Théorie du signal

Introduction aux signaux aléatoires - Notion de bruit

Laurent Oudre
laurent.oudre@univ-paris13.fr

Université Paris 13, Institut Galilée
Ecole d'ingénieurs Sup Galilée
Parcours Télécommunications et Réseaux - 1ère année
2019-2020

Plan du cours

Sommaire

Qu'est-ce qu'un signal aléatoire ?

Signaux déterministes, signaux aléatoires

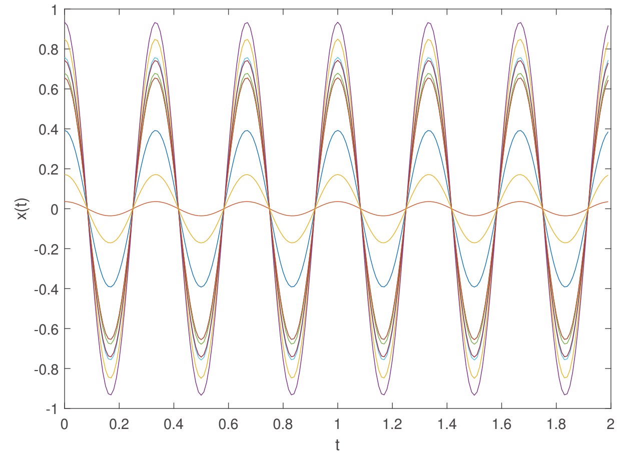
- Jusqu'à présent, nous disposions pour tous les signaux que nous avons étudiés d'une équation qui nous donnait l'expression de $x(t)$ en fonction de t on dit que ces signaux sont **déterministes**
- Dans la vraie vie en revanche, on ne dispose pas d'une équation pour tous les signaux !
- Certains phénomènes (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... On dit que ces signaux sont **aléatoires**

Modélisation des signaux aléatoires

- Les signaux aléatoires ne sont pas traitables uniquement grâce à des outils de traitement du signal. 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 $X(t)$ comme étant une **variable aléatoire** (réelle ou complexe), qui dépend donc d'événements aléatoires
- On appelle **réalisation** ou **trajectoire** d'un signal aléatoire $X(t)$, l'ensemble des valeurs prises par les v.a. $X(t)$ quand t varie.

Notion de trajectoire : exemple

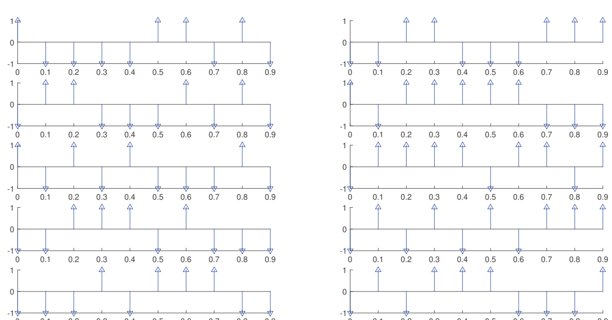
$X(t) = A\cos(2\pi f_0 t)$ avec A tiré aléatoirement de façon uniforme entre 0 et 1



10 réalisations (ou trajectoires) de ce signal aléatoire

Notion de trajectoire : exemple

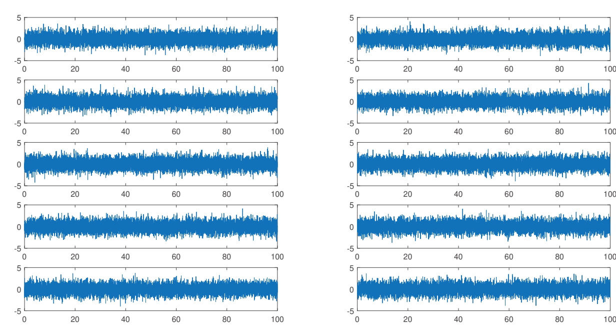
$$X(t) = \sum_{k=0}^{9} a_k \delta(t - kT_e) \text{ avec } a_k \text{ tiré aléatoirement de façon uniforme dans } \{-1, +1\}$$



10 réalisations (ou trajectoires) de ce signal aléatoire

Notion de trajectoire : exemple

$X(t)$ où chaque valeur $X(t)$ est tirée aléatoirement selon une loi gaussienne de moyenne 0 et de variance 1



10 réalisations (ou trajectoires) de ce signal aléatoire

Sommaire

Rappels de probabilités

Quels outils utiliser ?

