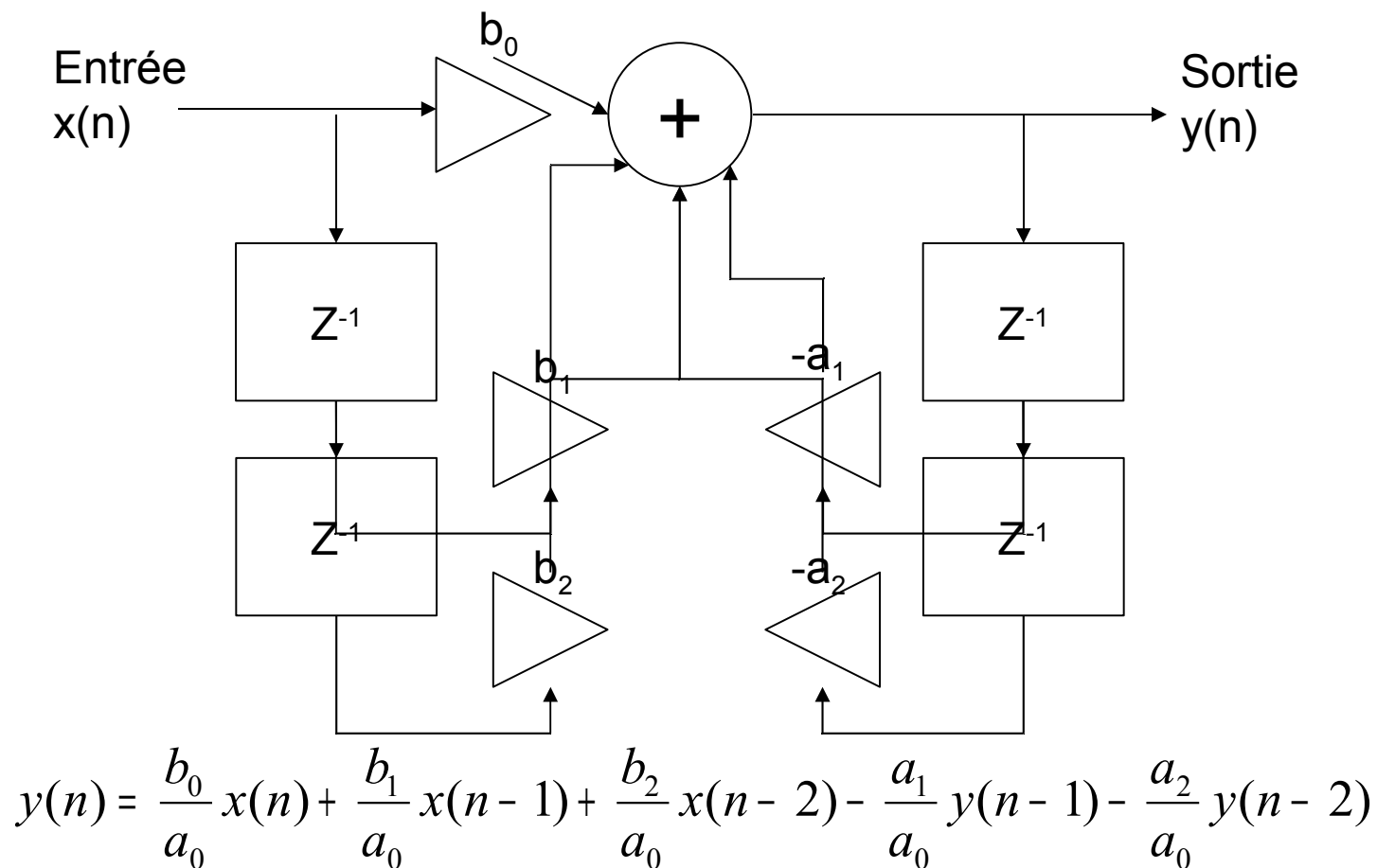
Filtres

Filtres récursifs

- Les filtres récursifs sont regroupés sous le nom générique des filtres à réponse impulsionnelle infinie (RII)

Cellules biquadratiques

- Ce sont des filtres RII du second ordre.



