

Traitement numérique du signal.  
Première partie : Bases mathématiques

J.IDIER  
H. PIET-LAHANIER  
G. LE BESNERAIS  
F. CHAMPAGNAT

Première version du document : 1993  
Date de la dernière remise à jour : mars 2004



# Table des matières

Remerciements . . . . .	7
<b>Introduction</b>	<b>9</b>
Notion de signal . . . . .	9
Notion de système . . . . .	9
<b>I Filtrage linéaire et transformée de Fourier des signaux déterministes à temps continu</b>	<b>11</b>
I.1 Introduction . . . . .	11
I.1.1 Analyse fréquentielle et transformée de Fourier . . . . .	11
I.1.2 Convolution . . . . .	14
I.2 Cas des signaux stables . . . . .	15
I.2.1 Définition et propriétés de la TF dans $L^1_{\mathbb{C}}(\mathbb{R})$ . . . . .	15
I.2.2 Convolution dans $L^1_{\mathbb{C}}(\mathbb{R})$ . . . . .	16
I.3 Cas des signaux d'énergie finie . . . . .	16
I.3.1 Définition et propriétés de la TF dans $L^2_{\mathbb{C}}(\mathbb{R})$ . . . . .	17
I.3.2 Convolution « $L^1_{\mathbb{C}}(\mathbb{R}) * L^2_{\mathbb{C}}(\mathbb{R})$ » . . . . .	18
I.4 Signaux harmonisables . . . . .	18
I.5 Filtrage linéaire . . . . .	19
I.5.1 Filtre convolutionnel . . . . .	19
I.5.2 Représentation fréquentielle . . . . .	20
Annexe A Fonctions régularisantes . . . . .	21
Annexe B TF inverse de $\tilde{x}$ stable . . . . .	22
<b>II Signaux déterministes à temps discret</b>	<b>25</b>
II.1 Introduction . . . . .	25
II.1.1 Formule sommatoire de Poisson . . . . .	26
II.2 Théorème d'échantillonnage . . . . .	26
II.2.1 Présentation . . . . .	26
II.2.2 Théorème de Shannon-Nyquist . . . . .	27
II.3 Transformée en $z$ . . . . .	30
II.3.1 Définition . . . . .	30
II.3.2 Propriétés de la transformée en $z$ . . . . .	31
II.4 Transformée de Fourier à temps discret . . . . .	32
II.4.1 Lien entre transformée en $z$ et transformée de Fourier . . . . .	32

II.4.2	Transformée de Fourier discrète (TFD)	32
<b>III</b>	<b>Filtres numériques</b>	<b>35</b>
III.1	Introduction	35
III.1.1	Stabilité	37
III.1.2	Causalité	37
III.1.3	Inversion d'un filtre numérique	37
III.1.4	Filtres rationnels, filtres dynamiques	37
III.1.5	Filtres à Réponse Impulsionnelle Finie (RIF)	39
III.1.6	Filtres Tous Pôles (FTP)	39
III.2	Exemples de filtres	40
III.2.1	Retard	40
III.2.2	Lissage	40
III.2.3	Filtre récursif d'ordre 1	40
III.2.4	Filtre récursif d'ordre 2	42
<b>IV</b>	<b>Événements, variables et vecteurs aléatoires</b>	<b>47</b>
IV.1	Introduction	47
IV.2	Intégration et probabilités	47
IV.2.1	Espaces et applications mesurables	47
IV.2.2	Notion de mesure	48
IV.2.3	Probabilités, événements	48
IV.3	Variables aléatoires	50
IV.3.1	Variables aléatoires	50
IV.3.2	Variables aléatoires réelles	50
IV.3.3	Espérance mathématique	52
IV.4	Couple de variables aléatoires réelles	54
IV.4.1	Caractérisation et propriétés	54
IV.4.2	Conditionnement	56
IV.5	Vecteurs aléatoires réels	57
IV.5.1	Généralisations	57
IV.5.2	Changement de variable	58
IV.5.3	Vecteurs aléatoires d'ordre deux	59
IV.5.4	Vecteurs aléatoires gaussiens	60
IV.6	L'espace de Hilbert $L^2(\Omega, \mathcal{F}, P)$	61
IV.6.1	Norme et produit scalaire	61
IV.6.2	Sous-espaces de Hilbert	61
IV.6.3	Espérance conditionnelle dans $L^2(\Omega, \mathcal{F}, P)$	63
<b>V</b>	<b>Signaux aléatoires à temps discret</b>	<b>65</b>
V.1	Introduction	65
V.2	Caractérisation	65
V.2.1	Définition	65

---

V.2.2	Loi temporelle . . . . .	65
V.2.3	Caractéristiques instantanées . . . . .	66
V.2.4	Caractéristiques à l'ordre deux . . . . .	67
V.3	Construction et exemples . . . . .	68
V.3.1	Bruits blancs . . . . .	68
V.3.2	Chaines de Markov . . . . .	69
V.3.3	Constructions indirectes . . . . .	70
V.4	Propriétés . . . . .	73
V.4.1	Convergence . . . . .	73
V.4.2	Stationnarité . . . . .	75
V.4.3	Ergodisme . . . . .	76
V.5	Les signaux gaussiens . . . . .	77
<b>VI</b>	<b>Signaux aléatoires stationnaires faibles</b>	<b>79</b>
VI.1	Introduction . . . . .	79
VI.2	Représentation harmonique . . . . .	79
VI.2.1	Fonction d'autocorrélation . . . . .	79
VI.2.2	Cas d'un bruit blanc numérique . . . . .	81
VI.2.3	Trajectoires . . . . .	81
VI.3	Filtrage linéaire convolutionnel . . . . .	82
VI.4	Analyse spectrale classique . . . . .	83
VI.4.1	Position du problème . . . . .	83
VI.4.2	Périodogramme . . . . .	83
VI.5	Analyse spectrale paramétrique . . . . .	85
VI.5.1	Définition . . . . .	88
VI.5.2	ARMA réguliers . . . . .	88
VI.5.3	Représentation spectrale des signaux autorégressifs . . . . .	89
VI.5.4	Équations de Yule-Walker . . . . .	89
VI.5.5	Estimation spectrale autorégressive . . . . .	90



# Avant-propos

## Remerciements

Les auteurs tiennent à remercier les personnes qui ont participé à ce cours à l'ENSTA et, en particulier, Yves Goussard, Stéphane Gautier et Jean-François Giovannelli ; ainsi que celles qui ont fait part de leurs remarques et corrections et contribué à l'amélioration du document écrit, en particulier Caroline Kulcsár.



# Introduction

# Signaux et systèmes

## Notion de signal

On désigne par *signal déterministe*, ou simplement par *signal*, une fonction d'une ou de plusieurs variables de temps ou d'espace, en général à valeurs réelles ou complexes. Un signal est un modèle commode pour décrire et étudier les variations continues d'un paramètre physique (tension, intensité lumineuse, pression...) : on parle alors de *signal analogique* ; ou encore pour modéliser un ensemble discret de valeurs indexé par des variations quantifiées (du temps, de variables d'espace...), et on parle alors de *signal numérique*. Tandis qu'un signal déterministe modélise les variations d'un paramètre au cours d'une expérience unique, la notion de *signal aléatoire* permet de rassembler des signaux déterministes différents (chaque *trajectoire* d'un signal aléatoire est un signal déterministe), mais dont on présume des propriétés communes. Par exemple les trajectoires de chacune des boules de loto mélangées dans une urne sont différentes mais interchangeable ; on pourrait caractériser leurs propriétés communes comme celles d'un unique signal aléatoire. Les signaux temporels sont seulement fonctions de la variable de temps. Lorsque le signal évolue de façon continue avec le temps, on parle de *signaux à temps continu*. Dans le cas où le signal est une fonction à variable discrète du temps, le signal est dit à *temps discret*. La plupart des concepts et des résultats présentés dans ce cours pour des signaux temporels admettent des extensions pour des signaux multivariés. Ainsi, le traitement des images fait largement appel aux extensions bi-dimensionnelles.

Très souvent, les signaux à temps discret sont obtenus en retenant les valeurs prises par un signal à temps continu en certains instants, à des fins de stockage ou de traitement sur ordinateur. Cette opération est appelée *échantillonnage* et, par extension, les signaux à temps discret sont souvent appelés *signaux échantillonnés*. L'échantillonnage est un préalable au *traitement numérique du signal*, dont l'importance pratique n'a cessé de croître depuis l'apparition des moyens de calculs informatiques. Désormais un simple PC doté d'une carte d'acquisition et d'un langage évolué peut souvent remplacer avantageusement les analyseurs de spectre, corrélateurs et autres filtres à transfert de charges... , dispositifs d'analyse et de traitement analogiques coûteux et fragiles. L'échantillonnage des signaux permet aussi le stockage en mémoires de masse numériques, disques durs, vidéodisques, supports magnétiques... C'est pourquoi nous insistons particulièrement dans ce cours sur la manipulation des signaux à temps discret.

## Notion de système

On peut définir les systèmes comme les modèles mathématiques des diverses transformations subies par les signaux : par exemple une onde qui se propage est modifiée suivant les caractéristiques de transmission du milieu qu'elle traverse ; une quantité physique mesurée par un

capteur subit une certaine altération, qui traduit l'effet de la « réponse » du capteur. Plus généralement, on caractérise la relation d'entrée-sortie d'un système quelconque par un opérateur mathématique qui associe à un signal d'entrée  $x(t)$  un signal de sortie  $y(t)$ . On s'intéressera plus spécifiquement aux cas particuliers des *filtres linéaires homogènes*, dont la caractéristique d'entrée-sortie est un opérateur *linéaire invariant dans le temps*. Ces filtres possèdent de nombreuses propriétés mathématiques et fournissent souvent un premier niveau de description satisfaisant (par exemple, pour des petites variations du signal d'entrée) de systèmes physiques beaucoup plus compliqués.

# Chapitre I

## Filtrage linéaire et transformée de Fourier des signaux déterministes à temps continu

### I.1 Introduction

L'objectif de ce chapitre est de construire et d'étudier les transformations linéaires fondamentales que sont la transformée de Fourier (TF) et le filtrage linéaire. Bien que l'accent soit ensuite porté sur les signaux discrets, l'étude préliminaire de ces transformations pour les signaux à temps continu permet une meilleure compréhension d'ensemble. Elle est aussi un préalable nécessaire à l'étude théorique de l'opération d'échantillonnage (voir le chapitre II).

#### I.1.1 Analyse fréquentielle et transformée de Fourier

##### Décomposition fréquentielle

La représentation fréquentielle ou spectrale d'un signal  $x$  s'impose de façon naturelle dans les domaines où les fréquences pures sont une réalité physique sous la forme des ondes monochromatiques en acoustique, en mécanique vibratoire ou en optique. Plus généralement, elle constitue un cadre pratique pour l'analyse des signaux (périodicité, pseudo-périodicité) et des systèmes linéaires. Formellement, l'analyse fréquentielle d'un signal  $x$  est la recherche d'une représentation du type

$$x(t) = \int_{\mathbb{R}} \tilde{x}(\nu) e^{2i\pi\nu t} d\nu, \quad \forall t \in \mathbb{R} \quad (\text{I.1})$$

c'est-à-dire d'une décomposition linéaire du signal sur l'ensemble des signaux complexes  $e^{2i\pi\nu t}$  de fréquence pure  $\nu$ .

##### Transformée de Fourier

$x$  étant une fonction à valeurs complexes de la variable réelle  $t$ , on appelle transformée de Fourier de  $x$  la fonction  $\tilde{x}$  à valeurs complexes de la variable réelle  $\nu$  ( $\nu$  est homogène à une fréquence quand  $t$  est homogène à un temps) :

$$x \xrightarrow{\text{TF}} \tilde{x}, \quad \tilde{x}(\nu) \triangleq \int_{\mathbb{R}} x(t) e^{-2i\pi\nu t} dt. \quad (\text{I.2})$$

**Remarque**

Dans certains cours d'intégration la TF de  $x$  est définie comme la fonction

$$\tilde{x}(\omega) = \int_{\mathbb{R}} x(t) e^{i\omega t} dt = \tilde{x}\left(-\frac{\omega}{2\pi}\right),$$

ce qui revient pour le praticien à mesurer des pulsations (en radians par seconde) plutôt que des fréquences (en Hertz). Cette variante a l'inconvénient de faire apparaître une constante qui dissymétrise la décomposition (I.1), qui devient

$$x(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{x}(\omega) e^{-i\omega t} d\omega.$$

**Propriétés**

Pourvu que les fonctions utilisées ci-dessous existent, on a les propriétés immédiates suivantes :

Linéarité :	$\lambda_1 x_1 + \lambda_2 x_2 \xrightarrow{\text{TF}} \lambda_1 \tilde{x}_1 + \lambda_2 \tilde{x}_2$	(I.3)
Retard :	$x(t - t_0) \xrightarrow{\text{TF}} e^{-2i\pi\nu t_0} \tilde{x}$	(I.4)
Modulation :	$e^{2i\pi\nu_0 t} x(t) \xrightarrow{\text{TF}} \tilde{x}(\nu - \nu_0)$	(I.5)
Affinité :	$x(at) \xrightarrow{\text{TF}} \frac{1}{ a } \tilde{x}\left(\frac{\nu}{a}\right)$	(I.6)
Conjugaison :	$x^*(t) \xrightarrow{\text{TF}} (\tilde{x}(-\nu))^*$	(I.7)
Dérivation :	$x'(t) \xrightarrow{\text{TF}} 2i\pi\nu \tilde{x}(\nu)$	(I.8)

De la propriété (I.7), on déduit que la TF d'un signal réel possède la symétrie hermitienne :

$$\tilde{x}(-\nu) = \tilde{x}(\nu)^*$$

et réciproquement, que la TF d'un signal à symétrie hermitienne est réelle. En associant les deux résultats, on obtient que la TF d'un signal symétrique réel est réelle et symétrique.

Il est fréquent chez les physiciens et les ingénieurs de considérer la représentation (I.1) comme une simple formule d'inversion de la TF du signal  $x$ . Encore faut-il garantir l'existence de la TF  $\tilde{x}$ , puis montrer que l'intégrale (I.1) converge... En quel sens ? Dans quel espace ? Pour tout  $t$  ? En fait les écritures (I.1) et (I.2) s'avèrent plus délicates que prévues, comme le montrent les deux exemples suivants.

**Exemple 1 : Inversibilité de la TF dans un cas simple**

Soit  $x(t) = \mathbf{1}_{[-\tau, \tau]}(t)$  le signal rectangle valant 1 dans l'intervalle  $[-\tau, \tau]$  et 0 ailleurs. Sa TF s'obtient sans difficulté sous la forme

$$\tilde{x}(\nu) = \frac{\sin 2\pi\nu\tau}{\pi\nu} = 2\tau \operatorname{sinc}(2\nu\tau),$$

où  $\operatorname{sinc}$  désigne le sinus cardinal défini par  $t \mapsto \sin(\pi t)/(\pi t)$ . Le sinus cardinal n'étant pas de module intégrable sur  $\mathbb{R}$ , la décomposition (I.1) se révèle problématique. On verra pourtant que cette décomposition est valide presque partout en tant que formule d'inversion de la TF dans l'espace de Hilbert des fonctions dont le carré du module est intégrable. Elle est même valide pour tout  $t$  sauf  $-\tau$  et  $\tau$ , comme le montre un calcul direct faisant appel au théorème des résidus.

### Exemple 2 : Décomposition des signaux périodiques en série de Fourier

Considérons l'ensemble des signaux périodiques de période 1 de carré intégrable sur leur période, sans faire la distinction entre des signaux presque partout (pp) égaux pour la mesure de Lebesgue. Muni de la mesure de Lebesgue et du produit scalaire

$$\langle x, y \rangle = \int_0^1 x(t) y^*(t) dt,$$

cet ensemble est l'espace de Hilbert  $L^2_{\mathbb{C}}([0, 1[)$ , admettant pour base dénombrable orthonormale la famille trigonométrique  $(e^{2i\pi n \cdot}, n \in \mathbb{Z})$ . Dans ces conditions, on définit le  $n$ -ième coefficient de Fourier d'un signal  $x$  de  $L^2_{\mathbb{C}}([0, 1[)$  par projection sur le  $n$ -ième vecteur de la base

$$\tilde{x}_n = \langle x, e^{2i\pi n \cdot} \rangle = \int_0^1 x(t) e^{-2i\pi n t} dt, \quad (\text{I.9})$$

et la série de Fourier  $\sum_{n \in \mathbb{Z}} \tilde{x}_n e^{2i\pi n t}$  converge vers  $x(t)$  dans  $L^2_{\mathbb{C}}([0, 1[)$  en moyenne quadratique, c'est-à-dire que la série est de carré sommable et que

$$\lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} \int_0^1 \left| \sum_{n=-M}^N \tilde{x}_n e^{2i\pi n t} - x(t) \right|^2 dt = 0.$$

En fait, si  $\sum_{n \in \mathbb{Z}} |\tilde{x}_n| < +\infty$ , on démontre aussi la convergence p.p. (c'est-à-dire pour presque tout  $t$ ) vers  $x(t)$  de sa série de Fourier, *i.e.*, on ne peut distinguer ces deux fonctions dans  $L^2_{\mathbb{C}}([0, 1[)$  :

$$x(t) = \sum_{n \in \mathbb{Z}} \tilde{x}_n e^{2i\pi n t}. \quad (\text{I.10})$$

L'expression (I.10) est une décomposition harmonique dont (I.9) est la formule d'inversion. Bien que ressemblant au couple (I.1)-(I.2) recherché, cette décomposition et son inverse ne se transposent pas sous la forme (I.1)-(I.2) sans soulever de nouvelles questions. Symboliquement, on peut obtenir la transposition grâce à la pseudo-fonction de Dirac  $\delta$ , dont la propriété essentielle réside en la formule symbolique suivante :

$$\int_{\mathbb{R}} \delta(u - a) \varphi(u) du = \varphi(a), \quad (\text{I.11})$$

qui donne

$$\int_{\mathbb{R}} \delta(\nu - n) e^{2i\pi \nu t} d\nu = e^{2i\pi n t},$$

ce qui permet d'identifier symboliquement la TF de  $x$  sous la forme d'un « peigne » de Dirac

$$\tilde{x}(\nu) = \sum_{n \in \mathbb{Z}} \tilde{x}_n \delta(\nu - n).$$

La théorie des distributions donne un sens mathématique à cette écriture formelle dans l'espace  $\mathcal{S}$  des distributions tempérées. La référence [4] donne une vision rigoureuse mais très accessible de la TF des distributions tempérées. La restriction de l'espace  $\mathcal{S}$  aux fonctions usuelles est très vaste : elle contient non seulement les espaces usuels  $L^p_{\mathbb{C}}(\mathbb{R})$ , mais aussi toutes les fonctions localement intégrables à croissance lente à l'infini (c'est-à-dire plus lente qu'un polynôme), et en particulier les fonctions périodiques de carré intégrable sur leur période, que l'on sait par ailleurs développer en série de Fourier (voir l'exemple 1 ci-dessus).

Dans le cadre de ce cours, nous nous contenterons d'une généralisation de la représentation fréquentielle (I.1) n'utilisant pas les distributions mais seulement la théorie de la mesure. Les signaux possédant cette représentation fréquentielle seront dits *harmonisables*.

Les deux exemples précédents suffisent à montrer qu'on ne peut se contenter d'écrire les formules (I.1) ou (I.2) en leur accordant une généralité qu'elles n'ont pas *a priori*, ni *a contrario* renoncer à ces représentations dès que les intégrales ne sont pas absolument convergentes. En fait la notion de représentation de Fourier, correctement envisagée, reste valable dans un grand nombre de situations.

Dans la présentation adoptée ici, on montre successivement l'existence et certaines propriétés de la TF d'un signal  $x$  :

— dans l'espace  $L^1_{\mathbb{C}}(\mathbb{R})$  des signaux *stables* à valeurs complexes, sans distinction entre des signaux égaux p.p. pour la mesure de Lebesgue :

$$x \in L^1_{\mathbb{C}}(\mathbb{R}) \iff \int_{\mathbb{R}} |x(t)| dt < +\infty; \quad (\text{I.12})$$

— dans l'espace  $L^2_{\mathbb{C}}(\mathbb{R})$  des signaux *d'énergie finie* à valeurs complexes, sans distinction entre des signaux égaux p.p. pour la mesure de Lebesgue :

$$x \in L^2_{\mathbb{C}}(\mathbb{R}) \iff \int_{\mathbb{R}} |x(t)|^2 dt < +\infty. \quad (\text{I.13})$$

— On définit ensuite la notion de signaux  $x$  harmonisables, c'est-à-dire auxquels on peut associer une mesure complexe  $\tilde{\mu}_x(d\nu)$  permettant une décomposition fréquentielle du type

$$x(t) = \int_{\mathbb{R}} e^{2i\pi\nu t} \tilde{\mu}_x(d\nu).$$

## I.1.2 Convolution

### Définition

En même temps que la TF, nous définirons la notion de *convolution*  $h * x$  de deux signaux  $h$  et  $x$  :

$$(h * x)(t) \triangleq \int_{\mathbb{R}} h(t-s) x(s) ds \quad (\text{I.14})$$

pour différents cas d'appartenance de  $h$  et  $x$  aux espaces  $L^1_{\mathbb{C}}(\mathbb{R})$  et  $L^2_{\mathbb{C}}(\mathbb{R})$ . D'un point de vue analytique, le problème sera là encore d'étudier la convergence d'une intégrale au sens de Lebesgue. Par changement de variable, remarquons que la convolution est commutative quand elle existe :

$$(h * x)(t) = \int_{\mathbb{R}} h(t-s) x(s) ds = \int_{\mathbb{R}} h(s') x(t-s') ds' = (x * h)(t).$$

### Lien avec le filtrage linéaire homogène

L'importance de la notion de convolution vient du fait qu'il s'agit d'un cas particulier de *filtrage linéaire homogène*, en considérant que  $h$  caractérise un filtre dont la sortie est  $(h * x)$  lorsque l'entrée est  $x$ . La linéarité est une conséquence immédiate de la linéarité de l'intégrale quand elle existe. Le terme homogène désigne l'invariance dans le temps : elle se déduit simplement en calculant la sortie pour une entrée décalée  $x_{\tau}(t) = x(t - \tau)$  : le résultat est la sortie  $(h * x)(t - \tau)$  décalée de la même quantité.

Cette remarque permet de construire un filtre linéaire homogène en spécifiant le signal  $h$ , constituant la *caractéristique temporelle* du filtre, ou encore sa *réponse impulsionnelle*, dénomination justifiée par le calcul formel de la sortie du filtre lorsque l'entrée est la pseudo-fonction de Dirac  $\delta$ , « impulsion » à l'instant  $t = 0$  :

$$(h * \delta)(t) = \int_{\mathbb{R}} h(s') \delta(t - s') ds' = h(t).$$

Notons que la spécification d'une réponse impulsionnelle caractérise un filtre linéaire homogène qui est seulement valable pour des entrées  $x$  telles que (I.14) existe.

## I.2 Cas des signaux stables

On considère ici les fonctions de l'espace de Banach (espace vectoriel normé complet)  $L_{\mathbb{C}}^1(\mathbb{R})$  comme des modèles de signaux. Par exemple, cet espace contient tous les signaux bornés de durée limitée à un intervalle de temps, tels le signal rectangle de l'exemple 1, et de façon plus générale, les signaux continus s'annulant à l'infini en décroissant plus vite que  $1/t$ . En restant dans cet espace, on pourra donc seulement modéliser et étudier des phénomènes transitoires, à l'exclusion de tout régime entretenu, périodique ou autre.

### I.2.1 Définition et propriétés de la TF dans $L_{\mathbb{C}}^1(\mathbb{R})$

#### Définition

L'espace  $L_{\mathbb{C}}^1(\mathbb{R})$  est celui dans lequel se définit le plus naturellement la TF (I.2), puisque cette dernière est une intégrale absolument convergente dans  $L_{\mathbb{C}}^1(\mathbb{R})$ , d'après la définition (I.12).

#### Propriétés

— La TF  $\tilde{x}$  de tout signal stable  $x$  est *bornée* :

$$|\tilde{x}(\nu)| \leq \int_{\mathbb{R}} |x(t)| dt < +\infty.$$

Dans le cas d'un signal stable positif, la borne est atteinte en  $\tilde{x}(0)$ .

—  $\tilde{x}$  est une fonction continue d'après le théorème de continuité issu de la convergence dominée de Lebesgue ( $x(t) e^{-2i\pi\nu t}$  est continue en  $\nu$  pour tout  $t$  fixé et majorée en module par  $|x(t)|$ ).

— Plus indirectement, on montre aussi que  $\tilde{x}$  tend vers zéro à l'infini. L'ensemble des fonctions continues nulles à l'infini est noté  $C_0(\mathbb{R})$  :  $\tilde{x} \in C_0(\mathbb{R})$ .

Malheureusement  $C_0(\mathbb{R})$  n'est pas inclus dans  $L_{\mathbb{C}}^1(\mathbb{R})$  ; les fonctions de  $C_0(\mathbb{R})$  sont seulement localement sommables (l'exemple 1 fournit un contre-exemple). On ne peut donc justifier la décomposition (I.1) aussi facilement que l'existence de la TF (I.2). Tout au plus peut-on assurer la validité de (I.1) dans le cas où  $\tilde{x}$  est stable. Dans ce cas, la TF inverse existe et elle est p.p. égale à  $x(t)$ , comme on pourra le montrer à l'aide d'une *fonction régularisante* dans l'Annexe B (ce procédé exploite les propriétés de la convolution dans  $L_{\mathbb{C}}^1(\mathbb{R})$  étudiées ci-dessous).

On en déduit l'*injectivité* de la TF des signaux stables : si deux signaux stables  $x$  et  $y$  ont même TF, ils sont égaux car leur différence est la TF inverse de la fonction nulle.

### I.2.2 Convolution dans $L^1_{\mathbb{C}}(\mathbb{R})$

Soient  $h$  et  $x$  deux signaux stables. Par application du théorème de Fubini,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |h(t-s)x(s)| dt ds = \int_{\mathbb{R}} |h(t)| dt \int_{\mathbb{R}} |x(t)| dt < +\infty, \quad (\text{I.15})$$

ce qui entraîne, pour presque tout  $t$ ,

$$\int_{\mathbb{R}} |h(t-s)x(s)| ds < +\infty.$$

Il est donc possible de définir pour presque tout  $t$  la fonction  $(h * x)(t) = \int_{\mathbb{R}} h(t-s)x(s) ds$ , stable d'après (I.15).

### Formule de convolution-multiplication

L'utilité pratique de la TF est en grande partie de faciliter l'étude du filtrage linéaire. Voyons le cas du filtrage convolutionnel stable. Considérons deux signaux  $x$  et  $h$  stables. Leur convolution  $y = h * x$  est stable, donc sa TF existe :

$$\begin{aligned} \tilde{y}(\nu) &= \int_{\mathbb{R}} \left( \int_{\mathbb{R}} h(t-s)x(s) ds \right) e^{-2i\pi\nu t} dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} h(t-s) e^{-2i\pi\nu(t-s)} x(s) e^{-2i\pi\nu s} ds dt \quad (\text{Fubini}) \\ &= \int_{\mathbb{R}} h(t) e^{-2i\pi\nu t} dt \int_{\mathbb{R}} x(s) e^{-2i\pi\nu s} ds \\ &= \tilde{h}(\nu) \tilde{x}(\nu), \end{aligned}$$

soit

$$\boxed{h * x \xrightarrow{\text{TF}} \tilde{h} \tilde{x} \quad \text{pour } x \text{ et } h \text{ stables.}} \quad (\text{I.16})$$

### Cas particulier : autocorrélation

Particularisons le résultat (I.16) en posant  $h(t) = x^*(-t)$  : d'après (I.6) et (I.7), on obtient

$$c_x(\tau) \triangleq \int_{\mathbb{R}} x(t)x^*(t-\tau) dt \xrightarrow{\text{TF}} |\tilde{x}(\nu)|^2.$$

La fonction d'*autocorrélation*  $c_x$  du signal  $x$  est définie presque partout dans  $L^1_{\mathbb{C}}(\mathbb{R})$  (que vaut  $c_x(0)$  ?). La fonction  $|\tilde{x}|^2$  est appelée *densité spectrale d'énergie* (ou spectre d'énergie) de  $x$ .

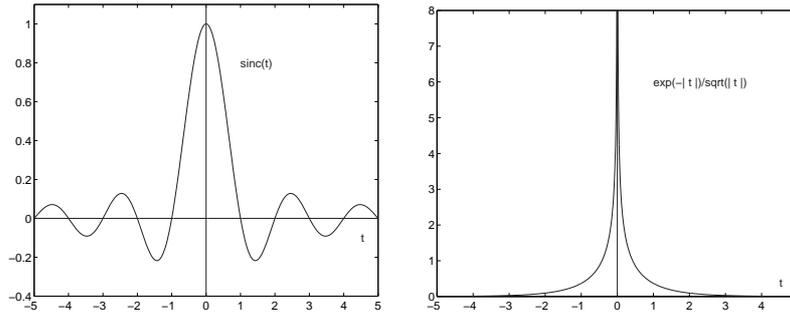
## I.3 Cas des signaux d'énergie finie

L'existence de la TF (I.2) des signaux stables est naturelle. En revanche, il est plus surprenant que la formulation inverse (I.1) ne soit pas toujours possible. Ce manque de symétrie frustrant disparaît quand on étudie la TF des signaux d'énergie finie, dont la définition est pourtant beaucoup moins directe.

Rappelons qu'il n'existe aucune relation d'inclusion entre  $L^1_{\mathbb{C}}(\mathbb{R})$  et  $L^2_{\mathbb{C}}(\mathbb{R})$ . Par exemple,

$$\begin{aligned} \text{sinc} &\in L^2_{\mathbb{C}}(\mathbb{R}) \quad \text{mais } \notin L^1_{\mathbb{C}}(\mathbb{R}) \\ e^{-|t|/\sqrt{|t|}} &\in L^1_{\mathbb{C}}(\mathbb{R}) \quad \text{mais } \notin L^2_{\mathbb{C}}(\mathbb{R}). \end{aligned}$$

Nous allons néanmoins pouvoir transporter dans  $L^2_{\mathbb{C}}(\mathbb{R})$  l'existence dans  $L^1_{\mathbb{C}}(\mathbb{R})$  de la TF et de la convolution, et montrer en plus que la formule d'inversion de la TF est systématique dans  $L^2_{\mathbb{C}}(\mathbb{R})$ .



### I.3.1 Définition et propriétés de la TF dans $L^2_{\mathbb{C}}(\mathbb{R})$

#### TF des signaux de $L^1_{\mathbb{C}}(\mathbb{R}) \cap L^2_{\mathbb{C}}(\mathbb{R})$

Soit  $x$  stable et d'énergie finie. Alors sa TF est d'énergie finie, égale à l'énergie du signal. C'est le théorème de Parseval

$$\int_{\mathbb{R}} |x(t)|^2 dt = \int_{\mathbb{R}} |\tilde{x}(\nu)|^2 d\nu, \quad (\text{I.17})$$

dont nous admettrons la démonstration, qui utilise à nouveau une fonction régularisante, voir Annexe A.

La TF est donc une application linéaire *isométrique* de  $L^1_{\mathbb{C}}(\mathbb{R}) \cap L^2_{\mathbb{C}}(\mathbb{R})$  vers  $L^2_{\mathbb{C}}(\mathbb{R})$ . Comme  $L^1_{\mathbb{C}}(\mathbb{R}) \cap L^2_{\mathbb{C}}(\mathbb{R})$  est dense dans  $L^2_{\mathbb{C}}(\mathbb{R})$ , espace complet, cette isométrie se prolonge par continuité en une application linéaire isométrique de  $L^2_{\mathbb{C}}(\mathbb{R})$  sur lui-même [6]. On continuera de noter  $\tilde{x}$  l'image de  $x$  par cette application, et on continuera de l'appeler la TF de  $x$ .

#### Propriétés de la TF dans $L^2_{\mathbb{C}}(\mathbb{R})$

— La TF dans  $L^2_{\mathbb{C}}(\mathbb{R})$  s'obtient comme une limite de TF d'éléments de  $L^1_{\mathbb{C}}(\mathbb{R}) \cap L^2_{\mathbb{C}}(\mathbb{R})$ . On ne peut donc pas écrire systématiquement (I.2) au sens de Lebesgue. En revanche,

$$x \xrightarrow{\text{TF}} \tilde{x}, \quad \tilde{x}(\nu) = \lim_{T \rightarrow +\infty} \int_{-T}^T x(t) e^{-2i\pi\nu t} dt$$

au sens  $L^2$ , c'est-à-dire en moyenne quadratique, et de la même façon, la TF est systématiquement inversible sous la forme

$$x(t) = \lim_{N \rightarrow +\infty} \int_{-N}^N \tilde{x}(\nu) e^{2i\pi\nu t} d\nu.$$

On se trouve dans le cas d'intégrales impropres au sens de Lebesgue, qui sont dites semi-convergentes [6]. En connaissance de cause, on conservera les écritures commodes (I.1) et (I.2).

— Quand on possède la TF  $\tilde{x}$  d'un signal  $x$  d'énergie finie, quelle est la TF du signal d'énergie finie  $\tilde{x}$ ? Il est facile de vérifier que le résultat est  $x(-t)$ . C'est pourquoi les tables de TF dans  $L^2_{\mathbb{C}}(\mathbb{R})$  peuvent se lire dans les deux sens. Par exemple, comme les fonctions régularisantes présentées en Annexe A sont paires, elles sont elles-mêmes la TF de leur TF.

#### Autocorrélation des signaux dans $L^2_{\mathbb{C}}(\mathbb{R})$

D'après l'inégalité de Schwarz dans  $L^2_{\mathbb{C}}(\mathbb{R})$ , les signaux d'énergie finie admettent une fonction d'autocorrélation :

$$|\langle x, x(\cdot - \tau) \rangle|^2 \leq \langle x, x \rangle \langle x(\cdot - \tau), x(\cdot - \tau) \rangle$$

c'est-à-dire

$$\left| \int_{\mathbb{R}} x(t)x^*(t-\tau) dt \right| \leq \int_{\mathbb{R}} |x(t)|^2 dt < +\infty,$$

soit encore  $|c_x(\tau)| \leq c_x(0) < +\infty$ .

