Intelligence Artificielle & Machine Learning pour la modélisation de séries temporelles et de signaux

Séance 2 : Quelques notions de traitement statistique du signal

Laurent Oudre
laurent.oudre@ens-paris-saclay.fr

Diplôme ARIA
ENS Paris Saclay
2026-27

Limites de l'approche déterministe

- Pour étudier un signal nous avons vu que l'on pouvait l'observer dans le domaine du temps ou le domaine des fréquences
- Pour une étude plus poussée, il faudrait avoir pour chaque signal $x[n]$ ou $x(t)$ une équation...
- Afin de pouvoir les étudier de façon théorique, on construit donc des **modèles** de signaux, qui n'existent pas nécessairement dans le monde physique, mais permettent de réaliser des tâches impossibles autrement (ex : prédiction de température, reconnaissance d'une note de musique)

Exemple de la sinusoïde

$$x(t) = A\sin(2\pi f_0 t + \phi)$$

- Permet de modéliser une note de musique idéale composée d'une seule fréquence : f_0 règle la hauteur de la note
- Dépend de plusieurs paramètres : A, f_0, ϕ
- Comment construire un algorithme de traitement qui fonctionne indépendamment de ces paramètres ?
- Nécessité de prendre en compte une certaine variabilité lors de la conception du traitement

Autres arguments

- Certains phénomènes (les perturbations par exemple) sont a priori inconnus et aléatoires. On n'est pas capable de dire précisément les valeurs que ces signaux vont prendre, mais on sait par exemple que leur amplitude sera de tel ordre de grandeur, ou que *en général* ils auront des fréquences dans une certaine bande, etc...
  Comment utiliser cette information pour l'intégrer au traitement ?
- En particulier, chaque mesure implique une certaine erreur de mesure, qui est nécessairement aléatoire
  Comment la prendre en compte et la modéliser ?

Signaux déterministes vs. aléatoires

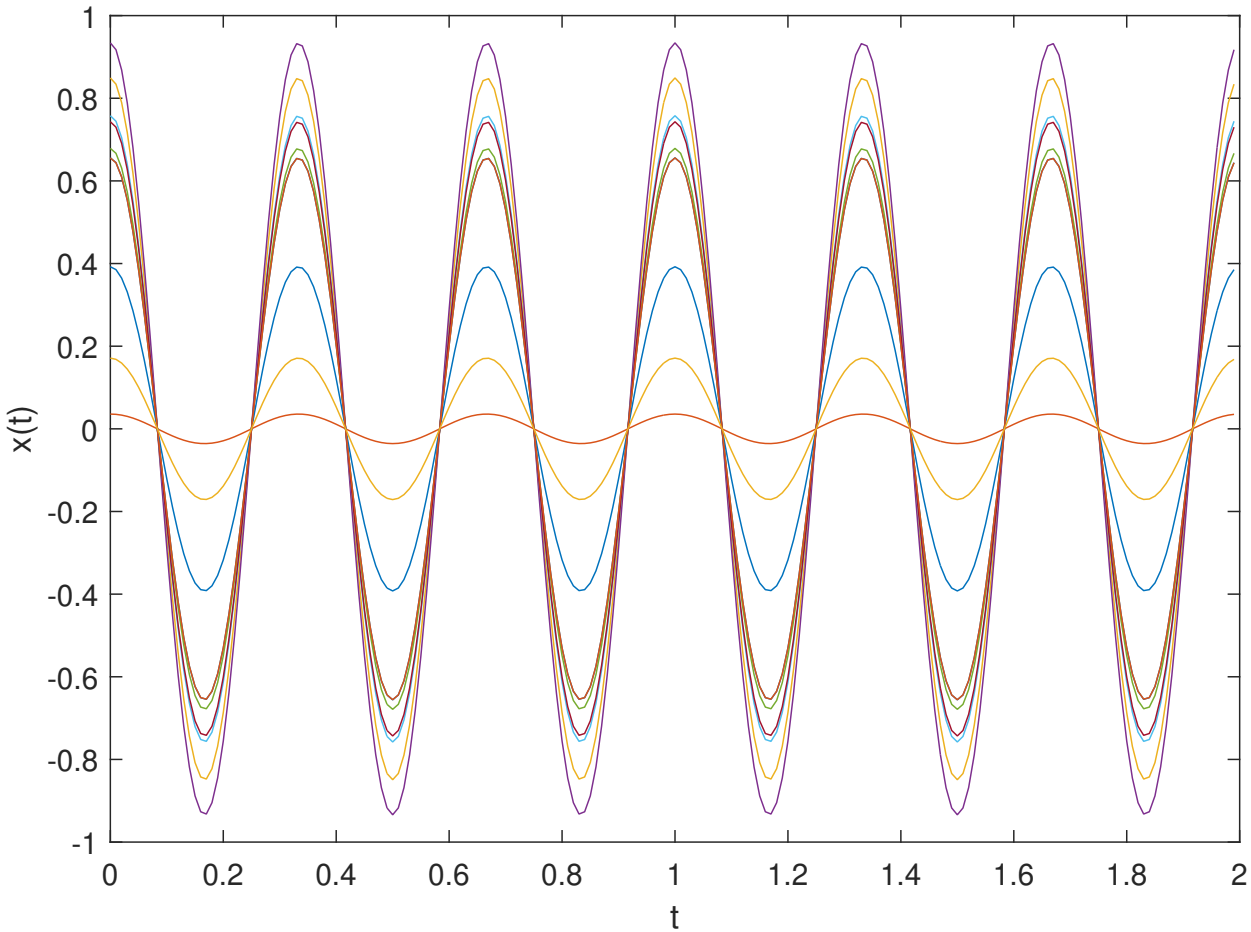
- On va donc étendre toutes les notions vues dans le premier cours à des signaux dits **aléatoires**
- Par opposition, tous les signaux vus précédemment sont dits **déterministes**
- Cette nouvelle vision intègre notamment des notions liées aux probabilités et aux statistiques, et permet de modéliser les relations entre les différents échantillons d'une série temporelle

Modélisation des signaux aléatoires

- Les signaux aléatoires ne sont pas traitables grâce aux outils vus précédemment. En effet, la valeur du signal n'est pas connue... alors comment faire ?
- Il va falloir utiliser la théorie des probabilités et des variables aléatoires pour travailler sur ces signaux
- On va considérer chaque valeur du signal $x[n]$ comme étant une **variable aléatoire** (réelle ou complexe), qui dépend donc d'événements aléatoires
- On appelle **réalisation** ou **trajectoire** d'un signal aléatoire $x[n]$, l'ensemble des valeurs prises par les v.a. $x[n]$ quand n varie.

Notion de trajectoire : exemple

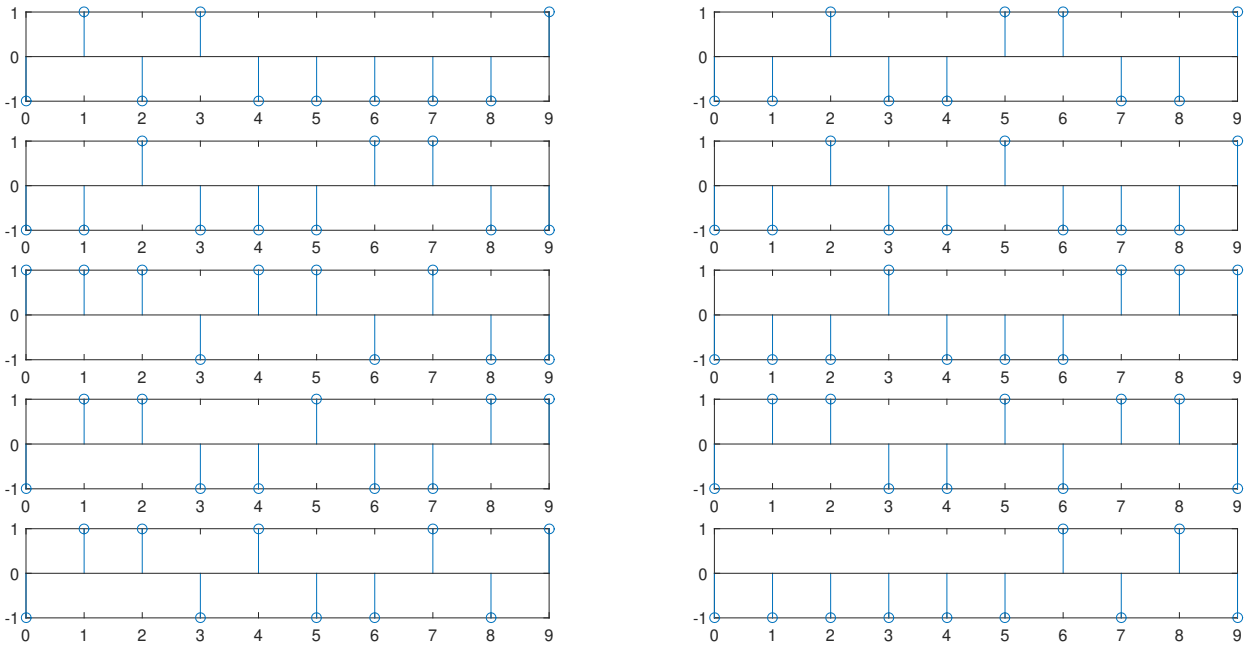
$x(t) = A\cos(2\pi f_0 t)$ avec A tiré aléatoirement de façon uniforme entre 0 et 1



