

Traitement du Signal

Auteurs: G. Chardon , R. Combes, J. Fiorina,
E. Lahalle, J. Picheral, P. Rodriguez,
C. Soussen, G. Valmorbida, A. Wautier

CentraleSupélec, Avril 2020

TABLE DES MATIÈRES

Résumé	7
Notations	9
1 BASES MATHÉMATIQUES	11
1.1 Espaces vectoriels de dimension finie	11
1.1.1 Définition	11
1.1.2 Applications linéaires	12
1.1.3 Norme, produit scalaire	12
1.2 Espaces de Banach	13
1.2.1 Espaces ℓ^p	14
1.2.2 Espaces L^p	14
1.3 Espaces de Hilbert	14
1.3.1 Séries de Fourier	16
1.4 Distributions	16
1.4.1 Distributions tempérées	19
1.4.2 Distributions à support compact	19
1.5 Variables et processus aléatoires	20
1.5.1 Brefs rappels de statistiques	24
2 SYSTÈMES LINÉAIRES INVARIANTS	25
2.1 Notion de signal	25
2.1.1 Définitions	25
2.1.2 Représentations énergétiques	26
2.2 Convolution	26
2.2.1 Cas continu : convolutions de fonctions	27
2.2.2 Cas discret	28
2.2.3 Signaux modélisés par des distributions	28
2.2.4 Propriétés du produit de convolution	30
2.3 Notion de système	30
2.3.1 Définition	30
2.3.2 Propriétés des systèmes	31
2.4 Filtres linéaires convolutifs	33
2.4.1 Définition	33
2.4.2 Filtres à réponse impulsionnelle finie (FIR)	34
2.4.3 Filtre à réponse impulsionnelle infinie (IIR)	36
2.4.4 Équations aux différences	36
2.5 Fonctions propres	37
2.6 Conclusion	38
3 TRANSFORMÉES DE FOURIER	39
3.1 Transformée de Fourier à Temps continu	40
3.1.1 Transformée de Fourier des signaux dans L^1	40
3.1.2 Transformée de Fourier des signaux dans L^2	42
3.1.3 Transformée de Fourier dans $\mathcal{S}'$	43

3.1.4	Quelques exemples classiques	46
3.2	Décomposition en série de Fourier	47
3.2.1	Signaux à support limité	47
3.2.2	Signaux périodiques	48
3.3	Transformée de Fourier à temps discret	48
3.3.1	Transformée de signaux dans ℓ^1	49
3.3.2	Transformée des signaux dans ℓ^2	50
3.3.3	Quelques exemples classiques	51
3.4	Transformée de Fourier Discrète	51
3.5	Application au filtrage linéaire	53
3.5.1	Filtrage dans le domaine de Fourier	53
3.5.2	Calcul numérique de la réponse fréquentielle d'un filtre FIR	55
3.6	Analyse spectrale	57
3.6.1	Effet du fenêtrage	57
3.6.2	Notion de résolution	58
3.6.3	Calcul numérique avec la TFD	58
3.7	Conclusion	60
4	ECHANTILLONNAGE	63
4.1	Effet de l'échantillonnage	63
4.2	Théorème de Shannon	66
4.2.1	Base orthogonale de PW_B	69
4.2.2	Convergence de la série	69
4.3	L'échantillonnage en pratique	70
4.3.1	Filtre anti-repliement	70
4.3.2	Reconstruction pratique	71
4.4	Conclusion	71
5	PROCESSUS ALÉATOIRES	73
5.1	Définition et caractérisation des processus aléatoires	73
5.1.1	Définition	74
5.1.2	Lois fini-dimensionnelles	75
5.1.3	Processus gaussien	76
5.1.4	Moyenne et autocorrélation	76
5.2	Processus stationnaires	78
5.2.1	Stationnarité stricte	78
5.2.2	Stationnarité au sens large (SSL)	78
5.2.3	Densité spectrale de puissance (DSP)	80
5.2.4	Bruit blanc	82
5.2.5	Filtrage de processus	84
5.2.6	Processus autorégressif (AR)	86
6	ESTIMATION	91
6.1	Estimateur de la moyenne et de l'autocorrélation	92
6.1.1	Estimateur de la moyenne	92
6.1.2	Estimateur de l'autocorrélation	93
6.2	Estimation de la DSP : périodogramme	96
6.2.1	Définition	96

6.2.2	Propriétés statistiques du périodogramme	98
6.2.3	Variantes du périodogramme	99
6.3	Estimation paramétrique	102
6.3.1	Estimation de la DSP : application à un modèle AR	102
6.4	Estimateur linéaire optimal	105
6.4.1	Définition et propriétés	106
6.4.2	Prédiction d'un processus aléatoire	107
6.4.3	Filtre de Wiener pour le débruitage	109
	Bibliographie	114

RÉSUMÉ

CONTEXTE ET ENJEUX

Le monde numérique produit des volumes importants de signaux de toutes sortes (audio, images, vidéo, mesures physiques) dans des domaines aussi variés que la santé, les télécommunications, l'industrie ou l'environnement. L'extraction d'information de ces données est de plus en plus nécessaire pour le diagnostic médical, la compression de données, la suppression de bruits parasites d'un signal audio, *etc.* Ce cours introduit des méthodes pour aborder différents problèmes comme le filtrage, la transmission, le débruitage de signaux, ou l'analyse spectrale.

ORGANISATION

Les méthodes présentées reposent sur des concepts mathématiques tels que les espaces de Hilbert, les distributions, et les vecteurs aléatoires. Ces concepts sont revisités au chapitre 1. Le chapitre 2 introduit les notions de signaux et de systèmes, en particulier les filtres linéaires convolutifs. Le chapitre 3 étudie les transformées Fourier, un outil incontournable du traitement du signal. Les différentes transformations sont introduites pour des signaux à temps continu et à temps discret, et utilisées pour analyser le contenu fréquentiel des signaux (analyse spectrale) et pour caractériser les filtres linéaires dans le domaine fréquentiel. La théorie de l'échantillonnage introduite au chapitre 4 définit les conditions sous lesquelles l'information contenue dans un signal est conservée lors de sa discrétisation. Le théorème de Shannon est le résultat fondamental permettant d'assurer la reconstruction exacte d'un signal à bande limitée. Enfin, les chapitres 5 et 6 introduisent un cadre probabiliste pour caractériser les propriétés statistiques de signaux, vus comme des processus aléatoires. L'analyse de signaux stationnaires repose sur la fonction d'auto-corrélation (domaine temporel) et la densité spectrale de puissance (domaine fréquentiel), introduites formellement au chapitre 5. Le chapitre 6 présente les techniques d'estimation de ces caractéristiques à partir de données expérimentales, avec des applications au débruitage, au filtrage, à la prédiction de données, et à l'analyse spectrale.

NOTATIONS

VECTEURS

$\mathbf{x} \in \mathbb{C}^n$	vecteur
$x[k]$	k -ème coordonnée du vecteur

REPRÉSENTATIONS FRÉQUENTIELLES

$(\mathcal{F}x)(f)$ ou $X(f)$	transformée de Fourier
$(\mathcal{F}x)(\nu)$ ou $X(\nu)$	transformée de Fourier à temps discret
$\mathbf{X} \in \mathbb{C}^n$	transformée de Fourier discrète d'un vecteur $\mathbf{x} \in \mathbb{C}^n$

VARIABLES ALÉATOIRES

X_k	variable aléatoire à l'instant k
-------	------------------------------------

PRODUITS SCALAIRES ET NORMES

$(x, y)_{L^2}$	produit scalaire dans $L^2(\mathbb{R})$, égal à $\int_{\mathbb{R}} x y^* d\lambda$
$(x, y)_{L^2}$	produit scalaire dans $L^2(0, T)$, égal à $\frac{1}{T} \int_{[0, T]} x y^* d\lambda$
(x, y)	produit scalaire dans $\ell^2(\mathbb{Z})$, égal à $\sum_{k \in \mathbb{Z}} x[k] y^*[k]$
$\ \cdot\ _{L^p}$	norme L^p , avec $\ x\ _{L^p}^p = \int_{\mathbb{R}} x(t) ^p dt$
$\ \cdot\ _p$	norme ℓ^p , avec $\ \mathbf{x}\ _p^p = \sum_k x[k] ^p$

DISTRIBUTIONS

$\langle T, \varphi \rangle$	image d'une fonction test φ par la distribution T
------------------------------	---

DIRAC vs. KRONECKER

δ_a	distribution de Dirac décentrée : $\langle \delta_a, \varphi \rangle = \varphi(a)$
$\delta = \delta_0$	distribution de Dirac centrée : $\langle \delta, \varphi \rangle = \varphi(0)$
$\mathbb{I}_T$	peigne de dirac : $\mathbb{I}_T = \sum_{k \in \mathbb{Z}} \delta_{kT}$
$\delta[k]$	fonction de Kronecker, $\delta[k] = 1$ si $k = 0$, 0 sinon

EXPONENTIELLES COMPLEXES

$e_f(t) = \exp(2i\pi ft)$	exponentielle complexe de fréquence f
$e_n(t) = \exp\left(2i\pi \frac{nt}{T}\right)$	exponentielle complexe utilisée pour les séries de Fourier [notation simplifiée pour $e_{n/T}$]
$e_\nu[k] = \exp(2i\pi \nu k)$	exponentielle complexe discrète de fréquence ν
$\mathbf{e}_n \in \mathbb{R}^N$	Version discrétisée de la fonction $\exp(2i\pi nt)$ sur $[0, 1[$: $\mathbf{e}_n = \left\{ \exp(2i\pi \frac{nk}{N}), k = 0, \dots, N-1 \right\}$

BASES MATHÉMATIQUES

Le but de ce chapitre est de poser les bases mathématiques de ce cours : espaces fonctionnels, opérations de base sur les signaux, rappels de probabilité, et de donner des liens avec les problématiques abordées dans le cours de traitement du signal.

On appellera **signal** une fonction, une distribution, ou une suite numérique, dépendant le plus souvent du temps (signal audio, télécommunications), de l'espace (image, données géographiques), ou des deux (vidéo).

1.1 ESPACES VECTORIELS DE DIMENSION FINIE

1.1.1 Définition

On appelle **espace vectoriel sur un corps $\mathbb{K}$** (ici, $\mathbb{R}$ ou $\mathbb{C}$) un ensemble E muni d'une loi de composition interne $+$ de $E \times E$ dans E , et d'une **loi de composition externe $\cdot$** de $\mathbb{K} \times E$ dans E .

Ces lois sont telles que $(E, +)$ est un groupe, que $\cdot$ vérifie des propriétés d'associativité et de distributivité par rapport à $+$, et que l'élément neutre de $\cdot$ soit l'élément neutre de la multiplication de $\mathbb{K}$.

Les espaces vectoriels en dimension finie admettent une base, c'est-à-dire une famille $\mathcal{B}$ de N vecteurs $(\mathbf{e}_i)_{1 \leq i \leq N}$ telle que tout vecteur $\mathbf{x}$ de E s'écrive comme une combinaison linéaire des éléments de la base **de façon unique**. On appelle coordonnées de $\mathbf{x}$ dans la base $\mathcal{B}$ les x_i tels que :

$$\mathbf{x} = \sum_{i=1}^N x_i \mathbf{e}_i.$$

Toutes les bases d'un même espace vectoriel contiennent le même nombre N de vecteurs. On appelle N la dimension de l'espace vectoriel.

APPLICATION EN TRAITEMENT DU SIGNAL : Les signaux en dimension finie seront membres d'un tel espace. **Pour un signal sonore, on peut penser à l'amplification ou à l'atténuation pour la multiplication par un scalaire, et au mixage de deux sons pour l'addition.**

1.1.2 Applications linéaires

On appelle application linéaire une application L de E dans F telle que pour tout $\mathbf{u}$ et $\mathbf{v}$ dans E , et λ dans $\mathbb{K}$,

$$\begin{aligned} L(\mathbf{u} + \mathbf{v}) &= L(\mathbf{u}) + L(\mathbf{v}) \\ L(\lambda \mathbf{u}) &= \lambda L(\mathbf{u}). \end{aligned}$$

La propriété de linéarité permet de simplifier cette description. En se donnant une base $\mathcal{B}_E$ de E , et $\mathcal{B}_F$ de F , on montre que les coefficients de $\mathbf{y} = L(\mathbf{x})$ sont donnés par :

$$y_n = \sum_{m=1}^N a_{nm} x_m, \quad (1)$$

pour N la dimension de E , et a_{nm} des éléments de $\mathbb{K}$, dépendant de L et du choix des bases. On appelle $\mathbf{A}$ la matrice de l'application L dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$. C'est le tableau à deux entrées contenant les éléments a_{nm} , et on définit le produit matriciel

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad (2)$$

par l'équation (1). On peut donc représenter L par sa matrice $\mathbf{A}$, simple tableau de nombres (aisément manipulable par un calculateur), et écrire le résultat de son application à un vecteur comme un produit matriciel (2), très simple à implémenter.

Certaines applications de E dans E sont de plus diagonalisables. Le résultat $\mathbf{y} = \mathbf{A}\mathbf{x}$ de leur application à un vecteur $\mathbf{x}$ se résume à multiplier les coefficients de $\mathbf{x}$ dans une base bien choisie (de vecteurs propres, tels que $L(\mathbf{u}_n) = \lambda_n \mathbf{u}_n$) :

$$y_n = \lambda_n x_n.$$

Une telle description permet une interprétation et un calcul plus simple de l'action de L sur des vecteurs de E .

APPLICATION EN TRAITEMENT DU SIGNAL : on s'intéressera en traitement du signal à une classe particulière d'applications linéaires (en dimension finie ou pas), qui en plus d'être linéaires, sont invariantes par translation, et qu'on appellera *filtres*. De la même façon qu'on décrit une application linéaire par une matrice, un filtre sera décrit par sa réponse impulsionnelle. L'équivalent du produit matriciel sera la convolution. La transformée de Fourier jouera le rôle de la diagonalisation.

1.1.3 Norme, produit scalaire

On appelle norme $N(\cdot)$ sur E une application de E dans $\mathbb{R}_+$, vérifiant les axiomes suivants :

- séparabilité : $N(\mathbf{x}) = 0 \Rightarrow \mathbf{x} = 0$,
- homogénéité : $N(\lambda \mathbf{x}) = |\lambda| N(\mathbf{x})$,
- inégalité triangulaire : $N(\mathbf{x} + \mathbf{y}) \leq N(\mathbf{x}) + N(\mathbf{y})$.

Un espace vectoriel muni d'une norme est appelé espace vectoriel normé.

Une norme permet de définir une topologie, et donc de parler de convergence, par exemple d'une suite de vecteurs $\mathbf{x}_n$ approximant un vecteur $\mathbf{x}$. On notera qu'en dimension finie, toutes les normes sont équivalents, et définissent donc la même topologie : la notion de convergence ne dépend pas alors de la norme choisie.

On rencontre souvent les normes suivantes :

- $\|\mathbf{x}\|_1 = \sum_{i=1}^N |x_i|$,
- $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^N |x_i|^2}$,
- $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq N} |x_i|$

Un espace de dimension finie muni d'un produit scalaire est dit *préhilbertien*. Un produit scalaire (hermitien, si $\mathbb{K} = \mathbb{C}$) est une application bilinéaire (sesquilinéaire) vérifiant les propriétés suivantes :

- symétrie : $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})^*$,
- positivité : $(\mathbf{x}, \mathbf{x}) \in \mathbb{R}_+$,
- définie : $(\mathbf{x}, \mathbf{x}) = 0 \Rightarrow \mathbf{x} = 0$.

Un produit scalaire induit une norme $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$. En particulier, le produit scalaire $(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N x_i y_i^*$ induit la norme $\|\cdot\|_2$. L'égalité de Cauchy-Schwarz nous indique que

$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\| \|\mathbf{y}\|. \quad (3)$$

On peut définir dans ces espaces des bases orthonormales, *i.e.*, des bases $(\mathbf{e}_i)_{1 \leq i \leq N}$ telles que $(\mathbf{e}_i, \mathbf{e}_j) = 1$ si $i = j$, et 0 sinon. Dans ce cas, les coefficients de décomposition de $\mathbf{x}$ sur la base $(\mathbf{e}_i)_{1 \leq i \leq N}$

$$\mathbf{x} = \sum_{i=1}^N x_i \mathbf{e}_i. \quad (4)$$

sont directement donnés par :

$$x_i = (\mathbf{x}, \mathbf{e}_i). \quad (5)$$

1.2 ESPACES DE BANACH

L'étude des espaces vectoriels ne pose pas de difficultés particulières en dimension finie : tout espace de dimension finie admet une base, on peut donc toujours représenter un vecteur par ses coordonnées, et toutes les normes définies sur un espace vectoriel donné sont équivalentes. De plus, tout espace vectoriel de dimension finie sur $\mathbb{R}$ ou $\mathbb{C}$ est complet.

La situation est bien sûr autrement plus complexe en dimension infinie. On appelle *espace de Banach* un espace vectoriel normé et complet. Nous utiliserons deux types d'espaces de Banach.

1.2.1 Espaces ℓ^p

Pour $p \leq 1$, l'espace ℓ^p est l'espace des suites $(x[n])_{n \in \mathbb{Z}}$ telles que $\sum_{n \in \mathbb{Z}} |x[n]|^p < \infty$. On s'intéressera en particulier :

- aux suites intégrables ℓ^1 , $\|x\|_1 = \sum_{n \in \mathbb{Z}} |x[n]|$,
- aux suites d'énergie finie ℓ^2 , $\|x\|_2^2 = \sum_{n \in \mathbb{Z}} |x[n]|^2$,
- aux suites bornées ℓ^∞ , $\|x\|_\infty = \sup_{i \in \mathbb{Z}} |x[i]|$.

On note $\delta[n]$ la suite définie par

$$\delta[n] = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{sinon} \end{cases} \quad (6)$$

δ est appelée **fonction de Kronecker** ou **impulsion de Kronecker**.

Remarque. La notation δ sera aussi utilisée pour désigner la distribution de Dirac, qu'on peut interpréter comme la limite d'une suite de fonctions définies sur $\mathbb{R}$. Au contraire de la distribution de Dirac, la fonction de Kronecker $\delta[k]$ est une suite définie sur $\mathbb{Z}$.

1.2.2 Espaces L^p

De même, on peut montrer que les espaces $L^p(I)$, pour $1 \leq p \leq \infty$ et I un intervalle, sont des espaces de Banach. On définit d'abord $\mathcal{L}^p(I)$, l'ensemble des fonctions mesurables x définies sur I telles que $\int_I |x|^p < \infty$. L'espace $L^p(I)$ est alors l'ensemble des classes d'équivalence de $\mathcal{L}^p(I)$ par la relation "égal presque partout". On s'intéressera en particulier :

- aux fonctions intégrables $L^1(\mathbb{R})$, $\|x\|_{L^1} = \int_{\mathbb{R}} |x| d\lambda$,
- aux fonctions d'énergie finie $L^2(\mathbb{R})$, $\|x\|_{L^2}^2 = \int_{\mathbb{R}} |x|^2 d\lambda$,
- aux fonctions essentiellement bornées $L^\infty(\mathbb{R}) = \text{esssup } |x|$.

On notera bien que ces espaces ne prennent pas en compte la régularité des fonctions (i.e., les fonctions de L^p n'ont aucune raison d'être continues), et qu'ils n'ont pas de relation d'inclusion entre eux (sauf si I est borné, auquel cas $L^p(I) \subset L^q(I)$ si $p > q$).

APPLICATION EN TRAITEMENT DU SIGNAL : les espaces ℓ^p serviront pour modéliser les signaux à temps discret, les espaces L^p pour les signaux à temps continu. En particulier, dans le cas $p = 2$ développé dans la section suivante, on parle de signaux d'énergie finie, où l'énergie est $\| \cdot \|_2^2$.

1.3 ESPACES DE HILBERT

Le cas particulier de $L^2(I)$ mérite un peu plus d'attention. Dans le cas où l'intervalle I n'est pas borné, on définit le produit scalaire dans $L^2(I)$ par :

$$(x, y)_{L^2} = \int_{\mathbb{R}} xy^* d\lambda. \quad (7)$$

Si I est borné, on choisit de normaliser le produit scalaire par la longueur de I ¹. Ainsi, le produit scalaire sur $L^2([0, T])$ est défini par :

$$(x, y)_{L^2} = \frac{1}{T} \int_{[0, T]} xy^* d\lambda. \quad (8)$$

Le produit scalaire vérifie évidemment l'inégalité de Cauchy-Schwarz.

$L^2(I)$ est alors un espace de Hilbert, *i.e.*, un espace de Banach muni d'un produit scalaire. De plus, L^2 est séparable (il existe un sous-ensemble dénombrable dense dans L^2). Il existe donc une base hilbertienne $(e_i)_{i \in \mathbb{Z}}$ de $L^2(I)$: toute fonction x de L^2 s'écrit comme la somme d'une série

$$x = \sum_{i \in \mathbb{Z}} c_i e_i, \quad (9)$$

où la convergence se fait en norme L^2 . De plus, la famille $(e_i)_{i \in \mathbb{Z}}$ est orthonormale. Les coefficients c_i sont donnés par

$$c_i = (x, e_i)_{L^2}. \quad (10)$$

Le théorème de représentation de Riesz est une propriété importante des espaces de Hilbert : une forme linéaire continue ℓ sur un espace de Hilbert H peut toujours s'écrire

$$\forall x \in H, \ell(x) = (x, x_\ell)_H$$

où $x_\ell \in H$. On peut donc, par l'intermédiaire du produit scalaire, confondre H et son dual.

On peut également définir les espaces de Sobolev H^m , dont les fonctions x sont telles que x et ses dérivées jusqu'à l'ordre m (au sens des distributions) soient dans L^2 . Ces espaces permettent de prendre en compte la régularité des fonctions. Pour le cas particulier de H^1 , le produit scalaire est :

$$(x, y)_{H^1} = (x, y)_{L^2} + (x', y')_{L^2}.$$

APPLICATION EN TRAITEMENT DU SIGNAL : la construction de bases d'espaces de Hilbert tels que $L^2(\mathbb{R})$ est très importante en traitement du signal, puisqu'elle est liée à la décomposition des signaux en temps continu. On peut citer par exemple les bases d'ondelettes discrètes, les bases de cosinus locaux, et pour les signaux à bande limitée, les bases de sinus cardinaux, qui permettront de les échantillonner sans perte d'information.

1. pour des raisons qui deviendront plus claires au paragraphe 1.3.1.

1.3.1 Séries de Fourier

L'espace $L^2([0, T])$ des fonctions de carré intégrable définies sur $[0, T]$ est un espace de Hilbert séparable. Une base orthonormale classique de $L^2([0, T])$ est la base de Fourier définie par :

$$e_n(t) = \exp\left(i2\pi \frac{nt}{T}\right), \quad n \in \mathbb{Z}. \quad (11)$$

Les coefficients c_n de la décomposition

$$x = \sum_{n \in \mathbb{Z}} c_n e_n \quad (12)$$

$$x(t) = \sum_{n \in \mathbb{Z}} c_n \exp\left(i2\pi \frac{nt}{T}\right) \quad (13)$$

sont donnés par :

$$\begin{aligned} c_n &= (x, e_n)_{L^2} \\ &= \frac{1}{T} \int_{[0, T]} x(t) \exp(-i2\pi nt) dt \end{aligned} \quad (14)$$

APPLICATION EN TRAITEMENT DU SIGNAL : la série de Fourier permet de décomposer un signal périodique de période T (signal de parole, note de musique, etc.) en sommes d'harmoniques. L'appartenance de la fonction à un espace de Sobolev H^m garantit la décroissance rapide des coefficients.

1.4 DISTRIBUTIONS

Comme vu plus haut, dans un espace de Hilbert H (par exemple L^2), il est possible d'identifier les vecteurs de H avec les formes linéaires continues sur H . C'était déjà le cas pour les espaces de dimension finie. Par contre, pour L^p en général, ce n'est plus le cas. On peut montrer que pour $p > 1$, le dual de L^p est isomorphe à L^q , avec $1/p + 1/q = 1$. Pour les espaces L^p définis sur un intervalle borné (par exemple $[0, 1]$), on remarque que pour $p > 2$, le dual L^p n'est pas égal à L^p , et est même strictement plus gros (en effet, c'est L^q , avec $q = p/(p-1) < p$). On peut donc faire grossir le dual en limitant l'espace de départ.

Cette idée est poussée à l'extrême par les distributions, qui forment l'espace $\mathcal{D}'(\mathbb{R})$ des formes linéaires continues sur l'espace $\mathcal{D}(\mathbb{R})$ des fonctions C^∞ à support compact, qu'on appelle fonctions test. L'espace de départ étant ici très restreint, son dual est absolument gigantesque, et contient par exemple tous les espaces L^p pour $p \geq 1$.

En effet, parmi les distributions, on trouve les distributions régulières associées aux fonctions localement intégrables (et donc, toutes

dual topologique de L^p = ensemble des applications linéaires continues de L^p dans $\mathbb{K}$

les fonctions de L^p , pour $1 \leq p \leq \infty$). Dans ce cas, l'image d'une fonction test φ par la distribution T_x associée à la fonction x est donné par

$$\langle T_x, \varphi \rangle = \int_{\mathbb{R}} x \varphi \, d\lambda. \quad (15)$$

Cependant, toutes les distributions ne sont pas associées à des fonctions. En particulier, la distribution δ de Dirac

$$\langle \delta, \varphi \rangle = \varphi(0) \quad (16)$$

ne peut pas s'écrire comme l'intégrale de φ contre une fonction. Plus généralement, on définit δ_a par

$$\langle \delta_a, \varphi \rangle = \varphi(a). \quad (17)$$

On a en particulier

$$\delta_0 = \delta. \quad (18)$$

Attention : **ne pas confondre la distribution de Dirac δ avec la fonction de Kronecker $\delta[k]$** définie en (6). Cette dernière est une suite numérique.

On définit également le **peigne de Dirac** comme la somme de distributions de Dirac espacées de a :

$$\mathbb{D}_a = \sum_{n \in \mathbb{Z}} \delta_{na}. \quad (19)$$

Égalité et convergence au sens des distributions sont définies par :

$$\begin{aligned} T = S &\Leftrightarrow \forall \varphi \in \mathcal{D}(\mathbb{R}), \langle T, \varphi \rangle = \langle S, \varphi \rangle, \\ T_n \rightarrow T &\Leftrightarrow \forall \varphi \in \mathcal{D}(\mathbb{R}), \langle T_n, \varphi \rangle \rightarrow \langle T, \varphi \rangle. \end{aligned}$$

Toute distribution est dérivable, et sa dérivée est encore une distribution, définie par

$$\langle T', \varphi \rangle = - \langle T, \varphi' \rangle, \quad (20)$$

qui étend l'intégration par parties quand T est une fonction suffisamment régulière. En particulier, la fonction de Heaviside $x = \mathbb{1}_{\mathbb{R}^+}$ n'est pas dérivable au sens classique, mais sa dérivée au sens des distributions est la distribution de Dirac δ (remarque : la valeur de $\mathbb{1}_{\mathbb{R}^+}$ en zéro n'a, du point de vue des distributions, aucune importance). Ce résultat peut se généraliser de la façon suivante : avec $I = [a, b]$ un intervalle de $\mathbb{R}$ et x une fonction dérivable, la dérivée de $x\mathbb{1}_I$ est donnée par

$$(x\mathbb{1}_I)' = x'\mathbb{1}_I + x(a)\delta_a - x(b)\delta_b. \quad (21)$$

Les discontinuités de la fonction font apparaître des distributions de Dirac dans la dérivée, avec des amplitudes égales à celles des sauts.

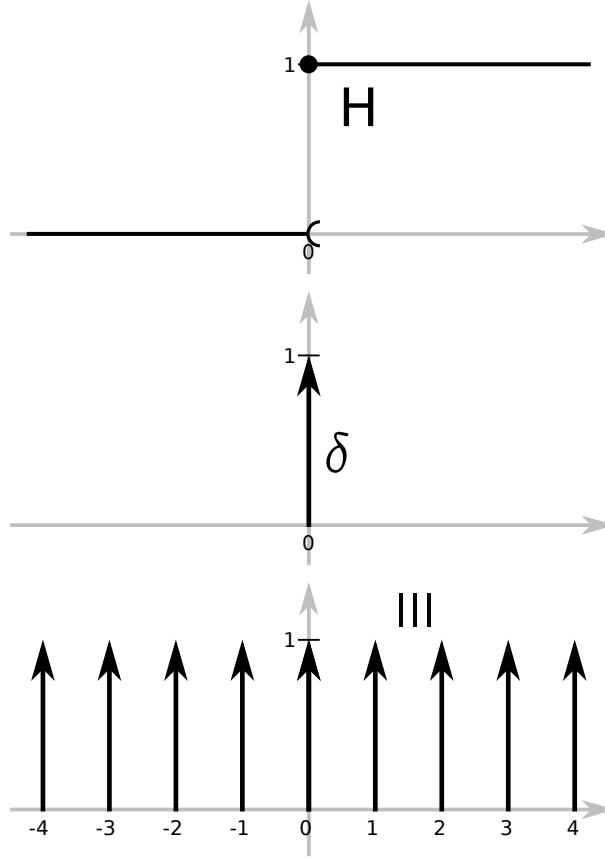


FIGURE 1 – Fonction de Heaviside $H = \mathbb{1}_{\mathbb{R}_+}$, distribution de Dirac $\delta = \delta_0$ et peigne de Dirac $\mathbb{I}_a$ pour $a = 1$.

Un changement de variable affine $\chi(x) = ax + b$ est défini de la façon suivante :

$$\langle T \circ \chi, \varphi \rangle = \frac{1}{|a|} \langle T, \varphi \circ \chi^{-1} \rangle. \quad (22)$$

Cette définition est justifié par le fait que si T est la distribution régulière associée à une fonction localement intégrable, elle est cohérente avec la formule de changement de variable dans le calcul de l'intégrale de $(x \circ \chi)\varphi$.

En particulier, pour la distribution de Dirac et $\chi(x) = ax - b$,

$$\delta_0 \circ \chi = \frac{1}{|a|} \delta_{b/a}. \quad (23)$$

APPLICATION EN TRAITEMENT DU SIGNAL : les distributions sont utiles pour modéliser les systèmes linéaires invariants dans le temps, définis plus tard dans le document, par des convolutions. En particulier, l'élément neutre de la convolution est la distribution δ . L'opérateur δ_a agit comme un retard sur un signal, avec un décalage

en temps égal à a . On utilisera parfois la notation $\delta(t - a)$ à la place de δ_a , pour rendre ce retard plus explicite. On gardera bien à l'esprit qu'une distribution n'est pas une fonction de t , et que cette notation n'est rien d'autre qu'un abus commode. Notons que la distribution de Dirac peut être vue comme la limite, au sens des distributions, d'une suite d'impulsions d'intégrale 1 et de supports de plus en plus concentrés autour de zéro.

Le peigne de Dirac sera important pour l'étude de l'opération d'échantillonnage (passage d'un signal en temps continu à un signal en temps discret).

1.4.1 Distributions tempérées

On définit l'espace de Schwartz $\mathcal{S}$ comme l'ensemble des fonctions $\varphi \in C^\infty$ sur $\mathbb{R}$ telles que

$$\forall \alpha, \beta \in \mathbb{N}, \sup_{t \in \mathbb{R}} |t^\alpha \varphi^{(\beta)}(t)| < \infty, \quad (24)$$

c'est-à-dire qu'elles, et leur dérivées, décroissent plus vite que n'importe quelle puissance négative de $|t|$.

Les membres de son dual $\mathcal{S}'$ sont appelées distributions tempérées. Les fonctions test étant dans l'espace de Schwartz ($\mathcal{D} \subset \mathcal{S}$), les distributions tempérées sont également des distributions ($\mathcal{S}' \subset \mathcal{D}'$).

Quelques exemples de distributions tempérées :

- fonctions localement intégrable à croissance au plus polynomiale
- distribution de Dirac
- peigne de Dirac

1.4.2 Distributions à support compact

Avant de définir les distributions à support compact, on rappelle que le support d'une fonction, noté $\text{supp}(\varphi)$, est défini comme l'adhérence de l'ensemble des valeurs x pour lesquelles $\varphi(x) \neq 0$.

Définition 1.1. Une distribution $T \in \mathcal{D}'(\mathbb{R})$ est dite nulle sur un ouvert $\mathcal{O}$ de $\mathbb{R}$ si pour tout $\varphi \in \mathcal{D}(\mathbb{R})$ telle que $\text{supp}(\varphi) \subset \mathcal{O}$, $\langle T, \varphi \rangle = 0$.

Par exemple, on peut facilement vérifier que δ est nulle sur $\mathbb{R}^*$ et que pour $x \in L^1(\mathbb{R})$, T_x est nulle sur $\mathcal{O}$ équivaut à $x|_{\mathcal{O}} = 0$ presque partout.

On étend à présent le concept de support aux distributions.

Définition 1.2 (Support d'une distribution). *Le support d'une distribution $T \in \mathcal{D}'(\mathbb{R})$ est le complémentaire du plus grand ouvert sur lequel T est nulle.*

Exemple 1.1.

- Si $a_i \in \mathbb{R}$ et $\alpha_i \in \mathbb{C}$, $\text{supp}(\sum_{i=1}^k \alpha_i \delta_{a_i}) = \{a_1, \dots, a_k\}$.
- Si $x \in L^1(\mathbb{R})$, $\text{supp}(T_x) = \text{supp}(x)$.

Définition 1.3 (Distributions à support compact). *On note $\mathcal{E}'(\mathbb{R})$ le sous-espace vectoriel de $\mathcal{D}'(\mathbb{R})$ formé des distributions à support compact.*

$\mathcal{E}'(\mathbb{R})$ contient les distributions de Dirac δ_a , leurs dérivées $\delta_a^{(k)}$ quel que soit $k \in \mathbb{N}$, et l'ensemble des fonctions $x \in L^1(\mathbb{R})$ à support compact (vues comme des distributions régulières). $\mathcal{E}'(\mathbb{R})$ contient en particulier l'espace des fonctions tests : $\mathcal{D}(\mathbb{R}) \subset \mathcal{E}'(\mathbb{R})$.