Dans le cas d'un signal  $x$  à valeurs réelles, les égalités :

$$\begin{aligned} \|x + x(\cdot - \tau)\|^2 &= 2(c_x(0) + c_x(\tau)) \\ \|x - x(\cdot - \tau)\|^2 &= 2(c_x(0) - c_x(\tau)) \end{aligned}$$

permettent en premier lieu de retrouver l'inégalité  $|c_x(\tau)| \leq c_x(0)$  ; elles permettent aussi d'interpréter qualitativement la fonction d'autocorrélation. Plus  $|c_x(\tau)|$  est faible, plus le signal décalé  $x(\cdot - \tau)$  diffère du signal  $x$ , au sens de la norme de  $L^2_{\mathbb{R}}(\mathbb{R})$ .

A la différence du cas d'un signal stable, la fonction d'autocorrélation  $c_x$  d'un signal d'énergie finie n'est pas forcément stable (ni d'énergie finie). En revanche on peut montrer sa continuité uniforme sur  $\mathbb{R}$ . Comme dans le cas des signaux stables, existe-t-il encore un lien avec la densité spectrale d'énergie  $|\tilde{x}|^2$  ? Celle-ci est maintenant dans  $L^1_{\mathbb{C}}(\mathbb{R})$ , donc on peut calculer l'intégrale  $\int_{\mathbb{R}} |\tilde{x}(\nu)|^2 e^{2i\pi\nu t} d\nu$ . En utilisant une fonction régularisante, on peut montrer que le résultat est

$$\forall t, \quad \int_{\mathbb{R}} |\tilde{x}(\nu)|^2 e^{2i\pi\nu t} d\nu = c_x(t). \quad (\text{I.18})$$

Ce résultat généralise le théorème de Parseval (I.17) (qui s'en déduit pour  $t = 0$ ) et constitue une décomposition spectrale de type (I.1) pour la fonction d'autocorrélation.

### I.3.2 Convolution « $L^1_{\mathbb{C}}(\mathbb{R}) * L^2_{\mathbb{C}}(\mathbb{R})$ »

#### Existence

Soient  $h$  un signal stable et  $x$  un signal d'énergie finie. On montre l'inégalité

$$\int_{\mathbb{R}} \left( \int_{\mathbb{R}} |h(s)x(t-s)| ds \right)^2 dt < \left( \int_{\mathbb{R}} |h(t)| dt \right)^2 \int_{\mathbb{R}} x(t)^2 dt$$

qui implique l'existence de la convolution  $(h * x)(t)$  pour presque tout  $t$  et l'appartenance de  $h * x$  à  $L^2_{\mathbb{C}}(\mathbb{R})$ . On peut ensuite obtenir l'essentiel des résultats du § I.2.2.

#### Formule de convolution-multiplication

Le théorème (I.16) a été obtenu pour des signaux stables. Il est *a fortiori* valable quand  $x$  est dans  $L^1_{\mathbb{C}}(\mathbb{R}) \cap L^2_{\mathbb{C}}(\mathbb{R})$ , et  $h * x$  est alors également dans  $L^1_{\mathbb{C}}(\mathbb{R}) \cap L^2_{\mathbb{C}}(\mathbb{R})$ . Si  $x$  est seulement dans  $L^2_{\mathbb{C}}(\mathbb{R})$ , le signal tronqué  $x_T(t) = \mathbf{1}_{[-T,T]}(t) x(t)$  est stable et on obtient (1.15) dans  $L^2_{\mathbb{C}}(\mathbb{R})$  comme une égalité de limite :

$$\begin{array}{ccc} y_T = h * x_T & \xrightarrow{\text{TF}} & \tilde{y}_T = \tilde{h} \tilde{x}_T \\ & & \downarrow L^2 \quad \downarrow L^2 \\ & & \tilde{y} = \tilde{h} \tilde{x} \\ & & \text{p.p.} \end{array}$$

## I.4 Signaux harmonisables

Un signal  $x$  est dit *harmonisable* s'il est possible de le mettre sous la forme

$$x(t) = \int_{\mathbb{R}} e^{2i\pi\nu t} \tilde{\mu}_x(d\nu), \quad (\text{I.19})$$

où  $\tilde{\mu}_x$ , mesure harmonique de  $x$ , est une mesure sur  $\mathbb{R}$  à *valeurs complexes* de variation totale  $|\tilde{\mu}_x| = |\int_{\mathbb{R}} \tilde{\mu}_x(d\nu)| < +\infty$ .

Reprenons l'exemple de la décomposition des signaux périodiques en série de Fourier. Soit  $x$  un signal périodique de période  $T$ , stable sur sa période. On définit le  $n$ -ième coefficient de Fourier du signal  $x$  par

$$\tilde{x}_n = \frac{1}{T} \int_0^T s(t) e^{-2i\pi nt/T} dt.$$

Si  $\sum_{n \in \mathbb{Z}} |\tilde{x}_n| < +\infty$  alors, pour presque tout  $t$ ,  $x(t) = \sum_{n \in \mathbb{Z}} \tilde{x}_n e^{2i\pi nt/T}$ . Dans ce cas, le signal  $x$  est harmonisable et sa mesure harmonique est  $\tilde{\mu}_x(d\nu) = \sum_{n \in \mathbb{Z}} \tilde{x}_n \delta_{n/T}(d\nu)$ , où  $\delta_a$  désigne la mesure de Dirac au point  $a$ .

De même, les signaux de  $L^1_{\mathbb{C}}(\mathbb{R})$  dont la TF est stable admettent la représentation harmonique (I.1), comme le montre l'Annexe B. Ces signaux sont harmonisables, de mesure harmonique à densité par rapport à la mesure de Lebesgue :  $\tilde{\mu}_x(d\nu) = \tilde{x}(\nu) d\nu$ .

## I.5 Filtrage linéaire

La notion de filtre permet de représenter tout système physique permettant de transformer un signal d'entrée  $x$  en un signal de sortie  $y$ . Comme indiqué dans l'introduction, nous nous limiterons, dans ce paragraphe, à l'étude des filtres *linéaires homogènes*, c'est-à-dire satisfaisant les *principes de linéarité et d'invariance dans le temps*. Ceci signifie que si  $y_1$  est la sortie associée au signal d'entrée  $x_1$  et  $y_2$  la sortie associée à  $x_2$ , le signal de sortie associé à l'entrée  $\lambda_1 x_1 + \lambda_2 x_2$  est  $\lambda_1 y_1 + \lambda_2 y_2$  et que la sortie de  $x_1(t - \tau)$  est  $y_1(t - \tau)$ .

### I.5.1 Filtre convolutionnel

Cas  $L^1_{\mathbb{C}}(\mathbb{R})$  ou  $L^2_{\mathbb{C}}(\mathbb{R})$

On a déjà mentionné au § I.2 la possibilité de caractériser un filtre par la convolution avec une réponse impulsionnelle  $h$ . Au § I.2.2 et au § I.3.2, nous avons effectivement pu définir p.p. la sortie  $y = h * x$  d'un filtre convolutionnel *stable*, c'est-à-dire associé à un signal  $h$  stable, dans le cas où  $x$  est stable ou d'énergie finie. La sortie est alors respectivement stable ou d'énergie finie.

#### Filtrage des signaux périodiques

De la même façon, on peut considérer la convolution d'une réponse impulsionnelle  $h$  stable par une entrée  $x$  périodique et stable sur sa période. Alors il est facile de montrer que la sortie  $y = h * x$  est définie, périodique de période  $T$  et stable sur sa période. Ses coefficients de Fourier sont  $\tilde{y}_n = \tilde{h}(n/T) \tilde{x}_n$ .

#### Extension aux signaux harmonisables

Montrons que le filtrage convolutionnel stable des signaux harmonisables est encore possible. Soit un signal  $x$  harmonisable de mesure harmonique  $\tilde{\mu}_x$  et  $h$  un signal stable. Le signal  $y = h * x$  est défini p.p., car

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |h(t-s) e^{2i\pi \nu t} \tilde{\mu}_x(d\nu)| ds \leq |\tilde{\mu}_x| \int_{\mathbb{R}} |h(s)| ds < +\infty,$$

ce qui entraîne, pour presque tout  $t$  (Fubini),

$$\int_{\mathbb{R}} |h(t-s) x(s)| ds < +\infty.$$

D'autre part on vérifie facilement que  $y$  admet la représentation harmonique

$$y(t) = \int_{\mathbb{R}} e^{2i\pi\nu t} \tilde{h}(\nu) \tilde{\mu}_x(d\nu),$$

c'est-à-dire que la mesure harmonique de  $y$  est

$$\tilde{\mu}_y(d\nu) = \tilde{h}(\nu) \tilde{\mu}_x(d\nu). \quad (\text{I.20})$$

Ce résultat peut être interprété comme une généralisation de la formule de convolution-multiplication (I.16).

### Causalité

Si le signal  $h(t)$  est égal à 0 pour  $t < 0$ , le filtre est dit *causal*. Dans la pratique, pour des signaux temporels, ceci signifie que le filtre ne produit de réponse qu'après avoir été sollicité par une entrée. En effet, la relation de convolution (I.14) peut s'écrire pour  $h$  causal

$$(h * x)(t) = \int_0^\infty h(s') x(t - s') ds',$$

ce qui montre que la sortie  $(h * x)(t)$  du filtre ne dépend (linéairement) que de la valeur présente de l'entrée  $x(t)$  ainsi que de ses valeurs passées.

Comme la causalité dans le temps est une condition nécessaire évidente pour qu'un filtre temporel soit *réalisable* sous la forme d'un système physique, certains ouvrages (par exemple spécialisés en communication) limitent la notion de filtre à des réalisations causales. Pour plusieurs raisons, ce point de vue nous semble trop restrictif dans le cadre d'un cours « généraliste » :

— d'une part, la contrainte de causalité n'existe pas dans le cas du filtrage spatial : un système optique isotrope est un opérateur linéaire dans le cadre de l'approximation de Gauss et sa réponse impulsionnelle est symétrique, donc évidemment « non causale » ;

— d'autre part, on peut « réaliser » formellement la sortie d'un filtre temporel non causal en tolérant un retard entre l'entrée et la sortie. L'exemple du traitement numérique hors ligne est révélateur : une fois les échantillons d'un signal stockés dans la mémoire d'un ordinateur, leur filtrage non causal n'est pas moins *réalisable* qu'un filtrage causal. On peut même « remonter le temps » en traitant les échantillons dans l'ordre inverse de leur acquisition !

## I.5.2 Représentation fréquentielle

### Généralités

Considérons un filtre convolutionnel de réponse impulsionnelle  $h$  stable et notons  $\tilde{H}$  sa transformée de Fourier (notée  $\tilde{h}$  jusqu'à présent). D'après la formule de convolution-multiplication, on obtient pour un signal d'entrée stable  $x$  la relation entrée/sortie dans le domaine de Fourier :

$$\tilde{y} = \tilde{H}\tilde{x}.$$

Évaluons d'autre part la sortie du filtre dans le cas d'un signal d'entrée exponentiel complexe  $x(t) = e^{2i\pi\nu_0 t}$ , pour  $\nu_0$  élément de  $\mathbb{R}$  :

$$\begin{aligned} y(t) &= \int_{\mathbb{R}} h(s) e^{2i\pi\nu_0(t-s)} ds \\ &= e^{2i\pi\nu_0 t} \int_{\mathbb{R}} h(s) e^{-2i\pi\nu_0 s} ds \\ &= x(t) \tilde{H}(\nu_0). \end{aligned}$$

En d'autres termes, les « fréquences pures » sont les signaux propres des filtres convolutionnels. Le coefficient  $\tilde{H}(\nu)$  est appelé transmittance complexe du filtre à la fréquence  $\nu$  et on peut l'écrire sous la forme suivante :

$$\tilde{H}(\nu) = A(\nu) e^{i\phi(\nu)},$$

avec  $A(\nu) = |\tilde{H}(\nu)|$  le *gain d'amplitude* et  $\phi(\nu) = \arg \tilde{H}(\nu)$  le *déphasage* (subis par le signal  $x(t) = e^{2i\pi\nu t}$  au passage dans le filtre).

Réciproquement, on peut chercher à définir un filtre linéaire homogène par sa transmittance complexe  $\tilde{H}$ . Formellement, il suffit pour cela de restreindre l'ensemble des entrées admissibles aux signaux  $x$  harmonisables pour lesquels

$$y(t) = \int_{\mathbb{R}} e^{2i\pi\nu t} \tilde{H}(\nu) \tilde{\mu}_x(d\nu)$$

est définie p.p. La mesure harmonique de la sortie est alors

$$\tilde{\mu}_y(d\nu) = \tilde{H}(\nu) \tilde{\mu}_x(d\nu).$$

### Exemple 1

Un filtre passe-bande idéal est caractérisé par la transmittance  $\tilde{H}(\nu) = \mathbf{1}_{[-B_2, -B_1]}(\nu) + \mathbf{1}_{[B_1, B_2]}(\nu)$ , c'est-à-dire qu'il « coupe » toute l'énergie du signal située en dehors de l'intervalle (la « bande ») de fréquence  $[B_1, B_2]$ . On peut définir sa sortie pour toute entrée harmonisable. On définit de même des filtres coupe-bande, passe-haut ( $B_2 = +\infty$ ), passe-bas ( $B_1 = 0$ ), etc. On peut s'intéresser à la réponse impulsionnelle de ces filtres (pourquoi existe-t-elle?) : un cas particulier très utile est le filtre passe-bas « idéal » ; pour  $B_1 = 0$  et  $B_2 = B$ , la réponse impulsionnelle  $h$  s'écrit

$$h(t) = 2B \operatorname{sinc}(2Bt).$$

On note immédiatement que cette réponse impulsionnelle est non causale et de support infini, ce qui signifie que la réalisation physique d'un tel filtre ne peut être qu'approchée (c'est pourquoi on l'appelle filtre passe-bas « idéal »).

### Exemple 2

Le filtre dérivateur est caractérisé par la transmittance  $\tilde{H}(\nu) = 2i\pi\nu$ . On peut le définir sur l'ensemble des entrées qui se mettent sous la forme :

$$x(t) = \int_{\mathbb{R}} e^{2i\pi\nu t} s(\nu) d\nu,$$

telles que  $s$  et  $\nu \rightarrow \nu s(\nu)$  soient des éléments de  $L^1_{\mathbb{C}}(\mathbb{R})$ . De tels signaux sont continus et dérivables et on voit facilement que le filtre dérivateur associe à  $x$  sa dérivée  $dx/dt$  :

$$\frac{d}{dt} x(t) = \int_{\mathbb{R}} 2i\pi\nu e^{2i\pi\nu t} s(\nu) d\nu.$$

## Annexe A Fonctions régularisantes

### Définition

Une fonction régularisante est une fonction  $\theta$  stable et normalisée :

$$\int_{\mathbb{R}} \theta(t) dt = \tilde{\theta}(0) = 1,$$

et dont la TF est également stable. En pratique on choisit des fonctions réelles positives et paires. D'autre part une fonction régularisante est majorée, continue et décroissante vers 0 sur  $\mathbb{R}_+$ .

## Exemples

Les fonctions régularisantes les plus courantes sont celles de Féjer-Cesaro, Cauchy-Poisson et Gauss-Weierstrass respectivement associées aux trois couples :

$$\theta(t) = \sin^2 \pi t / \pi^2 t^2 \quad \xrightarrow{\text{TF}} \quad \tilde{\theta}(\nu) = (1 - |\nu|) \mathbf{1}_{[-1,1]}(\nu) \quad (\text{I.21})$$

$$\theta(t) = 2/(1 + 4\pi^2 t^2) \quad \xrightarrow{\text{TF}} \quad \tilde{\theta}(\nu) = e^{-|\nu|} \quad (\text{I.22})$$

$$\theta(t) = \sqrt{\pi} e^{-\pi^2 t^2} \quad \xrightarrow{\text{TF}} \quad \tilde{\theta}(\nu) = e^{-\nu^2} \quad (\text{I.23})$$

## Propriétés

La caractéristique intéressante de ces fonctions est d'admettre la décomposition (I.1) par vérification directe. On peut alors « propager » cette propriété à d'autres fonctions, dont les fonctions stables à TF stable, comme nous allons le montrer maintenant.

## Annexe B TF inverse de $\tilde{x}$ stable

L'objet de cette partie est de montrer que dans le cas où la TF  $\tilde{x}$  d'un signal  $x$  stable est stable, la TF inverse (I.1) existe et est p.p. égale à  $x$ . Si  $x$  est stable, on a d'après (I.6) et (I.16) :

$$\lambda \theta(\lambda t) * x(t) \xrightarrow{\text{TF}} \tilde{\theta}\left(\frac{\nu}{\lambda}\right) \tilde{x}(\nu), \quad (\text{I.24})$$

c'est-à-dire (I.2) pour  $\lambda \theta(\lambda t) * x(t)$ . Montrons inversement (I.1) pour cette fonction régularisée :

$$\int_{\mathbb{R}} \tilde{\theta}\left(\frac{\nu}{\lambda}\right) \tilde{x}(\nu) e^{2i\pi\nu t} d\nu = \lambda \theta(\lambda t) * x(t) \text{ pour tout } t. \quad (\text{I.25})$$

D'après (I.14), on a

$$\lambda \theta(\lambda t) * x(t) = \int_{\mathbb{R}} \lambda \theta(\lambda(t-u)) x(u) du.$$

Soit  $\theta(\lambda(t-u)) = \theta(\lambda u) e_u(t)$ , avec par exemple  $e_u(t) = \exp(-\pi^2 \lambda^2 t(t-2u))$  dans le cas du facteur de Gauss-Weierstrass. Alors :

$$\lambda \theta(\lambda t) * x(t) = \int_{\mathbb{R}} \lambda \theta(\lambda u) x(u) e_u(t) du.$$

Par TF, en utilisant d'une part (I.24) et le théorème de Fubini d'autre part, on obtient

$$\tilde{\theta}\left(\frac{\nu}{\lambda}\right) \tilde{x}(\nu) = \int_{\mathbb{R}} \lambda \theta(\lambda u) x(u) \left( \int_{\mathbb{R}} e_u(t) e^{-2i\pi\nu t} dt \right) du = \int_{\mathbb{R}} \lambda \theta(\lambda u) x(u) \tilde{e}_u(\nu) du.$$

Toujours grâce au théorème de Fubini, en remarquant que  $|\lambda \theta(\lambda u) x(u) \tilde{e}_u(\nu)|$  est intégrable en  $du d\nu$ , on en déduit :

$$\begin{aligned} \int_{\mathbb{R}} \tilde{\theta}\left(\frac{\nu}{\lambda}\right) \tilde{x}(\nu) e^{2i\pi\nu t} d\nu &= \int_{\mathbb{R}} \lambda \theta(\lambda u) x(u) \left( \int_{\mathbb{R}} \tilde{e}_u(\nu) e^{2i\pi\nu t} d\nu \right) du \\ &= \int_{\mathbb{R}} \lambda \theta(\lambda u) x(u) e_u(t) du \\ &= \lambda \theta(\lambda t) * x(t). \end{aligned}$$

Dans le cas où  $\tilde{x}$  est également stable, le théorème de convergence dominée entraîne que le premier terme de (I.25) tend vers

$$\int_{\mathbb{R}} \tilde{x}(\nu) e^{2i\pi\nu t} d\nu$$

pour tout  $t$  quand  $\lambda$  tend vers  $+\infty$ , uniformément dans tout intervalle. D'autre part, si l'on peut démontrer que le deuxième membre de (I.25) tend simultanément vers  $x$  dans  $L^1_{\mathbb{C}}(\mathbb{R})$ , alors on aura la formule d'inversion (I.1) pour presque tout  $t$ . Pour démontrer ce dernier point, on forme en utilisant (I.24)

$$\int_{\mathbb{R}} |\lambda \theta(\lambda t) * x(t) - x(t)| dt = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} (x(t-u) - x(t)) \lambda \theta(\lambda u) du \right| dt.$$

Donc, en posant  $f(u) = \int_{\mathbb{R}} |x(t-u) - x(t)| dt$ ,

$$\int_{\mathbb{R}} |\lambda \theta(\lambda t) * x(t) - x(t)| dt \leq \int_{\mathbb{R}} f(u) \lambda \theta(\lambda u) du = \int_{\mathbb{R}} f\left(\frac{u'}{\lambda}\right) \theta(u') du'.$$

Comme  $f(u) = |f(u)|$  est majorée par  $2 \int_{\mathbb{R}} |x(t)| dt$ , le théorème de Lebesgue donne

$$\lim_{\lambda \rightarrow +\infty} \int_{\mathbb{R}} f\left(\frac{u'}{\lambda}\right) \theta(u') du' = \int_{\mathbb{R}} \lim_{\lambda \rightarrow +\infty} f\left(\frac{u'}{\lambda}\right) \theta(u') du' = 0$$

pourvu que  $f(u)$  tende vers zéro quand  $u$  tend vers zéro. D'après le théorème de Lebesgue, c'est vrai si  $x$  est continue à support compact. On passe ensuite au cas général en approchant dans  $L^1_{\mathbb{C}}(\mathbb{R})$  les fonctions stables par des fonctions continues à support compact.

### Remarque

Même si  $\tilde{x}$  n'est pas stable, on conserve quand même un résultat de convergence dans  $L^1_{\mathbb{C}}(\mathbb{R})$  :

$$\lim_{\lambda \rightarrow +\infty} \int_{\mathbb{R}} \tilde{x}(\nu) \tilde{\theta}\left(\frac{\nu}{\lambda}\right) e^{2i\pi\nu t} d\nu = x(t)$$

pour presque tout  $t$ , résultat valable pour les différentes fonctions régularisantes  $\theta$ .



## Chapitre II

# Signaux déterministes à temps discret

### II.1 Introduction

Les signaux à temps discret tiennent une place prépondérante dans les applications dans la mesure où de plus en plus de traitements sont effectués par voie numérique. Nous insisterons donc particulièrement dans ce cours sur ce type de signal, que ce soit dans le cadre déterministe ou aléatoire. Dans la pratique, une grande partie des signaux à temps discret correspondent à un échantillonnage d'un signal continu, mais il existe des processus qui sont par nature à valeurs discrètes comme par exemple des images reçues sur des mosaïques CCD.

Dans ce chapitre, nous étudierons en premier lieu le problème de la transformation d'un signal continu sous forme discrète. On utilise en général un échantillonnage régulier, par lequel on associe à un signal  $x$  la suite de valeurs numériques  $x_k = x(kT_e)$ ,  $k \in \mathbb{Z}$ , où  $T_e$  est le *pas* (ou *période*) d'échantillonnage. La première question à se poser lors de cette transformation est : dans quelle mesure la suite des échantillons  $\{x_k\}_{k \in \mathbb{Z}}$  décrit parfaitement le signal initial  $x$ ? Le *théorème de Shannon* permet de définir les conditions de conservation de l'information véhiculée par  $x$  dans le signal à temps discret<sup>1</sup>  $\{x_k\}_{k \in \mathbb{Z}}$ . Après avoir abordé cette question, nous présentons les outils les plus couramment utilisés pour les signaux à temps discret. Il s'agit de la *transformée en  $z$*  et de la *transformée de Fourier discrète*. Comme dans le cas continu, on se placera dans les deux cadres suivants :

— l'ensemble (espace de Banach) des signaux à temps discret stables

$$\ell^1 = \left\{ \{x_k\}_{k \in \mathbb{Z}} / \sum_{k \in \mathbb{Z}} |x_k| < +\infty \right\},$$

— l'espace de Hilbert des signaux à temps discret d'énergie finie,

$$\ell^2 = \left\{ \{x_k\}_{k \in \mathbb{Z}} / \sum_{k \in \mathbb{Z}} |x_k|^2 < +\infty \right\}.$$

Signalons la relation d'inclusion  $\ell^1 \subset \ell^2$  qui n'a pas cours dans le cas des signaux continus.

---

<sup>1</sup>Dans ce chapitre, un signal à temps discret est noté  $\{x_k\}_{k \in \mathbb{Z}}$  pour le distinguer du signal à temps continu  $x$ . Dans la suite du cours, les signaux à temps continu ne seront plus utilisés et la notation  $x$  désignera le signal à temps discret.

### II.1.1 Formule sommatoire de Poisson

Cette formule est utilisée en analyse pour calculer simplement des sommes de séries. Dans le cas du traitement du signal, cette formule est liée au théorème d'échantillonnage que nous allons présenter dans le paragraphe suivant et peut, d'autre part, être utilisée pour des calculs de spectres de signaux complexes.

#### Théorème

Soit  $s$  un signal complexe stable de transformée de Fourier  $\tilde{s}$  stable et soit  $T$  un réel positif quelconque. Alors pour presque tout  $(u, \nu) \in \mathbb{R}^2$

$$T \sum_{n \in \mathbb{Z}} s(nT - u) e^{2i\pi\nu n} = e^{2i\pi\nu u/T} \sum_{n \in \mathbb{Z}} \tilde{s}\left(\frac{n - \nu}{T}\right) e^{-2i\pi\nu u/T}. \quad (\text{II.1})$$

#### Démonstration

Sachant que  $\int_0^T \sum_{n \in \mathbb{Z}} |s(nT - u)| du = \int_{\mathbb{R}} |s(u)| du < +\infty$ , la fonction périodique  $\sum_{n \in \mathbb{Z}} |s(nT - u)|$  est presque partout finie. Donc

$$F(u, \nu) = \sum_{n \in \mathbb{Z}} s(nT - u) e^{2i\pi\nu(nT - u)/T} \quad (\text{II.2})$$

est une fonction définie presque partout en  $u$ , périodique et intégrable sur  $[0, T]$ .

Pour tout  $k \in \mathbb{Z}$ ,

$$\begin{aligned} \frac{1}{T} \int_0^T F(u, \nu) e^{-2i\pi k u/T} du &= \sum_{n \in \mathbb{Z}} \int_0^T s(nT - u) e^{2i\pi(\nu + k)(nT - u)/T} \\ &= \frac{1}{T} \tilde{s}\left(\frac{-\nu - k}{T}\right). \end{aligned}$$

Or, la transformée de Fourier est stable, donc pour presque tout  $\nu$ ,  $\sum_{k \in \mathbb{Z}} |\tilde{s}(\frac{-\nu - k}{T})| < +\infty$ . On peut donc utiliser les transformées de Fourier inverses et montrer que pour presque tout  $(u, \nu)$ ,

$$F(u, \nu) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \tilde{s}\left(\frac{k - \nu}{T}\right) e^{-2i\pi u k/T}.$$

En faisant  $u = \nu = 0$  dans (II.1), sous réserve d'existence, on obtient la *formule sommatoire de Poisson*,

$$\boxed{T \sum_{n \in \mathbb{Z}} s(nT) = \sum_{n \in \mathbb{Z}} \tilde{s}(n/T)} \quad (\text{II.3})$$

Notons que pour que (II.3) soit vrai, il faut que la série  $s(nT)$  soit absolument convergente, ce qui n'est pas équivalent à  $s$  stable.

## II.2 Théorème d'échantillonnage

### II.2.1 Présentation

Soit  $x$  un signal stable continu, dont la transformée de Fourier possède un support limité à la bande de fréquence  $[-B, B]$ . Cette transformée de Fourier est continue et puisqu'elle est à support borné, elle est intégrable et

$$\forall t \in \mathbb{R}, \quad x(t) = \int_{-B}^B \tilde{x}(\nu) e^{2i\pi\nu t} d\nu.$$

On dit alors que le signal  $x$  est à *bande limitée* sur  $[-B, B]$ .

En utilisant la formule de Poisson (II.1) en  $u = 0$  et  $\nu$  p.p., ce qui exige que

$$\sum_{n \in \mathbb{Z}} \left| x \left( \frac{n}{2B} \right) \right| < +\infty,$$

on peut écrire, en remplaçant  $\nu$  par  $-T\nu$  et  $T$  par  $1/2B$  dans (II.1),

$$\frac{1}{2B} \sum_{n \in \mathbb{Z}} x \left( \frac{n}{2B} \right) e^{-i\pi\nu n/B} = \sum_{n \in \mathbb{Z}} \tilde{x}(\nu + 2Bn) = \tilde{x}(\nu). \quad (\text{II.4})$$

Donc

$$x(t) = \int_{-B}^B \frac{1}{2B} \sum_{n \in \mathbb{Z}} x \left( \frac{n}{2B} \right) e^{-i\pi\nu n/B} e^{2i\pi\nu t} d\nu,$$

soit

$$x(t) = \frac{1}{2B} \sum_{n \in \mathbb{Z}} x \left( \frac{n}{2B} \right) \int_{-B}^B e^{-i\pi\nu n/B} e^{2i\pi\nu t} d\nu,$$

ou encore

$$x(t) = \frac{1}{2B} \sum_{n \in \mathbb{Z}} x \left( \frac{n}{2B} \right) h \left( t - \frac{n}{2B} \right),$$

où  $h$  est la réponse impulsionnelle du filtre passe-bas idéal de bande  $[-B, B]$  vue au chapitre précédent.

### II.2.2 Théorème de Shannon-Nyquist

Soit  $x$  un signal stable continu à bande limitée sur  $[-B, B]$ . Si on choisit une période d'échantillonnage  $T_e$  telle que  $T_e < 1/2B$ , et si

$$\sum_{n \in \mathbb{Z}} |x(nT_e)| < +\infty,$$

alors il est possible de reconstruire sans erreur le signal  $x$  à partir de ses échantillons  $x_n = x(nT_e)$  à l'aide de la formule suivante, dite *formule d'interpolation de Shannon* :

$$\boxed{x(t) = \sum_{n \in \mathbb{Z}} x(nT_e) \operatorname{sinc}(t/T_e - n)} \quad (\text{II.5})$$

#### Remarque 1

La position absolue des instants d'échantillonnage n'a pas d'importance ; c'est uniquement la cadence qui importe.

#### Remarque 2

Le théorème de reconstruction, énoncé pour  $K = 0$ , reste valable pour tout signal réel stable continu dont la TF est à support limité à des intervalles symétriques de type

$$\left[ \frac{-2K-1}{2T_e}, \frac{-2K+1}{2T_e} \right] \cup \left[ \frac{2K-1}{2T_e}, \frac{2K+1}{2T_e} \right],$$

pour tout  $K$  entier, pourvu que  $K$  soit connu. On obtient alors une version « haute fréquence » du théorème d'échantillonnage,

$$x_K(t) = x_0(t) \cos(2\pi K/T_e),$$

particulièrement utile pour l'échantillonnage des signaux modulés à haute fréquence.

### Interprétation du théorème d'échantillonnage. Phénomène de repliement

Lorsque les conditions du théorème d'échantillonnage ne sont pas satisfaites, il est important de comprendre qualitativement quel phénomène rend inapplicable la formule d'interpolation de Shannon.

Considérons donc un signal  $x$  qui n'est pas à bande limitée sur  $[-B, B]$ , et dont l'échantillonnage est effectué à la fréquence  $2B$ . Pour reconstruire le signal, on filtre le train d'impulsions obtenu par le filtre passe bas  $[-B, B]$ . Le signal reconstruit, après division par  $2B$ , est donc

$$\hat{x}(t) = \frac{1}{2B} \sum_{n \in \mathbb{Z}} x\left(\frac{n}{2B}\right) \int_{-B}^B e^{-2i\pi\nu n/2B} e^{2i\pi\nu t} d\nu = \frac{1}{2B} \int_{-B}^B e^{2i\pi\nu t} \left( \sum_{n \in \mathbb{Z}} x\left(\frac{n}{2B}\right) e^{-2i\pi\nu n/2B} \right) d\nu.$$

D'après la formule sommatoire de Poisson, l'expression entre parenthèses peut également s'écrire

$$\frac{1}{2B} \sum_{n \in \mathbb{Z}} x\left(\frac{n}{2B}\right) e^{-2i\pi\nu n/2B} = \sum_{n \in \mathbb{Z}} \tilde{x}(\nu - 2Bn)$$

ce qui, reporté dans l'expression précédente, permet d'écrire

$$\hat{x}(t) = \int_{-B}^B e^{2i\pi\nu t} \sum_{n \in \mathbb{Z}} \tilde{x}(\nu - 2Bn) d\nu.$$

La transformée de Fourier du signal reconstruit est donc

$$\tilde{\hat{x}}(\nu) = \sum_{n \in \mathbb{Z}} \tilde{x}(\nu - 2Bn) \mathbf{1}_{[-B, +B]}(\nu).$$

Cette transformée de Fourier est obtenue en superposant les spectres obtenus par décalage de multiples de  $2B$  à partir du spectre initial sur la plage de fréquences  $[-B, B]$ . Ce phénomène est dénommé *recouvrement de spectre* ou *aliasing*.

L'opération d'échantillonnage est illustrée par la figure II.1. L'encadré représente le signal  $x$  et sa transformée de Fourier à support limité sur  $[-B, B]$ .

— Dans la partie supérieure de la figure, l'échantillonnage du signal  $x$  est effectué à une cadence  $T_e < 1/2B$ . Dans le domaine de Fourier, conformément à (II.5),  $\tilde{g}(\nu)$  est obtenue par simple « périodisation » de  $\tilde{x}$ . Étant donné que  $2B < 1/T_e$ , il n'y a pas de chevauchement et on peut retrouver  $\tilde{x}$  en effectuant un filtrage passe-bas idéal. Dans le domaine temporel, l'équivalent de cette opération est la convolution par un sinus cardinal.

— Dans la partie inférieure de la figure, l'échantillonnage du signal  $x$  est effectué à une cadence  $T'_e > 1/2B$ . Dans ce cas, le calcul de  $\tilde{g}$  fait intervenir un *repliement* des motifs répliqués de  $\tilde{x}$  : un filtrage passe-bas idéal ne restituera pas le signal  $x$  original.