10 réalisations (ou trajectoires) de ce signal aléatoire

Notion de trajectoire : exemple

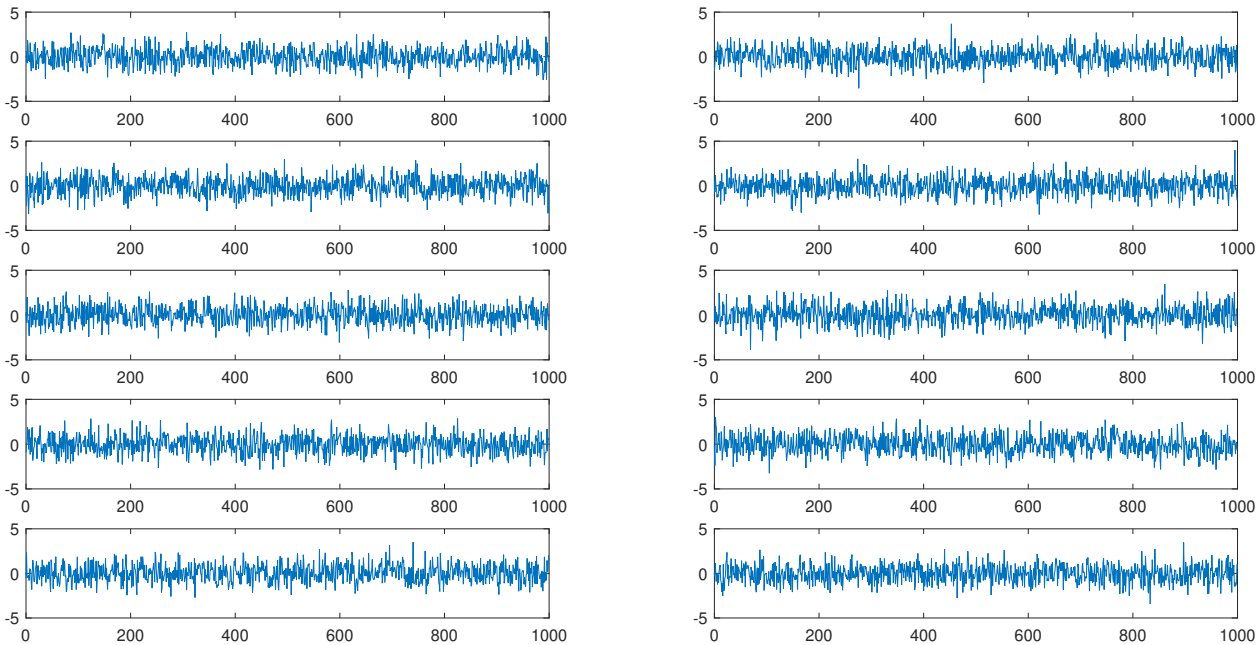
$x[n]$ où chaque échantillon est tiré aléatoirement de façon uniforme dans $\{-1, +1\}$



10 réalisations (ou trajectoires) de ce signal aléatoire

Notion de trajectoire : exemple

$x[n]$ où chaque échantillon est tiré aléatoirement selon une loi gaussienne de moyenne 0 et de variance 1



10 réalisations (ou trajectoires) de ce signal aléatoire

Programme de la séance

- Connaître les différents outils permettant l'exploration et la compréhension des signaux et séries temporelles
- Vision statistique : fonction d'autocorrélation, densité spectrale de puissance et spectrogramme

Session 2 : Introduction to Statistical Signal Processing

Plan du cours

1. Rappels de probabilités

2. Stationnarité et ergodicité

3. Fonction d'autocorrélation

4. Densité spectrale de puissance

5. Spectrogramme

Plan du cours

1. Rappels de probabilités

2. Stationnarité et ergodicité

3. Fonction d'autocorrélation

4. Densité spectrale de puissance

5. Spectrogramme

Quels outils utiliser ?

- Chaque trajectoire d'un signal aléatoire sera différente, alors comment caractériser un tel signal ?
- On va donc raisonner en termes de *moyenne*, mais pas de moyenne temporelle : il faudra utiliser la notion de **moyenne statistique**.
- On utilise donc l'espérance statistique $\mathbb{E}$ pour travailler sur ces signaux.

Variable aléatoire

- Une variable aléatoire réelle X est un réel qui dépend d'un événement aléatoire
  Exemple : tirage simultané de deux dés, et X = somme des deux valeurs
- Une variable aléatoire est associée à une densité de probabilité $p_X(x)$ telle que, pour $a < b$

$$\mathbb{P}(a \leq X \leq b) = \int_a^b p_X(x)dx$$

- On a en particulier

$$\int_{-\infty}^{+\infty} p_X(x)dx = \mathbb{P}(-\infty \leq X \leq +\infty) = 1$$

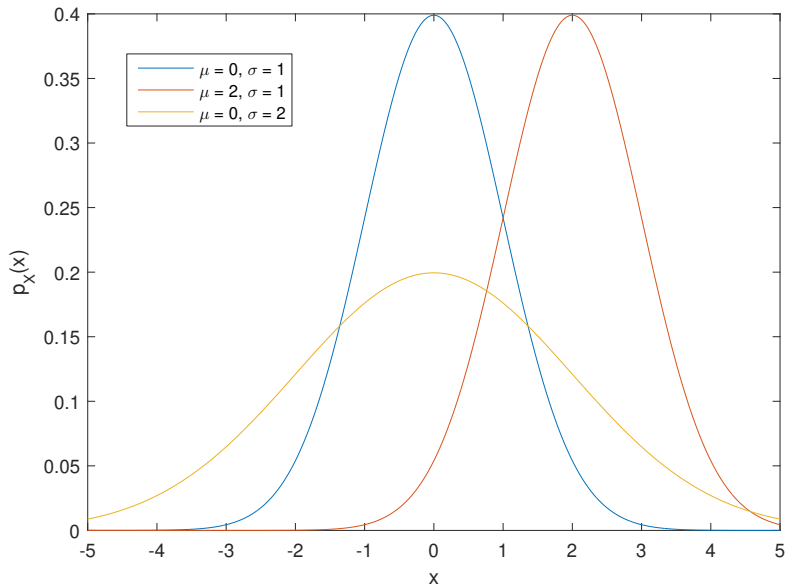
Quelques lois de probabilité

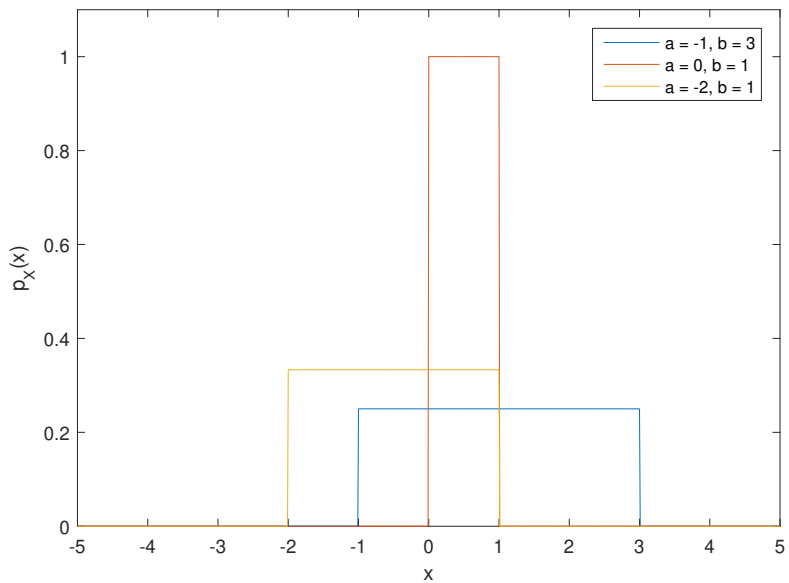
- Loi normale (ou Gaussienne) de moyenne μ et de variance σ^2

$$p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Loi uniforme sur l'intervalle $[a, b]$ avec $a < b$

$$p_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$$





Notion d'espérance

Soit X une variable aléatoire réelle, admettant une densité de probabilité $p_X(x)$. On définit l'**espérance statistique** de X et on note $\mathbb{E}[X]$ la quantité

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \, p_X(x) dx$$

- Il s'agit schématiquement de la valeur moyenne de X sur une infinité de réalisations
- L'espérance est la moyenne des valeurs de la variable aléatoire pondérées par leur probabilité d'apparition
- L'espérance est également appelée **moyenne statistique**

Propriétés de l'espérance

- L'espérance est linéaire : si X et Y sont deux v.a. et a et b deux réels, on a

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

- Si ψ est une fonction suffisamment régulière, $\psi(X)$ est également une variable aléatoire et on peut définir son espérance comme

$$\mathbb{E}[\psi(X)] = \int_{\mathbb{R}} \psi(x)\, p_X(x) dx$$

Covariance

- Etant données deux variables aléatoires X et Y on peut définir leur **covariance** $\text{cov}(X, Y)$ comme

$$\text{cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X])\,(Y - \mathbb{E}[Y])\right]$$

- Cette quantité exprime le lien qui existe entre les deux variables.
- Deux variables aléatoires sont dites non corrélées (ou décorrélées) si et seulement si

$$\text{cov}(X, Y) = 0$$

- En particulier, si X et Y sont indépendantes, alors X et Y sont décorrélées (réciproque fausse)

Variance

- La **variance** d'une variable aléatoire X, notée $\text{var}[X]$ est définie par

$$\text{var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

- On peut remarquer que

$$\text{var}[X] = \text{cov}(X, X)$$

- On a également

$$\text{var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$$

- La variance décrit à quel point la variable aléatoire a tendance à prendre des valeurs éloignées de son espérance.