On a de plus l'emboîtement suivant :

$$\mathcal{E}'(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R}). \quad (25)$$

APPLICATION EN TRAITEMENT DU SIGNAL : Les distributions tempérées forment le cadre le plus général pour la transformée de Fourier. Elle permettent notamment de calculer la transformée de Fourier d'une distribution de Dirac ou d'un peigne de Dirac, mais aussi d'une exponentielle complexe $e_\nu(t) = \exp(i2\pi\nu t)$, qui ne sont ni dans L^1 , ni dans L^2 .

1.5 VARIABLES ET PROCESSUS ALÉATOIRES

Étant donné un espace Ω muni d'une tribu $\mathcal{F}$ et d'une mesure $\mathbf{P}$ telle que $\mathbf{P}(\Omega) = 1$ (on parle d'espace probabilisé), une variable aléatoire X est une application mesurable de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B})$. Une variable aléatoire complexe X_c peut se décomposer comme la somme $X_c = X_r + iX_i$ où X_r et X_i sont des variables aléatoires réelles.

La loi de la variable aléatoire X est la mesure $\mathbf{P}_X$ sur $\mathbb{R}$ telle que $\mathbf{P}_X(A) = \mathbf{P}(X^{-1}(A))$. L'espérance de X est définie par :

$$\begin{aligned} \mathbb{E}(X) &= \int_{\Omega} X(\omega) d\mathbf{P}(\omega) \\ &= \int_{\mathbb{R}} x d\mathbf{P}_X(x), \end{aligned} \quad (26)$$

et sa variance $\text{Var}(X)$ par :

$$\begin{aligned} \text{Var}(X) &= \int_{\Omega} |X(\omega) - \mathbb{E}(X)|^2 d\mathbf{P}(\omega) \\ &= \int_{\mathbb{R}} |x - \mathbb{E}(X)|^2 d\mathbf{P}_X(x), \end{aligned} \quad (27)$$

On peut également caractériser une variable aléatoire par sa fonction de répartition :

$$F_X(x) = \mathbf{P}(X \leq x) \quad (28)$$

et sa fonction caractéristique

$$\varphi_X(t) = \mathbb{E}(\exp(itX)). \quad (29)$$

Si la loi $\mathbf{P}_X(x)$ est absolument continue par rapport à la mesure de Lebesgue (*i.e.*, $d\mathbf{P}_X = f_X d\lambda$), on parle de loi à densité. Un cas très important de loi à densité est la loi normale $\mathcal{N}(\mu, \sigma^2)$, de moyenne μ et de variance σ^2 , définie par :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad (30)$$

Les vecteurs aléatoires sont des variables aléatoires de Ω dans $\mathbb{R}^N$. Leur loi de probabilité est une mesure sur $\mathbb{R}^N$.

Pour un couple de variables aléatoires (X, Y) , on définit la covariance par :

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))^*) \\ &= \int_{\mathbb{R}^2} (x - \mathbb{E}(X))(y - \mathbb{E}(Y))^* d\mathbf{P}_{XY}(x, y). \end{aligned} \quad (31)$$

en notant $\mathbf{X}^*$ la conjuguée hermitienne de $\mathbf{X}$ (*i.e.*, la matrice transposée conjuguée).

La moyenne d'un vecteur aléatoire $\mathbf{X}$ est $\mu_{\mathbf{X}} = \mathbb{E}(\mathbf{X})$ et la matrice de covariance est :

$$\Sigma_{\mathbf{X}} = \mathbb{E}((\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^*), \quad (32)$$

dont les coefficients valent $\Sigma_{\mathbf{X}, nm} = \text{Cov}(X_n, X_m)$.

La matrice de covariance d'un vecteur aléatoire est toujours hermitienne semi-définie positive. En effet, quelque soit le vecteur $\mathbf{a} = (a_j)_{1 \leq j \leq N}$, $\mathbf{a}^* \Sigma_{\mathbf{X}} \mathbf{a} = \text{Var}((a, X)) \geq 0$.

Un vecteur aléatoire gaussien est, par définition, tel que toute combinaison linéaire de ses composantes suit une loi normale. Un vecteur aléatoire gaussien réel $\mathbf{X}$ est défini par sa moyenne μ et sa matrice de covariance Σ . Sa fonction caractéristique est :

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left(i\mu_{\mathbf{X}}^t \mathbf{t} - \frac{1}{2} \mathbf{t}^t \Sigma \mathbf{t}\right) \quad (33)$$

Si de plus Σ est inversible, ce vecteur aléatoire admet une densité donnée par :

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \mu_{\mathbf{X}})^t \Sigma^{-1} (\mathbf{x} - \mu_{\mathbf{X}})\right) \quad (34)$$

La figure 2 donne un exemple de vecteur gaussien. Il est à noter qu'il n'est pas suffisant que les lois marginales des composantes du

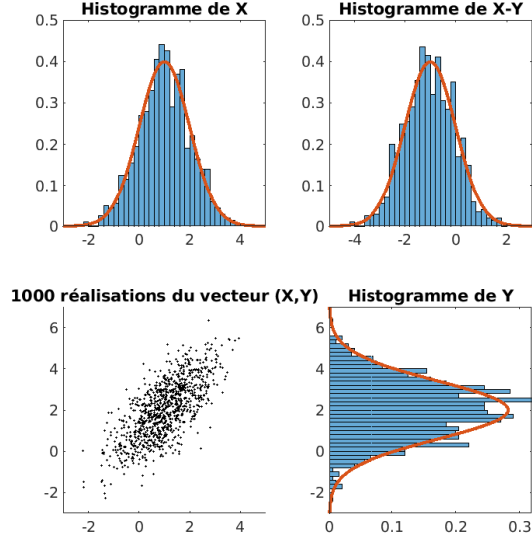


FIGURE 2 – Réalisations d'un vecteur gaussien (X, Y) , de moyenne $\mu = (1, 2)$, de matrice de covariance $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Les lois de toute combinaison linéaire X et Y sont normales.

vecteur soient normales pour que le vecteur soit gaussien, comme le montre l'exemple de la figure 3.

Un *processus aléatoire à temps discret* $(X_n(\omega))$ est une suite de variables aléatoires pour $n \in \mathbb{Z}$. Ce concept généralise celui de vecteurs aléatoires. Un processus est dit gaussien si tout vecteur extrait du processus est gaussien.

Des réalisations du processus $(X_n(\omega))$ (i.e., des valeurs du processus pour des ω donnés) sont aussi appelées *trajectoires*.

La figure 4 montre des réalisations d'un processus aléatoire simple, où les X_k prennent comme valeur -1 et 1, et où $X_{k+1} = X_k$ avec probabilité 9/10, $X_{k+1} = -X_k$ avec probabilité 1/10.

APPLICATION EN TRAITEMENT DU SIGNAL : de nombreux signaux peuvent être modélisés par des processus aléatoires, au sens où il est utile de les considérer comme des réalisations d'un processus aléatoire. Ce type de modélisation est, par exemple, très utile pour des signaux de parole non voisée (chuchotement, consonnes telles que /ch/, /s/, /f/), pour lesquels il est plus informatif de donner les caractéristiques du processus aléatoire les ayant produits, plutôt que les valeurs de leurs trajectoires spécifiques, compte tenu de la variabilité entre les trajectoires. En traitement du signal, on manipule l'autocorrélation d'un signal (entre deux instants n et m), égale à $\mathbb{E}(X_n X_m^*)$, qui coïncide avec l'élément $\Sigma_{X, nm}$ lorsque $\mathbf{X}$ est centré.

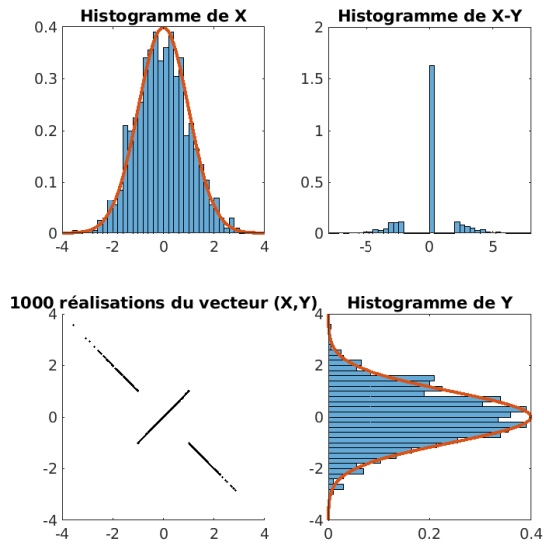


FIGURE 3 – Réalisations d'un vecteur aléatoire (X, Y) , où $Y = X\mathbb{1}_{|X|<1} - X\mathbb{1}_{|X|\geq 1}$, et X est une variable normale de loi $\mathcal{N}(0, 1)$. Les lois marginales de X et Y sont normales, mais $X - Y$ ne suis pas une loi normale. Ce n'est pas donc un vecteur gaussien.

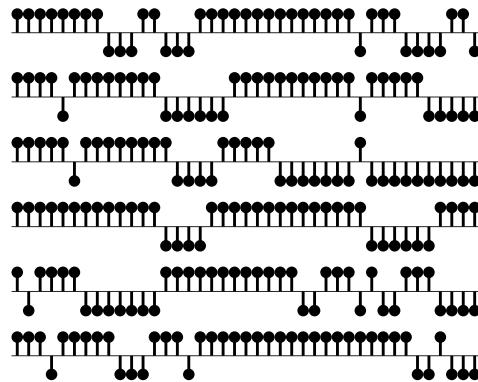


FIGURE 4 – 5 trajectoires d'un processus aléatoire à valeurs -1 ou 1 .

1.5.1 Brefs rappels de statistiques

Le but de la statistique est d'estimer des caractéristiques d'une variable aléatoire (moyenne, variance, paramètre d'une famille de lois de probabilité, etc.) à partir de réalisations de cette variable aléatoire. Par exemple, à partir d'un nombre fini de lancers de dé, on cherche à estimer la probabilité de chaque face du dé.

Voici un exemple simple d'estimation. La moyenne empirique $\hat{\mu}_N$ est un estimateur de la moyenne μ d'une variable aléatoire X . Avec des données X_n indépendantes identiquement distribuées pour $1 \leq n \leq N$,

$$\hat{\mu}_N = \frac{1}{N} \sum_{n=1}^N X_n. \quad (35)$$

Cet estimateur est également une variable aléatoire. On peut calculer son espérance et sa variance :

$$\mathbb{E}(\hat{\mu}_N) = \mathbb{E}(X) \quad (36)$$

$$\text{Var}(\hat{\mu}_N) = \frac{1}{N^2} \sum_{n=1}^N \text{Var}(X_n) \quad (37)$$

$$= \frac{1}{N} \text{Var}(X) \quad (38)$$

Cet estimateur est non biaisé (son espérance est l'espérance de X , quantité à estimer), et consistant (sa variance tend vers 0 quand la taille des données tend vers l'infini).

APPLICATION EN TRAITEMENT DU SIGNAL : comme vu plus haut, il est fréquent de considérer un signal comme la réalisation d'un processus aléatoire. Avec quelques hypothèses supplémentaires (stationnarité, ergodicité), nous verrons qu'il est possible **d'estimer les paramètres d'un processus aléatoire avec une unique réalisation**. Par exemple, on peut estimer les paramètres d'un modèle de production de parole à partir d'un enregistrement unique. Le problème de l'estimation de la fréquence d'une sinusoïde dans du bruit fera également appel à ce formalisme.

SYSTÈMES LINÉAIRES INVARIANTS

Ce chapitre étudie les filtres linéaires et invariants, qui sont des outils fondamentaux en traitement du signal. Après quelques définitions élémentaires pour décrire les signaux à temps continu et à temps discret, on introduit le produit de convolution de deux signaux et liste ses principales propriétés mathématiques. On étudie les propriétés des filtres linéaires convolutifs, et montre finalement que les signaux exponentiels complexes sont fonctions propres d'un filtre linéaire. A ce titre, ils caractérisent le filtre dans le domaine fréquentiel.

2.1 NOTION DE SIGNAL

2.1.1 Définitions

On désigne par signal déterministe, ou plus simplement signal, une fonction d'une ou plusieurs variables de temps ou d'espace, à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ en général. On distingue deux grandes classes de signaux, à temps continu et à temps discret.

Un signal (à temps) continu, appelé *signal analogique*, est une fonction d'une variable continue $t \in \mathbb{R}$. On la notera

$$\begin{aligned} x : \mathbb{R} &\rightarrow \mathbb{C} \\ t &\mapsto x(t) \end{aligned}$$

La variable t peut représenter le temps, mais aussi l'espace ($t = (x, y)$ dans le cas d'une image), la longueur d'onde, etc.

Un signal (à temps) discret est une suite numérique

$$x = \{x[k], k \in \mathbb{Z}\}$$

Très souvent, un signal discret est obtenu en *échantillonnant* les valeurs d'un signal à temps continu à différents instants $t_k, k \in \mathbb{Z}$. On a ainsi :

$$x[k] = x_c(t_k), \forall k \in \mathbb{Z}$$

en notant x_c le signal continu associé au signal discret x . Pour cette raison, les signaux discrets sont aussi appelés signaux échantillonnés.

Notons qu'un **signal numérique** est un signal discret et quantifié, i.e., ne pouvant prendre qu'un nombre fini de valeurs. La *quantification* est l'opération $q = Q(x)$ qui transforme un signal à valeurs

continues ($x[k] \in \mathbb{R}$) en un signal à valeurs discrètes. Par exemple, si q est à valeurs dans $\mathbb{N}$, $q[k]$ est l'entier le plus proche de $x[k]$.

Les signaux déterministes seront vus comme des objets mathématiques qui vivent dans des espaces fonctionnels comme $L^p(\mathbb{R})$ (temps continu) ou $\ell^p(\mathbb{Z})$ (temps discret). Par opposition aux signaux déterministes, on définit les *signaux stochastiques* comme des processus aléatoires. Ces représentations seront étudiées aux chapitres 5 et 6.

2.1.2 Représentations énergétiques

Soit x un signal déterministe à temps continu. On définit, quand elles existent, les grandeurs suivantes :

- La puissance moyenne $P_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} |x(t)|^2 dt$
- L'énergie $E_x = \int_{\mathbb{R}} |x(t)|^2 dt = \|x\|_{L^2}^2$.

Un signal d'énergie finie vérifie donc $x \in L^2(\mathbb{R})$, et $P_x = 0$. Des définitions analogues s'appliquent dans le cas discret, avec $E_x = \|x\|_2^2$.

On définit l'*intercorrélation entre deux signaux d'énergie finie* par :

$$\forall \tau \in \mathbb{R}, \gamma_{xy}(\tau) = \int_{\mathbb{R}} x(t) y^*(t - \tau) dt \quad (\text{cas continu}) \quad (39)$$

$$\forall n \in \mathbb{N}, \gamma_{xy}[n] = \sum_k x[k] y^*[k - n] \quad (\text{cas discret}) \quad (40)$$

L'*intercorrélation* est égale au produit scalaire entre x et la version retardée de y :

$$\gamma_{xy}(\tau) = (x, y(\cdot - \tau)) \quad (\text{cas continu})$$

$$\gamma_{xy}[n] = (x, y[\cdot - n]) \quad (\text{cas discret})$$

On définit le *Rapport Signal-Sur-Bruit* (RSB, ou SNR pour *Signal-to-Noise Ratio*) pour quantifier le niveau de perturbation d'un signal lors de sa transmission dans un réseau de communication. En notant u le signal utile (*i.e.*, non bruité) et $x = u + b$ le signal bruité (en supposant le bruit b additif), le SNR est respectivement défini par :

$$\text{SNR} = \left(\frac{P_u}{P_b} \right) \quad \text{ou} \quad \text{SNR} = \left(\frac{E_u}{E_b} \right) \quad (41)$$

pour des signaux de puissance et d'énergie finie. Le SNR est le plus souvent exprimé en dB :

$$\text{SNR}_{\text{dB}} = 10 \log_{10} (\text{SNR}). \quad (42)$$

2.2 CONVOLUTION

La convolution est une opération fondamentale en traitement du signal. Nous la définissons successivement dans les espaces L^p , pour les signaux discrets (espaces ℓ^p), et pour des distributions.

2.2.1 Cas continu : convolutions de fonctions

Définition 2.1 (Convolution de deux fonctions). Soient deux fonctions $x : \mathbb{R} \rightarrow \mathbb{C}$ et $y : \mathbb{R} \rightarrow \mathbb{C}$. Le produit de convolution de x et y , noté $x * y$, est la fonction, si elle est définie, définie par :

$$\forall t \in \mathbb{R}, (x * y)(t) = \int_{\mathbb{R}} x(\theta)y(t - \theta)d\theta \quad (43)$$

$$= \int_{\mathbb{R}} x(t - \theta)y(\theta)d\theta. \quad (44)$$

$(x * y)(t)$ est parfois noté $x * y(t)$. Dans ce polycopié, on privilégie l'écriture $(x * y)(t)$ pour bien marquer que $(x * y)(t)$ dépend de toutes les valeurs de $y(\theta)$.

On remarque que la convolution n'est pas toujours définie. Par exemple, si $x = y = \mathbb{1}_{\mathbb{R}}$, alors $\forall t, (x * y)(t) = \infty$. Par ailleurs, la convolution a un pouvoir régularisant (cf. cours de CIP). Par exemple, pour les fonctions discontinues $x = y = \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}$, où $T > 0$, $x * y$ est la fonction "triangle" (continue) de largeur $2T$ et d'amplitude T :

$$(x * y)(t) = (T - |t|) \mathbb{1}_{[-T, T]}(t). \quad (45)$$

La convolution est définie partout (éventuellement presque partout) en fonction de l'espace dans lequel vivent les signaux x et y .

Théorème 2.1. La convolution de deux fonctions de L^p et L^q est définie dans les cas suivants :

x	y	$x * y$	définie
L^1	L^1	L^1	<i>p.p.</i>
L^1	L^2	L^2	<i>p.p.</i>
L^1	L^∞	L^∞	<i>partout</i>
L^2	L^2	L^∞	<i>partout</i>

La norme de $x * y$ est bornée par le produit des normes de x et y .

Les démonstrations sont élémentaires et laissées en exercice.

Pour aller plus loin...

Le théorème 2.1 est généralisé par l'inégalité de Young. Avec $p, q, r \in [1, \infty]$ et

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r},$$

la convolution de $x \in L^p$ et de $y \in L^q$ est dans L^r , avec la borne

$$\|x * y\|_r \leq \|x\|_p \|y\|_q$$

Remarque. L'expression (43) ressemble à la définition de l'intercorrélation (39). Pour x et $y \in L^2(\mathbb{R})$, $\gamma_{xy}(\tau)$ est égal au produit scalaire entre x et le signal retardé $y^*(\cdot - \tau)$. En notant $\bar{y}(t) = y^*(-t)$, on a

$$\gamma_{xy} = x * \bar{y} \quad (46)$$

2.2.2 Cas discret

Définition 2.2 (Convolution dans le cas discret). Soient $x = \{x[k], k \in \mathbb{Z}\}$ et $y = \{y[k], k \in \mathbb{Z}\}$. Le produit de convolution de x et y est la suite $x * y$, si elle est définie, définie par :

$$\forall n \in \mathbb{Z}, (x * y)[n] = \sum_{k \in \mathbb{Z}} x[k] y[n - k] \quad (47)$$

$$= \sum_{k \in \mathbb{Z}} x[n - k] y[k]. \quad (48)$$

$(x * y)[n]$ est parfois noté $x * y[n]$. On privilégie l'écriture $(x * y)[n]$.

Les conditions d'existence sont similaires au cas continu. On peut donc écrire la convolution de deux signaux $\ell^1(\mathbb{Z})$, de deux signaux dans $\ell^2(\mathbb{Z})$, d'un signal $\ell^1(\mathbb{Z})$ avec un signal $\ell^\infty(\mathbb{Z})$, etc.

2.2.3 Signaux modélisés par des distributions

Il est nécessaire d'étendre la convolution aux distributions pour pouvoir définir certains filtres comme les filtres à retard (55) $y = \delta_\tau * x$. Le lecteur est redirigé vers le chapitre 1 pour une définition des espaces de distributions sur lesquels nous travaillons.

Convolution d'une fonction C^∞ et d'une distribution

On définit d'abord la convolution d'une fonction $C^\infty(\mathbb{R})$ avec une distribution comme extension directe de la définition (44).

Définition 2.3. Soit $x \in C^\infty(\mathbb{R})$ et $Y \in \mathcal{D}'(\mathbb{R})$. On suppose que l'une des trois hypothèses suivantes sont satisfaites :

1. $x \in \mathcal{D}(\mathbb{R})$ et $Y \in \mathcal{D}'(\mathbb{R})$;
2. $x \in \mathcal{S}(\mathbb{R})$ et $Y \in \mathcal{S}'(\mathbb{R})$;
3. $x \in C^\infty(\mathbb{R})$ et $Y \in \mathcal{E}'(\mathbb{R})$.

On définit la fonction $x * Y$ par

$$\forall t \in \mathbb{R}, (x * Y)(t) = \langle Y, x(t - \cdot) \rangle \quad (49)$$

où $x(t - \cdot)$ désigne la fonction $\theta \mapsto x(t - \theta)$.

Théorème 2.2. Sous les hypothèses précédentes, $x * Y \in C^\infty(\mathbb{R})$. De plus, pour tout $k \in \mathbb{N}$, $(x * Y)^{(k)} = x^{(k)} * Y = x * Y^{(k)}$.

Exemple 2.1. Pour $Y = \delta_a$ et $x \in C^\infty(\mathbb{R})$, (49) implique :

$$\forall t \in \mathbb{R}, (x * \delta_a)(t) = x(t - a). \quad (50)$$

Pour aller plus loin...

Définition 2.4 (Convolution de distributions). Soient $X \in \mathcal{E}'(\mathbb{R})$ et $Y \in \mathcal{D}'(\mathbb{R})$. Le produit de convolution $X * Y$ est la distribution $X * Y$ définie par :

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \langle X * Y, \phi \rangle = \langle X, \theta \mapsto \langle Y, \phi(\cdot + \theta) \rangle \rangle \quad (51)$$

$$= \langle Y, \theta \mapsto \langle X, \phi(\cdot + \theta) \rangle \rangle \quad (52)$$

Cette définition possède un lien direct avec la définition (43)-(44) de la convolution de deux fonctions x et y . Soient $x \in L^1(\mathbb{R})$ et $y \in L^1(\mathbb{R})$. D'après le théorème 2.1, $x * y \in L^1(\mathbb{R})$. La distribution régulière associée s'écrit :

$$\begin{aligned} \forall \phi \in \mathcal{D}(\mathbb{R}), \langle T_{x*y}, \phi \rangle &= \int_{\mathbb{R}} (x * y)(t) \phi(t) dt \\ &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} x(\theta) y(t - \theta) d\theta \right] \phi(t) dt \\ &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} y(t - \theta) \phi(t) dt \right] x(\theta) d\theta \\ &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} y(x) \phi(u + \theta) du \right] x(\theta) d\theta \\ &= \int_{\mathbb{R}} x(\theta) \langle T_y, \phi(\cdot + \theta) \rangle d\theta \end{aligned} \quad (53)$$

en utilisant le théorème de Fubini pour intervertir les deux intégrales. (53) s'identifie clairement avec (51). De façon similaire, on peut établir un lien direct entre (52) et (44).

Remarque. Le produit de convolution de deux distributions n'est donc pas toujours défini. En raison de l'emboîtement $\mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}'$, la définition (avec $X \in \mathcal{E}'$, $Y \in \mathcal{D}'$) couvre les cas suivants :

- $X \in \mathcal{D}$, $Y \in \mathcal{D}'$;
- X et $Y \in \mathcal{E}'$. On a alors $X * Y \in \mathcal{E}'$;
- $X \in \mathcal{E}'$ et $Y \in \mathcal{S}'$. On a alors $X * Y \in \mathcal{S}'$;

Exemple 2.2.

- Pour $X = \delta_a \in \mathcal{E}'(\mathbb{R})$, on a pour $Y \in \mathcal{D}'(\mathbb{R})$,

$$\forall \phi \in \mathcal{D}(\mathbb{R}), \langle \delta_a * Y, \phi \rangle = \langle Y, \phi(\cdot + a) \rangle.$$

Pour $a = 0$, on en déduit que $\delta * Y = Y$.

- Dans le cas particulier où $Y = y \in C^0(\mathbb{R})$, on a :

$$\langle \delta_a * y, \phi \rangle = \int_{\mathbb{R}} y(t) \phi(t + a) dt = \int_{\mathbb{R}} y(t - a) \phi(t) dt$$

On retrouve le résultat (50) : $\forall t, (\delta_a * y)(t) = y(t - a)$.

2.2.4 Propriétés du produit de convolution

Les propriétés sont énoncées quelque soit la nature des signaux x , y et z sous réserve que les quantités manipulées sont bien définies.

- **Commutativité** : $x * y = y * x$
- **Associativité** : $x * (y * z) = (x * y) * z$
- **Distributivité** : $x * (y + z) = (x * y) + (x * z)$
- **Élément neutre (cas continu)** : distribution de Dirac $\delta_0 = \delta$:
 $\delta * y = y$.
- **Élément neutre (cas discret)** : fonction de Kronecker $\delta[k]$:
 $\delta * y = y$.
- **Translation** : $x(t - \tau) = (x * \delta_\tau)(t)$
- **Dérivation (éventuellement au sens des distributions)** : $\frac{d}{dt}(x * y) = \frac{dx}{dt} * y = x * \frac{dy}{dt}$
- **Support (éventuellement au sens des distributions)** : $\text{supp}(x * y) \subset \text{supp}(x) + \text{supp}(y)$.
- **Intercorrélation** : $\gamma_{xy} = x * \bar{y}$ avec $\bar{y}(t) = y^*(-t)$.

2.3 NOTION DE SYSTÈME

2.3.1 Définition

Un système est défini comme une application $y = S(x)$ de l'espace des signaux d'entrée x vers celui des signaux de sortie y . Les espaces considérés sont typiquement les espaces L^p pour des signaux continus et ℓ^p pour des signaux discrets. Un système est représenté symboliquement par le schéma "boîte noire" de la figure 5.

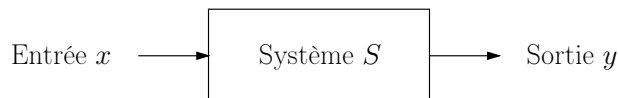


FIGURE 5 – Représentation d'un système.

Un système modélise des transformations subies par un signal. Il peut s'agir de l'altération d'un signal physique (signal audio, onde électromagnétique) lors de sa transmission à travers un milieu de propagation, mais aussi d'un traitement visant à améliorer le contenu du signal (débruitage d'un enregistrement sonore, déconvolution pour retirer le flou dans une image). Dans ce cours, on ne considère que les systèmes mono entrée - mono sortie.

2.3.2 Propriétés des systèmes

Définition 2.5 (Homogénéité). *Un système est dit homogène en entrée/sortie, si l'entrée et la sortie sont des signaux de même nature (continue ou discrète).*

Définition 2.6 (Linéarité). *Un système S est linéaire si pour toute combinaison linéaire*

$$x = a_1 x_1 + a_2 x_2, \quad a_1, a_2 \in \mathbb{R},$$

on a

$$S(x) = a_1 S(x_1) + a_2 S(x_2).$$

Ainsi, le signal de sortie $S(x)$ s'exprime comme la superposition des signaux $S(x_1)$ et $S(x_2)$ avec les mêmes poids a_1 et a_2 que dans la décomposition $x = a_1 x_1 + a_2 x_2$.

Remarque. Les systèmes physiques sont très rarement linéaires. Un modèle S est une équation qui représente le comportement attendu d'un phénomène conformément aux lois de la physique. On procède souvent à une linéarisation autour d'un point dit de fonctionnement (ou point d'équilibre) du système pour aboutir à un modèle linéaire.

Les définitions qui suivent sont données dans le cas de signaux à temps continu. Des définitions identiques s'appliquent aux signaux à temps discret.

On rappelle que $(\delta_\tau * x)(t) = x(t - \tau)$ est un décalage dans le temps du signal $x(t)$, aussi appelé translation. Pour $\tau > 0$, le signal $\delta_\tau * x$ est une version retardée de $x(t)$.

Définition 2.7 (Invariance par translation). *Un système S est dit invariant par translation si pour toute entrée x et pour tout $\tau \in \mathbb{R}$,*

$$S(\delta_\tau * x) = (S(x)) * \delta_\tau. \quad (54)$$

Autrement dit, la sortie du filtre associée au signal retardé coïncide avec la version retardée du signal de sortie $S(x)$, avec le même retard.

Définition 2.8 (Filtre linéaire homogène). *Un filtre linéaire est un système qui vérifie les deux propriétés de linéarité et d'invariance par translation.*

La convolution $y = h * x$ est un cas important de filtrage linéaire, où h (appelé réponse impulsionnelle) caractérise le filtre. On peut vérifier à partir des propriétés du produit de convolution (section 2.2.4),

que ce système est linéaire (trivial) et invariant par translation, car la sortie associée à un signal retardé s'écrit :

$$h * (\delta_\tau * x) = (h * x) * \delta_\tau.$$

Deux exemples remarquables s'obtiennent dans les cas où $h = \delta_\tau$ (système à retard) et $h = \mathbb{1}_{\mathbb{R}_+}$ (intégrateur).

Exemple 2.3.

— *Retard* : $y = \delta_\tau * x$ ou $y = \delta[\cdot - n] * x$:

$$y(t) = x(t - \tau), \quad \text{ou} \quad y[k] = x[k - n]. \quad (55)$$

— *Intégrateur* : $y = \mathbb{1}_{\mathbb{R}_+} * x$ ou $y = \mathbb{1}_{\mathbb{N}} * x$:

$$y(t) = \int_{-\infty}^t x(\tau) d\tau, \quad \text{ou} \quad y[k] = \sum_{n=-\infty}^k x[n] \quad (56)$$

Définition 2.9 (Système causal). Un système est causal si sa réponse à un instant donné ne dépend que des valeurs de l'entrée aux instants précédents (éventuellement de la valeur à l'instant présent).

Exemple 2.4. L'intégrateur défini en (56) est un système causal. En revanche, le calcul de la moyenne dans un intervalle centré

$$y(t) = \frac{1}{2T} \int_{t-T}^{t+T} x(\theta) d\theta \quad (57)$$

n'est pas causal, car il faut connaître la valeur de l'entrée $x(\theta)$ pour $\theta > t$.

Si la causalité semble naturelle pour des systèmes physiques (l'effet ne précède pas à la cause), il faut envisager l'existence de systèmes non causaux dans le cas de traitement en temps différé (*offline*). Par exemple, lorsque la totalité d'un signal a été enregistré ou stocké dans une mémoire tampon, on peut effectuer *a posteriori* des calculs de moyenne comme (57) sur une fenêtre centrée autour d'un instant t .

La stabilité d'un système peut être caractérisée de diverses manières. Nous retenons la définition au sens Entrée Bornée, Sortie Bornée (EBSB).

Définition 2.10 (Stabilité Entrée Bornée, Sortie Bornée). Le système S est stable si pour toute entrée $x \in L^\infty(\mathbb{R})$, $S(x) \in L^\infty(\mathbb{R})$.

2.4 FILTRES LINÉAIRES CONVOLUTIFS

Comme nous l'avons vu, un filtre linéaire est un système linéaire et invariant (SLI), que nous supposons homogène en entrée/sortie. On s'intéresse aux filtres définis par une convolution.

2.4.1 Définition

Définition 2.11 (Réponse impulsionnelle). Soit S un filtre linéaire dont la relation entrée-sortie est associée à une convolution :

$$y = S(x) = h * x, \quad (58)$$

(sous réserve que $h * x$ soit bien défini). Le signal h est appelé *réponse impulsionnelle du filtre*. C'est possiblement une distribution.

Comme la distribution de Dirac $\delta = \delta_0$ (contexte continu) ou la fonction de Kronecker $\delta[k]$ (contexte discret) est l'élément neutre de la convolution, on a

$$h = S(\delta), \quad (59)$$

D'où le nom de réponse impulsionnelle : h est la réponse du filtre à une impulsion (de Dirac ou de Kronecker suivant le contexte).

Interprétations du calcul

Prenons l'exemple d'un système discret et d'une entrée impulsionnelle ($x[k] = \delta[k]$). La sortie $y[k] = h[k]$ s'identifie avec la réponse impulsionnelle, comme le montre la figure 6.

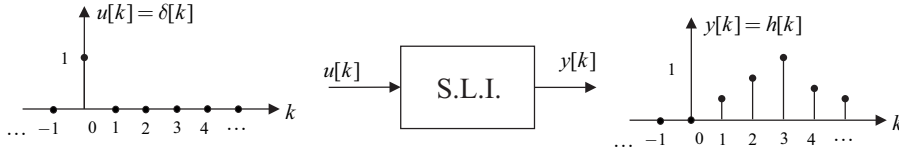


FIGURE 6 – Réponse impulsionnelle d'un SLI

Considérons maintenant une entrée quelconque $x[k]$. En connaissant la réponse impulsionnelle, la sortie

$$y[k] = (h * x)[k] = \sum_{n \in \mathbb{Z}} x[n] h[k - n]$$

s'obtient comme combinaison linéaire des versions retardées de h (les signaux $k \mapsto h[k - n]$) pondérées par les amplitudes des impulsions $x[n]$. Comme le produit de convolution est commutatif, on a aussi

$$y[k] = \sum_{n \in \mathbb{Z}} h[n] x[k - n]$$

$h[n]$ peut-être interprétée comme une séquence de pondération : le terme $h[n]$ pondère la valeur de l'entrée passée de n instants (c'est-à-dire $x[k - n]$) pour contribuer au calcul de y à l'instant k .

