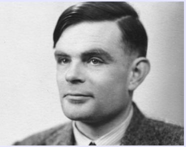
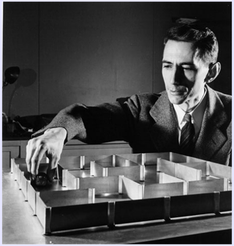
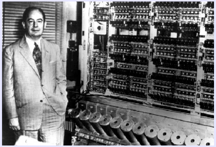
  

Information, Calcul et Communication

Module 2 : Information et Communication

Module 2: Information et Communication

Introduction

O. Lévêque – Faculté Informatique et Communications

EPFL

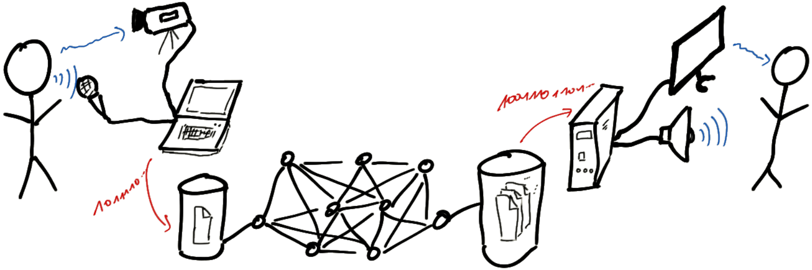
Introduction

Supposons que votre meilleur(e) ami(e) habite en Nouvelle-Zélande.

Avec un groupe d'amis, vous désirez lui jouer un sketch pour son anniversaire.

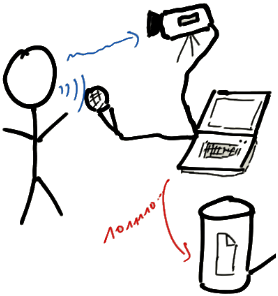
Il est désormais possible d'accomplir cette tâche en quelques minutes seulement.

Que se passe-t-il exactement lors d'une telle opération ?



Introduction

1. A l'aide de votre smartphone, vous enregistrez une vidéo amusante.
   - Ce faisant, un signal *analogique* est converti en sa représentation *numérique* au moyen d'un algorithme sophistiqué.
   - De plus, un algorithme de *correction d'erreurs* est utilisé pour stocker le fichier dans la mémoire.



Introduction

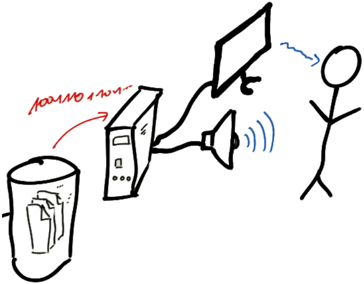
2. Vous téléchargez ensuite cette vidéo sur votre site web préféré, non sans en avoir réduit la taille au préalable, au moyen d'un algorithme de *compression*, pour que le téléchargement ne dure pas des heures.

- Lors du téléchargement, deux autres algorithmes de *correction* d'erreurs sont utilisés pour protéger la *transmission* des données a) jusqu'à votre borne wifi, b) sur internet.
- Si vous ne désirez pas que d'autres gens puissent profiter de votre sketch, un algorithme de *cryptage* est utilisé par le site web pour empêcher d'autres utilisateurs de visionner la vidéo.

Introduction

3. Puis votre ami(e) découvre cette vidéo sur son « mur » et la regarde.

- Un algorithme de *correction* d'erreurs est à nouveau utilisé ici...
- ...ainsi qu'un algorithme de *décryptage*,
- et le signal est *reconstruit* à partir des données numériques.



Introduction

En bref:

- Dans nos gestes quotidiens, nous utilisons, souvent sans nous en rendre compte, un grand nombre d'**algorithmes sophistiqués**.
- Ceci a (pour le meilleur ou pour le pire) considérablement changé notre manière de communiquer, de voyager, de voir le monde...
- Quelques **contributions fondamentales**, remontant pour certaines à plus d'une cinquantaine d'années, ont permis la réalisation de ces moyens de communication modernes.

Objectif principal de ce module:

Comprendre quelques-unes de ces contributions fondamentales.

Plan

Plan du module:

- leçons 2.1 et 2.2: échantillonnage de signaux
- leçons 2.3 et 2.4: compression de données

Et ce dont nous ne parlerons pas ou peu:

- correction d'erreur
- transmission de données / réseaux de communication
- cryptographie

Questions

Voici les questions auxquelles nous allons tenter de répondre dans ce module:

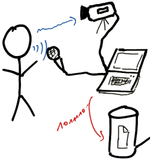
- Comment représenter / **capter** la réalité physique avec des bits ?
- Comment **restituer** cette réalité à partir de bits ?
- Comment **mesurer la quantité d'information** présente dans des données ?
- Comment **compresser** les données, c.-à-d. les *stocker en utilisant le moins d'espace* possible ?

