# TP 1 : Outils d'analyse spectrale des signaux déterministes et aléatoires

*Remarque* : les questions dont le numéro est suivi du signe † sont des questions théoriques, ne nécessitant pas l'utilisation de Matlab, qui doivent impérativement être résolues avant la séance de TP.

## I. Analyse spectrale des signaux déterministes

L'objectif de cette manipulation est d'illustrer les effets de l'utilisation de la Transformée de Fourier Discrète pour l'analyse spectrale des signaux déterministes. Avant d'arriver en séance, vous devez impérativement avoir assimilé les notions théoriques concernant :
– les hypothèses sous-jacentes à l'utilisation de la Transformée de Fourier Discrète,
– les calculs d'erreur en fréquence et d'erreur relative en amplitude,
– les techniques de bourrage de zéros (*zero padding*) et de fenêtrage,
– les problèmes de résolution en fréquence.

### I.1. Analyse spectrale d'une sinusoïde

On va effectuer l'analyse spectrale par TFD d'un signal sinusoïdal de différentes fréquences. Comme la TFD nécessite un nombre fini d'échantillons, on va échantillonner ce signal à 1 kHz et en prendre un nombre N = 100 d'échantillons.

1. † Pour quelles fréquences $f_0$ de la sinusoïde la fréquence d'échantillonnage respecte-t-elle le théorème de Shannon ?
2. † Montrer que, pour un signal de durée finie N, la TFD donne un échantillonnage de la TFSD. En déduire la précision fréquentielle que l'on peut obtenir par TFD.
3. † Dans le cas de la sinusoïde étudiée, à quoi correspond la TFSD et donc la TFD ?
4. Tracer la représentation temporelle d'un tel signal pour $f_0$ = 100 Hz. Tracer sa TFD. Peut-on ainsi retrouver la fréquence et l'amplitude de la sinusoïde ? L'amplitude maximale de la TFD correspond-elle à l'amplitude de la sinusoïde ? Comment pourrait-on normaliser la TFD pour retrouver cette amplitude ? Peut-on également retrouver sa phase à l'origine ?
5. Faire de même pour $f_0$ = 95 Hz. Commenter.
6. En déduire l'erreur relative en amplitude obtenue par l'analyse spectrale de la sinusoïde par TFD. Comparer avec la valeur théorique, calculée en TD.

On va maintenant effectuer du bourrage de zéros (*zero padding*), c'est-à-dire prolonger le signal étudié par des zéros avant d'en calculer sa TFD.

7. † Quelle est la relation entre la TFD de ce nouveau signal de longueur $N_0 > N$ et la TFSD du signal original ?
8. Construire un signal 10 fois plus long que le précédent en employant une telle technique et calculer sa TFD. Superposer le module de cette TFD au module de la TFD du signal original. Effectuer cette opération pour $f_0$ = 100 Hz et $f_0$ = 95 Hz et commenter.

### I.2. Fenêtrage

Afin d'obtenir une meilleure résolution en amplitude, on va effectuer un fenêtrage du signal avant de calculer sa TFD. De façon implicite, nous avions considéré jusqu'alors une fenêtre rectangulaire. On va maintenant étudier le cas de fenêtre de Hanning.

1. † Expliquer l'effet sur la TFSD de la pondération du signal par une fenêtre. En déduire l'effet sur la TFD et donc sur l'analyse spectrale effectuée.
2. † Quelles sont les propriétés souhaitées pour une telle fenêtre ?
3. Tracer la réponse en fréquence des fenêtres rectangulaire et Hanning et les comparer (on pourra par exemple tracer, en échelle logarithmique, le module de leur TFD après bourrage de zéros).
4. Normaliser ces fenêtres afin que dans le meilleur des cas l'amplitude maximale de la TFD soit égale à l'amplitude de la sinusoïde.
5. Fenêtrer le signal sinusoïdal et tracer le module de sa TFD. Commenter suivant la fenêtre prise en compte, pour $f_0 = 100$ Hz et $f_0 = 95$ Hz (avec $N = 100$).
6. Donner la précision relative en amplitude obtenue par Matlab pour l'analyse spectrale de la sinusoïde par TFD suivant la fenêtre prise en compte.