Stabilité

La stabilité EBSB du filtre est garantie si

$$\begin{aligned} h &\in L^1(\mathbb{R}) && (\text{cas continu}) \\ h &\in \ell^1(\mathbb{Z}) && (\text{cas discret}) \end{aligned}$$

En effet, d'après le théorème 2.1, ces conditions impliquent que si x est borné ($x \in L^\infty(\mathbb{R})$, respectivement $\ell^\infty(\mathbb{Z})$), alors $h * x$ est borné.

Causalité

En écrivant la définition de la convolution (ici dans le cas continu) :

$$y(t) = (h * x)(t) = \int_{\mathbb{R}} h(\theta)x(t - \theta)d\theta$$

il apparaît que le filtre est causal si et seulement si $h(t) = 0$ pour $t < 0$ (respectivement $h[k] = 0$ pour $k < 0$), c'est-à-dire

$$\begin{aligned} \text{supp}(h) &\subset \mathbb{R}_+ && (\text{cas continu}) \\ \text{supp}(h) &\subset \mathbb{N} && (\text{cas discret}) \end{aligned}$$

La relation entrée-sortie s'écrit alors

$$y(t) = \int_{\mathbb{R}_+} h(\theta)x(t - \theta) d\theta = \int_{]-\infty, t]} h(t - \theta)x(\theta) d\theta.$$

et prend la forme

$$y(t) = \int_{[0, t]} h(\theta)x(t - \theta) d\theta = \int_{[0, T]} h(t - \theta)x(\theta) d\theta.$$

si $x(t)$ est nulle pour $t < 0$. Des expressions similaires s'obtiennent dans le cas discret.

Remarque. Pour un système discret stable, $h[n]$ tend vers 0 quand $|n|$ tend vers l'infini : une valeur de l'entrée passée très éloignée de l'instant k contribue peu à la valeur de la sortie à cet instant.

2.4.2 Filtres à réponse impulsionnelle finie (FIR)

Définition 2.12. On appelle filtre **FIR** (*Finite Impulse Response*) un filtre dont le support de la réponse impulsionnelle est fini.

Exemple 2.5 (Filtre dérivateur). *Ce filtre réalise la différence entre deux points consécutifs, d'où son nom de dérivateur. La sortie y du filtre est :*

$$y[k] = x[k] - x[k - 1] \quad (60)$$

En identifiant cette équation avec

$$y[k] = \sum_{n \in \mathbb{Z}} h[n]x[k - n]$$

on obtient la réponse impulsionnelle :

$$h[n] = \begin{cases} 1 & \text{si } n = 0, \\ -1 & \text{si } n = 1, \\ 0 & \text{sinon.} \end{cases}$$

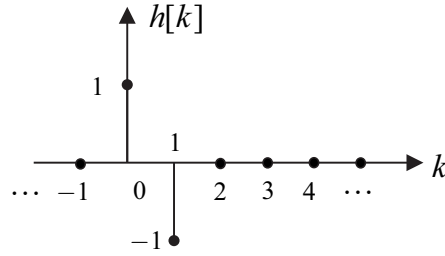


FIGURE 7 – Réponse impulsionnelle d'un filtre dérivateur.

Exemple 2.6 (Filtre à moyenne glissante). *Ce filtre calcule la valeurs du signal sur une fenêtre de $2K + 1$ points centrée sur le point suivant :*

$$y[k] = \frac{1}{2K + 1} \sum_{n=-K}^K x[k - n] \quad (61)$$

*Le filtre n'est pas causal car $y[k]$ dépend d'entrées postérieures à l'instant k . En identifiant (61) avec $(h * x)[k]$, il vient :*

$$h[n] = \begin{cases} 1 & \text{si } |n| \leq K, \\ 0 & \text{sinon.} \end{cases}$$

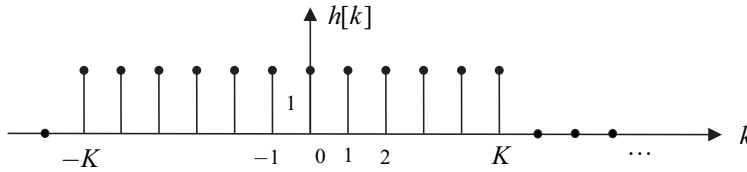


FIGURE 8 – Réponse impulsionnelle d'un filtre moyennneur.

2.4.3 Filtre à réponse impulsionnelle infinie (IIR)

Définition 2.13. On appelle **filtre IIR (Infinite Impulse response)** un filtre dont la réponse impulsionnelle est à support infini.

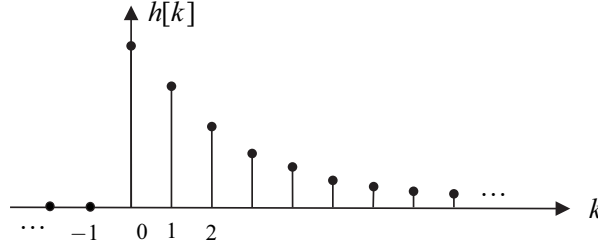


FIGURE 9 – Réponse impulsionnelle d'un filtre IIR causal.

Exemple 2.7. Lissage exponentiel causal

La figure 9 montre la réponse impulsionnelle d'un filtre pour lequel $h[n] = \alpha^n$ avec $0 < \alpha < 1$. La relation entrée-sortie s'écrit :

$$y[k] = \sum_{n=0}^{\infty} \alpha^n x[k-n] \quad (62)$$

2.4.4 Équations aux différences

La relation entrée/sortie d'un filtre peut-être décrite par une convolution, mais aussi par une équation réursive entre les sorties à différents instants. Prenons une équation aux dérivées ordinaires du type :

$$y(t) - \alpha_1 y'(t) - \dots - \alpha_p y^{(p)}(t) = \beta_0 x(t) + \dots + \beta_m x^{(m)}(t), \quad t \in \mathbb{R}_+$$

(avec p conditions initiales à $t = 0$). Sa discrétisation par la méthode de différences finies conduit à une équation du type :

$$y[k] = \sum_{i=1}^p a_i y[k-i] + \sum_{j=0}^m b_j x[k-j] \quad (63)$$

où les paramètres a_i et b_j sont indépendants de k . Cette équation est appelée **équation aux différences finies d'ordre p** .

Comme la réponse impulsionnelle h coïncide avec la sortie du filtre pour l'entrée $x[k] = \delta[k]$, il vient

$$\forall k > m, \quad h[k] = \sum_{i=1}^p a_i h[k-i] \quad (64)$$

$$\forall k \in \{0, \dots, m\}, \quad h[k] = \sum_{i=1}^p a_i h[k-i] + b_k. \quad (65)$$

En supposant le système causal, ces équations permettent de déduire h en résolvant un système linéaire.

2.5 FONCTIONS PROPRES

En dimension finie, les valeurs et vecteurs propres d'une application linéaire, donnent des informations capitales sur son action sur un vecteur. Le théorème suivant montre que tous les systèmes linéaires invariants stables admettent les signaux exponentiels complexes comme fonctions propres, i.e., telles que la sortie du filtre pour ces fonctions soit proportionnelle à l'entrée.

Théorème 2.3. Les exponentielles complexes définies par

$$e_f(t) = \exp(i2\pi ft)$$

sont fonctions propres des systèmes linéaires invariants stables.

Si le système admet une réponse impulsionnelle $h \in L^1$, les valeurs propres correspondantes sont données par la transformée de Fourier de h , appelée réponse fréquentielle :

$$H(f) = \int_{\mathbb{R}} h(t) \exp(-i2\pi ft) dt \quad (66)$$

Théorème 2.4. Les exponentielles complexes discrètes définies par

$$e_v[k] = \exp(i2\pi vk)$$

sont fonctions propres des systèmes linéaires invariants discrets stables.

Si le système admet une réponse impulsionnelle $h \in \ell^1$, les valeurs propres correspondantes sont données par la transformée de Fourier à temps discret de h :

$$H(v) = \sum_{k \in \mathbb{Z}} h[k] \exp(-i2\pi vk) \quad (67)$$

Démonstration. (Cas continu) Soit $\tau \in \mathbb{R}$. On remarque que $e_f * \delta_{-\tau} = \exp(i2\pi f\tau)e_f$. Par invariance temporelle et linéarité de S , on a donc :

$$S(e_f) * \delta_{-\tau} = S(e_f * \delta_{-\tau}) = \exp(i2\pi f\tau)S(e_f).$$

On en déduit que

$$\forall \tau, S(e_f)(0 + \tau) = \exp(i2\pi f\tau) S(e_f)(0)$$

c'est-à-dire $S(e_f) = S(e_f)(0)e_f$. Donc e_f est une fonction propre de S .

Dans le cas où S admet une réponse impulsionnelle $h \in L^1$, la valeur propre associée est donnée par :

$$S(e_f)(0) = (e_f * h)(0) = \int_{\mathbb{R}} h(t) \exp(-i2\pi ft) dt$$

□

On remarquera que dans le cas discret, les signaux e_ν et $e_{\nu+n}$ sont identiques pour $n \in \mathbb{Z}$, et que $H(\nu)$ est périodique de période 1.

Pour aller plus loin...

La réponse fréquentielle d'un filtre peut être mesurée en faisant agir le filtre sur une exponentielle complexe, et en mesurant le facteur selon lequel elle a été amplifiée ou atténuée, et déphasée. C'est en pratique impossible, une exponentielle complexe ayant un support infini, en particulier vers les temps négatifs. Cependant, dans le cas d'un filtre stable (et causal) de réponse impulsionnelle $h \in L^1$, la réponse à une exponentielle complexe limitée aux temps positifs $e_f \mathbb{1}_{\mathbb{R}_+}$ tend vers la réponse à l'exponentielle complexe e_f quand t tend vers l'infini.

En effet, la différence entre les deux réponses est donnée par :

$$S(e_f)(t) - S(e_f \mathbb{1}_{\mathbb{R}_+})(t) = S(e_f \mathbb{1}_{\mathbb{R}_-})(t) = \int_{\mathbb{R}} e_f(t - \tau) h(\tau) \mathbb{1}_{t, +\infty}(\tau) d\tau$$

qui tend vers 0 quand $t \rightarrow +\infty$ par convergence dominée.

On peut donc mesurer la réponse fréquentielle en introduisant une exponentielle complexe à partir de $t = 0$ dans le système, et en attendant l'extinction du régime transitoire.

2.6 CONCLUSION

Les liens entre systèmes linéaires invariants et transformée de Fourier ont une importance capitale en traitement du signal : ils justifient l'analyse fréquentielle des systèmes (par exemple, tracé de diagrammes de Bode, qui n'est rien d'autre que la détermination des valeurs propres d'un système). En effet, la connaissance de la réponse fréquentielle $H(f)$ d'un système permet, par transformée de Fourier inverse, de reconstituer sa réponse impulsionnelle, et donc de prédire son action sur n'importe quel signal d'entrée. Le chapitre 3 rappelle les propriétés importantes de la transformée de Fourier utiles à l'analyse des systèmes linéaires invariants.

Cependant, en temps continu, le lien entre réponse impulsionnelle et réponse fréquentielle n'est donné ici que quand la réponse impulsionnelle est une fonction intégrable. Or, de nombreux filtres admettent comme réponse impulsionnelle une distribution (comme le cas élémentaire des retards, dont la réponse impulsionnelle est une distribution de Dirac translatée). L'extension de la transformée de Fourier aux distributions tempérées, permettant l'analyse fréquentielle de ces systèmes, est également introduite dans le chapitre 3.

TRANSFORMÉES DE FOURIER

Dans ce chapitre, nous introduisons les transformées de Fourier, des transformées linéaires appliquées à des signaux à temps continu et temps discret. Cet outil permet de représenter les signaux dans le domaine fréquentiel. Par exemple, le contenu de signaux audio s'interprète aisément si les signaux sont exprimés dans le domaine fréquentiel. De plus, la représentation fréquentielle simplifie certains traitements. En particulier, une convolution dans le domaine temporel se traduit par un simple produit dans le domaine fréquentiel.

Le principe de l'analyse fréquentielle d'un signal est de rechercher une décomposition du type

$$x(t) = \int_{\mathbb{R}} X(f) \exp(2i\pi ft) df \quad (68)$$

qu'on peut interpréter comme une décomposition du signal x sur l'espace des signaux exponentiels $e_f : t \mapsto \exp(2i\pi ft)$. (68) généralise au cas de signaux non périodiques la décomposition en série de Fourier, bien connue dans le cas de signaux périodiques.

Organisation du chapitre

On présentera successivement les quatre transformées de Fourier :

- La transformée de Fourier à temps continu, pour un signal $x(t)$ à temps continu, de support infini.
- La décomposition en série de Fourier, pour un signal à temps continu et de support compact (ou pour un signal périodique).
- La transformée de Fourier à temps discret, pour un signal $x[k]$ à temps discret et de support infini.
- La transformée de Fourier discrète, pour un signal à temps discret de support fini, représenté comme un vecteur $\mathbf{x} \in \mathbb{C}^N$.

On présentera enfin des applications de ces résultats en traitement du signal dans le cadre du filtrage linéaire et de l'analyse spectrale.

3.1 TRANSFORMÉE DE FOURIER À TEMPS CONTINU

La transformée de Fourier (TF) d'un signal $x : \mathbb{R} \rightarrow \mathbb{C}$, $t \mapsto x(t)$ est la fonction $\mathcal{F}x : \mathbb{R} \rightarrow \mathbb{C}$ définie¹ par :

$$\forall f \in \mathbb{R}, \mathcal{F}x(f) = \int_{\mathbb{R}} e^{-i2\pi ft} x(t) dt \quad (69)$$

De même, la transformée de Fourier conjuguée d'un signal x est la fonction $\mathcal{F}^*x : \mathbb{R} \rightarrow \mathbb{C}$ définie par :

$$\forall f \in \mathbb{R}, \mathcal{F}^*x(f) = \int_{\mathbb{R}} e^{i2\pi ft} x(t) dt \quad (70)$$

Le module $|\mathcal{F}x(f)|$ est le plus souvent utilisé pour donner une représentation graphique de la transformée de Fourier, appelée **spectre du signal x** . Cependant, deux signaux peuvent avoir une TF de même module en étant sensiblement différents dans le domaine temporel.

Notations. $\mathcal{F}x(f)$ doit être compris comme $(\mathcal{F}x)(f)$. S'il est nécessaire d'expliciter la dépendance de x par rapport à t , on utilisera $\mathcal{F}\{x(t)\}(f)$. On utilisera très souvent la notation courte $X := \mathcal{F}x$. Ainsi,

$$\forall f \in \mathbb{R}, X(f) = \int_{\mathbb{R}} e^{-i2\pi ft} x(t) dt \quad (71)$$

Nous commençons par définir la transformée de Fourier dans $L^1(\mathbb{R})$, avant d'aborder les extensions à l'espace $L^2(\mathbb{R})$ et à l'espace $\mathcal{S}'(\mathbb{R})$ des distributions tempérées. Le lecteur est redirigé vers le polycopié de CIP pour les preuves des théorèmes.

3.1.1 Transformée de Fourier des signaux dans L^1

La définition de la TF est la plus simple pour $x \in L^1(\mathbb{R})$, car l'intégrale (69) est définie en tout point f . Notons que pour $f = 0$, on a :

$$X(0) = \int_{\mathbb{R}} x(t) dt. \quad (72)$$

Notons également que si x est à valeurs réelles, alors X est à symétrie hermitienne : $\forall f, X(-f) = X^*(f)$.

Théorème 3.1. Soit $x \in L^1$. Sa transformée de Fourier satisfait

- $X \in L^\infty$;
- $X \in \mathcal{C}^0$;
- (Riemann-Lebesgue) $\lim_{|f| \rightarrow \infty} X(f) = 0$.
- Si $x \in L^1$ et $\text{supp}(x)$ est borné, alors $X \in \mathcal{C}^\infty(\mathbb{R})$

1. On trouve parfois la définition alternative : $X(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp(-i\omega t) x(t) dt$.

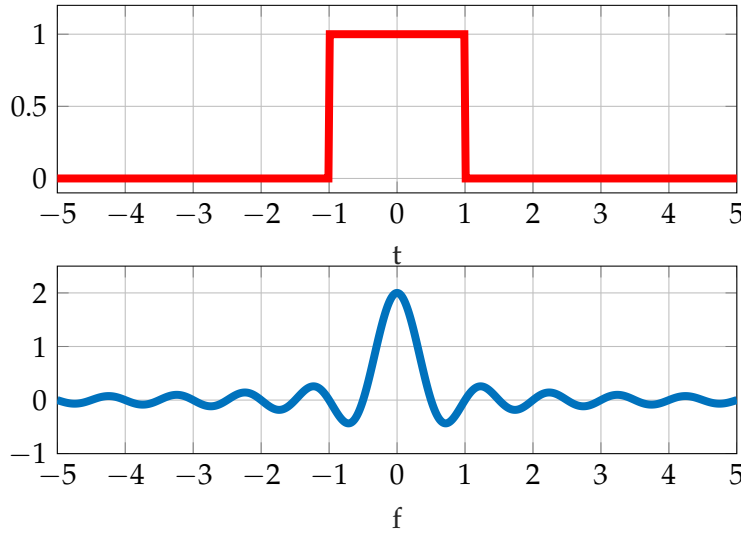


FIGURE 10 – Fenêtre rectangulaire de largeur 1 et sa TF. La TF est un sinus cardinal qui s’annule pour $f \in \mathbb{Z} \setminus \{0\}$.

D’après le théorème précédent, la TF d’un signal dans L^1 est continue et bornée. En revanche, X n’est pas nécessairement dans L^1 , comme le montre l’exemple ci-dessous.

Exemple 3.1 (Fenêtre rectangulaire). Soit le signal $x = \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}$. On calcule

$$X(f) = \int_{[-\frac{T}{2}, \frac{T}{2}]} e^{-i2\pi ft} dt = T \operatorname{sin}_c(T\pi f) \quad (73)$$

où la fonction sinus cardinal est définie par

$$\operatorname{sin}_c(\theta) := \begin{cases} \frac{\sin(\theta)}{\theta} & \text{si } \theta \neq 0 \\ 1 & \text{si } \theta = 0 \end{cases} \quad (74)$$

La TF de la fenêtre rectangulaire paire est donc à valeurs réelles, mais n’est pas dans $L^1(\mathbb{R})$. Elle s’annule pour $f = \frac{k}{T}$, $k \in \mathbb{Z}^*$. La valeur en zéro coïncide avec son intégrale, cf. (72), égale à T .

Propriétés

Les principales propriétés de la TF sont listées dans le tableau 1 et démontrées en annexe. On retiendra que la TF d’un signal dilaté est la version contractée de la TF du signal de départ, avec le même facteur de dilatation/contraction $(\alpha, 1/\alpha)$. On retiendra aussi que la TF d’un produit de convolution est égale au produit simple des TF. Inversement, la TF d’un produit de fonctions est égale au produit de convolution de leurs TF. Un cas d’application important de ce résultat est le suivant : la TF d’un signal retardé est égale à la version modulée de la TF du signal de départ (et inversement). Ces résultats se déduisent de ceux sur la TF du produit de convolution et du pro-

TABLE 1 – Propriétés de la transformée de Fourier, avec $e_{f_0}(t) = \exp(i2\pi f_0 t)$.

	$z(t)$	$Z(f)$
Linéarité	$(\alpha x + \beta y)(t)$	$\alpha X(f) + \beta Y(f)$
Dilatation / Contraction	$x(\alpha t)$	$\frac{1}{ \alpha } X\left(\frac{f}{\alpha}\right)$
Dualité de la TF	$(\mathcal{F}X)(t)$	$X(-f)$
Produit	$x(t)y(t)$	$(X * Y)(f)$
Convolution	$(x * y)(t)$	$X(f)Y(f)$
Retard	$x(t - \tau_0)$	$e_{-\tau_0}(f) X(f)$
Modulation	$e_{f_0}(t) x(t)$	$X(f - f_0)$
Dérivée	$x^{(n)}(t)$	$(i2\pi f)^n X(f)$
Produit par t^n	$t^n x(t)$	$\mathcal{F}x^{(n)}(f) / (-i2\pi)^n$

duit, en notant que $x(t - \tau_0) = (\delta_{\tau_0} * x)(t)$. Les TF de δ_{τ_0} et e_{f_0} sont données dans le tableau 2.

Une propriété fondamentale est que la transformée de Fourier conjuguée $\mathcal{F}^*$ est l'opérateur inverse de la transformée de Fourier.

Théorème 3.2. Si $x \in L^1(\mathbb{R})$ et $\mathcal{F}x \in L^1(\mathbb{R})$, alors on a $\mathcal{F}^* \mathcal{F}x(t) = x(t)$ presque partout.

La transformée de Fourier inverse s'écrit donc :

$$x(t) = \int_{\mathbb{R}} e^{2i\pi ft} X(f) df. \quad (75)$$

3.1.2 Transformée de Fourier des signaux dans L^2

La transformée de Fourier se prolonge en une isométrie de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$.

Définition 3.1. Soit $x \in L^2(\mathbb{R})$. On pose $x_n = x\mathbb{1}_{[-n,n]}$ pour $n \in \mathbb{Z}$. La TF de x est la limite dans L^2 :

$$X = \lim_{n \rightarrow \infty}^{L^2} X_n \quad (76)$$

de la suite X_n des transformées de Fourier (dans L^1) des fonctions x_n . La transformée de Fourier conjuguée se prolonge de façon similaire.

Le lecteur est redirigé vers le cours de CIP pour de plus amples détails. Les principales propriétés sont énoncées ci-dessous.

Théorème 3.3 (Plancherel-Parseval). *La TF est une isométrie de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$. Elle préserve la norme et le produit scalaire :*

- $\forall x \in L^2(\mathbb{R}), \|X\| = \|x\|$.
- $\forall x, y \in L^2(\mathbb{R}), (X, Y) = (x, y)$.

La première propriété s'interprète comme la conservation de l'énergie ($E_x = \|x\|^2 = E_X$) quelque soit la représentation, temporelle ou fréquentielle, choisie pour le signal.

Théorème 3.4 (Transformée de Fourier inverse). *Pour tout $x \in L^2(\mathbb{R}), \mathcal{F}\mathcal{F}^*x = \mathcal{F}^*\mathcal{F}x = x$ presque partout.*

3.1.3 Transformée de Fourier dans $\mathcal{S}'$

Les propriétés de la TF sur l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ sont à la base de la généralisation de la TF aux distributions. Tout d'abord, notons que $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}) \cap L^1(\mathbb{R})$. De plus, contrairement à $\mathcal{D}(\mathbb{R})$, $\mathcal{S}(\mathbb{R})$ est stable par transformée de Fourier (cf. polycopié du cours de CIP).

Lemme 3.1. $\forall \varphi \in \mathcal{S}(\mathbb{R}), \mathcal{F}\varphi \in \mathcal{S}(\mathbb{R})$.

C'est la raison pour laquelle on se place dans l'espace des distributions tempérées $T \in \mathcal{S}'(\mathbb{R})$ pour généraliser la TF.

Définition 3.2. *Soit $T \in \mathcal{S}'(\mathbb{R})$. La transformée de Fourier de T est la distribution tempérée $\mathcal{F}T \in \mathcal{S}'(\mathbb{R})$ définie par :*

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \mathcal{F}\{T\}, \varphi \rangle = \langle T, \mathcal{F}\varphi \rangle. \quad (77)$$

L'espace $\mathcal{E}'(\mathbb{R})$ des distributions à support compact étant inclus dans $\mathcal{S}'(\mathbb{R})$, la définition précédente est encore valable pour $T \in \mathcal{E}'(\mathbb{R})$.

Remarque. L'expression (77) est une généralisation du lemme suivant aux distributions.

Lemme 3.2. *Si x et $\varphi \in \mathcal{S}(\mathbb{R})$, alors $\mathcal{F}x \varphi$ et $\varphi \mathcal{F}x \in L^1(\mathbb{R})$, et*

$$\int_{\mathbb{R}} \mathcal{F}x \varphi = \int_{\mathbb{R}} x \mathcal{F}\varphi.$$

$\langle \mathcal{F}\{T\}, \varphi \rangle$ dans (77) joue le rôle de $\int_{\mathbb{R}} \mathcal{F}x \varphi$ ici.

Appliquons maintenant la définition au calcul de la TF de la distribution de Dirac.

Exemple 3.2 (TF d'un Dirac). *Soit $f_0 \in \mathbb{R}$. Pour $\varphi \in \mathcal{S}(\mathbb{R})$,*

$$\langle \mathcal{F}\delta_{f_0}, \varphi \rangle = \langle \delta_{f_0}, \mathcal{F}\varphi \rangle = (\mathcal{F}\varphi)(f_0) = \int_{\mathbb{R}} e^{-i2\pi f_0 t} \varphi(t) dt = \langle e_{-f_0}, \varphi \rangle.$$

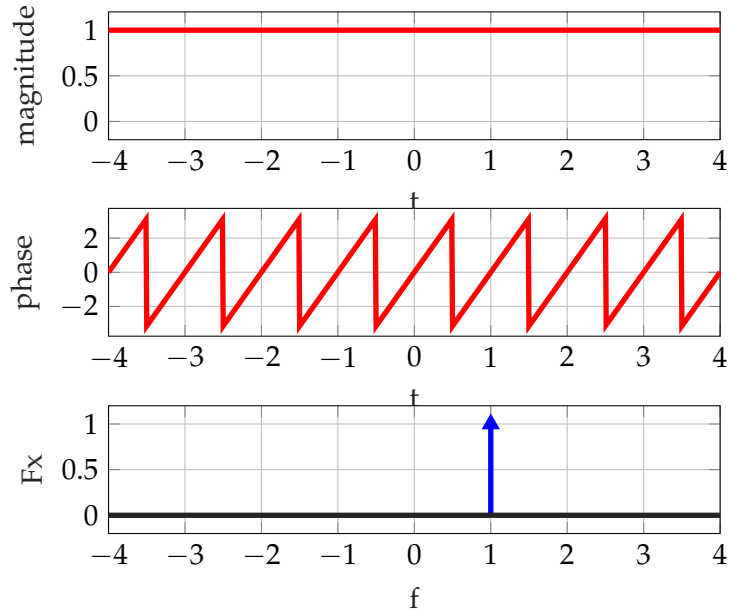


FIGURE 11 – Fonction $e_{f_0}(t) = \exp(2i\pi f_0 t)$ pour $f_0 = 1$: variations temporelles de l'amplitude et de la phase, tracé de sa TF.

On a donc

$$\mathcal{F}\delta_{f_0} = e_{-f_0}, \quad (78)$$

En particulier, pour $f_0 = 0$,

$$\mathcal{F}\delta = \mathbb{1}_{\mathbb{R}}. \quad (79)$$

Théorème 3.5 (Transformée de Fourier inverse). *La TF est un isomorphisme de $\mathcal{S}'(\mathbb{R})$ dans lui-même, et la transformée de Fourier inverse $\mathcal{F}^{-1} = \mathcal{F}^*$ est la distribution définie par :*

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \mathcal{F}^* \{T\}, \varphi \rangle = \langle T, \mathcal{F}^* \varphi \rangle.$$

Exemple 3.3 (TF de l'exponentielle complexe). Soit $f_0 \in \mathbb{R}$. En réitérant le calcul de l'exemple précédent pour la TF inverse, on trouve $\mathcal{F}^* \delta_{f_0} = e_{f_0}$, d'où

$$\mathcal{F}e_{f_0} = \mathcal{F}\mathcal{F}^* \delta_{f_0} = \delta_{f_0}. \quad (80)$$

Il en découle que les TF des distributions associés aux signaux trigonométriques sont données par

$$\mathcal{F} \{ \cos(2\pi f_0 t) \} = \mathcal{F} \left\{ \frac{e_{f_0} + e_{-f_0}}{2} \right\} = \frac{\delta_{-f_0} + \delta_{f_0}}{2}$$

$$\mathcal{F} \{ \sin(2\pi f_0 t) \} = \mathcal{F} \left\{ \frac{e_{f_0} - e_{-f_0}}{2i} \right\} = \frac{\delta_{-f_0} - \delta_{f_0}}{2i}.$$

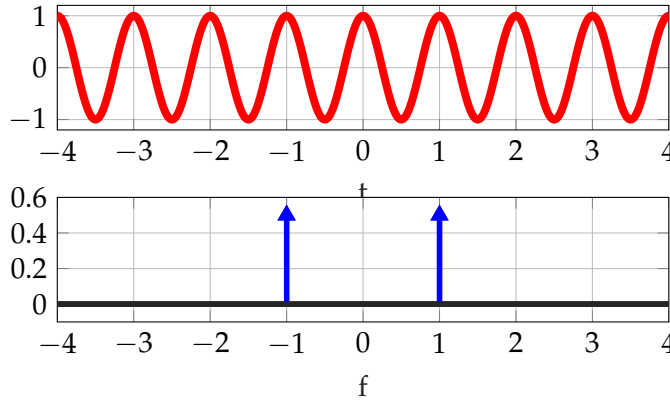


FIGURE 12 – Fonction $\cos(2\pi f_0 t)$ pour $f_0 = 1$: représentations temporelle et fréquentielle.

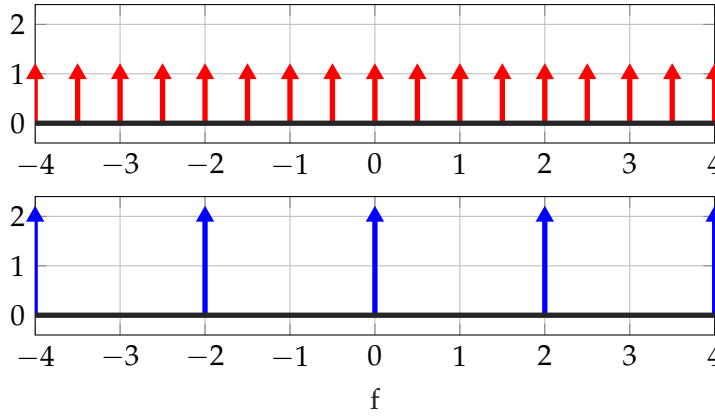


FIGURE 13 – Peigne de Dirac : représentations temporelle $\mathbb{I}_T$ et fréquentielle $\frac{1}{T}\mathbb{I}_{\frac{1}{T}}$ pour $T = \frac{1}{2}$.

Nous avons en particulier pour $f_0 = 0$:

$$\mathcal{F}\mathbb{I}_{\mathbb{R}} = \delta_0 = \delta.$$

Le calcul de la TF d'un peigne de Dirac repose sur le lemme suivant.

Lemme 3.3 (Formule sommatoire de Poisson). *Pour $\varphi \in \mathcal{S}(\mathbb{R})$ et $T > 0$, on a*

$$\frac{1}{T} \sum_{k \in \mathbb{Z}} \varphi\left(\frac{k}{T}\right) = \sum_{k \in \mathbb{Z}} \mathcal{F}\varphi(kT). \quad (81)$$

Théorème 3.6 (TF d'un peigne de Dirac). *Soit $T > 0$. Alors le peigne de Dirac $\mathbb{I}_T$ vérifie $\mathbb{I}_T \in \mathcal{S}'(\mathbb{R})$. De plus, on a :*

$$\mathcal{F}\mathbb{I}_T = \frac{1}{T}\mathbb{I}_{\frac{1}{T}} \quad (82)$$

Démonstration. En appliquant (81) aux fonctions test φ , on a :

$$\begin{aligned} \forall \varphi \in \mathcal{S}(\mathbb{R}), \quad \frac{1}{T} \langle \mathbb{W}_{\frac{1}{T}}, \varphi \rangle &= \frac{1}{T} \sum_{k \in \mathbb{Z}} \varphi\left(\frac{k}{T}\right) \\ &= \sum_{k \in \mathbb{Z}} (\mathcal{F}\varphi)(kT) \\ &= \langle \mathbb{W}_T, \mathcal{F}\varphi \rangle \\ &= \langle \mathcal{F}\mathbb{W}_T, \varphi \rangle. \end{aligned}$$

d'où l'égalité (82). □

3.1.4 Quelques exemples classiques

Les résultats les plus classiques sont récapitulés dans le tableau 2. On détaille maintenant le calcul très classique de la TF d'une fonction triangle. Soit

$$\triangle_{[-T,T]}(t) = \left(1 - \frac{|t|}{T}\right) \mathbb{1}_{[-T,T]}(t) \quad (83)$$

la fonction "triangle" centrée en 0, d'amplitude 1 et de support $[-T, T]$. Comme

$$\triangle_{[-T,T]} = \frac{1}{T} \left(\mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]} * \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]} \right), \quad (84)$$

on déduit de la propriété sur la TF du produit de convolution et de (73) que

$$\mathcal{F}\triangle_{[-T,T]}(f) = T \sin_c^2(T\pi f). \quad (85)$$

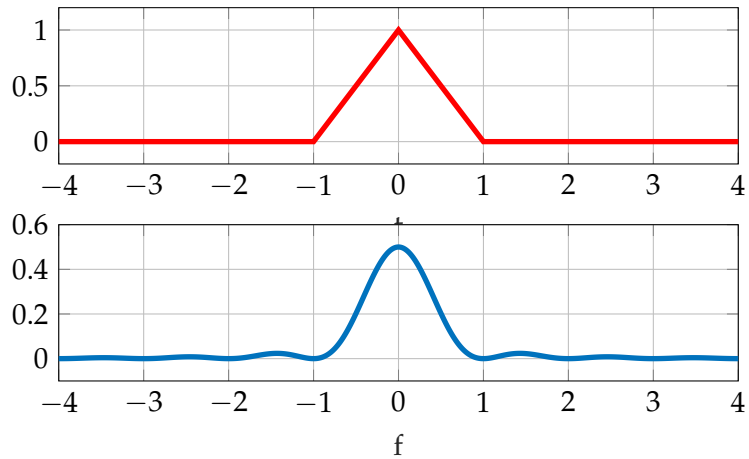


FIGURE 14 – Fonction $\triangle_{[-T,T]}$ pour $T = 1$: représentations temporelle et fréquentielle.