- Chaque trajectoire d'un signal aléatoire sera différente, alors comment caractériser un tel signal ?
- On va donc raisonner en terme 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 $a < b$

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

- On a en particulier

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

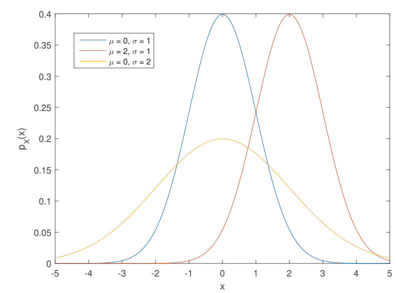
Quelques lois de probabilité

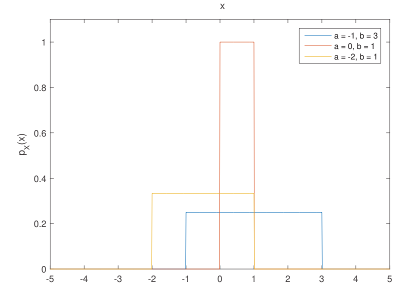
- Loi normale (ou Gaussienne) de moyenne μ et de variance σ^2

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

- 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

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

- 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

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

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

$$\boxed{\mathbb{E}\left[\psi(X)\right] = \int_{\mathbb{R}} \psi(x) \; p_X(x) dx}$$

Covariance

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

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

- 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

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

- 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.

Sommaire

Outils pour l'étude des signaux aléatoires

Stationnarité d'ordre 1

- Dans un signal aléatoire $X(t)$, chaque valeur $X(t_0)$ 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 $\mathbb{E}\left[X(t)\right]$
- En revanche, dans le cas général la quantité $\mathbb{E}\left[X(t)\right]$ n'est pas constante et varie en fonction du temps

$$\mathbb{E}\left[X(t)\right] = \mu(t)$$

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

$$\forall t, \quad \mathbb{E}\left[X(t)\right] = \mu$$

Fonction d'autocorrélation

- On peut également étudier les liens statistiques qui existent entre les variables aléatoires $X(t_1)$ et $X(t_2)$
- Etant donné un signal aléatoire $X(t)$, on appelle fonction d'autocorrélation et on note $R_X(t_1, t_2)$ la quantité

$$\boxed{R_X(t_1, t_2) = \mathbb{E}\left[X(t_1)X(t_2)\right]}$$

- Il s'agit ici d'une autocorrélation statistique, à ne pas confondre avec l'autocorrélation temporelle déjà vue pour les signaux déterministes

Fonction d'autocorrélation

- Dans le cas général la quantité $R_X(t_1, t_2)$ dépend des deux variables t_1 et t_2
- Lorsque cette quantité ne dépend que de l'écart $\tau = t_2 - t_1$ entre les instants t_1 et t_2, on dit que le signal est **stationnaire d'ordre 2**

$$\forall t_1, t_2, \quad R_X(t_1, t_2) = R_X\left(t_2 - t_1\right)$$

- Dans ce cas, il est courant d'écrire

$$R_X(\tau) = \mathbb{E}\left[X(t)X(t+\tau)\right]$$

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

Cas de deux signaux

- On peut étendre cette définition à deux signaux aléatoires $X(t)$ et $Y(t)$ en définissant une fonction d'intercorrélation
$$\boxed{R_{XY}(t_1, t_2) = \mathbb{E}\left[X(t_1)Y(t_2)\right]}$$
- On dira que de signaux sont **décorrélés** si et seulement si
$$\forall t_1, t_2, \qquad \mathrm{cov}\left(X(t_1), Y(t_2)\right) = 0$$
- Si les signaux $X(t)$ et $Y(t)$ sont à moyenne statistique nulle et décorrélés, alors leur fonction d'intercorrélation est toujours nulle. En effet, on a dans ce cas
$$R_{XY}(t_1, t_2) = \mathrm{cov}\left(X(t_1), Y(t_2)\right) = 0$$

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)
  $$\boxed{\forall t, \quad \mathbb{E}\left[X(t)\right] = \mu}$$
  2. Leur fonction d'autocorrélation ne dépend que de l'écart $\tau = t_2 - t_2$ entre deux instants (stationnarité d'ordre 2)
  $$\forall t_1, t_2, \quad R_X(t_1, t_2) = R_X\left(t_2 - t_1\right)$$
  $$\boxed{R_X(\tau) = \mathbb{E}\left[X(t)X(t+\tau)\right]}$$

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 !
- Pour un signal déterministe on a introduit la notion de densité spectrale d'énergie (DSE) ou densité spectrale d'énergie (DSP) $\Gamma_x(f)$, définie par :
$$\Gamma_x(f) = \mathcal{TF}\left\{\gamma_x(\tau)\right\}$$
- On va définir par analogie un outil similaire, également appelé **densité spectrale de puissance (DSP)**, permettant d'observer le contenu fréquentiel d'un signal aléatoire *en moyenne*

