

Traitement du Signal

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

CentraleSupélec, Mars 2025

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	15
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	23
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	32
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 REPRÉSENTATION FRÉQUENTIELLE	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	48
3.2.1	Signaux à support limité	48
3.2.2	Signaux périodiques	49
3.3	Transformée de Fourier à temps discret	49
3.3.1	Transformée de signaux dans ℓ^1	50
3.3.2	Transformée des signaux dans ℓ^2	51
3.3.3	Quelques résultats classiques	52
3.4	Transformée de Fourier Discrète	53
3.5	Application au filtrage linéaire	54
3.5.1	Filtrage dans le domaine de Fourier	54
3.5.2	Synthèse d'un filtre FIR par TFD	57
3.6	Analyse spectrale	59
3.6.1	Fenêtrage temporel de signaux à temps continu	59
3.6.2	Résolution spectrale	61
3.6.3	Fenêtrage temporel de signaux à temps discret	62
3.7	Conclusion	62
4	ECHANTILLONNAGE ET RECONSTRUCTION	67
4.1	Echantillonnage	67
4.2	Reconstruction	70
4.2.1	Signaux à bande limitée dans $[-B, B]$	70
4.2.2	Théorème de Shannon	71
4.2.3	Démonstration du théorème de Shannon	73
4.2.4	Signaux à bande limitée dans $[A, B]$	74
4.3	L'échantillonnage en pratique	75
4.3.1	Filtre anti-repliement	75
4.3.2	Reconstruction non idéale	76
4.4	Conclusion	77
5	SIGNAUX MODÉLISÉS PAR DES PROCESSUS ALÉATOIRES	79
5.1	Définition et caractérisation des processus aléatoires	79
5.1.1	Définition	79
5.1.2	Lois fini-dimensionnelles	81
5.1.3	Processus gaussien	82
5.1.4	Moyenne et autocorrélation	82
5.2	Processus stationnaires	84
5.2.1	Stationnarité stricte	84
5.2.2	Stationnarité au sens large	84
5.2.3	Densité spectrale de puissance	86
5.2.4	Bruit blanc	88
5.2.5	Filtrage linéaire de processus	90
5.2.6	Processus autorégressif	92
5.3	Conclusion	94
6	ESTIMATION BASÉE SUR DES MODÈLES ALÉATOIRES	95
6.1	Estimateur de la moyenne d'un processus aléatoire	96
6.2	Estimation de la fonction d'autocorrélation	97
6.3	Estimation de la DSP : périodogramme	100

6.3.1	Définition	100
6.3.2	Propriétés statistiques du périodogramme	102
6.3.3	Variantes de la méthode du périodogramme	103
6.4	Estimation paramétrique de la DSP	106
6.4.1	Principe	106
6.4.2	Cas d'un signal modélisé par un processus AR	106
6.5	Estimation linéaire optimale	108
6.5.1	Définition et propriétés	109
6.5.2	Prédiction d'un processus aléatoire	110
6.5.3	Filtre de Wiener pour le débruitage	112
6.6	Conclusion	115
	Bibliographie	118

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
$x[k]$ ou 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
$\mathbf{e}_n \in \mathbb{C}^N$	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 que pour des espaces de dimension finie sur les corps $\mathbb{R}$ ou $\mathbb{C}$, toutes les normes sont équivalentes 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 (sesquilineaire) vérifiant les propriétés suivantes :

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

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

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 autrement plus complexe en dimension infinie. On appelle espace de Banach un espace vectoriel normé et complet. Nous focaliserons notre attention sur deux types d'espaces de Banach.

1.2.1 Espaces ℓ^p

Pour $p \geq 1$, 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^∞ , $\|x\|_\infty = \sup_{n \in \mathbb{Z}} |x[n]|$.

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, 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 $\int_I |x|^p d\lambda < \infty$. 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 $L^\infty(\mathbb{R})$, $\|x\|_{L^\infty} = \sup |x|$.

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

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

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 :

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

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'est notamment 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, appelées fonctions test. L'espace de départ étant ici très restreint, son dual est gigantesque.

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

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

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.

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 consistante avec le fait que si T est la distribution régulière associée à une fonction localement intégrable f , (22)

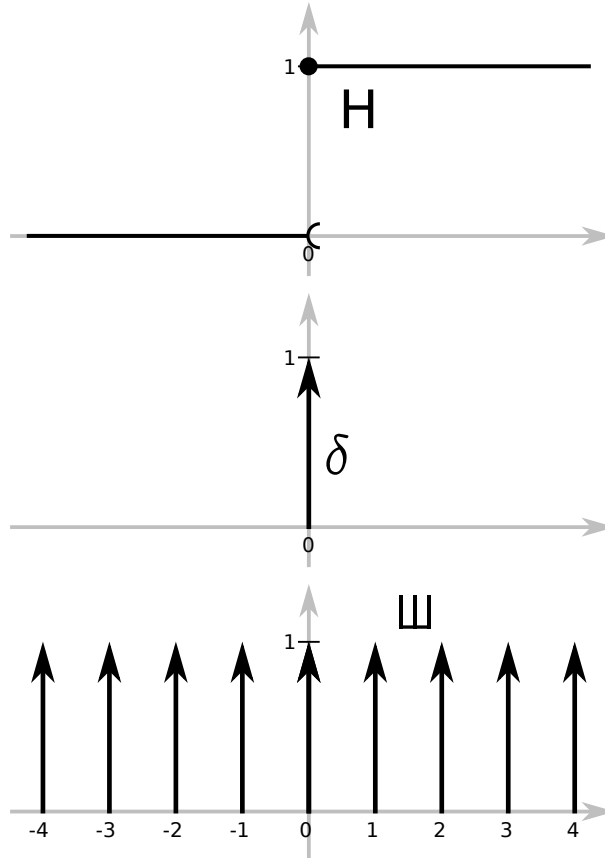


FIGURE 1 – Fonction de Heaviside $H = \mathbb{1}_{\mathbb{R}_+}$, distribution de Dirac $\delta = \delta_0$ et peigne de Dirac $\mathbb{W}_a$ pour $a = 1$.

s'identifie avec le calcul de l'intégrale de $(f \circ \chi)\varphi$ avec la formule de changement de variable.

Pour la distribution de Dirac et $\chi(x) = ax - b$, on a :

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

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-
quesoit $k \in \mathbb{N}$, 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 $e(t) = \exp(i2\pi f_0 t)$, 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 fonc-
tion 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 $\mathbf{a} = (a_j)_{1 \leq j \leq N}$, $\mathbf{a}^* \Sigma_{\mathbf{X}} \mathbf{a} = \text{Var}((\mathbf{a}, \mathbf{X})) \geq 0$.

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 entièrement 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 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 aléatoire. 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 obtenues 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.

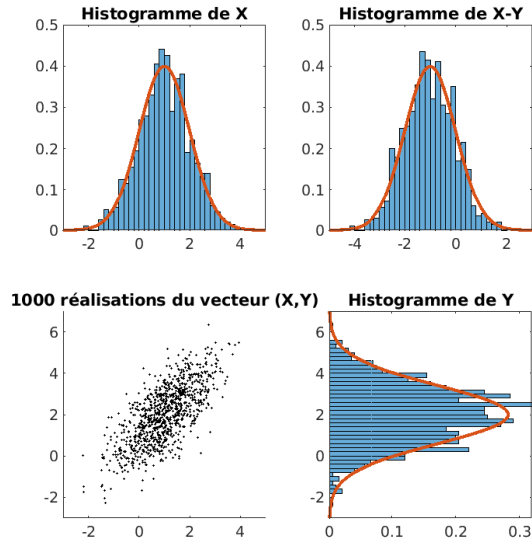


FIGURE 2 – Réalisations d'un vecteur gaussien (X, Y) , de moyenne $\mu = (1, 2)$, de matrice de covariance $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Les lois de toute combinaison linéaire X et Y sont normales.

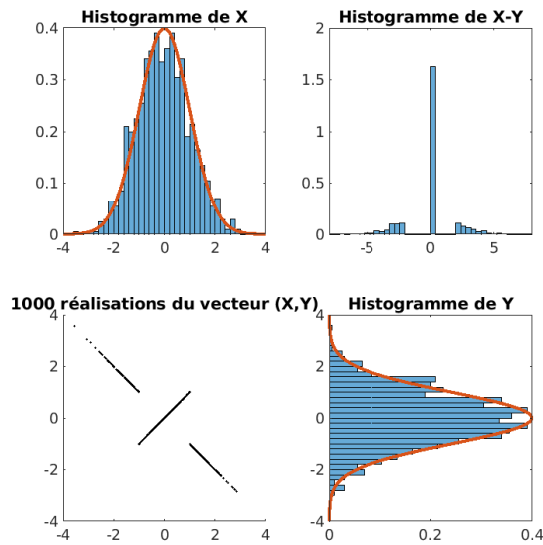


FIGURE 3 – Réalisations d'un vecteur aléatoire (X, Y) , où $Y = X\mathbb{1}_{|X|<1} - X\mathbb{1}_{|X|\geq 1}$, et X est une variable normale de loi $\mathcal{N}(0, 1)$. Les lois marginales de X et Y sont normales, mais $X - Y$ ne suis pas une loi normale. (X, Y) n'est pas donc un vecteur gaussien.

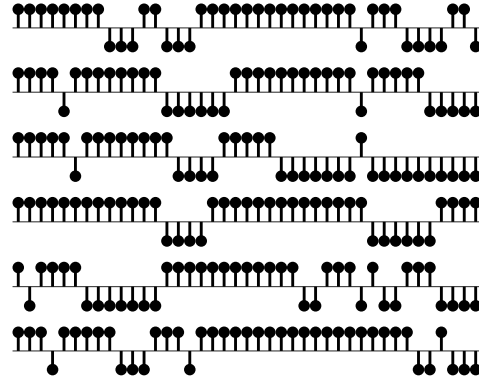


FIGURE 4 – Six trajectoires d'un processus aléatoire à valeurs -1 ou 1 .

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.

1.5.1 Brefs rappels de statistiques

Le but de la statistique est d'estimer des caractéristiques d'une variable aléatoire (moyenne, variance, paramètres 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 (mutuellement) indépendantes et identiquement distribuées (*i.i.d*) 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 le nombre de 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 :

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

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

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 $x * y(t)$. On privilégiera l'écriture $(x * y)(t)$ pour bien marquer que $(x * y)(t)$ dépend de toutes les valeurs de $y(\theta)$.

La convolution n'est pas toujours définie. Par exemple, $\forall t, (\mathbb{1}_{\mathbb{R}} * \mathbb{1}_{\mathbb{R}})(t) = \infty$. 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	presque partout
L^1	L^2	L^2	presque partout
L^1	L^∞	L^∞	partout
L^2	L^2	L^∞	partout

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

Les démonstrations, élémentaires, sont 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) est proche de 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 hypothèses suivantes est satisfaite :

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 $Y \in \mathcal{D}'(\mathbb{R})$. La convolution $X * Y$ est la distribution 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 :

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

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 a donc (au sens des distributions) $(\delta_a * y)(t) = y(t - a)$, qui généralise (50) à la convolution de δ_a et d'une fonction continue.

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) : $\text{supp}(x * y) \subset \text{supp}(x) + \text{supp}(y) = \{t_1 + t_2 : t_1 \in \text{supp}(x), t_2 \in \text{supp}(y)\}$.
- Intercorrélation : $\gamma_{xy} = x * \bar{y}$ avec $\bar{y}(t) = y^*(-t)$.

2.3 NOTION DE SYSTÈME

2.3.1 Définition

Un système est 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 par le schéma "boîte noire" de la figure 5.



FIGURE 5 – Représentation d'un système.

Un système modélise des transformations subies par un signal, comme l'altération d'un signal physique (signal audio, onde électromagnétique) lors de sa transmission, ou l'amélioration de son contenu après traitement (débruitage d'un enregistrement sonore, déconvolution pour retirer le flou dans une image). On ne considérera 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 $x = a_1 x_1 + a_2 x_2$ (avec $a_1, a_2 \in \mathbb{R}$), 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 $\theta > t$ pour calculer $y(t)$.

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 de la convolution

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.

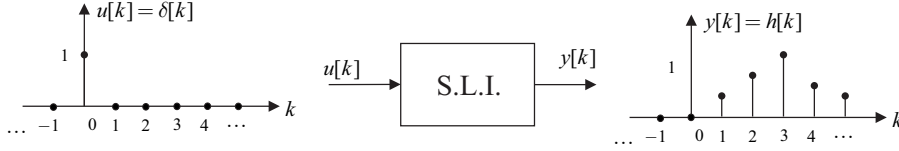


FIGURE 6 – Réponse impulsionnelle d'un SLI

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

$$h \in L^1(\mathbb{R}) \quad (\text{cas continu})$$

$$h \in \ell^1(\mathbb{Z}) \quad (\text{cas discret})$$

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

$$\text{supp}(h) \subset \mathbb{R}_+ \quad (\text{cas continu})$$

$$\text{supp}(h) \subset \mathbb{N} \quad (\text{cas discret})$$

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 $x(t) = 0$ 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*, ou filtre RIF en français) un filtre dont le support de la réponse impulsionnelle est borné (de durée finie).

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 :

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

Le support de h est $\{0, 1\}$. Il est borné, donc h est un filtre RIF.

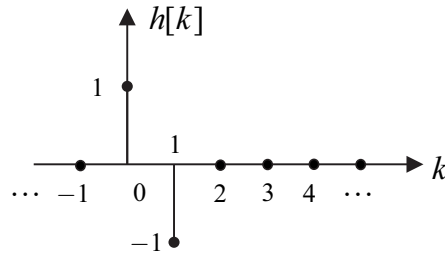


FIGURE 7 – Réponse impulsionnelle d'un filtre dérivateur.

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 courant :*

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

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 :

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

Le support de h est $\{-K, \dots, K\}$. Il est borné, donc h est un filtre RIF.

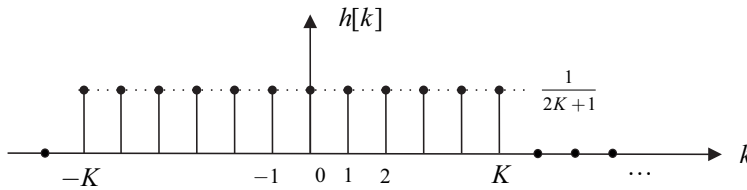


FIGURE 8 – Réponse impulsionnelle d'un filtre moyennneur.

2.4.3 Filtre à réponse impulsionnelle infinie (IIR)

Définition 2.13. On appelle filtre IIR (*Infinite Impulse Response*, ou filtre RII en français) un filtre dont la réponse impulsionnelle est à support infini.

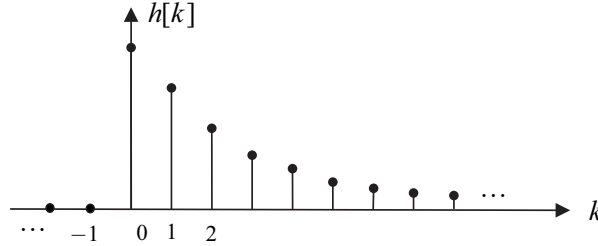


FIGURE 9 – Réponse impulsionnelle d'un filtre IIR causal.

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. *On considère un filtre stable de réponse impulsionnelle h . Le signal de sortie $y = h * e_f$, réponse du SLI à une exponentielle complexe*

$$e_f(t) = \exp(2i\pi ft)$$

est proportionnel à e_f :

$$y(t) = H(f) e_f(t).$$

Le facteur de proportionnalité est donné par la transformée de Fourier de h , appelée réponse fréquentielle :

$$H(f) = \int_{\mathbb{R}} h(t) \exp(-2i\pi ft) dt \quad (66)$$

Théorème 2.4. *On considère un filtre stable de signaux à temps discret, de réponse impulsionnelle h . Le signal de sortie $y = h * e_v$, réponse du SLI à l'exponentielle complexe*

$$e_v[k] = \exp(2i\pi vk)$$

est proportionnel à e_v :

$$y[k] = H(v) e_v[k].$$

Le facteur de proportionnalité est donné par la transformée de Fourier à temps discret de h :

$$H(v) = \sum_{k \in \mathbb{Z}} h[k] \exp(-2i\pi vk) \quad (67)$$

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 à e_f est donnée par :

$$S(e_f)(0) = (e_f * h)(0) = \int_{\mathbb{R}} h(t) \exp(-2i\pi ft) dt.$$

□

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 est amplifiée ou atténuée, et déphasée. C'est en pratique impossible, une exponentielle complexe ayant un support infini. 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, le tracé de diagrammes de Bode 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.

En temps continu, le lien entre réponse impulsionnelle et réponse fréquentielle n'a été donné ici que quand la réponse impulsionnelle est une fonction intégrable. Or, de nombreux filtres admettent comme réponse impulsionnelle une distribution (comme dans le cas élémentaire des systèmes à retard, 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.

REPRÉSENTATION FRÉQUENTIELLE

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 car une convolution dans le domaine temporel se traduit par un simple produit terme à terme 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 :

$$\forall f \in \mathbb{R}, \mathcal{F}x(f) = \int_{\mathbb{R}} e^{-2i\pi ft} x(t) dt. \quad (69)$$

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

$$\forall f \in \mathbb{R}, \mathcal{F}^*x(f) = \int_{\mathbb{R}} e^{2i\pi ft} x(t) dt \quad (70)$$

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,

$$\forall f \in \mathbb{R}, X(f) = \int_{\mathbb{R}} e^{-2i\pi ft} x(t) dt. \quad (71)$$

Remarque. Si x n'est pas dans $L^1(\mathbb{R})$, les intégrales (69) et (70) sont possiblement divergentes. Ces définitions ne sont donc plus valides. Cependant, nous verrons que la TF peut être définie sous une forme généralisée pour une large classe de signaux.

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 premiers 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 la TF d'une fonction paire ($x(-t) = x(t)$) est une fonction paire ($X(-f) = X(f)$). De même, la TF d'une fonction impaire est une fonction impaire. De plus, si x est à valeurs réelles, alors X est à symétrie hermitienne : $\forall f, X(-f) = X^*(f)$.

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

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

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

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 $x = \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}$. On calcule

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

où la fonction sinus cardinal est définie par

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

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 $f = \frac{k}{T}$, $k \in \mathbb{Z}^*$. $X(0)$ coïncide avec son intégrale, cf. (72), égale à T .

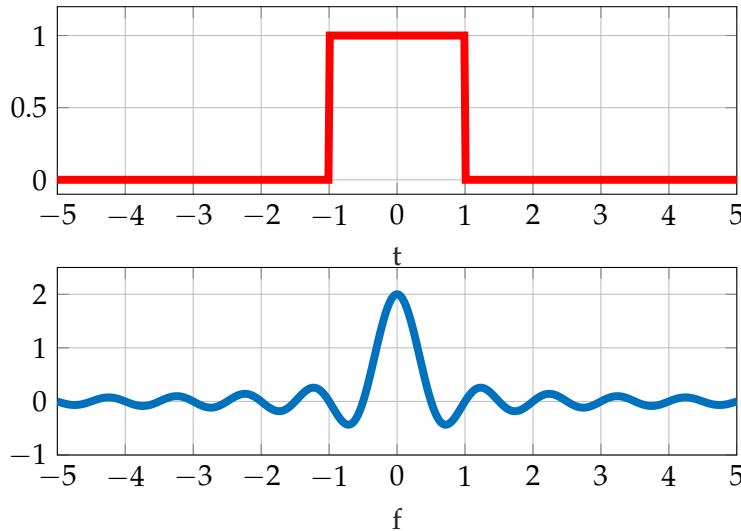


FIGURE 10 – Fenêtre rectangulaire de largeur 2. Sa TF est un sinus cardinal qui s'annule si f est multiple de $1/2$ et non nul.

TABLE 1 – Propriétés de la transformée de Fourier.