TABLE 2 – Transformées de Fourier usuelles. On note la symétrie des expressions (lignes 1-2, 3-4, 5-6, 7-8) du fait de la propriété de dualité de la TF : $(\mathcal{F}\mathcal{F}x)(t) = x(-t)$.

$x(t)$	$X(f)$
$\mathbb{1}_{[-T,T]}(t)$	$2T \operatorname{sinc}(2\pi T f)$
$\operatorname{sinc}(2\pi f_0 t)$	$\frac{1}{2f_0} \mathbb{1}_{[-f_0, f_0]}$
$\mathbb{1}_{\mathbb{R}}$	δ
δ	$\mathbb{1}_{\mathbb{R}}$
e_{f_0}	δ_{f_0}
δ_{τ_0}	$e_{-\tau_0}$
$\cos(2\pi f_0 t)$	$\delta_{-f_0} + \delta_{f_0}$
$\sin(2\pi f_0 t)$	$\frac{2}{\delta_{-f_0} - \delta_{f_0}}$
$\sqcup_T$	$\frac{1}{T} \sqcup_{\frac{1}{T}}$
$\frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{t^2}{2\sigma^2})$	$\exp(-2\sigma^2\pi^2 f^2)$

3.2 DÉCOMPOSITION EN SÉRIE DE FOURIER

3.2.1 Signaux à support limité

Le théorème de décomposition en série de Fourier est énoncé pour des signaux à support limité, dans l'intervalle $[0, T]$. L'espace $L^2([0, T])$ est muni du produit scalaire $(x, y) = \frac{1}{T} \int_{[0,T]} xy^* d\lambda$.

Théorème 3.7. *Tout signal $x \in L^2([0, T])$ est décomposable dans la base hilbertienne de $L^2([0, T])$ formée par les fonctions :*

$$e_k : [0, T] \rightarrow \mathbb{C}$$

$$t \mapsto \exp\left(2i\pi \frac{kt}{T}\right)$$

La décomposition s'écrit

$$x = \sum_{k \in \mathbb{Z}} c_k e_k \quad (86)$$

où

$$c_k = (x, e_k) = \frac{1}{T} \int_0^T \exp\left(-i2\pi \frac{kt}{T}\right) x(t) dt. \quad (87)$$

On a de plus l'égalité de Parseval :

$$\|x\|_{L^2(0,T)}^2 = \frac{1}{T} \int_{[0,T]} |x(t)|^2 dt = \sum_{k \in \mathbb{Z}} |c_k|^2. \quad (88)$$

La série (86) converge dans L^2 . Sous certaines hypothèses, on a des propriétés de convergence simple et uniforme.

Théorème 3.8 (Dirichlet). *Si x est $\mathcal{C}^1$ par morceaux, alors (86) est valable en tout point t où x est continue. Si x est de plus continue, la convergence de la série de Fourier vers x est uniforme.*

3.2.2 Signaux périodiques

On définit la transformée de Fourier de signaux périodiques au sens des distributions.

Considérons un signal T -périodique x et son motif période $x_T := x\mathbb{1}_{[0,T]}$. On peut alors écrire $\forall t, x(t) = \sum_{k \in \mathbb{Z}} x_T(t - kT)$, ou encore

$$x = x_T * \mathbb{I}_T. \quad (89)$$

On suppose que $x_T \in L^1(0, T)$, de telle sorte que $X_T \in \mathcal{C}^0(\mathbb{R})$. D'après le théorème 3.6, la transformée de Fourier de x s'écrit :

$$X = \mathcal{F} \{x_T * \mathbb{I}_T\} = X_T \frac{1}{T} \mathbb{I}_{\frac{1}{T}}.$$

On a donc

$$X = \sum_{k \in \mathbb{Z}} c_k \delta_{\frac{k}{T}} \quad (90)$$

où

$$\begin{aligned} c_k &= \frac{1}{T} X_T \left(\frac{k}{T} \right) \\ &= \frac{1}{T} \int_0^T \exp \left(-i2\pi \frac{kt}{T} \right) x(t) dt \end{aligned} \quad (91)$$

sont égaux aux coefficients de Fourier de x_T définis en (87).

On retiendra qu'une fonction périodique possède un spectre de raies conformément à (90). Inversement, on montre dans la prochaine section que la transformée de Fourier d'un signal à temps discret est périodique.

3.3 TRANSFORMÉE DE FOURIER À TEMPS DISCRET

La transformée de Fourier d'un signal à temps discret (TFTD) $x : \mathbb{Z} \rightarrow \mathbb{C}$ est définie par

$$\forall \nu \in \mathbb{R}, \mathcal{F}x(\nu) = X(\nu) = \sum_{k \in \mathbb{Z}} x[k] e^{-i2\pi \nu k} \quad (92)$$

De même que dans le cas continu, on utilise $X := \mathcal{F}x$, et la notation $\mathcal{F} \{x[k]\}(\nu)$ lorsqu'il est nécessaire d'expliciter la dépendance de x par rapport à k .

La transformée de Fourier à temps discret est un signal

$$\begin{aligned} X : \mathbb{R} &\rightarrow \mathbb{C} \\ \nu &\mapsto X(\nu) \end{aligned}$$

périodique de période 1. La variable ν est appelée *fréquence réduite*.

3.3.1 Transformée de signaux dans ℓ^1

Les propriétés de la TFTD dans le cas de signaux dans $\ell^1(\mathbb{Z})$ sont analogues à celles de la TF pour les signaux L^1 .

Théorème 3.9. Soit $x \in \ell^1$. La TFTD satisfait

- $X \in \ell^\infty$;
- $X \in \mathcal{C}^0$.

Nous avons également, à partir de la définition, que

$$X(0) = \sum_{k \in \mathbb{Z}} x[k]. \quad (93)$$

Exemple 3.4 (Fenêtre rectangulaire). Considérons la fenêtre (discrète) rectangulaire de largeur $2N + 1$ donnée par

$$\mathbb{1}_{[-N, N]}[k] = \sum_{n=-N}^N \delta[k - n]$$

La TFTD s'écrit :

$$\mathcal{F} \left\{ \mathbb{1}_{[-N, N]} \right\} (\nu) = \sum_{k=-N}^N e^{-i2\pi k \nu} = \begin{cases} \frac{\sin(\pi(2N+1)\nu)}{\sin(\pi\nu)} & \text{si } \nu \neq 0 \\ 2N+1 & \text{si } \nu = 0. \end{cases}$$

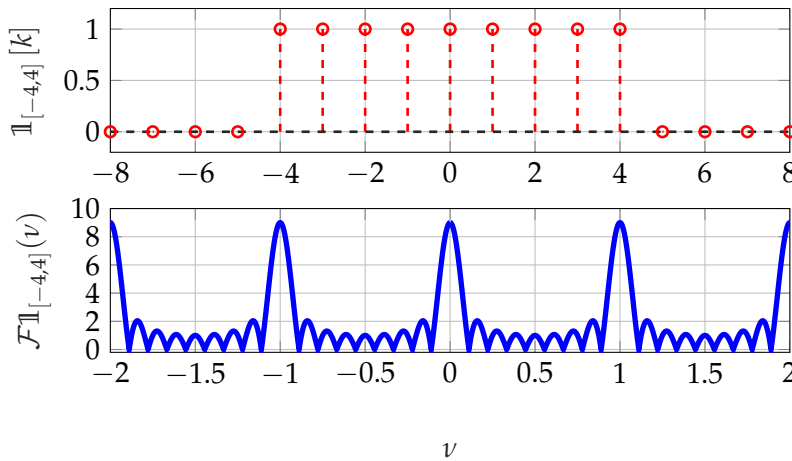


FIGURE 15 – Fenêtre rectangulaire discrète $\mathbb{1}_{[-N, N]}$: représentations temporelle et fréquentielle (en module). Comme pour tout signal discret, la TFTD est 1-périodique.

TABLE 3 – Propriétés de la transformée de Fourier à temps discret.

	$z[k]$	$Z(\nu)$
Linéarité	$\alpha x[k] + \beta y[k]$	$\alpha X(\nu) + \beta Y(\nu)$
Dualité	$\mathcal{F}\{X(\nu)\}[k]$	$X(-\nu)$
Retard	$x[k - k_0]$	$e^{-i2\pi\nu k_0} X(\nu)$
Modulation	$e^{i2\pi\nu_0 k} x[k]$	$X(\nu - \nu_0)$
Produit	$x[k] y[k]$	$(X \circledast Y)(\nu)$
Convolution	$(x * y)[k]$	$X(\nu) Y(\nu)$

On remarque que le sinus cardinal obtenu dans le cas continu, cf. (73) est remplacé ici par un rapport de sinus.

3.3.2 Transformée des signaux dans ℓ^2

Théorème 3.10. L'opérateur $\mathcal{F}$ défini par (92) est une isométrie bijective de $\ell_2(\mathbb{Z})$ dans $L^2([0, 1])$, ce qui se traduit par la conservation du produit scalaire et de l'énergie (égalité de Parseval) :

$$\forall x, y \in \ell_2(\mathbb{Z}), \sum_{k \in \mathbb{Z}} x[k] y^*[k] = \int_{[0,1]} X(\nu) Y^*(\nu) d\nu$$

$$\forall x \in \ell_2(\mathbb{Z}), \sum_{k \in \mathbb{Z}} |x[k]|^2 = \int_{[0,1]} |X(\nu)|^2 d\nu.$$

La transformation inverse est donnée par

$$x[k] = \int_{[0,1]} X(\nu) e^{2i\pi k\nu} d\nu \quad (94)$$

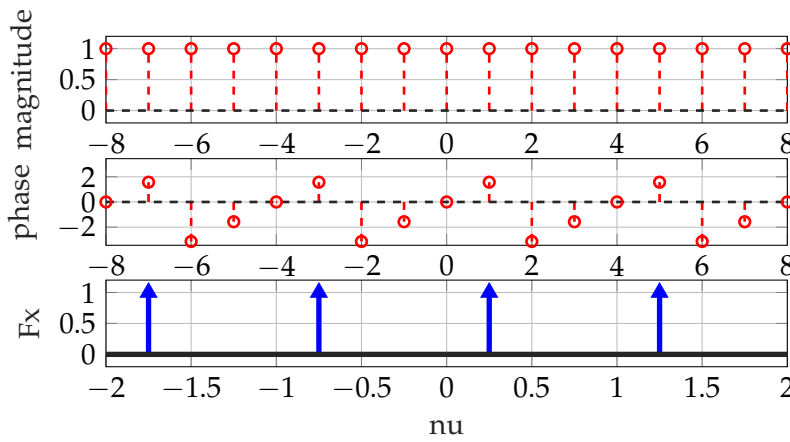
Remarque. (92) n'est rien d'autre que la décomposition en série de Fourier de la fonction (1-périodique) $X = \mathcal{F}x$. Les coefficients de cette décomposition s'expriment comme les produits scalaires de X avec les signaux exponentiels complexes, voir (94). X étant 1-périodique, son énergie est définie comme une intégrale sur une période, comme $[0, 1]$.

Les principales propriétés de la TFTD sont résumées dans le tableau 3. De même que pour la TF de signaux continus, la TFTD d'une convolution est le produit des TFTD de chaque signal discret. Inversement, la TFTD d'un produit est le produit de convolution circulaire des TFTD. X et Y étant des fonctions 1-périodiques, la convolution circulaire $\circledast$ est définie par :

$$\forall \nu \in \mathbb{R}, (X \circledast Y)(\nu) = \int_{[0,1]} X(\theta) Y(\nu - \theta) d\theta. \quad (95)$$

TABLE 4 – Récapitulatif des transformées de Fourier de signaux à temps discret.

$x[k]$	$X(\nu)$
$\mathbb{1}_{[-N,N]}[k]$	$\frac{\sin(\pi(2N+1)\nu)}{\sin(\pi\nu)}$
$\exp(i2\pi\nu_0 k)$	$\sum_{k \in \mathbb{Z}} \delta_{k+\nu_0}$
$\delta[k]$	$\mathbb{1}_{\mathbb{R}}(\nu)$
$\sum_{n=-\infty}^{\infty} \delta[k - nN]$	$\frac{1}{N} \coprod_{\frac{1}{N}}$
$\cos(2\pi\nu_0 k)$	$\frac{1}{2} \sum_{k \in \mathbb{Z}} (\delta_{k-\nu_0} + \delta_{k+\nu_0})$
$\sin(2\pi\nu_0 k)$	$\frac{1}{2i} \sum_{k \in \mathbb{Z}} (\delta_{k-\nu_0} - \delta_{k+\nu_0})$

FIGURE 16 – Exponentielle complexe $\exp(2i\pi\nu_0 k)$: représentations temporelle (amplitude et phase) et fréquentielle pour $\nu_0 = 0,25$. La TFTD est 1-périodique.

3.3.3 Quelques exemples classiques

Le tableau 4 récapitule des exemples classiques de résultats de transformée de Fourier. Certains signaux considérés comme les exponentielles complexes, ne se trouvent ni dans ℓ_1 ni dans ℓ_2 . Dans ce cas, leur TFTD est calculée au sens des distributions.

3.4 TRANSFORMÉE DE FOURIER DISCRÈTE

La transformée de Fourier discrète (à ne pas confondre avec la TFTD) est un outil très utile en pratique pour le traitement numérique du signal. Contrairement à la TFTD, elle s'applique à un **signal discret de longueur finie**, représenté par un vecteur $\mathbf{x} = \{x[0], \dots, x[N-1]\} \in \mathbb{C}^N$.

Définition 3.3 (Transformée de Fourier Discrète). *La Transformée de Fourier Discrète (TFD) est l'application linéaire*

$$\begin{aligned}\mathbb{C}^N &\rightarrow \mathbb{C}^N \\ \mathbf{x} &\mapsto \mathbf{X} = \mathbf{F}_N \mathbf{x}\end{aligned}$$

où $\mathbf{F}_N \in \mathbb{C}^{N \times N}$ est la matrice définie par

$$\mathbf{F}_N[n, k] = e^{-i2\pi \frac{nk}{N}} \quad (96)$$

pour n et $k \in \{0, \dots, N-1\}$.

La TFD permet de définir une représentation fréquentielle $X[n]$, $n \in \{0, \dots, N-1\}$ du signal $\mathbf{x}$ avec

$$\begin{aligned}X[n] &= \sum_{k=0}^{N-1} \mathbf{F}_N[n, k] x[k] \\ &= \sum_{k=0}^{N-1} e^{-i2\pi \frac{nk}{N}} x[k].\end{aligned} \quad (97)$$

On note que $\mathbf{F}_N$ est une matrice symétrique de Vandermonde, dont l'inverse est égale (à un facteur multiplicatif près) à sa conjuguée.

Théorème 3.11 (TFD inverse). *La Transformée inverse de la TFD est l'opération linéaire $\mathbf{x} = \mathbf{F}_N^{-1} \mathbf{X}$, où*

$$\mathbf{F}_N^{-1} = \frac{1}{N} \mathbf{F}_N^*. \quad (98)$$

Ainsi, la TFD inverse prend la forme

$$x[k] = \frac{1}{N} \sum_{n=0}^{N-1} e^{i2\pi \frac{nk}{N}} X[n] \quad (99)$$

pour $k \in \{0, \dots, N-1\}$.

Remarque. (99) s'interprète comme la décomposition de $\mathbf{x}$ dans la base orthogonale formée par les signaux exponentiels complexes de fréquence $\frac{n}{N}$, notés $\mathbf{e}_n = \{\exp(2i\pi \frac{nk}{N}), k = 0, \dots, N-1\}$:

$$\mathbf{x} = \sum_{n=0}^{N-1} \left(\frac{1}{N} X[n] \right) \mathbf{e}_n \quad (100)$$

Le théorème de Parseval garantit la conservation de l'énergie (à un facteur multiplicatif près).

Théorème 3.12 (Parseval). *Pour $\mathbf{x} \in \mathbb{C}^N$, on a $\|\mathbf{F}_N \mathbf{x}\|_2^2 = N \|\mathbf{x}\|_2^2$.*

Démonstration.

$$\|\mathbf{F}_N \mathbf{x}\|_2^2 = (\mathbf{F}_N \mathbf{x})^* (\mathbf{F}_N \mathbf{x}) = \mathbf{x}^* \mathbf{F}_N^* \mathbf{F}_N \mathbf{x} = \mathbf{x}^* (N\mathbf{I}) \mathbf{x} = N \|\mathbf{x}\|_2^2.$$

□

Notons que les calculs de produit matrice-vecteur $\mathbf{F}_N \mathbf{x}$ et $\mathbf{F}_N^{-1} \mathbf{X}$ coûtent en principe de l'ordre de N^2 opérations élémentaires (additions et multiplications). Cependant, on dispose d'algorithmes rapides (appelés FFT, pour *Fast Fourier Transform*) permettant d'effectuer ces calculs en $\mathcal{O}(N \log(N))$ opérations, en exploitant la structure particulière de la matrice $\mathbf{F}_N$.

3.5 APPLICATION AU FILTRAGE LINÉAIRE

3.5.1 Filtrage dans le domaine de Fourier

Considérons un filtre convolutionnel de réponse impulsionnelle h . La relation entrée-sortie s'écrit $y(t) = (h * x)(t)$ ou $y[k] = (h * x)[k]$. Cette relation est simplifiée dans le domaine fréquentiel car la TF (TFTD) d'un produit de convolution est le produit des TF (TFTD).

Filtres à temps continu

Définition 3.4 (Réponse fréquentielle). *La relation entrée-sortie d'un filtre s'écrit dans le domaine fréquentiel sous la forme :*

$$\forall f \in \mathbb{R}, \quad Y(f) = H(f) X(f) \quad (101)$$

où $H(f) = (\mathcal{F}h)(f)$ est appelée **réponse fréquentielle** ou **fonction de transfert**.

Les exponentielles complexes e_f sont fonctions propres des filtres (cf. section 2.5). La sortie associée s'écrit :

$$y = H(f) e_f. \quad (102)$$

Le module et la phase de $H(f)$ s'interprètent comme le gain en amplitude et le déphasage subis par le signal e_f au passage dans le filtre.

Exemple 3.5. Un filtre passe-bas idéal est caractérisé par la réponse fréquentielle $H(f) = \mathbb{1}_{[-f_c, f_c]}(f)$ en notant f_c la fréquence de coupure. Il a pour effet de couper les hautes fréquences correspondant aux discontinuités (sauts des valeurs du signal) et au bruit contenu dans le signal d'entrée. Il lisse ce signal. D'après le tableau 2, sa réponse impulsionnelle s'écrit

$$h(t) = 2f_c \operatorname{sinc}(2\pi f_c t).$$

Le filtre n'est ni causal, ni à réponse impulsionnelle finie, d'où son nom de passe-bas idéal. En pratique, on a recours à des techniques de fenêtrage

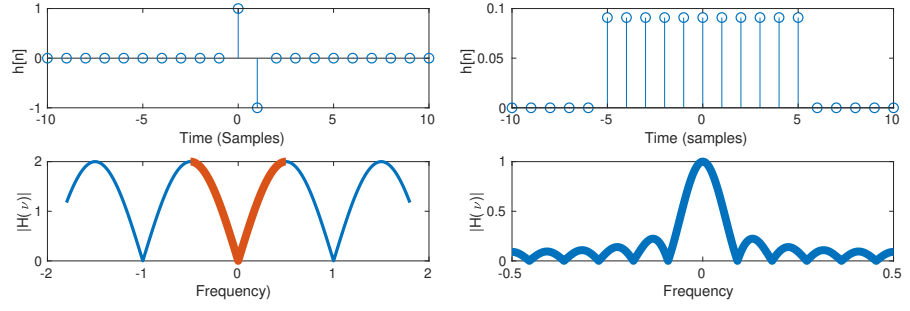


FIGURE 17 – Réponses impulsionnelle et fréquentielle des filtres dérivateur (gauche) et à moyenne glissante (droite).

pour concevoir un filtre causal et stable dont la réponse fréquentielle soit une bonne approximation de $\mathbb{1}_{[-f_c, f_c]}$.

Filtres à temps discret

Définition 3.5. La relation entrée-sortie d'un filtre à temps discret s'écrit dans le domaine fréquentiel sous la forme :

$$\forall \nu, Y(\nu) = H(\nu) X(\nu) \quad (103)$$

où $H(\nu) = (\mathcal{F}h)(\nu)$ est la réponse fréquentielle. Les TFTD étant périodiques, elle sont calculées sur un intervalle de longueur 1 (généralement, pour $\nu \in [0, 1[$ ou $\nu \in [-\frac{1}{2}, \frac{1}{2}[$).

Exemple 3.6. Le filtre dérivateur est défini par $y[n] = x[n] - x[n-1]$. Sa réponse impulsionnelle s'écrit

$$h[k] = \delta[k] - \delta[k-1] \quad (104)$$

(cf. paragraphe 2.4.2). Le support de h est donc réduit à $\{0, 1\}$. La réponse fréquentielle du filtre s'écrit :

$$\begin{aligned} H(\nu) &= h[0] + h[1] \exp(-2i\pi\nu) \\ &= 1 - \exp(-2i\pi\nu) \\ &= 2i \exp(-i\pi\nu) \sin(\pi\nu). \end{aligned} \quad (105)$$

C'est un filtre passe-haut car $H(0) = 0$ et les valeurs maximales de $|H(\nu)|$ sont atteintes pour $|\nu| = 1/2$.

Exemple 3.7. Le filtre à moyenne glissante est défini par

$$y[n] = \frac{1}{2L+1} \sum_{k=-L}^L x[n-k].$$

On montre que sa réponse fréquentielle s'écrit

$$H(\nu) = \frac{\sin(2\pi(2L+1)\nu)}{(2L+1)\sin(\pi\nu)}$$

C'est un filtre passe bas.

Équations aux différences

Certains filtres IIR sont caractérisés par des équations aux différences du type :

$$\forall k, y[k] = \sum_{\ell=1}^p a_{\ell} y[k-\ell] + x[k] \quad (106)$$

où p est l'ordre du filtre. On a donc l'égalité des signaux :

$$y = \sum_{\ell=1}^p a_{\ell} y[\cdot - \ell] + x.$$

La réponse impulsionnelle du filtre étant de support infini, son calcul n'est pas direct. En revanche, la réponse fréquentielle $H(\nu)$ se calcule très simplement. On applique la propriété de la TFTD de signaux retardés (tableau 3) :

$$\forall \nu, Y(\nu) = \sum_{\ell=1}^p a_{\ell} e^{-2i\pi\ell\nu} Y(\nu) + X(\nu). \quad (107)$$

Par identification avec (103), on déduit que :

$$\forall \nu, H(\nu) = \frac{1}{1 - \sum_{\ell=1}^p a_{\ell} e^{-2i\pi\ell\nu}}. \quad (108)$$

3.5.2 Calcul numérique de la réponse fréquentielle d'un filtre FIR

Considérons un **filtre FIR causal**, et notons p la **taille de la réponse impulsionnelle**. La relation entrée-sortie s'écrit

$$y[n] = \sum_{k=0}^{p-1} h[k] x[n-k]$$

Le support de h étant fini, on a recours à la TFD pour calculer la réponse fréquentielle $H(\nu)$, renotée $(\mathcal{F}h)(\nu)$ pour éviter les confusions.

La TFD opère sur le vecteur $\mathbf{h} = \{h[k], k = 0, \dots, p-1\} \in \mathbb{C}^p$. Le vecteur $\mathbf{H} = \mathbf{F}_p \mathbf{h}$ obtenu (cf. (97)) coïncide avec l'évaluation de $(\mathcal{F}h)(\nu)$ pour des fréquences multiples de $\frac{1}{p}$:

$$\forall n \in \{0, \dots, p-1\}, H[n] = (\mathcal{F}h)\left(\frac{n}{p}\right) \quad (109)$$

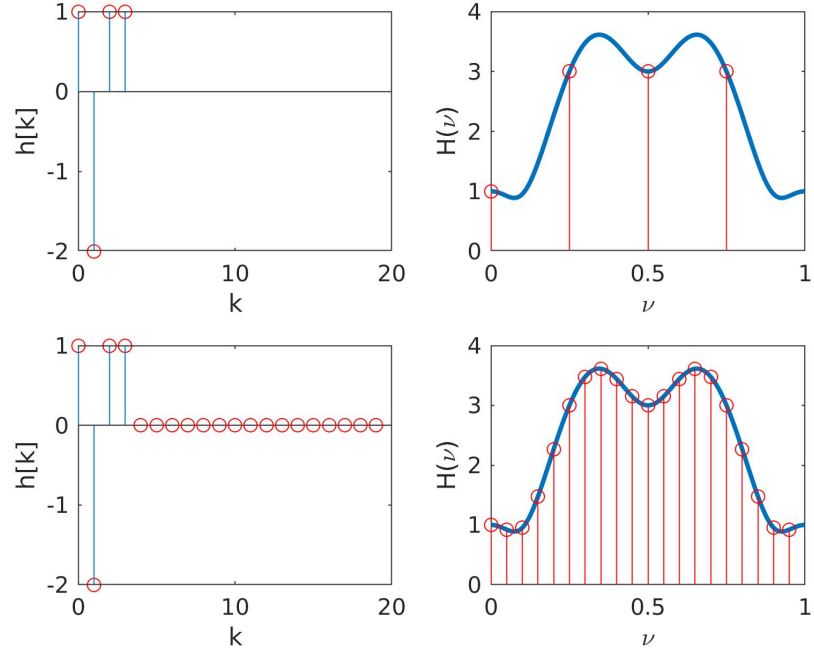


FIGURE 18 – Calcul de la réponse fréquentielle d’un filtre FIR. A droite, la TFTD $(\mathcal{F}h)(\nu)$ est représentée en trait continu, et son évaluation (pour 4 et 20 points fréquentiels, respectivement) par TFD est en rouge. Le bourrage de zéros (figures du bas) améliore la représentation fréquentielle.

Le pas de la grille fréquentielle (égal à l’espacement entre deux points sur cette grille) est donc $1/p$.

La technique de **bourrage de zéros** (ou **zero-padding**) consiste à rajouter artificiellement des 0 dans le vecteur $\mathbf{h}$ afin d’améliorer la représentation fréquentielle. Notons $\mathbf{h}' = \{\mathbf{h}, \mathbf{0}_{M-p}\} \in \mathbb{C}^M$ le vecteur formé en ajoutant $M - p$ zéros à la fin du vecteur $\mathbf{h}$ (avec $M > p$). La TFD $\mathbf{H}' = \mathbf{F}_M \mathbf{h}'$ évalue la réponse fréquentielle sur la grille de pas $1/M$:

$$\forall n \in \{0, \dots, M-1\}, H'[n] = (\mathcal{F}h)\left(\frac{n}{M}\right). \quad (110)$$

Pour pouvoir calculer la représentation fréquentielle de la sortie $y = h * x$ du filtre associée à une entrée x par la formule (103), on veillera à calculer la TFTD $X(\nu)$ sur la même grille que la grille de calcul de $(\mathcal{F}h)(\nu)$, c’est-à-dire sur une grille de pas $1/M$ en suivant (110).

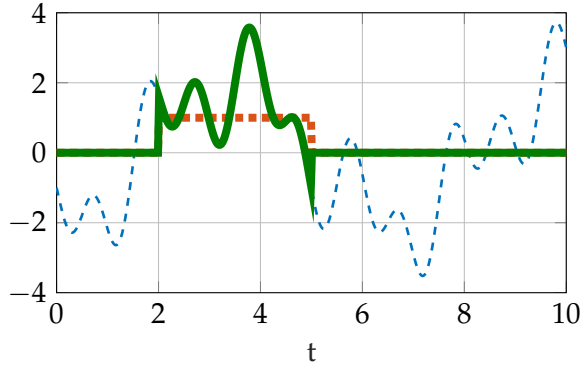


FIGURE 19 – Fenêtrage rectangulaire : $x(t) = s(t) \mathbb{1}_{[a,b]}(t)$. Le signal observé x est tracé en vert.

3.6 ANALYSE SPECTRALE

Le problème de l'analyse spectrale consiste à décomposer un signal comme une somme de sinusoïdes. C'est un problème très classique rencontré dans différents domaines allant de l'analyse du son à la localisation radar, à l'analyse de la lumière émise par un corps céleste en astrophysique. La transformée de Fourier est l'outil naturel pour effectuer cette décomposition.

3.6.1 Effet du fenêtrage

En pratique, le signal $x(t)$ qu'on cherche à analyser a été observé pendant une durée finie, par exemple pour $t \in [-\frac{T}{2}, \frac{T}{2}]$. On peut donc le modéliser comme le produit d'un signal idéal stationnaire $s(t)$, correspondant à un horizon infini, par une fenêtre rectangulaire :

$$x(t) = s(t) \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}(t) \quad (111)$$

On déduit des propriétés de la TF que :

$$\begin{aligned} X &= S * \mathcal{F} \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]} \\ &= S * \{T \operatorname{sinc}(\pi T f)\}. \end{aligned} \quad (112)$$

Prenons l'exemple du signal sinusoïdal $s(t) = \cos(2\pi f_0 t)$, dont la TF est égale à $\frac{1}{2}(\delta_{f_0} + \delta_{-f_0})$. (112) se réécrit :

$$\forall f, X(f) = \frac{T}{2} (\operatorname{sinc}(\pi T (f - f_0)) + \operatorname{sinc}(\pi T (f + f_0))) \quad (113)$$

Graphiquement, les impulsions de Dirac (aussi appelées raies) en $f = \pm f_0$ sont remplacées par des sinus cardinaux, toujours centrés en $\pm f_0$. Leurs lobes primaires et secondaires sont représentés sur la figure 20.

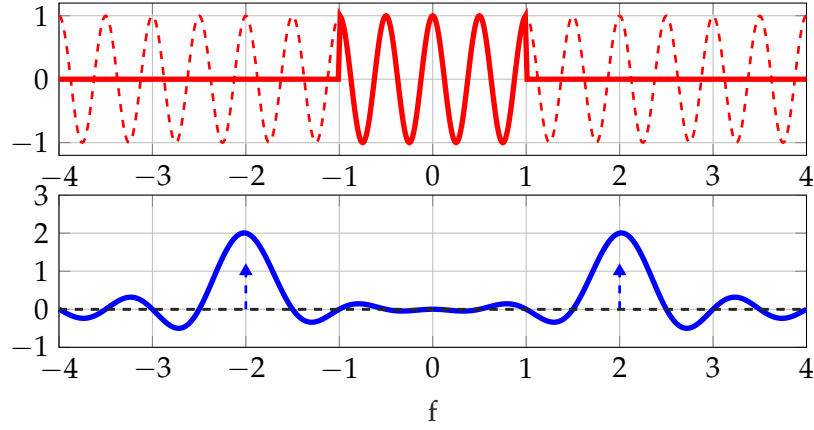


FIGURE 20 – Sinusoïde avec fenêtrage rectangulaire : représentations temporelle et fréquentielle.

3.6.2 Notion de résolution

La **résolution spectrale** est la capacité à distinguer des sinusoïdes proches. Dans le cas d'une **fenêtre rectangulaire**, elle est égale à la **largeur du lobe principal** du motif $T \sin_c(\pi T f)$. Elle vaut $2/T$ car $\sin_c(\pi T f)$ s'annule en $f = \pm \frac{1}{T}$. La résolution est donc d'autant **meilleure** que la **durée d'observation** est longue. De plus, l'amplitude des lobes secondaires décroît relativement lentement (en $1/f$).

Dans le cas d'un signal plus complexe qu'une simple sinusoïde, la présence de lobes multiples et leur superposition constitue une difficulté pour distinguer des fréquences proches.

On peut atténuer, dans une certaine mesure, les effets du fenêtrage rectangulaire en remplaçant (111) par

$$\forall t \in \mathbb{R}, x(t) = s(t) w_{[-\frac{T}{2}, \frac{T}{2}]}(t) \quad (114)$$

où $w_{[-\frac{T}{2}, \frac{T}{2}]}(t)$ est la **fenêtre triangulaire** $\triangle_{[-T/2, T/2]}$ définie en (83), ou encore une **fenêtre de forme plus spécifique**, comme celle Hann :

$$w_{[-\frac{T}{2}, \frac{T}{2}]}(t) = \cos^2\left(\frac{\pi t}{T}\right) \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}(t). \quad (115)$$

représentée Fig. 21. Ces fenêtres sont de largeur T et conduisent à une décroissance rapide des lobes secondaires (en $1/f^2$ ou $1/f^3$, respectivement) au prix d'une résolution spectrale double ($4/T$), voir Fig. 22.

3.6.3 Calcul numérique avec la TFD

Pour un signal discret quelconque $s[k]$, le principe du fenêtrage reste valable. Le signal fenêtré s'exprime par :

$$\forall k \in \mathbb{Z}, x[k] = s[k] w_{\{0, \dots, N-1\}}(k) \quad (116)$$

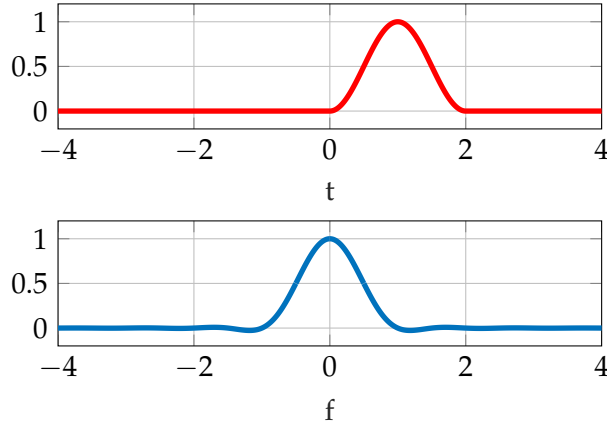


FIGURE 21 – Fenêtre de Hann : représentations temporelle et fréquentielle (en module).

