Traitement Numérique des Données

M1 – INF 2163

AIDN: Applications Interactives et Données Numériques

Sylvie Gibet

1

1

Objectifs du cours INF 2163 – TND Traitement numérique des données

- Thématique 1 : Données numériques
  - Programmation en Python et numpy
  - Théorie de l'échantillonnage : numériser les données
  - Structuration, exploration, visualisation (pandas)
- Thématique 2 : Traitement du signal
  - Série, transformée de Fourier discrète
  - applications au son et à l'image
- Thématique 3 : Machine learning
  - Apprentissage supervisé
  - Apprentissage non supervisé

2

2

I. Théorie de l'échantillonnage

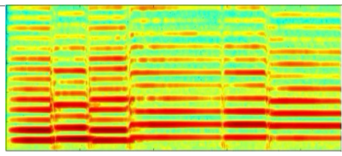
Théorie de l'échantillonnage

- Qu'est-ce qu'un signal ?
  - Signal à temps discret : les capteurs délivrent de l'information au cours du temps, soit à des instants réguliers, soit de manière irrégulière
  - Signal multidimensionnel discret : 2D, 3D, nd
  - Théorie de l'échantillonnage : comprendre les aspects propres à la numérisation du signal

- Programmes en Python
  - TP0 : se familiariser avec Python 3, numpy, les vecteurs et matrices
  - TP1 : générer un signal discret, lire un signal discret et le transformer (image)

Qu’est-ce qu’un signal ?

- Onde acoustique
  - Son : spectrogramme



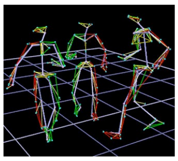
1D

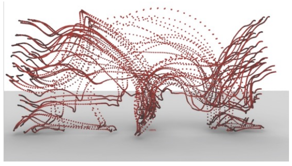
- Onde lumineuse
  - Image

2D

- Capteurs position
  - Mouvement

nd

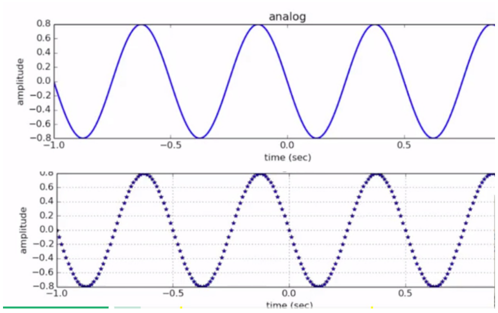


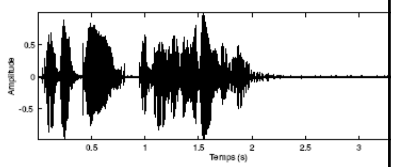


5

Signal : analogique vs. numérique

- son : pression de l’air en fonction du temps
  - Son pur : 1 seule fréquence (onde sinusoïdale)
  - Son quelconque : 1 ensemble de fréquences





6

Théorème de l'échantillonage

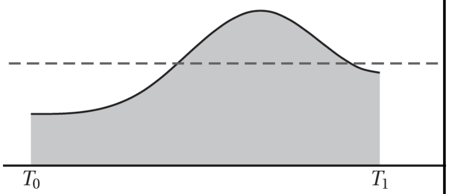
- Le **théorème de Nyquist/Shannon** : à la base du passage continu -> discret des signaux (voir plus loin)
- Echantillonnage (sampling) : observer et écrire les valeurs d'une grandeur physique à des temps discrets

7

7

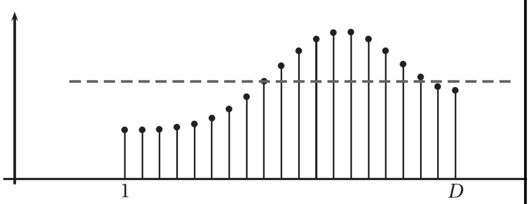
Temps discret vs temps continu

- Signal continu f(t) entre T_0 et T_1



- Signal discret échantillonné à la période T_s:

$$c_n = f(nT_s)$$

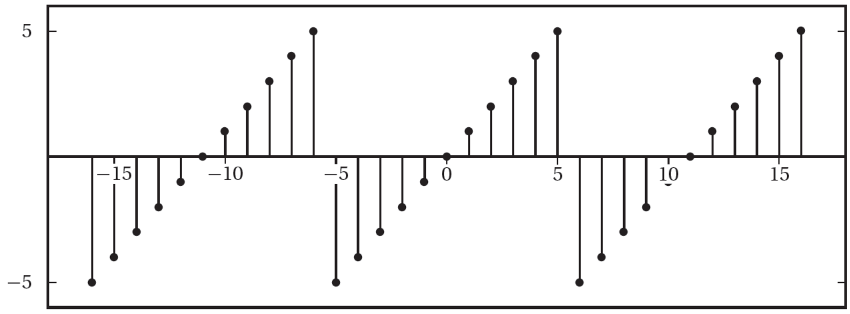


8

8

Signal à temps discret

- Définition : séquence de valeurs définie comme une fonction d'un entier d'index n, $n \in \mathbb{Z}$
  - C'est donc une collection infinie de valeurs (n négatif et positif)
  - Exemple

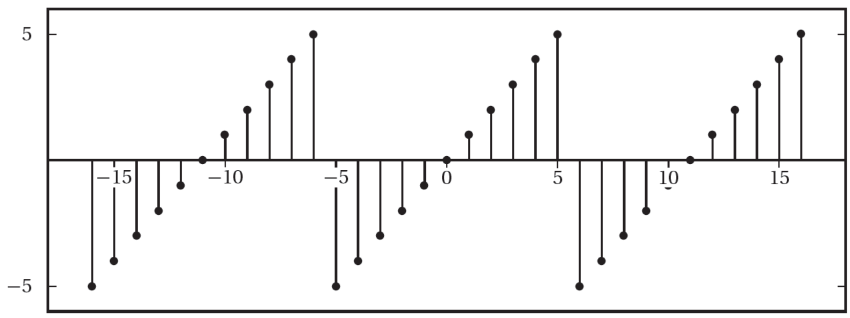


9

Signal à temps discret

- Définition : séquence de valeurs définie comme une fonction d'un entier d'index n, $n \in \mathbb{Z}$
  - C'est donc une collection infinie de valeurs (n négatif et positif)
  - Exemple

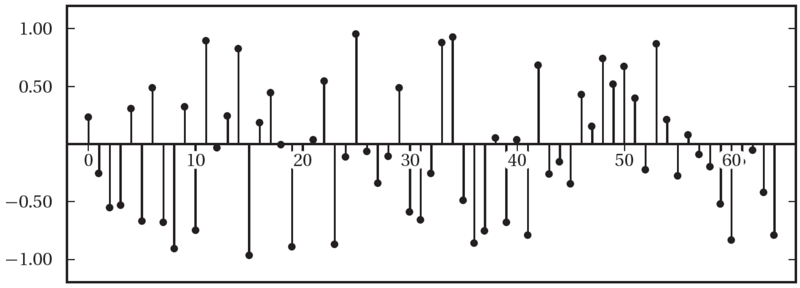
$$x[n] = ((n+5) \bmod 11) - 5$$



10

Signal à temps discret

- Autre exemple
  - x[n] = la n-ème sortie d'une source aléatoire U(-1,1)