Plan

Plan détaillé des deux leçons à venir:

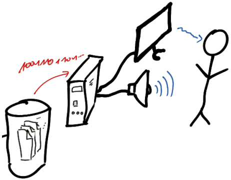
Aujourd'hui:

- signaux, fréquences, bande passante et spectre
- filtrage
- échantillonnage



La semaine prochaine:

- reconstruction
- théorème d'échantillonnage
- sous-échantillonnage



Signaux, fréquences et bande passante

Qu'est-ce qu'un signal ? C'est une fonction !

Exemples:

1. Une onde sonore ($X : \mathbb{R} \to \mathbb{R}$)
2. Une onde électromagnétique ($X : \mathbb{R} \to \mathbb{R}^3$)
3. Une photo noir-blanc ($X : \mathbb{R}^2 \to \mathbb{R}$)
4. Une photo couleur ($X : \mathbb{R}^2 \to \mathbb{R}^3$)
5. Une vidéo ($X : \mathbb{R}^3 \to \mathbb{R}^3$)

De manière générale, on peut définir un signal comme une fonction (continue, bornée) $X : \mathbb{R}^d \to \mathbb{R}^k$.

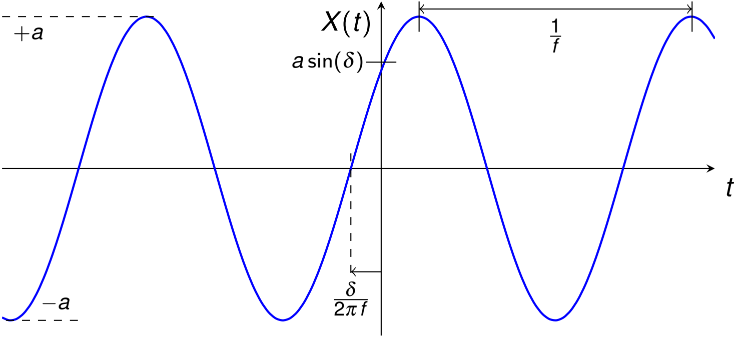
Dans le cadre de ce module, nous considérerons presque exclusivement des signaux unidimensionnels ($X : \mathbb{R} \to \mathbb{R}$), par souci de clarté et de simplification.

Exemples de signaux

Sinusoïde pure:

$$X(t) = a\,\sin(2\pi f\,t + \delta), \quad t \in \mathbb{R}$$

$a =$ amplitude, $f =$ fréquence, $\delta =$ déphasage



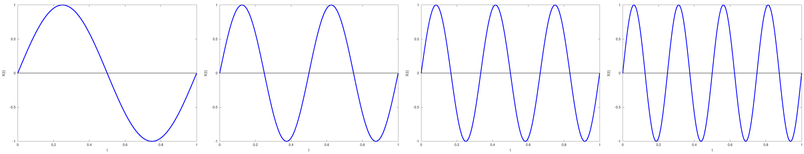
Exemples de signaux

Sinusoïde pure:

$$X(t) = a\sin(2\pi f t + \delta), \quad t \in \mathbb{R}$$

$a =$ amplitude, $f =$ fréquence, $\delta =$ déphasage

- $a = 1$, $f = 1$, $\delta = 0$: $X(t) = \sin(2\pi t)$
- $a = 1$, $f = 2$, $\delta = 0$: $X(t) = \sin(4\pi t)$
- $a = 1$, $f = 3$, $\delta = 0$: $X(t) = \sin(6\pi t)$
- $a = 1$, $f = 4$, $\delta = 0$: $X(t) = \sin(8\pi t)$



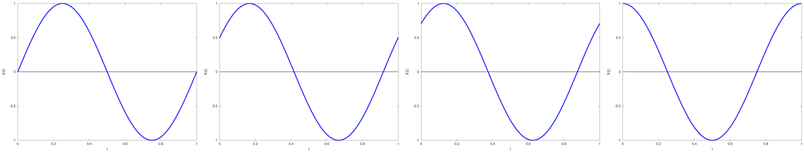
Exemples de signaux

Sinusoïde pure:

$$X(t) = a\sin(2\pi f t + \delta), \quad t \in \mathbb{R}$$

$a =$ amplitude, $f =$ fréquence, $\delta =$ déphasage

- $a = 1,\ f = 1,\ \delta = 0$: $X(t) = \sin(2\pi t)$
- $a = 1,\ f = 1,\ \delta = \pi/6$: $X(t) = \sin(2\pi t + \frac{\pi}{6})$
- $a = 1,\ f = 1,\ \delta = \pi/4$: $X(t) = \sin(2\pi t + \frac{\pi}{4})$
- $a = 1,\ f = 1,\ \delta = \pi/2$: $X(t) = \sin(2\pi t + \frac{\pi}{2}) = \cos(2\pi t)$