Vision statistique

- On considère une série temporelle $x[n]$, supposée aléatoire
- Nous allons utiliser la théorie des probabilités et des statistiques pour modéliser le comportement de cette série temporelle
- Pour cela, on va se baser notamment sur la notion d'espérance statistique, qui dépend de la loi de probabilité du signal
- Lorsque la loi n'est pas connue, il faudra estimer les différentes quantités statistiques à partir des données, donc de façon empirique

Plan du cours

1. Rappels de probabilités

2. Stationnarité et ergodicité

3. Fonction d'autocorrélation

4. Densité spectrale de puissance

5. Spectrogramme

Stationnarité d'ordre 1

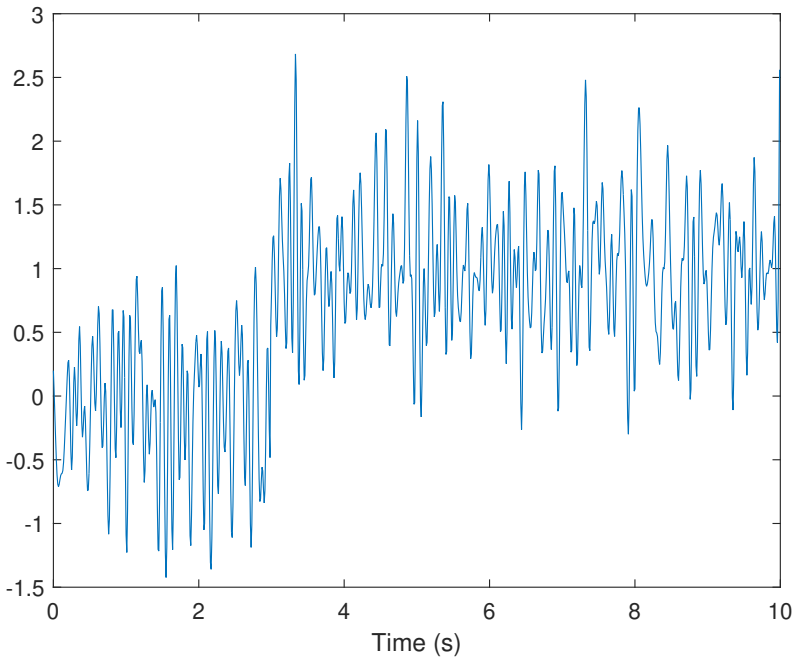
- Dans un signal aléatoire $x[n]$, chaque valeur $x[n_0]$ est une variable aléatoire, que l'on peut étudier grâce aux outils issus des probabilités
- Pour étudier la moyenne statistique d'un signal aléatoire il faut donc calculer $\mathbb{E}[x[n]]$
- En revanche, dans le cas général la quantité $\mathbb{E}[x[n]]$ n'est pas constante et varie en fonction du temps

$$\mathbb{E}[x[n]] = \mu[n]$$

- Lorsque cette quantité est constante on dit que le signal est **stationnaire d'ordre 1**

$$\forall n, \quad \mathbb{E}[x[n]] = \mu$$

Stationnarité vs. non stationnarité



- Si un signal subit un changement significatif de moyenne, on dit qu'il est non stationnaire
- Il s'agit souvent d'un vrai handicap pour le traitement de tels signaux
- On utilise dans ce cas en pré-traitement des algorithmes de segmentation ou **détection de ruptures**, qui vont permettre d'isoler les différents régimes présents dans le signal (voir Séance 8)

Ergodicité

- Dans le cas d'une série temporelle $x[n]$ dont on ne connaît pas la loi de probabilité, comment estimer la moyenne statistique $\mathbb{E}[x[n]]$?
- La première des hypothèses consiste à supposer que le signal est stationnaire, donc que la quantité $\mathbb{E}[x[n]]$ sera une constante
- Néanmoins, bien souvent, lorsque l'on mesure un signal, on n'a accès qu'à une seule réalisation (ou trajectoire). Il est donc a priori impossible de savoir quelles autres valeurs le signal aurait pu prendre...
- On se place dans ce cas sous une hypothèse d'**ergodicité** : on suppose que si l'on observe le signal suffisamment longtemps il finira par prendre toutes les valeurs qu'il peut statistiquement prendre, et dans les mêmes proportions que sur un unique échantillon

Ergodicité : On remplace l'estimation statistique par une estimation temporelle

Estimation empirique de la moyenne statistique

- Sous l'hypothèse de stationnarité d'ordre 1 et d'ergodicité, la quantité $\mathbb{E}[x[n]]$ est une constante μ, qui est également égale à la moyenne temporelle du signal
- On peut donc l'estimer de façon empirique à partir des échantillons $x[n]$ avec $n = 0 \dots N-1$, il s'agit simplement de la moyenne des échantillons

$$\hat{\mu} = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$$

Plan du cours

1. Rappels de probabilités

2. Stationnarité et ergodicité

3. Fonction d'autocorrélation

4. Densité spectrale de puissance

5. Spectrogramme

Fonction d'autocorrélation

- On peut également étudier les liens statistiques qui existent entre les variables aléatoires $x[n_1]$ et $x[n_2]$, donc entre deux instants du même signal
- Etant donné un signal aléatoire $x[n]$, on appelle fonction d'autocorrélation et on note $\gamma_x[n_1, n_2]$ la quantité

$$\gamma_x[n_1, n_2] = \mathbb{E}\left[x[n_1]x[n_2]\right]$$

- Cette fonction est indispensable pour modéliser les relations qui existent entre deux échantillons d'une série temporelle

Fonction d'autocorrélation

- Dans le cas général la quantité $\gamma_x[n_1, n_2]$ dépend des deux variables n_1 et n_2
- Lorsque cette quantité ne dépend que de l'écart $m = n_2 - n_1$ entre les instants n_1 et n_2, on dit que le signal est **stationnaire d'ordre 2**

$$\forall n_1, n_2, \quad \gamma_x[n_1, n_2] = \gamma_x[n_2 - n_1]$$

- Dans ce cas, il est courant d'écrire

$$\gamma_x[m] = \mathbb{E}\left[x[n]x[n+m]\right]$$

et la quantité ne dépend plus que d'une seule variable temporelle m

- On suppose alors qu'il y a la même corrélation entre le temps $n = 1$ et $n = 5$ qu'entre le temps $n = 2$ et $n = 6$, etc...

Stationnarité au sens large

- Afin de simplifier les calculs et les considérations, la plupart des signaux aléatoires que l'on traitera seront **stationnaires au sens large**, c'est à dire que

  1. Leur moyenne statistique ne dépend pas du temps (stationnarité d'ordre 1)

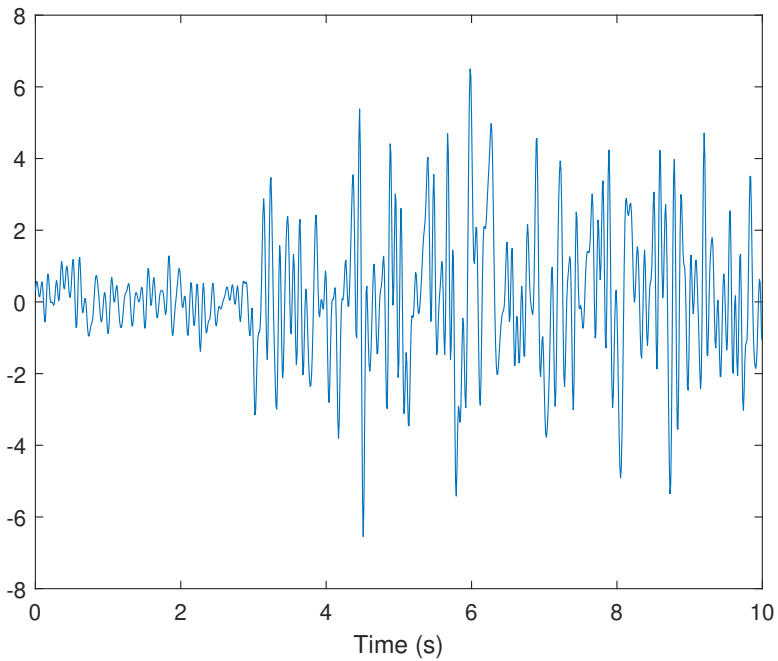
$$\forall n, \quad \mathbb{E}[x[n]] = \mu$$

  2. Leur fonction d'autocorrélation ne dépend que de l'écart $m = n_2 - n_1$ entre deux instants (stationnarité d'ordre 2)

$$\forall n_1, n_2, \quad \gamma_x[n_1, n_2] = \gamma_x[n_2 - n_1]$$

$$\gamma_x[m] = \mathbb{E}[x[n]x[n+m]]$$

Stationnarité vs. non stationnarité



- Un signal peut être stationnaire d'ordre 1 mais pas d'ordre 2
- On parle ici également de signal non stationnaire : plus généralement ce type de signaux est constitué de plusieurs régimes qui correspondent chacun à un modèle statistique différent
- Ce modèle peut affecter la valeur des échantillons, mais également les relations statistiques qui existent entre eux
- L'hypothèse de stationnarité au sens large est fondamentale : lorsqu'un signal ne vérifie pas cette hypothèse, on le segmente en trames avant de le traiter