### Filtre anti-repliement

Il est d'usage, lorsque l'on échantillonne un signal  $x$  à une cadence donnée  $T_e$ , de filtrer le signal au préalable avec un filtre passe-bas de fréquence de coupure  $1/2T_e$ , ce qui permet d'éviter les phénomènes de repliement. Notons  $x_{BF}$  le signal filtré, qui vérifie les hypothèses de Shannon pour l'échantillonnage à une cadence de  $T_e$  et peut donc être reconstruit sans erreur à partir de ses échantillons  $\{x_{BF}(nT_e)\}_{n \in \mathbb{Z}}$ . Le signal  $x_{BF}$  ne contient que les composantes basse fréquence du signal original, mais constitue cependant une approximation du signal  $x$  meilleure que le signal  $\hat{x}$  reconstruit à partir des échantillons  $\{x(nT_e)\}_{n \in \mathbb{Z}}$  (voir paragraphe précédent), au sens de la norme dans  $L^2_{\mathbb{C}}(\mathbb{R})$ . En effet, on démontre (simplement, en utilisant Parseval) l'inégalité :

$$\int_{\mathbb{R}} |x(t) - x_{BF}(t)|^2 dt \leq \int_{\mathbb{R}} |x(t) - \hat{x}(t)|^2 dt.$$

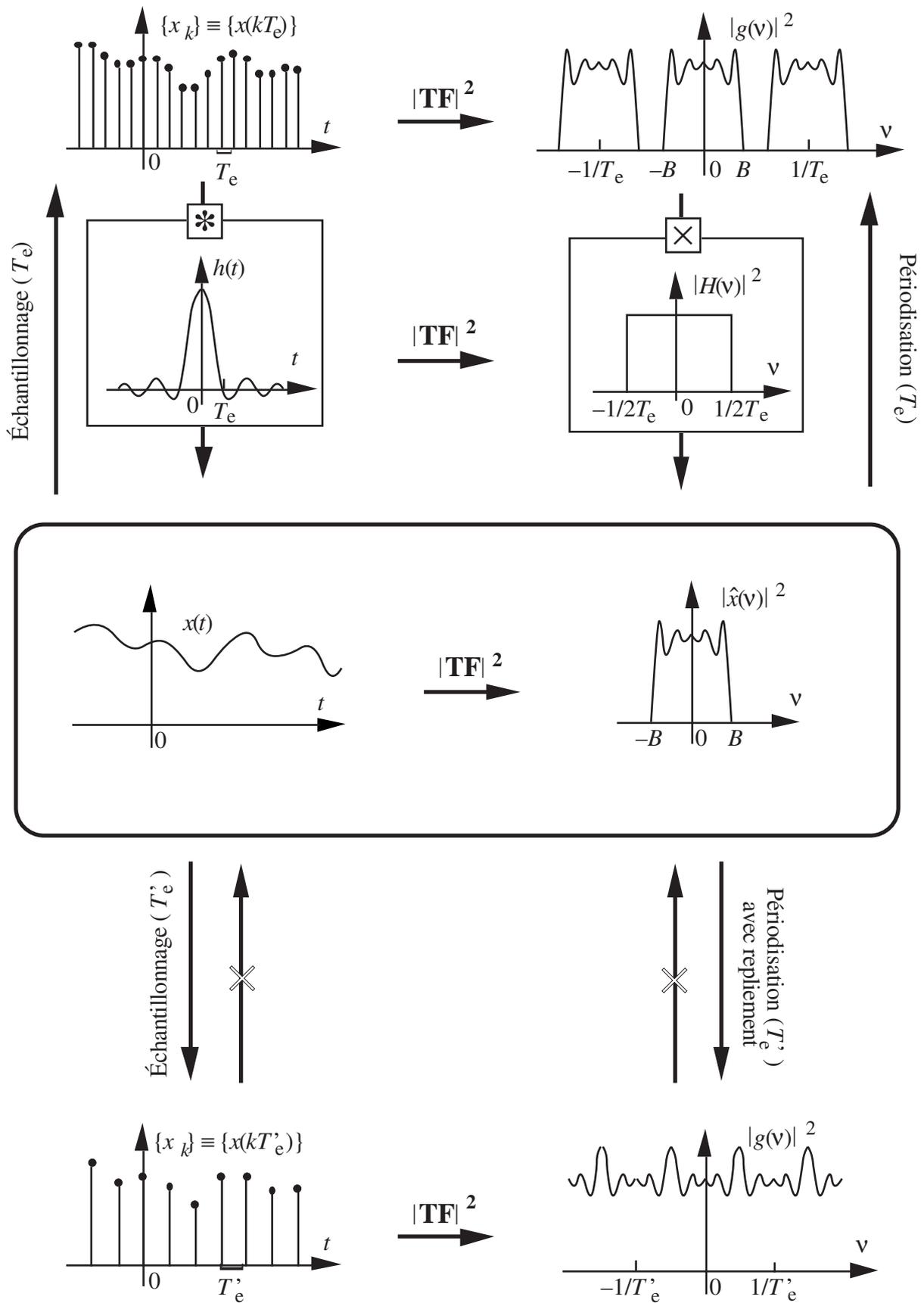


Fig. II.1. Échantillonnage d'un signal et repliement de sa TF (voir texte).

## II.3 Transformée en $z$

### II.3.1 Définition

La transformée en  $z$  (TZ) est l'analogie discrète de la transformée de Laplace des signaux continus. Elle sera particulièrement utilisée dans le chapitre suivant consacré aux filtres numériques. En effet, la transformée en  $z$  d'un filtre permet d'étudier de façon simple sa stabilité. Elle fournit également une représentation compacte des filtres AR ou ARMA qui sont fréquemment utilisés en traitement du signal.

On associe formellement à tout signal discret  $\{x_k\}_{k \in \mathbb{Z}}$  sa TZ définie par

$$X(z) = \sum_{k \in \mathbb{Z}} x_k z^{-k}, \quad z \in \mathbb{C}. \quad (\text{II.6})$$

Pour un signal donné, l'ensemble des valeurs de  $z$  pour lesquelles la série converge est la *région de convergence* (RDC). Pour la déterminer, on peut utiliser le critère de Cauchy sur la convergence des séries de puissance. On décompose la somme précédente sous la forme  $X(z) = X_1(z) + X_2(z)$  avec

$$X_1(z) = \sum_{k=1}^{+\infty} x_{-k} z^k \quad \text{et} \quad X_2(z) = \sum_{k=0}^{+\infty} x_k z^{-k}.$$

On montre alors que  $X_2(z)$  converge si  $|z| > R_{x-}$ , où

$$R_{x-} = \limsup_{k \rightarrow +\infty} |x_k|^{1/k}.$$

D'une manière similaire on peut montrer que  $X_1(z)$  converge si  $|z| < R_{x+}$ , où  $R_{x+}$  est défini par

$$\frac{1}{R_{x+}} = \limsup_{k \rightarrow +\infty} |x_{-k}|^{1/k}.$$

Dans le cas général, la série converge dans un anneau du plan complexe défini par

$$0 \leq R_{x-} < |z| < R_{x+} \leq +\infty.$$

Le problème suivant consiste à déterminer un signal connaissant sa TZ et la RDC associée. Pour cela on s'appuie sur des résultats de l'étude des fonctions de la variable complexe [2]. L'expression (II.6) est un développement en série de Laurent de la fonction  $X(z)$ , valable dans une couronne de convergence. Par conséquent la représentation en TZ du signal  $\{x_k\}_{k \in \mathbb{Z}}$  est la donnée de la fonction  $X(z)$  et de la RDC sur laquelle le développement en série de Laurent de  $X(z)$  restitue les coefficients  $\{x_k\}_{k \in \mathbb{Z}}$  :

$$\{x_k\}_{k \in \mathbb{Z}} \xrightarrow{TZ} X(z) \quad \text{et} \quad \text{RDC} = \{z / R_{x-} < |z| < R_{x+}\} \quad (\text{II.7})$$

Le calcul d'un signal à partir de sa TZ est donc un problème de développement en série de Laurent. On rappelle rapidement la démarche, qui fait appel au théorème de Cauchy sur l'intégration le long d'un contour dans le plan complexe. De ce théorème, on peut évaluer l'intégrale  $I$  :

$$I = \frac{1}{2i\pi} \oint_{\Gamma^+} z^{k-1} dz = \delta_k$$

où  $\Gamma^+$  est un contour fermé qui entoure l'origine des  $z$  parcouru dans le sens trigonométrique.  $\{\delta_k\}$  est le signal de Kronecker défini par  $\delta_0 = 1$  et  $\delta_k = 0$  si  $k \neq 0$ . En multipliant les deux membres de l'équation définissant la transformée en  $z$  par  $z^{l-1}/2i\pi$  et en intégrant le long d'un contour  $\Gamma^+$  entourant l'origine et contenu dans la RDC, on a

$$\frac{1}{2i\pi} \oint_{\Gamma^+} X(z) z^{k-1} dz = \frac{1}{2i\pi} \oint_{\Gamma^+} \sum_{l \in \mathbb{Z}} x_l z^{-l+k-1} dz.$$

L'intégrale étant calculée dans la RDC, la série est absolument convergente et on peut donc inverser l'intégration et la sommation, ce qui s'écrit

$$\frac{1}{2i\pi} \oint_{\Gamma^+} X(z) z^{k-1} dz = \sum_{-\infty}^{+\infty} \frac{x_l}{2i\pi} \oint_{\Gamma^+} z^{-l+k-1} dz.$$

Compte tenu de la remarque sur l'intégrale  $I$ , on obtient finalement la formule d'inversion de la TZ

$$x_k = \frac{1}{2\pi i} \oint_{\Gamma^+} X(z) z^{k-1} dz. \quad (\text{II.8})$$

Le calcul effectif de l'intégrale peut se faire de plusieurs manières, par exemple grâce à la méthode des résidus (voir [2]). Le terme  $x_k$  s'exprime alors comme la somme des résidus de la fonction  $X(z) z^{k-1}$  pour les pôles intérieurs au contour  $\Gamma^+$ . Lorsque  $k$  est négatif, on peut préférer sommer les résidus aux pôles extérieurs au contour  $\Gamma^+$ , ce qui revient à le parcourir dans le sens opposé au sens trigonométrique (et  $x_k$  est alors l'opposé de la somme de ces résidus). Cette pratique est intéressante lorsque  $|X(z) z^{k-1}|$  décroît suffisamment vite à l'infini.

### II.3.2 Propriétés de la transformée en $z$

— On appelle *pôles* de  $X(z)$ , les valeurs de  $z$  pour lesquelles la transformée  $X(z)$  tend vers l'infini, et *zéros* de  $X(z)$  les valeurs de  $z$  pour lesquelles  $X(z)$  s'annule.

— Soit un signal  $\{x_k\}_{k \in \mathbb{Z}}$  et soit la version décalée  $y_k = x_{k-k_0}$ , qui correspond à l'introduction d'un *retard*  $k_0$ . En prenant la transformée en  $z$ , on obtient :

$$Y(z) = z^{-k_0} X(z),$$

sous réserve que  $z$  appartienne à la RDC de  $\{x_k\}_{k \in \mathbb{Z}}$ .

— Le *produit de convolution à temps discret*  $x * h$  de deux signaux  $\{x_k\}_{k \in \mathbb{Z}}$  et  $\{h_k\}_{k \in \mathbb{Z}}$  est défini par

$$y_k = \sum_{n \in \mathbb{Z}} x_n h_{k-n}.$$

Son élément neutre est le signal de Kronecker  $\{\delta_k\}$  défini au paragraphe précédent. Notons qu'au contraire de la pseudo-fonction de Dirac du cas continu, ce signal ne pose aucun problème de définition et qu'il est élément de  $\ell^1$ . Comme dans le cas continu, sous certaines conditions, la transformée en  $z$  établit une correspondance entre produit de convolution et produit : la transformée en  $z$  de  $\{y_k\}_{k \in \mathbb{Z}}$  s'écrit

$$Y(z) = \sum_{k \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} x_n h_{k-n} z^{-k}.$$

Si  $z$  appartient à l'intersection des RDC de chacune de ces séries et en posant  $m = k - n$ , il vient

$$Y(z) = \sum_{n \in \mathbb{Z}} x_n z^{-n} \sum_{m \in \mathbb{Z}} h_m z^{-m} = X(z) H(z),$$

*i.e.*, le produit de convolution de deux signaux conduit à un produit simple par transformée en  $z$ . Notons que la RDC de  $\{y_k\}_{k \in \mathbb{Z}}$  peut s'avérer plus grande que l'intersection des RDC de  $\{x_k\}_{k \in \mathbb{Z}}$  et de  $\{h_k\}_{k \in \mathbb{Z}}$ .

## II.4 Transformée de Fourier à temps discret

### II.4.1 Lien entre transformée en $z$ et transformée de Fourier

Pour un signal à temps discret  $\{x_k\}_{k \in \mathbb{Z}}$ , la transformée de Fourier  $\tilde{X}(\nu)$  correspond à l'expression de sa transformée en  $z$  prise en  $z = e^{2i\pi\nu}$  :

$$\tilde{X}(\nu) = \sum_{k \in \mathbb{Z}} x_k e^{-2i\pi\nu k}, \quad \nu \in \mathbb{R}. \quad (\text{II.9})$$

Son existence dépend de la convergence absolue de la série (II.9) sur le cercle unité : si le signal  $\{x_k\}_{k \in \mathbb{Z}}$  est dans  $\ell^1$ , la RDC de la série (II.9) contient le cercle unité et la transformée de Fourier existe. Notons qu'alors la fonction  $\tilde{X}$  est périodique de période 1.

#### Formule d'inversion

Contrairement au cas continu, la stabilité permet de garantir l'existence d'une transformée de Fourier inverse. En multipliant (II.9) par  $e^{2i\pi n\nu}$  et en intégrant sur  $[0, 1]$ , on obtient  $x_n$  comme le  $n$ -ième coefficient de Fourier de  $\tilde{X}$  :

$$x_n = \int_0^1 \tilde{X}(\nu) e^{2i\pi n\nu} d\nu. \quad (\text{II.10})$$

#### Égalité de Plancherel-Parseval

L'analogue du théorème de Parseval vu au chapitre I, (I.16) existe en discret. Soit  $\{x_k\}_{k \in \mathbb{Z}}$  appartenant à  $\ell^2$  muni du produit hermitien

$$\langle x, y \rangle = \sum_{k \in \mathbb{Z}} x_k y_k^*.$$

Soit  $\tilde{X}(\nu)$  la transformée de Fourier associée à  $x_n$ , alors

$$\sum_{n \in \mathbb{Z}} |x_n|^2 = \int_0^1 |\tilde{X}(\nu)|^2 d\nu. \quad (\text{II.11})$$

### II.4.2 Transformée de Fourier discrète (TFD)

Comme dans le cas continu, on peut s'intéresser à la représentation de Fourier des signaux discrets *périodiques*  $\{x_k\}_{k \in \mathbb{Z}}$  de période  $N$ . Dans le cas discret, cette représentation est d'un intérêt pratique considérable car on dispose d'algorithmes rapides qui permettent de passer de la représentation temporelle  $\{x_k\}_{k=0, N-1}$  à une représentation fréquentielle équivalente  $\{X_l\}_{l=0, N-1}$ , et réciproquement.

Pour un signal discret périodique  $\{x_k\}_{k \in \mathbb{Z}}$  de période  $N$ , on définit sa TFD par

$$X_l = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{-2i\pi kl/N}, \quad l \in \mathbb{Z}. \quad (\text{II.12})$$

On obtient ainsi un nouveau signal discret  $\{X_l\}_{l \in \mathbb{Z}}$ , également périodique de période  $N$ .

De façon duale, on montre la formule d'inversion

$$x_k = \sum_{l=0}^{N-1} X_l e^{2i\pi kl/N}, \quad k \in \mathbb{Z}. \quad (\text{II.13})$$

Le signal  $\{X_l\}_{l \in \mathbb{Z}}$  est donc une représentation équivalente de  $\{x_k\}_{k \in \mathbb{Z}}$ . Du fait de la périodicité, la connaissance de  $N$  points successifs suffit à définir entièrement chacun de ces signaux. Les transformations linéaires (II.12) et (II.13) peuvent donc être résumées par la relation matricielle liant les vecteurs  $\mathbf{x} = [x_0, \dots, x_{N-1}]^t$  et  $\mathbf{X} = [X_0, \dots, X_{N-1}]^t$  :

$$\mathbf{X} = \frac{1}{N} W_N \mathbf{x}, \quad (\text{II.14})$$

où  $W_N$  est une matrice inversible d'éléments  $w_N^{kl} = e^{-2i\pi kl/N}$ , d'inverse  $W_N^*/N$ .

### Transformée de Fourier discrète rapide (TFR ou FFT)

Le calcul de la TFD du vecteur  $\mathbf{x}$  par la formule (II.14) nécessite environ  $N^2$  multiplications. L'algorithme de transformée de Fourier rapide est un moyen de calcul exact de la transformée de Fourier discrète qui permet de réduire notablement le coût de calcul. Le principe est de découper le problème initial en un grand nombre de problèmes auxiliaires de dimensions réduites, afin que le coût total de leur résolution (comprenant la somme des coûts de chacun d'entre eux et le raccordement des solutions élémentaires) soit sensiblement réduit. Ce découpage est rendu possible par la structure très particulière de la matrice  $W_N$ . De façon générale, la dimension des problèmes auxiliaires dépend de la décomposition de  $N$  en facteurs premiers. Le cas le plus favorable correspond à  $N = 2^K$ . Comme nous le verrons dans le dernier paragraphe, on se ramène souvent à ce cas, aussi est-ce le seul que nous détaillons. L'opération élémentaire est alors la décomposition d'une TFD d'ordre  $N$  en deux TFD, chacune d'ordre  $N/2$ .

On peut écrire (II.14) sous la forme

$$X_l = \frac{1}{N} \left( \sum_{j=0}^{N/2-1} x_{2j} w_N^{2jl} + \sum_{j=0}^{N/2-1} x_{2j+1} w_N^{(2j+1)l} \right).$$

Or  $w_N^{2k} = w_{N/2}^k$ , ce qui permet de transformer la relation précédente en

$$X_l = \frac{1}{N} \left( \sum_{j=0}^{N/2-1} x_{2j} w_{N/2}^{jl} + w_N^l \sum_{j=0}^{N/2-1} x_{2j+1} w_{N/2}^{jl} \right),$$

que l'on peut écrire sous la forme

$$X_l = Y_l + w_N^l Z_l \quad (0 \leq l \leq N-1),$$

où  $Y_l$  et  $Z_l$  correspondent aux TFD de  $y_i = x_{2i}$  et  $z_i = x_{2i+1}$  définies par II.12. En observant de plus que  $Y_{l+N/2} = Y_l$ ,  $Z_{l+N/2} = Z_l$  et  $w_N^{l+N/2} = -w_N^l$ , la formule précédente peut se séparer en

$$X_l = Y_l + w_N^l Z_l \quad (0 \leq l \leq N/2-1), \quad X_l = Y_l - w_N^l Z_l \quad (N/2 \leq l \leq N-1).$$

Le calcul de  $\mathbf{Y}$  et  $\mathbf{Z}$ , de dimension  $N/2$ , nécessite deux TFD d'ordre  $N/2$ . On a donc remplacé une TFD de longueur  $N$  par deux TFD d'ordre  $N/2$ , au prix de  $N/2$  multiplications complexes et  $N$  additions complexes supplémentaires. Dans le cas où  $N$  est une puissance de 2, on peut itérer le processus pour aboutir à des TFD de longueur 2 triviales. Le nombre d'itérations étant  $\log_2 N/2$ , on montre que la complexité totale de l'algorithme est de  $N/2 \log_2 N/2$  multiplications complexes et  $N \log_2 N/2$  additions complexes.

### Bourrage de zéros (*zero padding*)

Considérons un signal discret  $\{x_k\}_{k \in \mathbb{Z}}$  de support  $[0, N-1]$ . L'équation (II.9) nous permet de calculer sa transformée de Fourier

$$\tilde{X}(\nu) = \sum_{k=0}^{N-1} x_k e^{-2i\pi\nu k}, \quad \nu \in \mathbb{R},$$

qui est périodique de période 1. Notons  $\{x_k^M\}_{k \in \mathbb{Z}}$  le signal périodique de période  $M > N$ , construit à partir des échantillons non nuls du signal  $\{x_k\}_{k \in \mathbb{Z}}$  complétés par des zéros, c'est-à-dire défini par la période suivante :

$$x_0^M = x_0, \dots, x_{N-1}^M = x_{N-1}, x_N^M = 0, \dots, x_{M-1}^M = 0.$$

Sa transformée de Fourier discrète se calcule à partir de l'équation (II.12) :

$$X_l^M = \frac{1}{M} \sum_{k=0}^{N-1} x_k e^{-2i\pi k l / M}, \quad l = 0, \dots, M-1.$$

En examinant les deux équations précédentes on remarque que

$$\tilde{X}(l/M) = M X_l^M, \quad l = 0, \dots, M-1.$$

Ainsi la TFD de  $\{x_k^M\}_{k \in \mathbb{Z}}$  conduit à une version échantillonnée de la TF du signal à support limité  $\{x_k\}_{k \in \mathbb{Z}}$ , pour la période d'échantillonnage  $1/M$  (échantillonnage *fréquentiel*).

L'opération précédente est appelée « bourrage de zéro » (*zero padding*) dans la littérature de traitement du signal. Les calculs précédents montrent qu'elle permet :

— d'interpoler la transformée de Fourier d'un signal à support limité pour toutes les fréquences uniquement en utilisant des TFR,

— de remplacer le calcul d'une TFR sur  $N$  points par un calcul sur  $M = 2^K$  points ( $M > N$ ). On obtient alors la TF de  $\{x_k\}$  échantillonnée avec la période fréquentielle  $1/2^K$  au lieu de  $1/N$ . Ceci n'empêche pas d'étudier le comportement fréquentiel du signal  $\{x_k\}$ , et permet d'utiliser l'algorithme de TFR associé aux puissances de deux (voir le paragraphe précédent). Notons que cette opération est indispensable lorsque l'on veut faire la TFR d'une séquence dont la taille  $N$  est un (grand) nombre premier.

### Formule de convolution-multiplication

Soit deux signaux discrets périodiques  $\{x_k\}$  et  $\{h_k\}$  de même période  $N$  et de représentations fréquentielles respectives  $\{X_k\}_{k \in \mathbb{Z}}$  et  $\{H_k\}_{k \in \mathbb{Z}}$ . La formule de convolution-multiplication s'écrit :

$$y_k = \sum_{n=0}^{N-1} x_n h_{k-n}, \quad k \in \mathbb{Z} \quad \xrightarrow{\text{TFD}} \quad Y_k = H_k X_k, \quad k \in \mathbb{Z}$$

où  $\{y_k\}_{k \in \mathbb{Z}}$  et  $\{Y_k\}_{k \in \mathbb{Z}}$  sont des signaux périodiques de période  $N$ . Si l'on représente ces trois signaux périodiques par les vecteurs  $\mathbf{x} = (x_0, \dots, x_{N-1})^t$ ,  $\mathbf{h}$ ,  $\mathbf{y}$  et leurs représentations fréquentielles respectives par  $\mathbf{X}$ ,  $\mathbf{H}$  et  $\mathbf{Y}$  on peut écrire la formule précédente comme une relation matricielle

$$\mathbf{Y} = (H_0 X_0, \dots, H_{N-1} X_{N-1})^t = \frac{1}{N} W_N \mathbf{y}.$$

Le vecteur  $\mathbf{y}$  est alors la *convolution circulaire* des vecteurs  $\mathbf{x}$  et  $\mathbf{h}$ , définie par

$$\begin{cases} y_k &= \sum_{n=0}^k x_n h_{k-n} + \sum_{n=k+1}^{N-1} x_n h_{k+N-n}, \quad 0 \leq k \leq N-2; \\ y_{N-1} &= \sum_{n=0}^{N-1} x_n h_{k-n}. \end{cases} \quad (\text{II.15})$$

# Chapitre III

## Filtres numériques

### III.1 Introduction

Le signal de sortie d'un filtre linéaire homogène est donné par le produit de convolution de  $\{x_k\}$ , le signal d'entrée, et de  $\{h_k\}$ , la réponse impulsionnelle du filtre. La réponse impulsionnelle  $\{h_k\}$ , ou réponse percussive dans certains ouvrages, est définie comme la sortie du filtre pour le signal d'entrée  $\{\delta_k\}$ . La TZ de  $\{h_k\}$  est définie par :

$$\{h_k\}_{k \in \mathbb{Z}} \xrightarrow{\text{TZ}} H(z) \quad \text{et} \quad \text{RDC} = \{z/R_1 < |z| < R_2\}. \quad (\text{III.1})$$

On appelle *fonction de transfert* la fonction  $H(z)$ . Une autre représentation que nous serons souvent appelés à utiliser est la réponse fréquentielle du filtre  $\tilde{H}(\nu)$  qui est la transformée de Fourier de la réponse impulsionnelle.

Les filtres sont fréquemment utilisés afin de transformer les signaux d'entrée de façon à éliminer, ou tout du moins à atténuer fortement, certaines fréquences qui sont indésirables ou qui correspondent à des perturbations. Par exemple, il est souvent nécessaire de supprimer des fréquences proches de 50 ou 60 Hz qui peuvent induire des interférences avec le courant d'alimentation. Nous avons présenté au chapitre 1 certains types de filtres utilisés pour cet usage, qui sont :

— les filtres passe-bas qui atténuent fortement les fréquences situées au delà d'une limite définie sans affecter les fréquences plus basses.

— Les filtres passe-haut qui atténuent fortement les fréquences situées en deçà d'une limite définie sans affecter les fréquences plus hautes.

— Les filtres passe-bande qui préservent le signal d'entrée dans un intervalle de fréquence et atténuent ses composantes situées à l'extérieur de cet intervalle.

— Les filtres coupe-bande qui suppriment les composantes du signal d'entrée situées dans un intervalle de fréquence et n'affectent pas les composantes situées à l'extérieur de cet intervalle.

La figure III.1 présente pour chacun des types de filtre mentionnés ci-dessus la réponse fréquentielle associée à des filtres réels en comparant avec la réponse idéale souhaitée. La synthèse des filtres réels a été effectuée par l'approche type Butterworth [3] où l'on approxime la fonction à l'aide d'un polynôme de degré  $N$ . L'approximation est d'autant meilleure que  $N$  est élevé mais le filtre obtenu devient difficile à utiliser dans la pratique. Les filtres présentés ici correspondent à  $N = 8$ .

Nous allons à présent introduire les propriétés importantes de certains filtres numériques.

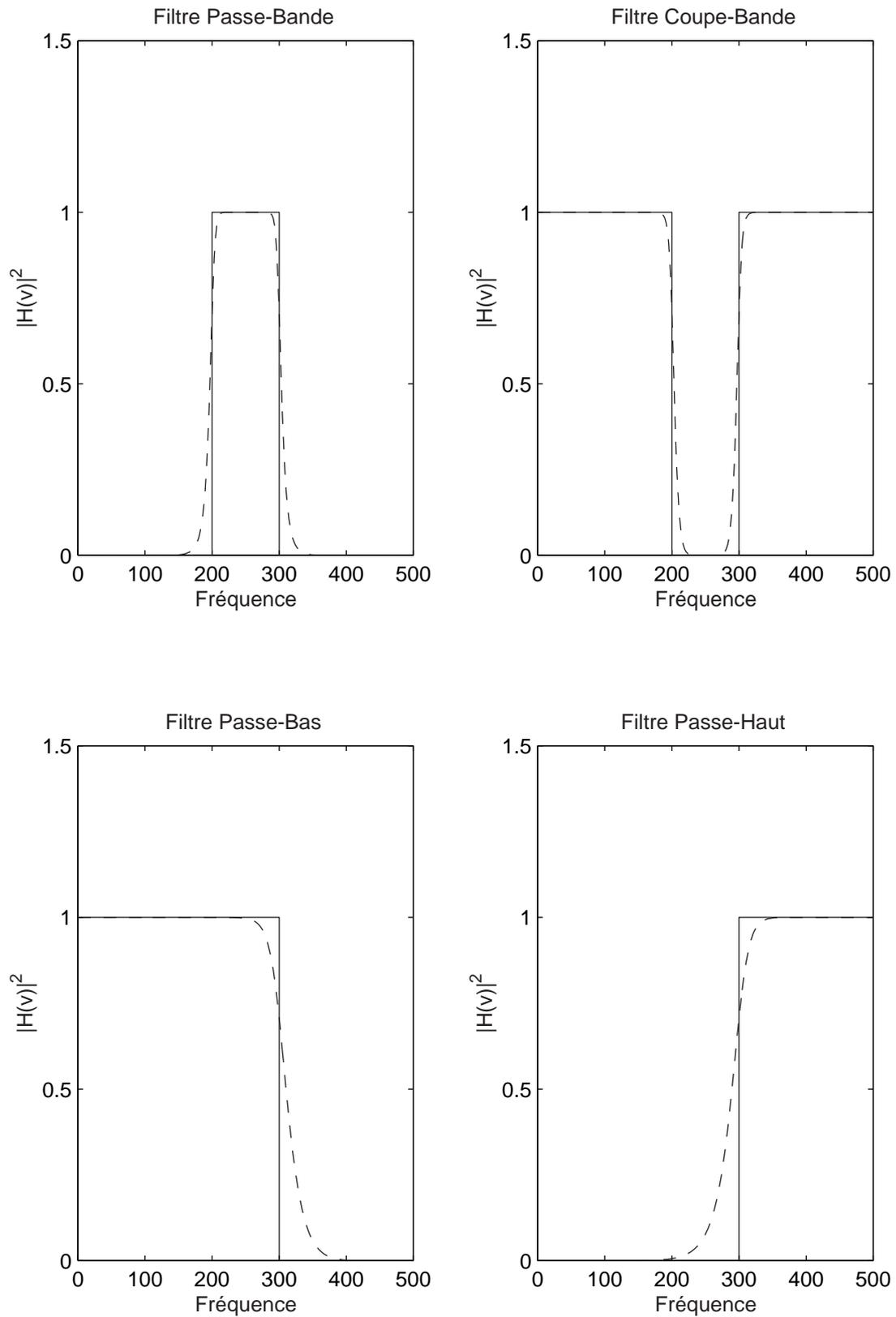


Fig. III.1. Exemples de modules de fonctions de transfert de filtre.

### III.1.1 Stabilité

Il existe plusieurs définitions de la stabilité d'un filtre ; nous ne retiendrons ici que la stabilité au sens déjà défini dans le cas continu, c'est-à-dire  $\{h_k\} \in \ell^1$ . On la désigne parfois par « stabilité au sens strict », ou « stabilité BIBO » (Bounded Input Bounded Output), car on montre que  $\{h_k\} \in \ell^1$  est une condition nécessaire et suffisante pour qu'à toute entrée bornée le filtre associe une sortie bornée. Comme indiqué précédemment, la stabilité stricte de  $\{h_k\}$  correspond au fait que le cercle unité appartienne à la RDC, ou encore à l'existence de la transformée de Fourier de  $\{h_k\}$  :  $\tilde{H}(\nu) = H(e^{2i\pi\nu})$ , appelée transmittance, comme à temps continu.

### III.1.2 Causalité

Un filtre est dit *causal* si  $h_k = 0$  pour tout  $k < 0$ . Cette propriété est équivalente pour la TZ du filtre à : la RDC associée à  $\{h_k\}$  est de la forme  $[R_1, +\infty[$ .

On peut facilement comprendre cette caractérisation si on se souvient de la méthode d'inversion de la TZ par le calcul des résidus : on a vu en effet que l'on pouvait calculer les valeurs de  $h_k$  pour  $k < 0$  en sommant les résidus de  $X(z)z^{k-1}$  sur les pôles extérieurs au contour d'intégration (situé dans la RDC). La condition précédente garantit qu'il n'existe pas de tels pôles, puisque la série entière est holomorphe sur sa RDC  $[R_1, +\infty[$ . Donc le résultat de l'intégration (II.8) est nul :  $h_k = 0$  pour  $k < 0$ .

Inversement, si  $h_k = 0$  pour tout  $k < 0$ , en reprenant les calculs du paragraphe II.3.1 permettant de définir la région de convergence de la TZ, on constate aisément que  $R_{h+}$  est égal à  $+\infty$  et que la RDC est définie par  $[R_1, +\infty[$ .

Un filtre causal et stable est dit *réalisable*.

### III.1.3 Inversion d'un filtre numérique

On considère un filtre numérique  $H$  défini par sa réponse impulsionnelle  $\{h_k\}$ . Soit  $\{x_k\}$  un signal discret appliqué en entrée du filtre  $H$ , et soit  $\{y_k\}$  le signal de sortie associé. On définit le *filtre inverse du filtre  $H$*  comme le filtre qui permet de retrouver  $\{x_k\}$  en filtrant  $\{y_k\}$ , pour une famille de signaux  $\{x_k\}$ . Ce filtre inverse est noté  $H^{-1}$ .