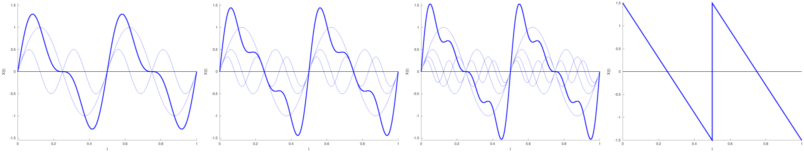
Exemples de signaux

Somme de sinusoïdes:

$$X(t) = a_1 \sin(2\pi f_1\, t + \delta_1) + \ldots + a_n \sin(2\pi f_n\, t + \delta_n), \quad t \in \mathbb{R}$$

$a_j =$ amplitudes, $f_j =$ fréquences, $\delta_j =$ déphasages

- Exemple: $a_j = 1/j$, $f_j = 2j$, $\delta_j = 0$, $n = 1, 2, 3, 4, \ldots$
- $n = 1: \; X(t) = \sin(4\pi t)$
- $n = 2: \; X(t) = \sin(4\pi t) + \frac{1}{2}\sin(8\pi t)$
- $n = 3: \; X(t) = \sin(4\pi t) + \frac{1}{2}\sin(8\pi t) + \frac{1}{3}\sin(12\pi t)$
- $n = 4: \; X(t) = \sin(4\pi t) + \frac{1}{2}\sin(8\pi t) + \frac{1}{3}\sin(12\pi t) + \frac{1}{4}\sin(16\pi t)$
- « $n = \infty$ »



Signaux en général

Affirmation: (à mettre en doute...)

« Tout signal est une somme de sinusoïdes ! »

Dans ce cours, nous ne considérerons que des signaux qui sont effectivement des sommes finies de sinusoïdes.

Fréquences: unité de mesure

La fréquence f contenue dans la sinusoïde pure $X(t) = a \sin(2\pi f t + \delta)$ s'exprime en hertz = Hz = $\frac{1}{\text{sec}}$.

Un signal dont la fréquence est de f Hz se répète toutes les $T = \frac{1}{f}$ sec.

Exemple: La note « La » à 440 Hz est une sinusoïde pure qui se répète toutes les $\frac{1}{440} = 2.2727...$ millisecondes.

Cette unité de mesure a été attribuée en l'honneur d'Heinrich Rudolf Hertz (1857-1894), à qui on doit:

- la vérification expérimentale de la théorie de Maxwell affirmant que la lumière est une onde électromagnétique;
- le premier système permettant la transmission et la réception d'ondes radio.

Fréquences: quelques ordres de grandeur

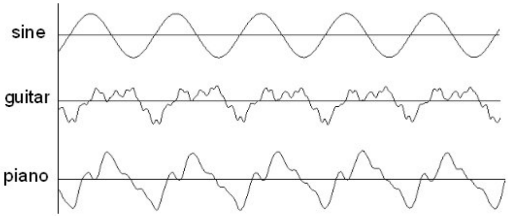
Ondes sonores:

- 20 Hz - 20 kHz: sons audibles
- 20 kHz +: ultrasons

Ondes électromagnétiques:

- 150 kHz - 3 GHz: ondes radio
- 3 GHz - 300 GHz: micro-ondes, radar
- 300 GHz - 4.3 x 10^{14} Hz: infrarouge
- 4.3 x 10^{14} Hz - 7.5 x 10^{14} Hz: lumière visible
- 7.5 x 10^{14} Hz - 3 x 10^{17} Hz: ultraviolet
- 3 x 10^{17} Hz +: rayons X, rayons γ, rayons cosmiques...

Tous les « las 440 » ne sont pas les mêmes!



EXEMPLE tiré de: `http://www.yuvalnov.org/temperament/`

- diapason (électronique)
- guitare
- piano

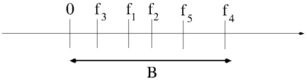
Bande passante

Revenons aux sommes de sinusoïdes:

$$X(t) = a_1 \sin(2\pi f_1\, t + \delta_1) + \ldots + a_n \sin(2\pi f_n\, t + \delta_n)$$

On définit comme suit la *bande passante* de ce signal:

$$B = f_{\max} = \max\{f_1, \ldots, f_n\}$$



Comme nous allons le voir, la bande passante joue un rôle primordial en traitement du signal.