Densité spectrale de puissance

- Soit $X(t)$ un signal aléatoire stationnaire au sens large, ayant une fonction d'autocorrélation $R_X(\tau)$. On appelle densité spectrale de puissance (DSP) et on note $S_X(f)$ la quantité
$$\boxed{S_X(f) = \mathcal{TF}\left\{R_X(\tau)\right\}}$$
- $S_X(f)$ donne la répartition de la puissance selon les fréquences f
- Pour un signal aléatoire on peut définir la notion de puissance moyenne P_X définie comme
$$\boxed{P_X = \int_{-\infty}^{+\infty} S_X(f)df}$$

Filtrage des signaux aléatoires

$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

$$X(f) \longrightarrow \boxed{H(f)} \longrightarrow Y(f)$$

- Dans le cas déterministe, on a $Y(f) = X(f)H(f)$
- Si l'on suppose maintenant que le signal $X(t)$ est aléatoire, que devient l'expression ?
- $h(t)$ est la réponse impulsionnelle d'un filtre, c'est donc une quantité déterministe dont on peut prendre la transformée de Fourier
- En revanche, $Y(t)$ dépend de $X(t)$, il est donc également aléatoire !

Formule de Wiener Lee

Considérons :

- Un signal aléatoire $X(t)$ de DSP $S_X(f)$ que l'on filtre avec un filtre linéaire de fonction de transfert $H(f)$.
- Alors la DSP $S_Y(f)$ de la sortie du filtre $Y(t)$ vérifie la relation suivante :

$$\boxed{S_Y(f) = |H(f)|^2 \; S_X(f)}$$

Sommaire

Notion de bruit

Capteurs et mesures

- Lorsque l'on mesure un phénomène physique (température, tension électrique, son...), il est impossible d'accéder exactement aux quantités pertinentes que l'on cherche à quantifier.
- On observe l'apparition de signaux parasites qui viennent se superposer au signal dit utile (i.e l'information que l'on souhaite récupérer). Ces signaux sont une gêne pour la compréhension de l'information que le signal transporte.
- Ces signaux parasites sont nécessairement aléatoires car non connus : on appelle ce type de perturbation un **bruit**

Modèle de signal et bruit

- Le modèle le plus courant pour exprimer ce phénomène est le suivant

$$Y(t) = X(t) + B(t)$$

  - $X(t)$ est le signal utile, celui que l'on veut mesurer
  - $B(t)$ est le bruit de mesure, aléatoire, que l'on connaît pas
  - $Y(t)$ est le signal mesuré
  - $X(t)$ et $B(t)$ sont décorrélés
- Bien souvent, le but est de retrouver $X(t)$ à partir de l'observation bruitée $Y(t)$

Pourquoi parler de bruit ?

$$Y(t) = X(t) + B(t)$$

Exemple historique de la parole dans la téléphonie

- On cherche à transmettre un signal de voix, des phrases et des mots. Il s'agit du signal utile qui contient de l'information
- Tout le reste (grésillements, bruit ambiant...) est considéré comme du bruit, qui parasite la communication
- La qualité de la communication va dépendre des niveaux relatifs du signal utile (la voix de l'interlocuteur) et du bruit (tous les autres sons)

Rapport signal sur bruit

- Etant donné un signal $X(t)$ déterministe corrompu par un bruit $B(t)$, le signal bruité $Y(t)$ s'écrit :

$$Y(t) = X(t) + B(t)$$

- On appelle le rapport signal sur bruit, la quantité :

$$\boxed{SNR = \frac{P_X}{P_B}}$$

  Plus cette quantité est élevée, moins le bruit n'a d'importance
- On exprime souvent cette quantité en décibels :

$$\boxed{SNR_{dB} = 10\log_{10}\left(\frac{P_X}{P_B}\right)}$$

Débruitage

- Une des tâches fondamentales en traitement du signal consiste à débruiter le signal mesuré, c'est à dire à retrouver le signal $X(t)$ à partir du signal $Y(t)$
- C'est une tâche impossible à réaliser de façon parfaite, car par définition on ne connaît pas $B(t)$
- Néanmoins, si on connait quelques propriétés du bruit et/ou du signal, on peut tout de même améliorer la qualité du signal
- Exemple
  - Signal en bande de base de largeur de bande $B = 10$ Hz
  - Bruit corrompant tous les fréquences jusqu'à 100 Hz
  - On sait qu'on peut supprimer toutes les fréquences au dessus de 10 Hz, qui ne seront liées qu'au bruit. On débruitera donc grâce à un filtre passe-bas.