	$z(t)$	$Z(f)$
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	$x(t - \tau_0)$	$e^{-2i\pi\tau_0 f} X(f)$
Modulation	$e^{2i\pi f_0 t} x(t)$	$X(f - f_0)$
Dérivée	$x^{(n)}(t)$	$(2i\pi f)^n X(f)$
Produit par t^n	$t^n x(t)$	$\mathcal{F}x^{(n)}(f) / (-2i\pi)^n$

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 produit, en notant que $x(t - \tau_0) = (\delta_{\tau_0} * x)(t)$. Les TF de δ_{τ_0} et $e_{f_0}(t) = e^{2i\pi f_0 t}$ 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 $\mathcal{S}(\mathbb{R})$ (voir définition (24)) 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 reste 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 $\varphi \in \mathcal{S}(\mathbb{R})$,

$$\langle \mathcal{F}\delta_{f_0}, \varphi \rangle = \langle \delta_{f_0}, \mathcal{F}\varphi \rangle = (\mathcal{F}\varphi)(f_0) = \int_{\mathbb{R}} e^{-2i\pi f_0 t} \varphi(t) dt = \langle e_{-f_0}, \varphi \rangle.$$

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

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

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

Nous avons en particulier pour $f_0 = 0$:

$$\mathcal{F}\mathbb{1}_{\mathbb{R}} = \delta_0 = \delta.$$

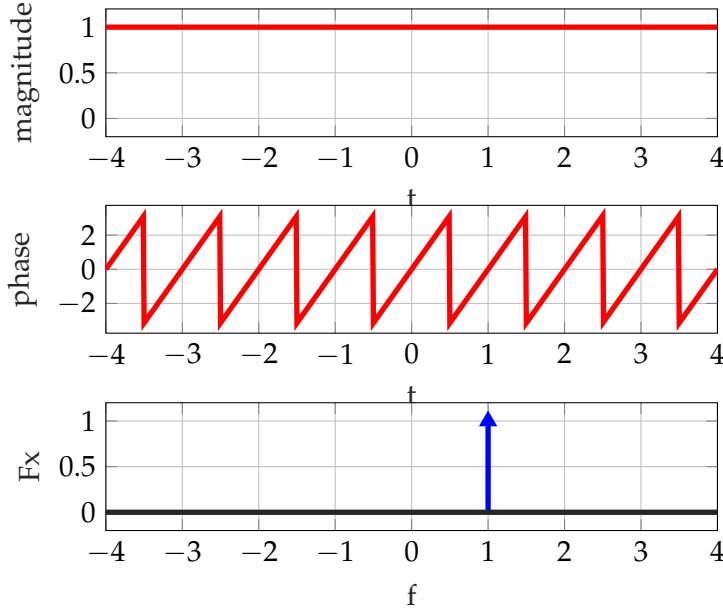


FIGURE 11 – Fonction $e_{f_0}(t) = \exp(2i\pi f_0 t)$ pour $f_0 = 1$: variations temporelles de l'amplitude et de la phase, tracé de sa TF.

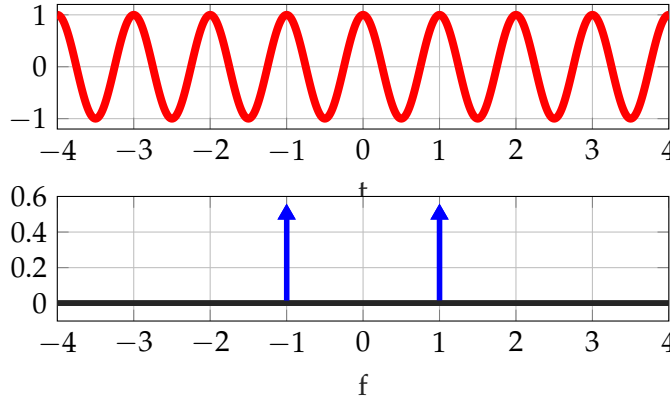


FIGURE 12 – Fonction $\cos(2\pi f_0 t)$ pour $f_0 = 1$: représentations temporelle et fréquentielle.

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*

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

Démonstration. La preuve est disponible dans le recueil de TD de CIP avec des hypothèses un peu plus faibles :

1. $\varphi \in L^1(\mathbb{R}) \cap \mathcal{C}^0(\mathbb{R})$,
2. $\exists M > 0, \exists \alpha > 1 : \forall x \in \mathbb{R}, |\varphi(x)|(1 + |x|^\alpha) \leq M$,

$$3. \sum_{k \in \mathbb{Z}} |\mathcal{F}\varphi(kT)| < \infty.$$

Ces hypothèses sont toutes vérifiées pour $\varphi \in \mathcal{S}(\mathbb{R})$, car :

1. $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R})$,
2. φ est à décroissance rapide, voir (24),
3. d'après le lemme 3.1, $\mathcal{F}\varphi \in \mathcal{S}(\mathbb{R})$. $\mathcal{F}\varphi$ est donc à décroissance rapide.

□

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

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

Démonstration. L'égalité (82) s'obtient en appliquant (81) aux fonctions test φ :

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

□

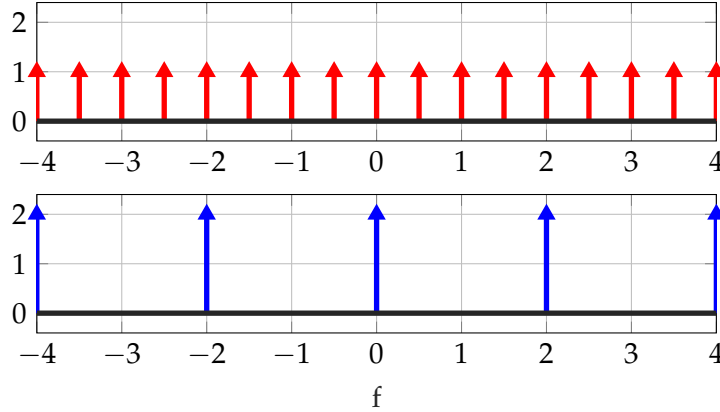


FIGURE 13 – Peigne de Dirac : représentations temporelle $\mathbb{W}_T$ et fréquentielle $\frac{1}{T} \mathbb{W}_{\frac{1}{T}}$ pour $T = \frac{1}{2}$.

3.1.4 Quelques exemples classiques

Les résultats les plus classiques sont récapitulés dans le tableau 2. On détaille maintenant deux calculs très classiques.

TABLE 2 – TF de fonctions et distributions usuelles. Certaines expressions (lignes 1 et 2 par exemple) sont symétriques du fait de la propriété de dualité : $(\mathcal{F}\mathcal{F}x)(t) = x(-t)$.

x	$x \in \dots$	X
$\mathbb{1}_{[-T,T]}$	$L^1(\mathbb{R})$	$f \mapsto 2T \sin_c(2\pi T f)$
$t \mapsto \sin_c(2\pi f_0 t)$	$L^2(\mathbb{R})$	$\frac{1}{2f_0} \mathbb{1}_{[-f_0, f_0]}$
$\mathbb{1}_{\mathbb{R}}$	$\mathcal{S}'(\mathbb{R})$	δ
δ	$\mathcal{S}'(\mathbb{R})$	$\mathbb{1}_{\mathbb{R}}$
$t \mapsto \exp(2i\pi f_0 t)$	$\mathcal{S}'(\mathbb{R})$	δ_{f_0}
δ_{τ_0}	$\mathcal{S}'(\mathbb{R})$	$f \mapsto \exp(-2i\pi \tau_0 f)$
$t \mapsto \cos(2\pi f_0 t)$	$\mathcal{S}'(\mathbb{R})$	$\frac{\delta_{-f_0} + \delta_{f_0}}{2}$
$t \mapsto \sin(2\pi f_0 t)$	$\mathcal{S}'(\mathbb{R})$	$\frac{\delta_{f_0} - \delta_{-f_0}}{2i}$
$\mathbb{1}_T$	$\mathcal{S}'(\mathbb{R})$	$\frac{1}{T} \mathbb{1}_{\frac{1}{T}}$
$t \mapsto \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{t^2}{2\sigma^2})$	$L^1(\mathbb{R})$	$f \mapsto \exp(-2\sigma^2\pi^2 f^2)$

Le premier est la TF d'une fonction triangle. Soit

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

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

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

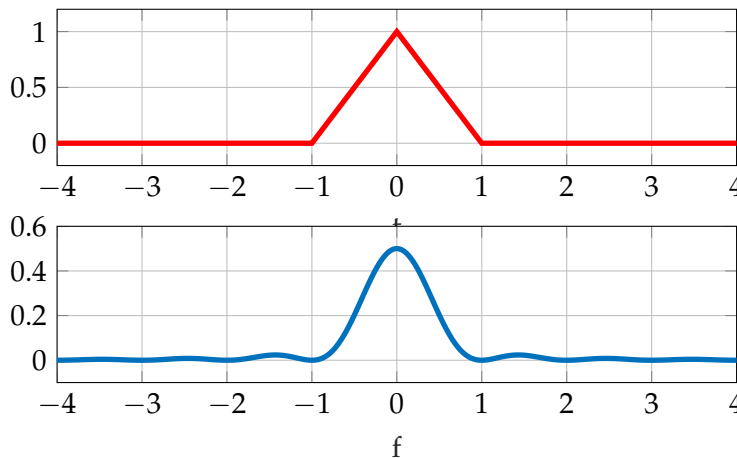


FIGURE 14 – Fonction $\Delta_{[-T,T]}$ pour $T = 1$: représentations temporelle et fréquentielle.

on déduit l'expression de la TF d'une porte, cf. (73), et de la propriété sur la TF du produit de convolution, que :

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

Le second calcul est la TF d'un sinus cardinal. On pose $s(t) = \sin_c(\pi t)$. D'après la propriété de dualité de la transformée de Fourier (voir tableau 2) et par parité de la fonction porte,

$$\mathcal{F}\mathcal{F}\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]} = \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}.$$

Comme $\mathcal{F}\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]} = s$, l'équation précédente se récrit $\mathcal{F}s = \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}$, c'est-à-dire

$$\mathcal{F}\{\sin_c(\pi t)\}(f) = \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(f). \quad (86)$$

En appliquant la propriété de dilatation/contraction avec un facteur $\alpha = 2f_0$, ce résultat se généralise ainsi (pour $f_0 > 0$) :

$$\begin{aligned} \mathcal{F}\{\sin_c(2\pi f_0 t)\}(f) &= \frac{1}{2f_0} \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]} \left(\frac{f}{2f_0} \right) \\ &= \frac{1}{2f_0} \mathbb{1}_{[-f_0,f_0]}(f). \end{aligned} \quad (87)$$

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 :*

$$\begin{aligned} e_k : [0, T] &\rightarrow \mathbb{C} \\ t &\mapsto \exp\left(2i\pi \frac{kt}{T}\right) \end{aligned}$$

La décomposition s'écrit

$$x = \sum_{k \in \mathbb{Z}} c_k e_k \quad (88)$$

où

$$c_k = (x, e_k) = \frac{1}{T} \int_0^T \exp\left(-2i\pi \frac{kt}{T}\right) x(t) dt. \quad (89)$$

On a de plus l'égalité de Parseval :

$$\|x\|_{L^2(0,T)}^2 = \frac{1}{T} \int_{[0,T]} |x(t)|^2 dt = \sum_{k \in \mathbb{Z}} |c_k|^2. \quad (90)$$

La série (88) 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 (88) 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

La transformée de Fourier de signaux périodiques est définie au sens des distributions. Soit x un signal T -périodique et $x_T := x\mathbb{1}_{[0,T]}$ son motif période. On écrit $\forall t, x(t) = \sum_{k \in \mathbb{Z}} x_T(t - kT)$, ou encore

$$x = x_T * \mathbb{I}_T. \quad (91)$$

On suppose que $x_T \in L^1(0, T)$, de telle sorte que $X_T \in \mathcal{C}^\infty(\mathbb{R})$. 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

$$X = \sum_{k \in \mathbb{Z}} c_k \delta_{\frac{k}{T}} \quad (92)$$

où

$$\begin{aligned} c_k &= \frac{1}{T} X_T\left(\frac{k}{T}\right) \\ &= \frac{1}{T} \int_0^T \exp\left(-2i\pi \frac{kt}{T}\right) x(t) dt \end{aligned} \quad (93)$$

sont égaux aux coefficients de Fourier de x_T définis en (89).

On retiendra qu'une fonction périodique possède un spectre de raies conformément à (92). 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

$$\forall \nu \in \mathbb{R}, \mathcal{F}x(\nu) = X(\nu) = \sum_{k \in \mathbb{Z}} x[k] e^{-2i\pi \nu k} \quad (94)$$

lorsque cette série converge (notamment pour $x \in \ell^1(\mathbb{Z})$). Dans le cas contraire, la définition (94) n'est plus valide. Ceci étant, la TFTD peut être définie pour des classes plus étendues de signaux (dans

$\ell^2(\mathbb{Z})$, pour des signaux périodiques) en utilisant le formalisme des distributions.

On notera $X := \mathcal{F}x$, et on utilisera $\mathcal{F}\{x[k]\}(\nu)$ lorsqu'il est nécessaire d'expliciter la dépendance de x par rapport à k .

Remarque (fréquence réduite, hautes et basses fréquences). *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.

$X(\nu)$ étant 1-périodique, les basses fréquences correspondent à $\nu \approx 0$, mais aussi à $\nu \approx \ell$ pour $\ell \in \mathbb{Z}$. Les hautes fréquences correspondent à $\nu \approx \frac{1}{2} + \ell$, $\ell \in \mathbb{Z}$.

$|X(\nu)|$ est généralement représenté sur une période, pour $\nu \in [0, 1]$ (avec la fréquence la plus haute $\nu = \frac{1}{2}$ au centre de la figure, et les fréquences les plus basses $\nu = 0$ et 1 aux extrémités) ou pour $\nu \in [-\frac{1}{2}, \frac{1}{2}]$ (avec la fréquence nulle au centre de la figure, et les hautes fréquences $\nu \approx \pm \frac{1}{2}$ aux extrémités).

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

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

Nous avons également, à partir de la définition,

$$X(0) = \sum_{k \in \mathbb{Z}} x[k]. \quad (95)$$

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 :

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

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

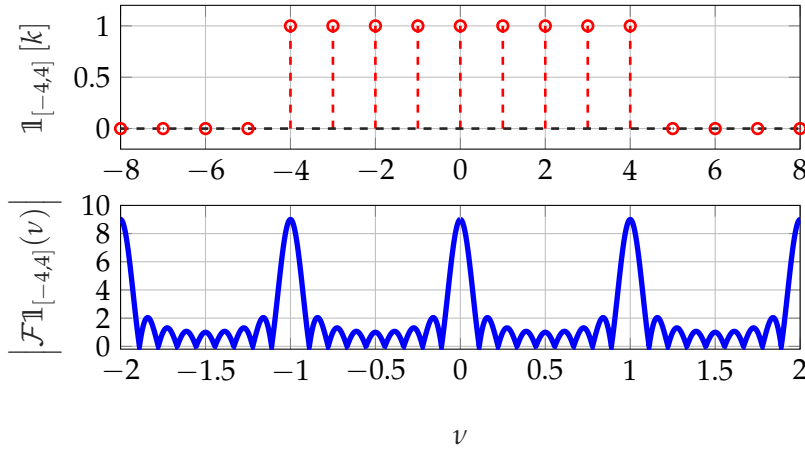


FIGURE 15 – Fenêtre rectangulaire discrète $\mathbb{1}_{[-N,N]}$: représentations temporelle et fréquentielle (en module). Comme pour tout signal discret, la TFTD est 1-périodique.

3.3.2 Transformée des signaux dans ℓ^2

Théorème 3.10. *L'opérateur $\mathcal{F}$ définissant la TFTD 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

$$x[k] = \int_{[0,1[} X(\nu) e^{2i\pi k\nu} d\nu \quad (96)$$

Remarque. (94) 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 (96). X étant 1-périodique, son énergie est définie comme l'intégrale de $|X(\nu)|^2$ 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 $\otimes$ est définie par :

$$\forall \nu \in \mathbb{R}, (X \otimes Y)(\nu) = \int_{[0,1[} X(\theta) Y(\nu - \theta) d\theta. \quad (97)$$

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	$x[k - k_0]$	$e^{-2i\pi\nu k_0} X(\nu)$
Modulation	$e^{2i\pi\nu_0 k} x[k]$	$X(\nu - \nu_0)$
Produit	$x[k] y[k]$	$(X \otimes Y)(\nu)$
Convolution	$(x * y)[k]$	$X(\nu) Y(\nu)$

TABLE 4 – Récapitulatif des transformées de Fourier de signaux à temps discret. La TFTD est à chaque fois 1-périodique.

$x[k]$	$x \in \dots$	X
$\delta[k]$ (fonction de Kronecker)	$\ell^1(\mathbb{Z})$	$\mathbb{1}_{\mathbb{R}}$
$\delta[k - k_0]$	$\ell^1(\mathbb{Z})$	$\nu \mapsto \exp(-2i\pi\nu k_0)$
$\mathbb{1}_{[-N,N]}[k] = \sum_{n=-N}^N \delta[k - n]$	$\ell^1(\mathbb{Z})$	$\nu \mapsto \frac{\sin(\pi(2N+1)\nu)}{\sin(\pi\nu)}$
$\mathbb{1}_{\mathbb{Z}}[k] = \sum_{n=-\infty}^{\infty} \delta[k - n]$	$\ell^\infty(\mathbb{Z})$	$\mathbb{1}_{\mathbb{Z}}$ (peigne de Dirac)
$\exp(2i\pi\nu_0 k)$	$\ell^\infty(\mathbb{Z})$	$\sum_{\ell \in \mathbb{Z}} \delta_{\nu_0 + \ell}$ (peigne de Dirac)
$\cos(2\pi\nu_0 k)$	$\ell^\infty(\mathbb{Z})$	$\frac{1}{2} \sum_{\ell \in \mathbb{Z}} (\delta_{-\nu_0 + \ell} + \delta_{\nu_0 + \ell})$
$\sin(2\pi\nu_0 k)$	$\ell^\infty(\mathbb{Z})$	$\frac{1}{2i} \sum_{\ell \in \mathbb{Z}} (\delta_{\nu_0 + \ell} - \delta_{-\nu_0 + \ell})$

3.3.3 Quelques résultats classiques

Le tableau 4 récapitule des résultats classiques de TFTD. Certains signaux considérés comme les exponentielles complexes, ne se trouvent ni dans ℓ^1 ni dans ℓ^2 , mais dans ℓ^∞ . La définition de la TFTD de tels signaux est (très) subtile et ne sera pas détaillée ici. Le lecteur intéressé est redirigé vers le livre [5]. On retiendra que **la TFTD d'une sinusoïde discrète est la somme de deux peignes de Dirac**, et que la TFTD d'une fonction porte n'est pas égale à un sinus cardinal mais à un rapport de deux sinus.

Remarque (Méthode de calcul). Pour les signaux dans ℓ^1 , la TFTD se calcule par la formule (94). Pour des signaux n'étant pas dans ℓ^1 , il convient de combiner les résultats classiques du tableau 4 avec les propriétés énoncées dans le tableau 3 : TFTD d'un signal retardé, modulé, d'une combinaison linéaire, etc.

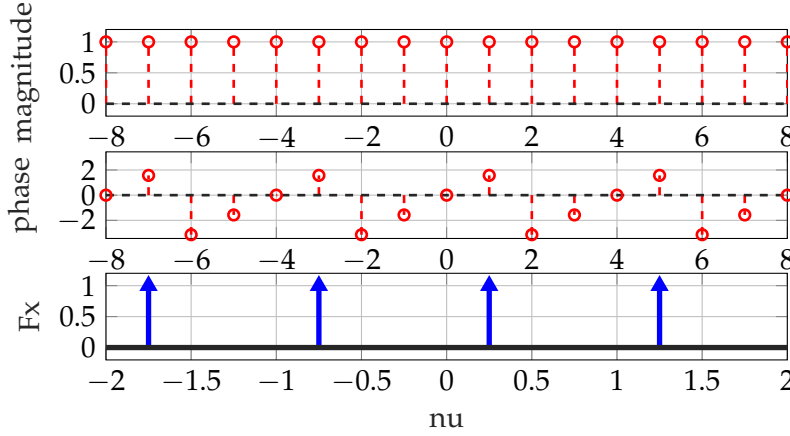


FIGURE 16 – Exponentielle complexe $\exp(2i\pi\nu_0k)$: représentations temporelle (amplitude et phase) et fréquentielle pour $\nu_0 = 0,25$. La TFTD est 1-périodique.