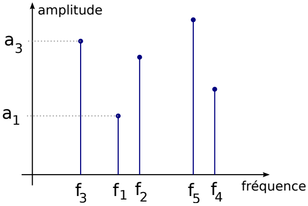
Spectre

Autre représentation graphique bien utile d'un signal : son **spectre** :

dans « l'espace des fréquences » :
axe horizontal = fréquences présentes
axe vertical = amplitude correspondante

Exemple avec une somme de sinusoïdes:

$$X(t) = a_1 \sin(2\pi f_1 t + \delta_1) + \ldots + a_n \sin(2\pi f_n t + \delta_n)$$



Filtrage d'un signal

De manière générale, lorsqu'un signal $(X(t),\, t \in \mathbb{R})$ passe par un **filtre**, il en ressort une version *déformée* $(\widehat{X}(t),\, t \in \mathbb{R})$:

$$X(t) \longrightarrow \boxed{\text{Filtre}} \longrightarrow \widehat{X}(t)$$

Pourquoi donc vouloir filtrer un signal ?
Le plus souvent, pour **supprimer** (ou du moins, atténuer) **le bruit** présent dans le signal.

Il existe bien sûr de multiples sortes de filtres.

Dans ce cours, nous allons voir une catégorie particulière de filtres:

les **filtres « passe-bas »**

Filtre passe-bas idéal

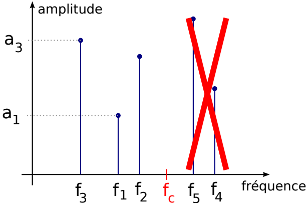
Un **filtre passe-bas idéal** est un filtre qui

supprime d'un signal toutes les fréquences supérieures à une **fréquence de coupure** f_c donnée

(c.-à-d. supprime les hautes fréquences, généralement sources de bruit).

Concrètement, si $X(t)$ est une somme de sinusoïdes, alors après le filtre, toutes les composantes de $X(t)$ dont la fréquence est plus grande que f_c ont disparu :

$$X(t) = a_1 \sin(2\pi f_1\, t + \delta_1) + \ldots \cancel{+ a_k \sin(2\pi f_k\, t + \delta_k)} + \ldots + a_n \sin(2\pi f_n\, t + \delta_n)$$



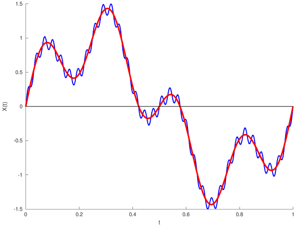
Filtre passe-bas idéal : exemple

- Considérons le signal (contenant les fréquences $f = 1, 4$ et 32 Hz):

$$X(t) = \sin(2\pi t) + \frac{1}{2}\sin(8\pi t) + \frac{1}{10}\sin(64\pi t)$$

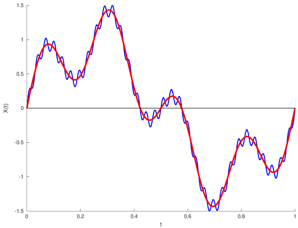
Filtre passe-bas idéal : exemple

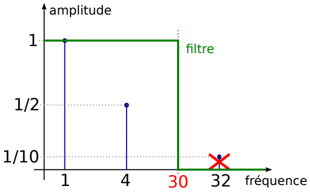
- Après passage au travers d'un filtre passe-bas avec fréquence de coupure $f_c = 30$ Hz, la composante du signal à 32 Hz disparaît, et le signal devient: $\widehat{X}(t) = \sin(2\pi t) + \frac{1}{2}\sin(8\pi t)$



Filtre passe-bas idéal : exemple

- Après passage au travers d'un filtre passe-bas avec fréquence de coupure $f_c = 30$ Hz, la composante du signal à 32 Hz disparaît, et le signal devient: $\widehat{X}(t) = \sin(2\pi t) + \frac{1}{2}\sin(8\pi t)$