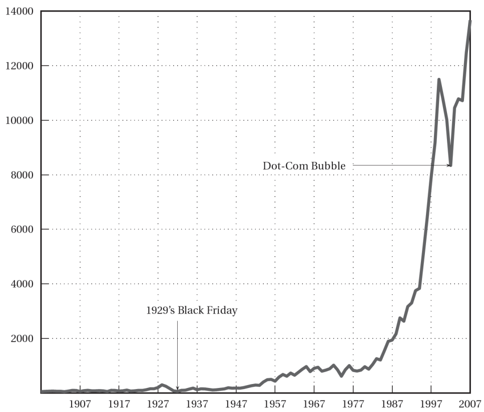


11

11

Signal à temps discret

- Autre exemple
  - x[n] = L'index de Dow-Jones moyen de l'année *n*

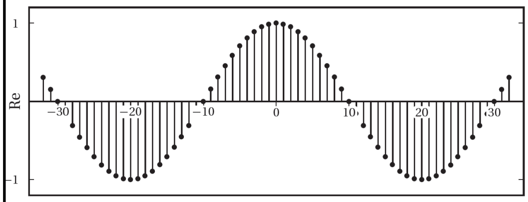


12

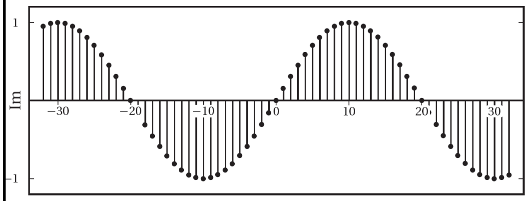
12

Signal à temps discret

- Autre exemple



$$x[n] = cos(\frac{\pi}{20}.n)$$



$$x[n] = sin(\frac{\pi}{20}.n)$$

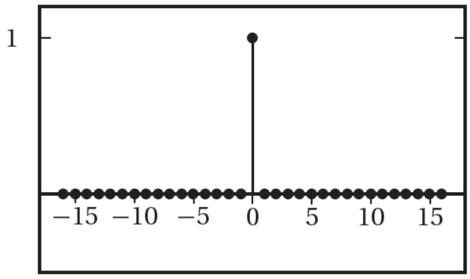
13

13

Signaux basiques

- Impulsion (aussi appelée fonction de Dirac)