Soit  $K$  le filtre inverse d'un filtre  $H$  défini par sa transformée en  $z$ ,  $H(z)$ , et sa région de convergence. Le filtre inverse admet comme transformée en  $z$

$$K(z) = \frac{1}{H(z)}.$$

Les pôles de cette fonction sont les zéros de  $H(z)$  et le domaine de convergence de  $K(z)$  sera déduit de l'étude de ces zéros. Si elle existe, la transmittance du filtre  $K$  est obtenue à partir de celle du filtre  $H$  par

$$K(e^{2i\pi\nu}) = \frac{1}{H(e^{2i\pi\nu})}.$$

### III.1.4 Filtres rationnels, filtres dynamiques

De même qu'un grand nombre de systèmes à temps continu sont régis par des équations différentielles, une large classe de systèmes à temps discret découlent d'*équations aux différences*, du type

$$\sum_{n=0}^N a_n y_{k-n} = \sum_{m=0}^M b_m x_{k-m} \quad (\text{III.2})$$

où  $\{y_k\}$  est considéré comme la sortie d'un système dont  $\{x_k\}$  est l'entrée. En prenant la TZ des deux membres de l'équation précédente on obtient la fonction de transfert du système correspondant :

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{m=0}^M b_m z^{-m}}{\sum_{n=0}^N a_n z^{-n}}. \quad (\text{III.3})$$

Partant de (III.3), il est facile de mettre  $H(z)$  sous la forme d'un quotient  $P(z)/Q(z)$  de deux polynômes en  $z$ , de degrés respectifs  $M_0$  et  $N_0$ . Comme les polynômes de la variable complexe sont entièrement définis par la position de leurs zéros à un facteur près,  $H(z)$  est entièrement définie par ses  $N_0$  pôles (zéros de  $Q(z)$ ), ses  $M_0$  zéros (zéros de  $P(z)$ ) et un facteur multiplicatif  $A$  :

$$H(z) = \frac{P(z)}{Q(z)} = A \frac{\prod_{m=1}^{M_0} (z - z_m)}{\prod_{n=1}^{N_0} (z - p_n)}. \quad (\text{III.4})$$

Les  $N_0$  zéros de  $Q(z)$  et les  $M_0$  zéros de  $P(z)$  définissent les pôles et les zéros *finis* de  $H(z)$ . Dans le cas d'un filtre réel, ces pôles et ces zéros sont réels ou deux à deux conjugués.

- Lorsque  $M_0$  est supérieur à  $N_0$ ,  $H(z)$  admet également  $M_0 - N_0$  pôles à l'infini.
- Lorsque  $N_0$  est supérieur à  $M_0$ ,  $H(z)$  admet également  $N_0 - M_0$  zéros à l'infini.

L'expression de  $H(z)$ , fonction de la variable complexe, est holomorphe en tout point sauf en ses  $N_0$  pôles. Elle admet un développement en série de Laurent différent en chaque couronne du type  $\{z/|z| \in ]R_1, R_2[ \}$ , où  $R_1$  et  $R_2$  sont les modules de deux pôles « successifs » (*i.e.*, consécutifs dans la liste des pôles rangés par module croissant), c'est-à-dire qu'en chacune de ces régions on peut calculer une réponse impulsionnelle d'un filtre décrit par les équations (III.2). Notons que l'existence de plusieurs filtres associés à l'équation (III.2) ne doit pas surprendre : elle est le pendant du choix des conditions initiales dans l'équation différentielle. Tous ces filtres sont appelés *filtres rationnels*.

Étudions maintenant à quelles conditions on peut trouver des filtres causaux et des filtres stables à partir de  $H(z)$ .

— La *causalité* correspond à une région de convergence du type  $\{z/|z| \in ]R_1, +\infty[ \}$ . Dans le cas où il existe des pôles à l'infini, aucune région de convergence de ce type ne peut être associée à  $H(z)$ , il n'existe pas de filtre causal. Dans le cas contraire où il n'y a pas de pôle à l'infini, c'est-à-dire que  $N_0$  est supérieur ou égal à  $M_0$ , alors le filtre correspondant à la RDC  $\{z/|z| \in ]R_{\max}, +\infty[ \}$ , où  $R_{\max}$  est le plus grand des modules des pôles de  $H(z)$ , est causal.

— La *stabilité* correspond à l'appartenance du cercle unité à la RDC. Il est clair que tant que la fonction de transfert  $H(z)$  ne possède pas de pôles sur le cercle unité, il est possible de trouver une RDC contenant celui-ci : le filtre obtenu sur cette région est stable.

— Pour être *réalisable* il faut réunir les deux conditions précédentes (RDC du type  $\{z/|z| \in ]R_1, +\infty[ \}$  contenant le cercle unité) : il faut et il suffit que tous les pôles de  $H(z)$  soient à l'intérieur du cercle unité.

On appelle *filtre dynamique* un filtre réalisable dont la fonction de transfert est une fraction rationnelle. L'étude du comportement en fréquence d'un filtre dynamique, c'est-à-dire l'étude en phase et module de la transmittance  $\tilde{H}(e^{2i\pi\nu})$ , peut être trouvée dans de nombreux ouvrages de référence plus détaillés que ces notes de cours.

Un filtre dynamique est à *phase minimale* ou à *déphasage minimal* si son inverse stable (*i.e.*, le filtre associé à  $H^{-1}(z)$  développée sur le cercle unité) est lui aussi réalisable.

Pour qu'il existe un filtre à phase minimale associé à  $H(z)$ , il faut qu'il existe :

- un filtre causal associé à  $H(z)$  donc que la fonction n'admette aucun pôle à l'infini donc que  $N_0$  soit supérieur ou égal à  $M_0$  ;

- un filtre causal associé à  $H^{-1}(z)$  donc que la fonction  $H(z)$  n'admette aucun zéro à l'infini donc que  $N_0$  soit inférieur ou égal à  $M_0$ .

De plus, pour assurer la causalité du filtre et de son inverse, tout en préservant la stabilité, les pôles et les zéros de  $H(z)$  doivent se trouver à l'intérieur du cercle unité.

En résumé, il existera un filtre à phase minimale associé à  $H(z)$ , si il existe *autant de pôles que de zéros, tous situés à l'intérieur du cercle unité*.

Résumons l'emboîtement des propriétés et leurs appellations usuelles :

$$\left. \begin{array}{l} \text{causal} \\ \text{stable} \end{array} \right\} = \left. \begin{array}{l} \text{réalisable} \\ \text{rationnel} \end{array} \right\} = \left. \begin{array}{l} \text{dynamique} \\ \text{d'inverse réalisable} \end{array} \right\} = \text{à phase minimale.}$$

Examinons maintenant deux cas particuliers importants de filtres rationnels, éventuellement réalisables.

### III.1.5 Filtres à Réponse Impulsionnelle Finie (RIF)

Ces filtres sont également dénommés *filtres transverses*. C'est le cas particulier de (III.4) obtenu pour  $Q(z) = z^M$ . L'équation d'entrée-sortie s'écrit alors :

$$y_k = \sum_{m=0}^M b_m x_{k-m}. \tag{III.5}$$

On constate que la réponse impulsionnelle de ce filtre est simplement égale à la suite des coefficients  $b_m$  pour  $m = 0, \dots, M$ , prolongée à gauche et à droite par des zéros. C'est pourquoi on parle de filtre à réponse impulsionnelle finie. La fonction de transfert étant réduite à un polynôme en  $z^{-1}$ , on peut l'écrire

$$H(z) = \frac{1}{z^M} \sum_{m=0}^M b_{M-m} z^m \tag{III.6}$$

et il n'y a pas de problème de choix de RDC puisque tous les pôles sont réunis en zéro. Par contre ce filtre n'est à phase minimale que si les zéros de  $H(z)$  sont à l'intérieur du cercle unité.

### III.1.6 Filtres Tous Pôles (FTP)

Ce type de filtre est également dénommé *filtre récursif*. C'est le cas particulier de (III.4) obtenu pour  $P(z) = z^N$ . L'équation d'entrée-sortie s'écrit alors :

$$\sum_{n=0}^N a_n y_{k-n} = x_k, \tag{III.7}$$

et la fonction de transfert se réduit à :

$$H(z) = \frac{z^N}{\sum_{n=0}^N a_{N-n} z^n}. \tag{III.8}$$

Le filtre est réalisable si tous les pôles sont dans le cercle unité. Il est alors aussi à phase minimale. La causalité du filtre se traduit par la possibilité de calculer *récursivement* la sortie  $\{y_k\}$  sous la forme suivante, déduite de (III.7) en supposant  $a_0 \neq 0$  :

$$y_k = \frac{1}{a_0} x_k - \sum_{n=1}^N \frac{a_n}{a_0} y_{k-n}.$$

## III.2 Exemples de filtres

### III.2.1 Retard

Soit la relation entrée-sortie entre  $\{x_k\}$  et  $\{y_k\}$  définie par

$$y_k = x_{k-n}. \quad (\text{III.9})$$

Cette relation peut être obtenue en filtrant le signal d'entrée  $\{x_k\}$  par le filtre homogène de réponse impulsionnelle  $\{h_k\}$  :

$$h_k = \begin{cases} 1 & \text{si } k = n, \\ 0 & \text{sinon} \end{cases}$$

qui a pour fonction de transfert

$$H(z) = z^{-n}$$

et pour transmittance

$$\tilde{H}(\nu) = e^{-2i\pi\nu n}.$$

### III.2.2 Lissage

On considère le filtre défini par la relation d'entrée-sortie où  $\{x_k\}$  est le signal d'entrée stable et  $\{y_k\}$  le signal de sortie

$$y_k = \sum_{m=-N}^N x_{k-m}. \quad (\text{III.10})$$

Ce filtre a pour réponse impulsionnelle

$$h_k = \begin{cases} 1 & \text{si } -N \leq k \leq N, \\ 0 & \text{sinon.} \end{cases}$$

Sa transmittance est alors

$$\tilde{H}(\nu) = \frac{\sin(2N+1)\pi\nu}{\sin\pi\nu}.$$

La figure III.2 présente la réponse impulsionnelle et la réponse fréquentielle d'un tel filtre pour  $N = 6$ . Le filtre ainsi défini est stable mais non causal. Pour le rendre causal, on introduit un retard de  $N$  échantillons et la relation (III.10) devient

$$y_k = \sum_{n=0}^{2N} x_{k-n}.$$

### III.2.3 Filtre récursif d'ordre 1

Soit un signal réel numérique discret stable  $\{x_k\}$  et le signal de sortie  $\{y_k\}$  réel défini par la relation de récurrence

$$y_k = \lambda y_{k-1} + \mu x_k \quad \forall k \in \mathbb{Z}, \quad \lambda, \mu \in \mathbb{R} \quad (\text{III.11})$$

Les transformées en  $z$  associées sont obtenues en multipliant les deux membres de l'équation par  $z^{-k}$  et en effectuant la somme pour  $z \in \mathbb{Z}$ . On obtient ainsi

$$Y(z) = H(z)X(z) = \frac{\mu}{1 - \lambda z^{-1}} X(z).$$

Le signal de sortie  $y_k$  est donc obtenu par filtrage du signal d'entrée  $x_k$  par le filtre numérique  $H(z)$  dont la transformée en  $z$  s'écrit

$$H(z) = \frac{\mu}{1 - \lambda z^{-1}} = \frac{\mu z}{z - \lambda}. \quad (\text{III.12})$$

Ce filtre est dénommé *filtre récursif d'ordre 1*.

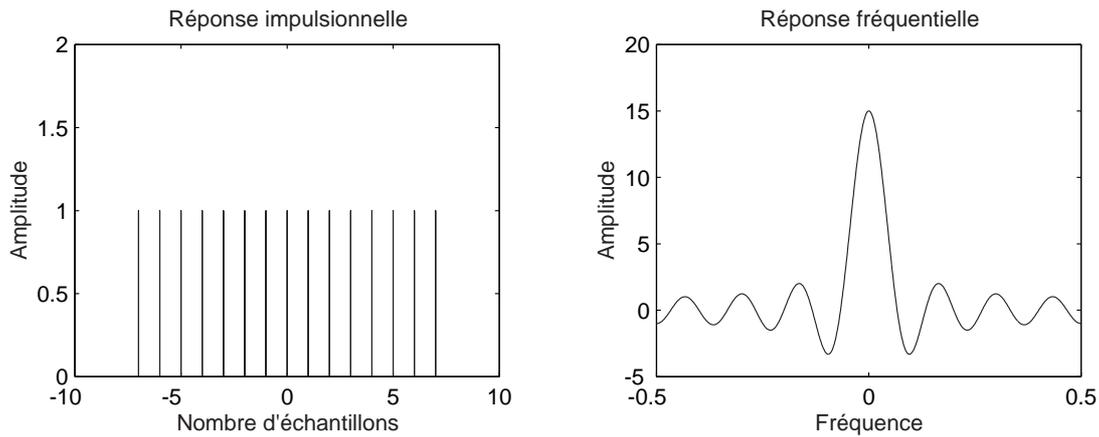


Fig. III.2. Réponse impulsionnelle et fréquentielle d'un filtre lisseur. Le filtre est de type passe-bas.

### Stabilité du filtre

Le filtre  $H(z)$  est stable et causal si et seulement si  $|\lambda| < 1$ . Dans ce cas, sa réponse impulsionnelle est définie par

$$h_k = \mu \lambda^k, \quad \forall k \in \mathbb{N}.$$

Sa transmittance est

$$\tilde{H}(\nu) = \frac{\mu}{1 - \lambda e^{-2i\pi\nu}}.$$

qui a pour module

$$|\tilde{H}(\nu)| = \frac{|\mu|}{\sqrt{1 + \lambda^2 - 2\lambda \cos 2\pi\nu}}$$

La figure III.3 présente ce module dans l'intervalle  $[-1/2, 1/2]$  pour  $\lambda = 0,7$  et  $\mu = 0,3$ . Le filtre récursif est donc de type passe-bas. Il est d'autant plus efficace que  $\lambda$  est voisin de 1 mais il devient alors d'autant moins stable. Pour illustrer cette propriété, la figure présente le signal de sortie  $\{y_k\}$  obtenu en filtrant un signal d'entrée  $\{x_k\}$  par  $H$ . Afin d'illustrer les propriétés fréquentielles du filtre, le signal d'entrée doit posséder un spectre couvrant la gamme de fréquences.

Le signal d'entrée choisi est du type

$$x_k = e^{-at_k^2} \sin \omega t_k$$

où  $\omega$  varie en fonction du temps sous la forme

$$\omega = 200 \sin 100t_k.$$

Le signal ainsi obtenu est stable quel que soit  $\omega$  et est dit *vobulé*.

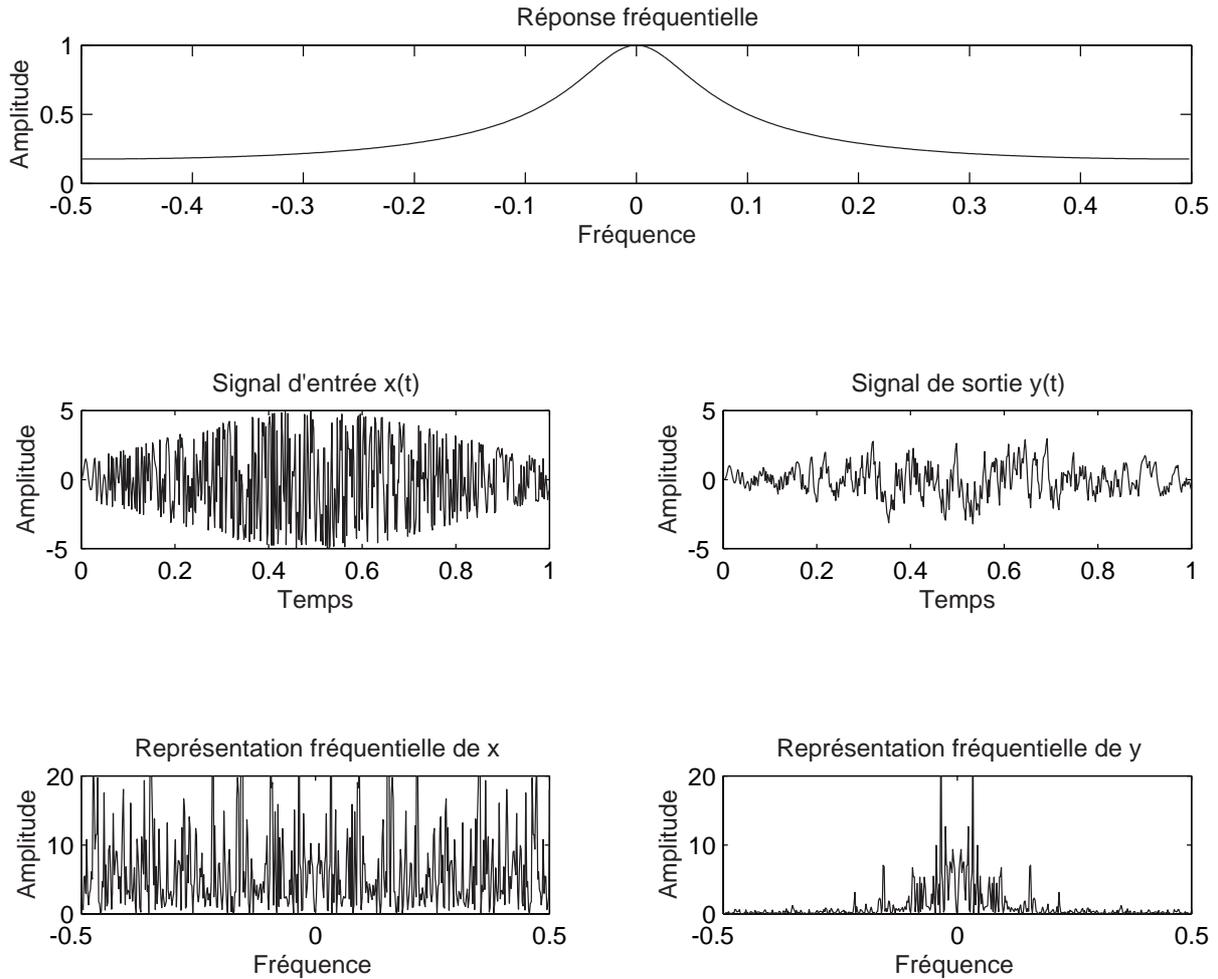
### Filtre inverse

Le filtre récursif d'ordre 1 est inversible et admet pour filtre inverse le filtre  $K$  de transformée en  $z$

$$K(z) = \frac{z - \lambda}{\mu z}.$$

Il existe un pôle simple situé à l'origine. Le filtre  $K$  est donc stable et a pour réponse impulsionnelle  $\{k_n\}$ ,

$$k_0 = \frac{1}{\mu}, \quad k_1 = -\frac{\lambda}{\mu}, \quad \forall n \geq 2, k_n = 0.$$



**Fig. III.3.** Représentation temps-fréquence d'un signal vobulé filtré par un filtre AR d'ordre 1.

Le signal  $\{x_k\}$  peut être obtenu par filtrage de  $\{y_k\}$  par  $K$ , ce qui s'écrit

$$x_k = \frac{1}{\mu} y_k - \frac{\lambda}{\mu} y_{k-1}, \quad \forall k \in \mathbb{N}.$$

Le filtre  $K$  effectue une *moyenne glissante* sur  $\{y_k\}$  pour produire le signal  $\{x_k\}$ . Il est de type passe-haut puisque le filtre récursif d'ordre 1,  $H$ , est de type passe-bas.

La figure III.4 présente la réponse fréquentielle du filtre inverse du filtre  $H$  défini précédemment. L'évolution du signal de sortie  $\{y_k\}$  obtenu par filtrage d'un signal d'entrée  $\{x_k\}$  de caractéristiques analogues à celles du filtre  $H$  illustre les variations en fréquence correspondantes.

### III.2.4 Filtre récursif d'ordre 2

Soit un signal numérique discret stable  $\{x_k\}$  et le signal de sortie  $\{y_k\}$  défini par la relation de récurrence

$$y_k = \lambda_1 y_{k-1} + \lambda_2 y_{k-2} + \mu x_k, \quad \forall k \in \mathbb{Z}, \lambda_1, \lambda_2, \mu \in \mathbb{R}. \quad (\text{III.13})$$

Par un processus analogue à celui exposé au paragraphe précédent, on montre que le signal de sortie  $\{y_k\}$  est obtenu par filtrage du signal d'entrée  $\{x_k\}$  par le filtre numérique  $H(z)$  défini par

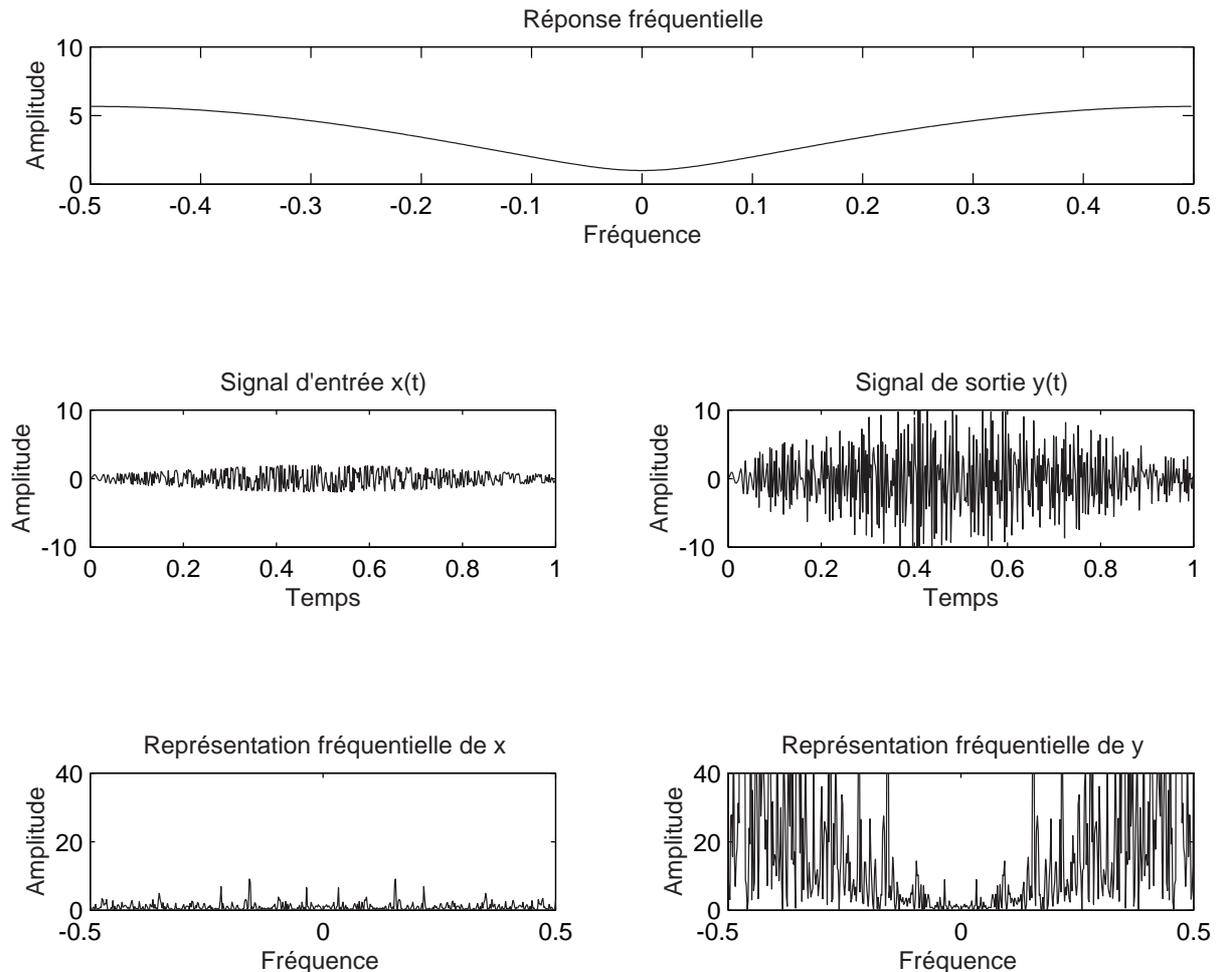


Fig. III.4. Représentation temps-fréquence d'un signal vobulé filtré par le filtre inverse.

$$H(z) = \frac{\mu z^2}{z^2 - \lambda_1 z - \lambda_2}. \quad (\text{III.14})$$

Ce filtre est dénommé *filtre récursif d'ordre 2*. Nous nous limiterons par la suite au cas où les pôles de  $H$  sont complexes conjugués. En effet, si les pôles de  $H$  sont réels, ce filtre peut être représenté sous la forme de la somme de deux filtres récursifs d'ordre 1 que nous venons d'étudier.

Dans le cas présent,  $H$  est un filtre dynamique si et seulement si  $|\lambda_2| < 1$ . Soient  $z_0$  et  $\bar{z}_0$  les pôles de  $H$ , qui s'expriment sous la forme  $z_0 = r e^{i\theta}$ ,  $\bar{z}_0 = r e^{-i\theta}$ , ce qui implique  $\lambda_1 = 2r \cos(\theta)$ ,  $\lambda_2 = -r^2$ . La réponse impulsionnelle du filtre  $H$  est alors définie par

$$h_n = \mu r^n \frac{\sin(n+1)\theta}{\sin \theta}, \quad \forall n \in \mathbb{N}.$$

Sa transmittance est

$$\tilde{H}(\nu) = \frac{\mu}{1 - \lambda_1 e^{-2i\pi\nu} - \lambda_2 e^{-4i\pi\nu}}$$

qui a pour module

$$|H(e^{2i\pi\nu})| = \frac{|\mu|}{\sqrt{(1 + r^2 - 2 \cos(\theta - 2\pi\nu))(1 + r^2 - 2 \cos(\theta + 2\pi\nu))}}.$$

Ce module passe par des extréma pour  $\nu = 0$  et  $\nu = \pm 1/2$ . De plus, on peut montrer que si on a la condition  $|\cos \theta| < 2r/(1 + r^2)$ , il existe un maximum pour la fréquence  $\nu = \pm \nu_0$  donnée par

$$\cos 2\pi\nu_0 = \frac{1 + r^2}{2r} \cos \theta$$

et le maximum est égal à

$$\frac{|\mu|}{1 - r^2} \frac{1}{\sqrt{1 - \cos^2 \theta}}.$$

La fréquence  $\nu_0$  est dite *fréquence de résonance* du filtre.

La figure III.5 présente le module d'un tel filtre dans l'intervalle  $[-1/2, 1/2]$  pour  $\lambda_1 = 1,5$ ,  $\lambda_2 = -0,7$  et  $\mu = 0,13$ . La fréquence de résonance associée est  $\nu_0 = 0,07$ . L'évolution en temps et en fréquence du signal de sortie  $\{y_k\}$  obtenu par filtrage d'un signal d'entrée vobulé  $\{x_k\}$  illustre les propriétés fréquentielles d'un tel filtre.

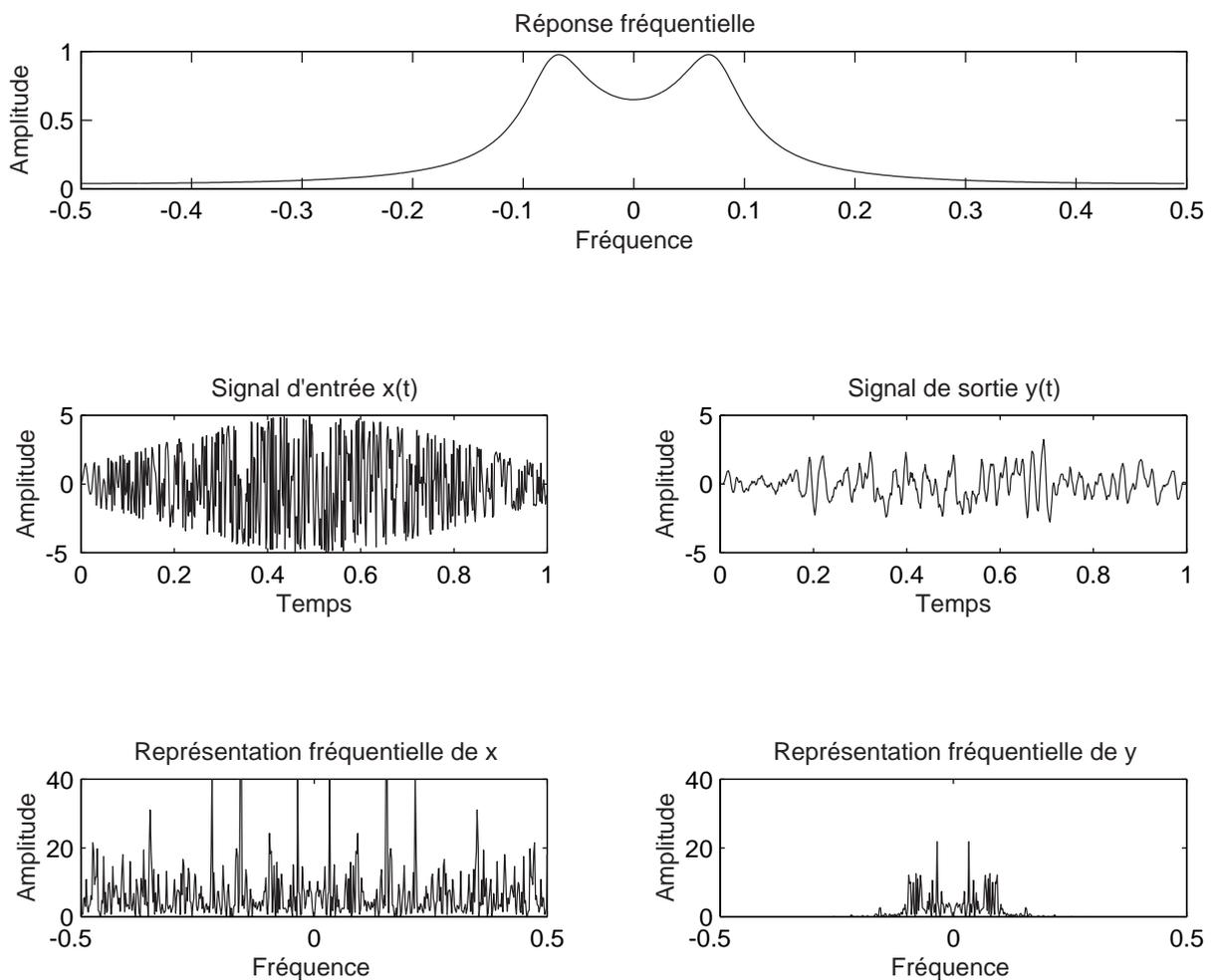


Fig. III.5. Représentation temps-fréquence d'un signal vobulé filtré par un filtre AR d'ordre 2.

**Filtre inverse**

Le filtre récursif d'ordre 2 est inversible et admet pour filtre inverse le filtre  $K$  de transformée en  $z$

$$K(z) = \frac{z^2 - \lambda_1 z - \lambda_2 z}{\mu z^2}.$$

Il existe un pôle double situé à l'origine. Le filtre  $K$  est donc stable et a pour réponse impulsionnelle  $\{k_n\}$ ,

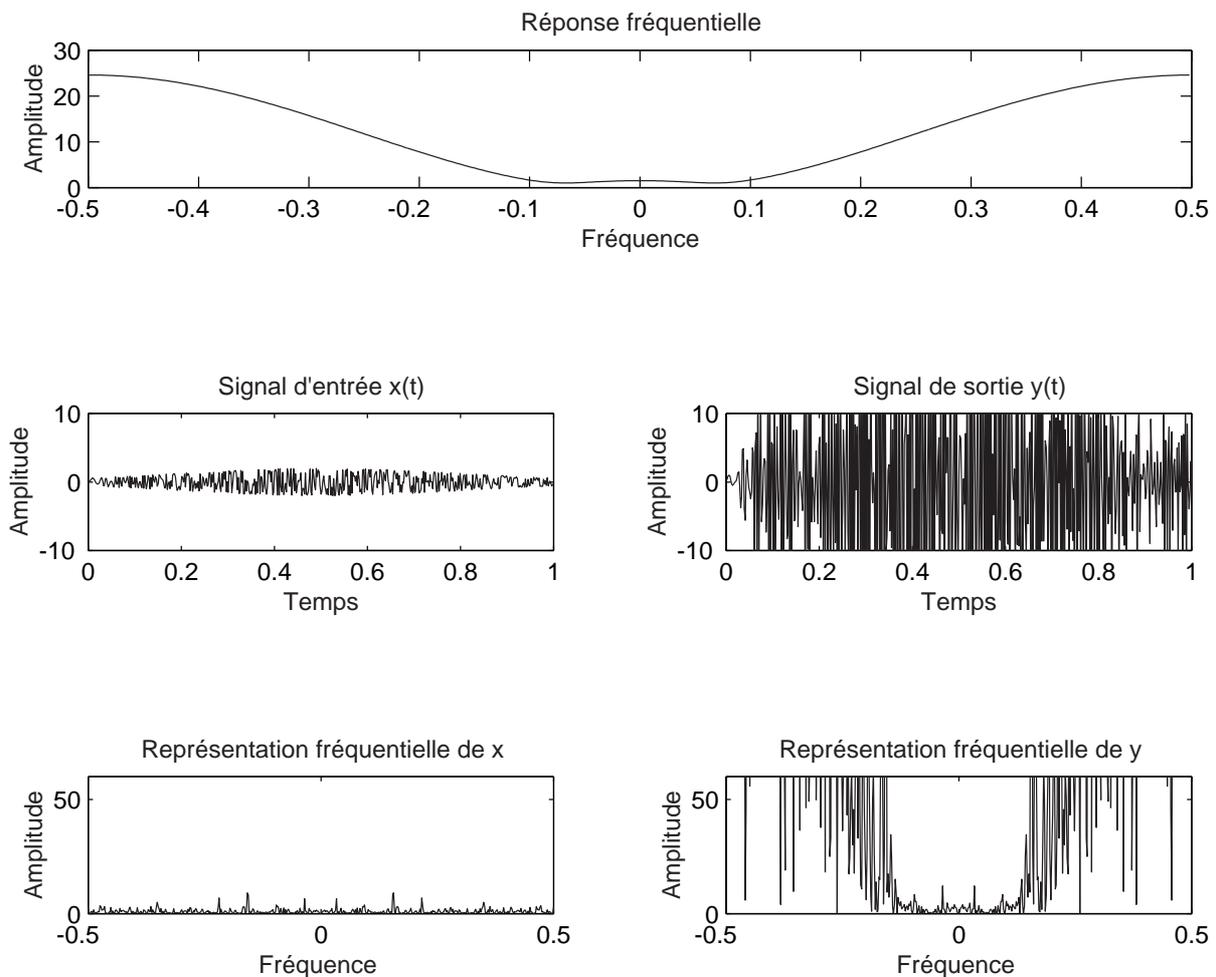
$$k_0 = \frac{1}{\mu}, k_1 = -\frac{\lambda_1}{\mu}, k_2 = -\frac{\lambda_2}{\mu}, \quad \forall n \geq 3, k_n = 0.$$

Le signal  $\{x_k\}$  peut être obtenu par filtrage de  $\{y_k\}$  par  $K$ , ce qui s'écrit

$$x_k = \frac{1}{\mu} y_k - \frac{\lambda_1}{\mu} y_{k-1} - \frac{\lambda_2}{\mu} y_{k-2}, \quad \forall k \in \mathbb{N}.$$

Le filtre  $K$  est donc un *filtre RIF d'ordre 3*.