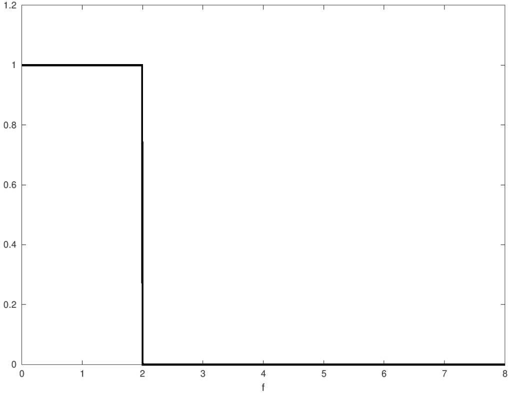




Spectre $\widehat{X}$ = Spectre $X \times$ filtre

Filtre passe-bas idéal dans l'espace des fréquences

Exemple :
un filtre passe-bas idéal avec fréquence de coupure $f_c = 2$ Hz



Filtre à moyenne mobile

Le signal $\widehat{X}(t)$ sortant à l'instant t d'un **filtre à moyenne mobile** est donné par:

$$\widehat{X}(t) = \frac{1}{T_c} \int_{t-T_c}^{t} X(s)\, ds$$

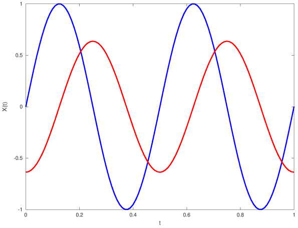
T_c est la période sur laquelle on moyenne le signal avant l'instant t.

Exemple: Qu'arrive-t-il à une sinusoïde pure qui passe par un tel filtre ? $X(t) = \sin(2\pi f\, t)$ devient:

$$\widehat{X}(t) = \frac{1}{T_c} \int_{t-T_c}^{t} \sin(2\pi f\, s)\, ds$$

$$= \frac{\cos(2\pi f\,(t - T_c)) - \cos(2\pi f\, t)}{2\pi f\, T_c}$$

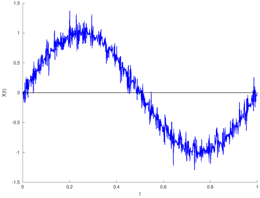
$$= \frac{\sin(\pi f\, T_c)}{\pi f\, T_c} \sin\left(2\pi f\, t - \pi f\, T_c\right)$$

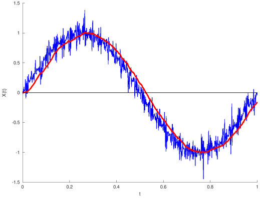


(ici, $f = 2$ Hz, $T_c = 0.25$ sec)

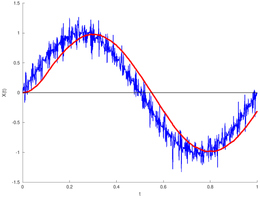
Filtre à moyenne mobile

Autre exemple: $X(t) \mapsto \widehat{X}(t) = \frac{1}{T_c}\int_{t-T_c}^{t} X(s)\,ds$





$T_c = 0.05$ sec

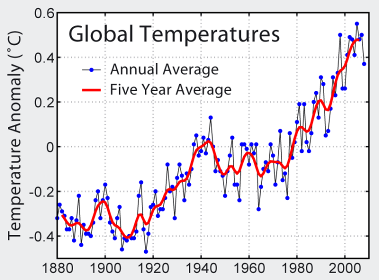


$T_c = 0.1$ sec

Plus T_c augmente, plus le signal sortant est régulier,
mais plus le déphasage est grand également.

Filtre à moyenne mobile

Autre exemple:



source: Global Warming Art

Note : sur cette figure le signal filtré a été resynchronisé avec le signal d'origine (suppression du déphasage)

Filtre à moyenne mobile

Revenons à la sinusoïde pure $X(t) = \sin(2\pi f t)$:

$$\widehat{X}(t) = \frac{1}{T_c} \int_{t-T_c}^{t} \sin(2\pi f s)\, ds = \frac{\sin(\pi f T_c)}{\pi f T_c} \sin\left(2\pi f t - \pi f T_c\right)$$

De cette expression, on déduit que

$$\max_{t \in \mathbb{R}} |\widehat{X}(t)| \leq \frac{1}{\pi f T_c}$$

On voit que si f est grand, alors le signal $\widehat{X}(t)$ est de faible amplitude.

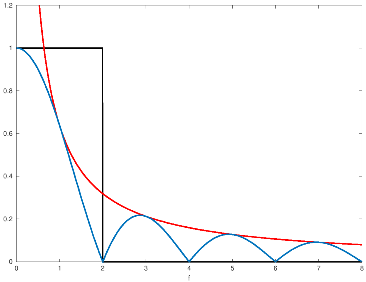
Après le passage à travers un filtre à moyenne mobile, les **hautes fréquences** d'un signal sont donc **fortement atténuées**.