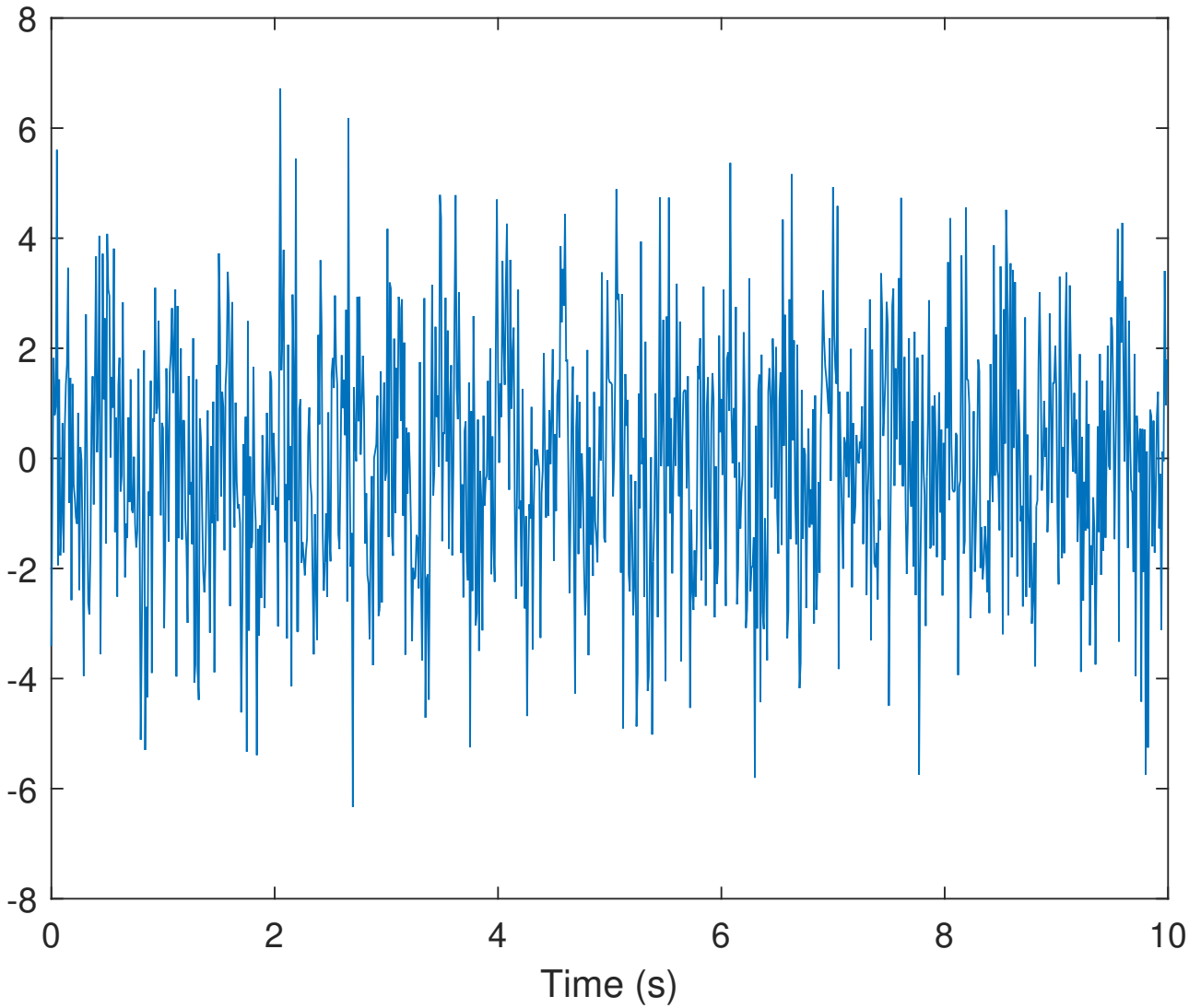
Estimation empirique de la fonction d'autocorrélation

- Sous l'hypothèse de stationnarité au sens large et d'ergodicité, la quantité $\gamma_x[m]$ est estimée de façon temporelle en étudiant les liens entre tous les échantillons $x[n]$ et $x[n+m]$
- On peut l'estimer de façon empirique à partir des échantillons $x[n]$ avec $n = 0 \dots N-1$. Deux estimateurs empiriques existent :

$$\hat{\gamma}_x^{\text{biais}}[m] = \frac{1}{N} \sum_{n=0}^{N-1} x[n]x[n+m] \quad \text{avec } x[n] = 0 \text{ si } n \notin [\![0, N-1]\!]$$

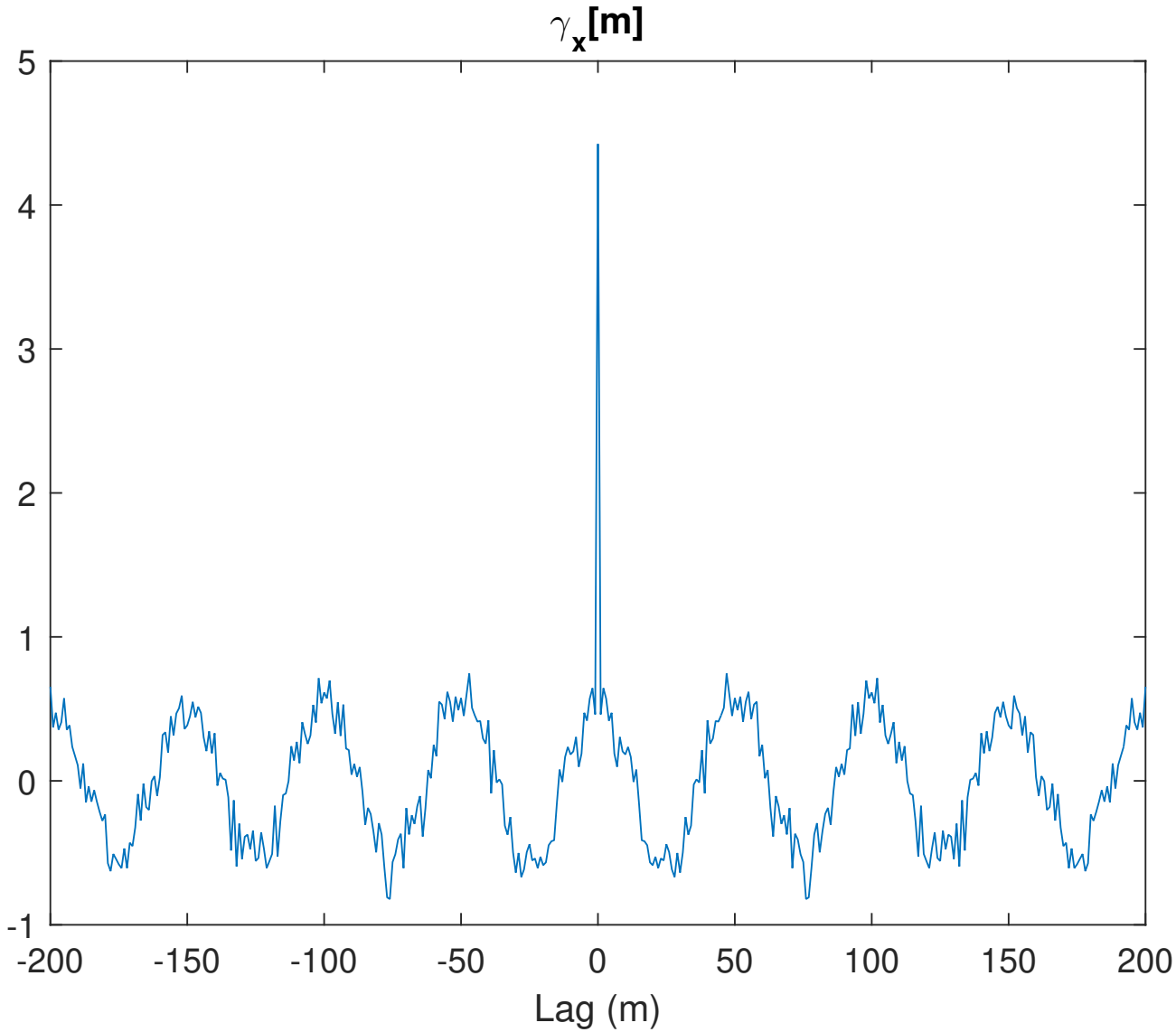
$$\hat{\gamma}_x^{\text{non biais}}[m] = \frac{1}{N-|m|} \sum_{n=0}^{N-1} x[n]x[n+m] \quad \text{avec } x[n] = 0 \text{ si } n \notin [\![0, N-1]\!]$$