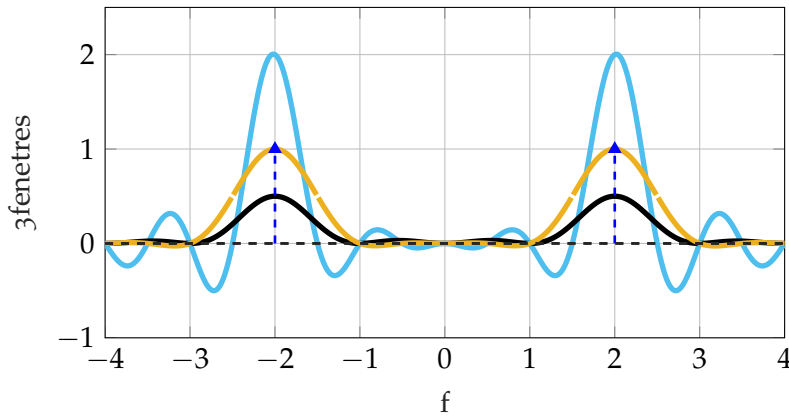


FIGURE 22 – Effets de fenêtrage sur une sinusoïde : fenêtrages rectangulaire (bleu), triangulaire (noir), de Hann (jaune).

où $w_{\{0,\dots,N-1\}}$ est une version discrète des fenêtres précédentes (rectangulaire, triangulaire ou Hann) de longueur N .

Après fenêtrage, on dispose d'un vecteur $\mathbf{x} = \{x[k], k = 0, \dots, N-1\}$ de longueur N . La technique du bourrage de zéros introduite dans la section 3.5.2 consiste à rajouter artificiellement $M - N$ zéros à ce vecteur, pour $M > N$. La TFD du vecteur $\mathbf{x}$ augmenté fournit une évaluation $\mathbf{X} \in \mathbb{C}^M$ de $\mathcal{F}x(\nu)$ sur la grille fréquentielle

$$\nu_n = \frac{n}{M}, \quad n = 0, \dots, M-1$$

de pas $\frac{1}{M}$. En localisant la position, l'amplitude et la phase des maxima locaux de $|X[n]|$, on déduit les paramètres des sinusoïdes contenues dans le signal de départ $s[k]$.

L'intérêt de la technique de bourrage de zéros est illustrée sur la Fig. 23 dans le cas où $M = 2N$. On observe un gain notable de précision pour localiser la fréquence de la sinusoïde discrète.

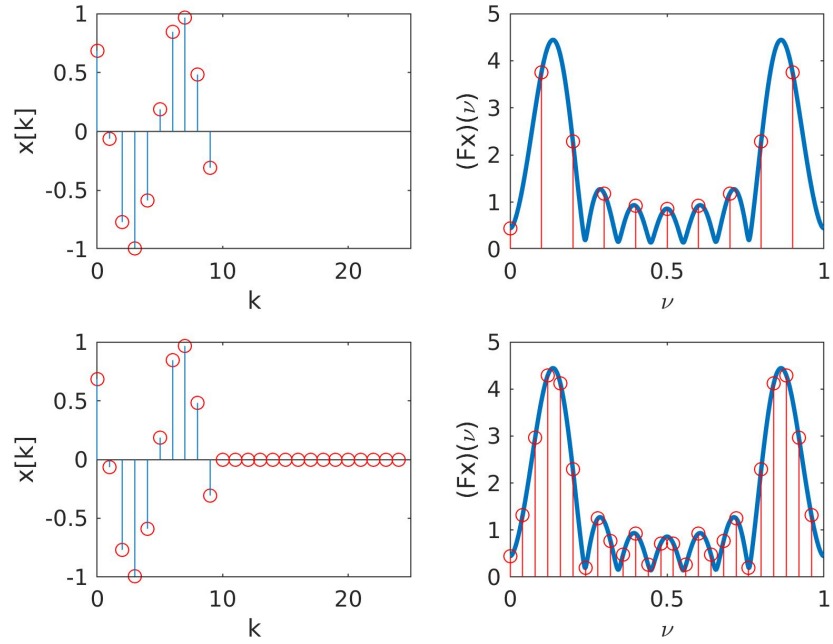


FIGURE 23 – Représentations fréquentielles d'un signal discret de longueur $N = 10$. A droite, la TFTD $X(\nu)$ est représentée en trait continu. Son évaluation par TFD est en rouge. Le calcul est fait pour N points fréquentiels (figures du haut). Pour un calcul avec M points fréquentiels ($M > N$), on procède à un bourrage de zéros en ajoutant $M - N$ zéros au signal $x[k]$ ($M = 25$ pour les figures du bas).

3.7 CONCLUSION

La transformée de Fourier est un outil fondamental permettant d'analyser des signaux stationnaires. De nombreux signaux ne sont pas stationnaires, on a alors recours à des outils plus sophistiqués comme la transformée de Fourier à court terme ou la transformée en ondelettes. Enfin, la TF joue un rôle majeur dans la compréhension de l'échantillonnage de signaux continus. Cette opération est nécessaire pour la transmission, le stockage et la reconstruction des signaux temporels et sera détaillée au prochain chapitre.

ANNEXE

Les propriétés sont présentées pour les TF de fonctions dans L^1 .

Linéarité. Soit $z = \alpha x + \beta y$.

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-i2\pi ft} (\alpha x(t) + \beta y(t)) dt \\ &= \alpha \int_{\mathbb{R}} e^{-i2\pi ft} x(t) dt + \beta \int_{\mathbb{R}} e^{-i2\pi ft} y(t) dt \\ &= \alpha X(f) + \beta Y(f) \end{aligned}$$

Dilatation. Soit $z(t) = x(\alpha t)$.

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-i2\pi ft} x(\alpha t) dt \\ &= \int_{\mathbb{R}} e^{-i2\pi f \frac{\theta}{\alpha}} x(\theta) \frac{1}{|\alpha|} d\theta \\ &= \frac{1}{|\alpha|} X\left(\frac{f}{\alpha}\right). \end{aligned}$$

Dualité de la TF. La transformée de Fourier inverse est donnée par $x(t) = \int_{\mathbb{R}} e^{i2\pi t\eta} X(\eta) d\eta$, nous avons donc

$$x(-f) = \int_{\mathbb{R}} e^{-i2\pi f\eta} X(\eta) d\eta = \mathcal{F}\{X\}(f)$$

Produit de convolution. Soit $z = x * y$.

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-i2\pi ft} \int_{\mathbb{R}} x(\theta) y(t - \theta) d\theta dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-i2\pi ft} x(\theta) y(t - \theta) d\theta dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-i2\pi f(t-\theta)} e^{-i2\pi f\theta} x(\theta) y(t - \theta) d\theta dt \\ &= \int_{\mathbb{R}} e^{-i2\pi f\theta} x(\theta) \int_{\mathbb{R}} e^{-i2\pi f(t-\theta)} y(t - \theta) dt d\theta \\ &= \int_{\mathbb{R}} e^{-i2\pi f\theta} x(\theta) \int_{\mathbb{R}} e^{-i2\pi f\tau} y(\tau) d\tau d\theta \\ &= \int_{\mathbb{R}} e^{-i2\pi f\theta} x(\theta) d\theta \int_{\mathbb{R}} e^{-i2\pi f\tau} y(\tau) d\tau \\ &= X(f) Y(f). \end{aligned}$$

Produit. Soit $z = xy$.

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-i2\pi ft} x(t) y(t) dt \\ &= \int_{\mathbb{R}} e^{-i2\pi ft} \left(\int_{\mathbb{R}} e^{i2\pi t\eta} \mathcal{F}x(\eta) d\eta \right) y(t) dt \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-i2\pi(f-\eta)t} y(t) dt \right) \mathcal{F}x(\eta) d\eta \\ &= \int_{\mathbb{R}} \mathcal{F}y(f - \eta) \mathcal{F}x(\eta) d\eta \\ &= (X * Y)(f) (\mathcal{F}y * \mathcal{F}x)(f) \end{aligned}$$

Retard. Soit $z(t) = x(t - \tau)$. On peut effectuer le calcul :

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-i2\pi f t} x(t - \tau) dt \\ &= \int_{\mathbb{R}} e^{-i2\pi f(\theta + \tau)} x(\theta) d\theta \\ &= e^{-i2\pi f \tau} \int_{\mathbb{R}} e^{-i2\pi f \theta} x(\theta) d\theta \\ &= e_{-\tau}(f) X(f). \end{aligned}$$

On arrive au même résultat de façon plus directe en notant que $z = \delta_{\tau} * x$, et en exploitant le résultat sur la TF du produit de convolution :

$$Z(f) = (\mathcal{F}\delta_{\tau})(f)X(f) = e_{-\tau}(f)X(f)$$

Modulation / Translation en fréquence.

$$\begin{aligned} \mathcal{F}(e_{f_0} x)(f) &= \int_{\mathbb{R}} e^{-i2\pi f t} \left(e^{i2\pi f_0 t} x(t) \right) dt \\ &= \int_{\mathbb{R}} e^{-i2\pi(f-f_0)t} x(t) dt \\ &= X(f - f_0). \end{aligned}$$

On arrive directement au résultat en exploitant le résultat sur la TF du produit de convolution :

$$\mathcal{F}(e_{f_0} x) = (\mathcal{F}e_{f_0}) * X = \delta_{f_0} * X.$$

Dérivée. La TF existe si $\lim_{t \rightarrow \infty} x(t) = 0$. Une intégration par parties conduit à :

$$\begin{aligned} \mathcal{F}\{x'\}(f) &= \int_{\mathbb{R}} e^{-i2\pi f t} x'(t) dt \\ &= \lim_{a \rightarrow \infty} \left(e^{-i2\pi f a} x(a) - e^{i2\pi f a} x(-a) \right) + \int_{\mathbb{R}} (2i\pi f e^{-i2\pi f t}) x(t) dt \\ &= (2i\pi f) X(f). \end{aligned}$$

Par récurrence on conclut que

$$\mathcal{F}\{x^{(n)}\}(f) = (i2\pi f)^n X(f).$$

Multiplication par t^n . La dérivée de la TF est

$$\begin{aligned} X'(f) &= \frac{d}{df} \int_{\mathbb{R}} e^{-i2\pi f t} x(t) dt \\ &= (-i2\pi) \frac{d}{df} \int_{\mathbb{R}} e^{-i2\pi f t} t x(t) dt \\ &= (-i2\pi) \mathcal{F}\{t x(t)\}(f) \end{aligned}$$

Nous avons donc $\mathcal{F}\{t x(t)\}(f) = \frac{1}{(-i2\pi)} X'(f)$. La propriété

$$\mathcal{F}\{t^n x(t)\}(f) = \frac{1}{(-i2\pi)^n} X^{(n)}(f)$$

se déduit par récurrence.

ECHANTILLONNAGE

L'application d'un traitement numérique à un signal provenant du monde réel, à temps continu, nécessite que le signal soit représenté sous forme discrète, *i.e.*, une suite de nombres (réels ou complexes), qu'il est possible de stocker dans la mémoire d'un ordinateur.

La notion de discrétisation a déjà été abordée dans les cours de mathématiques :

- en CIP, l'existence de bases orthogonales des espaces de Hilbert séparables, par exemple l'isomorphisme entre les fonctions L^2 sur $[0, 1]$ et les suites $\ell^2(\mathbb{Z})$ donné par la série de Fourier,
- en EDP, discrétisation par éléments finis ou différences finies, où la régularité des fonctions (au sens $\mathcal{C}^k$ ou des espaces de Sobolev) permet de montrer la convergence de la discrétisation vers la fonction approximée.

Dans ce cours, la discrétisation des signaux est abordée sous l'angle de l'échantillonnage, qui consiste à répondre à la question suivante : sous quelles conditions est-il possible de reconstituer un signal à temps continu $x(t)$ à partir de ses valeurs x_n à certains temps t_n ?

Nous verrons dans le cas d'un échantillonnage uniforme qu'une condition suffisante pour reconstituer un signal à partir de ses échantillons est que son spectre soit à support compact, et que le signal original peut-être reconstitué de façon simple à partir de ses échantillons.

4.1 EFFET DE L'ÉCHANTILLONNAGE

Définition 4.1 (Echantillonnage). On appelle *échantillonnage* l'opération qui consiste à construire un signal à temps discret $x_d[n]$ à partir d'un signal à temps continu $x(t)$. On se limitera ici à l'échantillonnage uniforme, c'est-à-dire, étant donnée une période d'échantillonnage $T_e > 0$,

$$x_d[n] = x(nT_e). \quad (117)$$

On définit également la fréquence d'échantillonnage $f_e = 1/T_e$.

Il est évidemment nécessaire que le signal x soit continu, sans quoi l'équation (117) n'aurait pas de sens. On décide de se placer dans le

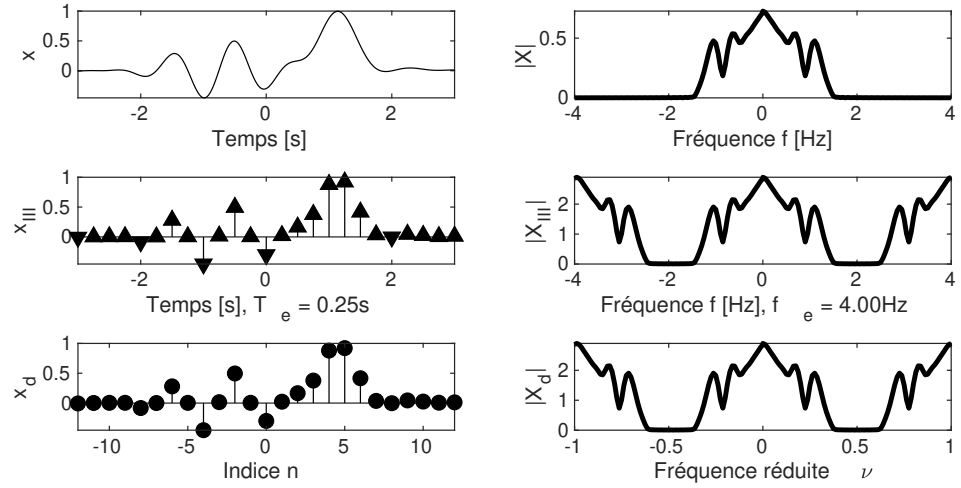


FIGURE 24 – Cas où le critère de Shannon est respecté. Gauche : Signaux continus x et x_{Δ} , et signal discret x_d . Droite : leurs spectres. Le spectre $X_{\Delta}(f)$ est f_e périodique, son motif période est $X(f)$ à un facteur multiplicatif près. De plus, x_d étant un signal discret, $X_d(\nu)$ est 1-périodique, et le graphe est identique à celui de $X_{\Delta}(f)$, à un changement d'échelle près.

cas où $x(t)$ est tel que sa transformée de Fourier est dans L^1 . On sait en effet que la transformée de Fourier d'une fonction L^1 est continue et bornée, et que les transformées de Fourier directe et inverse ont (quasiment) les mêmes propriétés, ce qui assure que $x(t)$ est continu et borné. Dans ce cas, le signal discret x_d est dans ℓ^∞ .

On définit le signal x_{Δ} par :

$$x_{\Delta} = \sum_{n \in \mathbb{Z}} x(nT_e) \delta_{nT_e}. \quad (118)$$

qu'on peut réécrire comme le produit de $x(t)$ et du peigne de Dirac Δ_{T_e} :

$$x_{\Delta} = \Delta_{T_e} x. \quad (119)$$

Ce signal modèle permet de faire le lien entre le signal $x(t)$ à temps continu, et le signal $x_d[n]$ à temps discret. En effet, on peut relier x_{Δ} à x_d par l'équation

$$x_{\Delta} = \sum_{n \in \mathbb{Z}} x_d[n] \delta_{nT_e}. \quad (120)$$

La colonne de gauche de la figure 24 donne une représentation de ces trois signaux dans un cas particulier.

Le théorème suivant relie les spectres de x , x_{Δ} et x_d :

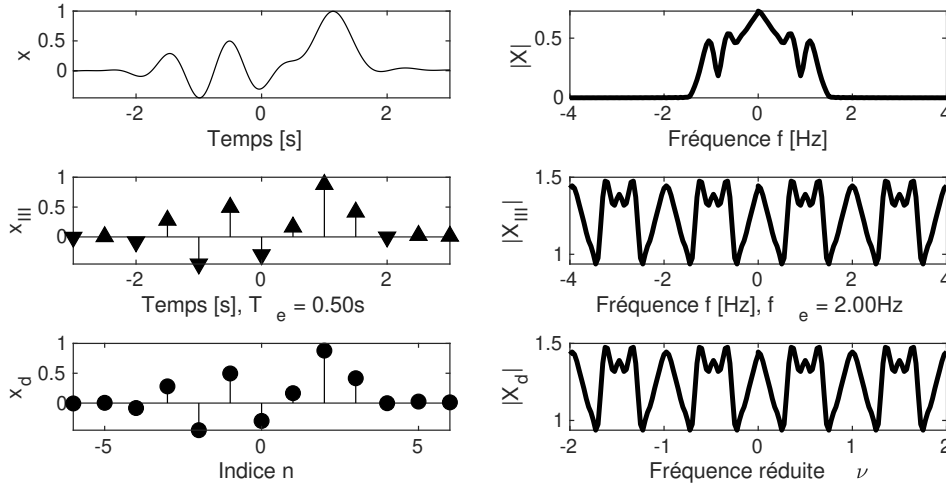


FIGURE 25 – Cas où le critère de Shannon n'est pas respecté. Gauche : Signaux continus x et x_{III} , et signal discret x_d . Droite : leurs spectres. Dans le spectre $X_{III}(f)$, les copies du spectre $X(f)$ se superposent et empêchent la reconstruction du signal.

Théorème 4.1. Soit $x(t)$ un signal tel que sa transformée de Fourier $X(f)$ soit dans L^1 , et x_{III} et x_d définis comme plus haut. La transformée de Fourier X_{III} de x_{III} est donnée par :

$$X_{III}(f) = f_e \sum_{n \in \mathbb{Z}} X(f - nf_e) \quad (121)$$

et la transformée de Fourier à temps discret de x_d est donnée par :

$$X_d(v) = f_e \sum_{n \in \mathbb{Z}} X(vf_e - nf_e) \quad (122)$$

Les séries convergent au sens L^1 sur l'intervalle $]-\frac{f_e}{2}, \frac{f_e}{2}[$, resp. $]-\frac{1}{2}, \frac{1}{2}[$

Démonstration. X_{III} est le spectre du produit de $\mathbb{I}_{T_e}$ et de x . C'est donc la convolution des spectres de ces deux signaux. D'après le théorème 3.6 sur la TF d'un peigne de Dirac,

$$X_{III}(f) = \left(\frac{1}{T_e} \mathbb{I}_{\frac{1}{T_e}} * X \right)(f) = f_e \sum_{n \in \mathbb{Z}} X(f - nf_e).$$

Dans $L^1(]-\frac{f_e}{2}, \frac{f_e}{2}[)$, la suite $N \rightarrow \sum_{n=-N}^N X(f - nf_e)$ est de Cauchy par le théorème

de convergence dominée, et convergente par complétude de $L^1(]-\frac{f_e}{2}, \frac{f_e}{2}[)$.

La deuxième égalité du théorème s'obtient en calculant la transformée de Fourier de (120) :

$$X_{III}(f) = \sum_{n \in \mathbb{Z}} x_d[n] e^{-i2\pi f n T_e} = X_d(f T_e),$$

en posant $v = f T_e = f / f_e$. $\square$

On retiendra de ce théorème que **la TF d'un signal échantillonné à la fréquence f_e est f_e -périodique. Inversement, la TF d'un signal périodique est un spectre de raies** (cf. section 3.2).

INTERPRÉTATION DU THÉORÈME 4.1 : l'opération d'échantillonnage en temps est équivalente à périodiser le spectre du signal. Il est donc, en général, impossible de reconstituer un signal à partir de ses échantillons, différentes copies du spectre pouvant se superposer et empêcher la reconstruction du spectre original. Ceci est visible sur la figure 25.

On remarque cependant que si les versions du spectre décalées de multiples de f_e ont des supports disjoints, comme c'est le cas sur la figure 24, il est possible de reconstituer le spectre du signal d'original.

Cette observation est formalisée dans la section suivante.

Pour aller plus loin...

Le théorème 4.1 est également valable dans le cas où $x(t)$ est tel que sa transformée de Fourier soit une distribution à support compact (par exemple, un signal combinaison linéaire finie de sinusoides). En effet, de telles fonctions sont continues (plus précisément, C^∞ à croissance au plus polynomiale). De plus, les séries du théorème se réduisent à des sommes finies et ne posent pas de problème de convergence.

4.2 THÉORÈME DE SHANNON

On montre dans cette partie, que si un **signal d'énergie finie** est à bande limitée, et qu'il est échantillonné avec une période suffisamment petite, il peut être reconstitué de manière exacte.

Définition 4.2 (Signal à bande limitée). *Un signal à bande limitée est tel que le support de son spectre est inclus dans $[-B, B]$ pour un B fini. On appelle PW_B l'ensemble des fonctions L^2 dont le support du spectre est inclus dans l'intervalle $[-B, B]$. PW_B est donc l'image réciproque par la transformée de Fourier des fonctions L^2 à support dans $[-B, B]$.*

Par exemple, la fonction définie par $x(t) = \sin_c(\pi t)$, où on rappelle la définition de $\sin_c$ donnée par l'équation (74) :

$$\sin_c(\theta) := \begin{cases} \frac{\sin(\theta)}{\theta} & \text{si } \theta \neq 0 \\ 1 & \text{si } \theta = 0 \end{cases}$$

vérifie $x \in PW_{\frac{1}{2}}$ car sa TF vaut $X = \mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]}$. En utilisant les propriétés de dilatation-contraction de la TF, la fonction définie $x_B(t) = \sin_c(2\pi Bt)$ pour $B > 0$ vérifie $x_B \in PW_B$ car sa TF vaut $X_B = \frac{1}{2B} \mathbb{1}_{[-B, B]}$.

Les fonctions de PW_B vérifient l'hypothèse du théorème 4.1 (en effet, la TF d'une fonction $x \in L^2$ est dans L^2 , de plus toute fonction $X \in L^2$ à support compact est L^1). Le **théorème de Shannon**, aussi appelé **théorème d'échantillonnage**, est un théorème fondamental qui établit les conditions de conservation de l'information contenue dans un signal $x(t)$ à bande limitée lors de son échantillonnage.

Théorème 4.2 (Théorème de Shannon). Soit $x \in PW_B$, et $T_e = 1/(2B)$. Alors pour tout $t \in \mathbb{R}$:

$$\begin{aligned} x(t) &= \sum_{n \in \mathbb{Z}} x(nT_e) \operatorname{sinc}(\pi(t - nT_e)/T_e) \\ &= \sum_{n \in \mathbb{Z}} x_d[n] \operatorname{sinc}(\pi(t - nT_e)/T_e) \end{aligned} \quad (123)$$

Remarque. Comme $x \in PW_B$ implique $x \in PW_{B'}$ avec $B' \geq B$, le théorème est bien évidemment valable pour tout $T_e \leq 1/2B$.

CRITÈRE DE SHANNON : Dans le cas de fonction d'énergie finie et à bande limité dans $[-B, B]$, il est suffisant d'échantillonner à la fréquence

$$f_e \geq 2B \quad (\text{critère de Shannon})$$

pour reconstruire exactement le signal $x(t)$ à partir de ses échantillons $x_d[n]$. Sans autre information (par ex., trous dans le spectre), cette condition est également nécessaire : un échantillonnage à la fréquence f_e ne peut caractériser de façon unique que les signaux à spectre limité dans la bande $[-f_e/2, f_e/2]$.

INTERPRÉTATION DU THÉORÈME DE SHANNON : L'équation (123) montre que $x(t)$ peut être interpolé à partir de ses échantillons $x_d[n]$. (123) peut s'interpréter comme la convolution de $x_{\square}$ et du sinus cardinal $\operatorname{sinc}(\pi t/f_e)$. Le spectre de cette convolution est le spectre de $x_{\square}$ multiplié par la fenêtre rectangulaire $\mathbb{1}_{[-B, B]}$ (transformée de Fourier du sinus cardinal). Dans la somme (121), et sous la condition que $x \in PW_B$,

- le support du terme $X(f)$ pour $n = 0$ est inclus dans le support de la fenêtre,
- les supports des termes $X(f - nf_e)$ pour $n \neq 0$ ne s'intersectent pas avec le support de la fenêtre.

La formule de reconstruction (123) peut donc être interprétée comme un filtrage des repliements du spectre du signal original introduits par la discrétisation.

Les figures 26 et 27 montrent des reconstructions de signaux dans les cas où le critère de Shannon est respecté ou non.

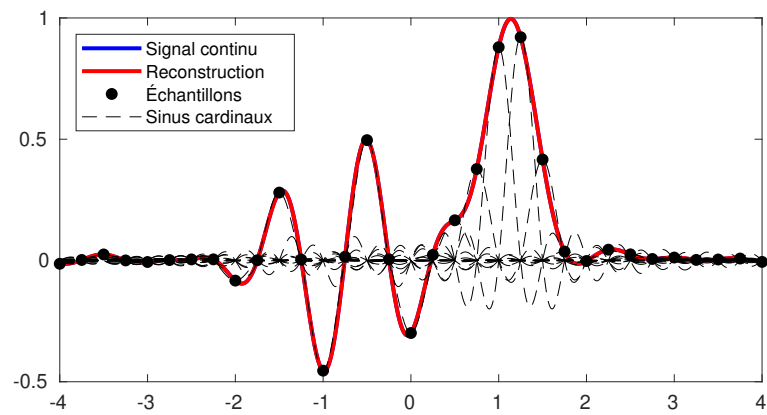


FIGURE 26 – Cas où le critère de Shannon est respecté. Reconstruction du signal par la somme (123).

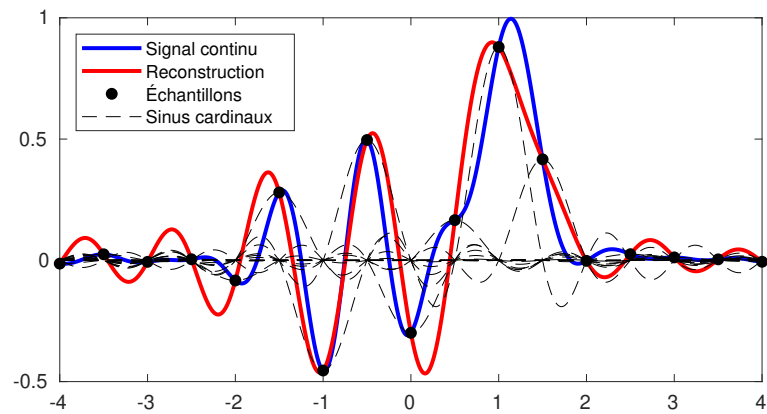


FIGURE 27 – Cas où le critère de Shannon n'est pas respecté. Reconstruction du signal par la somme (123).

Le théorème est démontré en deux étapes :

1. construction d'une base orthogonale de PW_B , permettant d'assurer la convergence de (123) dans L^2 .
2. analyse de la convergence simple de la série.

4.2.1 Base orthogonale de PW_B

Lemme 4.1. La famille de fonctions s_n pour $n \in \mathbb{Z}$ définie par

$$s_n(t) = \frac{1}{\sqrt{T_e}} \frac{\sin(\pi(t - nT_e)/T_e)}{\pi(t - nT_e)/T_e}. \quad (124)$$

est une base orthogonale de PW_B .

Les coefficients x_n d'une fonction $x \in PW_B$ dans cette base sont donnés par :

$$x_n = \sqrt{T_e} x(nT_e), \quad (125)$$

et

$$x = \sum_{n \in \mathbb{Z}} x_n s_n \quad (126)$$

au sens L^2 , où $T_e = 1/(2B)$.

Démonstration. La base de Fourier de $L^2([-B, B])$ est donnée par :

$$\forall n \in \mathbb{Z}, S_n(f) = \frac{1}{\sqrt{2B}} \exp\left(-i2\pi \frac{nf}{2B}\right) \mathbb{1}_{[-B, B]}(f).$$

La transformée de Fourier étant une isométrie, les transformées de Fourier inverses des S_n forment une base de PW_B . Or, les transformées inverses des S_n sont les s_n . Les coefficients X_n de $X(f)$ dans la base des S_n sont les mêmes que les coefficients x_n de x dans la base des s_n (Parseval).

Les coefficients de décomposition de X dans la base de Fourier sont donnés par

$$X_n = \frac{1}{\sqrt{2B}} \int_{[-B, B]} X(f) \exp\left(i2\pi \frac{nf}{2B}\right) df$$

où $X = \sum_{n \in \mathbb{Z}} X_n s_n$. En interprétation l'intégrale comme la transformée de Fourier inverse de $X(f)$, évaluée en $t = n/(2B) = nT_e$, on trouve :

$$x_n = X_n = \sqrt{T_e} x(nT_e)$$

□

4.2.2 Convergence de la série

On démontre maintenant la convergence simple de la série.

La famille s_n étant une base orthogonale, on observe que

$$\|x_n\|_2^2 = \|x\|_{L^2}^2, \quad (127)$$

ce qui montre que les coefficients x_n sont dans ℓ^2 . De plus, pour tout t , la suite de terme général $s_n(t)$ est de carré intégrable. En interprétant (123) comme le produit scalaire de deux suites de termes x_n et $s_n(t)$,

on en déduit la convergence de la série par l'inégalité de Cauchy-Schwarz.

On démontre, de plus, que la somme de la série est continue. Sur l'intervalle $t \in [-T_e, T_e]$:

- pour tout t , la suite $x_n s_n(t)$ est sommable
- pour tout n , $x_n s_n(t)$ est continue
- pour tout t dans l'intervalle, $|x_n s_n(t)| \leq |x_n| \max(1, (\pi(|n| - 1))^{-1})$, sommable comme produit de deux suites de ℓ^2 .

On en déduit la continuité de la série sur $[-T_e, T_e]$. On peut reprendre cette démonstration sur tout intervalle de la forme $[-T_e + kT_e, T_e + kT_e]$ pour démontrer que continuité sur $\mathbb{R}$ entier. La somme de la série étant continue, et égale à la fonction continue $x(t)$ au sens L^2 , la série convergence simplement vers $x(t)$ pour tout t .

Pour aller plus loin...

Dans le cas de signaux dont la transformée de Fourier est une distribution à support dans $[-B, B]$ (ce qui assure continuité et croissance au plus polynomiale), le théorème doit être légèrement modifié. La fréquence d'échantillonnage $f_e = 1/T_e$ doit être strictement plus grande que $2B$, et le signal $x(t)$ est reconstruit par la formule

$$\forall t \in \mathbb{R}, x(t) = \sum_{n \in \mathbb{Z}} x(nT_e) \phi(t - nT_e) \quad (128)$$

où ϕ est choisie telle que sa transformée de Fourier Φ vérifie les conditions suivantes :

- $\Phi \in C^\infty$,
- $\Phi(f) = 1/f_e$ pour $|f| \leq B$,
- $\Phi(f) = 0$ pour $|f| \geq f_e/2$.

La série (128) peut être écrite comme la convolution $x_{\square} * \phi$. Son spectre est donné par

$$X_{\square}(f) \Phi(f) = f_e \sum_{n \in \mathbb{Z}} X(f - nf_e) \Phi(f). \quad (129)$$

Les conditions sur les valeurs de Φ et le support de $X(f)$ permettent de sélectionner uniquement le terme de la série pour $n = 0$, i.e., de reconstituer exactement $X(f)$.

De plus, la régularité de Φ est nécessaire pour garantir la convergence simple de (128). En effet, les échantillons $x(nT_e)$ ne sont plus bornés, mais seulement à croissance au plus polynomiale. La régularité de Φ implique que ϕ décroît plus vite que tout polynôme, ce qui garantit la convergence de la série.

4.3 L'ÉCHANTILLONNAGE EN PRATIQUE

4.3.1 Filtre anti-repliement

Le théorème de Shannon suppose que le signal à échantillonner a un spectre à support compact. Ce n'est évidemment pas toujours le cas en pratique. De plus, même quand le spectre du signal à échantillonner est à support borné, il n'est pas toujours possible d'utiliser une fréquence d'échantillonnage assez élevée.

L'échantillonnage est donc toujours précédé par un filtrage passe-bas du signal, coupant la bande de fréquence au dessus de $f_e/2$. Il est à noter que ce filtre est forcément implémenté sous forme analogique.

4.3.2 Reconstruction pratique

Les formules de reconstruction (123) et (128) sont inutilisables en pratique. En effet, la reconstruction de x à un temps t nécessite la connaissance de tous les échantillons $x(nT_e)$, y compris ceux situés dans le futur.

Le **bloquage d'ordre zéro** est une technique simple permettant d'approximer le signal à reconstruire. Le signal bloqué $x_0(t)$ est le signal en escalier défini par :

$$x_0(t) = \sum_{n \in \mathbb{Z}} x_d[n] \mathbb{1}_{[nT_e, (n+1)T_e[}(t). \quad (130)$$

La reconstruction au temps t ne nécessite que la connaissance de l'échantillon à l'instant précédent nT_e , où $n = \lfloor t/T_e \rfloor$.

L'effet du bloquage d'ordre zéro sur le spectre peut être étudié en écrivant x_0 comme une convolution

$$x_0 = \mathbb{1}_{[0, T_e[} * x_{\square}. \quad (131)$$

Le spectre de x_0 est donc donné par

$$X_0(f) = X_{\square}(f) \left(T_e e^{-i\pi f T_e} \frac{\sin \pi f T_e}{\pi f T_e} \right) \quad (132)$$

Les effets du blocage d'ordre zéro sont :

- l'introduction d'un décalage temporel de durée $T_e/2$ (terme $e^{-i\pi f T_e}$),
- un filtrage du signal dans la bande utile ($(\sin \pi f T_e) / (\pi f T_e)$ non constant dans la bande utile),
- la présence de repliements du spectre atténués en $1/f$.