Remarquer aussi que si $f T_c$ est un nombre entier, alors on a même $\widehat{X}(t) = 0$ puisqu'alors $\sin(\pi f T_c) = 0$.

Comparaison des filtres

Comparons l'atténuation des fréquences entre:

- un filtre passe-bas idéal avec fréquence de coupure $f_c = 2$ Hz et
- un filtre à moyenne mobile de période $T_c = 1/f_c = 0.5$ sec.



(sur ce dernier graphe apparaît en rouge la borne supérieure $\frac{1}{\pi f T_c}$ qu'on vient de calculer)

Filtres: conclusion

- Un filtre passe-bas sert à supprimer ou atténuer les hautes fréquences dans un signal.
- On a vu *deux* filtres passe-bas: idéal ou « à moyenne mobile »
- La semaine prochaine, nous verrons une application importante des filtres passe-bas.

Echantillonnage d'un signal

Revenons maintenant à notre première question:

Comment représenter / **capter** la réalité physique avec des bits?

Les signaux qui nous entourent sont de nature *analogique*
(ondes sonores, électromagnétiques).

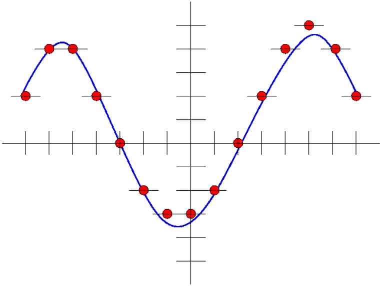
Or un ordinateur ne peut traiter que des données *numériques*.

Pour pouvoir traiter l'information contenue dans un signal $(X(t), t \in \mathbb{R})$,
il faut donc:

1. **échantillonner** le signal à des instants discrets; $X(nT_e)$
2. **quantifier** les valeurs du signal à ces instants. (virgule flottante)

Une question se pose naturellement: que perd-on du signal d'origine à travers ces deux opérations successives?

Echantillonnage d'un signal



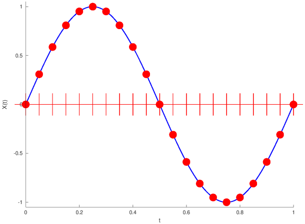
- signal échantillonné et quantifié

Échantillonnage

Pour représenter un signal dans l'ordinateur, on va n'en garder que quelques valeurs « bien choisies »

Pour simplifier, ces valeurs sont choisies de façon périodique : on parle d'**échantillonage**

Le signal $X(t)$ sera donc représenté par la suite de valeurs $X(n\,T_e)$



T_e s'appelle la « période d'échantillonnage »

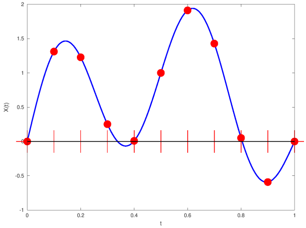
(Note : la représentation numérique des valeurs elles-mêmes (cf leçon I.4) s'appelle « quantification »)

Echantillonnage d’un signal

Nous nous concentrons ici sur la partie « échantilonnage »:

$$X(t) \longrightarrow \boxed{\text{Echantillonneur}} \longrightarrow X(nT_e)$$

signal entrant $(X(t), t \in \mathbb{R}) \quad \mapsto \quad$ signal échantilloné $(X(nT_e), n \in \mathbb{Z})$:



T_e = période d’échantillonnage, $f_e = \frac{1}{T_e}$ = fréquence d’échantillonnage

Période d'échantillonnage T_e

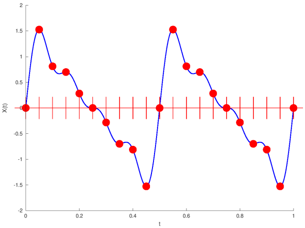
Quelle période d'échantillonnage T_e est la « bonne » ?

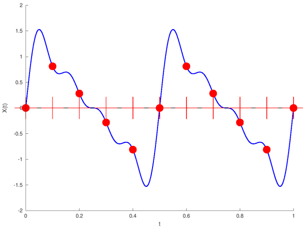
- T_e trop petite: trop d'information à traiter...
- T_e trop grande: de l'information est perdue...