Exemple d'utilisation



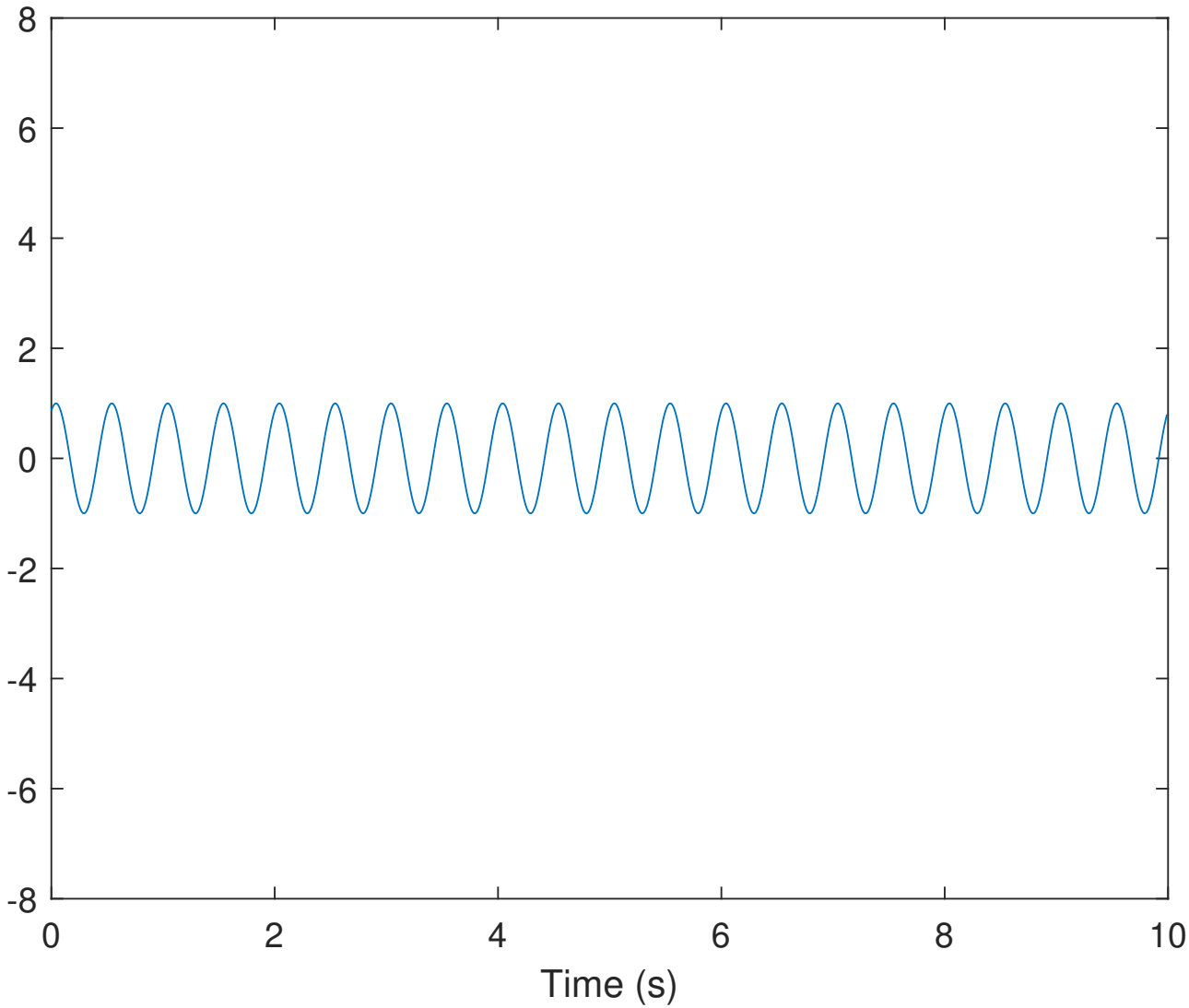
Signal échantillonné à 100 Hz

Exemple d'utilisation



Présence de pics pour des lags multiples de 50

Exemple d'utilisation



Un signal périodique de période $50 \times \frac{1}{100} = 0.5$ s se cachait !

Variance d'un signal

- La variance d'un signal peut s'exprimer par

$$\sigma^2 = \mathbb{E}\left[(x[n] - \mathbb{E}[x[n]])^2\right]$$

- Dans le cas où le signal est stationnaire au sens large et a une moyenne statistique nulle on a donc

$$\sigma^2 = \gamma_x[0]$$

Cas de deux signaux

- On peut étendre cette définition à deux signaux aléatoires $x[n]$ et $y[n]$ en définissant une fonction d'intercorrélation

$$\gamma_{xy}[n_1, n_2] = \mathbb{E}\left[x[n_1]y[n_2]\right]$$

- On dira que deux signaux sont **décorrélés** si et seulement si

$$\forall n_1, n_2, \qquad \mathrm{cov}\left(x[n_1], y[n_2]\right) = 0$$

- Si les signaux $x[n]$ et $y[n]$ sont à moyenne statistique nulle et décorrélés, alors leur fonction d'intercorrélation est toujours nulle. En effet, on a dans ce cas

$$\gamma_{xy}[n_1, n_2] = \mathrm{cov}\left(x[n_1], y[n_2]\right) = 0$$

Plan du cours

1. Rappels de probabilités

2. Stationnarité et ergodicité

3. Fonction d'autocorrélation

4. Densité spectrale de puissance

5. Spectrogramme

Transformée de Fourier ?

- Nous avons vu que pour les signaux déterministes, il était important de les étudier dans le domaine fréquentiel pour mieux comprendre leurs propriétés
- Comment faire pour un signal aléatoire ? Par définition, à chaque nouveau signal que l'on génère, on aura un signal différent et on aura une transformée de Fourier différente !
- On va définir par analogie un outil similaire, appelé **densité spectrale de puissance (DSP)**, permettant d'observer le contenu fréquentiel d'un signal aléatoire *en moyenne*

Densité spectrale de puissance

- Soit $x[n]$ un signal aléatoire stationnaire au sens large, ayant une fonction d'autocorrélation $\gamma_x[m]$. On appelle densité spectrale de puissance (DSP) et on note $\Gamma_x[k]$ la quantité

$$\Gamma_x[k] = \mathcal{TFD}\{\gamma_x[m]\}$$

- $\Gamma_x[k]$ donne la répartition de la puissance selon les fréquences $f[k] = k\frac{F_e}{N}$

Estimation empirique

- La véritable fonction d'autocorrélation n'étant pas connue, il faut estimer la DSP de façon empirique
- Deux principaux estimateurs existent :

$$\hat{\Gamma}_x^{corre}[k] = \mathcal{TFD}\left\{\hat{\gamma}_x[m]\right\}$$ corrélogramme

$$\hat{\Gamma}_x^{perio}[k] = \frac{1}{N}|X[k]|^2$$ périodogramme

- Selon les définitions prises (Matlab, Python...), les constantes multiplicatives peuvent différer
- Observer le spectre (module au carré de la TFD) revient donc à étudier la répartition de la puissance en fonction de la fréquence

Plan du cours

1. Rappels de probabilités

2. Stationnarité et ergodicité

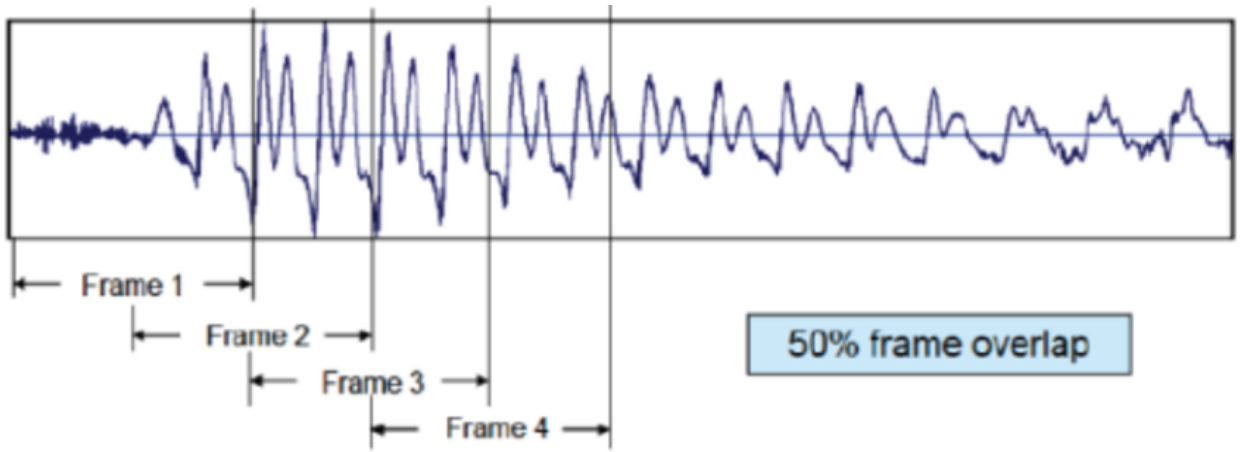
3. Fonction d'autocorrélation

4. Densité spectrale de puissance

5. Spectrogramme

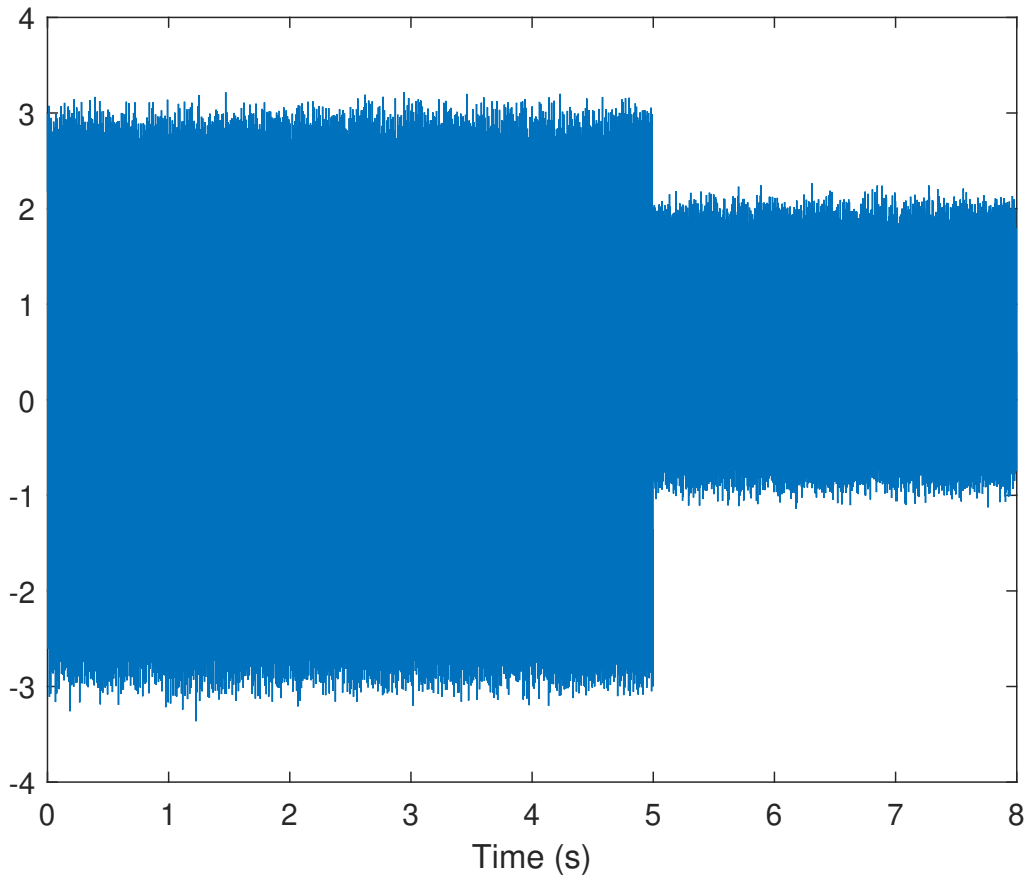
Cas des signaux non stationnaires

- Pour un signal non stationnaire, la fonction d'autocorrélation $\gamma_x[n_1, n_2]$ dépend des deux instants et ne se réduit pas à une fonction $\gamma_x[m]$ du seul écart : il est donc impossible de définir une DSP
- Dans ce cas on suppose qu'à une certaine échelle temporelle (souvent petite), le signal est stationnaire et on calcule la DSP sur cette petite fenêtre
- On fait ensuite glisser la fenêtre le long du signal (parfois avec un recouvrement) et pour chaque position de la fenêtre on recalcule la DSP