3.4 TRANSFORMÉE DE FOURIER DISCRÈTE

La transformée de Fourier discrète (à ne pas confondre avec la TFTD) est l'outil utilisé en pratique pour traiter un signal numérique. 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

$$\mathbf{F}_N[n, k] = e^{-2i\pi \frac{nk}{N}} \quad (98)$$

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

$$X_n = \sum_{k=0}^{N-1} \mathbf{F}_N[n, k] x_k = \sum_{k=0}^{N-1} e^{-2i\pi \frac{kn}{N}} x_k. \quad (99)$$

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*

$$\mathbf{F}_N^{-1} = \frac{1}{N} \mathbf{F}_N^*. \quad (100)$$

Ainsi, la TFD inverse prend la forme

$$x_k = \frac{1}{N} \sum_{n=0}^{N-1} e^{2i\pi \frac{nk}{N}} X_n \quad \text{pour } k \in \{0, \dots, N-1\}. \quad (101)$$

Remarque. (101) s'interprète comme la décomposition de $\mathbf{x}$ dans la base orthogonale formée par les signaux exponentiels complexes de fréquence $\frac{n}{N}$, notés $\mathbf{e}_n = \{\exp(2i\pi \frac{nk}{N}), k = 0, \dots, N-1\}$:

$$\mathbf{x} = \sum_{n=0}^{N-1} \left(\frac{1}{N} X_n \right) \mathbf{e}_n \quad (102)$$

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

Les calculs de produit matrice-vecteur du type $\mathbf{F}_N \mathbf{x}$ et $\mathbf{F}_N^{-1} \mathbf{X}$ coûtent habituellement $\mathcal{O}(N^2)$ opérations élémentaires (additions et multiplications). Cependant, on dispose d'algorithmes rapides (appelés FFT, pour *Fast Fourier Transform*) pour 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 :

$$\forall f \in \mathbb{R}, Y(f) = H(f) X(f) \quad (103)$$

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 :

$$y = H(f) e_f. \quad (104)$$

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 élimine les hautes fréquences correspondant aux discontinuités (sauts des valeurs du signal) et au bruit contenu dans le signal, et produit une version lissée du signal de départ. 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 appellation "idéal". En pratique, on a recours à des techniques de fenêtrage pour concevoir un filtre causal et stable dont la réponse fréquentielle est 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 :

$$\forall v, Y(v) = H(v) X(v) \quad (105)$$

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

Définition 3.6. Un filtre numérique est dit :

— passe-haut s'il laisse passer les hautes fréquences :

$$H(v) \neq 0 \text{ pour } v_c \leq |v| \leq \frac{1}{2}$$

et coupe les basses fréquences :

$$H(\nu) \approx 0 \text{ pour } |\nu| \leq \nu_c,$$

en notant $\nu_c \in (0, \frac{1}{2})$ la fréquence de coupure ;

- passe-bas s'il laisse passer les basses fréquences ($H(0) \neq 0$) et coupe les hautes fréquences ($H(\pm\frac{1}{2}) = 0$) ;
- passe-bande s'il ne laisse passer qu'un intervalle de fréquences ($H(\nu) \neq 0 \iff |\nu| \in [\nu_1, \nu_2]$, avec $0 < \nu_1 < \nu_2 < \frac{1}{2}$) ;
- coupe-bande s'il coupe les fréquences dans un certain intervalle ($H(\nu) \approx 0 \iff |\nu| \in [\nu_1, \nu_2]$, avec $0 < \nu_1 < \nu_2 < \frac{1}{2}$).

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

$$h[k] = \delta[k] - \delta[k-1] \quad (106)$$

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

$$\begin{aligned} H(\nu) &= h[0] + h[1] \exp(-2i\pi\nu) \\ &= 1 - \exp(-2i\pi\nu) \\ &= 2i \exp(-i\pi\nu) \sin(\pi\nu). \end{aligned} \quad (107)$$

h est un filtre passe-haut car $H(0) = 0$ et les valeurs maximales de $|H(\nu)|$ sont atteintes pour $|\nu| = \frac{1}{2}$.

Exemple 3.7. Le filtre à moyenne glissante est défini par

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

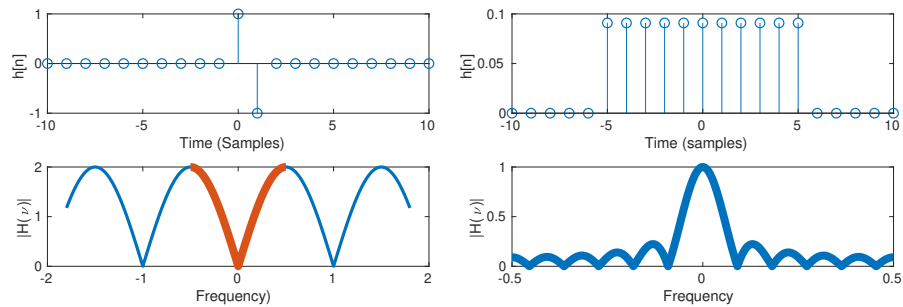


FIGURE 17 – Réponses impulsionnelle et fréquentielle des filtres dérivateur (gauche) et à moyenne glissante (droite). Ces filtres sont respectivement passe-haut et passe-bas.

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 :

$$\forall k, y[k] = \sum_{\ell=1}^p a_{\ell} y[k-\ell] + x[k] \quad (108)$$

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) :

$$\forall \nu, Y(\nu) = \sum_{\ell=1}^p a_{\ell} e^{-2i\pi\ell\nu} Y(\nu) + X(\nu). \quad (109)$$

Par identification avec (105), on déduit que :

$$\forall \nu, H(\nu) = \frac{1}{1 - \sum_{\ell=1}^p a_{\ell} e^{-2i\pi\ell\nu}}. \quad (110)$$

3.5.2 Synthèse d'un filtre FIR par TFD

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

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

On a recours à la TFD pour calculer la réponse fréquentielle du filtre, $(\mathcal{F}h)(\nu)$, en des points ν particuliers.

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. (99)) coïncide avec l'évaluation de $(\mathcal{F}h)(\nu)$ pour des fréquences multiples de $\frac{1}{p}$:

$$\forall n \in \{0, \dots, p-1\}, H_n = (\mathcal{F}h) \left(\frac{n}{p} \right). \quad (111)$$

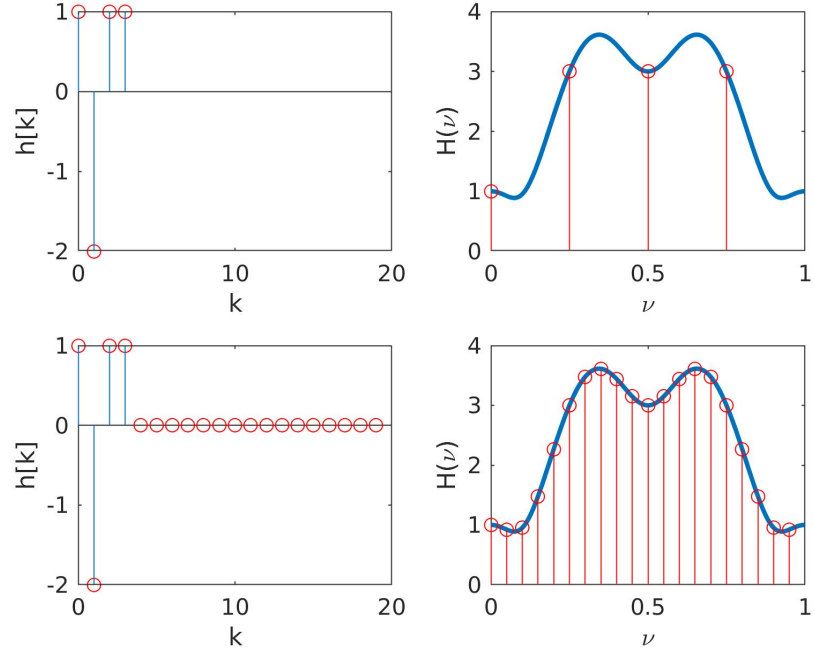


FIGURE 18 – Calcul de la réponse fréquentielle d'un filtre FIR. A droite, la TFTD $(\mathcal{F}h)(\nu)$ est représentée en trait continu, et son évaluation (pour 4 et 20 points fréquentiels, respectivement) par TFD est en rouge. Le bourrage de zéros (figures du bas) améliore la représentation fréquentielle.

Ainsi, $\mathcal{F}h$ est évaluée sur une grille fréquentielle dont le pas (égal à l'espacement entre deux points sur cette grille) est $\frac{1}{p}$.

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 une grille de pas $\frac{1}{M}$:

$$\forall m \in \{0, \dots, M-1\}, H'_m = (\mathcal{F}h) \left(\frac{m}{M} \right). \quad (112)$$

L'intérêt du bourrage de zéros est illustré sur la Fig. 18 : le nombre de points fréquentiels considérés est largement supérieur lorsque le bourrage de zéros est pratiqué. Il en résulte une précision du tracé (en rouge) beaucoup plus grande.

Une fois la réponse fréquentielle du filtre calculée, on peut calculer la représentation fréquentielle de la sortie $y = h * x$ associée à une entrée x donnée, par la formule $Y(\nu) = (\mathcal{F}h(\nu)) X(\nu)$. Pour cela, 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 $\frac{1}{M}$ en suivant (112).

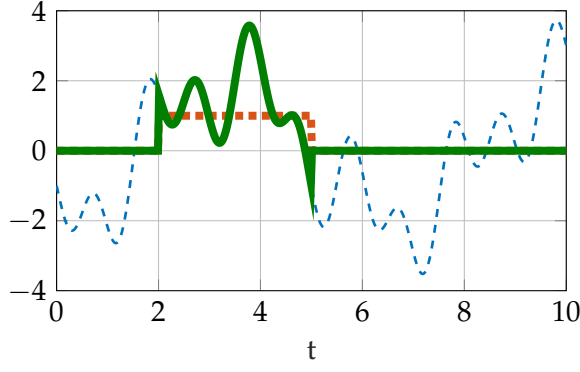


FIGURE 19 – Fenêtrage rectangulaire : $x(t) = s(t) \mathbb{1}_{[a,b]}(t)$. Le signal observé x est tracé en vert.

3.6 ANALYSE SPECTRALE

L'analyse spectrale consiste à décomposer un signal comme une somme de sinusôides. 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 Fenêtrage temporel de signaux à temps continu

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 :

$$x(t) = s(t) \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}(t) \quad (113)$$

On déduit des propriétés de la TF que :

$$\begin{aligned} X &= S * \mathcal{F} \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]} \\ &= S * \{T \operatorname{sinc}(\pi T f)\}. \end{aligned} \quad (114)$$

Prenons l'exemple du signal sinusoïdal $s(t) = \cos(2\pi f_0 t)$, dont la TF s'écrit $S = \frac{1}{2}(\delta_{f_0} + \delta_{-f_0})$. (114) se réécrit :

$$\forall f, X(f) = \frac{T}{2} (\operatorname{sinc}(\pi T (f - f_0)) + \operatorname{sinc}(\pi T (f + f_0))). \quad (115)$$

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.

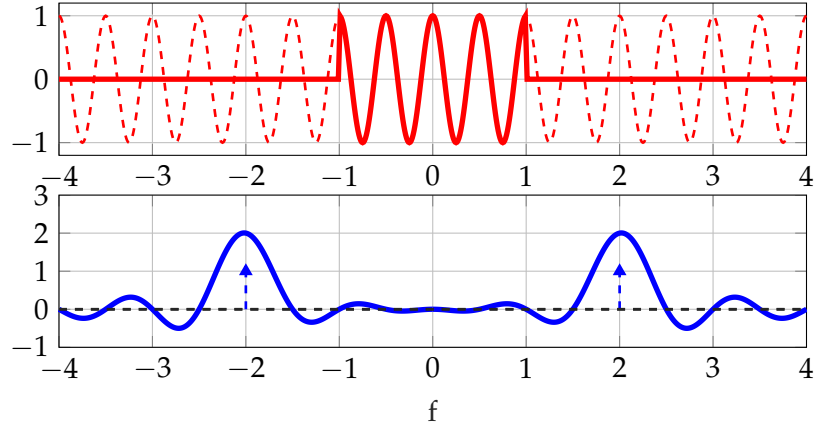


FIGURE 20 – Sinusoïde avec fenêtrage rectangulaire : représentations temporelle et fréquentielle.

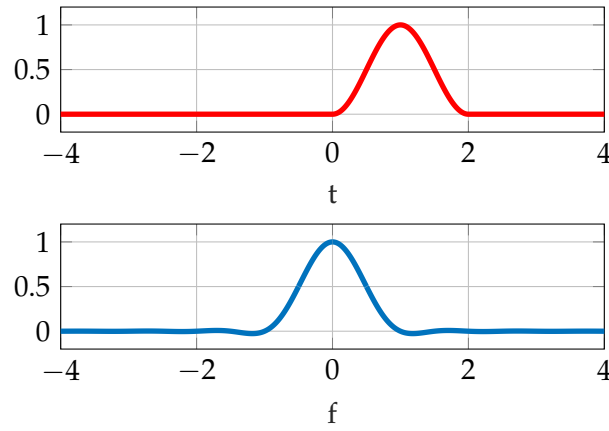


FIGURE 21 – Fenêtre de Hann : représentations temporelle et fréquentielle (en module).

Dans le cas d'un signal ayant un contenu plus riche 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 (113) par

$$\forall t \in \mathbb{R}, x(t) = s(t) w_{[-\frac{T}{2}, \frac{T}{2}]}(t) \quad (116)$$

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

$$w_{[-\frac{T}{2}, \frac{T}{2}]}(t) = \cos^2\left(\frac{\pi t}{T}\right) \mathbb{1}_{[-\frac{T}{2}, \frac{T}{2}]}(t) \quad (117)$$

représentée en Fig. 21. Ces fenêtres sont de largeur T et conduisent à une décroissance rapide des lobes secondaires (en $\frac{1}{f^2}$ ou $\frac{1}{f^3}$, respectivement) au prix d'un élargissement de la largeur du lobe principal

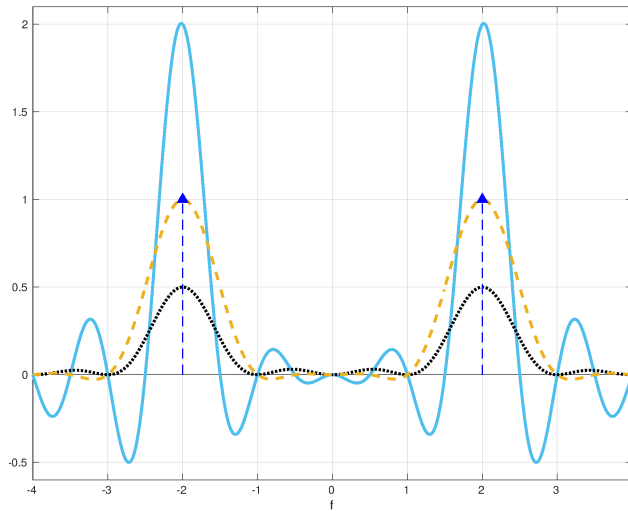


FIGURE 22 – Effets de fenêtrage sur une sinusoïde : fenêtrages rectangulaire (trait plein), triangulaire (pointillés noirs), de Hann (pointillés jaunes).

($\frac{4}{T}$), voir Fig. 22. En comparaison, avec une fenêtre rectangulaire, l'amplitude des lobes secondaires décroît relativement lentement (en $\frac{1}{f}$) et la largeur du lobe principale est de $\frac{2}{T}$.

3.6.2 Résolution spectrale

La notion de résolution spectrale est liée à la capacité à distinguer des sinusoïdes, ou plus précisément, des exponentielles complexes à des fréquences proches. La résolution spectrale est influencée par la fenêtre de troncature considérée.

Définition 3.7 (Résolution spectrale). Soit s_1 et s_2 deux exponentielles complexes de même amplitude, de fréquences respectives f_1 et f_2 , définies sur une fenêtre d'observation de durée T . La résolution spectrale est l'écart fréquentiel minimal pour lequel les deux sinusoïdes peuvent être distinguées par transformée de Fourier, quelque soit leur déphasage.

Pour une fenêtre rectangulaire, la résolution spectrale vaut $\frac{2}{T}$. Il suffit que

$$|f_1 - f_2| \geq \frac{2}{T} \quad (118)$$

pour pouvoir distinguer les deux exponentielles complexes.

Pour un mélange de plus de deux exponentielles complexes, la règle (118) est appliquée à l'écart minimal entre les fréquences de deux des exponentielles.

Le résultat (118) exprime le fait qu'on peut distinguer d'autant plus facilement deux sinusoïdes à des fréquences proches que la durée d'observation est grande. La formule (118) est liée à l'expression de la TF de la fenêtre rectangulaire : $T \sin_c(\pi T f)$, cf. (114). La largeur du lobe principal de ce motif vaut $\frac{2}{T}$ car $\sin_c(\pi T f)$ s'annule en $f = \pm \frac{1}{T}$.

3.6.3 Fenêtrage temporel de signaux à temps discret

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

$$\forall k \in \mathbb{Z}, x[k] = s[k] w_{\{0, \dots, N-1\}}[k] \quad (119)$$

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

Pour calculer numériquement la TTFD de x , on ne conserve que les points k pour lesquels $w_{\{0, \dots, N-1\}}[k] \neq 0$. On dispose donc d'un vecteur $\mathbf{x} = \{x_k, k = 0, \dots, N-1\}$ de longueur N , en notant $x_k = x[k]$. La technique du bourrage de zéros introduite dans la section 3.5.2 consiste à ajouter artificiellement $M - N$ zéros à ce vecteur, pour $M > N$. La TFD du vecteur $\mathbf{x}$ augmenté fournit une évaluation de $\mathcal{F}x(v)$ sur la grille fréquentielle

$$v_m = \frac{m}{M}, \quad m = 0, \dots, M-1$$

de pas $\frac{1}{M}$. Notons $\mathbf{X} = \text{TFD}(\mathbf{x})$. En localisant la position ($v = \frac{m}{M}$), l'amplitude et la phase des maxima locaux de $|X_m|$, 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.

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.

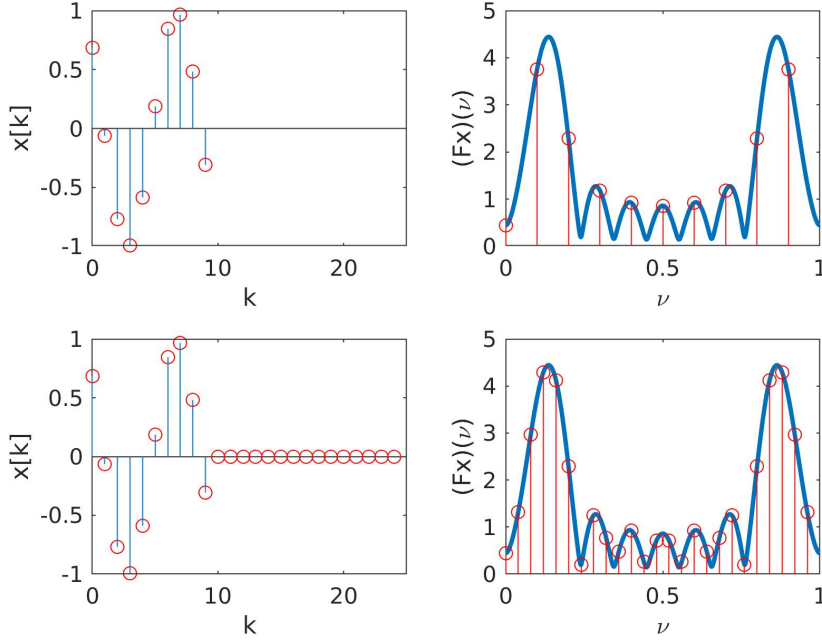


FIGURE 23 – Représentations fréquentielles d'un signal discret de longueur $N = 10$. A droite, la TFD $X(\nu)$ est représentée en trait continu. Son évaluation par TFD est en rouge. Le calcul est fait pour N points fréquentiels (figures du haut). Pour un calcul avec M points fréquentiels ($M > N$), on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

ANNEXE

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

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

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-2i\pi ft} (\alpha x(t) + \beta y(t)) dt \\ &= \alpha \int_{\mathbb{R}} e^{-2i\pi ft} x(t) dt + \beta \int_{\mathbb{R}} e^{-2i\pi ft} y(t) dt \\ &= \alpha X(f) + \beta Y(f) \end{aligned}$$

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

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-2i\pi ft} x(\alpha t) dt \\ &= \int_{\mathbb{R}} e^{-2i\pi f \frac{\theta}{\alpha}} x(\theta) \frac{1}{|\alpha|} d\theta \\ &= \frac{1}{|\alpha|} X\left(\frac{f}{\alpha}\right). \end{aligned}$$

Dualité de la TF. La transformée de Fourier inverse est donnée par

$$x(t) = \int_{\mathbb{R}} e^{2i\pi t\eta} X(\eta) d\eta.$$

Nous avons donc

$$x(-t) = \int_{\mathbb{R}} e^{-2i\pi t\eta} X(\eta) d\eta = (\mathcal{F}X)(t)$$

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

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-2i\pi ft} \int_{\mathbb{R}} x(\theta) y(t - \theta) d\theta dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-2i\pi ft} x(\theta) y(t - \theta) d\theta dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-2i\pi f(t-\theta)} e^{-2i\pi f\theta} x(\theta) y(t - \theta) d\theta dt \\ &= \int_{\mathbb{R}} e^{-2i\pi f\theta} x(\theta) \int_{\mathbb{R}} e^{-2i\pi f(t-\theta)} y(t - \theta) dt d\theta \\ &= \int_{\mathbb{R}} e^{-2i\pi f\theta} x(\theta) \int_{\mathbb{R}} e^{-2i\pi f\tau} y(\tau) d\tau d\theta \\ &= \int_{\mathbb{R}} e^{-2i\pi f\theta} x(\theta) d\theta \int_{\mathbb{R}} e^{-2i\pi f\tau} y(\tau) d\tau \\ &= X(f) Y(f). \end{aligned}$$

Produit. Soit $z = xy$.