$$\delta[n] = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}$$

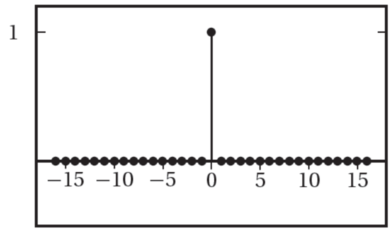


14

14

Exemple en Python

- Synthèse d'une impulsion en Python



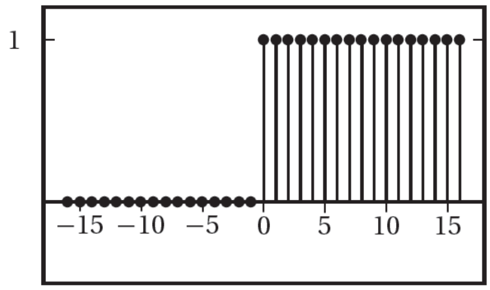
15

15

Signaux basiques

- Marche unitaire

$$u[n] = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases}$$

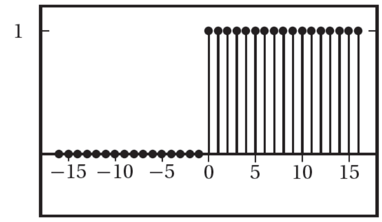


16

16

Exemple en Python

- Synthèse d'une marche unitaire



17

17

Signaux basiques

- Continu : signaux à exponentielle réelle

$$x(t) = Ce^{at} \quad avec \quad C\ et\ a\ réels$$

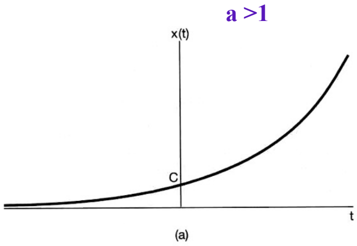
18

18

Signaux basiques

- Continu : signaux à exponentielle réelle

$$x(t) = Ce^{at} \quad avec \quad C \; et \; a \; réels$$



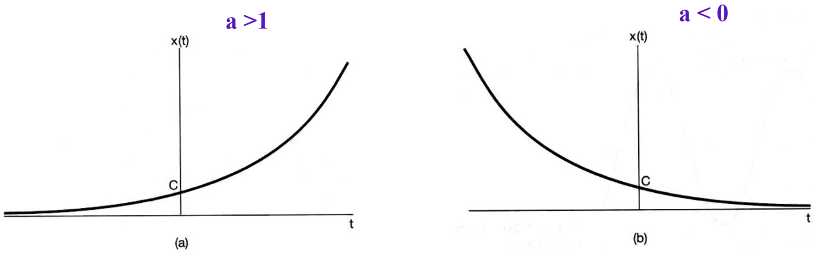
19

19

Signaux basiques

- Continu : signaux à exponentielle réelle

$$x(t) = Ce^{at} \quad avec \quad C \; et \; a \; réels$$

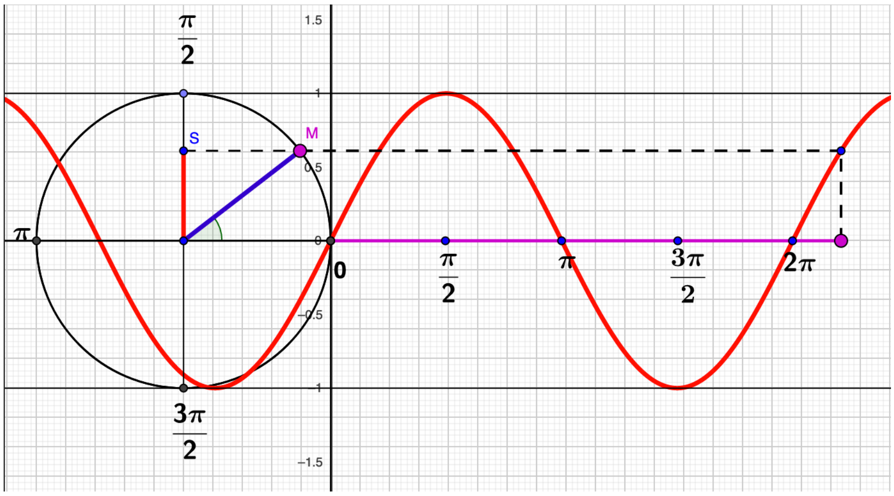


20

20

Signaux basiques

- Sinusoide (sin ou cos) : $f(t) = \sin(\omega_0.t) = \sin(2.\Pi.f_0.t)$ continu

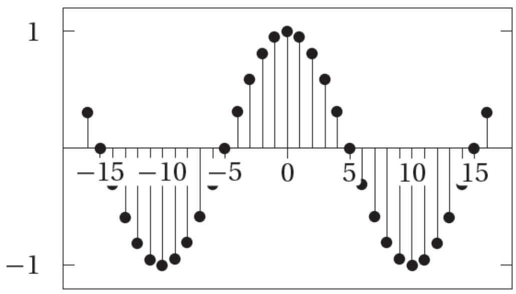


21

Signaux basiques

- Cosinus échantillonné : $f[n] = \cos(2.\Pi.f_0.(n.T_s)) = \cos((2.\Pi.(f_0/f_s).n)$
  - f_0 = fréquence du signal : $T_0 = 1/f_0$ = période du signal
  - f_s = fréquence d'échantillonnage (sampling frequency ou frame rate ou fps), $T_s = 1/f_s$
  - Ici $f_s = 20 * f_0$ (il faut que $f_s > 2*f_0$)

discret



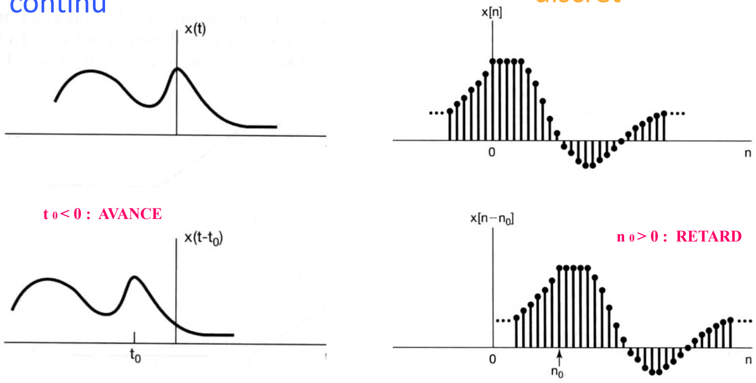
22

Principaux opérateurs

□ Décalage dans le temps : y[n] = x[n – k]