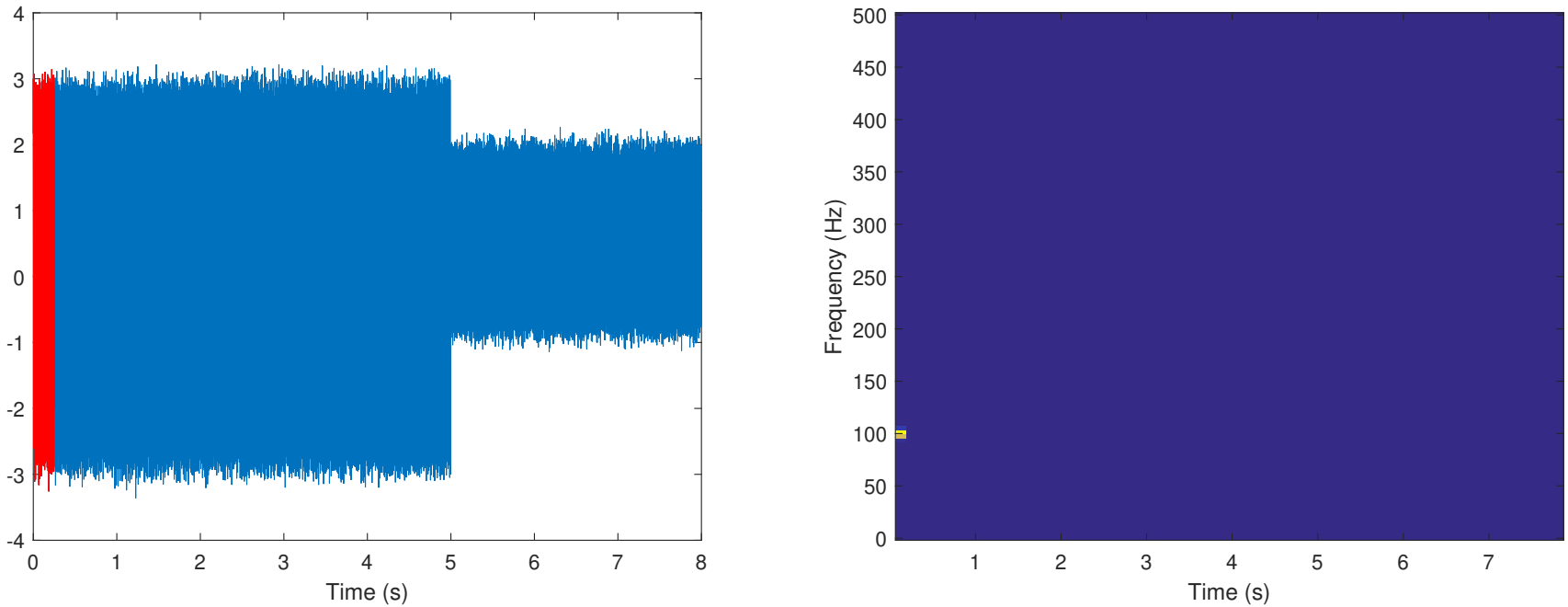
- On obtient une représentation temps-fréquence, appelée le **spectrogramme**

Exemple de spectrogramme



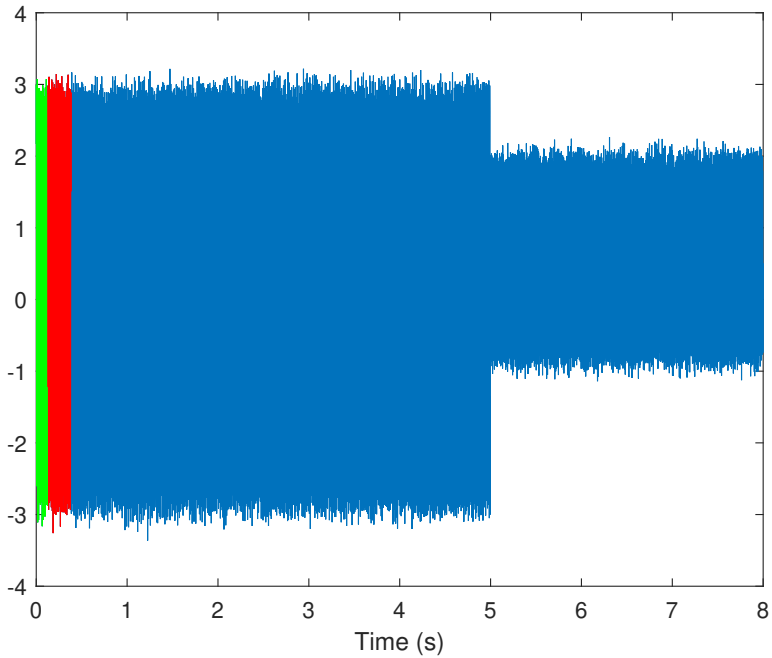
Signal non stationnaire (changement de fréquence fondamentale, d'amplitude et de valeur moyenne), échantillonné à 1000 Hz