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-2i\pi ft} x(t) y(t) dt \\ &= \int_{\mathbb{R}} e^{-2i\pi ft} \left(\int_{\mathbb{R}} e^{2i\pi t\eta} \mathcal{F}x(\eta) d\eta \right) y(t) dt \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-2i\pi(f-\eta)t} y(t) dt \right) \mathcal{F}x(\eta) d\eta \\ &= \int_{\mathbb{R}} \mathcal{F}y(f - \eta) \mathcal{F}x(\eta) d\eta \\ &= (X * Y)(f) (\mathcal{F}y * \mathcal{F}x)(f) \end{aligned}$$

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

$$\begin{aligned} Z(f) &= \int_{\mathbb{R}} e^{-2i\pi ft} x(t - \tau) dt \\ &= \int_{\mathbb{R}} e^{-2i\pi f(\theta + \tau)} x(\theta) d\theta \\ &= e^{-2i\pi f\tau} \int_{\mathbb{R}} e^{-2i\pi f\theta} x(\theta) d\theta \\ &= e^{-2i\pi f\tau} X(f). \end{aligned}$$

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 :

$$Z(f) = (\mathcal{F}\delta_{\tau})(f) X(f) = e^{-2i\pi f\tau} X(f)$$

Modulation / Translation en fréquence.

$$\begin{aligned} \mathcal{F}(e_{f_0} x)(f) &= \int_{\mathbb{R}} e^{-2i\pi ft} \left(e^{2i\pi f_0 t} x(t) \right) dt \\ &= \int_{\mathbb{R}} e^{-2i\pi(f-f_0)t} x(t) dt \\ &= X(f - f_0). \end{aligned}$$

On arrive directement au résultat en exploitant le résultat sur la TF du produit terme à terme :

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

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

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

Par récurrence on conclut que

$$\mathcal{F}\{x^{(n)}\}(f) = (2i\pi f)^n X(f).$$

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

$$\begin{aligned} X'(f) &= \frac{d}{df} \int_{\mathbb{R}} e^{-2i\pi ft} x(t) dt \\ &= (-2i\pi) \frac{d}{df} \int_{\mathbb{R}} e^{-2i\pi ft} t x(t) dt \\ &= (-2i\pi) \mathcal{F}\{tx(t)\}(f) \end{aligned}$$

Nous avons donc $\mathcal{F}\{tx(t)\}(f) = \frac{1}{(-2i\pi)} X'(f)$. La propriété

$$\mathcal{F}\{t^n x(t)\}(f) = \frac{1}{(-2i\pi)^n} X^{(n)}(f)$$

se déduit par récurrence.

ECHANTILLONNAGE ET RECONSTRUCTION

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 sa transformée de Fourier soit à support compact, et que le signal original peut-être reconstitué de façon simple à partir de ses échantillons.

4.1 ECHANTILLONNAGE

Définition 4.1 (Echantillonnage). *On appelle échantillonnage l'opération qui construit 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 $T_e > 0$,*

$$x_d[n] = x(nT_e). \quad (120)$$

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 (120) n'a pas de sens. On décide de se placer dans le cas

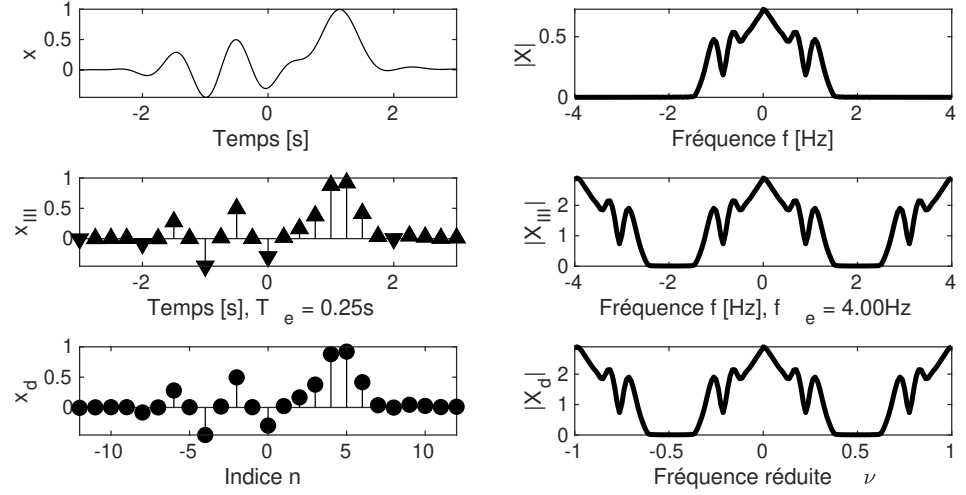


FIGURE 24 – Echantillonnage pour $f_e = 4$ Hz. Gauche : signaux x , $x_||$ et x_d . Droite : leurs spectres. $X_|| (f)$ est f_e -périodique, son motif période est $X(f)$ à un facteur multiplicatif près. $X_d(\nu)$ est 1-périodique. Les graphes de $|X_|| (f)|$ et $|X_d(\nu)|$ sont identiques à un changement d'échelle près.

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 $x_||$ par :

$$x_|| = \sum_{n \in \mathbb{Z}} x(nT_e) \delta_{nT_e} \quad (121)$$

qu'on peut réécrire comme le produit de x et du peigne de Dirac $\mathbb{I}_{T_e}$:

$$x_|| = \mathbb{I}_{T_e} x. \quad (122)$$

$x_||$ permet de relier le signal $x(t)$ à temps continu et le signal $x_d[n]$ à temps discret. En effet,

$$x_|| = \sum_{n \in \mathbb{Z}} x_d[n] \delta_{nT_e}. \quad (123)$$

La colonne de gauche de la figure 24 montre une représentation graphique de ces trois signaux.

Le théorème suivant relie les TF de x , $x_||$ et x_d .

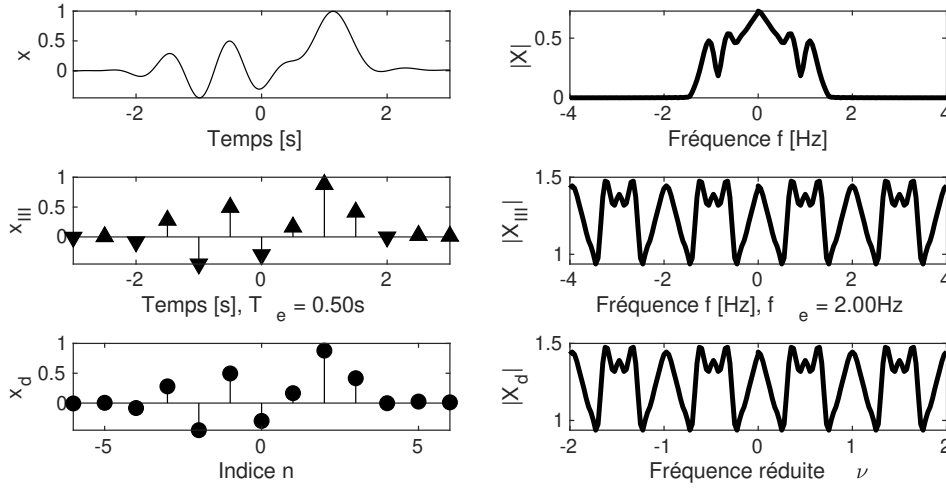


FIGURE 25 – Echantillonnage pour $f_e = 2$ Hz. Gauche : signaux x , x_{III} et x_d . Droite : leurs spectres. Dans le graphe de $|X_{III}(f)|$, les copies du motif $|X(f)|$ se superposent (repliement spectral).

Théorème 4.1. Soit $x(t)$ un signal dont la transformée de Fourier $X(f)$ est 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 :

$$X_{III}(f) = f_e \sum_{n \in \mathbb{Z}} X(f - nf_e) \quad (124)$$

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

$$X_d(v) = f_e \sum_{n \in \mathbb{Z}} X(vf_e - nf_e). \quad (125)$$

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

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

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

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 (123) :

$$X_{III}(f) = \sum_{n \in \mathbb{Z}} x_d[n] e^{-2i\pi f n T_e} = X_d(f T_e),$$

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

On retiendra que la TF d'un signal échantillonné à la fréquence f_e est f_e -périodique avec une périodisation du motif $X(f)$ tous les f_e . Inversement, la TF d'un signal périodique est un spectre de raies, cf. section 3.2.

Remarque (Cas d'une exponentielle complexe). Le théorème 4.1 ne s'applique directement que si la transformée de Fourier X est dans L^1 . On peut cependant obtenir des expressions généralisées si X est une distribution à support compact. Prenons le cas d'une exponentielle complexe $x(t) = \exp(2i\pi f_0 t)$. Alors, $X = \delta_{f_0} \notin L^1$. Comme $x_d[n] = x(nT_e) = \exp(2i\pi \frac{f_0}{f_e} n)$, on calcule (voir tableau 4) :

$$X_d = \sum_{n \in \mathbb{Z}} \delta_{\frac{f_0}{f_e} - n}. \quad (126)$$

Ce résultat s'apparente à (125) dans la mesure où la raie spectrale présente en $f = f_0$ dans le signal X induit une raie en $\frac{f_0}{f_e} - n$ pour tout $n \in \mathbb{Z}$ dans le signal X_d . On notera que le facteur multiplicatif f_e dans (125) n'apparaît plus ici.

Interprétation du théorème 4.1

L'échantillonnage temporel équivaut à périodiser le spectre du signal. Il est en général impossible de reconstituer un signal à partir de ses échantillons, car différentes copies du spectre peuvent se superposer (voir la figure 25).

Cependant, 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 original. Cette observation est formalisée dans la section suivante.

4.2 RECONSTRUCTION

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.

4.2.1 Signaux à bande limitée dans $[-B, B]$

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

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

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

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 par $x_B(t) = \text{sin}_c(2\pi Bt)$ pour $B > 0$ vérifie $x_B \in PW_B$ car $X_B = \frac{1}{2B} \mathbb{1}_{[-B, B]}$.

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 .

4.2.2 Théorème de Shannon

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 $t \in \mathbb{R}$:

$$\begin{aligned} x(t) &= \sum_{n \in \mathbb{Z}} x(nT_e) \text{sin}_c\left(\pi \frac{t - nT_e}{T_e}\right) \\ &= \sum_{n \in \mathbb{Z}} x_d[n] \text{sin}_c\left(\pi \frac{t - nT_e}{T_e}\right) \end{aligned} \quad (127)$$

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

Critère de Shannon pour un signal à bande limitée

Dans le cas d'une fonction d'énergie finie et à bande limitée 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 (ex., des 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 $[-f_e/2, f_e/2]$.

Sur l'exemple de la figure 24, $B = 1.5$ Hz et $f_e = 4$ Hz. Le critère de Shannon est respecté ($f_e \geq 2B$). Il n'est pas sur l'exemple de la figure 25, où $B = 1.5$ Hz et $f_e = 2$ Hz ($f_e < 2B$), d'où le phénomène de repliement spectral.

Interprétation du théorème de Shannon

Dans l'équation (127), $x(t)$ est interpolé à partir de ses échantillons $x_d[n]$. (127) s'interprète comme la convolution de $x_{\square}$ et de $\text{sinc}(\pi t / f_e)$. La TF de cette convolution est égale au produit de la TF de $x_{\square}$ par la fenêtre rectangulaire $\mathbb{1}_{[-B, B]}$ (TF du sinus cardinal). Dans la somme (124) et sous la condition $x \in PW_B$, on note que :

- Pour $n = 0$, le support de la fonction $f \mapsto X(f)$ est inclus dans $[-B, B]$.
- Pour $n \neq 0$, le support de $f \mapsto X(f - nf_e)$ est inclus dans $[nf_e - B, nf_e + B]$. Ce support n'intersecte pas $[-B, B]$.

La formule de reconstruction (127) peut donc s'interpréter comme un filtrage passe-bas qui élimine les repliements du spectre du signal original (induits par la discrétisation) pour les fréquences supérieures à B . Les figures 26 et 27 montrent des reconstructions de signaux dans les cas où le critère de Shannon est respecté ou non.

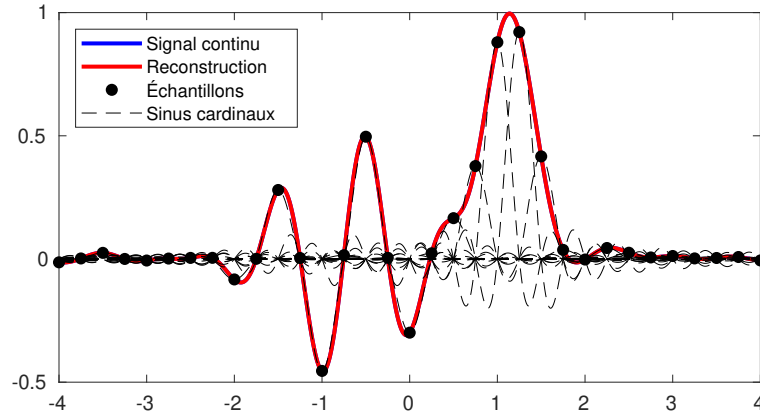


FIGURE 26 – Cas où le critère de Shannon est respecté. Reconstruction du signal par la somme (127).

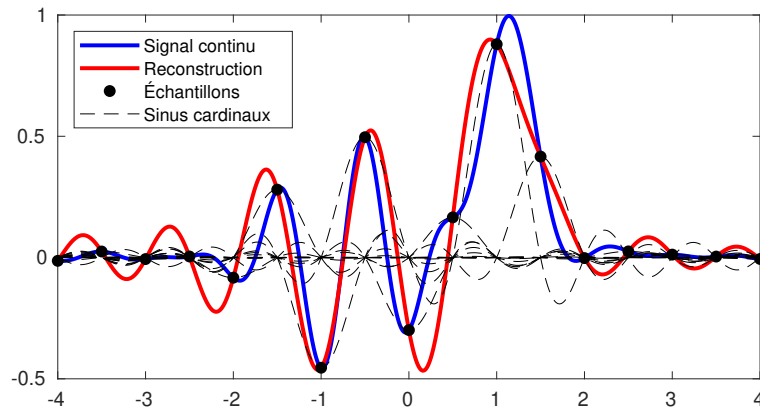


FIGURE 27 – Cas où le critère de Shannon n'est pas respecté. Reconstruction du signal par la somme (127).

4.2.3 Démonstration du théorème de Shannon

Le théorème peut être vu comme la décomposition de x dans une base orthogonale de PW_B . Sa démonstration comporte deux étapes :

1. construction d'une base orthogonale de PW_B , permettant d'assurer la convergence de (127) dans L^2 .
2. analyse de la convergence simple de la série.

Base orthogonale de PW_B

Lemme 4.1. La famille de fonctions s_n définie par

$$s_n(t) = \text{sinc}_c \left(\pi \frac{t - nT_e}{T_e} \right), \quad n \in \mathbb{Z} \quad (128)$$

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

$$x_n = x(nT_e), \quad (129)$$

et

$$x = \sum_{n \in \mathbb{Z}} x_n s_n \quad (130)$$

au sens L^2 , où $T_e = \frac{1}{2B}$.

Démonstration. On a $s_0(t) = \text{sinc}(\pi \frac{t}{T_e})$. Comme $s_n(t) = s_0(t - nT_e)$,

$$S_n(f) = S_0(f) e^{-2i\pi f n T_e}.$$

D'après (87), la TF de s_0 vaut

$$S_0(f) = \frac{1}{2B} \mathbb{1}_{[-B, B]}(f).$$

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

$$S_n(f) = \frac{1}{2B} \exp \left(-2i\pi \frac{nf}{2B} \right) \mathbb{1}_{[-B, B]}(f).$$

La transformée de Fourier étant une isométrie, les signaux obtenus par transformée de Fourier inverse des S_n forment une base de PW_B . Autrement dit, les signaux $\{s_n, n \in \mathbb{Z}\}$ forment une base de PW_B . D'après l'identité de Parseval,

$$\|s_n\|_{L^2}^2 = \|S_n\|_{L^2}^2 = \int_{\mathbb{R}} |S_n(f)|^2 df = \frac{1}{2B}.$$

De plus, les coefficients X_n de $X(f)$ dans la base $\{S_n\}$ de L^2 coïncident avec les coefficients x_n de x dans la base $\{s_n\}$ de PW_B . Ces coefficients sont donnés par :

$$X_n = \frac{(X, S_n)}{\|S_n\|^2} = \int_{[-B, B]} X(f) \exp \left(2i\pi \frac{nf}{2B} \right) df$$

où $X = \sum_{n \in \mathbb{Z}} X_n S_n$ dans L^2 . En interprétant l'intégrale comme la transformée de Fourier inverse de $X(f)$ évaluée en $t = \frac{n}{2B} = nT_e$, on trouve $x_n = X_n = x(nT_e)$. $\square$

Convergence de la série

On démontre maintenant la convergence simple de la série.

La famille s_n étant une base orthogonale, on observe que

$$\|x\|_{L^2}^2 = \sum_{n \in \mathbb{Z}} |x_n|^2 \|s_n\|_{L^2}^2 = \frac{1}{2B} \sum_{n \in \mathbb{Z}} |x_n|^2, \quad (131)$$

ce qui montre que la suite $(x_n)_{n \in \mathbb{Z}}$ est dans $\ell^2(\mathbb{Z})$. De plus, pour tout t , la suite $(s_n(t))_{n \in \mathbb{Z}}$ est de carré intégrable. En interprétant (127) comme le produit scalaire de $(x_n)_{n \in \mathbb{Z}}$ et $(s_n(t))_{n \in \mathbb{Z}}$, on 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 $n, t \mapsto x_n s_n(t)$ est continue,
- pour tout t , $|x_n s_n(t)| \leq |x_n| \max(1, (\pi(|n| - 1))^{-1})$, 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 la 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 converge simplement vers $x(t)$ (pour tout t).

Pour aller plus loin...

Pour les signaux dont la TF 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

$$\forall t \in \mathbb{R}, x(t) = \sum_{n \in \mathbb{Z}} x(nT_e) \phi(t - nT_e) \quad (132)$$

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

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

La série (132) peut être écrite comme la convolution $x_{\mathbb{W}} * \phi$. Sa TF vaut :

$$X_{\mathbb{W}}(f) \Phi(f) = f_e \sum_{n \in \mathbb{Z}} X(f - nf_e) \Phi(f). \quad (133)$$

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

De plus, la régularité de Φ est nécessaire pour garantir la convergence simple de (132). 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.

4.2.4 Signaux à bande limitée dans $[A, B]$

Le théorème d'échantillonnage (théorème 4.2) est énoncé pour des signaux à bande limitée dans $[-B, B]$. Les signaux dits « modulés »,

ayant un support inclus dans un intervalle non symétrique $[A, B]$, peuvent évidemment être considérés comme des signaux à bande limitée dans $[-C, C]$, avec $C = \max(|A|, |B|)$. On peut alors appliquer le théorème d'échantillonnage pour $f_e \geq 2C$. On montre dans ce qui suit que la borne $f_e \geq 2C$ peut être améliorée, permettant ainsi d'utiliser une fréquence d'échantillonnage inférieure à $2C$.

Critère de Shannon pour un signal modulé

Soit $x(t)$ un signal d'énergie finie et à bande limitée dans $[A, B]$. En supposant la fréquence centrale $\frac{A+B}{2}$ connue, il est possible de reconstruire $x(t)$ à partir du signal échantillonné $x_d[n]$ à une fréquence

$$f_e \geq B - A. \quad (134)$$

En effet, le signal en « bande de base » (*i.e.*, non modulé) s'écrit :

$$y(t) = x(t) e^{-2i\pi \frac{A+B}{2} t} = x(t) e^{-i\pi(A+B)t}. \quad (135)$$

En utilisant les propriétés de la transformée de Fourier (*cf.* tableau 1 pour la TF d'un signal modulé), on montre facilement que $y(t)$ est à bande limitée dans $[-\frac{B-A}{2}, \frac{B-A}{2}]$. Le théorème de Shannon appliqué à $y(t)$ garantit alors la reconstruction exacte de $y(t)$ pour $f_e \geq B - A$:

$$y(t) = \sum_{n \in \mathbb{Z}} y(nT_e) \operatorname{sinc}(\pi(t - nT_e)/T_e).$$

En utilisant (135), on déduit la formule de reconstruction de $x(t)$:

$$x(t) = \sum_{n \in \mathbb{Z}} x(nT_e) \operatorname{sinc}(\pi(t - nT_e)/T_e) e^{i\pi(A+B)(t-nT_e)}.$$

On remarque que cette formule fait intervenir $A + B$. Concrètement, il est nécessaire de connaître la valeur de la fréquence centrale $\frac{A+B}{2}$ pour pouvoir reconstruire le signal $x(t)$.

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. Même quand le spectre du signal à échantillonner est à support compact, 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.