Utilisation des biquads pour filtres passe-bande

Les formules de Bristow-Johnson

permettent de déterminer les coefficients d'une cellule biquad :

$$A = \sqrt{10^{\frac{\text{gain}_{dB}}{20}}}$$

$$C = \cos\left(\frac{2\pi f}{F_s}\right)$$

$$S = \sin\left(\frac{2\pi f}{F_s}\right)$$

$$\alpha = \frac{S}{2Q}$$

$$\alpha = S \cdot \sinh\left(\frac{\ln(2)}{2} \cdot W \cdot \frac{2\pi f}{S}\right)$$

Coefficients pour
filtre passe-bande

$$b_0 = 1 + (A\alpha)$$

$$b_1 = -2C$$

$$b_2 = 1 - (A\alpha)$$

$$a_0 = 1 + \left(\frac{\alpha}{A}\right)$$

$$a_1 = -2C$$

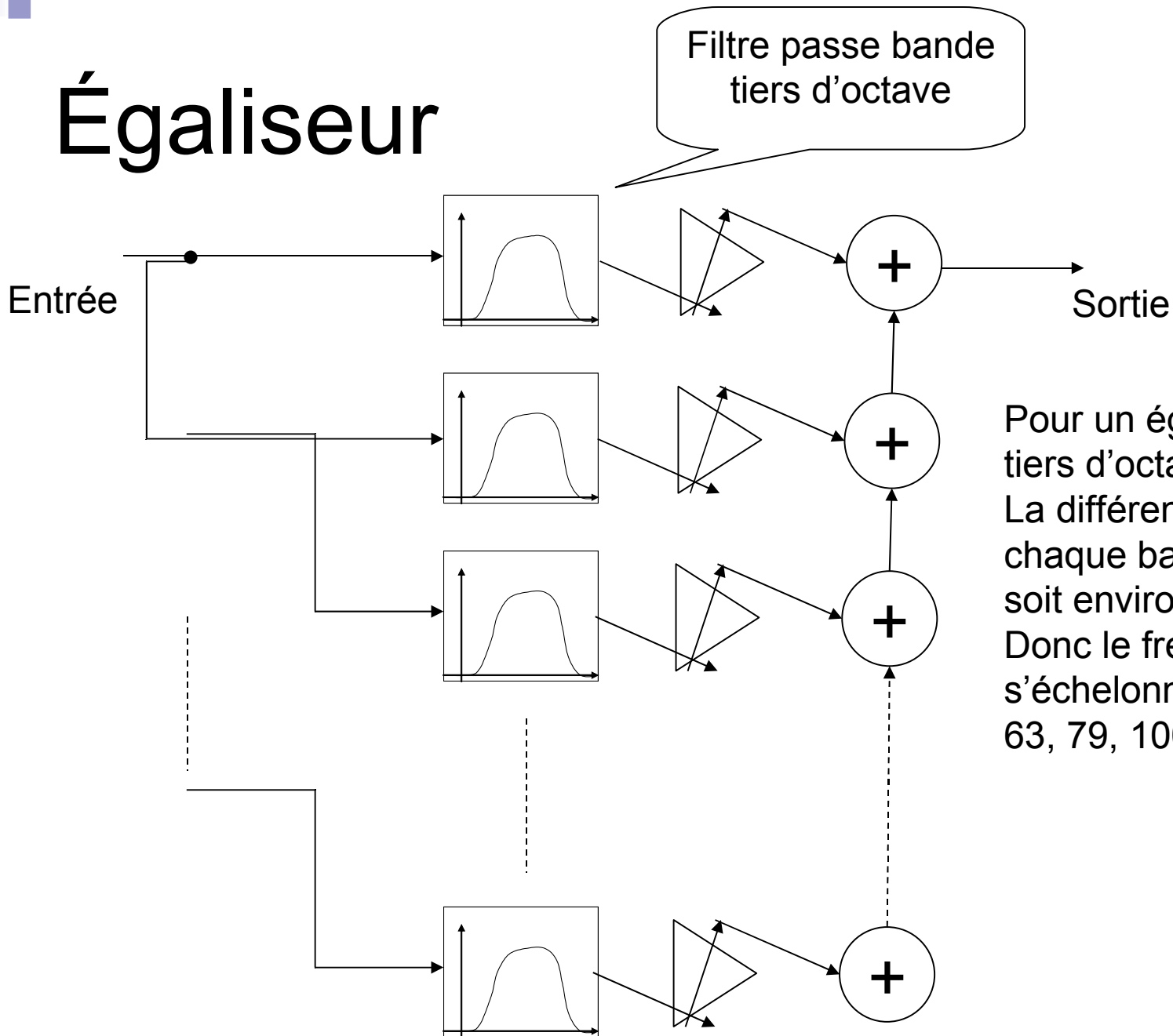
$$a_2 = 1 - \left(\frac{\alpha}{A}\right)$$