Notion de bruit blanc

On appelle **bruit blanc** un signal aléatoire $B(t)$ tel que

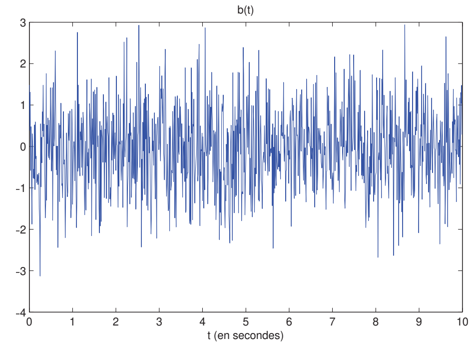
1. Le signal est stationnaire au sens large
2. Sa moyenne statistique est nulle

$$\forall t, \quad \mathbb{E}\left[B(t)\right] = 0$$

3. Deux valeurs $B(t_1)$ et $B(t_2)$ avec $t_1 \neq t_2$ sont décorrélées

$$\forall t_1, t_2 \text{ tels que } t_1 \neq t_2 \qquad \text{cov}\left(B(t_1), B(t_2)\right) = 0$$

Notion de bruit blanc



Un bruit blanc $B(t)$ est un signal aléatoire qui prend des valeurs au hasard à chaque instant t

- Les valeurs prises à chaque instant sont toutes décorrélées les unes des autres
- La moyenne de ce signal est égale à 0
- La variance du signal est égale à σ^2
- Plus σ^2 est élevé, plus le bruit blanc a tendance à prendre des valeurs éloignées de 0 et plus les amplitudes sont élevées (en valeur absolue)
- Très utilisé pour modéliser des perturbations dont on ne connaît que l'ordre de grandeur (paramétré par σ^2)

Fonction d'autocorrélation

- Comme un bruit blanc est stationnaire au sens large et de moyenne nulle on a

$$R_B(\tau) = \mathbb{E}\left[B(t)B(t+\tau)\right] = \text{cov}\left(B(t), B(t+\tau)\right)$$

- Donc la fonction d'autocorrélation est nulle pour tout $\tau \neq 0$, et pour $\tau = 0$ on a

$$R_B(0) = \mathbb{E}\left[B(t)^2\right] = \text{var}\left[B(t)\right] = \sigma^2$$

- Pour un bruit blanc on a donc

$$\boxed{R_B(\tau) = \sigma^2\delta(\tau)}$$

Densité spectrale de puissance

- Par définition de la densité spectrale de puissance, on a

$$S_B(f) = \mathcal{TF}\left\{R_B(\tau)\right\}$$

- Pour un bruit blanc on a donc

$$S_B(f) = \mathcal{TF}\left\{\sigma^2\delta(\tau)\right\} = \sigma^2$$

- La densité spectrale de puissance (DSP) d'un bruit blanc est donc

$$\boxed{S_B(f) = \sigma^2}$$

Densité spectrale de puissance

- Dans le cas d'un bruit blanc, comme tout est aléatoire et décorrélé, il n'y a pas de raison qu'une fréquence ait plus d'importance que les autres. La densité spectrale de puissance est donc constante.
- Pour un bruit blanc $B(t)$ de variance σ^2, on a donc :

$$\boxed{S_B(f) = \sigma^2}$$

- On remarque qu'on a donc $P_B = \int_{-\infty}^{+\infty} \sigma^2 df = +\infty$! Un bruit blanc n'existe donc pas dans la réalité physique, car sa puissance moyenne totale est infinie !
- Au lieu d'utiliser σ^2, il est courant d'introduire $N_0 = 2\sigma^2$ (en Watt par Hertz) et d'écrire :

$$\boxed{S_B(f) = \frac{N_0}{2}}$$

Interprétation

- Dans un bruit blanc, toutes les fréquences sont donc présentes avec des contributions égales. On l'appelle bruit blanc par analogie avec la lumière blanche qui mélange toutes les fréquences lumineuses.
- En théorie un bruit blanc ne peut pas exister car il a une puissance moyenne totale infinie. On suppose la plupart du temps que l'on travaille dans une certaine bande passante : dans ce cas la densité spectrale de puissance est la même pour toutes les fréquences de la bande passante.
- Dans le modèle signal/bruit, si on s'interesse à un système ayant une bande passante finie, on supposera donc que le bruit blanc additif aura une largeur de bande égale à cette bande passante