T_e ne peut pas être aussi grande qu'on veut

Exemple: reprenons un signal vu précédemment :

$$X(t) = \sin(4\pi t) + \frac{1}{2}\sin(8\pi t) + \frac{1}{3}\sin(12\pi t) + \frac{1}{4}\sin(16\pi t)$$



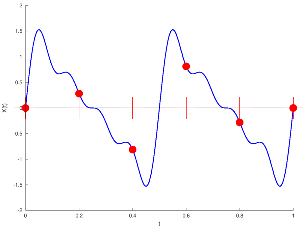


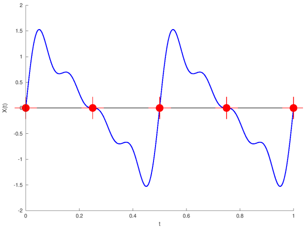
Période d'échantillonnage $T_e = 0.05$, 0.1 sec.

T_e ne peut pas être aussi grande qu'on veut

Exemple: reprenons un signal vu précédemment :

$$X(t) = \sin(4\pi t) + \frac{1}{2}\sin(8\pi t) + \frac{1}{3}\sin(12\pi t) + \frac{1}{4}\sin(16\pi t)$$



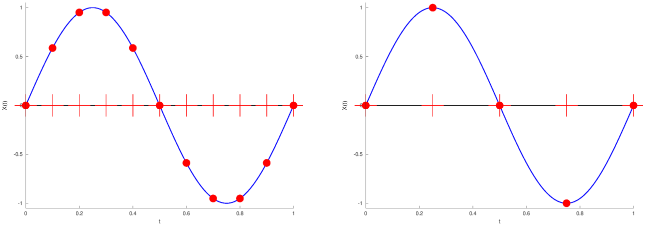


Période d'échantillonnage $T_e = 0.2,$ 0.25 sec.

Echantillonnage d'une sinusoïde pure

Autre exemple: sinusoïde pure

$$X(t) = \sin(2\pi t) \quad (f = 1\,\mathrm{Hz})$$

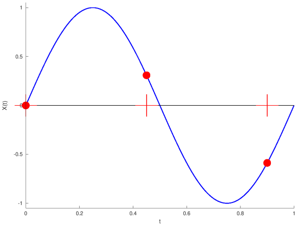


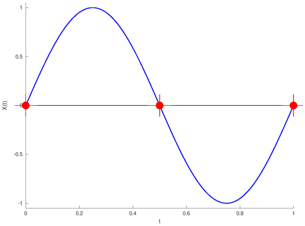
Période d'échantillonnage $T_e = 0.1,$ $\quad$ 0.25 sec.

Echantillonnage d'une sinusoïde pure

Autre exemple: sinusoïde pure

$$X(t) = \sin(2\pi t) \quad (f = 1\,\text{Hz})$$





Période d'échantillonnage $T_e = 0.45$, 0.5 sec.

Pour pouvoir reconstruire la sinusoïde à partir de l'échantillon, il est suffisant que $T_e < 0.5$ sec, autrement dit, que $f_e = \frac{1}{T_e} > 2$ Hz.

Echantillonnage d’une sinusoïde pure

De manière plus générale, on peut dire les choses suivantes:

Soit $X(t)$ une sinusoïde pure dont la fréquence est *plus petite ou égale à* f.

- Pour pouvoir reconstruire cette sinusoïde à partir de sa version échantillonnée à la fréquence f_e, il est suffisant que

$$f_e > 2f.$$

- Le *théorème d’échantillonnage* que nous verrons la semaine prochaine dit pour l’essentiel que cette condition est non seulement *suffisante* mais aussi (presque) *nécessaire*.

- Nous verrons également que ce théorème s’applique à tous les signaux, et pas seulement aux sinusoïdes.

Application

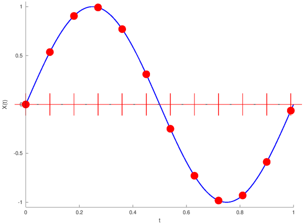
Sur un CD, le son est échantillonné à une fréquence de 44.1 kHz, car les sons au-dessus d'une fréquence de 22 kHz ne sont (en général) pas perçus par l'oreille humaine.

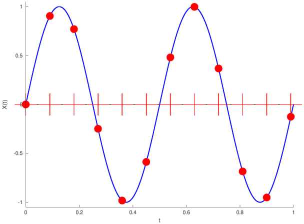
Et si $f_e < 2f$, que se passe-t-il ?

Que se passe-t-il lorsque la fréquence d'échantillonnage f_e est trop basse, c.-à-d. lorsque le signal est *sous-échantillonné* ?

Nous poursuivons ici avec l'exemple d'une sinusoïde pure: $X(t) = \sin(2\pi f\, t)$, échantillonnée avec une période $T_e = 0.09$ sec, donc $f_e = \frac{1}{0.09} = 11.11...$ Hz.