Fonction du facteur de qualité Q
ou de la largeur de bande W

Filtres d'égalisation

- Une des applications en audionumérique est l'égalisation en fréquences, c'est-à-dire d'atténuer ou d'amplifier une gamme de fréquence
- Égaliseur :
 - Utiliser une batterie de filtres passe-bas réparties sur le spectre.
 - Chaque filtre est suivi d'une cellule de gain réglable permettant de choisir le niveau de chaque bande de fréquence.

Égaliseur



Pour un égaliseur à tiers d'octave :
La différence entre chaque bande est $2^{1/3}$, soit environ 1,26.
Donc le fréquence s'échelonnent : 50Hz, 63, 79, 100, 126, 158...

Fonctions typiques de TNS

- Convolution, filtrage
 - $y = y + x.h$: MAC
- Adaptation
 - $Y_n = y_{n-1} + x_n.h$
- FFT, multiplication complexe
 - $X_r = x_r.w_r - x_i.w_i$;
 - $X_i = x_r.w_i + x_i.w_r$
- Viterbi
 - $a_1 = x_1 + x_2$;
 - $a_2 = y_1 + y_2$;
 - $y = (a_1 > a_2) ? a_1 : a_2$



Effets de distorsion

Effet de distorsion

- La distorsion correspond à un amplificateur poussé en **saturation** ce qui fait apparaître de nombreuses harmoniques issu des amplis tubes à vide (à lampes) dans les années 50 et 60.

Saturation numérique

- En numérique, la saturation consiste en un dépassement de capacité.
- Pour éviter l'overflow, les DSP proposent un mode de saturation activé par un registre d'état (bit ALUSAT de mode1)
 - Mode normal : 32500+300 dans un registre 16 bits signé = -32735
 - Mode saturation : 32767

Production d'un effet de distorsion

- Appliquer un gain suffisant sur un signal d'entrée en mode saturation
- + reproduire la non linéarité des composants magnétiques.
 - => appliquer un gain non linéaire au signal entrant
 - Si $|x| < a$
 - $Y = x$
 - Sinon
 - $y = a + (x - a / (1 + (x - a / (1 - a))))^2$