Les deux derniers effets peuvent être limités par l'utilisation d'un filtrage analogique passe-bas appliqué à x_0 .

4.4 CONCLUSION

La théorie de l'échantillonnage permet de relier les spectres d'un signal continu $x(t)$ et de sa version échantillonnée x_d . Nous retiendrons en premier lieu que l'échantillonnage d'un signal avec une période T_e se traduit par une périodisation du spectre avec la période $f_e = 1/T_e$. De plus, l'information contenue dans $x(t)$ est préservée si le critère de Shannon est satisfait, à savoir $f_e \geq 2B$ où B est la fréquence maximale contenue dans le signal x . Dans ce cas, on peut (théoriquement) reconstruire parfaitement $x(t)$ à partir des valeurs de $x_d[n]$, $n \in \mathbb{Z}$.

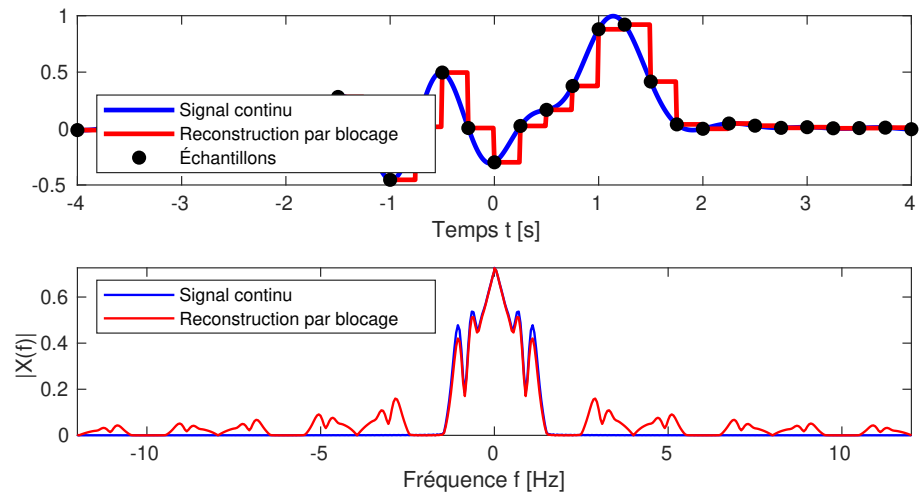


FIGURE 28 – Reconstruction par blocage d'ordre 0. Le filtrage du signal et les copies générées par le blocage sont visibles sur le spectre.

Si le critère de Shannon n'est pas satisfait, on utilise un filtre anti-repliement (passe-bas) en amont de l'échantillonnage. Dans ce cas, le signal reconstruit à partir des échantillons est une version approchée du signal $x(t)$, contenant ses composantes basse-fréquences.

PROCESSUS ALÉATOIRES

La théorie des signaux déterministes atteint ses limites pour traiter et modéliser des phénomènes physiques pour lesquels il n'est plus possible d'expliciter les valeurs des signaux pour chaque instant (typiquement à l'aide d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$). On utilise alors des processus aléatoires qui sont des modèles probabilistes permettant de caractériser les propriétés statistiques des signaux et de construire des traitements exploitant ces propriétés.

Prenons l'exemple d'une mesure physique entachée de bruit. Le cadre déterministe ne permet pas de proposer un modèle du bruit qui expliciterait la valeur du signal pour chaque instant. Grâce au processus aléatoire, il est possible de caractériser les propriétés statistiques du bruit et du signal utile afin de réaliser un traitement de débruitage pour réduire l'influence du bruit (un tel traitement sera détaillé à la section 6.4.3 du chapitre 6).

Considérons un deuxième exemple : en traitement de la parole, deux sons peuvent sembler identiques d'un point de vue auditif alors que les valeurs des deux signaux associés sont très différentes. L'étude des propriétés statistiques de ces signaux permet de classer les différents sons par exemple dans le cadre d'applications de reconnaissance vocale. A partir de ces propriétés statistiques, il est aussi possible de synthétiser de nouveaux signaux très similaires du point de vue auditif (par exemple pour la synthèse vocale). Ces aspects seront succinctement abordés dans le cadre d'un exemple illustrant la section 6.3.1 du chapitre 6.

5.1 DÉFINITION ET CARACTÉRISATION DES PROCESSUS ALÉATOIRES

Les processus aléatoires (à temps discrets) sont une généralisation des vecteurs aléatoires afin de construire des modèles de signaux à support infini contrairement aux vecteurs aléatoires, définis à partir de N variables aléatoires scalaires et sont donc de dimension finie.

5.1.1 Définition

Dans le cadre de ce cours, nous nous limiterons à l'étude des processus aléatoires à temps discrets.

Définition 5.1 (processus aléatoire discret). *Un processus aléatoire discret* noté X_n est une suite de variable aléatoires $\{X_n(\omega)\}_{n \in \mathbb{Z}}$ définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathcal{P})$.

En traitement du signal, l'indice n correspond à la variable d'évolution du signal (par exemple le temps ou l'espace).

On appelle **réalisation**¹ du processus X_n , le signal déterministe, noté $x[n]$, construit à partir d'une réalisation de la famille de variables aléatoires $\{X_n(\omega)\}_{n \in \mathbb{Z}}$. Ainsi pour un ω donné, on a :

$$\forall n \in \mathbb{Z}, \quad x[n] = X_n(\omega)$$

Il convient de bien distinguer deux notions complémentaires :

- d'une part le processus aléatoire qui est une suite de variables aléatoires, pour modéliser les signaux afin d'établir certaines propriétés (par ex. la loi de probabilité, la moyenne, la corrélation...)
- d'autre part la (ou les) réalisation(s) du processus qui sont les signaux observés.

En traitement du signal et en statistique, on utilise les réalisations pour estimer les propriétés du processus aléatoire. La spécificité du traitement du signal est que l'on exploite les propriétés du processus aléatoire pour traiter les signaux observés (les réalisations) afin d'améliorer leur qualité (par ex. suppression de bruit ou déconvolution pour supprimer le flou introduit par un appareil de mesure).

Remarque. On utilise couramment aussi les termes *processus stochastiques* ou *signaux aléatoires* pour désigner les processus aléatoires. La réalisation d'un processus aléatoire est aussi souvent appelée *trajectoire* ou *observation*.

Exemple 5.1. Voici trois exemples de processus avec une représentation de quelques-unes de leurs réalisations

- Exemple 1 : Suite de variables i.i.d
 $X_n = A_n$ avec $A_n \sim \mathcal{N}(0, \sigma)$ et A_n i.i.d
- Exemple 2 : Processus constant à valeur aléatoire
 $Y_n = A$ avec $A \sim \mathcal{N}(0, \sigma)$
- Exemple 3 : Modulation d'amplitude par des variables aléatoires
 $Z_n = A \cos(2\pi\nu_0 n) + B \sin(2\pi\nu_0 n)$ où $(A, B) \sim \mathcal{N}(\mathbf{0}, \Sigma)$.

1. En statistique on parle de données plutôt que de réalisations

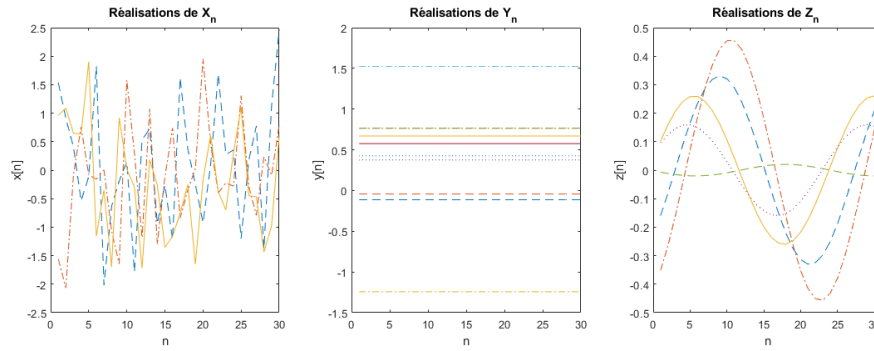


FIGURE 29 – Exemples de réalisations de processus aléatoires

Pour aller plus loin...

Processus aléatoire à temps continu.

Il existe aussi des processus continus. La définition de processus aléatoire au sens général est la suivante.

Définition 5.2. Un processus aléatoire $X(\omega, t)$ est une application définie par

$$\begin{aligned} X &: \Omega \times \mathcal{T} \longrightarrow K \\ (\omega, t) &\longmapsto X(\omega, t), \end{aligned}$$

où

- $(\Omega, \mathcal{E}, P)$ est un espace probabilisé,
- K est un espace topologique ou mesurable (en général un espace vectoriel réel ou complexe normé),
- $\mathcal{T}$ est un ensemble,

et telle que pour tout $t \in \mathcal{T}$, l'application partielle

$$\begin{aligned} \Omega &\longrightarrow K \\ \omega &\longmapsto X(\omega, t) \end{aligned}$$

soit une variable aléatoire.

La variable t est la variable d'évolution du signal, il s'agit en général du temps ou de l'espace.

Pour ω_0 fixé, $X(\omega_0, t)$ est une réalisation du processus. C'est un signal qui ne dépend plus que de la variable d'évolution t .

Selon l'ensemble défini par $\mathcal{T}$, on obtient un processus discret ou continu :

- $t \in \mathbb{R}$: processus continu
- $t \in \mathbb{Z}$: processus discret (on note plutôt $X(\omega, n)$ avec $n \in \mathbb{Z}$)

5.1.2 Lois fini-dimensionnelles

Les lois fini-dimensionnelles sont la généralisation pour un processus aléatoire, de la loi de probabilité d'un vecteur aléatoire.

Définition 5.3 (lois fini-dimensionnelle). *Pour tout $N > 0$, on définit les lois fini-dimensionnelles d'ordre N comme les lois de probabilité jointes définies pour tout $(n_1, n_2, \dots, n_N) \in \mathbb{Z}^N$ par*

$$P(X_{n_1} \in [a_1, b_1], X_{n_2} \in [a_2, b_2], \dots, X_{n_N} \in [a_N, b_N])$$

Un processus aléatoire est entièrement décrit par l'ensemble de ses lois fini-dimensionnelles. Cependant, il s'agit d'une caractérisation en général très complexe et difficile à expliciter. En pratique, on utilise donc rarement cette caractérisation.

5.1.3 Processus gaussien

Définition 5.4 (Processus gaussien). *Un processus est gaussien si ses lois fini-dimensionnelles sont normales :*

$$\forall N > 0, \forall (n_1, n_2, \dots, n_N) \in \mathbb{Z}^N,$$

$$(X_{n_1}, X_{n_2}, \dots, X_{n_N}) \sim \mathcal{N}(\mu_{n_1, n_2, \dots, n_N}, \Sigma_{n_1, n_2, \dots, n_N})$$

ou de manière équivalente si toute combinaison linéaire finie de X_n est une variable aléatoire gaussienne :

$$\forall N > 0, \forall (n_1, n_2, \dots, n_N) \in \mathbb{Z}^N, \forall (\alpha_1, \alpha_2, \dots, \alpha_N) \in \mathbb{R}^N,$$

$$\exists \mu, \sigma : \sum_{i=1}^N \alpha_i X_{n_i} \sim \mathcal{N}(\mu, \sigma)$$

Il est important de remarquer que la gaussianité des lois marginales de chaque échantillon X_n n'est pas suffisante pour assurer que le processus est gaussien. Ainsi on peut construire un processus X non gaussien bien que les variables aléatoires X_n soient gaussiennes pour tout n .

5.1.4 Moyenne et autocorrélation

Nous avons vu précédemment que la description complète d'un processus par ses lois fini-dimensionnelles est une tâche assez complexe. C'est pourquoi, on se contente couramment de caractériser partiellement les processus à l'aide de leurs deux premiers moments : la **moyenne** et l'**autocorrélation**.

En pratique cette caractérisation partielle est bien souvent suffisante pour de nombreuses applications. Par exemple, pour caractériser différents phonèmes de parole, l'étude des deux premiers moments du signal permettra de distinguer un son /f/, d'un son /s/.

Une deuxième motivation pour s'intéresser à ces deux premiers moments est que dans le cas particulier des processus gaussiens, ils caractérisent entièrement le processus. Il est bien sûr important de noter que dans le cas général, il s'agit d'une caractérisation partielle.

Définition 5.5 (Moyenne). On définit la **moyenne** (ou **moment d'ordre 1**) d'un processus X_n par

$$\mu_X[n] = \mathbb{E}(X_n)$$

$\mu_X[n]$ donne la valeur moyenne du processus à l'instant n (c'est un signal déterministe).

Définition 5.6 (Autocorrélation). On définit l'**autocorrélation** (ou **moment d'ordre 2**) d'un processus X_n par

$$r_X[n_1, n_2] = \mathbb{E}(X_{n_1} X_{n_2}^*)$$

$r_X[n_1, n_2]$ donne la corrélation du processus entre les instants n_1 et n_2 (c'est un signal déterministe à deux dimensions).

Pour la plupart des processus modélisant des phénomènes physiques, $r_X[n_1, n_2]$ décroît quand $|n_2 - n_1|$ devient grand. Intuitivement, cela correspond au fait que deux échantillons pris à des instants proches sont en général plus corrélés que deux échantillons très éloignés.

Définition 5.7 (Intercorrélation). On définit l'**intercorrélation** entre deux processus X_n et Y_n par

$$r_{XY}[n_1, n_2] = \mathbb{E}(X_{n_1} Y_{n_2}^*)$$

Remarque. On peut remarquer que $r_{XX}[n_1, n_2] = r_X[n_1, n_2]$, c'est à dire que l'intercorrélation d'un processus avec lui-même correspond à son autocorrélation.

Définition 5.8 (Décorrélation). Les processus X_n et Y_n sont dit **décorrélés** si

$$\forall n_1, n_2, \quad r_{XY}[n_1, n_2] - \mu_X[n_1] \mu_Y^*[n_2] = 0,$$

ou encore :

$$\forall n_1, n_2, \quad \mathbb{E}((X_{n_1} - \mu_X[n_1])(Y_{n_2}^* - \mu_Y^*[n_2])) = 0.$$

Remarque. si $Z = X + Y$ avec X et Y décorrélés et centrés alors :

$$r_Z[n_1, n_2] = r_X[n_1, n_2] + r_Y[n_1, n_2]$$

Pour aller plus loin...

On peut aussi définir le moment d'ordre 2 centré, dénommé **autocovariance** : $\mathbb{E}((X_{n_1} - \mu_X[n_1])(Y_{n_2}^* - \mu_Y^*[n_2]))$. De la même manière, on définit aussi l'**intercovariance** $\mathbb{E}((X_{n_1} - \mu_X[n_1])(Y_{n_2}^* - \mu_Y^*[n_2]))$.

Si X_n et Y_n sont centrés, leur auto/inter-covariances coïncident avec leur auto/inter-corrélation.

5.2 PROCESSUS STATIONNAIRES

La stationnarité est une propriété extrêmement importante qui permet de nombreuses simplifications calculatoires pour étudier les processus. La stationnarité se traduit par le fait que les propriétés du processus ne changent pas au cours du temps. Un exemple trivial de processus non stationnaire est un processus dont la moyenne $\mu_X[n]$ n'est pas constante en fonction de n .

5.2.1 Stationnarité stricte

Définition 5.9 (Stationnarité stricte). Un processus aléatoire est dit **strictement stationnaire** si les lois fini-dimensionnelles du processus sont invariantes par translation. C'est à dire que, $\forall k \in \mathbb{Z}, \forall N > 0$, les vecteurs aléatoires $(X_{n_1}, \dots, X_{n_N})$ et $(X_{n_1+k}, \dots, X_{n_N+k})$ ont des lois identiques.

La stationnarité au sens strict est en général difficile à établir et d'un intérêt assez limité². C'est pourquoi par la suite, nous nous intéresserons uniquement à la stationnarité au sens large.

5.2.2 Stationnarité au sens large (SSL)

Définition 5.10 (Stationnarité au sens large (SSL)). Un processus est dit **stationnaire au sens large (SSL)** (ou **du second ordre** ou **stationnaire au sens faible**) si :

- la moyenne est constante (elle ne dépend pas du temps) :

$$\forall n, \quad \mu_X[n] = \mu_X$$

- $\mathbb{E}(|X_n|^2) < \infty$

2. elle est souvent surdimensionnée, au sens où une propriété de stationnarité moins forte (la stationnarité au sens large, définie plus loin) est suffisante pour mener à bien la plupart des calculs.

- l'autocorrélation ne dépend que l'écart de temps entre les deux instants considérés :

$$\forall n_1, n_2 \quad r_X[n_1, n_2] = \gamma_X[n_1 - n_2]$$

On peut aussi remarquer que si X_n est SSL, $\mathbb{E}(|X_n|^2) = r_X[n, n]$ est aussi constant.

Pour l'autocorrélation d'un processus stationnaire, on utilisera donc la notation simplifiée :

$$\gamma_X[k] := r_X[n + k, n] = \mathbb{E}(X_{n+k}X_n^*).$$

Remarque. En raison de la stationnarité du processus, on peut écrire de manière équivalente

$$\gamma_X[k] = \mathbb{E}(X_{n+k}X_n^*) = \mathbb{E}(X_nX_{n-k}^*) = \mathbb{E}(X_{n+1+k}X_{n+1}^*) = \dots$$

De même, l'intercorrélation entre deux processus conjointement stationnaires s'écrit :

$$\gamma_{XY}[k] = \mathbb{E}(X_{n+k}Y_n^*).$$

Pour résumer, il convient de distinguer les deux fonctions :

- la fonction d'autocorrélation à deux variables (définie pour tous les processus) : $r_X[n_1, n_2] = \mathbb{E}(X_{n_1}X_{n_2}^*)$
- la fonction d'autocorrélation à une variable (définie uniquement pour les processus stationnaires) : $\gamma_X[k] = \mathbb{E}(X_{n+k}X_n^*) = r_X[n + k, n]$

On supposera dans la suite que tous les processus sont stationnaires au sens large, en particulier cela implique que leur moment d'ordre deux sont finis et on utilisera donc les notations γ_X et γ_{XY} .

Définition 5.11 (Variance). On définit la **variance** d'un processus stationnaire X_n par :

$$\sigma_X^2 = \mathbb{E}(|X_n - \mu_X|^2) = \mathbb{E}(|X_n|^2) - |\mu_X|^2$$

Définition 5.12 (Rapport signal-sur-bruit). La **rapport signal-sur-bruit** quantifie la perturbation d'un processus stationnaire $X_n = U_n + W_n$, où U_n est le signal utile et W_n le bruit perturbateur. Il est défini par le rapport de variances σ_U^2/σ_W^2 . Il est le plus souvent exprimé en dB :

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left(\frac{\sigma_U^2}{\sigma_W^2} \right). \quad (133)$$

Définition 5.13 (Puissance). La **puissance d'un processus stationnaire** X_n est la valeur de l'autocorrélation en zéro :

$$\mathbb{E}(|X_n|^2) = \gamma_X[0] = \sigma_X^2 + |\mu_X|^2$$

Propriété. L'autocorrélation d'un processus stationnaire est finie et maximum en zéro :

$$\forall k \in \mathbb{Z}, |\gamma_X[k]| \leq \gamma_X[0]$$

Démonstration. Le résultat est trivial en utilisant l'inégalité de Cauchy-Schwarz avec $\gamma_X[k] = \langle X_{n+k}, X_n \rangle$ (en notant $\langle \cdot, \cdot \rangle$ le produit scalaire entre variables aléatoires). Par ailleurs, par définition d'un processus stationnaire au sens large, $\gamma_X[0]$ est fini. $\square$

Propriété. L'autocorrélation est à symétrie hermitienne :

$$\forall k \in \mathbb{Z}, \gamma_X[k] = \gamma_X^*[-k]$$

Démonstration.

$$\gamma_X[-k] = \mathbb{E}(X_{n-k} X_n^*) = \mathbb{E}(X_{n-k}^* X_n)^* = \gamma_X^*[k].$$

$\square$

5.2.3 Densité spectrale de puissance (DSP)

Pour de nombreuses applications, il est important de pouvoir caractériser un processus aléatoire dans le domaine fréquentiel. Par exemple, les signaux de télécommunication sont couramment modélisés par des processus aléatoires et il est important de pouvoir quantifier la bande spectrale utilisée par ces signaux pour optimiser les canaux de transmissions.

L'outil principal qui permet de caractériser fréquemment les processus aléatoires stationnaires est la **Densité Spectrale de Puissance (DSP)**.

Définition 5.14. La **Densité Spectrale de Puissance (DSP)** d'un processus aléatoire stationnaire X_n est définie par :

$$\forall \nu \in \mathbb{R}, \Gamma_X(\nu) = \lim_{N \rightarrow \infty} \mathbb{E} \left(\frac{1}{2N+1} \left| \sum_{k=-N}^N X_k e^{-i2\pi k \nu} \right|^2 \right)$$

Théorème 5.1 (Wiener-Khinchin). *La densité spectrale de puissance d'un processus X_n est la transformée de Fourier de l'autocorrélation du processus :*

$$\forall \nu \in \mathbb{R}, \quad \Gamma_X(\nu) = \sum_{k \in \mathbb{Z}} \gamma_X[k] e^{-i2\pi \nu k}$$

Démonstration. Par définition de la DSP :

$$\Gamma_X(\nu) = \lim_{N \rightarrow \infty} \mathbb{E} \left(\frac{1}{2N+1} \left| \sum_{k=-N}^N X_k \exp(-i2\pi k\nu) \right|^2 \right) \quad (134)$$

On écrit :

$$\begin{aligned} \Gamma_X^N(\nu) &= \mathbb{E} \left(\frac{1}{2N+1} \left| \sum_{k=-N}^N X_k \exp(-i2\pi k\nu) \right|^2 \right) \\ &= \frac{1}{2N+1} \mathbb{E} \left(\left(\sum_{k=-N}^N X_k \exp(-i2\pi k\nu) \right)^* \left(\sum_{k=-N}^N X_k \exp(-i2\pi k\nu) \right) \right) \\ &= \frac{1}{2N+1} \mathbb{E} \left(\sum_{k=-N}^N \sum_{n=-N}^N X_k^* X_n \exp(-i2\pi(n-k)\nu) \right) \\ &= \frac{1}{2N+1} \sum_{k=-N}^N \sum_{n=-N}^N \mathbb{E}(X_k^* X_n) \exp(-i2\pi(n-k)\nu) \\ &= \frac{1}{2N+1} \sum_{k=-N}^N \sum_{n=-N}^N \gamma_X[n-k] \exp(-i2\pi(n-k)\nu) \\ &= \sum_{m=-2N}^{2N} \frac{2N+1-|m|}{2N+1} \gamma_X[m] \exp(-i2\pi m\nu) \\ &= \sum_{m \in \mathbb{Z}} \max \left(0, \frac{2N+1-|m|}{2N+1} \right) \gamma_X[m] \exp(-i2\pi m\nu) \end{aligned}$$

Dans le cas où $\gamma_X \in L^1$, on peut borner les termes de la série par $|\gamma_X[m]|$ et appliquer le théorème de convergence dominée pour intervertir limite et somme. On trouve :

$$\Gamma_X(\nu) = \lim_{N \rightarrow \infty} \Gamma_X^N(\nu) = \sum_{m \in \mathbb{Z}} \gamma_X[m] \exp(-i2\pi m\nu)$$

de plus, $\Gamma_X(\nu)$ est continue.

Dans le cas où γ_X n'est pas intégrable, il faut calculer la limite au sens des distributions pour montrer que

$$\lim_{N \rightarrow \infty} \langle P_X^N, \phi \rangle = \left\langle \sum_{m \in \mathbb{Z}} \gamma_X[m] \delta_m, \Phi \right\rangle$$

où est ϕ une fonction test et Φ sa transformée de Fourier. $\square$

En pratique, on utilise plus fréquemment l'expression de la densité spectrale de puissance donnée par le théorème de Wiener-Khinchin car il est, en général, plus simple de calculer la transformée de Fourier de l'autocorrélation du processus que d'appliquer le passage à la limite et l'espérance à partir de la définition de la DSP.

Propriété. La *densité spectrale de puissance est réelle et positive* :

$$\forall \nu, \Gamma_X(\nu) \in \mathbb{R}$$

$$\forall \nu, \Gamma_X(\nu) \geq 0$$

Ces propriétés se déduisent directement de la définition.

Propriété. La *puissance d'un processus est l'intégrale de la densité spectrale de puissance* :

$$\gamma_X[0] = \int_{-1/2}^{1/2} \Gamma_X(\nu) d\nu$$

Démonstration. D'après le théorème de Wiener-Khinchin, l'autocorrélation est la transformée de Fourier inverse de la DSP :

$$\gamma_X[n] = \int_{-1/2}^{1/2} \Gamma_X(\nu) e^{i2\pi\nu n} d\nu$$

Pour $n = 0$ on retrouve l'expression de la puissance. □

Ce résultat justifie la dénomination de densité spectrale de puissance puisqu'en intégrant la densité dans le domaine spectral, on obtient bien la puissance du processus. On peut, bien sûr, intégrer la DSP dans une bande fréquentielle restreinte pour obtenir la puissance du processus dans cette bande de fréquence.

Méthodologie de calcul

En pratique, la méthodologie usuelle pour calculer analytiquement la DSP d'un processus aléatoire est la suivante :

1. Calculer la moyenne $\mu_X[n] = \mathbb{E}(X_n)$
2. Calculer l'autocorrélation : $r_X[k_1, k_2] = \mathbb{E}(X_{k_1} X_{k_2}^*)$
3. Vérifier que le processus est SSL
 - $\mu_X[k] = \mu_X$,
 - $r_X[k_1, k_2] = \gamma_X[k_1 - k_2]$
4. Vérifier la cohérence des calculs à l'aide des propriétés de $\gamma_X[k]$:
 - $\gamma_X[0] \geq |\gamma_X[k]|$,
 - $\gamma_X[k] = \gamma_X^*[-k]$.
5. Calcul de la densité spectrale de puissance $\Gamma_X(\nu) = \mathcal{F}\gamma_X(\nu)$.

5.2.4 Bruit blanc

Le *bruit blanc* est un processus aléatoire couramment utilisé pour modéliser des phénomènes physiques, par exemple pour établir un modèle mathématique du bruit de mesure.

Définition 5.15 (bruit blanc au sens large). Un processus W_n est un **bruit blanc au sens large** de variance σ^2 s'il vérifie les propriétés suivantes :

- W_n est un processus centré et stationnaire au sens large,
- $\gamma_W[n] = \sigma^2 \delta[n]$.

On dit parfois qu'il s'agit d'un bruit blanc au sens faible.

Définition 5.16 (bruit blanc au sens strict). Un processus W_n est un **bruit blanc au sens strict** s'il vérifie les propriétés suivantes :

- W_n est un processus stationnaire au sens large,
- $\gamma_W[n] = \sigma^2 \delta[n]$,
- W_n et W_k sont indépendants $\forall k \neq n$.

Remarque. Un bruit blanc gaussien est toujours un bruit blanc au sens strict car la décorrélation implique l'indépendance pour les processus gaussiens.

Propriété. Un bruit blanc au sens large vérifie les propriétés suivantes :

- W_n et W_m sont **décorrélés** pour $n \neq m$
- La **DSP d'un bruit blanc est constante** : $\Gamma_W = \sigma^2 \mathbb{1}_{\mathbb{R}}$

Le bruit blanc tire son nom de cette dernière propriété par analogie avec la lumière blanche dont le spectre contient toutes les composantes fréquentielles dans les mêmes proportions.

Exemple 5.2. Soit deux processus : un bruit blanc (à gauche sur la figure 30), et un bruit coloré (à droite). La figure 30 représente les réalisations, les autocorrélations et les DSP de ces processus.

- **Réalisations** : D'un point de vue empirique, en observant une réalisation de chaque processus, on peut remarquer qu'à gauche, pour le bruit blanc, les valeurs successives semblent décorrélées entre elles ; tandis qu'à droite, on remarque une alternance positive/négative des valeurs traduisant une certaine corrélation.

Du point de vue théorique, on peut confirmer les commentaires précédents à partir des caractéristiques des processus :

- **Autocorrélation** : l'autocorrélation du bruit blanc est de forme impulsionnelle, tandis que l'autocorrélation du bruit coloré fait ressortir la corrélation positive entre les échantillons espacés d'un nombre de points pair et négative pour les intervalles impairs.
- **Densité Spectrale de Puissance** : la DSP du bruit blanc est constante tandis que la DSP du bruit coloré comporte principalement des hautes fréquences ce qui est cohérent avec les variations rapides observées sur la réalisation.

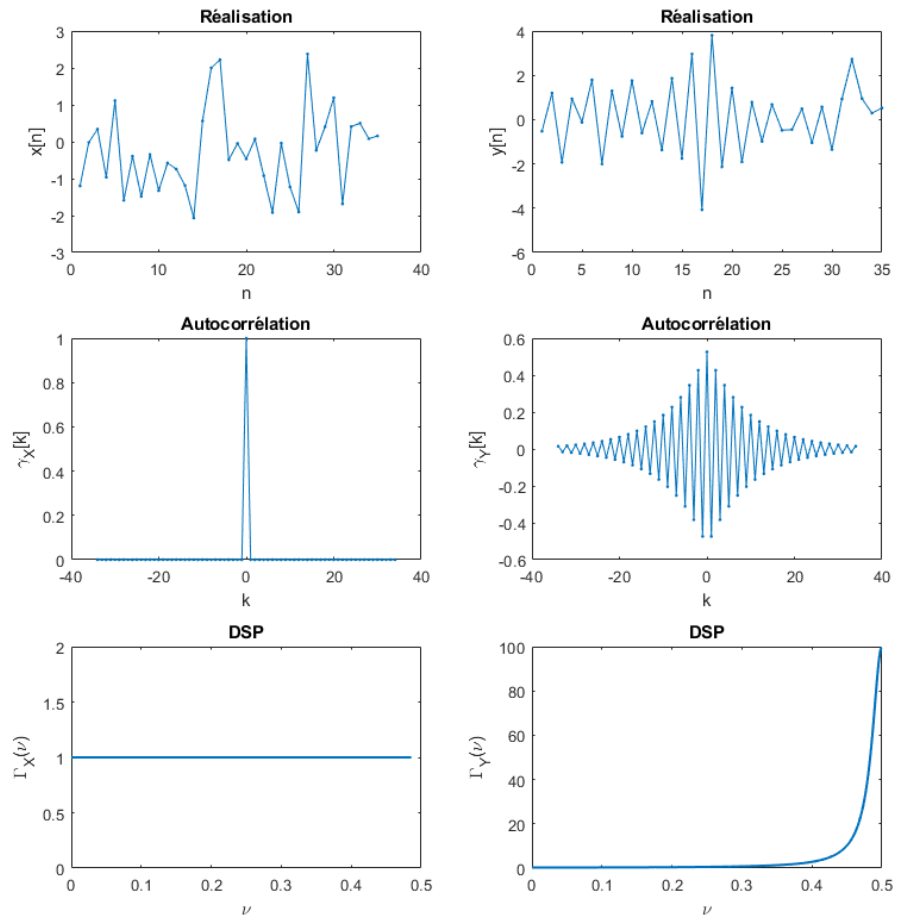


FIGURE 30 – Comparaison des caractéristiques de deux processus : un bruit blanc et un bruit coloré

5.2.5 Filtrage de processus

Le filtrage de signaux est l'une des principales opérations utilisées en traitement du signal, l'objet de cette section est de déterminer les propriétés d'un processus filtré.

Formule des interférences

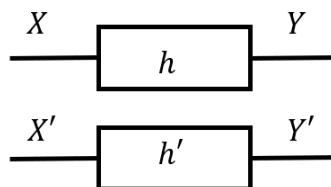


FIGURE 31 – Filtrage de deux processus (formule des interférences)

Théorème 5.2 (Formule des interférences). *Soit :*

- X_n et X'_n deux processus aléatoires stationnaires au sens large,
- h et h' les réponses impulsionnelles (ℓ^1) de deux filtres stables
- les processus filtrés Y_n et Y'_n sont définis par

$$Y_n = (h * X)_n = \sum_{k \in \mathbb{Z}} h[k] X_{n-k}$$

$$Y'_n = (h' * X')_n = \sum_{k \in \mathbb{Z}} h'[k] X'_{n-k}$$

Alors :

- les processus filtrés Y_n et Y'_n sont stationnaires,
- l'intercorrélation entre Y_n et Y'_n est donnée par :

$$\gamma_{YY'} = h * \bar{h}' * \gamma_{XX'}$$

$$\text{avec } \bar{h}'[k] = h'^*[-k].$$

Démonstration.

$$\begin{aligned} \gamma_{YY'}[k] &= \mathbb{E}(Y_{n+k} Y_n'^*) \\ &= \mathbb{E} \left(\sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} h[i] h'^*[j] X_{n+k-i} X_{n-j}'^* \right) \\ &= \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} h[i] h'^*[j] \mathbb{E}(X_{n+k-i} X_{n-j}'^*) \\ &= \sum_{j \in \mathbb{Z}} h'^*[j] \sum_{i \in \mathbb{Z}} h[i] \gamma_{XX'}[j+k-i] \\ &= \sum_{j \in \mathbb{Z}} h'^*[j] g[j+k] \quad \text{avec } g = h * \gamma \\ &= \sum_{j \in \mathbb{Z}} \bar{h}'[-j] g[j+k] \\ &= \sum_{j \in \mathbb{Z}} \bar{h}'[j] g[k-j] \\ &= \bar{h}' * g[k] \\ &= \bar{h}' * h * \gamma_{XX'}[k] \end{aligned}$$

□

Ce théorème (appelé **formule des interférences**) est très important car il permet d'exprimer les moments d'ordre 2 des processus filtrés en fonction des réponses impulsionnelles des filtres et des moments des processus d'entrée.

Ce théorème se décline sur de nombreuses applications pratiques dont quelques cas particuliers sont détaillés ci-dessous.