Si $n_0 > 0$, décalage à droite, si $n_0<0$, décalage à gauche

continu discret



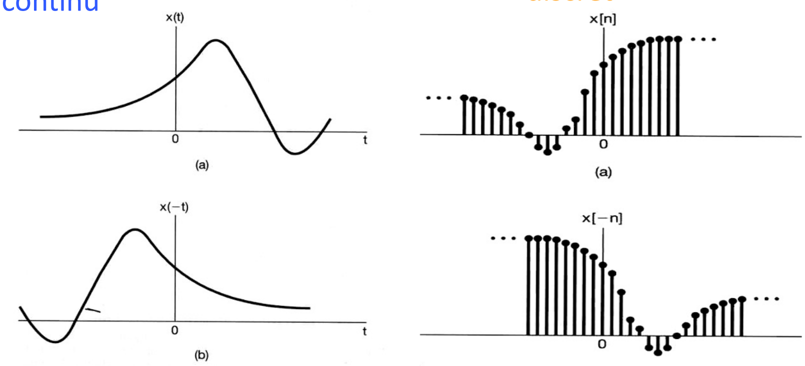
23

23

Principaux opérateurs

■ Inversion temporelle

continu discret



24

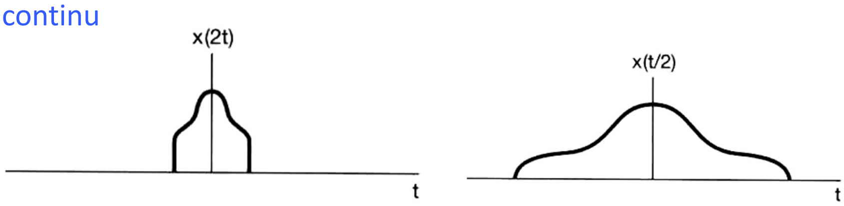
Principaux opérateurs

- Amplification ou atténuation (scaling) : y[n] = a.x[n]

a est un scalaire

- Changement d'échelle (déformation temporelle du signal)

continu



25

25

Principaux opérateurs

- Somme de 2 séquences : y[n] = x[n] + w[n]

- Scaling : $\alpha.\ (x[n] + w[n]) = \alpha.\ x[n] + \alpha.w[n]$

discret

- Retard de k échantillons :

$$F_{Retard}(x[n] + w[n]) = x[n-k] + w[n-k]$$

26

26

Principaux opérateurs

- Multiplication par une fenêtre : y[n] = x[n]. w[n]

discret

27

27

Principaux opérateurs

- Intégration

discret

$$y[n] = \sum_{k=-\infty}^{n} x[k]$$

28

28

Principaux opérateurs

- Différentiation : approximation par une méthode aux différences finies au premier ordre :

  y[n] = x[n] – x[n – 1]

29

29

Principaux opérateurs

- Énergie quadratique $E_x = \|x\|_2^2 = \sum_{n=-\infty}^{\infty} |x[n]|^2$

- Puissance: énergie moyenne sur une période

$$P_x = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|^2$$

30

30

Échantillonnage

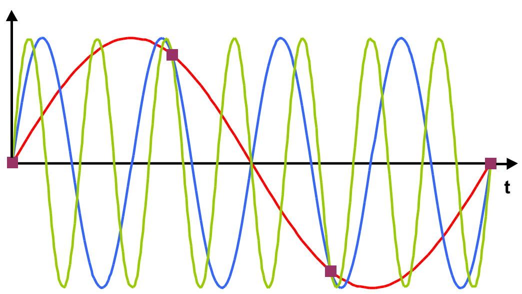
- Problème : à quelle vitesse doit on échantillonner un signal continu pour préserver l'information qu'il contient ?

- Ex : roues d'un train dans une vidéo
  - 25 images par seconde (fréquence d'échantillonnage)
  - Le train démarre : les roues tournent dans le sens des aiguilles d'une montre
  - Le train accélère : les roues tournent dans le sens inverse !

POURQUOI ?