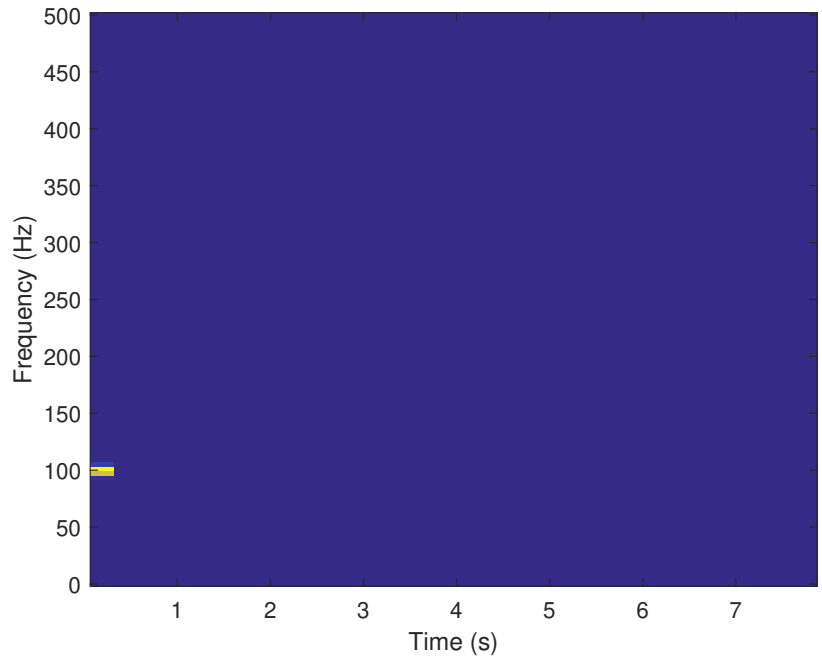
Exemple de spectrogramme



Trame de $N_w = 256$ échantillons, et calcul de la DSP

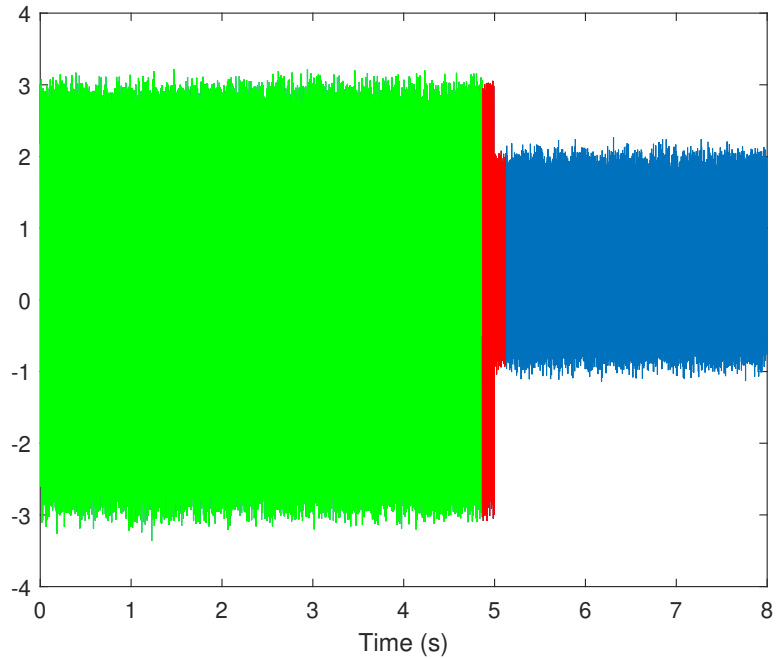
Exemple de spectrogramme

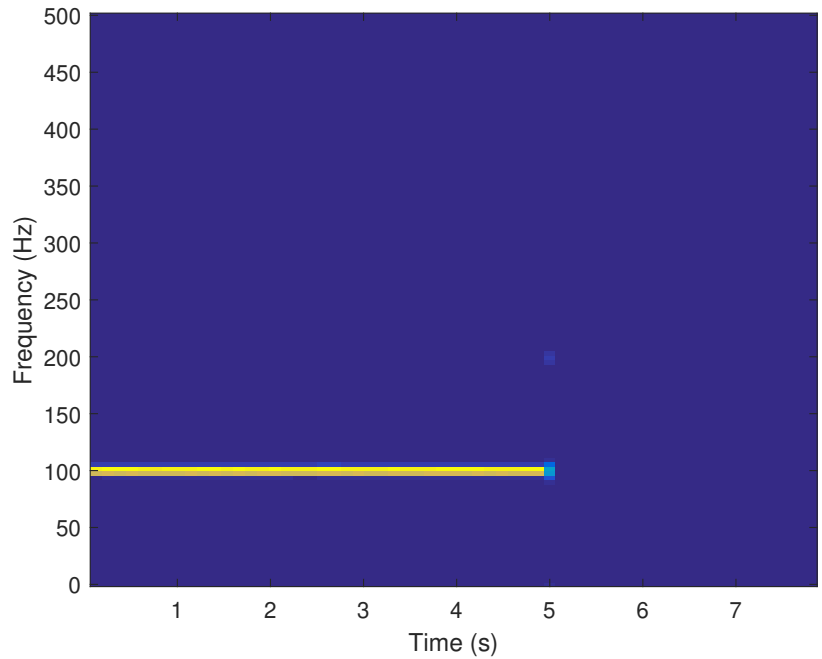




Deuxième trame $N_w = 256$ échantillons avec un recouvrement (overlap) de 50% avec la trame précédente, et calcul de la DSP

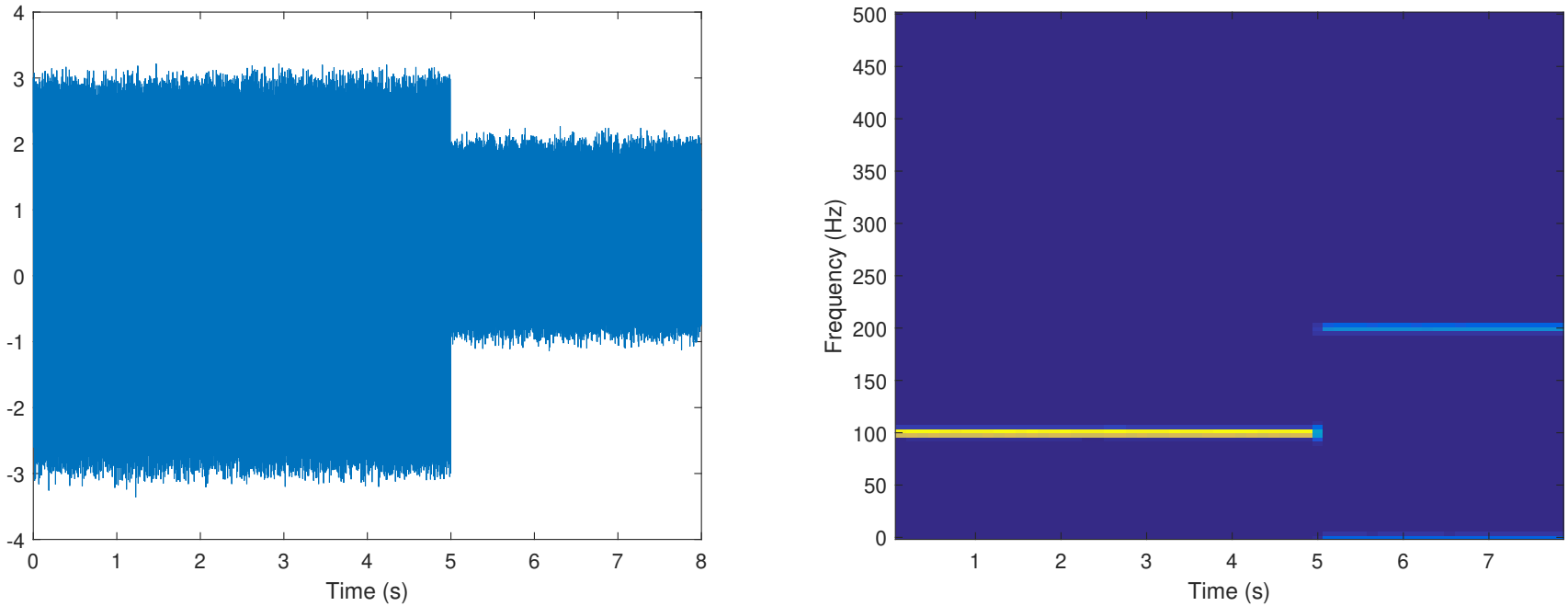
Exemple de spectrogramme





A l'instant de la discontinuité, l'hypothèse de stationnarité n'est pas vérifiée

Exemple de spectrogramme

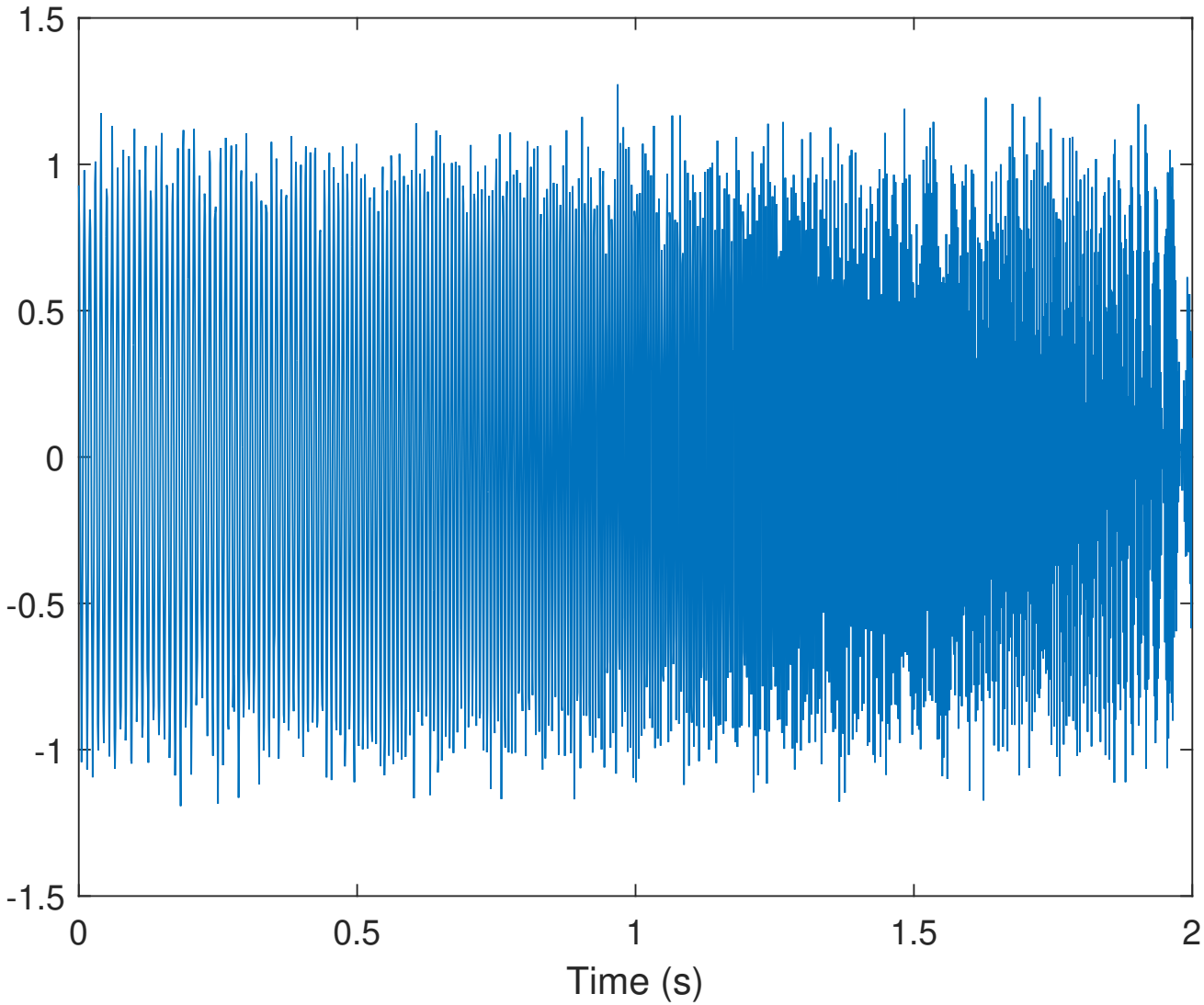


Représentation temps-fréquence du signal

DSP vs. spectrogramme

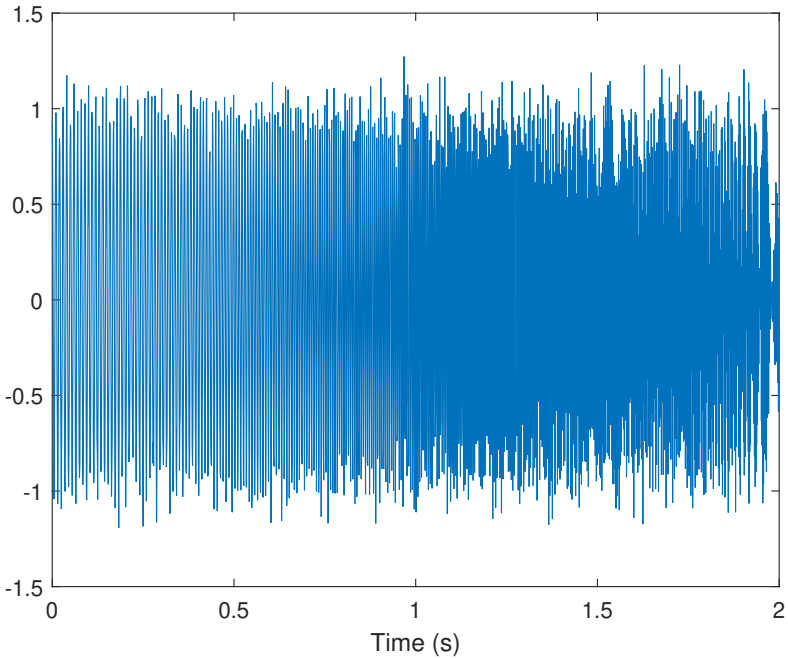
- L'utilisation de la DSP ne doit se faire que si l'on est assuré que le signal est bien stationnaire
- Il suffit parfois d'une moindre discontinuité pour fausser totalement l'interprétation du spectre
- Il est donc toujours plus prudent de commencer par observer le spectrogramme pour repérer les éventuels changements