La figure III.6 présente le module du filtre  $K$  inverse du filtre  $H$  présenté précédemment. L'évolution en temps et en fréquence du signal de sortie  $\{y_k\}$  obtenu par filtrage d'un signal d'entrée vobulé  $\{x_k\}$  illustre les propriétés fréquentielles d'un tel filtre.



**Fig. III.6.** Représentation temps-fréquence d'un signal vobulé filtré par le filtre inverse.



# Chapitre IV

## Événements, variables et vecteurs aléatoires

### IV.1 Introduction

Comme on l'a vu dans les chapitres précédents, les signaux déterministes permettent de modéliser des phénomènes limités dans le temps, ou bien périodiques, oscillatoires amortis... De nombreux phénomènes physiques présentent des fluctuations permanentes non périodiques, non amorties, mais qui manifestent pourtant un certain degré de « régularité ». Par exemple, la météorologie est fondée sur l'analyse de signaux de pluviométrie, d'ensoleillement... , qui présentent un caractère pseudo-périodique (saisons, alternance jour-nuit, cycle des marées...), mais dont la partie « utile » est noyée dans un ensemble de « perturbations » d'origines diverses et de forme irrégulière. Les signaux déterministes étudiés dans les chapitres précédents sont trop limités pour intégrer facilement de telles caractéristiques, tandis que les modèles aléatoires fournissent un cadre mathématique plus adapté.

La modélisation aléatoire des signaux est présentée au chapitre V, uniquement à temps discret par souci de simplicité ; cette présentation est largement suffisante dans le cadre d'un cours sur le traitement numérique du signal. Certaines notions concernant les signaux aléatoires à temps continu (ou encore, *processus stochastiques*) sont présentées dans [7].

Le présent chapitre est consacré à la théorie des probabilités, sur laquelle repose la modélisation aléatoire des signaux. Pour commencer, on rappelle le vocabulaire de la théorie de la mesure, essentielle dans la suite puisqu'une probabilité est un cas particulier de mesure finie et une variable aléatoire, un cas particulier d'application mesurable. Dans les sections suivantes, on introduit successivement les variables puis les vecteurs aléatoires.

### IV.2 Intégration et probabilités

#### IV.2.1 Espaces et applications mesurables

On appelle *espace mesurable*  $(\Omega, \mathcal{F})$  le couple formé d'un ensemble  $\Omega$  et d'une *tribu*  $\mathcal{F}$ , c'est-à-dire une famille de parties de  $\Omega$  vérifiant :

- (i)  $\Omega \in \mathcal{F}$ ,
- (ii) la stabilité par complémentation :  $\forall A \in \mathcal{F}, A^c \in \mathcal{F}$ ,
- (iii) la stabilité par union dénombrable : pour toute famille dénombrable  $\{A_i \in \mathcal{F}, i \in I\}$ , on a  $(\cup_I A_i) \in \mathcal{F}$ .

On déduit facilement de (i)-(ii) que  $\emptyset \in \mathcal{F}$  et de (ii)-(iii) que  $\mathcal{F}$  est stable par intersection

dénombrable.

Pour toute famille  $\mathcal{C}$  de parties de  $\Omega$ , on peut définir la tribu engendrée par  $\mathcal{C}$  comme la plus petite tribu contenant  $\mathcal{C}$ . Ainsi, on munit généralement  $\mathbb{R}^d$  de la tribu  $\mathcal{B}^d$  des *boréliens*, qui est la tribu engendrée par les pavés ouverts.

Soient  $(\Omega, \mathcal{F})$  et  $(E, \mathcal{A})$  deux espaces mesurables. On appelle *application mesurable* toute application  $X$  de  $(\Omega, \mathcal{F})$  dans  $(E, \mathcal{A})$  telle que  $\forall A \in \mathcal{A}, X^{-1}(A) \in \mathcal{F}$ .

## IV.2.2 Notion de mesure

### Mesure sur $(\Omega, \mathcal{F})$

On appelle *mesure* sur un espace mesurable  $(\Omega, \mathcal{F})$  toute application  $\mu$  de  $\mathcal{F}$  dans  $\mathbb{R}^+$  qui possède la propriété de  $\sigma$ -additivité, c'est-à-dire que pour toute famille dénombrable  $\{A_i, i \in I\}$  d'éléments deux à deux disjoints de  $\mathcal{F}$ , on a

$$\mu(\cup_I A_i) = \sum_I \mu(A_i). \quad (\text{IV.1})$$

Le triplet  $(\Omega, \mathcal{F}, \mu)$  s'appelle un *espace mesuré*.

Soit  $(\Omega, \mathcal{F}, \mu)$  un espace mesuré et  $X$  une application mesurable de  $(\Omega, \mathcal{F})$  dans  $(E, \mathcal{A})$ .  $X$  permet de construire sur  $(E, \mathcal{A})$  une mesure notée  $\mu_X$  définie par :

$$\forall A \in \mathcal{A}, \mu_X(A) = \mu(X^{-1}(A)).$$

La mesure  $\mu_X$  est appelée *mesure image* de  $\mu$  par  $X$ .

### $\mu$ -négligeable, $\mu$ -presque partout

Une partie de  $\Omega$  est dite  $\mu$ -négligeable si elle est incluse dans un élément de mesure nulle de la tribu  $\mathcal{F}$ . Une fonction sur  $\Omega$  est dite  $\mu$ -négligeable si elle est nulle en dehors d'un ensemble  $\mu$ -négligeable.

Une propriété sur les points de  $E$  est dite vraie  $\mu$ -presque partout ( $\mu$ -p.p.) si elle est vraie pour tous les points de  $E$  exceptés ceux d'un ensemble  $\mu$ -négligeable. En particulier, deux fonctions sur  $\Omega$  sont égales  $\mu$ -p.p. si elles sont égales en dehors d'un ensemble négligeable.

### Mesure de Lebesgue sur $\mathbb{R}^d$

On montre qu'il existe une mesure  $\sigma$ -finie unique  $\mu_0$  sur  $(\mathbb{R}^d, \mathcal{B}^d)$  telle que

$$\mu_0\left(\prod_{i=1}^d [a_i, b_i]\right) = \prod_{i=1}^d (b_i - a_i), \quad \forall a_i < b_i, 1 \leq i \leq d.$$

Cette mesure est appelée *mesure de Lebesgue* sur  $(\mathbb{R}^d, \mathcal{B}^d)$ . Sur  $\mathbb{R}$ , elle mesure les longueurs, sur  $\mathbb{R}^2$ , les surfaces etc.

## IV.2.3 Probabilités, événements

### Vocabulaire

$P$  est une *mesure de probabilité*, ou encore une probabilité sur  $(\Omega, \mathcal{F})$ , si c'est une mesure telle que  $P(\Omega) = 1$ . Le triplet  $(\Omega, \mathcal{F}, P)$  constitue alors un *espace probabilisé*. En langage probabiliste,

tout élément de  $\mathcal{F}$  est appelé *événement*<sup>1</sup>. Une conjonction d'événements s'écrit indifféremment  $A \cap B = A.B = AB = (A \text{ et } B) = (A, B)$ . Pour une propriété vraie  $P$ -p.p, on dit plutôt presque sûrement (p.s.). Enfin, les éléments de  $\Omega$  sont appelés des *épreuves*.

### Qu'est-ce que le hasard ?

Dans une interprétation classique de la notion de hasard, chaque épreuve  $\omega$  est le résultat possible d'une « expérience aléatoire », au cours de laquelle tout événement  $A$  est *réalisé* si  $\omega \in A$ . Dans ce contexte formel,  $P(A)$  désigne le taux de réalisation de  $A$  quand on répète l'expérience un grand nombre de fois.

Cette interprétation permet de guider l'intuition, mais il faut reconnaître qu'elle ne constitue pas une définition mathématique de la notion d'« expérience aléatoire » ou de « tirage aléatoire ». En fait la théorie des probabilités *ne dit rien* sur la notion de hasard, ni sur la façon de le générer, ni sur la façon de l'authentifier. D'où l'utilisation récurrente de guillemets dans ce cours pour signaler des expressions faisant appel à l'intuition plutôt qu'à un concept précis.

A cet égard, on peut citer l'exemple de l'écriture décimale de nombres transcendants tels que  $\pi$ , qui a toutes les apparences de la suite des épreuves d'une expérience consistant à tirer un chiffre « au hasard », alors que le sens commun n'y voit qu'une succession déterministe. A l'inverse, on a l'habitude d'accorder à une suite de pile ou face un caractère aléatoire, bien qu'on admette qu'une pièce de monnaie respecte les équations déterministes de la physique classique.

### Événements indépendants

Deux événements  $A$  et  $B$  sont dits *indépendants* si  $P(A.B) = P(A)P(B)$ . Plus généralement,  $A_1, A_2, \dots, A_I$  sont indépendants dans leur ensemble si

$$\forall i, j, \dots, r \text{ (différents entre eux) }, \quad P(A_i.A_j.\dots.A_r) = P(A_i)P(A_j)\dots P(A_r).$$

Des événements indépendants deux à deux ne le sont pas nécessairement dans leur ensemble (il est facile de vérifier que le contraire est vrai).

### Probabilités conditionnelles

Soient deux événements  $A$  et  $B$  tels que  $P(B) > 0$ . La *probabilité conditionnelle de  $A$  sachant  $B$*  est définie par

$$P(A|B) = \frac{P(A.B)}{P(B)}. \quad (\text{IV.2})$$

Notons que l'indépendance de  $A$  et  $B$  équivaut à  $P(A|B) = P(A)$  quand  $P(B) > 0$ .

La probabilité conditionnelle est en fait une nouvelle mesure de probabilité définie sur l'espace probabilisé  $(B, \mathcal{F}_B)$ , où  $\mathcal{F}_B$  est une tribu sur  $B$  :  $\mathcal{F}_B = \{A.B, A \in \mathcal{F}\}$ .

Par récurrence sur (IV.2),  $\forall A_1, A_2, \dots, A_I \in \mathcal{F}^I$ , on a

$$P(A_1.A_2.\dots.A_I) = P(A_1)P(A_2|A_1)P(A_3|A_1.A_2)\dots P(A_I|A_1.\dots.A_{I-1}).$$

Soit  $\{A_i, i \in I\}$  une partition de  $\Omega$  constituée d'événements, et soit  $B$  un événement tel que  $P(B) > 0$ . Alors  $\{A_i.B, i \in I\}$  est une partition de  $B$  en événements (de  $\mathcal{F}_B$ ). De (IV.1) et (IV.2), on déduit alors *la règle des causes totales* :

$$P(B) = \sum_{i=1}^I P(B.A_i) = \sum_{i=1}^I P(A_i)P(B|A_i)$$

<sup>1</sup>Remarquer l'accent aigu : événement...

ainsi que la formule de Bayes :

$$P(A_j | B) = \frac{P(A_j \cdot B)}{P(B)} = \frac{P(A_j) P(B | A_j)}{\sum_{i=1}^I P(A_i) P(B | A_i)}.$$

### IV.3 Variables aléatoires

Dans toute la suite, on se place dans le cas d'un espace probabilisé  $(\Omega, \mathcal{F}, P)$ .

#### IV.3.1 Variables aléatoires

##### Définition

On appelle *variable aléatoire* (v.a.) toute application mesurable  $X$  sur un espace probabilisé  $(\Omega, \mathcal{F}, P)$ . L'image par  $X$  d'une épreuve  $\omega$ ,  $X(\omega)$ , est appelée une *réalisation* de  $X$ . La mesure image de  $P$  par  $X$ , notée  $P_X$ , est appelée *loi de  $X$* .

La distinction faite entre une variable aléatoire et sa loi paraît sans objet lorsqu'on utilise une seule variable aléatoire  $X$ , car on s'intéresse à la probabilité d'apparition des différentes réalisations de  $X$  et il est inutile de distinguer quelles épreuves donnent telles ou telles réalisations de  $X$ . Autrement dit, tout ce qui nous est utile au sujet de  $X$  est contenu dans sa loi. Mais dans la plupart des problèmes on a affaire à plusieurs variables aléatoires définies sur le même espace de probabilité et on s'intéresse entre autres aux relations qui existent entre ces variables. Il est alors tout à fait essentiel de bien voir que des variables aléatoires peuvent avoir la même loi sans être égales. Insistons encore sur le fait que lorsqu'on a deux v.a., par exemple à valeurs réelles,  $X_1$  et  $X_2$ , dont on connaît seulement pour chacune la loi, on ne peut rien dire en général sur les v.a.  $X_1 + X_2$ ,  $\min(X_1, X_2)$ ... tant qu'on n'a pas défini précisément les v.a.  $X_1$  et  $X_2$ , ou bien seulement la loi du couple  $(X_1, X_2)$  conformément au paragraphe IV.3.3.

##### Existence

Le paragraphe IV.3.2 est consacré à la caractérisation de v.a. réelles. Réciproquement, les problèmes d'existence de v.a. de loi donnée (et plus généralement de processus aléatoires de loi donnée) sont liés à des problèmes fondamentaux d'existence et de construction de mesures, qui sont résolus par des théorèmes de théorie de la mesure dépassant largement le cadre de ce cours. On peut dire simplement que ces théorèmes (en particulier les théorèmes d'extension de Hopf et de Kolmogorov [5]) permettent de construire des mesures et de probabiliser des espaces mesurables tels que  $(\mathbb{R}^N, \mathcal{B}^{\otimes N})$  et même  $(\mathbb{R}^{\mathbb{R}}, \mathcal{B}^{\otimes \mathbb{R}})$ , suffisamment vastes pour permettre la construction d'autant de variables aléatoires puis d'autant de signaux aléatoires qu'on le souhaite. C'est pourquoi en pratique, on laisse en suspens la spécification de l'espace d'épreuves  $\Omega$ , de sa tribu  $\mathcal{F}$  et de la probabilité « générique »  $P$ , en supposant que  $(\Omega, \mathcal{F}, P)$  est assez vaste pour engendrer simultanément tous les modèles nécessaires.

#### IV.3.2 Variables aléatoires réelles

##### Définition

Le cas particulier des v.a. *réelles* (v.a.r.) désigne les applications mesurables  $X$  à valeurs dans l'espace  $(\mathbb{R}, \mathcal{B})$ . Les v.a.r. sont les modèles les plus courants pour décrire des grandeurs « incertaines ». Les v.a. à valeurs complexes sont également utiles ; on peut les considérer comme des couples de v.a.r. (voir le paragraphe IV.3.4).

### Première caractérisation de la loi d'une v.a.r. : fonction de répartition

La loi d'une v.a.r.  $X$  est donc une mesure de probabilité sur les boréliens de  $\mathbb{R}$ . En particulier, elle permet de calculer la probabilité de tous les événements de la forme  $\{\omega, X(\omega) \leq x\}$ , abrégés en  $(X \leq x)$ . Réciproquement, comme les demi-droites  $] -\infty, x]$  génèrent les boréliens de  $\mathbb{R}$ , on obtient la caractérisation suivante :

La loi d'une v.a.r.  $X$  est entièrement caractérisée par la *fonction de répartition* de  $X$ , *croissante entre 0 et 1 et continue à droite* :

$$F_X(x) \triangleq P(X \leq x).$$

On en déduit par exemple  $P(a < X \leq b) = F_X(b) - F_X(a)$ .

Selon les caractéristiques de la fonction de répartition  $F_X(x)$ , on distingue principalement deux types « classiques » de v.a.r. Ces deux types, variables à *densité* et variables *discrètes*, ne recouvrent pas l'ensemble des v.a.r., mais correspondent à un très grand nombre de situations.

### Variables aléatoires discrètes

Dans le cas de v.a. discrètes,  $X$  prend un nombre fini ou dénombrable de valeurs  $x_i$  avec les probabilités  $P(X = x_i) = p_i$ , où nécessairement  $\sum_i p_i = 1$ . Il s'agit en particulier des v.a.r. construites par dénombrement ( $x_i \in \mathbb{N}$ ). Dans ces conditions,  $F_X$  est la fonction en escalier :

$$F_X(x) = \sum_{i: x_i \leq x} p_i,$$

et par conséquent la donnée des  $p_i$  (et des  $x_i$ ) caractérise la loi de  $X$ .

### Variables aléatoires à densité

La loi de  $X$  est dite à *densité* s'il existe une fonction  $f_X(x) \geq 0$ ,  $dx$ -intégrable, telle que pour tout réel  $x$  :

$$F_X(x) = \int_{-\infty}^x f_X(s) ds. \quad (\text{IV.3})$$

La fonction  $f_X(x)$  est appelée *densité de probabilité* de la variable aléatoire  $X$ . Une condition suffisante pour que  $X$  admette une densité est que  $F_X(x)$  soit continûment dérivable pour tout  $x$ . Dans ce cas, (IV.3) montre que la densité caractérise la loi de  $X$ .

*Remarque.* Il ne faut pas confondre la densité de probabilité  $f_X(x)$  avec la mesure de probabilité d'un événement du type  $(X = x)$ . Par exemple, dans le cas d'un espace  $\Omega = [0, 1]$  muni de sa tribu borélienne et de la mesure de Lebesgue, la v.a.r.  $X(\omega) = \cos \pi\omega$  admet une densité de probabilité

$$f_X(x) = \begin{cases} \frac{1}{\pi\sqrt{1-x^2}} & \text{pour } x \in ]-1, 1[, \\ 0 & \text{ailleurs.} \end{cases}$$

En effet,

$$\begin{aligned} F_X(x) &= P(X \leq x), \\ &= P(\omega \geq \arccos(x))/\pi \quad \forall x \in ]-1, 1[, \\ &= 1 - \arccos(x)/\pi. \end{aligned}$$

Cette fonction de répartition  $F_X(x)$  est continûment dérivable sur  $] -1, 1[$ ; sa dérivée est la densité de probabilité  $f_X(x)$  ci-dessus. Il est facile de voir que  $f_X(x)$  prend des valeurs arbitrairement grandes quand  $x$  se rapproche des extrémités de l'intervalle  $] -1, 1[$ , ce qui n'est pas

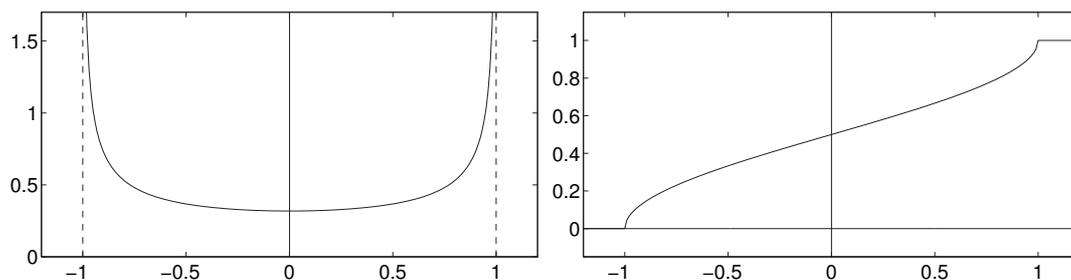


Fig. IV.1. Densité de probabilité (à gauche) et fonction de répartition (à droite) d'une variable aléatoire (voir texte)

incompatible avec le fait que tout événement  $(X = x)$  est de probabilité nulle ( $y$  compris pour  $x = \pm 1$ ).

### Changement de variable

Bien souvent il faut calculer la densité de probabilité d'une variable aléatoire réelle  $Y$  définie comme la transformation  $Y = g(X)$  d'une autre v.a.r.  $X$  dont la densité  $f_X$  est connue. Ce genre de problème peut souvent être résolu par de simples manipulations des fonctions de répartition de  $Y$  et  $X$ , et non pas directement des densités. Prenons un exemple : soit  $X$  une v.a. de fonction de répartition  $F_X$  bijective. Calculons la loi de la v.a.  $Y = F_X(X)$ . Il est clair que  $Y$  est à valeurs dans l'intervalle  $[0, 1]$  et pour  $y \in [0, 1]$ , il vient :

$$F_Y(y) = P(Y \leq y) = P(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y.$$

Donc  $Y$  est uniformément répartie sur  $[0, 1]$ . Cet exemple est intéressant car il fournit un moyen de tirer « au hasard » des nombres suivant la loi de  $X$  à l'aide d'un simple générateur de nombres pseudo-aléatoires suivant la loi uniforme sur  $[0, 1]$  (c'est la fonction *rand* disponible sur de nombreux calculateurs). Le calcul précédent montre qu'il suffit d'appliquer aux nombres générés par *rand* la fonction  $F_X^{-1}$  pour obtenir des nombres distribués suivant la loi de  $X$ .

Pour en revenir aux changements de variable, il existe sous certaines conditions de régularité sur la transformation  $Y = g(X)$  une formule plus générale de changement de variable qui sera donnée dans le cas vectoriel au paragraphe IV.5.2

### IV.3.3 Espérance mathématique

#### Définition

Pour toute v.a.r.  $X$   $P$ -intégrable ( $\int_{\Omega} |X(\omega)| dP(\omega) < +\infty$ ), on définit l'espérance mathématique par :

$$\boxed{E(X) \triangleq \int_{\Omega} X(\omega) dP(\omega) = \int_{\mathbb{R}} x dP_X(x)} \quad (\text{IV.4})$$

L'espérance mathématique est un réel qui correspond à une moyenne de la v.a.r.  $X$  sur l'ensemble des épreuves  $\omega$ , pondérée par la probabilité  $dP(\omega)$ . La deuxième égalité traduit l'intégration pour la mesure image. Pour des v.a.r., on peut noter indifféremment le terme  $dP_X(x)$  sous la forme  $dF_X(x)$ , notation qu'on adopte dans la suite.

— Dans le cas de variables discrètes, on obtient pour l'espérance une forme classique de

moyenne pondérée par la probabilité des événements :

$$E(X) = \sum_i p_i x_i.$$

— Si la loi de  $X$  est à densité  $f_X(x)$ , on peut remplacer  $dF_X(x)$  par  $f_X(x) dx$  et se ramener ainsi aux notations classiques de l'intégration de fonctions de  $\mathbb{R}$  sur  $\mathbb{R}$  par rapport à la mesure de Lebesgue :

$$E(X) = \int_{\mathbb{R}} x dF_X(x) = \int_{\mathbb{R}} x f_X(x) dx.$$

L'espérance mathématique, lorsqu'elle est définie, possède les propriétés de linéarité et de monotonie ( $X \leq Y$  P-p.s.  $\Rightarrow E(X) \leq E(Y)$ ) de l'intégrale de Lebesgue.

La définition (IV.4) permet aussi d'exprimer, quand elle existe, l'espérance d'une v.a.  $\varphi(X)$  fonction d'une autre v.a.  $X$ , sous la forme

$$E(\varphi(X)) = \int_{\Omega} \varphi(X(\omega)) dP(\omega) = \int_{\mathbb{R}} \varphi(x) dF_X(x). \quad (\text{IV.5})$$

### Deuxième caractérisation de la loi d'une v.a.r. : fonction caractéristique

Pour toute v.a.r.  $X$  et tout  $u \in \mathbb{R}$ , on peut définir la quantité complexe  $E(e^{iuX})$ , intégrale absolument convergente :

La *fonction caractéristique* de  $X$  est l'application  $\Phi_X$  définie de  $\mathbb{R}$  dans  $\mathbb{C}$  par

$$u \rightarrow \Phi_X(u) \triangleq E(e^{iuX}) = \int_{\mathbb{R}} e^{iux} dF_X(x).$$

Dans le langage de l'intégration, cette fonction est (à une affinité  $u = -2\pi\nu$  près) la transformée de Fourier de la mesure de  $X$ . La transformée de Fourier des mesures étant injective, il suffit de connaître la fonction caractéristique de  $X$  pour connaître la loi de  $X$ . Comme la fonction de répartition, la fonction caractéristique permet donc de caractériser la loi d'une v.a.r. (d'où son nom).

### Variable aléatoire et moments d'ordre $p$

La v.a.r.  $X$  est dite d'ordre  $p$ , entier positif, si

$$E(|X|^p) = \int_{\Omega} |X(\omega)|^p dP(\omega) = \int_{\mathbb{R}} |x|^p dF_X(x) < +\infty.$$

On peut alors définir  $m_p = E(X^p)$  appelé *moment d'ordre  $p$*  de la variable aléatoire  $X$ .

On définit les espaces  $L^p(\Omega, \mathcal{F}, P)$ ,  $p \geq 1$ , comme les ensembles des classes d'équivalence des fonctions  $X$  égales P-p.p. et telles que  $|X|^p$  est intégrable. La norme  $\|\cdot\|_p$  définie par  $\|X\|_p = (E(|X|^p))^{1/p}$  confère aux espaces  $L^p(\Omega, \mathcal{F}, \mu)$  une structure de Banach. L'existence du moment d'ordre  $p$  entraîne celle de tous les moments d'ordre inférieur, du fait de l'emboîtement des espaces  $L^p(\Omega, \mathcal{F}, P)$  :

$$L^1(\Omega, \mathcal{F}, P) \supseteq L^2(\Omega, \mathcal{F}, P) \supseteq \dots \supseteq L^p(\Omega, \mathcal{F}, P) \supseteq \dots \supseteq L^\infty(\Omega, \mathcal{F}, P).$$

A la section IV.6, nous étudierons plus en détail le cas de  $L^2(\Omega, \mathcal{F}, P)$ , qui possède la structure d'un espace de Hilbert.

### Lien entre la fonction caractéristique et les moments d'une v.a.r.

L'existence des moments jusqu'à l'ordre  $p$  est liée à la régularité en zéro de la fonction caractéristique : si une variable aléatoire possède un moment d'ordre  $p$ , alors sa fonction caractéristique est  $p$  fois dérivable. Par identification du développement de Taylor-Mac Laurin à l'ordre  $p$ , on trouve

$$\Phi_X^{(p)}(0) = i^p m_p.$$

Réciproquement, la dérivabilité d'ordre  $p$  de  $\Phi_X$  en 0 entraîne l'existence de moments d'ordre  $p-1$  (et même d'ordre  $p$  si  $p$  est pair). En pratique, la fonction caractéristique est donc un outil très puissant pour calculer la suite des moments d'une variable aléatoire, s'ils existent.

Avec les mêmes conditions d'existence que celles des moments, on définit aussi les moments centrés  $m_{cp} = E((X - m_1)^p)$ . Le cas des variables d'ordre deux donne lieu à des appellations spécifiques :

Moyenne :	$m = E(X) = m_1$
Variance :	$\text{var}(X) = m_{c2} = E((X - m)^2) = E(X^2) - E(X)^2 = m_2 - m^2$
Écart-type :	$\sigma = \sqrt{\text{var}(X)}$

### Variable aléatoire gaussienne

Une variable aléatoire obéit à la *loi de Gauss*  $\mathcal{N}(m, \sigma^2)$  si sa densité de probabilité est de la forme

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

On vérifie qu'une telle variable est du second ordre, de moyenne  $m$  et d'écart-type  $\sigma$ . La variable est dite réduite si  $m = 0$  et  $\sigma = 1$ . La fonction caractéristique associée à une variable aléatoire gaussienne est de la forme

$$\Phi_X(u) = \exp\left(imu - \frac{u^2 \sigma^2}{2}\right).$$

## IV.4 Couple de variables aléatoires réelles

### IV.4.1 Caractérisation et propriétés

#### Fonction de répartition

Soient  $X$  et  $Y$  deux v.a.r. La loi du couple  $(X, Y)$  est une probabilité sur  $\mathbb{R}^2$  qui peut être caractérisée par la fonction de répartition

$$F_{XY}(x, y) \triangleq P(X \leq x, Y \leq y).$$

Étant donné la loi du couple  $(X, Y)$ , on déduit les *lois marginales* de  $X$  et de  $Y$  par les fonctions de répartition

$$F_X(x) = \lim_{y \rightarrow +\infty} F_{XY}(x, y) \text{ et } F_Y(y) = \lim_{x \rightarrow +\infty} F_{XY}(x, y).$$

Mais il est clair que la donnée des lois marginales de  $X$  et de  $Y$  ne suffit pas à reconstituer la loi *jointe* de  $(X, Y)$ .

### Densité

La loi de  $(X, Y)$  est à *densité* s'il existe une fonction  $f_{XY}(x, y) \geq 0$ , intégrable, telle que pour tout couple  $(x, y)$  de réels :

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(s, t) ds dt.$$

Si la fonction de répartition  $F_{XY}$  est deux fois continûment différentiable, la densité de probabilité du couple  $(X, Y)$  est définie par

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}}{\partial x \partial y}(x, y).$$

On peut alors vérifier que les lois marginales sont à densité, et que les *densités marginales* de  $X$  et de  $Y$  s'écrivent :

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy \quad \text{et} \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx.$$

### Fonction caractéristique

La notion de fonction caractéristique se généralise également :

$$\Phi_{XY}(u, v) = \mathbb{E}(e^{i(uX+vY)}) = \int_{\mathbb{R}^2} e^{i(ux+vy)} dF_{XY}(x, y);$$

les fonctions caractéristiques marginales sont évidemment  $\Phi_{XY}(u, 0)$  et  $\Phi_{XY}(0, v)$ .

### Variables indépendantes

Les v.a.r.  $X$  et  $Y$  sont dites indépendantes ssi  $\forall (A, B) \in \mathcal{B}^2$ , les événements  $\{X(\omega) \in A\}$  et  $\{Y(\omega) \in B\}$  sont indépendants. Mais cette condition est malcommode. Une condition équivalente est la séparabilité de la fonction de répartition  $F_{XY}(x, y) = F_X(x)F_Y(y)$ , ou encore de celle de la fonction caractéristique  $\Phi_{XY}(u, v) = \Phi_X(u)\Phi_Y(v)$ . Si  $(X, Y)$  est à densité, alors la condition d'indépendance s'écrit aussi

$$f_{XY}(x, y) = f_X(x)f_Y(y).$$

### Espérance

La notion d'espérance définie pour une v.a.r. au paragraphe IV.3.3 se généralise au cas d'un couple de v.a.r. de façon naturelle :

$$\begin{aligned} \mathbb{E}(X, Y) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (x, y) dF_{XY}(x, y) \\ &= \left( \int_{\mathbb{R}} x dF_X(x), \int_{\mathbb{R}} y dF_Y(y) \right) = (\mathbb{E}(X), \mathbb{E}(Y)) \quad (\text{Fubini}), \end{aligned}$$

c'est-à-dire que l'espérance d'un couple de v.a.r. s'obtient comme le couple des espérances *marginales*, pourvu qu'elles existent, que les v.a.r. soient indépendantes ou non.

### Covariance ou corrélation

Pour  $X$  et  $Y$  du second ordre, on définit la *covariance*, appelée également *corrélation*

$$\begin{aligned}\operatorname{cov}(X, Y) &\triangleq E((X - E(X))(Y - E(Y))) \\ &= E(XY) - E(X)E(Y).\end{aligned}$$

D'après l'inégalité de Schwarz dans  $L^2(\Omega, \mathcal{F}, P)$  démontrée plus loin (§ IV.6), il est facile de voir que le coefficient de corrélation

$$\rho_{XY} \triangleq \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}}$$

varie entre  $-1$  (atteint pour  $X = -Y$ ) et  $1$  (pour  $X = Y$ ), en passant par le cas de la *décorrélacion* entre  $X$  et  $Y$  :

$$X \text{ et } Y \text{ décorrélées} \iff \operatorname{cov}(X, Y) = 0. \quad (\text{IV.6})$$

L'indépendance des composantes d'un couple d'ordre deux implique leur décorrélacion, mais la réciproque est fautive.

#### IV.4.2 Conditionnement

##### Définition

La notion de conditionnement d'une variable aléatoire par une autre est très générale et très utile en statistique, en théorie de l'estimation, en traitement du signal. Mais elle pose des problèmes de définition inattendus, dont la résolution fait appel à des résultats très puissants de la théorie des probabilités. Nous nous contenterons ici de présenter cette notion dans deux cas particuliers : celui des v.a.r. à valeurs discrètes et celui des v.a.r. à densité.

Soit un couple de variables aléatoires discrètes  $(X, Y)$  à valeurs  $(x_i, y_j)$  dans un sous-ensemble dénombrable de  $\mathbb{R}^2$ . Conformément à la définition de la probabilité conditionnelle (IV.2), la loi de  $X$  conditionnellement à  $Y = y_j$  est entièrement décrite par

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} \quad \text{si } P(Y = y_j) > 0, \quad 0 \quad \text{sinon.}$$

Pour chaque  $y_j$  fixé tel que  $P(Y = y_j) > 0$ ,  $P(X = x_i | Y = y_j)$  définit une nouvelle mesure de probabilité sur  $\{x_i\}$ .

Dans le cas d'un couple  $(X, Y)$  de variables aléatoires à densité, on souhaite à nouveau conditionner les valeurs prises par  $X$  par des événements du type  $(Y = y)$ . Mais dans ce cas, ces événements sont de mesure nulle et (IV.2) n'est pas applicable. Pour résoudre cette difficulté, on peut tout d'abord définir

$$P(X \leq x | y < Y \leq y + dy) = \frac{F_{XY}(x, y + dy) - F_{XY}(x, y)}{F_Y(y + dy) - F_Y(y)}.$$

Comme  $(X, Y)$  admet une densité, cette quantité converge quand  $dy$  tend vers 0 ; la limite définit la *fonction de répartition de  $X$  conditionnelle à  $Y$*  :

$$F_{X|Y}(x | y) \triangleq \frac{1}{f_Y(y)} \frac{\partial F_{XY}}{\partial y}(x, y),$$

symboliquement,

$$F_{X|Y}(x | y) = P(X \leq x | Y = y).$$

Cette fonction de répartition définit une mesure de probabilité à densité. La densité est appelée *densité conditionnelle de  $X$  par rapport à  $Y$*  et notée  $f_{X|Y}(x|y)$ . En dérivant  $F_{X|Y}(x|y)$  par rapport à  $x$ , on obtient :

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}. \quad (\text{IV.7})$$

### Espérance conditionnelle

Plaçons-nous dans le cas d'un couple  $(X, Y)$  de variables aléatoires à densité, et reprenons les notations du paragraphe précédent. Pour  $X$  du premier ordre et pour tout réel  $y$  tel que  $f_Y(y) > 0$ , on définit l'*espérance conditionnelle de  $X$  sachant  $Y = y$*  :

$$E_{X|Y}(X|Y=y) \triangleq \int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx,$$

qui est une fonction déterministe de  $y$  (en cas d'ambiguïté, l'indice de l'espérance rappelle la loi par rapport à laquelle l'intégration est effectuée). L'espérance conditionnelle, en tant qu'opérateur sur la v.a.r.  $X$ , possède les propriétés de linéarité et de monotonie de l'intégrale de Lebesgue. En tant que fonction de  $y$  elle permet de définir une *variable aléatoire*  $E(X|Y)$  qui est une fonction déterministe de  $Y$ .