### I.3. Résolution spectrale

Soit le signal suivant constitué de deux sinusoïdes :

$$x[n] = A_1 \sin(2\pi n f_1/f_e + \varphi_1) + A_2 \sin(2\pi n f_2/f_e + \varphi_2).$$

L'étude d'un tel signal permet de faire apparaître les problèmes liés à la résolution de l'analyse. On se place dans un cas où la première sinusoïde a une amplitude unité $A_1 = 1$ et une fréquence $f_1 = 95$ Hz, la fréquence d'échantillonnage restant à 1 kHz, le nombre d'échantillons étant également fixé à $N = 100$. On fera varier l'amplitude et la fréquence de la seconde sinusoïde afin de mettre en avant les problèmes rencontrés lors de l'analyse spectrale par TFD de signaux réels.

1. En fixant l'amplitude $A_2 = A_1$, suivant la fenêtre de pondération utilisée, jusqu'à quelle fréquence $f_2$ peut-elle être abaissée tout en distinguant clairement les deux fréquences ? Quelles sont alors les amplitudes correspondantes ?
2. En fixant la fréquence $f_2 = 140$ Hz, suivant la fenêtre de pondération utilisée, jusqu'à quelle amplitude $A_2$ peut-elle être abaissée tout en distinguant clairement les deux fréquences ? Quelles sont alors les amplitudes correspondantes ?

## II. Analyse spectrale des signaux aléatoires

L'objectif de cette partie est d'étudier l'analyse spectrale des signaux aléatoires. En cours et en TD, on a pu caractériser de façon théorique des estimateurs de la fonction d'autocorrélation et de la densité spectrale de puissance (DSP) de signaux aléatoires. On va maintenant étudier en pratique, sur des signaux classiques simulés, l'utilisation de tels estimateurs.

Avant d'arriver en séance, vous devez impérativement avoir assimilé les notions théoriques concernant :

- les caractéristiques statistiques des estimateurs (biais, variance, …),
- les caractéristiques statistiques d'ordre 2 des signaux aléatoires et en particulier l'autocorrélation et son estimation,
- les caractéristiques fréquentielles des signaux aléatoires stationnaires au sens large et en particulier la densité spectrale de puissance et son estimation.

### II.1. Estimation de l'autocorrélation

On rappelle que l'autocorrélation d'un signal réel à temps discret stationnaire $X[n]$ est définie par

$$R_X[m] = E\{X[n+m]X[n]\}$$

D'un point de vue théorique, calculer cette fonction nécessite la connaissance de la loi de probabilité de $X$. En pratique, on ne dispose qu'un nombre fini d'échantillons d'une seule réalisation du signal, notée $x[n]$, $n=1\ldots N$, et l'on doit estimer l'autocorrélation à partir de ces échantillons.

On connaît deux estimateurs empiriques de l'autocorrélation, le premier est théoriquement non-biaisé :

$$\hat{R}_X^{nb}[m] = \frac{1}{N-m}\sum_{n=1}^{N-m} x[n+m]x[n] \qquad m \geq 0 \qquad \text{et} \qquad \hat{R}_X^{nb}[-m] = \hat{R}_X^{nb}[m]$$

le second est théoriquement biaisé :

$$\hat{R}_X^{b}[m] = \frac{1}{N}\sum_{n=1}^{N-m} x[n+m]x[n] \qquad m \geq 0 \qquad \text{et} \qquad \hat{R}_X^{b}[-m] = \hat{R}_X^{b}[m]$$

On va étudier l'utilisation pratique de ces estimateurs pour l'estimation de l'autocorrélation du signal $X_1[n] = B[n] + 1$, où $B[n]$ est un bruit blanc gaussien centré de variance unitaire.

1. † Rappeler l'autocorrélation théorique de ce signal.
2. † Rappeler les formules de biais pour les estimateurs non-biaisé et biaisé de l'autocorrélation. Quelles sont les valeurs théoriques de ces biais pour le signal ci-dessus ?
3. En utilisant la fonction Matlab *xcorr*, calculer les estimateurs biaisé et non-biaisé de l'autocorrélation pour une centaine de réalisations du signal ci-dessus. Chaque réalisation contient $N$=50 échantillons temporels. Estimer de façon empirique, pour cette centaine de réalisations, la valeur moyenne de l'estimateur ainsi que sa variance.
4. Tracer sur un même graphique la valeur moyenne ainsi que la valeur moyenne plus ou moins l'écart-type de ces estimateurs.
5. Commenter ces graphiques en termes de biais et variance des différents estimateurs pour les différentes valeurs de $m$.