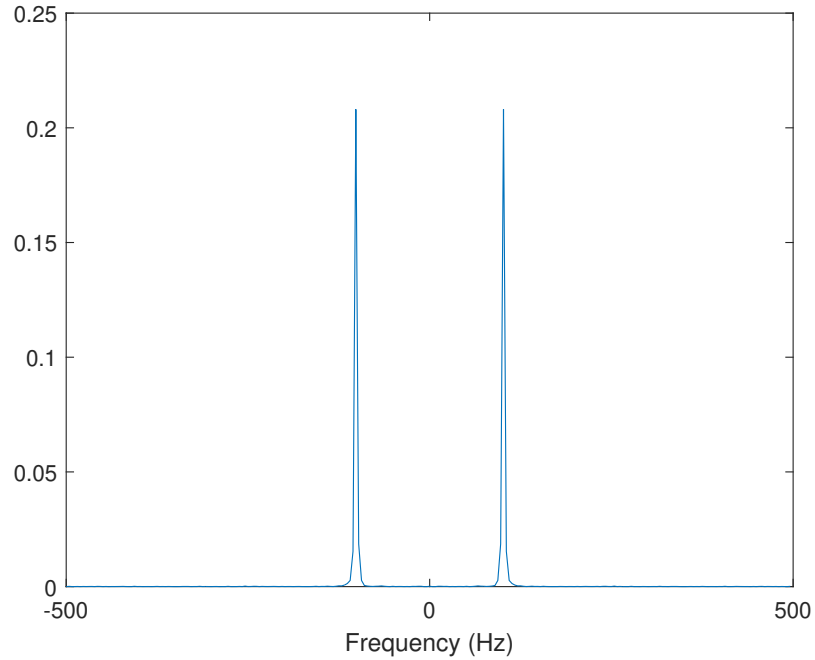
DSP vs. spectrogramme



Signal qui semble correspondre à un phénomène périodique (sinusoïde)

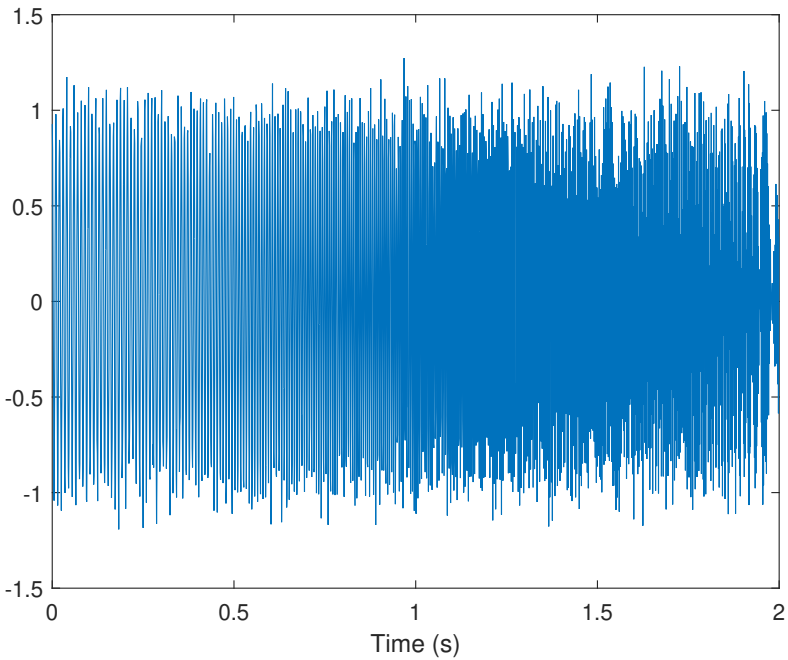
DSP vs. spectrogramme

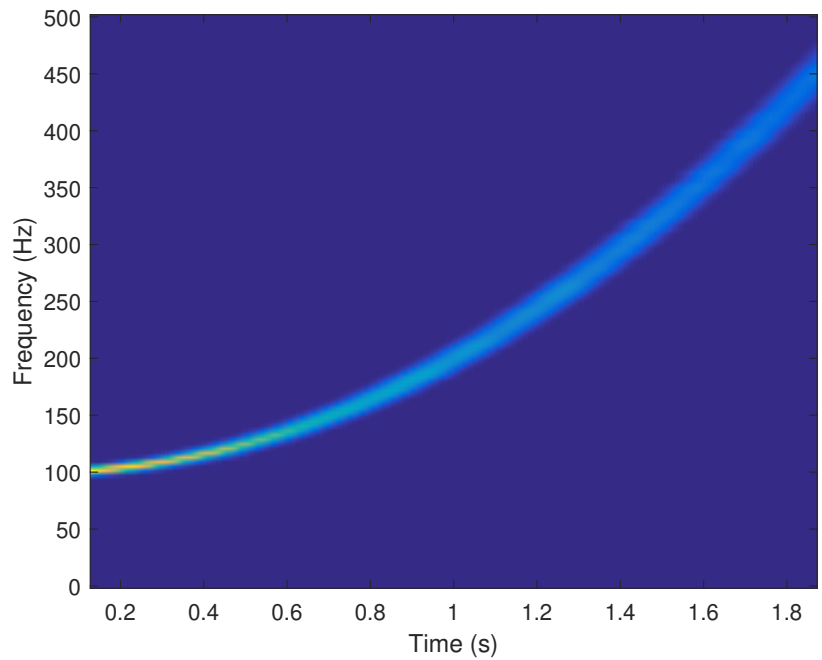




La DSP suggère un phénomène sinusoïdal autour de la fréquence 100 Hz

DSP vs. spectrogramme





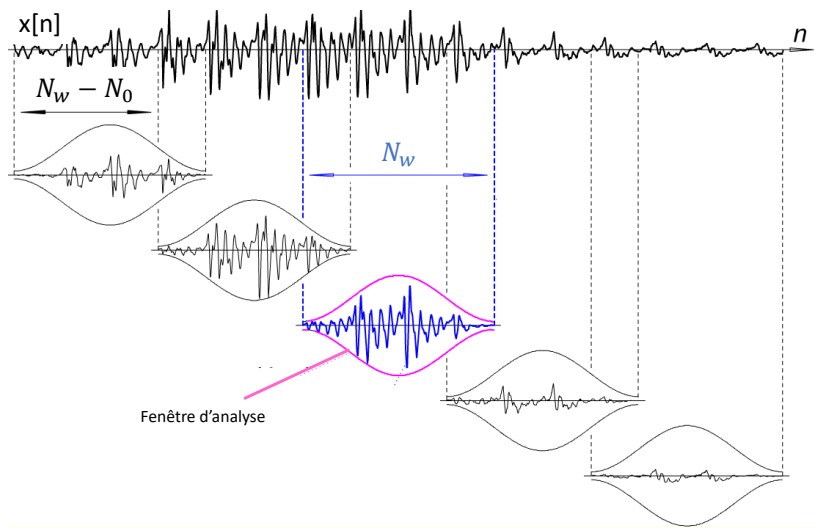
En réalité la fréquence fondamentale n'arrêtait pas de changer, et variait entre 100 et 500 Hz!!

Paramètres du spectrogramme

Le spectrogramme dépend d'un certain nombre de paramètres

- N_w : la durée d'une trame (en échantillons)
  On la prend souvent sous la forme d'une puissance de 2 et elle donne également le nombre de fréquences qui seront évaluées. Elle doit donc être choisie petite pour être sur que le signal est stationnaire mais suffisamment grande pour avoir une bonne résolution en fréquence...
- N_o : le recouvrement entre deux trames (en échantillons)
  On choisit en général un overlap de 50% ou 75% de la durée de la trame. Si on veut une bonne résolution en temps, on peut prendre $N_o = N_w - 1$

On peut également, avant de prendre la DSP de la trame, la multiplier par une fenêtre d'analyse qui permettra de se focaliser sur les échantillons au centre de la trame : fenêtres d'Hanning, Hamming, Blackman...



Références

- S.M. Kay. Fundamentals of statistical signal processing. Prentice Hall PTR, 1993.
- R.M. Gray, L.D. Davisson. An introduction to statistical signal processing. Cambridge University Press (2004).
- M. Barret. Traitement statistique du signal, Supélec https://hal.archives-ouvertes.fr/hal-00396998/
- S. Mammar. Traitement statistique du signal, Université d'Evry-Val d'Essonne http://aramis.iup.univ-evry.fr:8080/~smam/polycopies/ts71/ts31.pdf