*Remarques :*

— Le passage par l'espérance conditionnelle est souvent commode pour calculer une espérance simple telle que  $E(\varphi(X, Y))$ , où  $\varphi$  est une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$   $P_{XY}$ -intégrable. En effet, on a la formule du *conditionnement successif* :

$$E(\varphi(X, Y)) = E_Y(E_{X|Y}(\varphi(X, Y)|Y)) = E_X(E_{Y|X}(\varphi(X, Y)|X)). \quad (\text{IV.8})$$

Montrons ce résultat dans le cas d'un couple à densité :

$$\begin{aligned} E(\varphi(X, Y)) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi(x, y) f_{XY}(x, y) dx dy \\ &= \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \varphi(x, y) f_{X|Y}(x|y) dx \right) f_Y(y) dy, \end{aligned}$$

d'après (IV.7) et le théorème de Fubini, ce qui correspond à la première égalité dans (IV.8). En particulier, ce résultat montre que  $E(X|Y)$  est  $P_Y$ -intégrable et que son espérance vaut :

$$E_Y(E_{X|Y}(X|Y)) = E(X).$$

— Si  $X$  et  $Y$  sont indépendantes on a :  $E_{X|Y}(X|Y) = E(X)$ , quantité déterministe.

— Si  $g$  est une application mesurable de  $(\mathbb{R}, \mathcal{B}, P_Y) \rightarrow \mathbb{R}$  :  $E_{Y|Y}(g(Y)|Y) = g(Y)$ .

— On a vu que l'espérance conditionnelle de  $X$  sachant  $Y$  est linéaire en  $X$  ; par contre, en général,  $E_{X|Y}(X|Y)$  est une fonction non linéaire (et même non affine) de  $Y$ .

Le paragraphe IV.6 donne une interprétation « géométrique » fondamentale de l'espérance conditionnelle dans l'espace de Hilbert  $L^2(\Omega, \mathcal{F}, P)$ .

## IV.5 Vecteurs aléatoires réels

### IV.5.1 Généralisations

Un vecteur aléatoire réel  $\mathbf{X}$  est constitué de  $n$  v.a.r.  $\mathbf{X} = (X_1, \dots, X_n)^t$ <sup>2</sup>. C'est une application mesurable de  $\Omega$  dans  $\mathbb{R}^n$  (donc c'est aussi une v.a.). On peut remarquer aussi que si

<sup>2</sup>La transposée d'une matrice  $M$  est notée ici  $M^t$ , et non pas  ${}^tM$ .

$\mathbf{X}_c$  est un vecteur aléatoire à valeurs dans  $\mathbb{C}^m$ ,  $m \geq 1$ ,  $\mathbf{X}_c$  peut s'écrire sous la forme  $\mathbf{X}_r + i\mathbf{X}_i$  avec  $\mathbf{X}_r$  et  $\mathbf{X}_i$  deux vecteurs aléatoires à valeurs dans  $\mathbb{R}^m$ . On peut donc identifier  $\mathbb{C}$  à  $\mathbb{R}^2$ ,  $\mathbf{X}_c$  au vecteur aléatoire  $(\mathbf{X}_r^t, \mathbf{X}_i^t)^t$  à valeurs dans  $\mathbb{R}^{2m}$ , et ramener l'étude des vecteurs aléatoires à valeurs complexes à celle des vecteurs aléatoires réels.

La plupart des notions définies pour des couples de variables aléatoires se généralisent aux vecteurs aléatoires.

— La *fonction de répartition* d'un vecteur aléatoire  $\mathbf{X} = (X_1, \dots, X_n)^t$  est une généralisation de la définition de cette fonction pour un couple de v.a.r. aléatoires :

$$F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}}(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

— Un vecteur aléatoire  $\mathbf{X}$  admet une *densité*  $f_{\mathbf{X}}(\mathbf{x})$  si sa fonction de répartition s'exprime sous la forme

$$F_{\mathbf{X}}(\mathbf{x}) = \int_B f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x},$$

avec  $f_{\mathbf{X}}(\mathbf{x}) \geq 0$ ,  $d\mathbf{x} = dx_1 \dots dx_n$  est la mesure de Lebesgue sur  $\mathbb{R}^n$  et  $B = \prod_{i=1}^n ]-\infty, x_i]$ .

— La *fonction caractéristique* du vecteur aléatoire  $\mathbf{X}$  est la fonction  $\mathbf{u} \rightarrow \Phi_{\mathbf{X}}(\mathbf{u})$  définie sur  $\mathbb{R}^n$  à valeurs dans  $\mathbb{C}$  par

$$\Phi_{\mathbf{X}}(\mathbf{u}) = E(e^{i\mathbf{u}^t \mathbf{X}}) = \int_{\mathbb{R}^n} e^{i\mathbf{u}^t \mathbf{X}} P_{\mathbf{X}}(d\mathbf{x}).$$

— Deux vecteurs aléatoires  $\mathbf{X}$  et  $\mathbf{Y}$ , respectivement à valeurs dans  $\mathbb{R}^m$  et  $\mathbb{R}^n$ , sont indépendants si la fonction de répartition du vecteur  $\mathbf{Z} = (\mathbf{X}^t, \mathbf{Y}^t)^t$  est séparable sous la forme

$$F_{\mathbf{Z}}(\mathbf{x}, \mathbf{y}) = F_{\mathbf{X}}(\mathbf{x}) F_{\mathbf{Y}}(\mathbf{y}),$$

ou, de manière équivalente, si la fonction caractéristique  $\Phi_{\mathbf{Z}}(\mathbf{u}, \mathbf{v})$  de  $\mathbf{Z}$  s'écrit

$$\Phi_{\mathbf{Z}}(\mathbf{u}, \mathbf{v}) = \Phi_{\mathbf{X}}(\mathbf{u}) \Phi_{\mathbf{Y}}(\mathbf{v}).$$

— Pour un vecteur  $\mathbf{Z} = (\mathbf{X}^t, \mathbf{Y}^t)^t$  admettant une densité, on peut introduire la notion de conditionnement du vecteur  $\mathbf{X}$  par le vecteur  $\mathbf{Y}$ . En particulier, la densité conditionnelle de  $\mathbf{X}$  sachant  $\mathbf{Y}$  généralise (IV.7) :

$$f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{f_{\mathbf{Z}}(\mathbf{x}, \mathbf{y})}{f_{\mathbf{Y}}(\mathbf{y})}$$

et l'espérance conditionnelle s'écrit pour  $\mathbf{X} = (X_1, \dots, X_n)^t$  du premier ordre et pour tout  $\mathbf{Y}$  tel que  $f_{\mathbf{Y}}(\mathbf{y}) > 0$  :

$$E(\mathbf{X} | \mathbf{Y} = \mathbf{y}) = (E(X_1 | \mathbf{Y} = \mathbf{y}), \dots, E(X_n | \mathbf{Y} = \mathbf{y}))^t,$$

$$E(X_k | \mathbf{Y} = \mathbf{y}) = \int_{-\infty}^{+\infty} x f_{X_k|\mathbf{Y}}(x|\mathbf{y}) dx.$$

## IV.5.2 Changement de variable

Soit un vecteur aléatoire  $\mathbf{X} = (X_1, \dots, X_n)^t$  à valeurs dans un ouvert  $U$  de  $\mathbb{R}^n$ , admettant la densité de probabilité  $f_{\mathbf{X}}$  par rapport à la mesure de Lebesgue, telle que

$$P(\mathbf{X} \in U) = \int_U f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = 1.$$

Soit  $g$  une fonction de classe  $C^1$  bijective de  $U$  dans un ouvert  $V$  de  $\mathbb{R}^n$ . On note  $J_g(\mathbf{x})$  son déterminant jacobien, défini en tout  $\mathbf{x}$  de  $U$  comme le déterminant de la matrice carrée d'ordre

$n$  de terme général  $(\partial g_i / \partial x_j)(\mathbf{x})$ . On suppose que  $J_g$  ne s'annule pas sur  $U$ . On suppose de plus que l'application réciproque  $g^{-1}$  est de classe  $C^1$  de  $V$  dans  $U$  et on note  $J_{g^{-1}}(\mathbf{y})$  le déterminant jacobien de  $g^{-1}$ , pour  $\mathbf{y} \in V$ .

Sous ces hypothèses, on montre que le vecteur aléatoire  $\mathbf{Y} = g(\mathbf{X})$  admet une densité de probabilité  $f_{\mathbf{Y}}$ . Pour  $h$  une v.a. positive sur  $V$  on a :

$$\mathbb{E}(h(\mathbf{Y})) = \int_V h(\mathbf{y}) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} = \mathbb{E}(h(g(\mathbf{X}))) = \int_U h(g(\mathbf{x})) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}.$$

On effectue dans la seconde intégrale le changement de variable  $\mathbf{x} \rightarrow \mathbf{y}$  pour obtenir

$$\int_V h(\mathbf{y}) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} = \int_V h(\mathbf{y}) f_{\mathbf{X}}(g^{-1}(\mathbf{y})) |J_{g^{-1}}(\mathbf{y})| d\mathbf{y},$$

et on identifie la densité de  $\mathbf{Y}$  :

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{y})) |J_{g^{-1}}(\mathbf{y})|, \quad \text{pour } \mathbf{y} \in V. \quad (\text{IV.9})$$

*Cas particulier : transformation affine d'un vecteur aléatoire.* Lorsque  $\mathbf{Y} = \mathbf{G}\mathbf{X} + \mathbf{b}$ , avec  $\mathbf{G}$  une matrice carrée d'ordre  $n$  régulière et  $\mathbf{b}$  un vecteur de  $\mathbb{R}^n$ , la densité de  $\mathbf{Y}$  s'écrit :

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{G}^{-1}(\mathbf{y} - \mathbf{b})) |\det(\mathbf{G})|^{-1}. \quad (\text{IV.10})$$

### IV.5.3 Vecteurs aléatoires d'ordre deux

Le vecteur aléatoire  $\mathbf{X}$  est dit d'ordre deux si

$$\mathbb{E}(\mathbf{X}^t \mathbf{X}) = \int_{\mathbb{R}^n} \|\mathbf{x}\|^2 P_{\mathbf{X}}(d\mathbf{x}) < +\infty,$$

ou, de façon équivalente, si chacune de ses composantes est une variable aléatoire d'ordre deux.

— La *moyenne* de  $\mathbf{X}$  est le vecteur de  $\mathbb{R}^n$   $\mathbf{m}_{\mathbf{X}} = \mathbb{E}(\mathbf{X})$ , dont les composantes sont

$$\mathbf{m}_{\mathbf{X}}(j) = \int_{\mathbb{R}^n} x_j P_{\mathbf{X}}(d\mathbf{x}) = \int_{\mathbb{R}} x_j P_{X_j}(dx_j) = \mathbb{E}(X_j).$$

— La *matrice de covariance*  $\Gamma_{\mathbf{X}}$  est une matrice  $n \times n$  définie par

$$\Gamma_{\mathbf{X}} = \mathbb{E}(\mathbf{X} \mathbf{X}^t), \quad (\text{IV.11})$$

dont les éléments sont

$$\Gamma_{\mathbf{X}}(j, k) = \mathbb{E}(X_j X_k) = \int_{\mathbb{R}^n} x_j x_k P_{\mathbf{X}}(d\mathbf{x}) = \int_{\mathbb{R}^2} x_j x_k dP_{X_j X_k}(x_j, x_k).$$

On peut vérifier aisément que la matrice de covariance est symétrique, définie non négative et que sa trace est égale à  $\mathbb{E}(\mathbf{X}^t \mathbf{X})$ .

— La *matrice de corrélation, ou covariance centrée* est la matrice de covariance du vecteur aléatoire centré  $\mathbf{X} - \mathbf{m}_{\mathbf{X}}$  :

$$\mathbf{R}_{\mathbf{X}} = \mathbb{E}((\mathbf{X} - \mathbf{m}_{\mathbf{X}})(\mathbf{X} - \mathbf{m}_{\mathbf{X}})^t), \quad (\text{IV.12})$$

dont les composantes sont  $\mathbf{R}_{\mathbf{X}}(j, k) = \text{cov}(X_j, X_k)$ .

— Soit deux vecteurs aléatoires  $\mathbf{X}$  et  $\mathbf{Y}$  d'ordre deux, respectivement à valeurs dans  $\mathbb{R}^m$  et  $\mathbb{R}^n$ . La *matrice d'intercorrélation* est la matrice  $m \times n$  définie par

$$\mathbf{R}_{\mathbf{X}\mathbf{Y}} = \mathbb{E}((\mathbf{X} - \mathbf{m}_{\mathbf{X}})(\mathbf{Y} - \mathbf{m}_{\mathbf{Y}})^t). \quad (\text{IV.13})$$

Les vecteurs  $\mathbf{X}$  et  $\mathbf{Y}$  sont dits décorrélés si  $R_{\mathbf{X}\mathbf{Y}} = 0$ .

*Remarque.* Suivant les auteurs, les termes « covariance » et « corrélation » sont souvent utilisés l'un pour l'autre. Dans le cas fréquent de variables centrées, il est inutile de les distinguer. Dans les autres cas, il appartient au lecteur d'en vérifier la définition.

— Considérons à nouveau les *transformations affines d'un vecteur aléatoire*, cette fois dans le cas d'un vecteur  $\mathbf{X}$  d'ordre deux :  $\mathbf{Y} = G\mathbf{X} + \mathbf{b}$ , en reprenant les notations du paragraphe IV.5.2. Alors

$$\mathbf{m}_{\mathbf{Y}} = G \mathbf{m}_{\mathbf{X}} + \mathbf{b} \quad (\text{IV.14})$$

et

$$\begin{aligned} R_{\mathbf{Y}} &= E((G\mathbf{X} + \mathbf{b} - \mathbf{m}_{\mathbf{Y}})(G\mathbf{X} + \mathbf{b} - \mathbf{m}_{\mathbf{Y}})^t) \\ &= G R_{\mathbf{X}} G^t, \end{aligned} \quad (\text{IV.15})$$

et le résultat est évidemment valable pour une matrice  $G$  quelconque (éventuellement rectangulaire) : les caractéristiques d'ordre deux se « propagent » par transformation affine beaucoup plus facilement que la loi complète. Nous exploiterons souvent le résultat (IV.14)-(IV.15) dans la suite.

#### IV.5.4 Vecteurs aléatoires gaussiens

Les vecteurs gaussiens jouent un rôle considérable en traitement du signal en raison de leur facilité de manipulation, de leurs propriétés de stabilité et du fait qu'ils modélisent de nombreux phénomènes physiques mettant en jeu un très grand nombre de contributions élémentaires.

Un vecteur aléatoire  $\mathbf{X}$  est dit gaussien si toute combinaison affine de ses composantes est une variable aléatoire gaussienne.  $\mathbf{X}$  est donc d'ordre deux.

La fonction caractéristique d'un vecteur aléatoire gaussien est

$$\Phi_{\mathbf{X}}(\mathbf{u}) = E(e^{i\mathbf{u}^t \mathbf{X}}) = \exp\left(i\mathbf{u}^t \mathbf{m}_{\mathbf{X}} - \frac{1}{2} \mathbf{u}^t R_{\mathbf{X}} \mathbf{u}\right) \quad (\text{IV.16})$$

où  $R_{\mathbf{X}}$  est la matrice de covariance de  $\mathbf{X}$ .  $\mathbf{X}$  est donc entièrement défini à l'aide de ses caractéristiques à l'ordre deux. Réciproquement, la donnée d'un vecteur  $\mathbf{m}$  quelconque et d'une matrice  $R$  définie non négative permet de définir complètement un vecteur gaussien, et l'on note en abrégé :  $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, R)$  pour «  $\mathbf{X}$  vecteur gaussien de moyenne  $\mathbf{m}$  et de covariance  $R$  ».

— Lorsque  $R$  est inversible,  $\mathbf{X}$  admet une densité de probabilité  $f_{\mathbf{X}}$  :

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(R)}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m})^t R^{-1}(\mathbf{x} - \mathbf{m})\right). \quad (\text{IV.17})$$

— Deux vecteurs  $\mathbf{X}$  et  $\mathbf{Y}$  sont dits gaussiens dans leur ensemble si le vecteur joint  $\mathbf{Z} = (\mathbf{X}^t, \mathbf{Y}^t)^t$  est gaussien. Si deux vecteurs gaussiens dans leur ensemble  $\mathbf{X}$  et  $\mathbf{Y}$  sont décorrélés, alors ils sont indépendants. Ceci est très facile à voir à partir de la fonction caractéristique : supposons les deux vecteurs centrés, et notons

$$R_{\mathbf{Z}} = \left[ \begin{array}{c|c} R_{\mathbf{X}} & R_{\mathbf{X}\mathbf{Y}} \\ \hline R_{\mathbf{Y}\mathbf{X}} & R_{\mathbf{Y}} \end{array} \right] = \left[ \begin{array}{c|c} R_{\mathbf{X}} & \mathbf{0} \\ \hline \mathbf{0} & R_{\mathbf{Y}} \end{array} \right], \quad \mathbf{w} = \left[ \begin{array}{c} \mathbf{u} \\ \mathbf{v} \end{array} \right]$$

alors

$$\Phi_{\mathbf{Z}}(\mathbf{w}) = \exp\left(-\frac{1}{2} \mathbf{w}^t R_{\mathbf{Z}} \mathbf{w}\right) = \exp\left(-\frac{1}{2} \mathbf{u}^t R_{\mathbf{X}} \mathbf{u}\right) \exp\left(-\frac{1}{2} \mathbf{v}^t R_{\mathbf{Y}} \mathbf{v}\right) = \Phi_{\mathbf{X}}(\mathbf{u}) \Phi_{\mathbf{Y}}(\mathbf{v}).$$

D'après le paragraphe IV.5.1, la séparabilité de la fonction caractéristique en  $\mathbf{u}$  et  $\mathbf{v}$  équivaut à l'indépendance des vecteurs  $\mathbf{X}$  et  $\mathbf{Y}$ .

— Une propriété très importante des vecteurs gaussiens est liée au conditionnement de vecteurs gaussiens dans leur ensemble : soit  $\mathbf{X}$  et  $\mathbf{Y}$  deux vecteurs aléatoires gaussiens dans leur ensemble, alors l'espérance conditionnelle de  $\mathbf{X}$  conditionnellement à  $\mathbf{Y}$ ,  $E_{\mathbf{X}|\mathbf{Y}}(\mathbf{X} | \mathbf{Y})$  est une fonction affine de  $\mathbf{Y}$ . Plus précisément, avec les notations du paragraphe précédent, et  $R_{\mathbf{Y}} > 0$ , on obtient (exercice)

$$E_{\mathbf{X}|\mathbf{Y}}(\mathbf{X} | \mathbf{Y}) = \mathbf{m}_{\mathbf{X}} + R_{\mathbf{X}\mathbf{Y}} R_{\mathbf{Y}}^{-1} (\mathbf{Y} - \mathbf{m}_{\mathbf{Y}}). \quad (\text{IV.18})$$

De plus si nous notons  $\tilde{\mathbf{X}} = \mathbf{X} - E_{\mathbf{X}|\mathbf{Y}}(\mathbf{X} | \mathbf{Y})$ , alors :

$$R_{\tilde{\mathbf{X}}} = R_{\mathbf{X}} - R_{\mathbf{X}\mathbf{Y}} R_{\mathbf{Y}}^{-1} R_{\mathbf{Y}\mathbf{X}}. \quad (\text{IV.19})$$

## IV.6 L'espace de Hilbert $L^2(\Omega, \mathcal{F}, P)$

### IV.6.1 Norme et produit scalaire

Comme on l'a dit au § IV.3.3,  $\sqrt{E(X^2)}$  est une norme pour l'espace vectoriel complet  $L^2(\Omega, \mathcal{F}, P)$ . De plus cette norme est associée au produit scalaire  $\langle X, Y \rangle = \int_{\Omega} XY dP = E(XY)$ , ce qui confère à l'espace  $L^2(\Omega, \mathcal{F}, P)$  la structure d'espace de Hilbert. L'inégalité de Cauchy-Schwarz s'écrit :

$$(E(XY))^2 \leq E(X^2) E(Y^2). \quad (\text{IV.20})$$

De la même façon, les vecteurs aléatoires du second ordre de taille  $n$  forment un espace de Hilbert muni du produit scalaire  $\langle \mathbf{X}, \mathbf{Y} \rangle = E(\mathbf{X}^t \mathbf{Y})$  et de la norme  $\|\mathbf{X}\|^2 = E(\mathbf{X}^t \mathbf{X})$ .

### Démonstration de l'inégalité de Cauchy-Schwarz dans $L^2(\Omega, \mathcal{F}, P)$

Pour tout  $a, b$  réels et tout couple de v.a.r.  $X, Y$ , on a

$$E((aX + bY)^2) = a^2 E(X^2) + 2ab E(XY) + b^2 E(Y^2) \geq 0.$$

Prenons  $a = E(Y^2)$  et  $b = -E(XY)$ ; alors

$$E(Y^2)^2 E(X^2) - E(Y^2) E(XY)^2 \geq 0.$$

Si  $E(Y^2) > 0$ , on obtient directement (IV.20). Si  $E(Y^2) = 0$ ,  $Y = 0$  (p.s.) et  $E(XY) = 0$  : on a l'égalité  $0 = 0$ . Plus généralement, les cas d'égalité correspondent à des couples  $X, Y$  p.s. liés linéairement.

### Inégalité de Minkovski

On déduit facilement de (IV.20) l'inégalité de Minkovski

$$\|X + Y\| \leq \|X\| + \|Y\|$$

qui constitue l'inégalité triangulaire pour la norme de  $L^2(\Omega, \mathcal{F}, P)$ .

### IV.6.2 Sous-espaces de Hilbert

Dans un espace de Hilbert  $\mathcal{H}$ , tout sous-espace vectoriel fermé  $\mathcal{M}$  est un sous-espace hilbertien. Il est alors possible de définir pour tout élément  $X$  de  $\mathcal{H}$  sa *projection orthogonale* sur  $\mathcal{M}$ , qui minimise dans  $\mathcal{M}$  la distance à  $X$ .

Dans le cas de  $L^2(\Omega, \mathcal{F}, P)$ , deux types de sous-espaces présentent un intérêt particulier dans le contexte du traitement du signal.

### Sous-espaces de combinaisons affines

Considérons l'ensemble des combinaisons affines d'une v.a.r. d'ordre deux  $Y$  non constante. Il est évident que cet ensemble est un plan vectoriel de  $L^2(\Omega, \mathcal{F}, P)$ , de base orthonormée  $\{1, (Y - m_Y)/\sigma_Y\}$ . Il est aussi facile de montrer que cet espace est fermé, donc complet : il s'agit donc d'un sous-espace de Hilbert  $\mathcal{M}_Y$ . La projection  $\hat{X}(Y)$  d'une v.a.r.  $X$  sur ce sous-espace s'écrit sans difficulté (exercice) :

$$\hat{X}(Y) = m_X + \frac{\text{cov}(X, Y)}{\text{var}(Y)} (Y - m_Y).$$

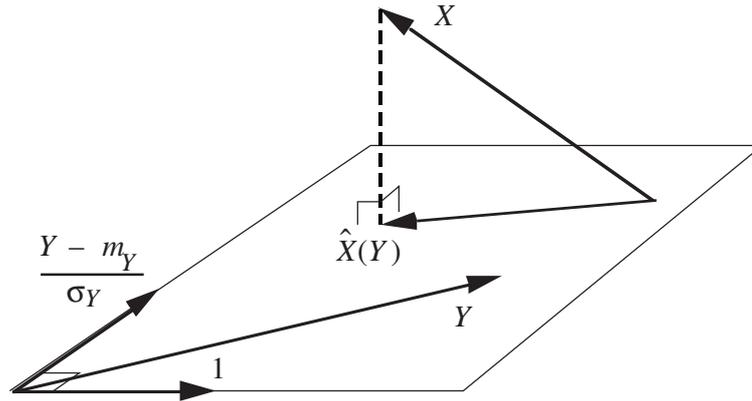


Fig. IV.2. Projection dans  $L^2(\Omega, \mathcal{F}, P)$

De façon plus générale, l'ensemble des v.a.r. obtenues par combinaison affine des composantes d'un vecteur d'ordre deux  $\mathbf{Y} = (Y_1, \dots, Y_n)^t$  est un sous-espace de Hilbert  $\mathcal{M}_{\mathbf{Y}}$  de dimension  $n + 1$  au plus. La projection d'une v.a.r.  $X$  d'ordre deux sur  $\mathcal{M}_{\mathbf{Y}}$  est la combinaison affine qui minimise la distance  $\|a_0 + a_1 Y_1 + \dots + a_n Y_n - X\|$ , où la norme est celle de  $L^2(\Omega, \mathcal{F}, P)$ . Dans le contexte de la théorie de l'estimation, cette projection est appelée *l'estimée linéaire d'erreur moyenne quadratique minimale* (ELMQ) de  $X$  étant donné  $\mathbf{Y}$ , dont nous verrons l'expression générale dans la deuxième partie du cours.

En général, la projection sur un sous-espace hilbertien séparable ou de taille finie est facilitée par la connaissance d'une famille génératrice ou d'une base du sous-espace. La recherche d'une base orthonormée de  $\mathcal{M}_{\mathbf{Y}}$  passe par la *factorisation* de la matrice de corrélation  $\mathbf{R}_{\mathbf{Y}}$ , c'est-à-dire par l'écriture de  $\mathbf{R}_{\mathbf{Y}}$  sous la forme

$$\mathbf{R}_{\mathbf{Y}} = \mathbf{A} \mathbf{A}^t, \quad (\text{IV.21})$$

où  $\mathbf{A}$  est une matrice rectangulaire telle que  $\mathbf{A}^t \mathbf{A}$  est inversible. En effet, soit  $\{1, X_1, \dots, X_m\}$  une base orthogonale de  $\mathcal{M}_{\mathbf{Y}}$  et  $\mathbf{X} = (X_1, \dots, X_m)^t$  : alors nécessairement  $\mathbf{m}_{\mathbf{X}} = \mathbf{0}$  et  $\mathbf{R}_{\mathbf{X}} = \mathbf{I}_m$  (matrice identité  $m \times m$ ). En remarquant que  $\mathbf{Y}$  est engendré par transformation affine de  $\mathbf{X}$ , on obtient (IV.21) comme une forme particulière de (IV.15).

Comme l'orthonormalisation de  $\mathbf{Y}$ , la factorisation de  $\mathbf{R}_{\mathbf{Y}}$  admet plusieurs solutions. Deux formes jouent un rôle prééminent dans les algorithmes d'estimation linéaire en traitement du signal :

— en tant que matrice symétrique réelle,  $\mathbf{R}_{\mathbf{Y}}$  admet une *décomposition spectrale*  $\mathbf{R}_{\mathbf{Y}} = \mathbf{M} \mathbf{\Lambda} \mathbf{M}^t$ , où  $\mathbf{\Lambda}$  est une matrice diagonale des valeurs propres ( $\geq 0$  car  $\mathbf{R}_{\mathbf{Y}} \geq 0$ ) et où  $\mathbf{M}$  est une matrice orthogonale ( $\mathbf{M}^t \mathbf{M} = \mathbf{I}_n$ ) de vecteurs propres assemblés en colonnes. De la décomposition spectrale, on déduit facilement une forme particulière de (IV.21) (exercice), qui réalise une

« double orthonormalisation » de  $\mathbf{Y}$ , au sens où  $\mathbf{X}$  est composé de v.a.r. orthonormées dans  $L^2(\Omega, \mathcal{F}, P)$ , tandis que les vecteurs colonnes de  $\mathbf{M}$  sont orthonormés dans  $\mathbb{R}^n$  !

— en effectuant une orthonormalisation pas à pas de  $\mathbf{Y}$ , on obtient  $\mathbf{A}$  sous la forme d'une matrice triangulaire inférieure (méthode de Gram-Schmidt). Le résultat est la *factorisation de Cholesky* de  $\mathbf{R}_{\mathbf{Y}}$ . Dans le cas où  $Y_1, \dots, Y_n$  modélisent les échantillons successifs d'un signal temporel, cette forme réalise l'orthonormalisation *causale* de la suite des échantillons du signal. La plupart des algorithmes récurrents d'estimation linéaire exploitent ce résultat (voir la seconde partie du cours).

### Sous-espace engendré par une ou plusieurs v.a.r.

Considérons non plus seulement les combinaisons affines d'une v.a.r. d'ordre deux  $X$ , mais toutes les v.a.r. d'ordre deux s'exprimant comme fonction déterministe de  $Y$

$$\mathcal{H}_Y = \{\nu(Y) \in L^2(\Omega, \mathcal{F}, P), \nu \text{ application mesurable de } (\mathbb{R}, \mathcal{B}) \rightarrow (\mathbb{R}, \mathcal{B})\}.$$

Il est évident que  $\mathcal{H}_Y$  est encore un sous-espace vectoriel de  $L^2(\Omega, \mathcal{F}, P)$ . En fait c'est un sous-espace de Hilbert de  $L^2(\Omega, \mathcal{F}, P)$  :  $Y$  induit un nouvel espace probabilisé,  $(\Omega, Y^{-1}(\mathcal{B}), P_Y)$  isomorphe à  $\mathcal{H}_Y$  ; or  $L^2(\Omega, Y^{-1}(\mathcal{B}), P_Y)$  est un espace de Hilbert au même titre que  $L^2(\Omega, \mathcal{F}, P)$ .

De la même façon, l'espace  $\mathcal{H}_{\mathbf{Y}}$  des v.a.r. d'ordre deux s'exprimant comme fonction déterministe d'un vecteur d'ordre deux est aussi un sous-espace de Hilbert de  $L^2(\Omega, \mathcal{F}, P)$ .

La projection d'une v.a.r.  $X$  d'ordre deux sur  $\mathcal{H}_{\mathbf{Y}}$  est la v.a.r. d'ordre deux  $\nu(\mathbf{Y})$  qui minimise la distance  $\|\nu(\mathbf{Y}) - X\|$ , où la norme est celle de  $L^2(\Omega, \mathcal{F}, P)$ . Dans le contexte de la théorie de l'estimation, cette projection est l'*estimée d'erreur moyenne quadratique minimale* (EMQ) de  $X$ , étant donné  $\mathbf{Y}$ . En remarquant  $\mathcal{M}_{\mathbf{Y}} \subset \mathcal{H}_{\mathbf{Y}}$ , on a  $\|\text{EMQ} - X\| \leq \|\text{ELMQ} - X\|$ . Le paragraphe suivant met en évidence un résultat essentiel : l'*EMQ de  $X$  étant donné  $\mathbf{Y}$*  n'est autre que l'espérance conditionnelle  $E(X | \mathbf{Y})$  ; ce résultat se généralise sans difficulté au cas de vecteurs aléatoires  $\mathbf{X}$  et  $\mathbf{Y}$ .

### IV.6.3 Espérance conditionnelle dans $L^2(\Omega, \mathcal{F}, P)$

Soit  $X$  une v.a.r. de  $L^2(\Omega, \mathcal{F}, P)$ . Alors on a le résultat suivant :

L'espérance conditionnelle  $E_{X|Y}(X | Y)$  réalise la meilleure approximation de  $X$  par une fonction de  $Y$ , au sens de la norme de  $L^2(\Omega, \mathcal{F}, P)$  :

$$\|X - E(X | Y)\| = \min_{\nu(Y) \in L^2(\Omega, \mathcal{F}, P)} \|X - \nu(Y)\|.$$

La structure d'espace de Hilbert de  $L^2(\Omega, \mathcal{F}, P)$  permet d'identifier alors  $E_{X|Y}(X | Y)$  à la *projection orthogonale* de la v.a.  $X$  sur le sous-espace hilbertien  $\mathcal{H}_Y$  de  $L^2(\Omega, \mathcal{F}, P)$ .

*Démonstration dans le cas d'un couple  $(X, Y)$  à densité :*

Commençons par montrer que  $E(X | Y)$  est un élément de  $\mathcal{H}_Y$ . A  $y$  fixé, l'inégalité de Cauchy-Schwarz pour les deux fonctions de carré intégrable  $x \rightarrow x(f_{X|Y}(x | y))^{1/2}$  et  $x \rightarrow (f_{X|Y}(x | y))^{1/2}$  fournit la majoration :

$$(E(X | Y = y))^2 \leq \int x^2 f_{X|Y}(x | y) dx.$$

En utilisant la monotonie de l'intégrale, on obtient alors :

$$\|E(X | Y)\|^2 \leq \int \left( \int x^2 f_{X|Y}(x | y) dx \right) f_Y(y) dy = \|X\|^2 < +\infty.$$

Notons désormais  $h(Y)$  la projection orthogonale de  $X$  sur  $\mathcal{H}_Y$  et soit  $\nu(Y)$  un élément quelconque de  $\mathcal{H}_Y$ . Par définition, on a

$$\langle h(Y) - X, \nu(Y) \rangle = \iint (h(y) - x) \nu(y) f_{XY}(x, y) dx dy = 0.$$

En développant  $f_{XY}(x, y) = f_{X|Y}(x|y) f_Y(y)$ , il vient

$$\int \left( h(y) - \int x f_{X|Y}(x|y) dx \right) \nu(y) f_Y(y) dy = \langle h(Y) - E_{X|Y}(X|Y), \nu(Y) \rangle = 0.$$

Cette égalité est valable pour tout  $\nu(Y)$  de  $\mathcal{H}_Y$ . En particulier pour  $\nu(Y) = h(Y) - E(X|Y)$  on obtient  $\|h(Y) - E_{X|Y}(X|Y)\|^2 = 0$ , ce qui conclut la démonstration.