31

Échantillonnage



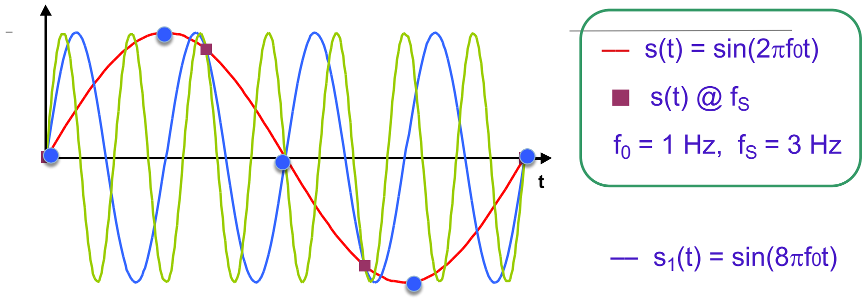
— $s(t) = \sin(2\pi f_0 t)$

■ $s(t)$ @ f_S

$f_0 = 1$ Hz, $f_S = 3$ Hz

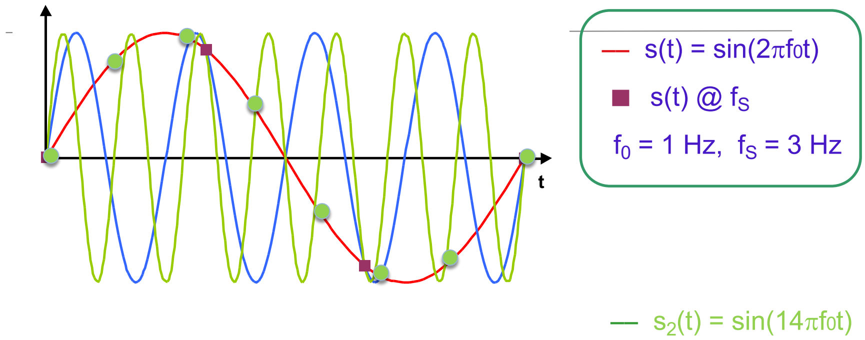
32

Échantillonnage



33

Échantillonnage



34

Théorème de l'échantillonage ou de Nyquist-Shannon

- Le **théorème de Nyquist-Shannon** énonce que pour représenter correctement un signal numérisé, la fréquence d'échantillonnage d'un signal doit être égale ou supérieure au double de la fréquence maximale contenue dans ce signal, afin de convertir ce signal d'une forme continue à une forme discrète.
- Ce théorème est à la base du passage continu -> discret des signaux

35

35

Théorème de l'échantillonnage

Theo* $\boxed{\text{Un signal } s(t) \text{ de fréquence max } f_{MAX} \text{ peut être reconstitué s'il est échantillonné avec } f_S > 2\, f_{MAX}\,.}$

* Des auteurs multiples: Whittaker(s), Nyquist, Shannon, Kotel'nikov.

Exemple

$$s(t) = \underbrace{3 \cdot \cos(50\,\pi\,t)}_{F_1} + \underbrace{10 \cdot \sin(300\,\pi\,t)}_{F_2} - \underbrace{\cos(100\pi\,t)}_{F_3}$$

Condition sur f_S?

F_1=25 Hz, F_2 = 150 Hz, F_3 = 50 Hz

36

36

Théorème de l'échantillonnage

Theo* Un signal s(t) de fréquence max f_{MAX} peut être reconstitué s'il est échantillonné avec $f_S > 2\, f_{MAX}$.

* Des auteurs multiples: Whittaker(s), Nyquist, Shannon, Kotel'nikov.

Exemple

$$s(t) = 3\cdot \underbrace{\cos(50\,\pi\, t)}_{F_1} + 10\cdot \underbrace{\sin(300\,\pi\, t)}_{F_2} - \underbrace{\cos(100\pi\, t)}_{F_3}$$

F_1=25 Hz, F_2 = 150 Hz, F_3 = 50 Hz

f_{MAX}

Condition sur f_S?

$\boxed{f_S > 300 \text{ Hz}}$

37

37

Échantillonnage : exemple

- Exemples
  - CD audio échantillonnés : ?
  - Téléphone :
  - Systèmes HiFi : piètre qualité pour les animaux domestiques (les chats ou chiens entendent des ultra-sons à une fréquence d'environ 30 000 Hz !)
  - Vision : limitée par le nombre de cellules visuelles dans la rétine

38

Échantillonnage

- Point de vue perceptuel
  - Perte d'information importante ? NON, car nos sens n'ont qu'une précision limitée et agissent comme des filtres passe-bas.
  - Exemple : on ne peut pas percevoir des sons au delà de 20 000 Hz (limite inférieure des ultra-sons) -> pour l'oreille humaine, il suffit pour reconstruire des sons d'échantillonner le signal au-delà de 40 000 Hz

39

Traitement discret vs. continu

Traitement numérique (TNI)

Avantages

- Plus flexible.
- Souvent plus facile à mettre à jour.
- Données facilement enregistrées et stockées.
- Meilleurs contrôle des performances.
- Reproductibilité.

Limitations

- Vitesse des CA/D & processeurs de T.S. : les signaux à larges bandes spectrales toujours difficiles à traiter (systèmes temps-réels).
- Effet de la longueur finie des représentations digitales.
- Obsolescence (électronique analogique vieillit moins vite).

40