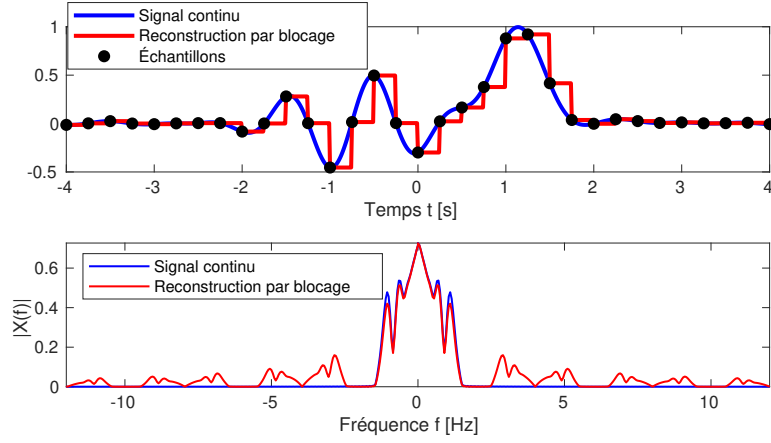


FIGURE 28 – Reconstruction par blocage d’ordre 0. Le filtrage du signal dans la bande utile et les copies du spectre atténuées par le bloqueur sont visibles sur le spectre.

4.3.2 Reconstruction non idéale

Les formules de reconstruction (127) et (132) posent plusieurs problèmes de mise en œuvre 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 blocage d’ordre zéro est une technique simple permettant d’approximer le signal à reconstruire. Le signal bloqué $\tilde{x}(t)$ est le signal en escalier défini par :

$$\tilde{x}(t) = \sum_{n \in \mathbb{Z}} x_d[n] \mathbb{1}_{[nT_e, (n+1)T_e[}(t). \quad (136)$$

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

Le bloqueur d’ordre zéro peut être analysé en écrivant :

$$\tilde{x} = \mathbb{1}_{[0, T_e[} * x_{\sqcup}. \quad (137)$$

Le spectre de $\tilde{x}$ est donc donné par

$$\tilde{X}(f) = X_{\sqcup}(f) \left(T_e e^{-i\pi f T_e} \frac{\sin \pi f T_e}{\pi f T_e} \right). \quad (138)$$

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 ($(\sin \pi f T_e)/(\pi f T_e)$ n’est pas 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 $\tilde{x}$.

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 $x_d[n]$, $n \in \mathbb{Z}$. 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-fréquence.

SIGNAUX MODÉLISÉS PAR DES PROCESSUS ALÉATOIRES

La théorie des signaux déterministes atteint ses limites pour traiter et modéliser des signaux dont les valeurs ne sont pas parfaitement connues (on ne peut alors pas décrire le signal par une fonction de $\mathbb{R}$ dans $\mathbb{R}$). Les processus aléatoires 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 qui expliciterait la valeur du bruit à chaque instant. Grâce aux processus aléatoires, il est possible de caractériser les propriétés statistiques du bruit et du signal utile afin de débruiter le signal et réduire l'influence du bruit (un tel traitement sera détaillé à la section 6.5.3).

Considérons un deuxième exemple : en traitement de la parole, deux sons peuvent sembler identiques alors que les valeurs des deux signaux associés sont très différentes. L'étude des propriétés statistiques de ces signaux permet de classifier différents sons dans le cadre de la reconnaissance vocale. À 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 à nouveau abordés dans la section 6.4.2.

5.1 DÉFINITION ET CARACTÉRISATION DES PROCESSUS ALÉATOIRES

Les processus aléatoires (à temps discret) 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 qui sont donc de dimension finie.

5.1.1 Définition

Nous nous limiterons aux processus aléatoires à temps discret.

Définition 5.1 (processus aléatoire à temps discret). *Un processus aléatoire discret noté X_n est une suite de variables 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 ω donné, on a :

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

Il convient de bien distinguer deux notions complémentaires :

- le processus aléatoire qui est une suite de variables aléatoires modélisant les signaux,
- la/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 statistiques du processus aléatoire. La spécificité du traitement du signal est que les propriétés du processus aléatoire sont exploitées pour traiter les signaux observés afin d'améliorer leur qualité (par ex. suppression de bruit, déconvolution pour supprimer le flou introduit par un appareil de mesure).

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

Exemple 5.1. Voici trois exemples de processus avec une représentation de quelques-unes de leurs réalisations, voir Fig. 29 :

1. Suite de variables i.i.d $X_n = A_n$ avec $A_n \sim \mathcal{N}(0, \sigma^2)$
2. Processus constant à valeur aléatoire : $Y_n = A$ avec $A \sim \mathcal{N}(0, \sigma^2)$
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.

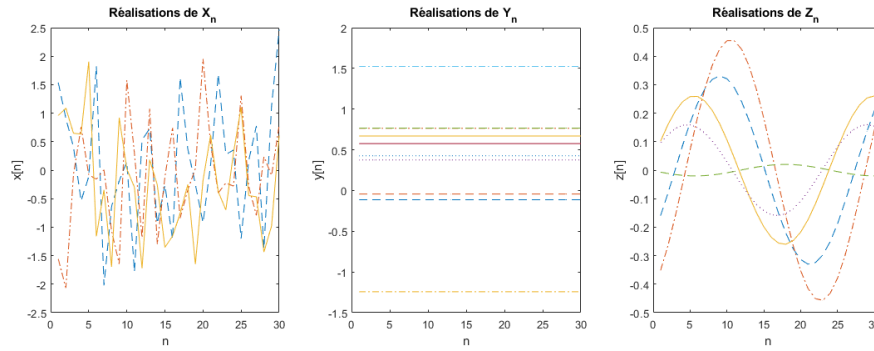


FIGURE 29 – Quelques réalisations de trois processus aléatoires.

Pour aller plus loin...

Processus aléatoire à temps continu.

La définition la plus générale d'un processus aléatoire 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 à temps 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-dimensionnelles). *Pour $N > 0$, les lois fini-dimensionnelles d'ordre N sont 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 complexe et difficile à expliciter. Elle est rarement utilisée en pratique.

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

$$\exists \mu, \sigma : \sum_{i=1}^N \alpha_i X_{n_i} \sim \mathcal{N}(\mu, \sigma^2).$$

Notons que la gaussianité des lois marginales de X_n n'implique pas que le processus est gaussien. Il existe des processus non gaussiens dont les variables aléatoires X_n sont gaussiennes pour tout n .

5.1.4 Moyenne et autocorrélation

Nous avons vu que la description complète d'un processus par ses lois fini-dimensionnelles est une tâche difficile. On se contente souvent de caractériser partiellement un processus par ses deux premiers moments : la **moyenne** et l'**autocorrélation**. Par exemple, pour caractériser différents phonèmes de parole, l'étude des deux premiers moments du signal permet 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 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

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

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

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

Remarque. Les variables aléatoires X_n appartiennent à $L^2(\Omega, \mathcal{F}, \mathcal{P})$ si $\mathbb{E}(|X_n|^2) < \infty$. L'autocorrélation $r_X[n_1, n_2] = \mathbb{E}(X_{n_1} X_{n_2}^*)$ coïncide avec le produit scalaire de X_{n_1} et X_{n_2} dans $L^2(\Omega, \mathcal{F}, \mathcal{P})$.

μ_X et r_X sont des signaux déterministes. $\mu_X[n]$ donne la valeur moyenne du processus à l'instant n , et $r_X[n_1, n_2]$ donne la corrélation du processus entre les instants n_1 et n_2 .

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 plus corrélés que deux échantillons 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. L'intercorrélation d'un processus avec lui-même correspond à son autocorrélation : $r_{XX}[n_1, n_2] = r_X[n_1, n_2]$.

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

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

ou encore :

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

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-corrélations.

5.2 PROCESSUS STATIONNAIRES

La stationnarité est une propriété clé qui permet de nombreuses simplifications calculatoires. Elle se traduit par le fait que les propriétés statistiques 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 est dit **strictement stationnaire** si ses lois fini-dimensionnelles sont invariantes par translation : pour tous $N > 0$ et $k \in \mathbb{Z}$, 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 pratique limité. C'est pourquoi on s'intéresse uniquement à la stationnarité au sens large.

5.2.2 Stationnarité au sens large

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

— *la moyenne ne dépend pas du temps :*

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

— *il est de puissance finie : $\forall n, \mathbb{E}(|X_n|^2) < \infty$*

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

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

Notons que si X_n est SSL, $\mathbb{E}(|X_n|^2) = r_X[n, n]$ est constant, et $r_X[n + k, n] = \mathbb{E}(X_{n+k}X_n^*) = \mathbb{E}(X_nX_{n-k}^*)$. On utilisera la notation simplifiée

$$\gamma_X[k] := r_X[n + k, n] = \mathbb{E}(X_{n+k}X_n^*) \quad (139)$$

pour l'autocorrélation. De même, l'intercorrrelation entre deux processus conjointement stationnaires sera notée :

$$\gamma_{XY}[k] := \mathbb{E}(X_{n+k}Y_n^*). \quad (140)$$

En résumé, il convient de distinguer la fonction d'autocorrélation à deux variables $r_X[n_1, n_2]$ (définie pour tous les processus) et la fonction d'autocorrélation $\gamma_X[k]$, définie uniquement pour les processus stationnaires. Dans ce qui suit, tous les processus sont supposés stationnaires au sens large. Cela implique que leurs moments d'ordre deux sont finis. On utilisera 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 (Puissance). La **puissance** d'un processus stationnaire X_n est la valeur de l'autocorrélation en zéro :

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

Pour un processus centré, la puissance est égale à la variance.

Définition 5.13 (Rapport signal-sur-bruit). Le **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 des puissances de U_n et W_n . Il est le plus souvent exprimé en dB :

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left(\frac{\mathbb{E}(|U_n|^2)}{\mathbb{E}(|W_n|^2)} \right). \quad (141)$$

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

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

Démonstration. Le résultat découle de 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 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. $\gamma_X[-k] = \mathbb{E}(X_{n-k}X_n^*) = \mathbb{E}(X_{n-k}^*X_n)^* = \gamma_X^*[k]$. □

5.2.3 Densité spectrale de puissance

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 on souhaite quantifier la bande spectrale utilisée par ces signaux pour optimiser les canaux de transmissions. L'outil principal pour caractériser les processus aléatoires stationnaires dans le domaine fréquentiel 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 :

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

Théorème 5.1 (Wiener-Khinchin). La densité spectrale de puissance 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 :

$$\Gamma_X(\nu) = \lim_{N \rightarrow \infty} \mathbb{E} \left(\frac{1}{2N+1} \left| \sum_{k=-N}^N X_k \exp(-i2\pi k \nu) \right|^2 \right). \quad (142)$$

On écrit :

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

Si $\gamma_X \in L^1$, 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.

Si $\gamma_X \notin L^1$, il faut calculer la limite au sens des distributions pour montrer que

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

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

En pratique, on utilise plus fréquemment l'expression de la DSP donnée par le théorème de Wiener-Khinchin car il est 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 :

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

Ces propriétés se déduisent directement de la définition.

Remarque. Si le processus X_n est réel, alors sa DSP $\Gamma_X(\nu)$ est paire :

$$\forall \nu, \Gamma_X(\nu) = \Gamma_X(-\nu).$$

En effet, si X_n est réel, alors $\gamma_X[k]$ est réel. Comme γ_X est à symétrie hermitienne, on a $\gamma_X[-k] = \gamma_X[k]$ pour tout $k \in \mathbb{N}$. D'après les propriétés de la transformée de Fourier, $\Gamma_X(\nu)$ est paire.

Propriété. La puissance d'un processus stationnaire est égale à 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. $\square$

Ce résultat justifie la dénomination de densité spectrale de puissance : en intégrant la densité dans le domaine spectral, on obtient la puissance du processus. En intégrant la DSP dans une bande fréquentielle restreinte, on obtient la puissance du processus dans cette bande de fréquence.

Méthodologie de calcul

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. Calculer 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 différentes sources d'erreur :

- les erreurs de mesure liées à l'imperfection des capteurs (bruit de mesure);
- l'inexactitude du modèle mathématique utilisé pour représenter les données (bruit de modèle), lorsque ce modèle est trop approximatif.

Le bruit blanc modélise alors la partie nuisible (non informative) contenue dans les données. Dans certaines situations, il permet au contraire de décrire la partie utile (informative) des données. Lorsque le bruit blanc est placé en entrée d'un filtre, le signal aléatoire obtenu en sortie décrit la diversité des signaux observables en pratique.

Définition 5.15 (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 centré et stationnaire au sens large,
- $\gamma_W[n] = \sigma^2 \delta[n]$,
- W_n et W_k sont indépendants si $k \neq n$.

Définition 5.16 (bruit blanc au sens large). *Un processus W_n est un **bruit blanc au sens large** (ou au sens faible) de variance σ^2 s'il vérifie les propriétés suivantes :*

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

Remarque.

- Un bruit blanc étant centré, sa puissance est égale à sa variance : $P_W = \sigma^2$.
- Un bruit blanc gaussien est un bruit blanc au sens strict car la décorrélation implique l'indépendance pour les processus gaussiens.

Propriété. Si W_n est un bruit blanc au sens large, alors :

- W_n et W_m sont **décorrélés** pour $n \neq m$.
- Sa **DSP 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.

Convention. Dans toute la suite du polycopié, la dénomination courte « bruit blanc » fait référence à un bruit blanc au sens large.

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 une réalisation de chacun, leurs autocorrélations et leurs DSP.

- **Réalisations** : Du point de vue empirique, en observant une réalisation de chaque processus, on remarque qu'à gauche, pour le bruit blanc, les valeurs successives semblent décorrélées entre elles ; tan-

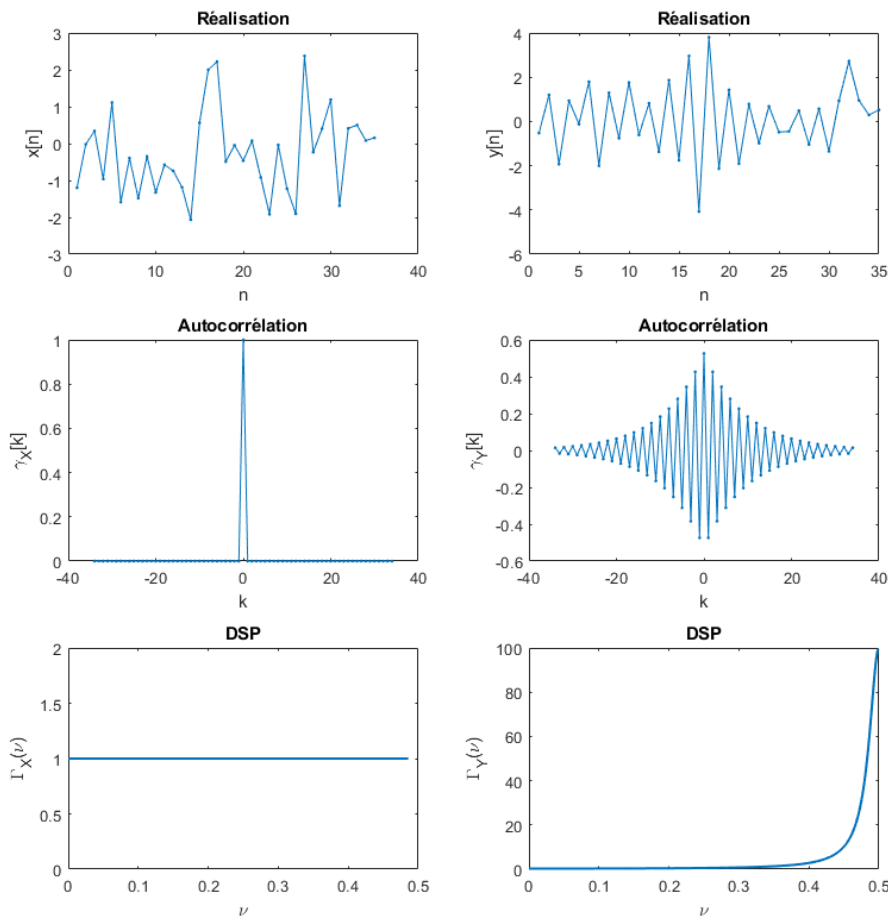


FIGURE 30 – Comparaison des caractéristiques d'un bruit blanc (gauche) et d'un bruit coloré (droite).

dis qu'à droite, on observe 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. Celle du bruit coloré comporte principalement des hautes fréquences ce qui est cohérent avec les variations rapides observées sur la réalisation.

5.2.5 Filtrage linéaire de processus

Le filtrage linéaire convolutif 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

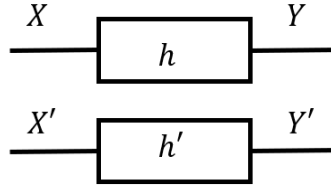


FIGURE 31 – Filtrage linéaire de deux processus aléatoires.

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 :

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

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

Démonstration.

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

□

Ce théorème (appelé **formule des interférences**) est très important car il permet d'exprimer les moments d'ordre deux 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.

Autocorrélation et DSP d'un processus filtré

Corollaire 5.1. *Soit*

- X_n un processus stationnaire au sens large,
- h la réponse impulsionnelle (ℓ^1) 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. On applique la formule des interférences avec : $X' = X$ et $h' = h$, puis on utilise la propriété de la transformé de Fourier : $\mathcal{F}\bar{h}(\nu) = H^*(\nu)$. □

Ce corollaire important permet d'expliciter la DSP (et l'autocorrélation) du processus filtré en fonction de celle du processus d'entrée.

Identification de la réponse d'un filtre linéaire

L'identification de la réponse impulsionnelle d'un filtre linéaire inconnu est un problème classique. Une méthode consiste à introduire un bruit blanc en entrée du filtre, et à calculer l'intercorrélacion entre ce bruit blanc et le signal mesuré en sortie du filtre.

Corollaire 5.2. Soit

- W_n un bruit blanc,
- h la réponse impulsionnelle (ℓ^1) d'un filtre stable.

L'intercorrélation entre le bruit en entrée W_n et le processus filtré $Y_n = h * W_n$ est proportionnelle à la réponse impulsionnelle du filtre :

$$\gamma_{YW}[k] = \sigma^2 h[k].$$

Démonstration. On applique la formule des interférences avec les entrées $X_n := W_n$ et $X'_n := W_n$, et les filtres $h[n]$ et $h'[n] := \delta[n]$.

Les signaux en sortie des filtres s'écrivent $Y_n = (h * W)_n$ et $W_n = (h' * W)_n$.

Finalement, $\gamma_{YW} = h * \bar{h}' * \gamma_{WW} = h * \delta * \gamma_W = \sigma^2 h$. $\square$

5.2.6 Processus autorégressif

Les processus autorégressifs (AR) sont des modèles couramment utilisés pour représenter des séries chronologiques aussi variées que des mesures de pollution, des valeurs financières ou des phonèmes de parole. Ils permettent de prédire des données futures à partir d'observations passées. 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.

Définition

Définition 5.17. (Processus autorégressif) Un processus autorégressif (AR) d'ordre p est un processus obtenu par filtrage d'un bruit blanc avec un filtre stable de fonction de transfert :

$$H(v) = \frac{1}{1 - \sum_{k=1}^p a_k e^{-i2\pi kv}}, \quad \text{où } a_1, \dots, a_p \in \mathbb{R}.$$

Relation temporelle : équation aux différences

D'après les résultats de la section 3.5.1, on peut définir un processus AR dans le domaine temporel à l'aide d'une équation aux différences.

Théorème 5.3. Un processus autorégressif Y_n d'ordre p s'écrit :

$$Y_n = \sum_{k=1}^p a_k Y_{n-k} + W_n \quad (143)$$

où W_n est un bruit blanc.

Remarque. La réponse impulsionnelle du filtre est de durée infinie. Pour exprimer la relation entrée-sortie, il est 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 :

- la variance σ^2 du bruit blanc générateur W_n ,
- les 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 du corollaire 5.1 :

$$\Gamma_Y(\nu) = \sigma^2 \left| \frac{1}{1 - \sum_{k=1}^p a_k e^{-i2\pi k\nu}} \right|^2. \quad (144)$$

Autocorrélation d'un modèle AR

Propriété. L'autocorrélation d'un processus AR satisfait :

$$\forall k \geq 0, \quad \gamma_Y[k] = \sum_{i=1}^p a_i \gamma_Y[k-i] + \sigma^2 \delta[k] \quad (145)$$

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 (143) et du fait que W_n est un bruit blanc ($\mathbb{E}(W_i W_j) = \sigma^2 \delta[i-j]$). En effet, d'après (143), Y_{n-k} ne dépend que de W_i aux instants $i \leq n-k$. On obtient :

- pour $k > 0$, $\mathbb{E}(W_n Y_{n-k}) = 0$;
- pour $k = 0$,

$$\begin{aligned} \mathbb{E}(W_n Y_{n-k}) &= \mathbb{E}(W_n Y_n) \\ &= \mathbb{E} \left(W_n \left(\sum_{i=1}^p a_i Y_{n-i} + W_n \right) \right) \\ &= \sum_{i=1}^p a_i \mathbb{E}(W_n Y_{n-i}) + \mathbb{E}(W_n^2) \\ &= 0 + \sigma^2. \end{aligned}$$

□

Les équations (145) obtenues pour $0 \leq k \leq p$ sont appelées équations de **Yule-Walker**. A partir de ces $p + 1$ équations linéaires, on peut estimer les $p + 1$ paramètres du modèle, $\{a_1, a_2, \dots, a_p, \sigma^2\}$. On reviendra sur ce sujet au chapitre 6.