Dans le cas d'un couple  $(X, Y)$  du second ordre mais sans densité de probabilité, nous n'avons pas encore défini l'espérance conditionnelle  $E(X|Y)$ . En revanche, la projection orthogonale de  $X$  sur  $\mathcal{H}_Y$  existe : dans ce cas on peut définir  $E(X|Y)$  comme étant cette projection. En fait l'espérance conditionnelle existe même sous des conditions plus générales.

# Chapitre V

## Signaux aléatoires à temps discret

### V.1 Introduction

Par souci de simplicité, on se limite dans la suite aux signaux aléatoires discrets, le plus souvent à valeurs réelles ou à états finis ou dénombrables. De fait, l'essentiel des algorithmes de traitement statistique du signal manipulent des modèles discrets, même si certains algorithmes admettent une extension à temps continu. C'est le cas du filtrage de Kalman, souvent présenté à temps continu dans les cours d'automatique parce que les systèmes à commander sont des systèmes physiques, tandis que le traitement du signal analyse le plus souvent des données échantillonnées dès leur acquisition (échantillonneurs-bloqueurs, caméras CCD), voire intrinsèquement discrètes dans un domaine tel que les communications numériques.

### V.2 Caractérisation

#### V.2.1 Définition

Un signal aléatoire à temps discret  $X$  est une suite  $\{X_k\}$  de variables aléatoires, indexée par  $\mathbb{N}$  ou  $\mathbb{Z}$ . De la même façon, un signal aléatoire à temps continu est une famille de variables aléatoires indexée par  $\mathbb{R}$ .

— A  $k$  fixé,  $X_k$  est une variable aléatoire et un  $n$ -uplet  $(X_{k_1}, \dots, X_{k_n})$  forme un vecteur aléatoire, auxquels pourront s'appliquer les outils et les propriétés mis en évidence dans le chapitre précédent. Un tel vecteur sera dénommé *vecteur extrait* de  $X$ .

— Pour une épreuve  $\omega$  « tirée au hasard » dans un espace probabilisé  $(\Omega, \mathcal{F}, P)$ ,  $X_k(\omega)$  est un signal déterministe à temps discret, dont on pourra étudier la stabilité, la représentation harmonique... Chaque réalisation particulière d'un signal aléatoire est communément appelée une *trajectoire*. Par exemple, on a représenté ci-dessous en figure V.1 trois fragments de trajectoires « également possibles » de la suite aléatoire binaire résultant d'un tirage à pile ou face équiprobable.

Comme dans le cas des variables et des vecteurs aléatoires, on renonce très souvent à définir explicitement un signal aléatoire pour ne s'intéresser qu'à sa loi.

#### V.2.2 Loi temporelle

##### Définition

La loi temporelle d'un signal aléatoire  $X$  correspond à la spécification des lois conjointes de  $n$  variables aléatoires  $X_{k_1}, X_{k_2}, \dots, X_{k_n}$  pour tout  $n$  entier et pour tout  $n$ -uplet d'instant



la fonction de répartition instantanée  $F_k(x) = P(X_k \leq x)$  et les moments de  $X_k$ , qui sont des fonctions déterministes du temps (sous réserve d'existence). On définit essentiellement la moyenne

$$E(X_k) = \int_{\mathbb{R}} x dF_k(x) = m(k) \quad (\text{V.3})$$

et le moment d'ordre deux

$$E(X_k^2) = \int_{\mathbb{R}} x^2 dF_k(x) = m^2(k) + \text{var}(X_k), \quad (\text{V.4})$$

qu'on appelle *puissance moyenne* à l'instant  $k$ .

## V.2.4 Caractéristiques à l'ordre deux

### Définition

Dans la perspective de travailler dans l'espace de Hilbert  $L^2(\Omega, \mathcal{F}, P)$ , il est intéressant de définir les caractéristiques d'un signal  $X$  à l'ordre deux, pourvu qu'il soit d'ordre deux, c'est-à-dire de puissance moyenne partout finie :  $E(X_k^2) < +\infty$  pour tout  $k$ . Il s'agit d'abord de la fonction moyenne  $m_X(k)$ , déjà définie par (V.3) à partir des caractéristiques instantanées. On définit aussi la *fonction d'autocorrélation*

$$\begin{aligned} r_X(i, j) &= \text{cov}(X_i, X_j) = E((X_i - m_X(i))(X_j - m_X(j))) \\ &= E(X_i X_j) - m_X(i) m_X(j). \end{aligned} \quad (\text{V.5})$$

Cette fonction est symétrique ( $r_X(j, i) = r_X(i, j)$ ) et *définie non négative* : pour tout  $n > 1$ ,  $k_1, \dots, k_n$  entiers et  $\lambda_1, \dots, \lambda_n$  réels,

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j r_X(k_i, k_j) = E\left(\left(\sum_{i=1}^n \lambda_i (X_{k_i} - m_X(k_i))\right)^2\right) \geq 0.$$

D'autre part, d'après l'inégalité de Schwarz, la fonction d'autocorrélation atteint son maximum en 0. Ce résultat est analogue à celui présenté au §I.3.1.

Un signal  $X$  est *décorrélé*, ou *blanc à l'ordre deux*, s'il est constitué de variables décorrélées :  $r_X(i, j) = r_X(i, i) \delta_{i-j}$ .

Enfin, pour deux signaux  $X$  et  $Y$  d'ordre deux, on définit à partir de la loi temporelle jointe la *fonction d'intercorrélacion*

$$\begin{aligned} r_{XY}(i, j) &= \text{cov}(X_i, Y_j) = E((X_i - m_X(i))(Y_j - m_Y(j))) \\ &= E(X_i Y_j) - m_X(i) m_Y(j). \end{aligned} \quad (\text{V.6})$$

$X$  et  $Y$  sont *décorrélés* s'ils sont constitués de variables décorrélées :  $r_{XY}(i, j) = 0$ .

### Sous-espaces de $L^2(\Omega, \mathcal{F}, P)$

Nous avons déjà signalé au chapitre précédent l'importance de la projection des v.a.r. de  $L^2(\Omega, \mathcal{F}, P)$  sur certains de ses sous-espaces hilbertiens, en particulier ceux qui sont engendrés par combinaisons affines à partir d'une v.a.r. ou d'un vecteur aléatoire. De la même façon, l'*histoire linéaire* d'un signal aléatoire  $X$  est l'ensemble engendré par combinaisons linéaires des  $X_k$ . A l'instant  $k$ , le *passé linéaire* est l'ensemble engendré par  $X_{k-1}, X_{k-2} \dots$ . Ces espaces sont séparables. On les construit comme la fermeture dans  $L^2(\Omega, \mathcal{F}, P)$  des ensembles des combinaisons linéaires finies.

## V.3 Construction et exemples

La spécification d'une loi temporelle, plus simple que la caractérisation complète de l'application mesurable  $X$ , se révèle néanmoins peu pratique pour *construire* des lois de signaux aléatoires, donc pour fabriquer des modèles adaptés à un contexte particulier. Le but de ce paragraphe n'est pas de fournir un catalogue exhaustif de modèles, ce qui serait à la fois illusoire et sans intérêt, mais plutôt d'introduire des modèles fondamentaux et de montrer comment on peut « enrichir » le catalogue en définissant des modèles plus compliqués à partir de modèles simples.

### V.3.1 Bruits blancs

#### Définition

On entend par *bruit blanc*, en référence à sa décomposition harmonique uniforme dans la bande des fréquences analysée (voir chapitre VI), un signal aléatoire constitué d'une suite de variables aléatoires *indépendantes*. La loi temporelle se sépare donc sous la forme caractéristique simple

$$F_{k_1, k_2, \dots, k_n}(x_1, x_2, \dots, x_n) = F_{k_1}(x_1) \dots F_{k_n}(x_n), \quad (\text{V.7})$$

pour tout entier  $n$  et pour tout  $n$ -uplet d'instants  $(k_1, \dots, k_n)$  et de valeurs  $(x_1, \dots, x_n)$ . La loi du signal est donc décrite entièrement par la suite des lois instantanées. Si ces lois sont identiques, le bruit blanc est une suite de v.a.r. *indépendantes identiquement distribuées* (i.i.d.); il est dit *homogène*.

Dans  $L^2(\Omega, \mathcal{F}, P)$ , il est fréquent de ne considérer que les caractéristiques à l'ordre deux d'un signal aléatoire. C'est ainsi qu'un *bruit blanc au sens faible* est une suite de variables aléatoires décorréées, mais pas forcément indépendantes.

#### Exemples

La suite des « pile ou face » proposée en introduction est un bruit blanc homogène binaire. Un autre exemple très utile est celui du bruit blanc homogène gaussien, dont la loi instantanée est décrite par la moyenne et la puissance (ou la variance) des v.a.r. gaussiennes i.i.d.  $X_k$ . La figure V.2 représente 50 points d'une trajectoire d'un bruit blanc gaussien réduit (moyenne nulle, variance unité).

En pratique, on utilise souvent des modèles homogènes blancs pour décrire la partie indésirable d'un signal observé, dans le cas où cette partie est trop irrégulière pour une modélisation déterministe. D'où l'appellation « négative » de « bruit » blanc, qu'on peut se représenter comme le « souffle » d'une cassette audio.

Ainsi présenté, un bruit blanc semble le contraire d'un signal « utile ». Mais l'utilité d'un signal n'est pas une caractéristique intrinsèque. C'est au contraire une notion subjective dépendant de ce que l'observateur cherche dans le signal. D'une part, tous les signaux « nuisibles » ne sont pas des bruits blancs : les courants sinusoïdaux à 50 Hz délivrés par Électricité de France sont globalement utiles, mais leurs résidus peuvent perturber des systèmes d'alimentation autonome ou de transmission d'information.

Réciproquement, tous les bruits blancs ne sont pas « nuisibles ». Par exemple, en théorie de l'information, les bruits blancs permettent de maximiser le débit d'information d'un canal de transmission, car leur caractère non prédictible assure qu'aucune redondance ne subsiste dans une suite de variables i.i.d. Le but de la compression de données est justement de transformer des signaux quelconques en signaux proches d'un bruit blanc homogène par une transformation inversible, pour ne stocker ou ne transmettre que le signal « blanchi », en augmentant ainsi la

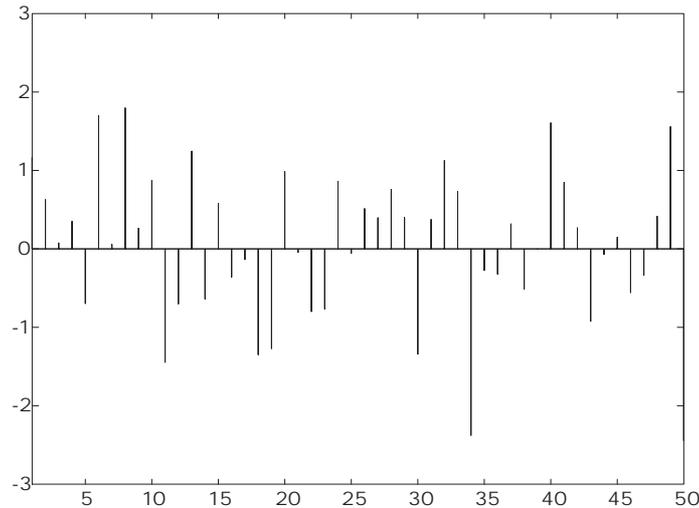


Fig. V.2. Exemple 2 : bruit blanc gaussien

capacité de stockage ou le débit de transmission. Nous reviendrons plus loin sur la recherche du « blanchiment » d'un signal, qui rejoint dans le contexte des signaux de puissance moyenne finie la notion d'orthogonalisation de la suite des échantillons du signal, évoquée à la fin du chapitre précédent (§IV.6.2).

### V.3.2 Chaines de Markov

#### Définition

Une chaîne de Markov est un signal aléatoire construit sur  $\mathbb{N}$  et à valeurs dans  $E$ , espace fini ou dénombrable. Traitons le cas d'un espace fini  $E = \{e_1, \dots, e_M\}$ .

Si un signal aléatoire  $X$  défini vers  $E^{\mathbb{N}}$  vérifie pour tout  $n \in \mathbb{N}^*$ ,  $i_0, i_1, \dots, i_{n-1}, i, j \in E$  :

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i) = p_{ij}, \quad (\text{V.8})$$

on dit que  $X$  est une chaîne de Markov (homogène) de matrice de transition  $P = \{p_{ij}\}_{i,j \in E}$ . La loi initiale de cette chaîne définie par le vecteur  $\mathbf{q} = (q_1, \dots, q_M)^t$  des probabilités initiales  $q_i = P(X_0 = i)$ .

Le couple  $(\mathbf{q}, P)$  caractérise la chaîne, au sens où  $(\mathbf{q}, P)$  détermine entièrement sa loi temporelle. En effet, la loi de Bayes (IV.2) donne sous sa forme récurrente

$$P(X_0 = i_0, \dots, X_k = i_k) = q_{i_0} p_{i_0 i_1} \cdots p_{i_{k-1} i_k}. \quad (\text{V.9})$$

En particulier, (V.9) permet de déduire la loi instantanée de la chaîne  $\{P(X_k = i)\}_{i \in E}$  à l'instant  $k$  : par sommations successives sur  $i_0, \dots, i_{k-1}$ , on obtient  $P(X_k = e_m)$  comme la  $m$ -ième composante du vecteur  $\mathbf{q}_k = (P^k)^t \mathbf{q}$ .

#### Construction

Réciproquement, soit un couple  $(\mathbf{q}, P)$ , où

- $\mathbf{q} = (q_1, \dots, q_M)^t$  est un vecteur de probabilités exhaustives, c'est-à-dire vérifiant  $0 \leq q_m \leq 1$  pour tout  $m$  et  $\sum_m q_m = 1$  ;
- $P$  est une matrice *stochastique* de taille  $M \times M$ , c'est-à-dire une matrice dont chaque ligne est un vecteur de probabilités exhaustives.

Alors en utilisant un théorème de construction de mesure (le théorème d'extension de Kolmogorov [5]), on montre que  $(\mathbf{q}, P)$  caractérise une loi de probabilité de signal aléatoire sur  $\mathbb{N}$  d'après (V.9), donc une chaîne de Markov. Par conséquent il est beaucoup plus simple de décrire la loi d'une chaîne de Markov homogène qu'une loi temporelle quelconque.

### Exemple

Les descriptions markoviennes sont commodes pour une multitude de signaux qui possèdent des propriétés « locales » de dépendance statistique. Par exemple, la marche aléatoire à  $d$  dimensions modélise la trajectoire d'un point dans  $E = \mathbb{Z}^d$  qui saute à chaque instant d'un pas de façon équiprobable dans une des directions parallèles aux axes. La figure V.3 représente 50 points d'une trajectoire de marche aléatoire  $X$  à une dimension ( $E = \mathbb{Z}$ ) dont la position initiale est  $X_0 = 0$ . Les caractéristiques de la chaîne sont donc

$$q_m = \delta_{0m}, m \in \mathbb{Z}$$

$$p_{ij} = \begin{cases} 1/2 & \text{si } i = j \pm 1, i, j \in \mathbb{Z}, \\ 0 & \text{sinon.} \end{cases}$$

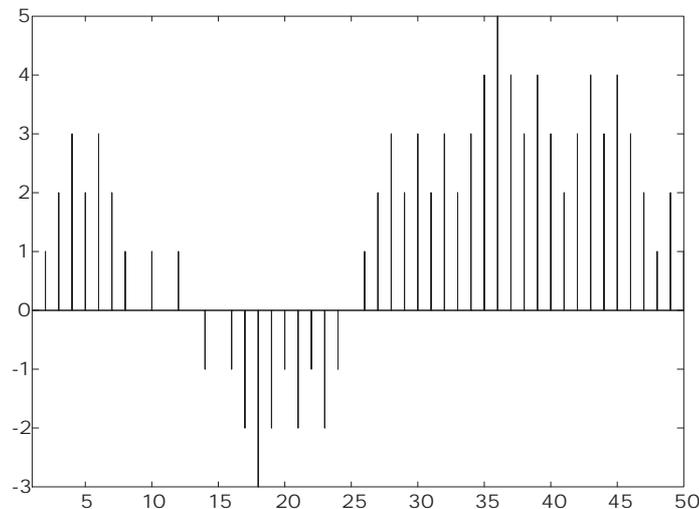


Fig. V.3. Exemple 3 : marche aléatoire

### V.3.3 Constructions indirectes

On peut définir de façon simple et très efficace des nouveaux signaux aléatoires comme fonctions déterministes d'autres modèles aléatoires, variables ou signaux.

### Signaux aléatoires fonctions de variables aléatoires

Sans rentrer dans les généralités, traitons le cas d'une sinusoïde de fréquence connue, mais d'amplitude et de phase aléatoires. Par exemple, ce modèle est classique pour décrire la composante parasite à 50 Hz évoquée au § V.3.1.

On considère un couple  $(A, \Phi)$  de variables aléatoires indépendantes à densité :  $f_{A,\phi}(a, \Phi) = f_A(a)f_\phi(\Phi)$ , où  $A$  suit une loi de Rayleigh sur  $\mathbb{R}_+$  de paramètre  $\sigma^2$

$$f_A(a) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right)$$

et  $\Phi$  est uniformément répartie sur  $]0, 2\pi[$ . On définit le signal aléatoire  $X$  comme l'application mesurable (de  $\Omega$  dans  $\mathbb{R}^{\mathbb{Z}}$ ) dont la composante à l'instant  $k$  est la v.a.r.  $X_k$  :

$$X_k = A \cos(2\pi\nu k + \Phi). \quad (\text{V.10})$$

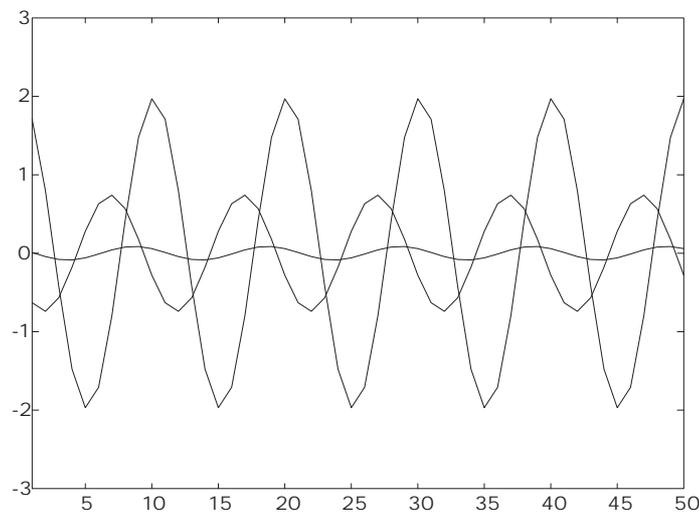


Fig. V.4. Exemple 4 : 3 trajectoires d'une sinusoïde à phase et amplitude aléatoires

Comme le montre la figure V.4, chaque trajectoire de ce modèle est une sinusoïde ; remarquons au passage qu'il est faux de penser que les signaux aléatoires ont toujours des trajectoires « irrégulières », « imprédictibles ». Au contraire, il est facile de vérifier que  $X$  satisfait l'équation de récurrence

$$X_k = 2 \cos(2\pi\nu) X_{k-1} - X_{k-2} \quad (\text{V.11})$$

qui montre que le signal aléatoire  $X$  est parfaitement prédictible par combinaison linéaire finie des deux derniers échantillons passés. Il s'agit d'un cas extrême de prédictibilité (le signal aléatoire est dit *singulier*).

Déterminons les caractéristiques instantanées de  $X$ . Posons  $Y_k = A \sin(2\pi\nu k + \Phi)$ . L'application  $g(a, \varphi) = (a \cos(2\pi\nu k + \varphi), a \sin(2\pi\nu k + \varphi))$  est bijective de  $\mathbb{R}_+ \times ]0, 2\pi[$  dans  $\mathbb{R}^2$  pour tout  $k$ , donc la formule de changement de variable (IV.9) permet de trouver la densité jointe du couple  $(X_k, Y_k)$ . Le calcul donne

$$f_{X_k, Y_k}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) = f_{X_k}(x) f_{Y_k}(y),$$

ce qui montre que  $X_k$  et  $Y_k$  sont des variables indépendantes de même loi gaussienne  $\mathcal{N}(0, \sigma^2)$  indépendante de  $k$  et de  $\nu$  !

## Fonctions de signaux aléatoires

Il s'agit des transformations déterministes d'un signal aléatoire  $X : Y = \varphi(X)$ , tel que  $Y$  est mesurable. Deux cas présentent un intérêt particulier :

— Les transformations *instantanées* d'un signal aléatoire  $X$ , c'est-à-dire les transformations instant par instant :  $Y$  est la suite des variables aléatoires  $Y = \varphi_k(X_k)$ .

— Les transformations affines d'un signal aléatoire. Ce cas est essentiel puisqu'il contient le filtrage linéaire. En particulier, on peut définir de nombreux signaux par filtrage linéaire des bruits blancs.

Par exemple, la marche aléatoire monodimensionnelle  $X = \{X_k\}_{k \in \mathbb{N}}$  de l'exemple 3 est la sortie d'un filtre « intégrateur » dont l'entrée est un bruit blanc binaire  $B = \{B_k\}_{k \in \mathbb{N}}$  à valeurs  $\{+1, -1\}$  du type de l'exemple 1 :

$$\begin{cases} X_0 = 0 \\ X_k = X_{k-1} + B_k, \quad k \geq 1. \end{cases} \quad (\text{V.12})$$

La TZ du filtre intégrateur est  $H(z) = z/(z-1)$  : il s'agit d'un filtre récursif (qui n'est pas stable ; voir le chapitre III).

## ARMA causaux

Plus généralement, toute fraction rationnelle en  $z$   $P(z)/Q(z)$  admettant un développement causal (c'est-à-dire telle que  $d^\circ(P) \leq d^\circ(Q) = N$  ; voir le chapitre III) permet de définir des signaux aléatoires sur  $\mathbb{N}$  par filtrage causal des bruits blancs homogènes, de la façon suivante :

- soit  $F_B$  la loi instantanée du bruit d'entrée  $B = \{B_k\}$  ;
- étant donnée la loi du vecteur  $(X_0, \dots, X_{N-1})^t$  définissant les conditions initiales, indépendant du bruit  $B$ ,
- pour tout  $n \geq N$ ,  $X_n$  est la sortie du filtre causal de fonction de transfert  $P(z)/Q(z)$ , qu'on peut obtenir sous forme récursive :

$$X_k = \frac{1}{a_0} \left( \sum_{m=0}^M b_m B_{k-m} - \sum_{n=1}^N a_n X_{k-n} \right) \quad (\text{V.13})$$

avec les notations du chapitre III.

Ces signaux aléatoires définis sur  $\mathbb{N}$  sont dits *ARMA causaux*.

— Pour  $M = 0$ , le filtre générateur est purement récursif et le signal est dit AR (pour *autorégressif*).

— Pour  $N = 0$ , le filtre est à réponse impulsionnelle finie et le signal est un MA (pour « moving average », *moyenne mobile*).

## ARMA stationnaires

Par filtrage linéaire causal des bruits blancs homogènes, on a seulement défini des signaux aléatoires sur  $\mathbb{N}$ . En fait on peut aussi obtenir des signaux aléatoires sur  $\mathbb{Z}$  par filtrage linéaire. Mais cette fois l'important est de garantir la stabilité, et non plus la causalité (et *a fortiori*, le filtrage stable donnera aussi des signaux aléatoires sur  $\mathbb{N}$  par restriction). Il faut aussi que le bruit soit homogène d'ordre deux. Alors on a le résultat suivant :

Le filtrage d'un bruit blanc homogène d'ordre deux  $B = \{B_k\}_{k \in \mathbb{Z}}$  par un filtre  $h = \{h_k\}_{k \in \mathbb{Z}}$  linéaire stable définit un signal  $X$  aléatoire d'ordre deux sur  $\mathbb{Z}$  :

$$X_n = \sum_{k \in \mathbb{Z}} h_{n-k} B_k. \quad (\text{V.14})$$

Ce résultat est une conséquence d'une propriété des systèmes orthonormés dans un espace de Hilbert :  $(B - m_B)/\sigma_B$  est un système orthonormé de  $L^2(\Omega, \mathcal{F}, P)$ , donc toute série  $\sum_{k \in \mathbb{Z}} c_k B_k$  converge dans  $L^2(\Omega, \mathcal{F}, P)$  (et même dans l'histoire linéaire de  $B$ ) pourvu que  $\sum_{k \in \mathbb{Z}} c_k^2 < +\infty$ , et définit une v.a.r. de puissance  $E(B_k^2) \sum_{k \in \mathbb{Z}} c_k^2$ . Or  $h$  est stable, donc d'énergie finie, donc le résultat s'applique à (V.14).

On déduit facilement de (V.14) l'invariance des caractéristiques statistiques d'ordre deux de  $X$  au cours du temps :

$$E(X_k) = m_B \sum_{k \in \mathbb{Z}} h_k, \quad (\text{V.15})$$

$$E(X_k^2) = \sigma_B^2 \sum_{k \in \mathbb{Z}} h_k^2 + m_B^2 \left( \sum_{k \in \mathbb{Z}} h_k \right)^2 < +\infty, \quad (\text{V.16})$$

$$E(X_n X_{n+m}) = \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} h_{n-k} h_{n+m-l} E(B_k B_l) \quad (\text{V.17})$$

$$= \sigma_B^2 \sum_{k \in \mathbb{Z}} h_k h_{k+m} + m_B^2 \left( \sum_{k \in \mathbb{Z}} h_k \right)^2. \quad (\text{V.18})$$

Donc la moyenne  $E(X_k)$  et l'autocorrélation  $E(X_k X_{k+m}) - m_X(k) m_X(k+m)$  ne dépendent pas de l'instant  $k$ . Le signal  $X$  est dit *stationnaire à l'ordre deux*, notion sur laquelle nous reviendrons au §V.4.2, puis au chapitre VI. Dans le cas où la TZ du filtre linéaire stable est une fraction rationnelle, on dit que  $X$  est un *ARMA stationnaire à l'ordre deux*.

Un filtre générateur à la fois stable et causal engendre par (V.14) un ARMA stationnaire à l'ordre deux et causal sur  $\mathbb{Z}$ . La causalité du filtre a pour conséquence évidente la décorrélation entre les sorties passées et les entrées futures du filtre :

$$E(X_n B_m) = 0 \text{ pour tout } m > n.$$

## V.4 Propriétés

Après avoir défini une gamme assez vaste de signaux aléatoires, nous allons pouvoir étudier certaines de leurs propriétés. Une des plus importantes, la *décomposition harmonique*, valable pour les modèles stationnaires à l'ordre deux seulement, sera traitée spécifiquement dans le chapitre VI.

Dans la suite, nous étudions différentes notions de convergence qui peuvent s'appliquer aux signaux aléatoires (§V.4.1), puis nous définissons la stationnarité, au sens strict et au sens large (§V.4.2) et les propriétés d'ergodisme de certains signaux stationnaires (§V.4.3).

### V.4.1 Convergence

La convergence d'un signal aléatoire réel  $\{X_n\}_{n \in \mathbb{N}}$  s'identifie à la convergence de la suite des v.a.r.  $X_0, X_1, \dots, X_k, \dots$ , c'est-à-dire que seules les caractéristiques instantanées du signal sont concernées. Il est important de voir que cette notion de convergence de v.a.r. peut prendre des sens très différents suivant que l'on s'intéresse seulement à la loi des variables aléatoires, ou bien aux valeurs qu'elles prennent épreuve par épreuve, ou encore à une mesure globale sur toutes les épreuves. On distingue quatre types de convergence.

#### Convergence en loi

C'est la notion la plus simple, mais aussi la convergence la plus faible. En fait elle ne désigne pas une convergence de variables aléatoires mais seulement la convergence de lois de probabilité.

On dit que la suite  $\{X_k\}$  converge *en loi* vers une v.a.r.  $X$  si la probabilité  $P_{X_k}$  converge vers  $P_X$  au sens des mesures sur  $\mathbb{R}$ . Une condition nécessaire et suffisante s'obtient sous la forme d'une convergence simple de fonction de répartition point par point :  $\{X_k\}$  converge *en loi* vers  $X$  ssi en tout point de continuité de la fonction de répartition de  $X$ , on a

$$\lim_{k \rightarrow \infty} F_{X_k}(x) = F_X(x). \quad (\text{V.19})$$

Une condition équivalente porte de la même façon sur la fonction caractéristique.

— Exemple : théorème central limite

Soit  $X = \{X_k\}_{k \in \mathbb{N}}$  un bruit blanc homogène au sens strict et d'ordre deux sur  $\mathbb{N}$ . Considérons le signal aléatoire  $S = \{S_k\}_{k \in \mathbb{N}}$ ,

$$S_k = \frac{1}{k} \sum_{i=0}^{k-1} X_i.$$

Il est évident que  $S_k$  est une v.a.r. d'ordre deux, de moyenne  $m_S = m_X$  et de variance  $\sigma_S^2 = \sigma_X^2/k$ . Le théorème central limite concerne la convergence en loi de la variable  $S_k$  une fois réduite : la loi de  $(S_k - m_S)/\sigma_S$  converge vers une gaussienne réduite.

### Convergence en probabilité (ou convergence stochastique)

Elle est obtenue ssi

$$\forall \varepsilon > 0, \lim_{k \rightarrow \infty} P(|X_k - X| \geq \varepsilon) = 0. \quad (\text{V.20})$$

Elle peut s'interpréter en disant que pour  $k$  grand, l'ensemble des  $\omega$  tels que  $X_k(\omega)$  s'écarte de  $X(\omega)$  de plus de  $\varepsilon$  est de mesure arbitrairement faible, pour tout  $\varepsilon$ . Il s'agit donc d'une convergence « globale », en probabilité sur l'ensemble des trajectoires.

### Convergence presque sûre

Elle est obtenue ssi l'ensemble des  $\omega$  ne vérifiant pas

$$\forall \varepsilon > 0, \exists N : \forall k > N, |X_k(\omega) - X(\omega)| < \varepsilon \quad (\text{V.21})$$

est de mesure nulle. Cette convergence est très forte puisqu'elle concerne séparément chaque épreuve  $\omega$  (donc chaque trajectoire du signal) au sens déterministe, sauf un ensemble négligeable de trajectoires. En fait, la limite d'une suite de v.a.r.  $X_k$  convergeant presque sûrement définit la variable aléatoire  $X$ . En pratique, les cas les plus fréquents de convergence presque sûre font intervenir une limite presque sûre : presque toutes les trajectoires sont convergentes au sens déterministe, vers la même limite.

— Exemple : loi des grands nombres

Soit  $X = \{X_k\}_{k \in \mathbb{N}}$  un bruit blanc homogène d'ordre deux sur  $\mathbb{N}$ . Alors il existe une version de la loi des grands nombres adaptée à cette hypothèse. Il s'agit d'un résultat de convergence presque sûre vers une quantité déterministe :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} X_k = m_X. \quad (\text{V.22})$$

### Convergence en moyenne quadratique

Il s'agit de la convergence dans l'espace de Hilbert des v.a.r. du second ordre  $L^2(\Omega, \mathcal{F}, P)$ , déjà évoquée pour fermer des sous-espaces linéaires (§V.2.4) ou pour justifier l'appartenance à

$L^2(\Omega, \mathcal{F}, P)$  des filtrées stables d'un bruit blanc homogène d'ordre deux (fin de § V.3.3) :

$$\lim_{k \rightarrow \infty} \|X_k - X\| = 0,$$

c'est-à-dire

$$\lim_{k \rightarrow \infty} E((X_k - X)^2) = 0. \quad (\text{V.23})$$

C'est une convergence globale (c'est-à-dire définie sur l'ensemble des épreuves) très forte, qui joue un rôle essentiel dans l'étude des processus stochastiques.

Il n'existe aucun lien systématique entre la convergence presque sûre et la convergence en moyenne quadratique. Cependant toutes deux impliquent la convergence en probabilité qui entraîne la convergence en loi.

## V.4.2 Stationnarité

Chaque fois qu'un signal a pour origine un phénomène permanent à l'échelle du temps de mesure, sans changement de comportement prévisible, il est intéressant de le décrire par un modèle aléatoire dont les caractéristiques statistiques possèdent des propriétés d'invariance adéquates. Cette situation est particulièrement fréquente dans le cas où l'observateur ne déclenche pas le phénomène qu'il mesure, et que ce dernier existe de façon permanente indépendamment de l'observation.

### Stationnarité forte. Exemple des chaînes de Markov

Un processus aléatoire  $X$  est dit *fortement stationnaire* (ou *stationnaire strict*, ou *stationnaire*) si sa loi temporelle est invariante par toute translation du temps, c'est-à-dire que les vecteurs aléatoires  $(X_{k_1}, X_{k_2}, \dots, X_{k_n})^t$  et  $(X_{k_1+k}, X_{k_2+k}, \dots, X_{k_n+k})^t$  ont des lois de probabilités identiques pour tout  $n \in \mathbb{N}^*$ , tout  $n$ -uplet  $(k_1, k_2, \dots, k_n) \in \mathbb{Z}^n$  et pour tout  $k \in \mathbb{Z}$ . Par restriction, on peut aussi définir la stationnarité des signaux aléatoires sur  $\mathbb{N}$ .