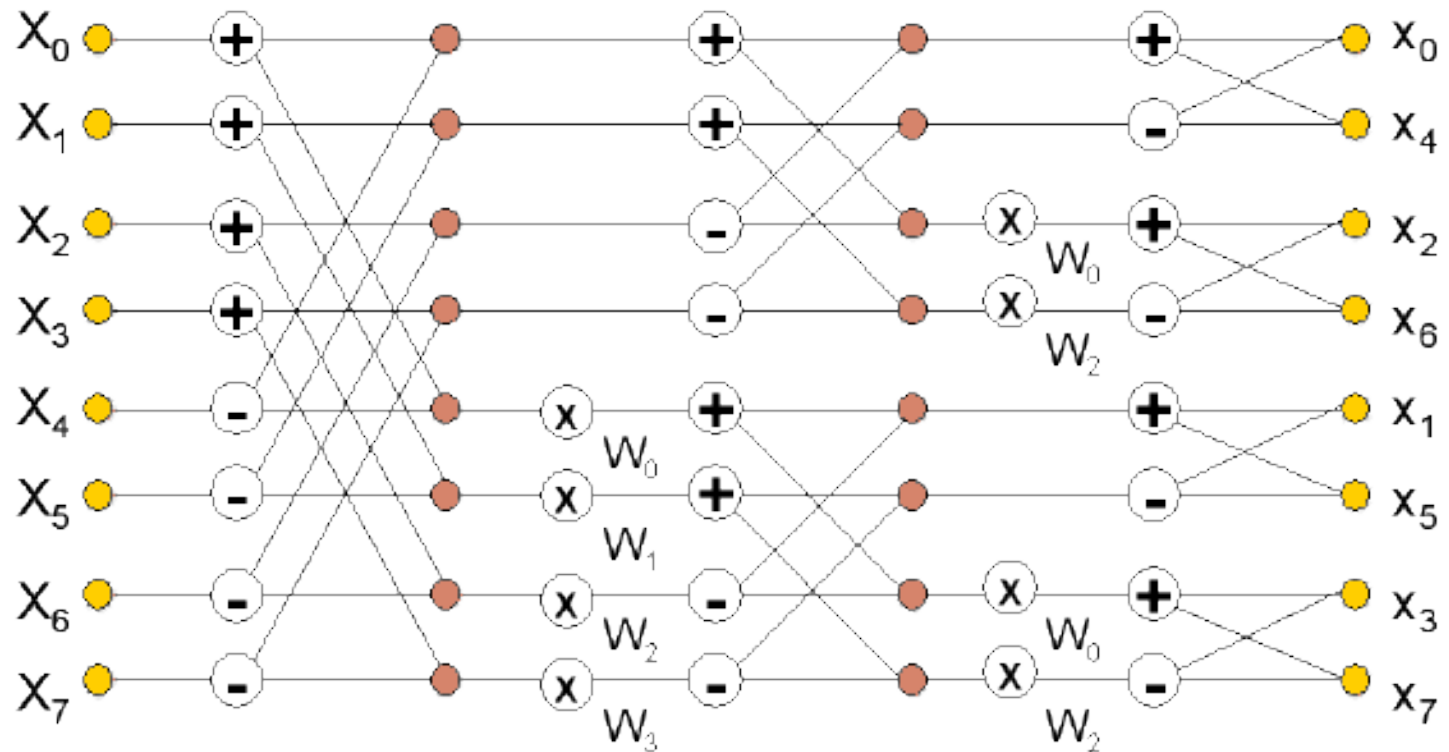
Transformée de Fourier Rapide

FFT

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \cdot e^{-2 \cdot j \cdot \pi \frac{nk}{N}}$$

$$\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \dots \\ X_{N-1} \end{bmatrix} = \frac{1}{N} \bullet \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & W & W^2 & W^3 & \dots & W^{N-1} \\ 1 & W^2 & W^4 & W^6 & \dots & W^{2(N-1)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & W^{N-1} & W^{2(N-1)} & \dots & \dots & W^{(N-1)(N-1)} \end{bmatrix} \bullet \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \dots \\ x_{N-1} \end{bmatrix}$$

Exemple de FFT sur 8 points



$$W = e^{-j \frac{2 \cdot \pi}{N}}$$

Application et effets

- Analyse spectrale
 - Calcul du module
 - Calcul de l'argument
- Filtrage par une courbe de réponse
 - Calcul FFT
 - Convolution par la courbe du filtre
 - Transformée inverse
 - Grande précision mais calculs lourds

$$|a + bi| = \rho = \sqrt{a^2 + b^2}$$

$$\theta = \arctan(a / b)$$

Transposition

- Cet effet permet de transposer (ou de décaler) le spectre d'un signal vers les fréquences hautes ou vers les fréquences basses.
- Cette opération peut se réaliser grâce à la transformée de Fourier sans modification de la durée du son :
 - FFT
 - Décalage de toute (ou certaines) fréquences
 - Transformée inverse