5.3 CONCLUSION

On retiendra en premier lieu la définition d'un processus à temps discret stationnaire et sa caractérisation par ses deux premiers moments (moyenne et fonction d'autocorrélation) dans le domaine temporel, et par sa densité spectrale de puissance dans le domaine fréquentiel.

Un autre concept central est celui de bruit blanc. On retiendra que la fonction d'autocorrélation d'un bruit blanc est impulsionnelle, et que sa DSP est constante quelles que soient les fréquences.

Enfin, on retiendra que le filtrage d'un bruit blanc génère un processus stationnaire (par exemple, un modèle AR). La formule des interférences est un outil important pour caractériser le processus généré dans les domaines temporel et fréquentiel.

ESTIMATION BASÉE SUR DES MODÈLES ALÉATOIRES

Au chapitre précédent, nous avons défini une panoplie d'outils mathématiques pour modéliser les signaux à l'aide de processus aléatoires. En pratique, on dispose d'une ou plusieurs réalisations d'un processus, à partir desquelles on souhaite estimer les paramètres du processus : moyenne, autocorrélation, DSP ou autres paramètres du modèle. L'objet du chapitre étant de définir des estimateurs, ce chapitre fait appel à des concepts communs avec le cours de statistiques.

A titre d'exemple, considérons le traitement de la parole, pour lequel on souhaite identifier la nature des sons élémentaires (appelés phonèmes) afin de réaliser la reconnaissance automatique ou la compression de signaux audio. Certains phonèmes comme le /s/ ou le /f/ peuvent être modélisés par un processus aléatoire autorégressif (AR). En pratique, on dispose de l'enregistrement d'un signal audio qui correspond à une seule réalisation. L'objectif est d'estimer les paramètres du modèle AR à partir de cette réalisation afin d'identifier le type de phonème.

Dans l'ensemble de ce chapitre, **on ne considère que les processus à temps discret, stationnaires au sens large**. Précisons les notations que nous serons amenés à manipuler :

- un processus aléatoire est noté d'une lettre majuscule avec en indice la variable d'évolution : X_n , $n \in \mathbb{Z}$.
- le signal déterministe correspondant à une réalisation de ce processus est noté avec une lettre minuscule : $x[n]$, $n \in \mathbb{Z}$.
- l'estimateur (ou la statistique) est noté avec un chapeau circonflexe : $\hat{\mu}$. Notons qu'il s'agit d'une variable (ou d'un processus) aléatoire.

Remarque. $x[k]$ est appelé le k -ème échantillon (au sens de l'échantillonnage temporel, cf. chapitre 4) du signal x , c'est-à-dire la valeur du signal x mesurée à l'instant k . De même, X_k sera appelé le k -ème échantillon du processus aléatoire X_n ($n \in \mathbb{Z}$). C'est la variable aléatoire X_n pour $n = k$. Notons qu'en statistique, un « échantillon » d'une variable aléatoire Y fait référence à une réalisation $Y(\omega)$ pour un ω donné. Ce terme a donc un sens complètement différent de l'échantillonnage temporel. Pour évi-

ter toute confusion, par échantillonnage, nous ferons toujours référence à l'échantillonnage temporel.

6.1 ESTIMATEUR DE LA MOYENNE D'UN PROCESSUS ALÉATOIRE

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. 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. Evidemment, cette approche n'a de sens que pour les signaux stationnaires pour lesquels les grandeurs à estimer ne dépendent pas du temps.

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.*

Démonstration. Le biais est défini par $\text{biais}(\hat{\mu}_X^N) = \mathbb{E}(\hat{\mu}_X^N) - \mu_X$. On montre aisément que $\mathbb{E}(\hat{\mu}_X^N) = \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}(X_n) = \mu_X$. Donc $\text{biais}(\hat{\mu}_X^N) = 0$. $\square$

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 les réalisations du processus par une moyenne temporelle sur les valeurs d'une unique réalisation.

Exemple 6.1.

— un bruit blanc est ergodique à l'ordre 1. En effet,

$$\begin{aligned}\text{Var}(\hat{\mu}_X^N) &= \frac{1}{N^2} \sum_{n'=0}^{N-1} \sum_{n=0}^{N-1} \mathbb{E}(X_n X_{n'}^*) \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} \mathbb{E}(|X_n|^2) \quad \text{car} \quad \mathbb{E}(X_n X_{n'}^*) = 0, \forall n \neq n' \\ &= \frac{\gamma_X[0]}{N} \xrightarrow{N \rightarrow \infty} 0.\end{aligned}$$

— le processus à valeur aléatoire constante défini par $\forall n, X_n = A$ avec $A \sim \mathcal{N}(0, \sigma^2)$ n'est pas ergodique. En effet,

$$\begin{aligned}\text{Var}(\hat{\mu}_X^N) &= \frac{1}{N^2} \sum_{n'=0}^{N-1} \sum_{n=0}^{N-1} \mathbb{E}(X_n X_{n'}^*) \\ &= \frac{1}{N^2} \sum_{n'=0}^{N-1} \sum_{n=0}^{N-1} \mathbb{E}(A^2) \\ &= \sigma^2\end{aligned}$$

ne converge pas vers 0.

6.2 ESTIMATION DE LA FONCTION D'AUTOCORRÉLATION

Définition 6.3 (Estimateur non biaisé de l'autocorrélation).

L'estimateur temporel non biaisé de l'autocorrélation est défini par :

$$\begin{aligned}\hat{\gamma}_X^{N,nb}[k] &= \frac{1}{N-k} \sum_{\ell=0}^{N-k-1} X_{\ell+k} X_{\ell}^* \quad \text{pour} \quad 0 \leq k < N, \\ \hat{\gamma}_X^{N,nb}[k] &= \left(\hat{\gamma}_X^{N,nb}[-k] \right)^* \quad \text{pour} \quad -N < k < 0.\end{aligned}$$

Propriété. Cet estimateur est non biaisé pour tout k tel que $|k| < N$.

Démonstration. On vérifie que $\mathbb{E}(\hat{\gamma}_X^{N,nb}[k]) = \gamma_X[k]$ pour $0 \leq k < N$. Puis on étend ce résultat aux valeurs négatives de k en exploitant que les fonctions γ_X et $\gamma_X^{N,nb}$ sont à symétrie hermitienne. $\square$

$\hat{\gamma}_X^{N,nb}[k]$ fournit une estimation de la fonction d'autocorrélation uniquement pour $|k| < N$. Il est rarement utilisé car sa variance est très élevée si k est proche de N . Dans le cas extrême $|k| = N - 1$, la moyenne arithmétique qui intervient dans la définition de $\hat{\gamma}_X^{N,nb}[k]$ ne comporte qu'un terme (ce qui implique une variance élevée). Au contraire, pour $k = 0$, la moyenne arithmétique comporte N termes. Ainsi, la variance de $\hat{\gamma}_X^{N,nb}[0]$ est raisonnablement faible si N est grand, voir les rappels de statistique 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 :*

$$\begin{aligned}\hat{\gamma}_X^N[k] &= \frac{1}{N} \sum_{\ell=0}^{N-k-1} X_{\ell+k} X_{\ell}^* \quad \text{pour } 0 \leq k < N, \\ \hat{\gamma}_X^N[k] &= \left(\hat{\gamma}_X^N[-k] \right)^* \quad \text{pour } -N < k < 0.\end{aligned}$$

La différence avec l'estimateur non biaisé est le facteur de normalisation, égal à $\frac{1}{N}$ au lieu de $\frac{1}{N-|k|}$. Cet estimateur est biaisé car :

$$\mathbb{E}(\hat{\gamma}_X^N[k]) = \mathbb{E} \left(\frac{N-|k|}{N} \hat{\gamma}_X^{N,nb}[k] \right) = \frac{N-|k|}{N} \gamma_X[k]. \quad (146)$$

Le biais vaut :

$$\begin{aligned}\text{biais}(\hat{\gamma}_X^N[k]) &= \mathbb{E}(\hat{\gamma}_X^N[k]) - \gamma_X[k] \\ &= \frac{N-|k|}{N} \gamma_X[k] - \gamma_X[k] \\ &= -\frac{|k|}{N} \gamma_X[k].\end{aligned}$$

Il est nul pour $k = 0$ et devient important pour $|k|$ proche de N . La variance de cet estimateur est plus faible que celle de l'estimateur non biaisé :

$$\text{Var}(\hat{\gamma}_X^N[k]) = \left(\frac{N-|k|}{N} \right)^2 \text{Var}(\gamma_X^{N,nb}[k]). \quad (147)$$

Elle est proche de 0 pour k non nul et N suffisamment grand.

En statistique, l'erreur quadratique moyenne (EQM) est couramment utilisée pour comparer des estimateurs. Elle s'écrit comme la somme du biais au carré et de la variance. L'EQM des estimateurs non biaisé et biaisé de la fonction d'autocorrélation s'écrit :

$$EQM(\hat{\gamma}_X^{N,nb}[k]) = 0 + \text{Var}(\hat{\gamma}_X^{N,nb}[k]) \quad (148)$$

et

$$EQM(\hat{\gamma}_X^N[k]) = \left(\text{biais}(\hat{\gamma}_X^N[k]) \right)^2 + \text{Var}(\hat{\gamma}_X^N[k]) \quad (149)$$

$$= \frac{k^2}{N^2} \left(\gamma_X[k] \right)^2 + \text{Var}(\hat{\gamma}_X^N[k]). \quad (150)$$

L'estimateur biaisé est souvent privilégié car son EQM est plus faible que celle de l'estimateur non biaisé. En effet, la plupart des processus modélisant des phénomènes physiques ont une fonction d'autocorrélation telle que $|\gamma_X[k]|$ décroît rapidement en fonction de $|k|$. Cela correspond au fait que deux échantillons pris à des instants proches sont plus corrélés que deux échantillons très éloignés. Dans ce cas, le terme de biais au carré qui intervient dans (150) est faible :

- pour $|k|$ faible, $\frac{k^2}{N^2}$ est proche de 0;
 - Pour des valeurs modérées ou grandes de $|k|$, $(\gamma_X[k])^2$ est faible.
- Comme la variance de l'estimateur biaisé est sensiblement plus faible que celle de l'estimateur non biaisé, ceci explique que l'EQM de l'estimateur biaisé est la plus faible.

Pour aller plus loin...

Soit $X^N = (X_0, X_1, \dots, X_{N-1})$ le processus aléatoire tronqué sur N échantillons temporels. C'est un vecteur aléatoire qui contient les variables aléatoires utilisées par les estimateurs définis ci-dessus.

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 $\forall k \in \{0, \dots, N\}$, $\bar{X}_{-k}^N := (X_k^N)^*$.

Démonstration. Pour $0 \leq k < N$,

$$\begin{aligned} \hat{\gamma}_X^N[k] &= \frac{1}{N} \sum_{\ell=0}^{N-k-1} X_{\ell+k} X_{\ell}^* \\ &= \frac{1}{N} \sum_{m=k}^{N-1} X_m X_{m-k}^* \quad \text{avec } m = \ell + k \\ &= \frac{1}{N} \sum_{m=k}^{N-1} X_m^N \bar{X}_{k-m}^N \\ &= \frac{1}{N} (X^N * \bar{X}^N)[k]. \end{aligned}$$

De plus, la fonction $\hat{\gamma}_X^N[k]$ est à symétrie hermitienne. En exploitant la définition de $\bar{X}^N[k]$ pour $k < 0$, on vérifie facilement que $X^N * \bar{X}^N$ est aussi à symétrie hermitienne. Ainsi, si $-N < k < 0$,

$$\begin{aligned} \hat{\gamma}_X^N[k] &= \left(\hat{\gamma}_X^N[-k] \right)^* \\ &= \frac{1}{N} \left(X^N * \bar{X}^N[-k] \right)^* \\ &= \frac{1}{N} (X^N * \bar{X}^N)[k]. \end{aligned}$$

□

Définition 6.5 (Ergodicité d'ordre 2). Le processus stationnaire X_n est dit ergodique d'ordre 2 si l'estimateur non biaisé de sa fonction d'autocorrélation converge en moyenne quadratique vers la fonction d'autocorrélation du processus :

$$\forall k \in \mathbb{N}, \quad \hat{\gamma}_X^{N,nb}[k] \xrightarrow[N \rightarrow \infty]{L^2} \gamma_X[k]$$

ou de manière équivalente, si

$$\forall k \in \mathbb{N}, \quad \text{Var}(\hat{\gamma}_X^{N,nb}[k]) \xrightarrow[N \rightarrow \infty]{} 0.$$

L'ergodicité d'ordre 2 est une propriété indispensable pour estimer la fonction d'autocorrélation à l'aide des estimateurs temporels.

6.3 ESTIMATION DE LA DSP : PÉRIODOGRAMME

L'estimation de la DSP d'un processus à partir de données se réduisant à une unique réalisation du processus est un problème classique. Le but est de caractériser un phénomène dans le domaine fréquentiel.

Exemple 6.2. *En climatologie, une question consiste à retrouver la périodicité des variations climatiques liées aux oscillations de l'axe de la terre. On cherche notamment à distinguer les variations climatiques liées à l'activité humaine par rapport aux autres facteurs.*

L'une des mesures les plus utilisées est la concentration en deutérium dans les carottes glaciaires datant de plusieurs milliers d'années. On modélise le signal de concentration en deutérium en fonction du temps par un processus aléatoire. Evidemment, on ne dispose que d'une seule réalisation de ce processus. En estimant la DSP du processus, on peut retrouver les principales périodicités qui correspondent aux pics de la DSP.

6.3.1 Définition

L'estimateur classique de la DSP est appelé périodogramme.

Définition 6.6 (Périodogramme). *Le périodogramme sur N échantillons temporels est un estimateur de la DSP défini comme la transformée de Fourier de l'estimateur biaisé de l'autocorrélation :*

$$\hat{f}_X^N(\nu) = \sum_{k=-N+1}^{N-1} \hat{\gamma}_X^N[k] e^{-i2\pi k\nu}$$

ou de manière équivalente :

$$\hat{f}_X^N(\nu) = \frac{1}{N} \left| \sum_{k=0}^{N-1} X_k e^{-i2\pi k\nu} \right|^2.$$

Démonstration. Notons $X^N = (X_0, X_1, \dots, X_{N-1})$ la troncature du processus aléatoire X_n sur N échantillons temporels. L'équivalence entre les deux définitions s'obtient en utilisant $\hat{\gamma}_X^N = \frac{1}{N} X^N * \bar{X}^N$:

$$\begin{aligned} \hat{f}_X^N(\nu) &= \mathcal{F}\gamma_X^N(\nu) \\ &= \frac{1}{N} \mathcal{F}\{X^N * \bar{X}^N\}(\nu) \\ &= \frac{1}{N} \mathcal{F}X^N(\nu) \mathcal{F}\bar{X}^N(\nu) \\ &= \frac{1}{N} \left(\mathcal{F}X^N(\nu) \right) \left(\mathcal{F}X^N(\nu) \right)^* \\ &= \frac{1}{N} \left| \mathcal{F}X^N(\nu) \right|^2. \end{aligned}$$

□

Rappelons la définition de la densité spectrale de puissance donnée à la section 5.2.3 :

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

On remarque que l'espérance et le passage à la limite ont été supprimés dans la définition du périodogramme. Ces deux simplifications ont des conséquences importantes :

- **Résolution spectrale limitée (biais)** : Supprimer le passage à la limite revient à réaliser une troncature sur N échantillons. Le périodogramme est donc un estimateur biaisé. Ce biais est asymptotiquement nul quand N devient grand.
- **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. On peut montrer que ce n'est pas un estimateur consistant, car sa variance ne tend pas vers 0 quand $N \rightarrow \infty$.

Pour calculer le périodogramme à partir d'une réalisation $x[n]$ du processus sur N échantillons temporels, on procède en deux temps. On réalise la TFD de $x[n]$ (avec en général un bourrage de zéros pour affiner le tracé, voir section 3.6), puis on calcule le module au carré de la TFD.

Exemple 6.3 (Variance du périodogramme). *On considère un bruit blanc filtré dont la DSP théorique est tracée en traits gras sur la figure 32. Trois réalisations du périodogramme sont aussi représentées, obtenues en filtrant trois réalisations différentes du bruit blanc. La longueur du signal utilisé est de $N = 40$ échantillons sur la figure de gauche, et de $N = 400$ sur celle de droite. On remarque que :*

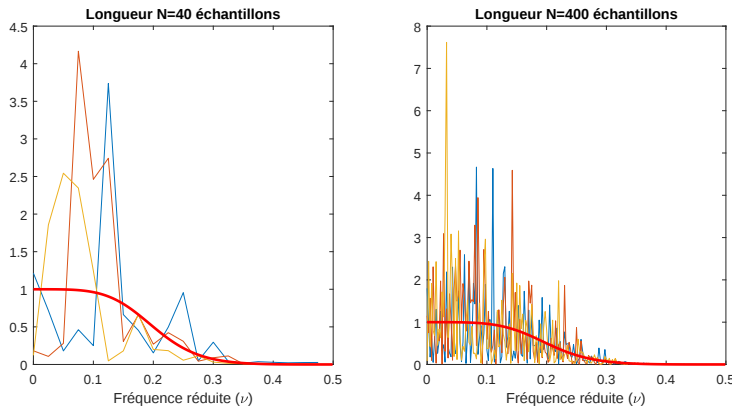


FIGURE 32 – Variance du périodogramme : comparaison de la DSP théorique (trait gras) avec des réalisations du périodogramme pour $N = 40$ et $N = 400$ échantillons.

- l'estimation de la DSP par 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.

6.3.2 Propriétés statistiques du périodogramme

Biais et résolution spectrale

Propriété. Le périodogramme est **biaisé** :

$$\mathbb{E}(\hat{\Gamma}_X^N(\nu)) = (\Gamma_X \circledast T_N)(\nu)$$

où $T_N(\nu) = \frac{1}{N} \frac{\sin^2(\pi N \nu)}{\sin^2(\pi \nu)}$ est la transformée de Fourier de la fenêtre triangulaire $t_N[k] = \frac{N-|k|}{N} \mathbb{1}_{|k| \leq N}$.

Démonstration. Par définition du périodogramme,

$$\hat{\Gamma}_X^N(\nu) = \sum_{k=-N+1}^{N-1} \hat{\gamma}_X^N[k] e^{-i2\pi k \nu}.$$

D'après (146), on déduit que :

$$\begin{aligned} \mathbb{E}(\hat{\Gamma}_X^N(\nu)) &= \sum_{k=-N+1}^{N-1} \mathbb{E}(\hat{\gamma}_X^N[k]) e^{-i2\pi k \nu} \\ &= \sum_{k=-N+1}^{N-1} \frac{N-|k|}{N} \gamma_X[k] e^{-i2\pi k \nu} \\ &= \sum_{k \in \mathbb{Z}} t_N[k] \gamma_X[k] e^{-i2\pi k \nu}. \end{aligned}$$

$\mathbb{E}(\hat{\Gamma}_X^N(\nu))$ s'écrit donc comme la TFTD du produit des fonctions t_N et γ_X , égale au produit de convolution circulaire de leurs TFTD d'après la section 3.3.2. $\square$

On retiendra que :

- le biais est lié à la troncature sur N points,
- le biais limite la résolution spectrale à $\frac{2}{N}$ (pour rappel, la résolution spectrale est définie comme la largeur du lobe principal, voir section 3.6.2. Elle vaut $\frac{2}{N}$ car T_N s'annule pour $\nu = \pm \frac{1}{N}$),
- 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 .

Pour aller plus loin...

Loi asymptotique de l'estimateur

Propriété. Soit X_n un processus gaussien dont la DSP est non nulle : $\forall \nu \in [-\frac{1}{2}, \frac{1}{2}], \Gamma_X(\nu) \neq 0$. Dans le régime asymptotique ($N \rightarrow \infty$), le périodogramme suit une loi exponentielle de paramètre $\frac{1}{\Gamma_X(\nu)}$:

$$p_{\hat{\Gamma}_X^\infty(\nu)}(x) = \frac{1}{\Gamma_X(\nu)} e^{-\frac{x}{\Gamma_X(\nu)}} \mathbf{1}_{x \geq 0}.$$

La moyenne et la variance asymptotique du périodogramme valent :

$$\mu_{\hat{\Gamma}_X^\infty(\nu)} = \Gamma_X(\nu), \quad \text{Var}(\hat{\Gamma}_X^\infty(\nu)) = \Gamma_X(\nu)^2.$$

Comme l'écart-type de l'estimateur est égal à la valeur que l'on cherche à estimer, on conclut que l'estimateur du périodogramme est très imprécis.