## II.2. Analyse spectrale non-paramétrique

La densité spectrale de puissance (DSP) d'un signal aléatoire est définie comme la transformée de Fourier de sa fonction d'autocorrélation. D'un point de vue empirique, on peut fabriquer deux types d'estimateur de la DSP d'un signal aléatoire : les premiers estimant l'autocorrélation à partir d'une réalisation du signal puis calculant la transformée de Fourier de cette corrélation ; les seconds estimant directement la DSP à partir de la transformée de Fourier de la réalisation du signal. Dans la suite, on s'intéresse surtout à ce deuxième type d'estimateur, à savoir le périodogramme et ses dérivées.

### II.2.1. Périodogramme

Le périodogramme est défini par $\hat{S}_X^p(f) = \frac{1}{N}\left|\hat{x}(f)\right|^2$ avec $\hat{x}(f) = \sum_{n=1}^{N} x[n]e^{-j2\pi fn}$ la transformée de Fourier de la réalisation du signal.

1. † Rappeler la DSP théorique d'un bruit blanc centré de variance unitaire.
2. † Le périodogramme est-il un estimateur non-biaisé de la DSP ? Quel est le biais de cet estimateur dans le cas général ? Quel est son biais dans le cas particulier d'un bruit blanc ?
3. Ecrire un programme Matlab calculant le périodogramme pour 100 réalisations d'un bruit blanc centré gaussien de puissance unitaire. Chaque réalisation contient $N$=128 échantillons temporels. Tracer sur un même graphique la valeur moyenne ainsi que la valeur moyenne plus ou moins l'écart-type de cet estimateur. Répéter le test avec 1000 réalisations. Conclure sur le biais du périodogramme d'un bruit blanc.

## II.2.2. Périodogramme moyenné

Pour réduire la variance du périodogramme, on peut utiliser l'approche suivante :

- on divise les $N$ échantillons du signal observé en $L$ sections $x_l[n]$, chacune de $M=N/L$ échantillons,
- on calcule le périodogramme sur chaque section,
- on calcule la moyenne des périodogrammes :

$$\overset{o}{S}{}_X^{pm}(f) = \frac{1}{L}\sum_{l=1}^{L} \overset{o}{S}{}_{X_l}^{p}(f) \quad \text{avec} \quad \overset{o}{S}{}_{X_l}^{p}(f) = \frac{1}{M}\left|\sum_{n=0}^{M-1} x_l[n]e^{-j2\pi fn}\right|^2$$

On va étudier l'utilisation pratique du périodogramme moyenné pour l'estimation de la DSP des signaux suivants :

$X_1[n]$ : bruit blanc centré de variance unitaire.
$X_2[n] = \cos(2\pi f_0 n+\varphi_0) + \cos(2\pi f_1 n+\varphi_1)$ avec $f_0=0.1$, $f_1=0.15$ et $\varphi_0$ et $\varphi_1$ deux variables aléatoires uniformément distribuées sur $[0, 2\pi]$.

1. Générer une réalisation de N=1024 échantillons de chacun des signaux ci-dessus.
2. Calculer et tracer le périodogramme moyenné de chaque signal avec différentes valeurs de $L$ dans la formule du périodogramme moyenné : $L=1$ (périodogramme simple), $L=16$, $L=32$, $L=128$. Conclure sur le meilleur choix de $L$ pour estimer la DSP de chacun des deux signaux.

## II.2.3. Périodogramme de Welch

Le périodogramme de Welch combine le moyennage et le lissage pour réduire la variance de l'estimateur :

- on divise le signal en $L$ sections de $M$ échantillons,
- chaque section est multipliée par une fenêtre $w[n]$,
- on calcule le périodogramme de chaque section fenêtrée,
- on calcule la moyenne de ces périodogrammes.

On va comparer le périodogramme simple et le périodogramme de Welch pour l'estimation de la DSP du signal ci-dessous :

$X[n] = \cos(2\pi f_0 n+\varphi_0)+B[n]$ avec $f_0=0.1$, $\varphi_0$ une variable aléatoire uniformément distribuée sur $[0, 2\pi]$ et $B[n]$ un bruit blanc de variance 4.

1. Générer une réalisation de $N=1024$ échantillons de ce signal.
2. Tracer les périodogrammes simple et de Welch de ce signal en utilisant les fonctions *periodogram* et *pwelch* de Matlab (utiliser les versions par défaut[1]). Commenter.

---

[1] La version par défaut de la fonction *pwelch* choisit $L=8$ sections avec un chevauchement de 50% et une fenêtre de Hamming.