*Filtrage d'un processus***Corollary 5.1.** *Soit*

- X_n un processus stationnaire au sens large,
 - h la réponse impulsionnelle d'un filtre stable,
- le processus filtré $Y_n = h * X_n$ possède les propriétés suivantes :
- l'autocorrélation de Y_n est : $\gamma_Y = h * \bar{h} * \gamma_X$,
 - la DSP de Y_n est : $\Gamma_Y(\nu) = |H(\nu)|^2 \Gamma_X(\nu)$.

Démonstration. La démonstration est triviale en utilisant la formule des interférences avec

- $X' = X$
- $h' = h$

NB. On utilise aussi la propriété de la transformé de Fourier : $\mathcal{F}\bar{h}(\nu) = H^*(\nu)$ $\square$

Cette formule est très importante puisqu'elle permet d'expliciter la DSP (et donc l'autocorrélation) du processus filtré en fonction de celle du processus d'entrée.

Identification d'un filtre

Il est courant de chercher à identifier la réponse impulsionnelle d'un système linéaire invariant (un filtre) inconnu afin de pouvoir le caractériser. Une méthode consiste à introduire un bruit blanc en entrée du filtre. Il est alors possible de retrouver la réponse impulsionnelle du filtre à partir de l'intercorrélation entre l'entrée et la sortie.

Corollary 5.2. *Soit*

- W_n un bruit blanc,
 - h la réponse impulsionnelle d'un filtre stable.
- L'intercorrélation entre le bruit en entrée W_n et le processus filtré $Y_n = h * W_n$ est égale à la réponse impulsionnelle du filtre :

$$\gamma_{YW}[k] = h[k].$$

Démonstration. La démonstration est triviale en utilisant la formule des interférences avec

- $X_n = X'_n = W_n$
- $h'[n] = \delta[n]$

On a alors $Y_n = (h * W)_n$ et $Y'_n = W_n$, d'où d'une part $\gamma_{YY'} = \gamma_{YW}$. D'autre part d'après la formule des interférences :

$$\gamma_{YY'} = h * \bar{h}' * \gamma_{XX'} = h * \delta * \gamma_W = h$$

$\square$

5.2.6 Processus autorégressif (AR)

Les processus aléatoires autorégressifs sont des modèles très couramment utilisés car ils permettent de modéliser de nombreux phénomènes. On peut par exemple citer la modélisation des phonèmes de parole, l'étude de séries chronologiques aussi variées que des mesures de pollution ou de valeurs financières.

Une deuxième raison justifiant l'utilisation des modèles AR est que l'estimation de leurs paramètres est relativement simple comme nous le verrons dans le chapitre 6 portant sur l'estimation.

Définition

Définition 5.17. (*Processus autorégressif*) Un processus autorégressif (AR) d'ordre p est un processus filtré, plus précisément il s'agit d'un bruit blanc filtré par un filtre stable dont la fonction de transfert est donnée par :

$$H(v) = \frac{1}{1 - \sum_{k=1}^p a_k e^{-i2\pi kv}}$$

où l'on suppose les a_k réels.

Relation temporelle : équation aux différences finies

D'après les résultats de la section 3.5.1, on peut définir un processus autorégressif dans le domaine temporel à l'aide de son équation aux différences de la manière suivante.

Théorème 5.3. Un processus autorégressif Y_n d'ordre p s'écrit sous la forme :

$$Y_n = \sum_{k=1}^p a_k Y_{n-k} + W_n \quad (135)$$

où W_n est un bruit blanc.

Remarque. Pour exprimer la relation entrée-sortie du filtre, on peut utiliser la réponse impulsionnelle du filtre, mais dans ce cas, elle est de durée infinie. Il est donc plus pertinent d'utiliser l'équation aux différences finies qui ne fait intervenir qu'un nombre fini de termes Y_{n-k} (et donc de paramètres) pour décrire le filtre.

Paramètres d'un modèle AR

Un processus AR d'ordre p est donc entièrement décrit par les paramètres suivants :

- la variance σ^2 du bruit blanc générateur W_n ,
- les p coefficients $\{a_k\}_{1 \leq k \leq p}$ de l'équation aux différences finies.

DSP d'un modèle AR

Un processus AR étant un bruit blanc filtré, sa DSP s'exprime aisément à l'aide des relations vues à la section précédente :

$$\Gamma_Y(\nu) = \sigma^2 \left| \frac{1}{1 - \sum_{k=1}^p a_k e^{-i2\pi k\nu}} \right|^2$$

Autocorrélation d'un modèle AR

Propriété. L'autocorrélation d'un processus AR s'écrit :

$$\forall k \geq 0, \quad \gamma_Y[k] = \sum_{i=1}^p a_i \gamma_Y[k-i] + \sigma^2 \delta[k] \quad (136)$$

avec $\gamma_Y[k] = \gamma_Y[-k]$ pour $k < 0$.

Démonstration.

$$\begin{aligned} \gamma_Y[k] &= \mathbb{E}(Y_n Y_{n-k}) \\ &= \mathbb{E} \left(\left(\sum_{i=1}^p a_i Y_{n-i} + W_n \right) Y_{n-k} \right) \\ &= \sum_{i=1}^p a_i \mathbb{E}(Y_{n-i} Y_{n-k}) + \mathbb{E}(W_n Y_{n-k}) \\ &= \sum_{i=1}^p a_i \gamma_Y[k-i] + \sigma^2 \delta[k] \end{aligned}$$

La relation $\mathbb{E}(W_n Y_{n-k}) = \sigma^2 \delta[k]$ se déduit de l'équation (135) et du fait que W_n est un bruit blanc ($\mathbb{E}(W_i W_j) = \sigma^2 \delta[i-j]$). En effet, en remarquant que d'après (135), Y_{n-k} ne dépend que des W_i tel que $i \leq n-k$, on obtient :

- pour $k > 0$, $\mathbb{E}(W_n Y_{n-k}) = 0$;
- pour $k = 0$,

$$\begin{aligned} \mathbb{E}(W_n Y_{n-k}) &= \mathbb{E}(W_n Y_n) \\ &= \mathbb{E} \left(W_n \left(\sum_{i=1}^p a_i Y_{n-i} + W_n \right) \right) \\ &= \sum_{i=1}^p a_i \mathbb{E}(W_n Y_{n-i}) + \mathbb{E}(W_n^2) \\ &= 0 + \sigma^2 \end{aligned}$$

□

Ainsi, si on prend les équations précédentes (136) pour $0 \leq k \leq p$, on obtient un système linéaire de $p+1$ équations qui permet d'estimer les $p+1$ paramètres du modèle, à savoir $\{a_1, a_2, \dots, a_p, \sigma^2\}$ à partir de l'autocorrélation. On reviendra sur ce sujet dans le chapitre suivant. Les équations qui forment ce système linéaire sont appelées équations de **Yule-Walker**.

À retenir du chapitre

- définition d'un processus aléatoire discret
- définition d'un processus gaussien
- définition des deux premiers moments : moyenne et autocorrélation
- définition de la stationnarité
- définition de la densité spectrale de puissance (DSP)
- méthode de calcul de la DSP
- définition d'un bruit blanc
- filtrage de processus : connaître et savoir utiliser la formule des interférences et son expression fréquentielle
- définition d'un processus autorégressif : expression temporelle, fréquentielle

ESTIMATION

Le chapitre précédent a permis de définir un certain nombre d'outils mathématiques permettant de modéliser et caractériser les signaux à l'aide de processus aléatoires.

En pratique, on dispose de signaux qui correspondent à des réalisations de ces processus et on souhaite estimer les quantités qui caractérisent les processus. Typiquement, à partir de mesures, on souhaite estimer la moyenne, l'autocorrélation, la DSP ou les paramètres du modèle. L'objet de chapitre est donc principalement de définir les estimateurs de ces grandeurs. Ce chapitre fera donc appel à des concepts communs avec le cours de statistiques.

A titre d'exemple, considérons par exemple une application en traitement de la parole, pour laquelle on souhaite identifier la nature des sons élémentaires (appelés phonèmes) afin de réaliser de la reconnaissance automatique ou de la compression. Certains phonèmes comme le /s/ ou le /f/ peuvent être modélisés par un processus aléatoire de type autorégressif (AR). En pratique, on dispose de l'enregistrement d'un signal audio qui correspond à une réalisation du processus AR. L'objectif est d'estimer les paramètres du modèle AR à partir de cette unique réalisation afin d'identifier le type de phonème.

Précisons, du point de vue des notations, les différents objets que nous serons amenés à manipuler dans ce chapitre :

- un processus aléatoire est noté d'une lettre majuscule avec en indice la variable d'évolution : X_n
- le signal correspondant à une réalisation de ce processus est noté avec une lettre minuscule : $x[n]$
- l'estimateur (ou la statistique) est noté avec un chapeau circonflexe : $\hat{\mu}$

NB. il s'agit d'une variable (ou d'un processus) aléatoire

Dans l'ensemble de ce chapitre on ne considérera que les processus à temps discret, stationnaires au sens large et à valeurs réelles.

Pour aller plus loin...

Dans un cadre plus général, on peut considérer un ensemble de N réalisations, ce qui amène à définir le concept d'échantillons aléatoire et d'échantillons d'observations^a.

Soit un processus aléatoire $X_n(\omega)$, on définit l'échantillon aléatoire de taille N comme le processus aléatoire de dimension N défini par :

$$(\omega_1, \omega_2, \dots, \omega_N) \mapsto (X_n(\omega_1), X_n(\omega_2), \dots, X_n(\omega_N))$$

L'échantillon d'observation $(X_n(\omega_1), X_n(\omega_2), \dots, X_n(\omega_N))$ est alors noté $(x^{(1)}[n], x^{(2)}[n], \dots, x^{(N)}[n])$

Dans la plupart des applications en traitement du signal, on ne dispose que d'une réalisation. Dans ce cas l'échantillon aléatoire est identique au processus X_n et l'échantillon d'observation correspond à la réalisation notée $x[n]$.

a. il s'agit d'une généralisation du concept vu en cours de statistiques pour les variables aléatoires

6.1 ESTIMATEUR DE LA MOYENNE ET DE L'AUTOCORRÉLATION

L'une des difficultés de l'estimation en traitement du signal est liée au fait qu'on ne dispose, en général, que d'une seule réalisation du processus aléatoire. En d'autres termes, il n'est possible de réaliser qu'une seule mesure du signal qui correspond donc à une observation du modèle. Si on fait le parallèle avec une variable aléatoire, cela reviendrait à essayer d'estimer sa moyenne ou sa variance à partir d'une seule réalisation !

C'est pourquoi on ne traitera que le cas des signaux ergodiques pour lesquels on peut remplacer l'estimateur empirique qui réalise la moyenne statistique sur les réalisations par un estimateur utilisant la moyenne temporelle des échantillons du processus.

Il est évident qu'une telle approche n'a de sens que pour les signaux stationnaires pour lesquels les grandeurs à estimer ne dépendent pas du temps.

6.1.1 Estimateur de la moyenne

Définition 6.1 (Estimateur de la moyenne). L'estimateur temporel de la moyenne à partir de N instants est défini par :

$$\hat{\mu}_X^N = \frac{1}{N} \sum_{n=0}^{N-1} X_n.$$

Propriété. Cet estimateur n'est pas biaisé.

biaisé : espérance différente de la moyenne des résultats à estimer

On montre en effet aisément que $\mathbb{E}(\hat{\mu}_X^N) = \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}(X_n) = \mu_X$.

Définition 6.2 (Ergodicité d'ordre 1). Le processus stationnaire X_n est dit ergodique d'ordre 1 si l'estimateur temporel de la moyenne converge dans L^2 vers le moment d'ordre 1 (i.e., la moyenne) du processus :

$$\hat{\mu}_X^N = \frac{1}{N} \sum_{n=0}^{N-1} X_n \xrightarrow[N \rightarrow \infty]{L^2} \mu_X$$

ou de manière équivalente car l'estimateur n'est pas biaisé :

$$\text{Var}(\hat{\mu}_X^N) \xrightarrow[N \rightarrow \infty]{} 0$$

En pratique, on calcule une réalisation de la variable aléatoire $\hat{\mu}_X^N$ à partir d'une unique réalisation $x[n] = X_n(\omega)$ d'un processus stationnaire et ergodique X_n . Cette réalisation s'écrit : $\hat{\mu}_X^N(\omega) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$. L'ergodicité du processus permet donc de remplacer l'opérateur de moyenne sur plusieurs réalisations par un opérateur de moyenne temporelle sur les valeurs d'une unique réalisation.

Exemple 6.1. Soit deux exemples de processus stationnaires, le premier est ergodique tandis que le deuxième ne l'est pas :

— un bruit blanc est ergodique à l'ordre 1. En effet

$$\begin{aligned} \text{Var}(\hat{\mu}_X^N) &= \frac{1}{N^2} \sum_{n'=0}^{N-1} \sum_{n=0}^{N-1} \mathbb{E}(X_n X_{n'}) \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} \mathbb{E}(X_n^2) \quad \text{car} \quad \mathbb{E}(X_n X_{n'}) = 0, \forall n \neq n' \\ &= \frac{\gamma_X[0]}{N} \end{aligned}$$

— le processus aléatoire à valeur aléatoire constante défini par $X_n = A$ avec $A \sim \mathcal{N}(0, \sigma)$ n'est pas ergodique. En effet :

$$\begin{aligned} \text{Var}(\hat{\mu}_X^N) &= \frac{1}{N^2} \sum_{n'=0}^{N-1} \sum_{n=0}^{N-1} \mathbb{E}(X_n X_{n'}) \\ &= \frac{1}{N^2} \sum_{n'=0}^{N-1} \sum_{n=0}^{N-1} \mathbb{E}(A^2) \\ &= \sigma^2 \end{aligned}$$

6.1.2 Estimateur de l'autocorrélation

Définition 6.3 (Estimateur non biaisé de l'autocorrélation). *L'estimateur temporel non biaisé de l'autocorrélation est défini par :*

$$\hat{\gamma}_X^{N,nb}[k] = \frac{1}{N-k} \sum_{\ell=0}^{N-k-1} X_{\ell+k} X_{\ell}^* \quad \text{pour } 0 \leq k < N.$$

avec de plus $\hat{\gamma}_X^{N,nb}[k] = \hat{\gamma}_X^{N,nb*}[-k]$ pour $-N < k < 0$.

Propriété. Cet estimateur est non biaisé pour les valeurs de k pour lesquelles il est défini.

On montre en effet facilement que $\mathbb{E}(\hat{\gamma}_X^{N,nb}[k]) = \gamma_X[k]$, $\forall |k| < N$.

On peut remarquer que cet estimateur fournit une estimation de l'autocorrélation uniquement pour les $2N - 1$ valeurs telles que $|k| < N$.

Cet estimateur est rarement utilisé car sa variance est très élevée pour les valeurs de k proches de N . En effet, dans le cas extrême $k = N - 1$, la moyenne ne porte plus que sur une unique valeur (ce qui implique une variance élevée). Par contre, pour $k = 0$, la moyenne porte sur N valeurs (la variance peut donc être supposée raisonnablement faible si N est suffisamment grand - cf. rappels de statistiques dans la section 1.5.1 du chapitre 1).

Définition 6.4 (Estimateur biaisé de l'autocorrélation). *L'estimateur temporel biaisé de l'autocorrélation est défini par :*

$$\hat{\gamma}_X^N[k] = \frac{1}{N} \sum_{\ell=0}^{N-k-1} X_{\ell} X_{\ell+k} \quad \text{pour } 0 \leq k < N.$$

avec de plus $\hat{\gamma}_X^N[k] = \hat{\gamma}_X^{N*}[-k]$ pour $-N < k < 0$.

La différence avec l'estimateur non biaisé est liée au facteur de normalisation en $\frac{1}{N}$ au lieu de $\frac{1}{N-|k|}$.

Cet estimateur est biaisé puisque :

$$\mathbb{E}(\hat{\gamma}_X^N[k]) = \mathbb{E}\left(\frac{N-|k|}{N} \hat{\gamma}_X^{N,nb}[k]\right) = \frac{N-|k|}{N} \gamma_X[k]$$

On peut remarquer que le biais est faible pour $k \approx 0$ et devient plus important pour k proche de N .

Cet estimateur possède une variance plus faible que celle de l'estimateur non biaisé puisque :

$$\text{Var}(\hat{\gamma}_X^N[k]) = \left(\frac{N-|k|}{N}\right)^2 \text{Var}(\gamma_X^{nb}[k])$$

Les termes pour k proche de N qui sont estimés sur peu d'échantillons sont forcés à des valeurs proches de zéros (par la normalisation en $1/N$) afin d'en réduire la variance.

On retiendra que cet estimateur est privilégié par rapport à l'estimateur non biaisé car il possède, en général, une erreur quadratique¹ plus faible que l'estimateur non biaisé².

Soit $X^N = (X_0, X_1, \dots, X_{N-1})$ le processus aléatoire tronqué sur N points. Il s'agit d'un vecteur aléatoire de dimension N qui contient l'ensemble des instants du processus utilisés par les estimateurs temporels définis ci-dessus.

Pour aller plus loin...

Propriété. On peut écrire l'estimateur biaisé de l'autocorrélation comme un produit de convolution :

$$\hat{\gamma}_X^N = \frac{1}{N} X^N * \bar{X}^N$$

avec $\forall k \in \mathbb{Z}, \quad \bar{X}_{-k}^N = X_k^{N*}$

Démonstration.

$$\begin{aligned} \gamma_X^N[k] &= \frac{1}{N} \sum_{\ell=0}^{N-k-1} X_{\ell+k} X_{\ell}^* \\ &= \frac{1}{N} \sum_{\ell=0}^{N-k-1} X_{\ell+k} \bar{X}_{-\ell} \\ &= \frac{1}{N} \sum_{\ell'=0}^{N-k-1} X_{\ell'} \bar{X}_{k-\ell'} \quad \text{avec } \ell' = \ell + k \\ &= \frac{1}{N} \sum_{\ell'=0}^{N-k-1} X_{\ell'}^N \bar{X}_{k-\ell'}^N \end{aligned}$$

d'où

$$\gamma_X^N = \frac{1}{N} X^{N*} \bar{X}^N$$

□

1. L'erreur quadratique moyenne (EQM) est couramment utilisée en statistique pour comparer les estimateurs en prenant en compte simultanément le biais et la variance, elle est définie par :

$$EQM(\hat{\gamma}) = \mathbb{E}(\hat{\gamma})^2 + \text{Var}(\hat{\gamma})$$

2. La plupart des processus modélisant des phénomènes physiques ont une autocorrélation qui décroît rapidement (intuitivement, cela correspond au fait que deux échantillons pris à des instants proches sont en général plus corrélés que deux échantillons très éloignés).

Dans ce cas, l'EQM de l'estimateur biaisé est plus faible que celle de l'estimateur non biaisé car on peut négliger le terme biais au biais. En effet, pour les petites valeurs de k le biais est faible car l'estimateur biaisé est proche de l'estimateur non biaisé $\left(\frac{N-|k|}{N} \approx \frac{1}{N}\right)$, et pour les grandes valeurs de k , le biais $\left(\frac{N-|k|}{N} \gamma_X[k]\right)$ reste faible car $\gamma_X[k]$ est petit.

Définition 6.5 (Ergodicité d'ordre 2). Le processus stationnaire X_n est dit ergodique d'ordre 2 si l'estimateur non biaisé de son moment d'ordre 2 (l'autocorrélation) converge en moyenne quadratique vers l'autocorrélation du processus :

$$\hat{\gamma}_X^{N,nb}[k] \xrightarrow[N \rightarrow \infty]{L^2} \gamma_X[k]$$

ou de manière équivalente :

$$\text{Var}(\gamma_X^{N,nb}[k]) \xrightarrow[N \rightarrow \infty]{} 0$$

L'ergodicité d'ordre 2 d'un processus est donc une propriété indispensable pour pouvoir estimer son autocorrélation à l'aide des estimateurs temporels.

6.2 ESTIMATION DE LA DSP : PÉRIODOGRAMME

L'estimation de la densité spectrale de puissance d'un processus à partir de données se résument à une unique réalisation du processus est un problème classique en traitement du signal afin de caractériser un phénomène dans le domaine fréquentiel.

Exemple 6.2. Ainsi par exemple en climatologie, une question consiste à retrouver la périodicité des variations climatiques liées aux oscillations de l'axe de la terre. L'objectif étant de distinguer plus finement les variations climatiques liées à l'activité humaine par rapport aux autres facteurs. L'une des mesures les plus utilisées est la concentration de deutérium dans les carottes glaciaires datant de plusieurs milliers d'années. On modélise le signal correspondant à la concentration en deutérium en fonction du temps par un processus aléatoire, bien évidemment, on ne dispose que d'une seule réalisation de ce processus. En estimant la DSP du processus à partir de cette unique réalisation, on peut alors retrouver les principales périodicités qui correspondent aux pics de la DSP.

6.2.1 Définition

L'estimateur classique de la densité spectrale de puissance est appelé périodogramme, il est défini de la manière suivante :

Définition 6.6 (Périodogramme). Le périodogramme sur N points est un estimateur de la DSP défini comme la transformée de Fourier de l'estimateur biaisé de l'autocorrélation :

$$\hat{f}_X(\nu) = \sum_{k=-N+1}^{N-1} \hat{\gamma}_X^N[k] e^{-i2\pi k\nu}$$

ou de manière équivalente :

$$\hat{\Gamma}_X(v) = \frac{1}{N} \left| \sum_{k=-N+1}^{N-1} X_k^N e^{-i2\pi kv} \right|^2$$

où $X^N = (X_0, X_1, \dots, X_{N-1})$ est une troncature du processus aléatoire X_n sur N points.

Démonstration. L'équivalence entre les deux définitions s'obtient en utilisant $\hat{\gamma}_X^N = \frac{1}{N} X^N * \bar{X}^N$:

$$\begin{aligned} \hat{\Gamma}_X(v) &= \mathcal{F}\gamma_X^N(v) \\ &= \frac{1}{N} \mathcal{F}\{X^N * \bar{X}^N\}(v) \\ &= \frac{1}{N} \mathcal{F}X^N(v) \mathcal{F}\bar{X}^N(v) \\ &= \frac{1}{N} \left(\mathcal{F}X^N(v) \right) \left(\mathcal{F}X^N(v) \right)^* \\ &= \frac{1}{N} \left| \mathcal{F}X^N(v) \right|^2 \end{aligned}$$

□

Rappelons la définition de la densité spectrale de puissance donnée à la section 5.2.3 du chapitre 5,

$$\forall v \in \mathbb{R}, \quad \Gamma_X(v) = \lim_{N \rightarrow \infty} \mathbb{E} \left(\frac{1}{2N+1} \left| \sum_{k=-N}^N X_k e^{-i2\pi kv} \right|^2 \right)$$

Si on compare cette expression à celle du périodogramme, on peut remarquer que l'espérance et le passage à la limite ont été supprimés pour pouvoir définir l'estimateur. On retiendra que ces deux simplifications ont des conséquences importantes :

- **Résolution fréquentielle limitée (biais)** : Supprimer le passage à la limite revient à réaliser une troncature sur N points : cela a des conséquences sur la résolution spectrale et les lobes secondaires de l'estimateur, cf. section 3.6.2 du chapitre 3.
- **Variance élevée** : supprimer l'espérance revient à réaliser une moyenne empirique sur une unique réalisation : il en résulte une variance très élevée de l'estimateur qui le rend inutilisable dans la plupart des cas. Ce qui justifie les variantes du périodogramme (présentées dans la section suivante) et dont l'objectif est de réduire la variance.

D'un point de vue pratique, pour obtenir une estimation de la DSP à partir d'une réalisation $x[n]$ du processus tronquée sur N points, on réalise la TFD des $x[n]$ (en général en effectuant un bourrage de zéros pour affiner le tracé comme introduit à la section 3.6 du chapitre 3).

Exemple 6.3. Afin d'illustrer la variance du périodogramme, on considère un bruit blanc filtré dont la DSP théorique est tracée en traits épais.

La figure 32 représente la DSP du processus ainsi que trois réalisations du périodogramme obtenues à partir de trois réalisations différentes du processus.

Sur la figure de gauche, la longueur du signal utilisé est de $N = 40$ échantillons tandis que sur la figure de droite $N = 400$.

On peut remarquer que :

- l'estimation de la DSP obtenue par le périodogramme est très différente de la DSP théorique, cela illustre la variance élevée de l'estimateur
- la dispersion des valeurs est très importante et ne diminue pas sur la figure de droite lorsqu'on utilise plus d'échantillons, cela illustre le fait que l'estimateur n'est pas consistant.

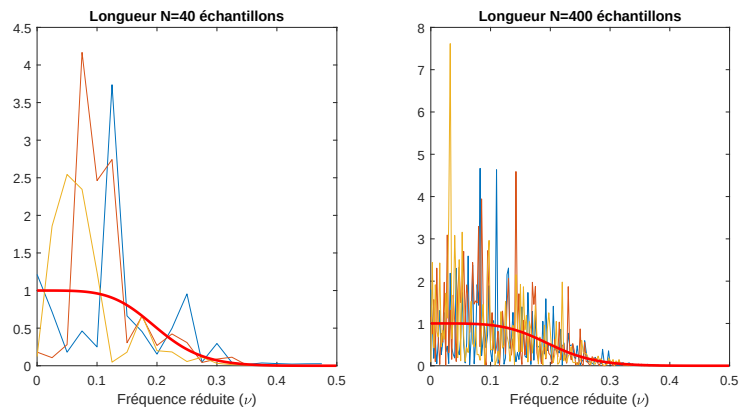


FIGURE 32 – Illustration de la variance du périodogramme : comparaison de la DSP théorique (trait gras) avec des réalisations du périodogramme pour $N = 40$ et $N = 400$ échantillons

Pour aller plus loin...

6.2.2 Propriétés statistiques du périodogramme

Biais et résolution

Propriété. Le périodogramme est **biaisé** :

$$\mathbb{E}(\Gamma_N(\nu)) = (\Gamma_X \otimes W_N)(\nu)$$

où $W(\nu)$ est la transformée de Fourier de la fenêtre triangle : $W_n(\nu) = \frac{1}{N} \frac{\sin^2(\pi N \nu)}{\sin^2(\pi \nu)}$

Démonstration.

$$\mathbb{E}(\hat{\gamma}_X^N[k]) = \frac{N - |k|}{N} \gamma_X[k]$$

$$\mathbb{E}(\Gamma_N(\nu)) = (\Gamma_X \otimes W_N)(\nu)$$

□

On peut retenir que :

- le biais est lié à la troncature sur N points,
- le biais limite la résolution spectrale à $1/N$,
- le biais est asymptotiquement nul (quand $N \rightarrow \infty$) : pour améliorer la résolution, il suffit d'augmenter la durée de l'observation (c'est à dire le nombre de points N)

Loi asymptotique de l'estimateur

Propriété. Si X_n est processus gaussien dont la DSP est $\Gamma_X(\nu)$

- Dans le cas asymptotique ($N \rightarrow \infty$), le périodogramme suit une loi exponentielle :

$$p_{\hat{\Gamma}(\nu)}(x) = \frac{1}{\Gamma(\nu)^2} e^{-\frac{x}{\Gamma(\nu)^2}}$$

- Moyenne asymptotique du périodogramme :

$$\mu_{\hat{\Gamma}(\nu)} = \Gamma(\nu)$$

- Variance asymptotique du périodogramme :

$$\text{Var}(\hat{\Gamma}(\nu)) = \Gamma(\nu)^2$$

On retiendra principalement de ce résultat que l'écart-type de l'estimateur est égal à la valeur que l'on cherche à estimer, ce qui limite fortement l'intérêt d'un tel estimateur !

6.2.3 Variantes du périodogramme

Comme nous l'avons vu précédemment, la variance du périodogramme est très élevée, les variantes du périodogramme proposent différentes méthodes pour la réduire par moyennage³. On présente succinctement le périodogramme de Bartlett.

Comme illustré sur la figure 33, le principe du périodogramme de Bartlett consiste à diviser le signal en K fenêtres, et à moyenner les périodogrammes obtenus sur chacune de ces fenêtres.

Définition 6.7. Soit une troncature X^N de N échantillons du processus X_n , avec $N = KL$, le périodogramme de Bartlett sur K fenêtres de longueur L est défini par :

$$\hat{\Gamma}^{\text{Bart}}(\nu) = \frac{1}{K} \sum_{k=0}^{K-1} \Gamma^{(k)}(\nu)$$

$$\text{avec } \Gamma^{(k)}(\nu) = \frac{1}{L} \left| \sum_{\ell=0}^{L-1} X_{\ell+kL}^N e^{-i2\pi\nu\ell} \right|^2.$$

3. En complément du moyennage, certaines variantes permettent aussi d'appliquer une fenêtre de troncature de forme particulière afin de limiter les lobes secondaires.

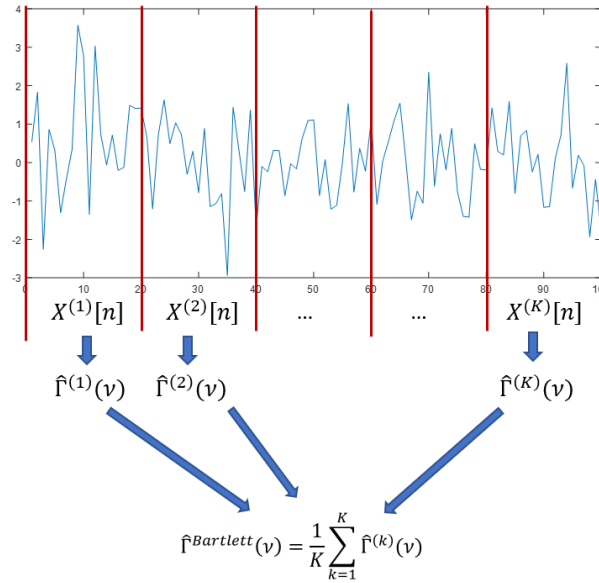


FIGURE 33 – Principe du périodogramme de Bartlett.

On pourra retenir les deux propriétés suivantes du périodogramme de Bartlett comparé au périodogramme standard :

- **Réduction de la variance** : moyenner K périodogrammes permet de réduire la variance de l'estimateur de la DSP⁴.
- **Augmentation du biais** : en réduisant la longueur de la fenêtre d'un facteur K , l'effet de troncature est amplifié, ce qui se traduit par une dégradation de la résolution spectrale (elle est multipliée par K) comme vu à la section 3.6 du chapitre 3.

Le choix du nombre de fenêtres K est donc un compromis entre le biais et la variance.

Exemple 6.4. *Considérons un processus dont la DSP est tracée en gras sur la figure 34. On dispose de l'observation de ce processus sur une durée de 10 000 échantillons.*

- *La figure du haut présente trois réalisations du périodogramme standard dont on peut constater que la variance est trop importante pour estimer correctement la DSP.*
- *La figure du centre présente les réalisations du périodogramme de Bartlett en moyennant sur 50 fenêtres de longueur 200. Bien que l'estimation demeure bruitée, elle se rapproche de la valeur théorique. Par ailleurs, le biais reste très faible (l'espérance tracée en pointillé est quasiment confondue avec la valeur théorique). Cet estimateur apparaît comme un bon compromis.*
- *la figure du bas présente les réalisations du périodogramme de Bartlett en moyennant sur 500 fenêtres de longueur 20. La variance est très faible, mais en contrepartie, le biais n'est plus négligeable.*

4. Dans le cas d'un bruit blanc les échantillons des fenêtres sont décorrélés, on peut donc dire que la variance est réduite d'un facteur K .

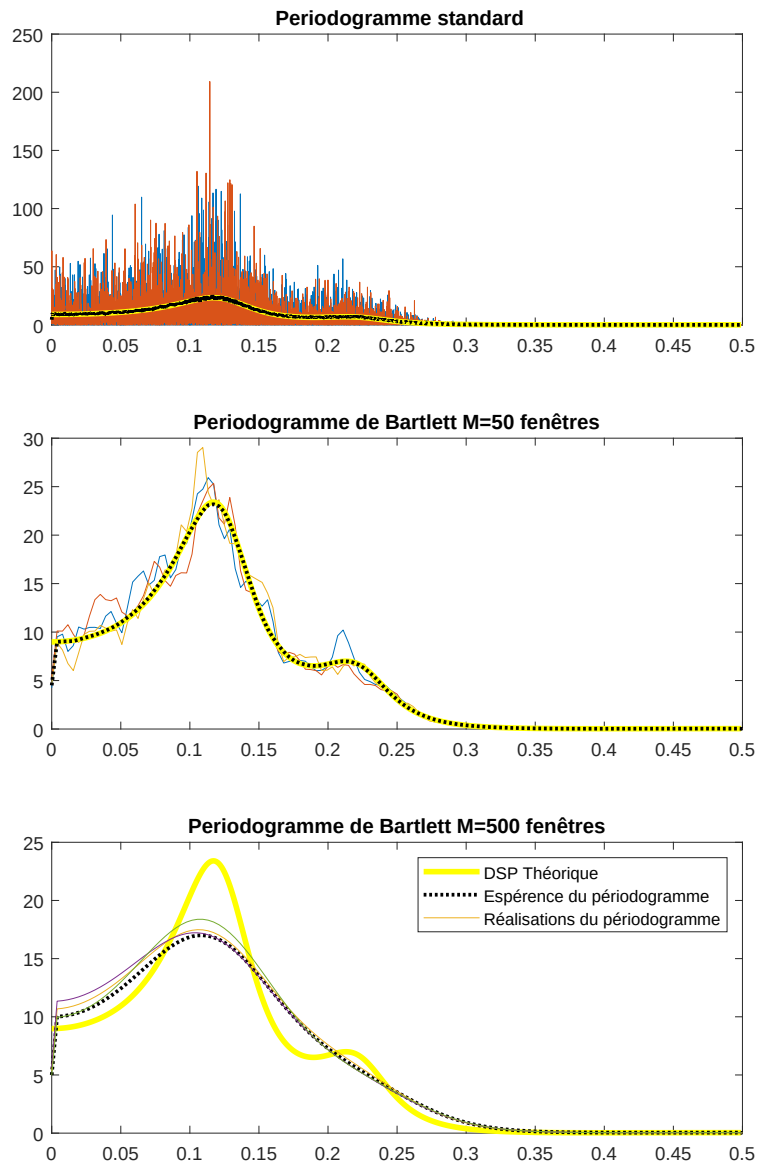


FIGURE 34 – Périodogramme de Bartlett : illustration de la réduction de la variance et de l'augmentation du biais

6.3 ESTIMATION PARAMÉTRIQUE

Cette partie porte sur l'estimation paramétrique et diverses applications qui en découlent. On parle d'estimation paramétrique lorsqu'au lieu d'estimer directement l'objet, on estime les paramètres du modèle qui représente l'objet.