6.3.3 Variantes de la méthode du périodogramme

Nous avons vu que le périodogramme est biaisé et possède une variance élevée. Il existe différentes techniques pour limiter l'un ou l'autre de ces effets. Dans ce qui suit, on présente le périodogramme de Bartlett dont la variance est sensiblement réduite. Son principe est de diviser le signal en K fenêtres, et à moyenner les périodogrammes obtenus sur chacune de ces fenêtres, voir figure 33.

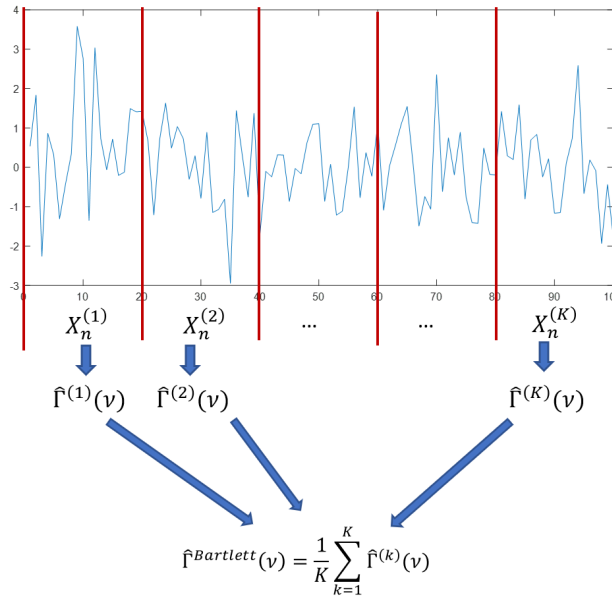


FIGURE 33 – Principe du périodogramme de Bartlett.

Définition 6.7. Soit une troncature $X_0, X_1, \dots, X_{N-1}$ de N échantillons du processus X_n , avec $N = KL$ ^(a). Le périodogramme de Bartlett sur K fenêtres de longueur L est défini par :

$$\hat{\Gamma}_X^{Bart}(\nu) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{\Gamma}^{(k)}(\nu) \quad \text{avec} \quad \hat{\Gamma}^{(k)}(\nu) = \frac{1}{L} \left| \sum_{\ell=0}^{L-1} X_{\ell+kL} e^{-i2\pi\nu\ell} \right|^2.$$

a. sans perte de généralité, on suppose que l'instant initial est $n = 0$.

Calculons le biais de cet estimateur. Le processus X_n étant stationnaire, $\mathbb{E}(\hat{\Gamma}^{(k)}(\nu))$ ne dépend pas de k . Donc $\forall k, \forall \nu, \mathbb{E}(\hat{\Gamma}^{(k)}(\nu)) = \mathbb{E}(\hat{\Gamma}^{(0)}(\nu))$. Le biais s'écrit :

$$\begin{aligned} \text{biais}(\hat{\Gamma}_X^{Bart}(\nu)) &= \mathbb{E}(\hat{\Gamma}_X^{Bart}(\nu)) - \Gamma_X(\nu) \\ &= \mathbb{E}(\hat{\Gamma}^{(0)}(\nu)) - \Gamma_X(\nu) \\ &= \text{biais}(\hat{\Gamma}_X^L(\nu)) \end{aligned}$$

et coïncide avec le biais du périodogramme (non fenêtré) sur L échantillons. D'après la section 6.3.2, **la résolution spectrale est de $\frac{2}{L}$. Elle est dégradée par rapport à celle du périodogramme calculé sur N échantillons ($\frac{2}{N}$).** Cette dernière valeur est K fois plus faible : $\frac{2}{N} = \frac{1}{K} \frac{2}{L}$. Ce résultat illustre que l'effet de troncature est amplifié dans la méthode de Bartlett, car on calcule des périodogrammes sur des fenêtres réduites de taille $L = \frac{N}{K}$.

En revanche, **moyenner K périodogrammes permet en général de réduire sensiblement la variance de l'estimateur de la DSP.** Pour fixer les idées, prenons le cas d'un bruit blanc au sens strict. Pour tous ν et $j \neq k$, les variables aléatoires $\hat{\Gamma}^{(j)}(\nu)$ et $\hat{\Gamma}^{(k)}(\nu)$ sont indépendantes et ont des variances égales. Donc :

$$\begin{aligned} \text{Var}(\hat{\Gamma}_X^{Bart}(\nu)) &= \frac{1}{K^2} \sum_{k=0}^{K-1} \text{Var}(\hat{\Gamma}^{(k)}(\nu)) \\ &= \frac{1}{K^2} \sum_{k=0}^{K-1} \text{Var}(\hat{\Gamma}^{(0)}(\nu)) \\ &= \frac{1}{K} \text{Var}(\hat{\Gamma}^{(0)}(\nu)) \\ &= \frac{\text{Var}(\hat{\Gamma}_X^L(\nu))}{K}. \end{aligned}$$

D'après la section 6.3.2, $\text{Var}(\hat{\Gamma}_X^L(\nu)) \rightarrow \Gamma_X(\nu)^2$ quand $L \rightarrow \infty$, donc $\text{Var}(\hat{\Gamma}_X^{Bart}(\nu)) \rightarrow \frac{\Gamma_X(\nu)^2}{K}$. Dans le régime asymptotique (K fixé, $L \rightarrow \infty$), la variance du périodogramme de Bartlett est donc K fois plus faible que celle du périodogramme non fenêtré sur N échantillons.

En pratique, le choix du nombre de fenêtres K est un compromis entre biais et variance, comme illustré dans l'exemple ci-dessous.

Exemple 6.4. On observe 10 000 échantillons temporels du processus dont la DSP est tracée en gras sur la figure 34.

- La figure du haut présente trois réalisations du périodogramme standard. On constate 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. La variance est réduite, et le biais reste très faible : la courbe de $\hat{\Gamma}_X^{Bart}(v)$, tracée en pointillés, est quasiment confondue avec celle de $\Gamma_X(v)$. 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.

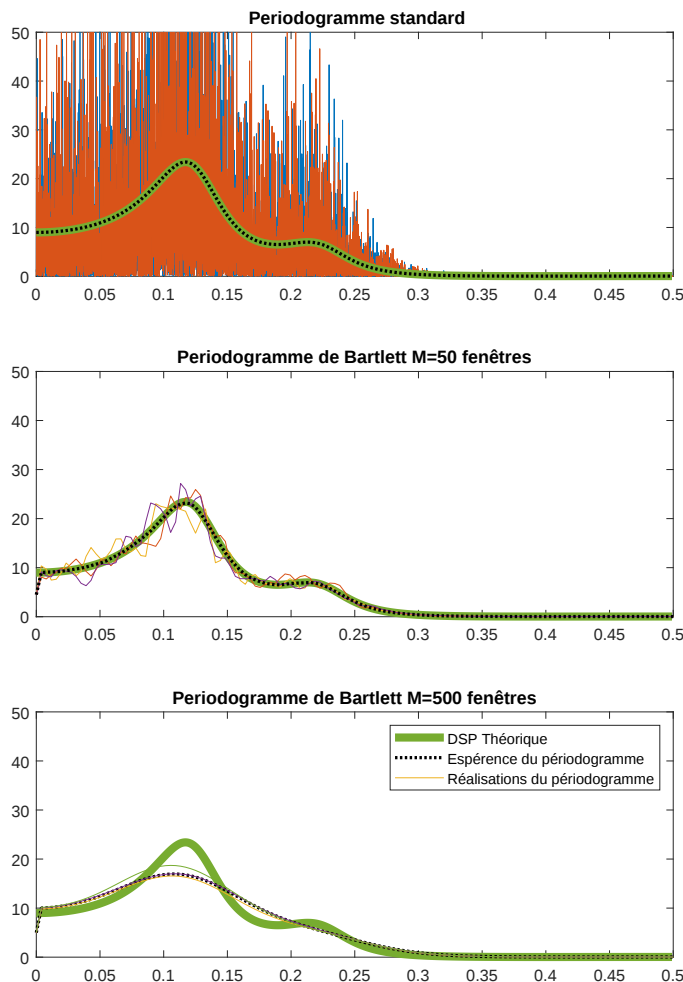


FIGURE 34 – Périodogramme de Bartlett : augmenter le nombre de fenêtres permet de réduire la variance au détriment du biais.

6.4 ESTIMATION PARAMÉTRIQUE DE LA DSP

6.4.1 Principe

Le périodogramme est un estimateur non paramétrique de variance très élevée. L'approche paramétrique consiste à décrire le processus stationnaire X_n par un modèle ayant quelques paramètres (par exemple, le modèle autorégressif). En estimant ces paramètres à partir d'une réalisation $x[n]$, on peut en déduire une estimation de la DSP. Cette approche nécessite de faire une hypothèse sur le modèle décrivant le processus, mais conduit à des estimateurs de variance réduite.

L'estimation paramétrique s'applique par exemple au traitement de la parole avec des modèles AR, ou à l'analyse vibratoire *via* l'estimation des paramètres de sinusoides bruitées.

6.4.2 Cas d'un signal modélisé par un processus AR

D'après le théorème 5.3, le modèle AR d'ordre p s'écrit sous la forme récursive :

$$X_n = \sum_{k=1}^p a_k X_{n-k} + W_n \quad (151)$$

où W_n est un bruit blanc générateur de variance σ^2 . Il y a donc $p + 1$ paramètres : $a_1, \dots, a_p, \sigma^2$.

Estimation des paramètres du modèle AR à partir de données expérimentales

L'approche la plus classique s'appuie sur l'estimation de la fonction d'autocorrélation du processus AR. Les équations de Yule-Walker (145) permettent de relier linéairement les paramètres du modèle aux valeurs de l'autocorrélation :

$$\forall k \geq 0, \quad \gamma_X[k] = \sum_{i=1}^p a_i \gamma_X[k-i] + \sigma^2 \delta[k].$$

Rassemblons ces équations pour $k = 0, \dots, p$. Comme la fonction $\gamma_X[k]$ est paire, on obtient le système linéaire :

$$\begin{pmatrix} \gamma_X[1] & \gamma_X[2] & \cdots & \gamma_X[p] & 1 \\ \gamma_X[0] & \gamma_X[1] & \cdots & \gamma_X[p-1] & 0 \\ \gamma_X[1] & \gamma_X[0] & \cdots & \gamma_X[p-2] & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ \gamma_X[p-1] & \gamma_X[p-2] & \cdots & \gamma_X[0] & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \\ \sigma^2 \end{pmatrix} = \begin{pmatrix} \gamma_X[0] \\ \gamma_X[1] \\ \gamma_X[2] \\ \vdots \\ \gamma_X[p] \end{pmatrix}$$

Concrètement, la fonction d'autocorrélation $\gamma_X[k]$ est remplacée par son estimateur biaisé $\hat{\gamma}_X^N[k]$, voir définition 6.4. La résolution du système linéaire fournit une estimation des paramètres $\hat{a}_k$ et $\hat{\sigma}^2$.

Estimation de la DSP

La DSP d'un processus AR s'exprime analytiquement à partir de ses paramètres, voir (144). Ainsi, l'estimateur de la DSP s'écrit :

$$\hat{\Gamma}_X^{\text{AR}}(\nu) = \frac{\hat{\sigma}^2}{\left| 1 - \sum_{k=1}^p \hat{a}_k e^{-2\pi i k \nu} \right|^2}.$$

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.

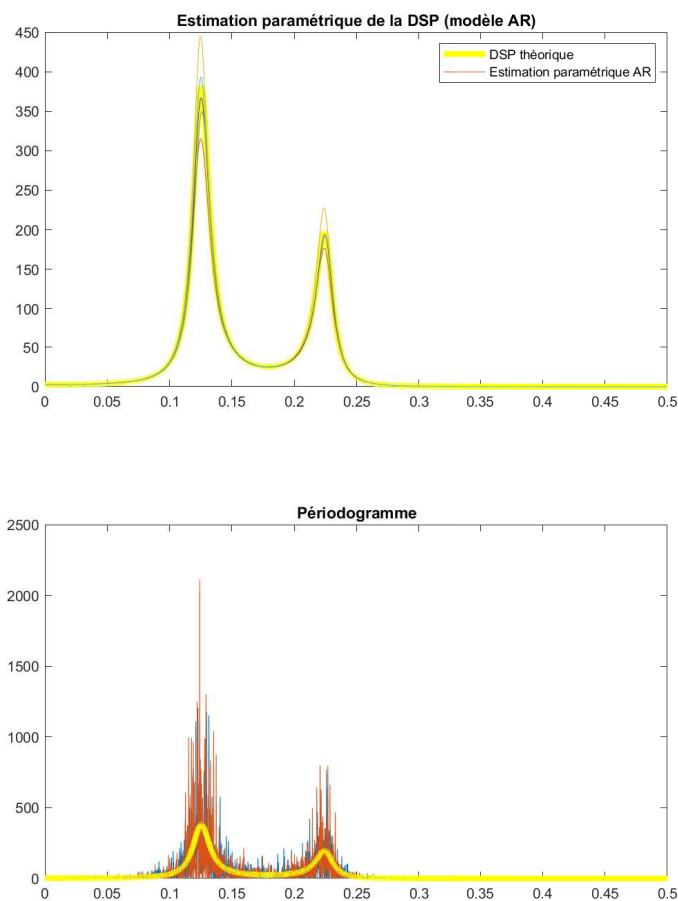


FIGURE 35 – Comparaison de l'estimation paramétrique de la DSP avec le périodogramme standard.

Exemple 6.5. Application au traitement de la parole

Certains sons de parole sont bien représentés par un modèle AR. Il s'agit des phonèmes non voisés, pour lesquels les cordes vocales ne vibrent pas. Les principaux sons non voisés sont les consonnes fricatives : /f/, /s/, etc. et les voyelles chuchotées. Le modèle AR peut être interprété ainsi :

- les turbulences de l'air à l'entrée du conduit vocal correspondent à un bruit blanc,
- les résonances du conduit vocal correspondent aux pôles du modèle AR.

L'estimation des paramètres du modèle AR associé à un phonème permet l'identification du phonème (reconnaissance vocale) ou la synthèse de nouveaux échantillons (débruitage et synthèse vocale). La figure 36 illustre le traitement complet.

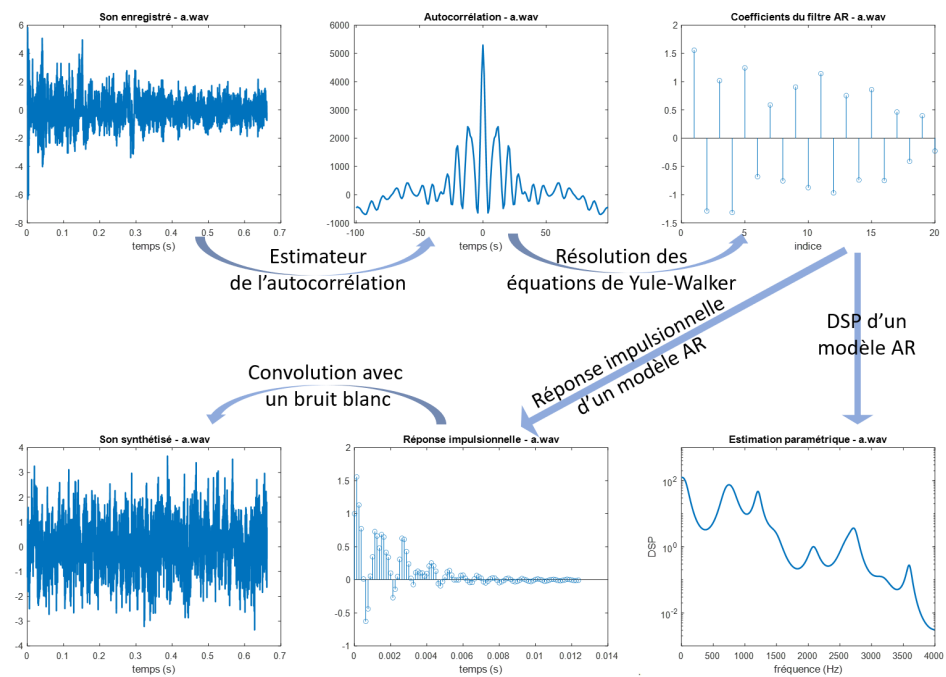


FIGURE 36 – Traitement de la parole et synthèse vocale.

6.5 ESTIMATION LINÉAIRE OPTIMALE

Dans de nombreuses applications, on cherche à estimer un processus aléatoire X à l'instant n à partir de P échantillons temporels d'un processus aléatoire Y , modélisant les observations disponibles. Pour que l'estimation soit possible, il est indispensable que les processus X et Y soient corrélés. On supposera connues la fonction d'autocorrélation de Y et la fonction d'intercorrélation entre X et Y . De plus, les processus seront supposés centrés et à valeurs réelles.

Parmi les applications les plus courantes, on peut citer :

- la prédiction : prédire la valeur du processus X_n à l'instant n à partir de P observations antérieures : $X_{n-1}, \dots, X_{n-P}$,
- le débruitage : estimer X_n à l'instant n à partir d'observations bruitées $Y_k = X_k + W_k$ à différents instants k ,
- la déconvolution : estimer X_n à l'instant n à partir de mesures entachées d'un flou $Y_k = (h * X)_k$, où h est la réponse impulsionnelle d'un filtre passe-bas, modélisant le flou.

Dans la section suivante, on utilise les notations simplifiées $U = X_n \in \mathbb{R}$ pour la variable aléatoire à estimer et $\mathbf{Z} \in \mathbb{R}^P$ pour le vecteur aléatoire rassemblant les P variables aléatoires extraites du processus Y_n , correspondant aux observations.

6.5.1 Définition et propriétés

Définition 6.8 (Estimateur linéaire optimal). Soit $\mathbf{Z}$ un vecteur aléatoire centré à valeurs dans $\mathbb{R}^P$ et U une variable aléatoire réelle centrée. Un **estimateur linéaire** de U à partir de $\mathbf{Z}$ s'écrit

$$\hat{U}(\mathbf{h}) = \mathbf{h}^t \mathbf{Z} = \sum_{p=1}^P h_p Z_p \quad (152)$$

où $\mathbf{h}$ regroupe les poids h_p .

Un **estimateur linéaire optimal** s'écrit $\hat{U}(\mathbf{h}_{\text{opt}}) = \mathbf{h}_{\text{opt}}^t \mathbf{Z}$ où $\mathbf{h}_{\text{opt}}$ minimise la variance de l'erreur d'estimation :

$$\mathbf{h}_{\text{opt}} = \underset{\mathbf{h}}{\operatorname{argmin}} \mathbb{E} \left[\left(U - \hat{U}(\mathbf{h}) \right)^2 \right]. \quad (153)$$

Théorème 6.1. Soit $\Sigma_{\mathbf{Z}} = \mathbb{E}(\mathbf{Z}\mathbf{Z}^t)$ la matrice de covariance $\mathbf{Z}$, supposée inversible. Le vecteur de poids optimal est donné par :

$$\mathbf{h}_{\text{opt}} = \Sigma_{\mathbf{Z}}^{-1} \mathbb{E}(\mathbf{Z}U). \quad (154)$$

Démonstration. La variance de l'erreur d'estimation s'écrit :

$$\begin{aligned} \mathbb{E} \left[\left(U - \hat{U}(\mathbf{h}) \right)^2 \right] &= \mathbb{E} \left((U - \mathbf{h}^t \mathbf{Z})^2 \right) \\ &= \mathbb{E} (\mathbf{h}^t \mathbf{Z} \mathbf{Z}^t \mathbf{h}) - 2\mathbb{E} (\mathbf{h}^t \mathbf{Z} U) + \mathbb{E} (U^2) \\ &= \mathbf{h}^t \Sigma_{\mathbf{Z}} \mathbf{h} - 2\mathbf{h}^t \mathbb{E}(\mathbf{Z}U) + \mathbb{E}(U^2). \end{aligned}$$

Sa minimisation conduit à résoudre un problème d'optimisation quadratique :

$$\mathbf{h}_{\text{opt}} = \underset{\mathbf{h}}{\operatorname{argmin}} \left\{ J(\mathbf{h}) := \mathbf{h}^t \Sigma_{\mathbf{Z}} \mathbf{h} - 2\mathbf{h}^t \mathbb{E}(\mathbf{Z}U) + \mathbb{E}(U^2) \right\}.$$

Par hypothèse, la matrice $\Sigma_{\mathbf{Z}}$ est inversible. En tant que matrice de covariance, elle est symétrique définie positive. La fonction J est donc strictement convexe, et le problème d'optimisation possède une unique solution, qui satisfait :

$$\nabla_{\mathbf{h}} J(\mathbf{h} = \mathbf{h}_{\text{opt}}) = \mathbf{0} \iff 2\Sigma_{\mathbf{Z}} \mathbf{h}_{\text{opt}} - 2\mathbb{E}(\mathbf{Z}U) = \mathbf{0},$$

d'où $\mathbf{h}_{\text{opt}} = \Sigma_{\mathbf{Z}}^{-1} \mathbb{E}(\mathbf{Z}U)$. □

La solution $\hat{U}$ associée au vecteur de poids $\mathbf{h}_{\text{opt}}$ s'interprète comme la projection orthogonale de $U \in L^2(\Omega, \mathcal{F}, \mathcal{P})$ sur le sous-espace vectoriel de $L^2(\Omega, \mathcal{F}, \mathcal{P})$ engendré par les variables $Z_1, \dots, Z_P$. Le corollaire 6.1 exprime que le résidu de la projection, $U - \hat{U}$, est orthogonal à ce sous-espace au sens du produit scalaire de $L^2(\Omega, \mathcal{F}, \mathcal{P})$.