- $f = 1$ Hz
- $f = 2$ Hz



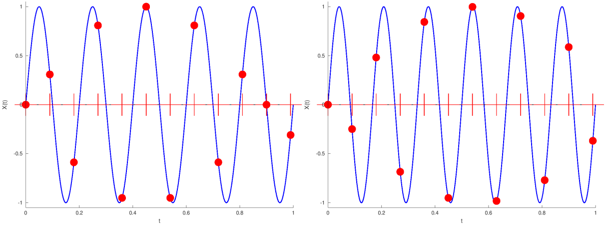


Et si $f_e < 2f$, que se passe-t-il ?

Que se passe-t-il lorsque la fréquence d'échantillonnage f_e est trop basse, c.-à-d. lorsque le signal est *sous-échantillonné* ?

Nous poursuivons ici avec l'exemple d'une sinusoïde pure: $X(t) = \sin(2\pi f\, t)$, échantillonnée avec une période $T_e = 0.09$ sec, donc $f_e = \frac{1}{0.09} = 11.11...$ Hz.

- $f = 5$ Hz
- $f = 6$ Hz

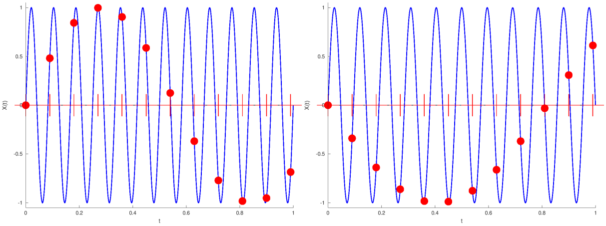


Et si $f_e < 2f$, que se passe-t-il ?

Que se passe-t-il lorsque la fréquence d'échantillonnage f_e est trop basse, c.-à-d. lorsque le signal est *sous-échantillonné* ?

Nous poursuivons ici avec l'exemple d'une sinusoïde pure: $X(t) = \sin(2\pi f t)$, échantillonnée avec une période $T_e = 0.09$ sec, donc $f_e = \frac{1}{0.09} = 11.11...$ Hz.

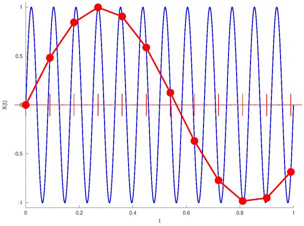
- $f = 12$ Hz
- $f = 10.5$ Hz

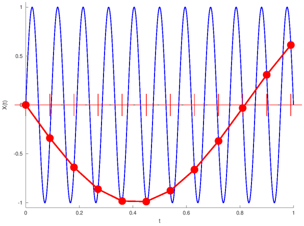


Et si $f_e < 2f$, que se passe-t-il ?

Dans les deux derniers cas, nous avons vu apparaître:

- une sinusoïde avec une fréquence plus lente;
- une autre sinusoïde, également avec une fréquence plus lente, qui part d'abord vers le bas.





Ce phénomène s'appelle l'**effet stroboscopique** et survient donc lorsqu'on *sous-échantillonne* un signal. Nous y reviendrons en détail la semaine prochaine.

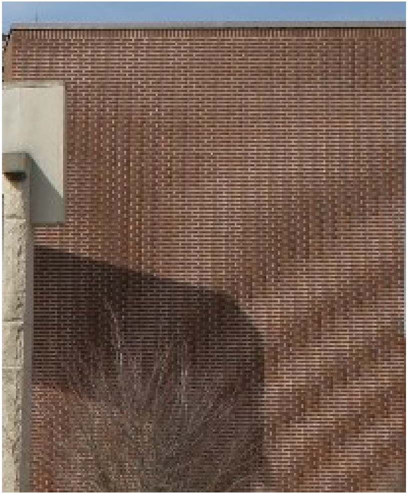
Effet stroboscopique: illustrations

EXEMPLES VIDEO :

http://www.youtube.com/watch?v=jHS9JGkEOmA
http://www.youtube.com/watch?v=r3hs8pPCQmo
http://www.youtube.com/watch?v=LVwmtwZLG88

Effet stroboscopique: illustrations

EXEMPLE VISUEL:



Conclusion temporaire

- signaux / sinusoïdes
- « tout signal est une somme de sinusoïdes ! »
- fréquence(s) présente(s) dans un signal, bande passante, spectre
- filtrage et échantillonnage
- condition suffisante pour pouvoir reconstruire le signal: $f_e > 2f$

La semaine prochaine:

- comment reconstruire un signal à partir d'un échantillon donné ?
- théorème d'échantillonnage
- sous-échantillonnage