Étirement temporel

- C'est l'opposé de la transposition, il permet d'augmenter ou de diminuer la durée d'un signal sans en modifier le spectre.
 - Après calcul de la transformée de Fourier et obtention des coefficients des sinusoides composant le signal,
 - Modification de l'argument des sinusoides
 - Transformée inverse

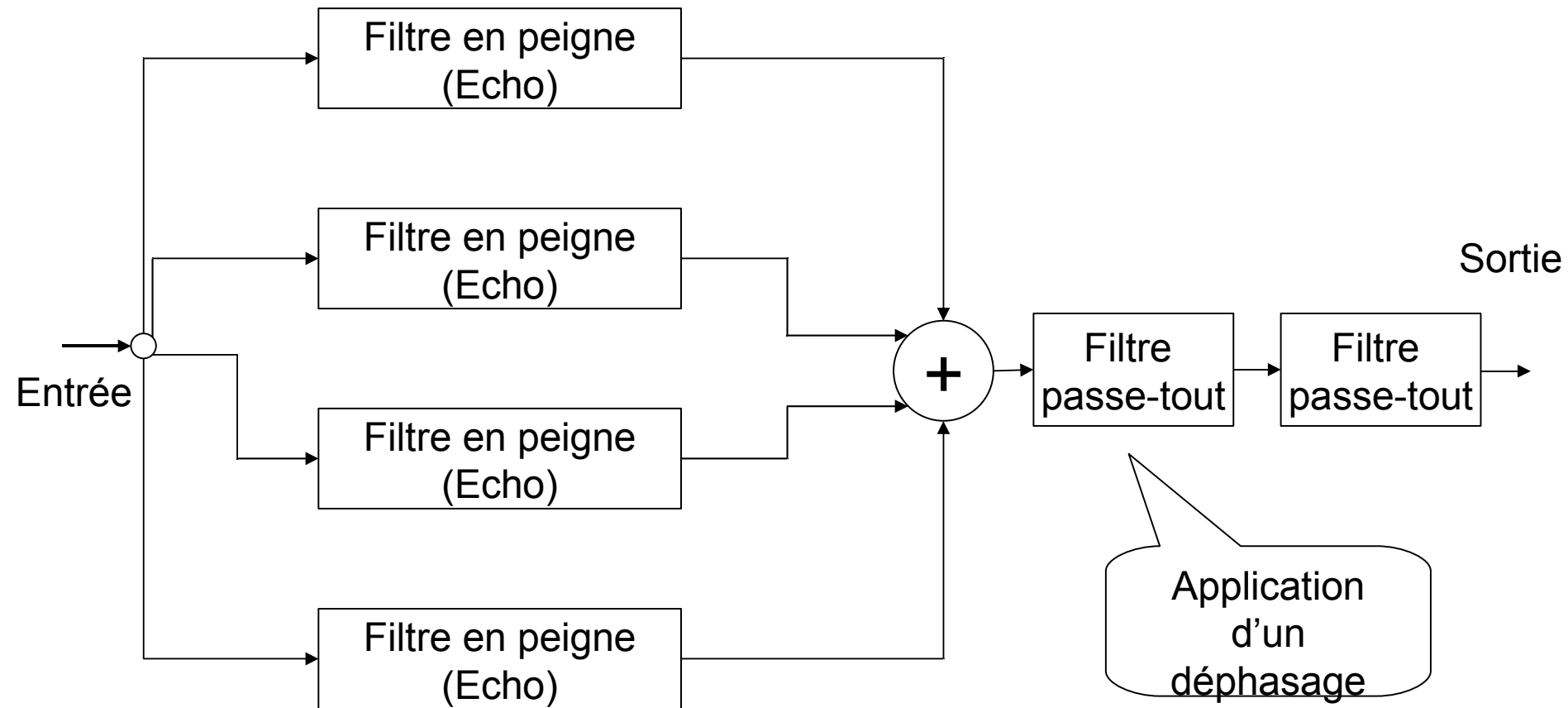


Effets de réverbération

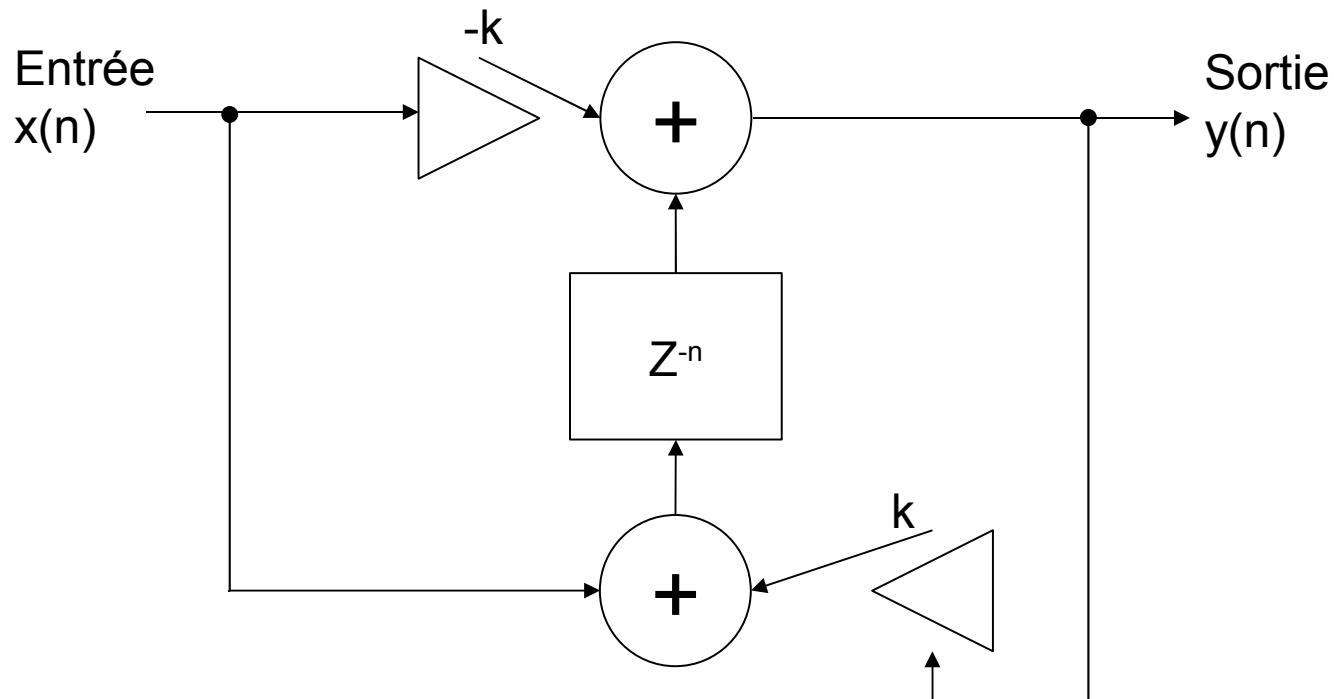
Réverbération

- L'énergie d'une onde sonore qui rencontre un obstacle se divise en trois parties :
 - Une partie absorbée par l'obstacle,
 - Une autre est transmise à l'environnement de l'obstacle,
 - La dernière est réfléchi.
- Les différentes composantes spectrales ne se comportent pas de la même manière.
- Une salle d'enregistrement est faite pour absorber les ondes (chambre anéchoïque ou chambre sourde)
- Il est donc nécessaire de rajouter artificiellement de la réverbération

Réverbération parallèle de Schroeder



Filtre passe-tout pour implémentation en virgule fixe



Autres implémentations

- Algorithme de Moorer
 - Corrige les défauts de Schroeder
 - Utilisation de réverbérateurs passe-bas
 - Atténue plus rapidement les fréquences élevées
- Réverbération à blocage
- Réverbération inversée
- ...