Prenons l'exemple de l'estimation de la DSP détaillé dans la section suivante. Dans la section précédente, nous avons étudié le périodogramme qui est un estimateur non paramétrique de la DSP avec une variance très élevée. Un estimateur paramétrique consiste par exemple à faire l'hypothèse que le processus dont on cherche à estimer la DSP est un processus AR de paramètres $\{a_1, \dots, a_p\}$ et à estimer les paramètres du processus. La DSP s'exprime alors analytiquement à partir des paramètres estimés $\hat{a}_k$.

Pour résumer, l'estimateur paramétrique de la DSP :

- nécessite une hypothèse sur le modèle du processus,
- fournit une estimation avec une variance plus faible.

Plusieurs applications illustrent l'estimation paramétrique, par exemple en traitement de la parole pour l'utilisation de modèles AR, ou en analyse vibratoire, pour l'estimation des paramètres d'une sinusoïde bruitée.

6.3.1 Estimation de la DSP : application à un modèle AR

Le principe de l'estimation paramétrique de la DSP est le suivant (illustré ici dans le cas d'un processus autorégressif défini à la section 5.2.6 du chapitre 5) :

- Étape 1 - **Choix d'un modèle paramétrique** : les données sont supposées générées par un processus aléatoire, décrit par un modèle avec peu de paramètres.

exemple : un modèle AR : $X_n = \sum_{k=1}^p a_k X_{n-k} + W_n$ avec W_n un bruit blanc de variance σ_W .

- Étape 2 - **Estimation des paramètres** à partir des données (une réalisation $x[k]$ du processus),

exemple : estimation de $\hat{a}_p$ et $\hat{\sigma}_W$

- Étape 3 - **Calcul de la DSP** à partir du modèle estimé

exemple : $\Gamma_X(\nu) = \hat{\sigma}_W^2 \left| \frac{1}{1 - \sum_{k=1}^p \hat{a}_k e^{-2\pi i k \nu}} \right|^2$

L'approche la plus usuelle pour l'estimation des paramètres du modèle (étape 2) consiste à utiliser la méthode des moments qui permet d'estimer les paramètres du modèle à partir de l'estimation des moments du processus. Elle se divise en trois sous-étapes :

- Étape 2a : établir la relation entre les moments et les paramètres du modèle
- Étape 2b : estimer les moments du processus,
- Étape 2c : résoudre la relation liant les moments aux paramètres.

Exemple : modèle AR

Etape 1 : supposons que les données dont on souhaite estimer la DSP sont générées par un processus AR d'ordre p tel que défini à la section 5.2.6 du chapitre 5 :

$$X_n = \sum_{k=1}^p a_k X_{n-k} + W_n$$

où W_n est processus blanc de puissance σ^2 .

Les paramètres qui définissent ce processus (et que l'on souhaite estimer) sont donc au nombre de $p + 1$: les p coefficients a_k et σ^2 .

Etape 2a : D'après les résultats de la section 5.2.6, les équations de Yule-Walker permettent de relier linéairement les paramètres du modèle à l'autocorrélation du processus AR :

$$\forall k \geq 0, \quad \gamma_X[k] = \sum_{i=1}^p a_i \gamma_X[k-i] + \sigma^2 \delta[k]$$

ou encore sous forme matricielle pour $k = 1, \dots, p$:

$$\begin{pmatrix} \gamma_X[0] & \gamma_X[1] & \cdots & \gamma_X[p-1] \\ \gamma_X[1] & \gamma_X[0] & \cdots & \gamma_X[p-2] \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_X[p-1] & \gamma_X[p-2] & \cdots & \gamma_X[0] \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \gamma_X[1] \\ \gamma_X[2] \\ \vdots \\ \gamma_X[p] \end{pmatrix}$$

et pour $k = 0$:

$$\sigma^2 = - \sum_{i=1}^p a_i \gamma_X[i] + \gamma_X[0]$$

Etape 2b : l'autocorrélation est estimée à l'aide de l'estimateur biaisé $\hat{\gamma}[k]$ défini à la section 6.1.2

Etape 2c : la résolution du système linéaire (après avoir remplacé $\gamma[k]$ par son estimation $\hat{\gamma}[k]$) fournit une estimation des paramètres du processus.

La figure 35 illustre la réduction de la variance obtenue grâce à l'estimation paramétrique : la DSP du processus est tracée en trait épais, le graphique du haut représente des réalisations de l'estimateur paramétrique tandis que celui du bas représente des réalisations du périodogramme standard.

Exemple 6.5. Application au traitement de la parole

Certains sons (phonèmes) de la parole sont bien représentés par un modèle AR, il s'agit des sons non voisés (c'est à dire ceux pour lesquels les cordes vocales ne vibrent pas). Les principaux sons non voisés sont :

- consonnes fricatives : /f/, /s/ ...
- voyelles chuchotées

Physiquement, le modèle AR peut être interprété de la manière suivante :

- les turbulences de l'air à l'entrée du conduit vocal correspondent un bruit blanc,

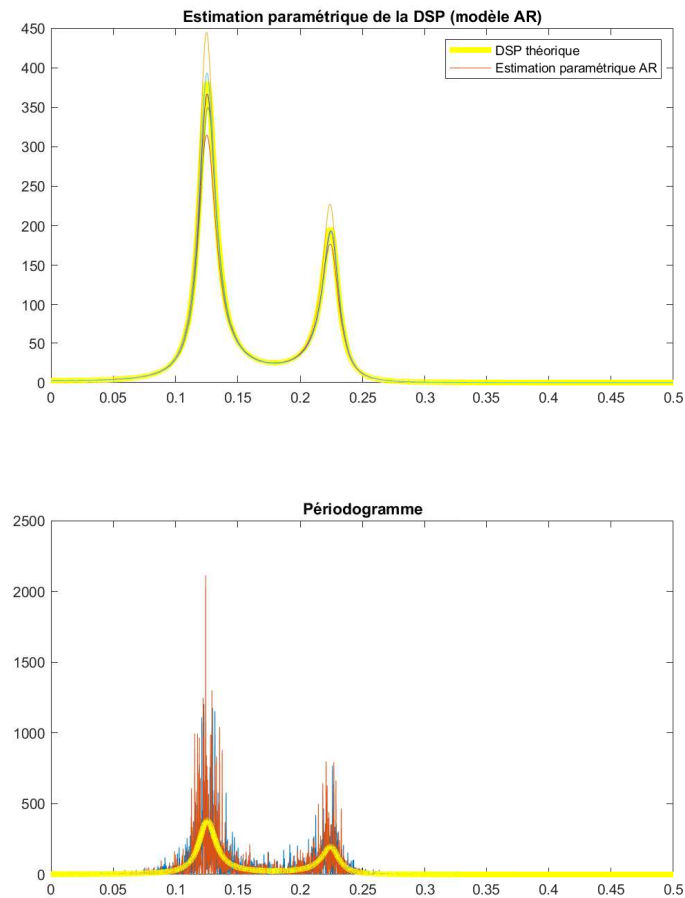


FIGURE 35 – Comparaison de l'estimation paramétrique de la DSP avec le périodogramme standard

— les résonances du conduit vocal correspondent aux pôles du modèle AR.

L'estimation des paramètres du modèle AR associé à un phonème permet de nombreuses applications comme l'identification de phonème (application à la reconnaissance vocale) ou de synthétiser de nouveaux échantillons (applications au débruitage, à la synthèse vocale...).

La figure 36 illustre le traitement complet.



Un exemple plus complet est disponible sur la page édunao du cours avec le code correspondant.

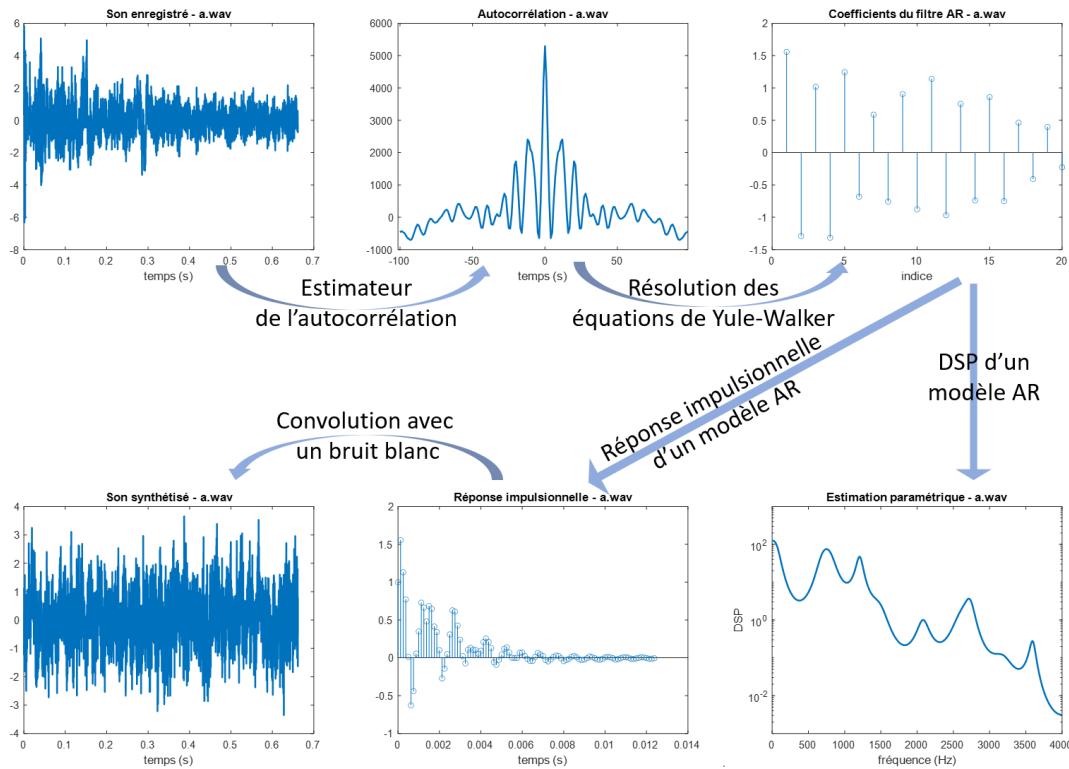


FIGURE 36 – Traitement de la parole et synthèse vocale

6.4 ESTIMATEUR LINÉAIRE OPTIMAL

Nous nous intéressons ici à un problème présentant de nombreuses applications dont certaines sont détaillées plus loin. Il s'agit d'estimer une variable aléatoire X à partir de P échantillons d'un processus aléatoire Y_n , c'est à dire à partir du vecteur aléatoire $\mathbf{Z}$ formé par P variables aléatoires du processus Y_n . Par exemple, $\mathbf{Z} = (Y_{P-1}, Y_{P-2}, \dots, Y_0)$.

Bien évidemment, pour que l'estimation soit possible il est indispensable que le processus Y_n soit corrélé avec X . On suppose connues l'autocorrélation de $\mathbf{Z}$ et l'intercorrrelation entre X et $\mathbf{Z}$.

Parmi les applications les plus courantes, on peut citer :

- la prédiction : prédire la valeur du processus X_n à partir des P valeurs précédentes de $Y_{n-k} = X_{n-k}$, $k = 1, \dots, P$,
- le débruitage : estimer le signal utile X_n à partir de l'observation bruitée $Y_n = X_n + W_n$,
- la déconvolution : estimer le signal utile X_n à partir de mesures entachées d'un flou $Y_n = (X * h)_n + W_n$

Dans cette section, on supposera les processus centrés et réels.

6.4.1 Définition et propriétés

Nous nous intéresserons ici uniquement aux estimateurs linéaires minimisant la variance de l'erreur d'estimation, ils sont définis de la manière suivante

Définition 6.8 (Estimateur linéaire optimal). *Un estimateur linéaire optimal vérifie les propriétés suivantes :*

— *l'estimateur est une combinaison linéaire des données :*

$$\hat{X} = \mathbf{h}_{\text{opt}}^t \mathbf{Z} = \sum_{p=0}^{P-1} h_p Z_p$$

où $\mathbf{h}_{\text{opt}}$ regroupe les coefficients de l'estimateur.

— *l'estimateur minimise la variance de l'erreur d'estimation :*

$$\mathbf{h}_{\text{opt}} = \underset{\mathbf{h}}{\operatorname{argmin}} \mathbb{E}((X - \hat{X})^2)$$

Théorème 6.1. *L'estimateur linéaire optimal de X à partir de $\mathbf{Z}$ est donné par $\hat{X} = \mathbf{h}_{\text{opt}}^t \mathbf{Z}$ avec :*

$$\mathbf{h}_{\text{opt}} = \boldsymbol{\Sigma}_{\mathbf{Z}}^{-1} \boldsymbol{\gamma}_{\mathbf{Z}X}.$$

Il s'écrit en fonction de :

- *l'intercorrélation entre X et $\mathbf{Z}$: $\boldsymbol{\gamma}_{\mathbf{Z}X} = \mathbb{E}(X\mathbf{Z})$,*
- *l'autocorrélation de $\mathbf{Z}$: $\boldsymbol{\Sigma}_{\mathbf{Z}} = \mathbb{E}(\mathbf{Z}\mathbf{Z}^t)$.*

Démonstration. La variance de l'erreur d'estimation que l'on cherche à minimiser s'écrit :

$$\begin{aligned} \mathbb{E}((X - \hat{X})^2) &= \mathbb{E}((X - \mathbf{h}^t \mathbf{Z})^2) \\ &= \mathbb{E}(\mathbf{h}^t \mathbf{Z} \mathbf{Z}^t \mathbf{h}) - 2\mathbb{E}(\mathbf{h}^t X \mathbf{Z}) + \mathbb{E}(X^2) \\ &= \mathbf{h}^t \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{h} - 2\mathbf{h}^t \boldsymbol{\gamma}_{\mathbf{Z}X} + \mathbb{E}(X^2) \end{aligned}$$

On obtient alors un problème d'optimisation quadratique :

$$\mathbf{h}_{\text{opt}} = \underset{\mathbf{h}}{\operatorname{argmin}} \left\{ J(\mathbf{h}) := \mathbf{h}^t \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{h} - 2\mathbf{h}^t \boldsymbol{\gamma}_{\mathbf{Z}X} + \mathbb{E}(X^2) \right\}.$$

On obtient l'optimum en annulant le gradient :

$$\nabla_{\mathbf{h}} J(\mathbf{h} = \mathbf{h}_{\text{opt}}) = \mathbf{0} \iff 2\boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{h}_{\text{opt}} - 2\boldsymbol{\gamma}_{\mathbf{Z}X} = \mathbf{0},$$

d'où $\mathbf{h}_{\text{opt}} = \boldsymbol{\Sigma}_{\mathbf{Z}}^{-1} \boldsymbol{\gamma}_{\mathbf{Z}X}$

□

Pour construire l'estimateur linéaire optimal au sens de la minimisation de l'erreur d'estimation, il suffit donc de connaître l'autocorrélation du processus $\boldsymbol{\Sigma}_{\mathbf{Z}} = \mathbb{E}(\mathbf{Z}\mathbf{Z}^t)$ ayant généré les données et l'intercorrélation $\boldsymbol{\gamma}_{\mathbf{Z}X} = \mathbb{E}(X\mathbf{Z})$ entre la variable à estimer et les données.

Théorème 6.2. *L'erreur d'estimation est décorrélée des observations :*

$$\mathbb{E}((X - \hat{X})Z_p) = 0 \quad \forall p \in \mathbb{Z}$$

Démonstration. L'estimateur minimise la variance de l'erreur $J(\mathbf{h}) = \mathbb{E}((X - \hat{X}(\mathbf{h}))^2)$ avec $\hat{X}(\mathbf{h}) = \mathbf{h}^t \mathbf{Z}$. On a donc :

$$\begin{aligned} \nabla_{\mathbf{h}} J(\mathbf{h} = \mathbf{h}_{\text{opt}}) &= 0 \\ \mathbb{E}(2(X - \hat{X}) \nabla_{\mathbf{h}} \hat{X}(\mathbf{h} = \mathbf{h}_{\text{opt}})) &= 0 \\ \mathbb{E}((X - \hat{X})\mathbf{Z}) &= 0. \end{aligned}$$

En effet, $\nabla_{\mathbf{h}} \hat{X}(\mathbf{h} = \mathbf{h}_{\text{opt}}) = \mathbf{Z}$ car $\hat{X}(\mathbf{h}) = \mathbf{h}^t \mathbf{Z}$. $\square$

6.4.2 Prédiction d'un processus aléatoire

Nous considérons ici une application de l'estimation linéaire optimale dédiée à la prédiction d'un processus aléatoire.

Le problème consiste à prédire la valeur du processus X_n à l'instant n à partir de P valeurs passées du processus X_n . En d'autres termes, il s'agit d'estimer X_n pour chaque n , à partir des P précédents échantillons de X_n que l'on notera $\mathbf{Z}^{(n)} = (X_{n-1}, X_{n-2}, \dots, X_{n-P})$. On suppose connue γ_X , l'autocorrélation du processus à débruiter.

Un tel problème présente des applications dans de nombreux domaines par exemple pour la prédiction de valeur financières modélisées par des processus aléatoires.

Théorème 6.3. *Le prédicteur linéaire $\hat{X}_n$ minimisant la variance de l'erreur est donné par :*

$$\hat{X}_n = \sum_{p=0}^{P-1} h_p X_{n-p-1}$$

où $\mathbf{h}$ s'écrit :

$$\mathbf{h} = \mathbf{\Sigma}_{\mathbf{Z}}^{-1} \mathbf{g}_X$$

avec : $\mathbf{g}_X := (\gamma_X[1], \dots, \gamma_X[P])$.

On note que la matrice $\mathbf{\Sigma}_{\mathbf{Z}}$ ne dépend que de $\gamma_X[0], \dots, \gamma_X[P-1]$.

Démonstration. Le prédicteur linéaire $\hat{X}_n$ est l'estimateur optimal de X_n à partir de $\mathbf{Z}^{(n)}$, soit

$$\hat{X}_n = \sum_{p=0}^{P-1} h_p Z_p^{(n)} = \sum_{p=0}^{P-1} h_p X_{n-p-1}$$

L'expression de $\mathbf{h}$ découle du théorème 6.1, avec $\boldsymbol{\gamma}_{\mathbf{Z}\mathbf{X}} = \mathbf{g}_X$. $\square$

Remarque. *Le prédicteur s'écrit comme une relation de convolution, le prédicteur peut donc être interprété comme un filtre linéaire dont la réponse impulsionnelle vaut $h[k] = h_k$ pour $0 \leq k \leq P-1$, et 0 ailleurs.*

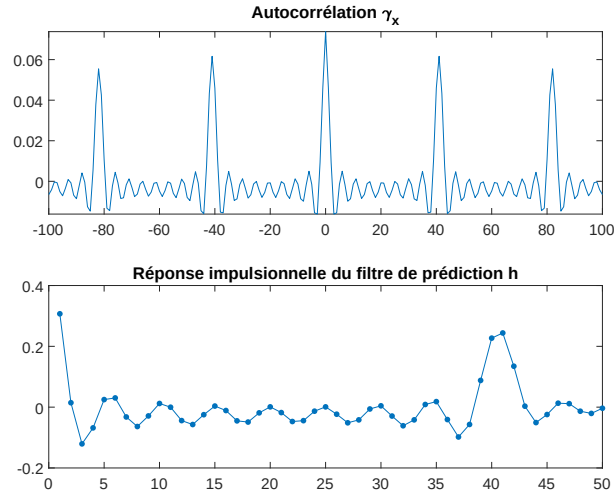


FIGURE 37 – Haut : autocorrélation du processus (γ_x) Bas : réponse impulsionnelle du filtre de prédiction (h)

Exemple 6.6. Sur la figure 37, le tracé du haut représente l'autocorrélation du processus pour $k \geq 0$. On peut remarquer que les échantillons espacés d'environ 40 points sont fortement corrélés. Le tracé du bas représente les coefficients h du prédicteur pour $P = 50$.

La figure 38 représente une réalisation du processus à prédire tracée en gras. Les prédictions réalisées avec $P = 5$ (haut) et $P = 50$ échantillons (bas) sont tracées en traits fins. Naturellement, l'usage de 50 échantillons permet d'améliorer la prédiction notamment en tirant parti de la corrélation élevée avec les échantillons placés 40 points avant celui à prédire.

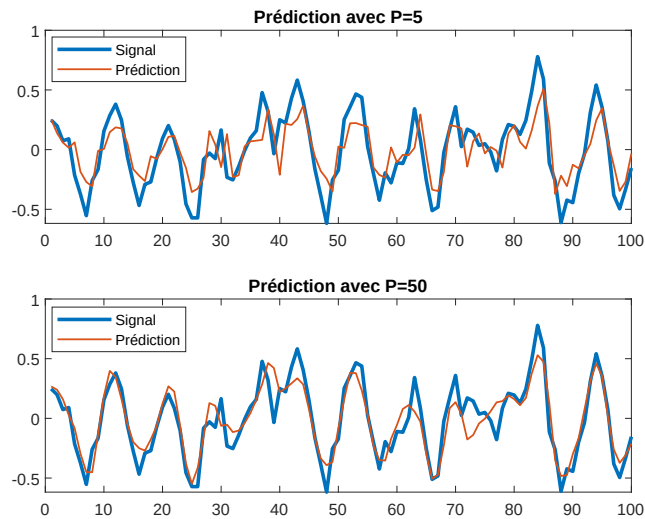


FIGURE 38 – Prédiction du processus en utilisant $P = 5$ (haut) et $P = 50$ (bas) échantillons

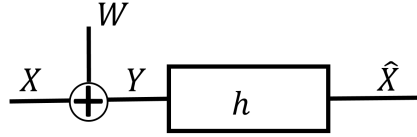


FIGURE 39 – Modèle pour le débruitage

6.4.3 Filtre de Wiener pour le débruitage

Dans cette section, nous allons étudier une application de l'estimation linéaire optimale pour le débruitage du signal.

Considérons un signal modélisé par un processus Y_n résultant de la somme d'un processus X_n contenant l'information utile et d'un processus W_n inconnu correspondant à un bruit :

$$Y_n = X_n + W_n$$

L'objectif est de réduire le bruit d'observation afin d'améliorer la qualité du signal. L'ensemble de la chaîne de traitement est représentée sur la figure 39.

En d'autres termes, il s'agit d'estimer X_n à partir des P précédents échantillons de Y_n regroupés dans le vecteur aléatoire $\mathbf{Z}^{(n)} = (Y_n, Y_{n-1}, \dots, Y_{n-P+1})$. On fait les hypothèses suivantes :

- W et X sont décorrélés,
- les autocorrélations γ_X et γ_W sont connues.

Pour chaque n , on retrouve donc un problème d'estimation linéaire optimale consistant à définir l'estimateur $\hat{X}_n$ de X_n à partir de $\mathbf{Z}^{(n)}$. On peut montrer que cet estimateur s'écrit sous la forme d'un filtre linéaire, c'est l'objet du théorème ci-dessous.

Théorème 6.4 (Filtre de Wiener). *L'estimateur linéaire optimal pour le débruitage correspond au **filtre Wiener** dont l'équation de convolution est donnée par :*

$$\hat{X}_n = \sum_{p=0}^{P-1} h_p Y_{n-p}$$

et dont la fonction de transfert est :

$$H(v) = \frac{\Gamma_X(v)}{\Gamma_X(v) + \Gamma_W(v)} = \frac{1}{1 + \Gamma_W(v)/\Gamma_X(v)}$$

Démonstration. L'écriture sous la forme d'un produit de convolution est triviale puisque pour chaque n l'estimateur linéaire optimal est par définition :

$$\begin{aligned}\hat{X}_n &= \sum_{p=0}^{P-1} h_p Z_p^{(n)} \\ &= \sum_{p=0}^{P-1} h_p Y_{n-p}\end{aligned}$$

car $\mathbf{Z}^{(n)} = (Y_n, Y_{n-1}, \dots, Y_{n-P+1})$.

Dans un deuxième temps, établissons l'expression de h en fonction de γ_X et γ_W . $\hat{X}_n$ est l'estimateur linéaire optimal donc l'erreur est décorrelée des données :

$$\begin{aligned}\mathbb{E}((X_n - \hat{X}_n)Y_p) &= 0 \quad \forall p, n \\ \mathbb{E}(X_n Y_p) - \mathbb{E}(\hat{X}_n Y_p) &= 0 \\ \mathbb{E}(X_n(X_p + W_p)) - \mathbb{E}(\hat{X}_n Y_p) &= 0 \\ \mathbb{E}(X_n X_p) + \mathbb{E}(X_n W_p) - \mathbb{E}(\hat{X}_n Y_p) &= 0 \\ \gamma_X[k] + 0 - \gamma_{\hat{X}Y}[k] &= 0 \quad \text{avec } k = n - p\end{aligned}$$

donc $\gamma_X = \gamma_{\hat{X}Y}$. Par ailleurs d'après la formule des interférences $\gamma_{\hat{X}Y}[k] = (h * \gamma_Y)[k]$. Il vient alors :

$$\begin{aligned}\gamma_X &= \gamma_{\hat{X}Y} \\ &= h * \gamma_Y \\ &= h * (\gamma_X + \gamma_W)\end{aligned}$$

car $Y_n = X_n + W_n$ avec W_n et X_n décorrelés.

On obtient donc la relation :

$$\gamma_X = h * (\gamma_X + \gamma_W)$$

En appliquant la transformée de Fourier à l'expression précédente, on obtient :

$$\Gamma_X(\nu) = H(\nu)(\Gamma_X(\nu) + \Gamma_W(\nu))$$

d'où l'expression de $H(\nu)$. □

Intuitivement, cette réponse fréquentielle s'interprète de la manière suivante : le filtre atténue le signal pour les fréquences où la DSP du bruit est plus grande que la DSP du signal et vice versa. Ceci est illustré par l'exemple ci-dessous.

L'une des principales limites de cette approche est qu'elle nécessite la connaissance de la DSP du bruit et celle du processus à débruiter.

Exemple 6.7. Soit un signal bruité $Y_n = X_n + W_n$. On connaît la DSP $\Gamma_X(\nu)$ du signal utile et celle du bruit ($\Gamma_W(\nu)$), représentées en figure 40. La réponse fréquentielle ($H(\nu)$) du filtre de Wiener utilisé pour le débruitage est représentée sur la même figure en pointillé. On remarque qu'elle s'adapte au rapport signal sur bruit dans chaque bande de fréquence : lorsque la puissance du signal est importante, le gain du filtre est élevé pour préserver le signal et par contre lorsque la puissance du bruit est plus importante que celle du signal le gain du filtre devient faible pour atténuer le bruit sans trop affaiblir le signal.

La figure 41 représente les signaux temporels : en haut, le signal bruité est comparé au signal de utile de référence ; en bas, le signal débruité est

comparé au signal de référence. On remarque que le signal après filtrage est visuellement plus proche du signal de référence. Pour quantifier l'efficacité du filtrage, on mesure le rapport signal sur bruit avant et après filtrage^a. Numériquement, sur cet exemple on obtient un incrément d'environ 5 dB, c'est à dire que la puissance du bruit a été divisée par trois !



Les codes correspondant à cet exemple sont disponibles sur édunao ainsi qu'un exemple appliqué au débruitage d'un signal audio.

a. Le rapport signal sur bruit avant filtrage correspond à la variance de X_n divisée par celle de W_n . Après filtrage, le rapport signal sur bruit correspond au rapport entre la variance de X_n et celle de $X_n - \hat{X}_n$

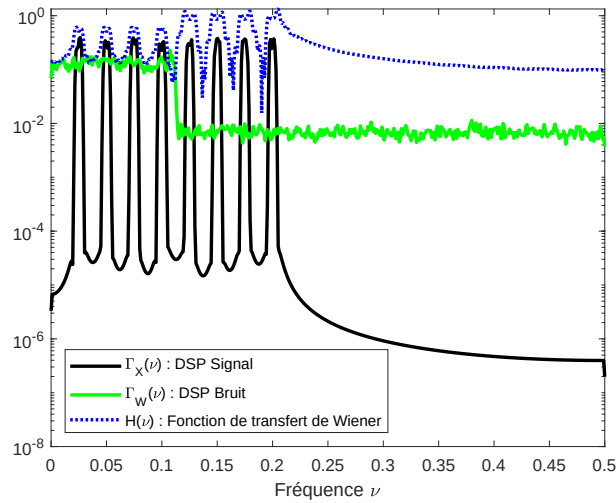


FIGURE 40 – DSP du signal ($\Gamma_X(\nu)$), du bruit ($\Gamma_W(\nu)$) et réponse fréquentielle du filtre de Wiener ($H(\nu)$)

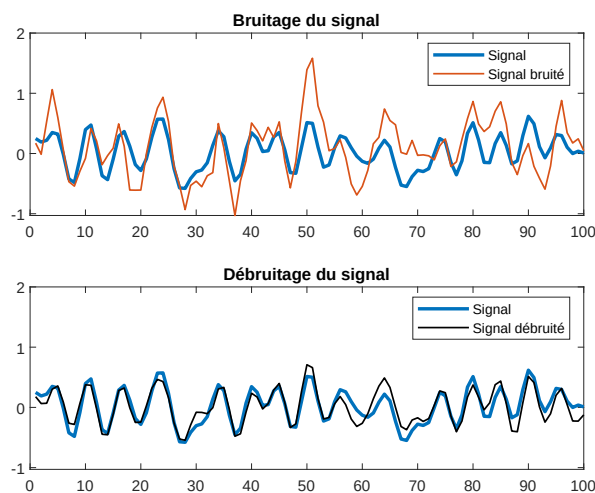


FIGURE 41 – Débruitage à l'aide du filtre de Wiener

À retenir du chapitre

- définition de l'estimateur de la moyenne d'un processus
- définition des estimateurs biaisé et non biaisé de l'auto-corrélation d'un processus
- définition de l'ergodicité à l'ordre 1 et 2
- définition du périodogramme et du periodogramme de Bartlett
- les limitations du périodogramme : biais (résolution spectrale) et variance
- principe de l'estimation paramétrique de DSP pour un modèle AR
- principe de l'estimateur linéaire optimal

INDEX

- autocorrélation, 77, 94
- autocovariance, 78
- base hilbertienne, 15
- bloquage d'ordre zéro, 71
- bourrage de zéros, 56, 59, 97
- bruit blanc, 82
- causalité, 32
- convolution, 26
- débruitage, 109
- décomposition en série de
 - Fourier, 47
- densité spectrale de puissance,
 - 80, 96
- distributions, 16
- échantillonnage, 63
- énergie, 26
- équation aux différences finies,
 - 36, 87
- équation de Yule-Walker, 88
- ergodicité, 93
- espace ℓ^p , 14
- espace L^p , 14
- espace de Schwartz, 19
- estimateur de
 - l'autocorrélation, 93
- estimateur de la moyenne, 92
- estimateur linéaire optimal,
 - 105
- estimation paramétrique, 102
- fenêtrage, 57
- filtrage de processus, 84, 110
- filtre anti-repliement, 70
- filtre de Wiener, 109
- filtre FIR, 34
- filtre IIR, 36
- fonction de Kronecker, 14
- fonction de transfert, 53
- formule des interférences, 84
- formule sommatoire de
 - Poisson, 45
- intercorrélation, 26, 77
- intercovariance, 78
- invariance par translation, 31
- lois fini-dimensionnelles, 75
- moyenne, 77
- observation, 74
- périodogramme, 96
- périodogramme de Bartlett, 99
- peigne de Dirac, 17, 45
- prédiction, 107
- processus aléatoire, 21
- processus autorégressif (AR),
 - 86, 103
- processus gaussien, 76
- puissance, 26, 80
- quantification, 25
- réalisation, 74
- réponse fréquentielle, 53
- réponse impulsionnelle, 33, 86
- résolution spectrale, 97
- rapport signal-sur-bruit, 26, 79,
 - 111
- signal à temps continu, 25
- signal à temps discret, 25
- signal numérique, 25
- signal stochastique, 26, 74
- spectre d'un signal, 40
- stabilité EBSB, 32
- stationnarité, 78
- support, 19
- théorème de Shannon, 66

traitement de la parole, 103

trajectoire, 74

transformée de Fourier, 40

transformée de Fourier à
temps discret, 48

transformée de Fourier

discrète, 51, 97

variance, 79

variante du periodogramme,
99

vecteur aléatoire, 21

BIBLIOGRAPHIE

- [1] S. Mallat. *A wavelet tour of signal processing*. Elsevier, Amsterdam, troisième édition, déc. 2008.
- [2] Jutten C. *et al.* *L'héritage de Fourier, 250 ans après*. Interstices, revue scientifique en ligne, <https://interstices.info/1-heritage-de-fourier>, Le Chesnay, nov. 2018.
- [3] L. Schwartz. *Théorie des distributions*. Hermann, 1966.
- [4] F. Golse. *Distributions, analyse de Fourier, équations aux dérivées partielles*. Éditions de l'École Polytechnique, Palaiseau, oct. 2012.
- [5] M. Vetterli, J. Kovacevic, and V. K. Goyal. *Foundations of Signal Processing*. Cambridge Univ. Press, Cambridge, UK, oct. 2014.