Corollaire 6.1. *L'erreur d'estimation est décorrélée des observations :*

$$\forall p \in \{1, \dots, P\}, \quad \mathbb{E}[(U - \hat{U}(\mathbf{h}_{\text{opt}})) Z_p] = 0.$$

Démonstration. En remarquant que $\mathbf{h}_{\text{opt}}$ est constant (déterministe), on a :

$$\begin{aligned} \mathbb{E}(\mathbf{Z}(U - \hat{U}(\mathbf{h}_{\text{opt}}))) &= \mathbb{E}(\mathbf{Z}U - \mathbf{Z}\mathbf{h}_{\text{opt}}^t \mathbf{Z}) \\ &= \mathbb{E}(\mathbf{Z}U - \mathbf{Z}\mathbf{Z}^t \mathbf{h}_{\text{opt}}) \\ &= \mathbb{E}(\mathbf{Z}U) - \mathbb{E}(\mathbf{Z}\mathbf{Z}^t) \mathbf{h}_{\text{opt}} \\ &= \mathbb{E}(\mathbf{Z}U) - \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{h}_{\text{opt}} \\ &= 0 \end{aligned}$$

d'après (154). □

6.5.2 Prédiction d'un processus aléatoire

On applique la technique d'estimation linéaire optimale à la prédiction d'un processus aléatoire. Dans différents contextes (par exemple, la prédiction de l'évolution de valeurs financières), on est amené à prédire la valeur d'un processus X_n à l'instant n ($U = X_n$) à partir de P valeurs passées, notées $\mathbf{Z} = (X_{n-1}, X_{n-2}, \dots, X_{n-P})$. La fonction d'autocorrélation γ_X est supposée connue.

Théorème 6.2. *Le prédicteur linéaire optimal de X_n à partir de $\mathbf{Z} = (X_{n-1}, X_{n-2}, \dots, X_{n-P})^t$ est donné par :*

$$\hat{X}_n = \mathbf{h}_{\text{opt}}^t \mathbf{Z} = \sum_{p=1}^P [h_{\text{opt}}]_p X_{n-p} \quad (155)$$

où

$$\mathbf{h}_{\text{opt}} = \begin{pmatrix} \gamma_X[0] & \gamma_X[1] & \cdots & \gamma_X[P-1] \\ \gamma_X[1] & \gamma_X[0] & \cdots & \gamma_X[P-2] \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_X[P-1] & \gamma_X[P-2] & \cdots & \gamma_X[0] \end{pmatrix}^{-1} \begin{pmatrix} \gamma_X[1] \\ \gamma_X[2] \\ \vdots \\ \gamma_X[P] \end{pmatrix}.$$

Démonstration. L'expression de $\mathbf{h}_{\text{opt}}$ découle du théorème 6.1 et de la relation $\gamma_X[k] = \mathbb{E}(X_{n+k}X_n) = \gamma_X[-k]$, X étant à valeurs réelles. Notons que comme X_n est stationnaire, $\mathbb{E}(\mathbf{Z}X_n)$ et $\boldsymbol{\Sigma}_{\mathbf{Z}} = \mathbb{E}[\mathbf{Z}\mathbf{Z}^t]$ ne dépendent pas de n . □

Remarque. (155) est une relation de convolution. Le prédicteur s'interprète donc comme un filtre linéaire de réponse impulsionnelle $\mathbf{h}_{\text{opt}}$.

Exemple 6.6. Sur la figure 37, le tracé du haut représente $\gamma_X[k]$. On remarque que les échantillons espacés d'environ 40 points sont fortement corrélés. Le tracé du bas représente les coefficients $\mathbf{h}_{\text{opt}}$ pour $P = 50$. La figure 38 représente une réalisation $x[n]$ du processus X_n , tracée en gras. Les prédictions de $x[n]$ réalisées à partir de $P = 5$ et $P = 50$ échantillons temporels antérieurs sont tracées en trait fin. L'usage d'un plus grand nombre d'échantillons antérieurs améliore la prédiction en tirant parti de la corrélation élevée avec les échantillons temporels lointains.

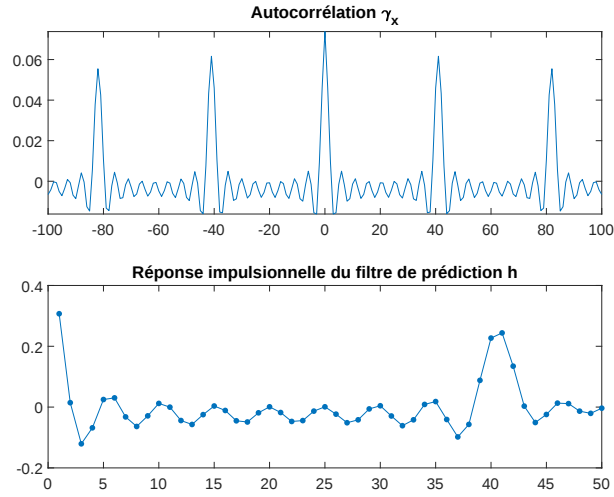


FIGURE 37 – Haut : autocorrélation du processus (γ_X). Bas : réponse impulsionnelle du filtre de prédiction ($\mathbf{h}_{\text{opt}}$).

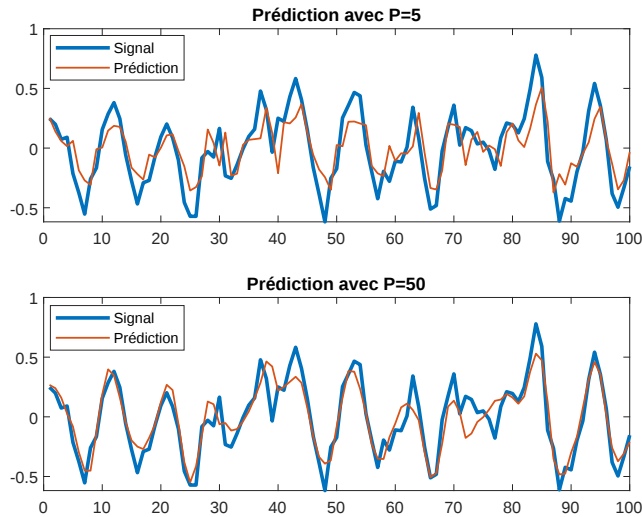


FIGURE 38 – Prédiction en utilisant $P = 5$ et $P = 50$ échantillons.

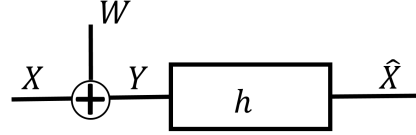


FIGURE 39 – Modèle pour le débruitage : recherche du filtre optimal h à appliquer aux données bruitées Y_n .

6.5.3 Filtre de Wiener pour le débruitage

Dans cette section, l'estimation linéaire optimale est appliquée au débruitage d'un signal. Le signal bruité est modélisé par un processus Y_n résultant de la somme d'un processus X_n (inconnu, non bruité) et d'un processus W_n inconnu correspondant au bruit :

$$Y_n = X_n + W_n$$

L'objectif est de reconstruire X_n en filtrant le signal Y_n . La chaîne de traitement est représentée en Fig. 39. On fait les hypothèses suivantes :

- W et X sont décorrélés,
- les fonctions d'autocorrélation γ_X et γ_W sont connues.

Cas d'un filtre à réponse impulsionnelle finie

On traite ici l'estimation de X_n à partir de P observations passées. On cherche donc à appliquer au processus Y un filtre h de durée P :

$$\hat{X}_n = (h * Y)_n = \sum_{p=0}^{P-1} h_p Y_{n-p}.$$

Cette équation se réécrit sous la forme $\hat{X}_n = \mathbf{h}^t \mathbf{Z}$ avec $\mathbf{Z} = (Y_n, Y_{n-1}, \dots, Y_{n-P+1})^t$. D'après le théorème 6.1, le filtre optimal est donné par $\mathbf{h}_{\text{opt}} = \mathbf{\Sigma}_Z^{-1} \mathbb{E}(\mathbf{Z}X)$, avec

$$\mathbf{\Sigma}_Z = \begin{pmatrix} \gamma_Y[0] & \gamma_Y[1] & \cdots & \cdots & \gamma_Y[P-1] \\ \gamma_Y[1] & \gamma_Y[0] & \cdots & \cdots & \gamma_Y[P-2] \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \gamma_Y[P-1] & \gamma_Y[P-2] & \cdots & \cdots & \gamma_Y[0] \end{pmatrix}$$

et

$$\mathbb{E}(\mathbf{Z}X) = \begin{pmatrix} \gamma_{YX}[0] \\ \gamma_{YX}[-1] \\ \vdots \\ \gamma_{YX}[1-P] \end{pmatrix}.$$

Comme W et X sont décorrélés, les éléments de ces matrices s'expriment directement en fonction de γ_X et γ_W : pour $k = 0, \dots, P-1$, $\gamma_Y[k] = \gamma_X[k] + \gamma_W[k]$ et $\gamma_{YX}[-k] = \gamma_X[-k] = (\gamma_X[k])^*$. Comme X_n est à valeurs réelles, on a donc $\gamma_{YX}[-k] = \gamma_X[k]$.

Cas d'un filtre à réponse impulsionnelle infinie

On étend le résultat précédent à un filtre de réponse impulsionnelle infinie. Autrement dit, on cherche à estimer X_n à partir d'un nombre infini d'observations temporelles, Y_{n-p} pour $p \in \mathbb{Z}$. On montre que la réponse fréquentielle du filtre optimal s'exprime directement en fonction des densités spectrales de puissance de X et W .

Théorème 6.3 (Filtre de Wiener). *Le filtre linéaire optimal pour le débruitage, appelé **filtre de Wiener**, a pour équation de convolution :*

$$\hat{X}_n = \sum_{p \in \mathbb{Z}} [\mathbf{h}_{\text{opt}}]_p Y_{n-p}$$

où la fonction de transfert est donnée par :

$$H_{\text{opt}}(\nu) = \frac{\Gamma_X(\nu)}{\Gamma_X(\nu) + \Gamma_W(\nu)} = \frac{1}{1 + \Gamma_W(\nu)/\Gamma_X(\nu)}.$$

Esquisse de démonstration. Le cas d'une réponse impulsionnelle infinie ne correspond pas parfaitement au cadre du théorème 6.1, où le nombre d'observations Y_{n-p} est fini. Ceci étant, certains raisonnements restent valables, comme le fait que $\hat{X}_n$ est la projection orthogonale de X_n sur l'adhérence du sous-espace de $L^2(\Omega, \mathcal{F}, \mathcal{P})$ engendré par les variables Y_{n-p} . Ainsi, la décorrélation entre l'erreur d'estimation et les observations (corollaire 6.1) est assurée. On a donc :

$$\begin{aligned} \mathbb{E}((X_n - \hat{X}_n)Y_{n-p}) &= 0 \quad \forall p, n \in \mathbb{Z} \\ \mathbb{E}(X_n Y_{n-p}) - \mathbb{E}(\hat{X}_n Y_{n-p}) &= 0 \\ \mathbb{E}(X_n(X_{n-p} + W_{n-p})) - \mathbb{E}(\hat{X}_n Y_{n-p}) &= 0 \\ \mathbb{E}(X_n X_{n-p}) + \mathbb{E}(X_n W_{n-p}) - \mathbb{E}(\hat{X}_n Y_{n-p}) &= 0 \\ \gamma_X[p] + 0 - \gamma_{\hat{X}Y}[p] &= 0, \end{aligned}$$

donc $\gamma_X = \gamma_{\hat{X}Y}$. D'après la formule des interférences, $\gamma_{\hat{X}Y} = h_{\text{opt}} * \gamma_Y$. De plus, $\gamma_Y = \gamma_X + \gamma_W$ car $Y_n = X_n + W_n$ avec W_n et X_n décorrélés. Finalement,

$$\gamma_X = h_{\text{opt}} * \gamma_Y = h_{\text{opt}} * (\gamma_X + \gamma_W).$$

En appliquant la TFTD à l'expression précédente, on obtient :

$$\Gamma_X(\nu) = H_{\text{opt}}(\nu)(\Gamma_X(\nu) + \Gamma_W(\nu))$$

d'où l'expression de $H_{\text{opt}}(\nu)$. □

L'expression de la réponse fréquentielle s'interprète de la manière suivante : le filtre atténue le signal Y_n principalement dans les bandes de fréquence où la DSP du bruit est plus grande que celle du signal X . **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 processus bruité $Y_n = X_n + W_n$. On suppose connues les DSP $\Gamma_X(\nu)$ et $\Gamma_W(\nu)$, représentées en figure 40(a). La réponse fréquentielle $H_{\text{opt}}(\nu)$ du filtre de Wiener est représentée en pointillé. Elle s'adapte au rapport signal-sur-bruit dans chaque bande de fréquence : lorsque la puissance du signal est importante par rapport à celle du bruit, le gain du filtre est élevé pour préserver le contenu fréquentiel de Y . Dans le cas contraire, le gain du filtre devient faible et le bruit est atténué. La figure 40(b) représente des trajectoires temporelles des différents processus sur un exemple simulé. En haut, le processus bruité Y_n est représenté avec le signal de référence (processus non bruité X_n utilisé pour générer les données simulées Y_n) ; en bas, le processus débruité $\hat{X}_n$ est comparé avec X_n .

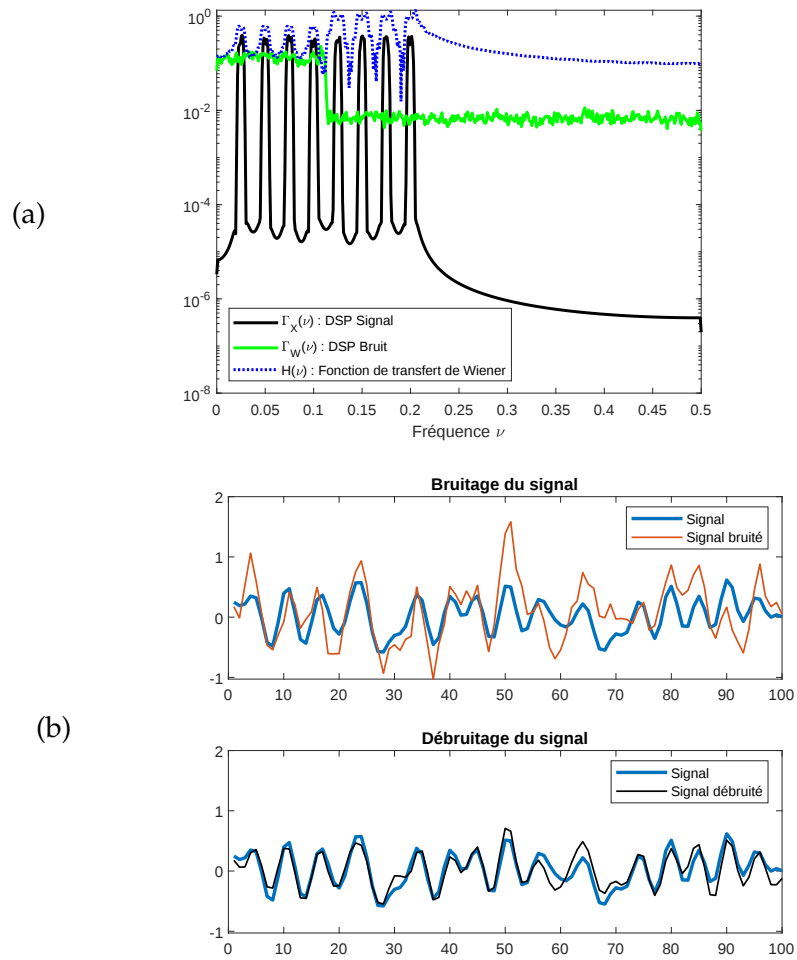


FIGURE 40 – Débruitage à l'aide du filtre de Wiener. (a) DSP du signal et du bruit (Γ_X et Γ_W) et réponse fréquentielle du filtre de Wiener (H_{opt}). (b) Représentation temporelle des processus Y_n (bruité), X_n (signal inconnu) et $\hat{X}_n$ (version débruitée de Y_n).

6.6 CONCLUSION

On retiendra en premier lieu les expressions de l'estimateur de la moyenne d'un processus, des estimateurs biaisé et non biaisé de la fonction d'autocorrélation, et des estimateurs de la DSP, à savoir le périodogramme standard et le périodogramme de Bartlett. La méthode du périodogramme standard possède des limitations importantes en termes de biais (lié à la résolution spectrale) et de variance, qui rendent l'estimation imprécise. Grâce à son effet de fenêtrage, la méthode de Bartlett permet de réduire sensiblement la variance de l'estimateur. Les techniques d'estimation paramétrique de la DSP sont très efficaces lorsque le processus d'intérêt est modélisable à l'aide de quelques paramètres (par exemple, pour les processus autorégressifs). Enfin, les techniques de filtrage linéaire optimal sont utiles pour aborder divers problèmes de prédiction, débruitage et déconvolution. Leur principe est de rechercher le meilleur filtre au sens de l'erreur quadratique moyenne, ce qui conduit à résoudre un problème d'optimisation dont les inconnues sont les coefficients du filtre. Ce problème d'optimisation étant quadratique, sa résolution se résume à l'inversion d'un système linéaire.

Les différentes techniques présentées sont illustrées par des simulations fournies sur la page edunao du cours, dans le cadre de la synthèse du son et du débruitage d'un signal audio. Les codes correspondant à ces exemples sont disponibles sur edunao.

INDEX

- autocorrélation, 83, 97
- autocovariance, 84
- base hilbertienne, 15
- bloquage d'ordre zéro, 76
- bourrage de zéros, 58, 62, 101
- bruit blanc, 88
- causalité, 32
- convolution, 26
- débruitage, 112
- décomposition en série de
 - Fourier, 48
- densité spectrale de puissance,
 - 86, 100
- distributions, 16
- échantillonnage, 67
- énergie, 26
- équation aux différences finies,
 - 36, 92
- équation de Yule-Walker, 94
- ergodicité, 96
- espace ℓ^p , 14
- espace L^p , 14
- espace de Schwartz, 19
- estimateur de
 - l'autocorrélation, 97
- estimateur de la moyenne, 96
- estimateur linéaire optimal,
 - 108
- estimation paramétrique, 106
- fenêtrage, 59
- filtrage de processus, 90, 113
- filtre anti-repliement, 75
- filtre de Wiener, 112
- filtre FIR, 34
- filtre IIR, 36
- fonction de Kronecker, 14
- fonction de transfert, 55
- formule des interférences, 90
- formule sommatoire de
 - Poisson, 45
- intercorrélation, 26, 83
- intercovariance, 84
- invariance par translation, 31
- lois fini-dimensionnelles, 81
- moyenne, 83
- observation, 80
- périodogramme, 100
- périodogramme de Bartlett,
 - 103
- peigne de Dirac, 17, 46
- prédiction, 110
- processus aléatoire, 21
- processus autorégressif (AR),
 - 92
- processus gaussien, 82
- puissance, 26, 85
- quantification, 25
- réalisation, 80
- réponse fréquentielle, 55
- réponse impulsionnelle, 33, 91
- résolution spectrale, 61
- rapport signal-sur-bruit, 26, 85
- signal à temps continu, 25
- signal à temps discret, 25
- signal numérique, 25
- signal stochastique, 26, 80
- spectre, 40
- stabilité EBSB, 32
- stationnarité, 84
- support, 19
- théorème de Shannon, 70

traitement de la parole, 108
trajectoire, 80
transformée de Fourier, 40
transformée de Fourier à
temps discret, 49

transformée de Fourier
discrète, 53, 101
variance, 85
variante du périodogramme,
103
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.