Reprenons l'exemple des chaînes de Markov : les couples  $(\mathbf{q}, P)$  qui vérifient  $P^t \mathbf{q} = \mathbf{q}$  définissent des modèles stationnaires stricts sur  $\mathbb{N}$ . En effet, cette propriété a pour conséquence immédiate que le vecteur  $\mathbf{q}_k = (P^k)^t \mathbf{q}$  des caractéristiques instantanées de la chaîne à l'instant  $k$  est égal à  $\mathbf{q}$ . A partir de ce résultat, l'application de la loi de Bayes ayant donné (V.9) donne le même résultat quel que soit l'instant :

$$P(X_n = i_0, \dots, X_{n+k} = i_k) = q_{i_0} p_{i_0 i_1} \dots p_{i_{k-1} i_k} = P(X_0 = i_0, \dots, X_k = i_k),$$

ce qui suffit pour prouver la stationnarité stricte.

Toutes les matrices de transition  $P$  n'admettent pas de vecteur  $\mathbf{q}$  initial tel que  $P^t \mathbf{q} = \mathbf{q}$ , tandis que d'autres en admettent plusieurs. La théorie des chaînes de Markov étudie en particulier ces questions d'existence et d'unicité. Dans le cas de la marche aléatoire, les vecteurs  $\mathbf{q}$  invariants sont tous colinéaires au vecteur constant  $\mathbf{1}$ . Mais aucun d'entre eux n'est un vecteur de probabilité : en fait la chaîne finit par « se diluer » uniformément sur tous les entiers relatifs quel que soit son point de départ. La chaîne ne converge pas en loi vers une variable aléatoire.

### Stationnarité faible

La stationnarité forte porte sur l'ensemble de la loi temporelle du processus. On est amené à définir un autre type de stationnarité dite *faible*, ou *au sens large*, ou *à l'ordre deux* car elle fait seulement intervenir l'invariance dans le temps des moments d'ordre un et deux du processus aléatoire.

Un processus aléatoire (d'ordre deux)  $X$  est dit *faiblement stationnaire* lorsque sa valeur moyenne est constante et sa fonction d'autocorrélation  $r_X(k_1, k_2)$  n'est fonction que de la différence  $n = k_1 - k_2$ . Ceci s'écrit

$$\forall k \in \mathbb{Z}, E(X_k) = m_X, \forall (k, n) \in \mathbb{Z}^2, r_X(k, k+n) = C_X(n).$$

En particulier, la puissance moyenne de  $X$  est une constante :

$$\forall k \in \mathbb{Z}, E(X_k^2) = m_X^2 + C_X(0).$$

Soit  $(X_k, \dots, X_{k+n})$  un vecteur extrait de  $X$ . La matrice d'autocorrélation de ce vecteur a pour composante  $a_{ij} = r_X(k+i-1; k+j-1)$ . La stationnarité faible confère à cette matrice une structure particulière dite *de Toeplitz* où les coefficients sont constants suivant chaque diagonale.

Il est évident que la stationnarité forte implique la stationnarité faible pourvu que les deux premiers moments existent. La réciproque est en général fausse, mais elle est vraie dans le cas des processus gaussiens.

Il est évident qu'un bruit blanc homogène au sens strict (respectivement au sens faible) est stationnaire au sens strict (resp. au sens faible). Par filtrage stable des bruits blancs stationnaires d'ordre deux, nous avons montré à la fin du § V.3.3 qu'on pouvait définir de nouveaux signaux stationnaires d'ordre deux. On peut démontrer [1] un résultat plus général :

Le filtrage par un filtre stable d'un signal stationnaire à l'ordre deux est un signal stationnaire à l'ordre deux.

### Retour sur les ARMA stationnaires faibles

La stationnarité large des signaux obtenus par filtrage stable d'un bruit blanc stationnaire est un résultat essentiel, car elle indique comment construire des modèles stationnaires possédant une structure de corrélation particulière. En effet, dans le cas d'un bruit blanc réduit, (V.15) et (V.18) impliquent

$$C_X(n) = \sum_{k \in \mathbb{Z}} h_k h_{k+n}, \quad (\text{V.24})$$

c'est-à-dire que l'autocorrélation de la réponse impulsionnelle  $h$  du filtre définie au sens des signaux déterministes stables (donc d'énergie finie) se transporte sur la fonction d'autocorrélation du signal aléatoire. En particulier, l'énergie du filtre donne la puissance moyenne du signal aléatoire d'après (V.16), ce dernier étant d'énergie infinie. Nous verrons dans le chapitre suivant que ce résultat est transposable sous forme de représentation harmonique : la densité spectrale d'énergie du filtre devient la densité spectrale de puissance du signal de sortie.

Ces résultats sont évidemment valables pour les filtres rationnels, qui nous ont permis de définir les ARMA stationnaires.

### V.4.3 Ergodisme

Dans le paragraphe précédent, nous avons souligné l'utilité des modèles stationnaires quand le phénomène mesuré existe de façon permanente indépendamment de son observation. En fait, dans cette situation, on ne peut souvent observer qu'un fragment d'une unique trajectoire (on peut considérer par exemple l'imagerie de l'activité solaire au cours du temps). Dans ces conditions, comment (et pourquoi) choisir ou identifier un modèle aléatoire ? La solution de ce problème passe souvent par la restriction à des modèles aléatoires *ergodiques*, c'est-à-dire dont on peut identifier entièrement ou en partie la loi temporelle par des moyennes temporelles sur une seule trajectoire.

— Prenons l'exemple de la moyenne  $m_X = E(X_k)$  d'un signal  $X$  faiblement stationnaire :  $X$  est *ergodique pour sa moyenne* ssi

$$\lim_{n \rightarrow \infty} \frac{1}{2n+1} \sum_{i=-n}^n X_i = m_X$$

au sens de la convergence presque sûre. Nous avons déjà vu un cas simple d'ergodisme pour la moyenne : la loi des grands nombres (V.22) assure l'ergodisme pour la moyenne d'un bruit blanc stationnaire d'ordre deux.

— Plus généralement,  $X$  est ergodique pour une fonction  $\varphi$  si

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \varphi(X_{k_{1+i}}, \dots, X_{k_{N+i}}) = E(\varphi(X_{k_1}, \dots, X_{k_N})) \quad (\text{V.25})$$

presque sûrement pour tout  $N$ , tout  $N$ -uplet d'instants distincts  $(k_1, \dots, k_N)$ .

—  $X$  est ergodique *au sens strict* si (V.25) est vérifié pour toute fonction  $\varphi$  telle que  $E(\varphi(X_{k_1}, \dots, X_{k_N}))$  ait un sens.

—  $X$  est ergodique *à l'ordre deux* si (V.25) est vraie pour la moyenne et pour l'autocorrélation.

Les principaux exemples de signaux stationnaires présentés dans ce chapitre sont ergodiques : chaînes de Markov stationnaires, ARMA stationnaires d'ordre deux.

## V.5 Les signaux gaussiens

Un signal gaussien est un signal aléatoire dont on ne peut extraire que des vecteurs gaussiens. On a vu dans le chapitre précédent que toute transformation affine d'un vecteur gaussien donnait un vecteur gaussien. Cette propriété se transpose donc aux signaux gaussiens. C'est une première propriété de stabilité, mais il y en a une autre qui est la stabilité en moyenne quadratique : la limite d'un signal gaussien convergeant en moyenne quadratique est une variable gaussienne. Pour démontrer cette propriété, il suffit de vérifier que la limite des fonctions caractéristiques est la fonction caractéristique d'une gaussienne.

D'après ce résultat, on peut dire que l'histoire ou le passé linéaire d'un signal gaussien sont des espaces de Hilbert gaussiens. Dans les espaces gaussiens, les variables sont d'ordre deux et les lois de probabilité sont déterminées entièrement par des moyennes et des corrélations. Stationnarité faible et forte d'un signal gaussien sont équivalentes.



# Chapitre VI

## Signaux aléatoires stationnaires faibles

### VI.1 Introduction

La représentation fréquentielle des signaux aléatoires s'avère un outil précieux puisqu'elle fournit des informations importantes sur le caractère périodique ou pseudo-périodique d'un signal. Nous verrons dans la suite de ce chapitre que de même que pour les signaux déterministes harmonisables, on peut définir une telle représentation pour les signaux aléatoires stationnaires faibles, en commençant par associer à leur fonction d'autocorrélation une *mesure* ou éventuellement une *densité spectrale de puissance*.

Comme dans le cas déterministe, nous précisons le lien essentiel entre la représentation fréquentielle et le filtrage linéaire convolutionnel pour les signaux aléatoires stationnaires faibles.

L'*analyse spectrale* consiste à chercher les caractéristiques spectrales d'un signal supposé stationnaire faible à partir d'un extrait de ce signal. Nous présenterons successivement l'analyse spectrale non paramétrique, utilisant le périodogramme, et l'analyse spectrale paramétrique qui exploite les propriétés des signaux ARMA introduits au chapitre V. Ces méthodes constitueront une première introduction aux problèmes d'estimation qui feront l'objet du prochain module de cours.

### VI.2 Représentation harmonique

L'étude de la représentation harmonique des signaux a déjà été abordée dans les chapitres I et II pour le cas des signaux déterministes. Dans le cas des signaux stationnaires faibles, deux résultats essentiels seront énoncés. Tout d'abord, le *théorème de Herglotz* permet de garantir l'existence d'une mesure spectrale de puissance définie, de même que pour les signaux déterministes, comme la transformée de Fourier de la fonction d'autocorrélation (voir (I.18) du chapitre I). En général, l'analyse harmonique du signal par l'intermédiaire de sa fonction d'autocorrélation est d'ailleurs suffisante dans les problèmes d'identification, d'estimation et de filtrage. Nous nous intéresserons ensuite à la représentation de Fourier du signal aléatoire lui-même. Le *théorème de Cramer-Khintchine* assure l'existence et l'unicité d'une telle représentation pour presque toute trajectoire du signal aléatoire.

#### VI.2.1 Fonction d'autocorrélation

D'après le chapitre précédent, un signal aléatoire à temps discret stationnaire du second ordre admet une valeur moyenne constante  $m_X = E(X_n)$  (*i.e.*, indépendante de  $n$ ), et une fonction

d'autocorrélation de la forme

$$r_X(n, p) = C_X(n - p).$$

On appellera par la suite  $C_X(k)$  sa fonction d'autocorrélation et on supposera le signal centré ( $m_X = 0$ ).

L'existence de la mesure spectrale de la fonction d'autocorrélation est garantie par le théorème de Herglotz, dont l'énoncé est le suivant.

—Théorème de Herglotz

Soit  $C : \mathbb{Z} \rightarrow \mathbb{C}$  une fonction définie non négative, c'est-à-dire telle que

$$\sum_{i=0}^{N-1} \sum_{k=0}^{N-1} C(i-k) \lambda_i \lambda_k^* \geq 0$$

pour tout  $N \geq 1$ ,  $(\lambda_0, \dots, \lambda_{N-1}) \in \mathbb{C}^N$ . Alors il existe une mesure finie  $\mu$  sur l'intervalle  $\Pi = ] - 1/2, 1/2 ]$  telle que

$$C(k) = \int_{\Pi} e^{2i\pi k\nu} \mu(d\nu)$$

et cette mesure est unique.

La démonstration du théorème de Herglotz est présentée dans [1]. On en déduit que la fonction d'autocorrélation du signal  $X$  admet une représentation sous la forme

$$C_X(k) = \int_{\Pi} e^{2i\pi k\nu} \mu_X(d\nu) \quad (\text{VI.1})$$

où  $\mu_X$  est unique. Pour  $k = 0$ , cette égalité reste vraie, donc

$$C_X(0) = \int_{\Pi} \mu_X(d\nu) = \mu_X \left( \left] -\frac{1}{2}, \frac{1}{2} \right] \right)$$

qui est la puissance moyenne  $E(X_n^2)$  du signal. La mesure  $\mu_X$  est dite *mesure spectrale de puissance* du signal  $X$ . Si cette mesure est à densité vis-à-vis de la mesure de Lebesgue, c'est-à-dire s'il existe une fonction réelle non négative  $\phi_X(\nu)$  telle que

$$C_X(k) = \int_{\Pi} e^{2i\pi k\nu} \phi_X(\nu) d\nu,$$

$\phi_X(\nu)$  est appelée la *densité spectrale de puissance* de  $X$ .

Si, de plus,  $C_X$  est dans  $\ell^2$  et admet donc une transformée de Fourier, on a alors

$$\phi_X(\nu) = \sum_{k \in \mathbb{Z}} C_X(k) e^{-2i\pi k\nu}.$$

Lorsqu'il existe une mesure ou une densité spectrale de puissance, il est donc possible d'analyser en fréquence la répartition d'énergie du signal aléatoire en utilisant la transformée de Fourier discrète de sa fonction d'autocorrélation.

Si la mesure  $\mu_X$  est une somme pondérée de masses ponctuelles,

$$\mu_X = \sum_{k \in \mathbb{Z}} a_k \delta(\nu - \nu_k) \quad (\text{VI.2})$$

où  $\delta(\nu - \nu_k)$  est la mesure de Dirac au point  $\nu_k$ , on dit que le signal  $X_k$  admet un spectre de raies,  $a_k$  étant l'intensité de la raie  $\nu_k$  ( $a_k > 0$ ).

### VI.2.2 Cas d'un bruit blanc numérique

Soit un bruit blanc  $\{B_n\}$  stationnaire faible centré, donc tel que

$$E(B_{n+k} B_n) = 0, \quad \forall k \in \mathbb{Z}^*.$$

Notons  $\sigma^2 = E(B_n^2)$ ; la fonction d'autocorrélation de  $\{B_n\}$  s'écrit :

$$C_B(k) = \sigma^2 \delta(k) = \int_{\Pi} e^{2i\pi k\nu} \sigma^2 d\nu,$$

et donc  $\{B_n\}$  admet une densité spectrale de puissance constante

$$\phi_B(\nu) = \sigma^2. \quad (\text{VI.3})$$

### VI.2.3 Trajectoires

Après avoir étudié le problème de la représentation spectrale de la fonction d'autocorrélation, nous allons nous intéresser au problème de la représentation spectrale des trajectoires du signal  $\{X_n\}$ . Pour chaque trajectoire du signal, l'existence de cette décomposition devient un problème déterministe qui a été traité dans les premiers chapitres dans les cas stables, d'énergie finie ou harmonisable.

Considérons tout d'abord un signal stationnaire ergodique, l'équation (V.24) nous permet d'écrire que p.s.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n |X_k| = E(|X_1|)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n |X_k|^2 = E(|X_1|^2).$$

Donc si  $X$  n'est pas identiquement nul,  $E(|X|)$  et  $E(|X|^2)$  sont des quantités strictement positives, ce qui implique que p.s.

$$\sum_{k=1}^{\infty} |X_k| = \infty \quad \text{et} \quad \sum_{k=1}^{\infty} |X_k|^2 = \infty.$$

Pour presque tout  $\omega$ , la trajectoire  $X(k, \omega)$  n'est pas dans  $\ell^2$  ni dans  $\ell^1$  et l'existence systématique de la transformée de Fourier n'est pas assurée.

Il existe pourtant une forme de relation de Fourier entre les trajectoires de processus stochastiques. Pour l'introduire considérons le cas particulier d'un processus à spectre de raies défini par

$$X_k = \sum_{j=1}^N A_j e^{2i\pi\nu_j k}$$

où les  $A_j$  sont des variables aléatoires du second ordre. Il apparaît immédiatement que le processus ainsi défini n'est stationnaire du second ordre que si les  $A_j$  sont des variables non corrélées, puisque :

$$E(X_n X_p^*) = \sum_{j=1}^N E(|A_j|^2) e^{2i\pi\nu_j(n-p)} + \sum_{j=1}^N \sum_{l=1, l \neq j}^N E(A_j A_l^*) e^{2i\pi\nu_j n} e^{-2i\pi\nu_l p}.$$

Lorsque cette condition est satisfaite, il est naturel de considérer la suite des  $A_j$  comme la représentation spectrale du signal  $X_k$  stationnaire d'ordre deux.

La généralisation de cette représentation à l'ensemble des processus stationnaires faibles est l'objet du théorème de Cramer-Khintchine.

### Théorème de Cramer-Khintchine

Soit  $X_k$  un signal stationnaire du second ordre et  $\mu_X$  sa mesure spectrale de puissance. Il existe p.s. une représentation spectrale de toute trajectoire du signal aléatoire, c'est-à-dire que p.s., on peut écrire

$$X(k) = \int_{\Pi} e^{2i\pi k\nu} dx(\nu). \quad (\text{VI.4})$$

La représentation  $\{x(\nu)\}_{\nu \in \mathbb{R}}$  appelée *représentation spectrale de Cramer-Khintchine* est unique. C'est un signal aléatoire à accroissements non corrélés centrés de mesure structurelle  $\mu_X$ . Ceci signifie que, pour tout  $(\nu_1, \nu_2) \in \mathbb{R}^2$

$$\begin{aligned} \mathbb{E}(|x(\nu_2) - x(\nu_1)|^2) &< \infty \\ \mathbb{E}(x(\nu_2) - x(\nu_1)) &= 0 \end{aligned}$$

et pour tout  $]\nu_1, \nu_2] \subset \mathbb{R}, ]\nu_3, \nu_4] \subset \mathbb{R}$ ,

$$\mathbb{E}\left(\left(x(\nu_2) - x(\nu_1)\right)\left(x(\nu_4) - x(\nu_3)\right)\right) = \mu_X(]\nu_1, \nu_2] \cap ]\nu_3, \nu_4]).$$

L'intégrale (VI.4) est une intégrale de Wiener dont l'étude dépasse le cadre de ce cours. Le lecteur intéressé peut se rapporter à [1] ou à [7].

Dans le cas où  $X(k)$  est gaussien de mesure spectrale  $\mu_X$ , sa représentation spectrale de Cramer-Khintchine  $x(\nu)$  est un processus gaussien.

## VI.3 Filtrage linéaire convolutionnel

Soit  $\{X_n\}$  un signal aléatoire centré discret stationnaire du second ordre de mesure spectrale de puissance  $\mu_X$ , et  $\{h_n\}$  la réponse impulsionnelle d'un filtre convolutionnel stable. Montrons que pour tout  $n \in \mathbb{Z}$ , la somme

$$\sum_{k \in \mathbb{Z}} h_{n-k} X_k \quad (\text{VI.5})$$

converge p.s. vers une v.a.r.  $Y_n$ . En effet,  $(|X_n| - 1)^2 \geq 0$ , donc

$$\mathbb{E}(|X_n|) \leq 1 + \mathbb{E}(|X_n|^2) = 1 + C_X(0) = K < \infty,$$

où  $C_X(0)$  est la puissance de  $X$ . Ceci implique que

$$\mathbb{E}\left(\sum_{k \in \mathbb{Z}} |h_{n-k}| |X_k|\right) = \sum_{k \in \mathbb{Z}} |h_{n-k}| \mathbb{E}(|X_k|) \leq K \sum_{k \in \mathbb{Z}} |h_{n-k}| < \infty$$

ce qui assure la convergence presque sûre de la somme (VI.5). On définit ainsi un signal aléatoire  $\{Y_n\}$  qui est la sortie du filtre ayant pour signal d'entrée  $\{X_n\}$ . Le signal  $\{X_n\}$  étant centré,  $\mathbb{E}(Y_n) = 0$ . Notons  $\tilde{H}(\nu)$  la transformée de Fourier associée à la réponse impulsionnelle du filtre

$$\tilde{H}(\nu) = \sum_{k \in \mathbb{Z}} h_k e^{-2i\pi k\nu}.$$

On peut montrer qu'alors le signal  $\{Y_n\}$  est stationnaire faible et sa mesure spectrale de puissance s'exprime sous la forme

$$\mu_Y = |\tilde{H}(\nu)|^2 \mu_X(d\nu). \quad (\text{VI.6})$$

Lorsque le signal d'entrée  $X$  admet une densité spectrale de puissance  $\phi_X(\nu)$ ,  $Y$  admet également une densité spectrale et le résultat précédent se met sous la forme

$$\boxed{\phi_Y(\nu) = |\tilde{H}(\nu)|^2 \phi_X(\nu)}. \quad (\text{VI.7})$$

Notons enfin que le filtrage linéaire d'un signal gaussien conserve le caractère gaussien.

## VI.4 Analyse spectrale classique

### VI.4.1 Position du problème

Soit  $\{X_n\}$  un signal aléatoire discret stationnaire du second ordre centré de fonction d'auto-corrélation  $C_X(k)$ , admettant la densité spectrale de puissance

$$\phi_X(\nu) = \sum_{k \in \mathbb{Z}} C_X(k) e^{-2i\pi k\nu}. \quad (\text{VI.8})$$

Dans la pratique, on travaille la plupart du temps avec un certain nombre de réalisations (souvent une seule) d'un extrait  $X_0, \dots, X_{N-1}$  du signal  $\{X_n\}$ . La question que l'on examinera ici est la suivante : peut-on approcher la densité spectrale de puissance  $\phi_X(\nu)$  du signal  $\{X_n\}$  à partir de ces échantillons ? Ce problème, qui constitue ce que l'on appelle l'*analyse spectrale* du signal  $\{X_n\}$ , est un des plus anciens problèmes du traitement de données expérimentales qui consiste, en clair, à rechercher des pseudo-périodes dans une suite de données au comportement irrégulier.

### VI.4.2 Périodogramme

#### Définition

L'analyse spectrale a été naturellement conduite à l'origine par analogie avec l'analyse fréquentielle des signaux déterministes. Lorsque l'on dispose d'une suite de  $N$  échantillons  $x_0, \dots, x_{N-1}$ , la répartition de l'énergie de cette séquence suivant les fréquences est donnée par le module au carré de la transformée de Fourier :

$$\frac{1}{N} \left| \sum_{k=0}^{N-1} x_k e^{-2i\pi\nu k} \right|^2.$$

La suite  $x_0, \dots, x_{N-1}$  est maintenant considérée comme la réalisation d'un extrait  $X_0, \dots, X_{N-1}$  du signal aléatoire stationnaire faible  $\{X_n\}$  vérifiant les hypothèses de l'introduction. L'expression précédente est alors la réalisation de la fonction aléatoire

$$\hat{\phi}_N(\nu) = \frac{1}{N} \left| \sum_{k=0}^{N-1} X_k e^{-2i\pi\nu k} \right|^2. \quad (\text{VI.9})$$

Cette fonction est appelé le *périodogramme*. Notons que cette fonction est en pratique calculée sur les fréquences multiples de  $1/N$  fois la fréquence d'échantillonnage par un algorithme de transformée de Fourier rapide (TFR ou FFT, présentée au chapitre II). Si l'on désire obtenir les valeurs de  $\hat{\phi}_N$  sur d'autres fréquences, il est possible d'interpoler l'expression précédente en effectuant du « bourrage de zéro » (voir le chapitre II).

Dans la suite, nous allons montrer les qualités et les défauts de l'estimation de la densité spectrale de puissance par le périodogramme. Pour cela, nous ferons usage d'une autre expression du périodogramme.

#### Transformée de Fourier du périodogramme

En développant le module dans l'expression du périodogramme (VI.9), on a

$$\hat{\phi}_N(\nu) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} X_l X_m e^{-2i\pi(l-m)\nu}.$$

En posant  $k = l - m$ , l'expression précédente devient :

$$\widehat{\phi}_N(\nu) = \frac{1}{N} \sum_{k=1-N}^{N-1} e^{-2i\pi k\nu} \sum_{m=0}^{N-1-|k|} X_m X_{m+|k|}.$$

Notons  $\widehat{C}_N(k)$  la suite définie par :

$$\widehat{C}_N(k) = \begin{cases} \frac{1}{N} \sum_{j=0}^{N-|k|-1} X_j X_{j+|k|} & \text{si } |k| < N, \\ 0 & \text{sinon.} \end{cases} \quad (\text{VI.10})$$

On a alors :

$$\widehat{\phi}_N(\nu) = \sum_{k \in \mathbb{Z}} \widehat{C}_N(k) e^{-2i\pi k\nu} \quad (\text{VI.11})$$

qui est analogue à (VI.8).

La suite  $\widehat{C}_N(k)$ , qui est appelée *autocorrélation empirique biaisée*, possède la propriété suivante

$$\mathbb{E} \left( \widehat{C}_N(k) \right) = \left( 1 - \frac{|k|}{N} \right) C(k), \quad (\text{VI.12})$$

qui implique que pour  $N$  tendant vers l'infini,  $\mathbb{E}(\widehat{C}_N(k))$  tend vers la fonction d'autocorrélation du processus  $C(k)$ . Cette propriété et la relation (VI.11) justifient intuitivement le choix du périodogramme pour approximer la densité spectrale de puissance. Elle permet de plus d'étudier précisément la qualité de cette approximation, comme nous allons le voir dans la suite.

### Caractéristiques de l'analyse spectrale par périodogramme

En utilisant l'expression précédente, on peut calculer la moyenne du périodogramme  $\widehat{\phi}_N(\nu)$  :

$$\mathbb{E}(\widehat{\phi}_N(\nu)) = \sum_{k=1-N}^{N-1} \left( 1 - \frac{|k|}{N} \right) C(k) e^{-2i\pi k\nu}$$

ce qui montre que pour  $N$  fixé, il existe une différence entre la valeur moyenne  $\mathbb{E}(\widehat{\phi}_N(\nu))$  et la densité spectrale  $\phi(\nu)$ . Cette différence est appelée un *biais*. On s'intéresse ensuite à l'évolution de ce biais lorsque le nombre d'échantillons  $N$  tend vers l'infini. En supposant

$$\sum_{k \in \mathbb{Z}} |C(k)| < \infty,$$

on montre, à l'aide du théorème de convergence dominée de Lebesgue

$$\lim_{N \rightarrow \infty} \mathbb{E}(\widehat{\phi}_N(\nu)) = \phi(\nu).$$

On étudie la convergence en moyenne quadratique de la suite des v.a.  $\widehat{\phi}_N(\nu)$  vers la valeur de la densité spectrale  $\phi(\nu)$ , pour une fréquence  $\nu$  fixée quelconque, c'est-à-dire que l'on s'intéresse à la variance asymptotique

$$\lim_{N \rightarrow \infty} \mathbb{E} \left( (\widehat{\phi}_N(\nu) - \phi(\nu))^2 \right).$$

On peut montrer que cette valeur est en général non nulle. Si on considère par exemple un bruit blanc gaussien de variance  $\sigma^2$ , dont la densité spectrale est constante  $\phi_X(\nu) = \sigma^2$ , alors on a

$$\sigma^{-4} \mathbb{E} \left( (\widehat{\phi}_N(\nu) - \sigma^2)^2 \right) = 1 - \frac{1}{N} + \frac{2}{N^2} \left| \frac{1 - e^{2i\pi\nu}}{1 - e^{2i\pi N\nu}} \right|^2,$$

expression qui ne tend pas vers 0 lorsque  $N$  tend vers l'infini. Il n'y a pas convergence en moyenne quadratique du périodogramme vers la densité spectrale de puissance. Pratiquement, ceci constitue un très grave défaut pour l'analyse spectrale. En effet, grossièrement, un périodogramme même calculé avec un très grand nombre d'échantillons conserve un aspect très chahuté, du fait de la variance résiduelle non nulle. Cet aspect gêne considérablement l'interprétation du résultat. Ci-après, nous présentons une technique classique de réduction de cette variance asymptotique qui est obtenue au prix d'une perte de résolution.

La figure VI.1 illustre les résultats obtenus à l'aide du périodogramme simple pour le signal de sortie d'un filtre récursif d'ordre 2 dont l'entrée est un bruit blanc gaussien. Les caractéristiques du filtre utilisé sont celles du filtre présenté dans le §III.2.4.

### Périodogramme moyenné

Pour remédier aux limitations du périodogramme, on définit le périodogramme moyenné de la façon suivante : la séquence d'observations de  $N$  échantillons est divisée en  $K$  sous-séquences chacune de longueur  $M$  avec  $N = KM$ . Sur chacune de ces sous-séquences  $\{kM, (k+1)M - 1\}$  on peut calculer un périodogramme

$$\widehat{\phi}_{KN}^k(\nu) = \frac{1}{M} \left| \sum_{n=0}^{M-1} X_{kM+n} e^{-2i\pi n\nu} \right|^2. \quad (\text{VI.13})$$

Le périodogramme moyenné est alors défini comme la moyenne empirique des périodogrammes calculés sur les  $K$  sous-séquences. Sous hypothèse d'ergodicité, on peut alors montrer que

$$\begin{aligned} \lim_{K, N \rightarrow \infty} \mathbb{E} \left( \widehat{\phi}_{KN}^k(\nu) - \phi(\nu) \right) &= 0 \\ \lim_{K, N \rightarrow \infty} \mathbb{E} \left( (\widehat{\phi}_{KN}^k(\nu) - \phi(\nu))^2 \right) &= 0 \end{aligned}$$

donc le périodogramme moyenné converge en moyenne quadratique vers la densité spectrale de puissance du signal.

La figure VI.2 montre la densité spectrale de puissance obtenue à l'aide du périodogramme moyenné sur un signal de caractéristiques identiques à celles utilisées pour le périodogramme simple. On peut constater l'effet de lissage de la courbe obtenu lorsque le paramètre  $K$  augmente.

### Conclusion

La méthode du périodogramme présente l'avantage d'être conceptuellement simple et de pouvoir exploiter les algorithmes rapides de mise en œuvre de la transformée de Fourier discrète (FFT). Lorsqu'on dispose d'un nombre important d'échantillons, la méthode du périodogramme converge en moyenne vers la densité spectrale de puissance et en moyenne quadratique si on utilise l'approche par périodogramme moyenné. Cependant elle présente un certain nombre d'inconvénients en pratique et en particulier ne permet pas d'obtenir une bonne discrimination entre des raies spectrales ou des pics très rapprochés. D'autre part, si le nombre d'échantillons dont on dispose est faible, l'estimation peut s'avérer médiocre. Différentes méthodes permettant de pallier cet inconvénient pour certaines classes de signaux et en particulier pour les signaux AR et ARMA, ont été développées. Dans la suite de ce chapitre on se limitera au cas des AR pour exposer les principes de base de ces autres approches.

## VI.5 Analyse spectrale paramétrique

Les modèles autorégressifs à moyenne mobile (ARMA), et parmi ceux-ci les modèles autorégressifs (AR), ont été présentés au chapitre V pour illustrer la construction de signaux aléatoires.

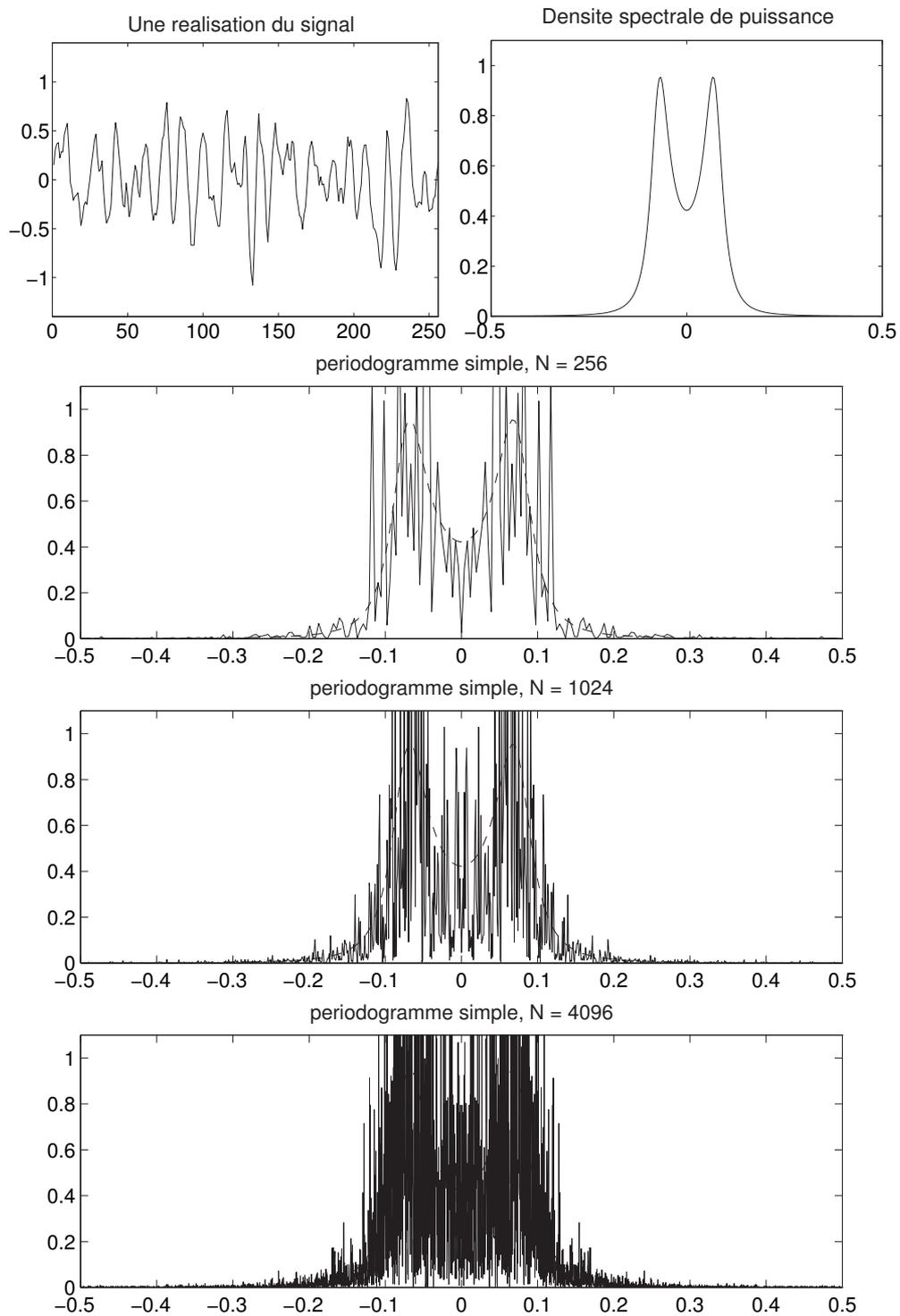


Fig. VI.1. Périodogramme simple

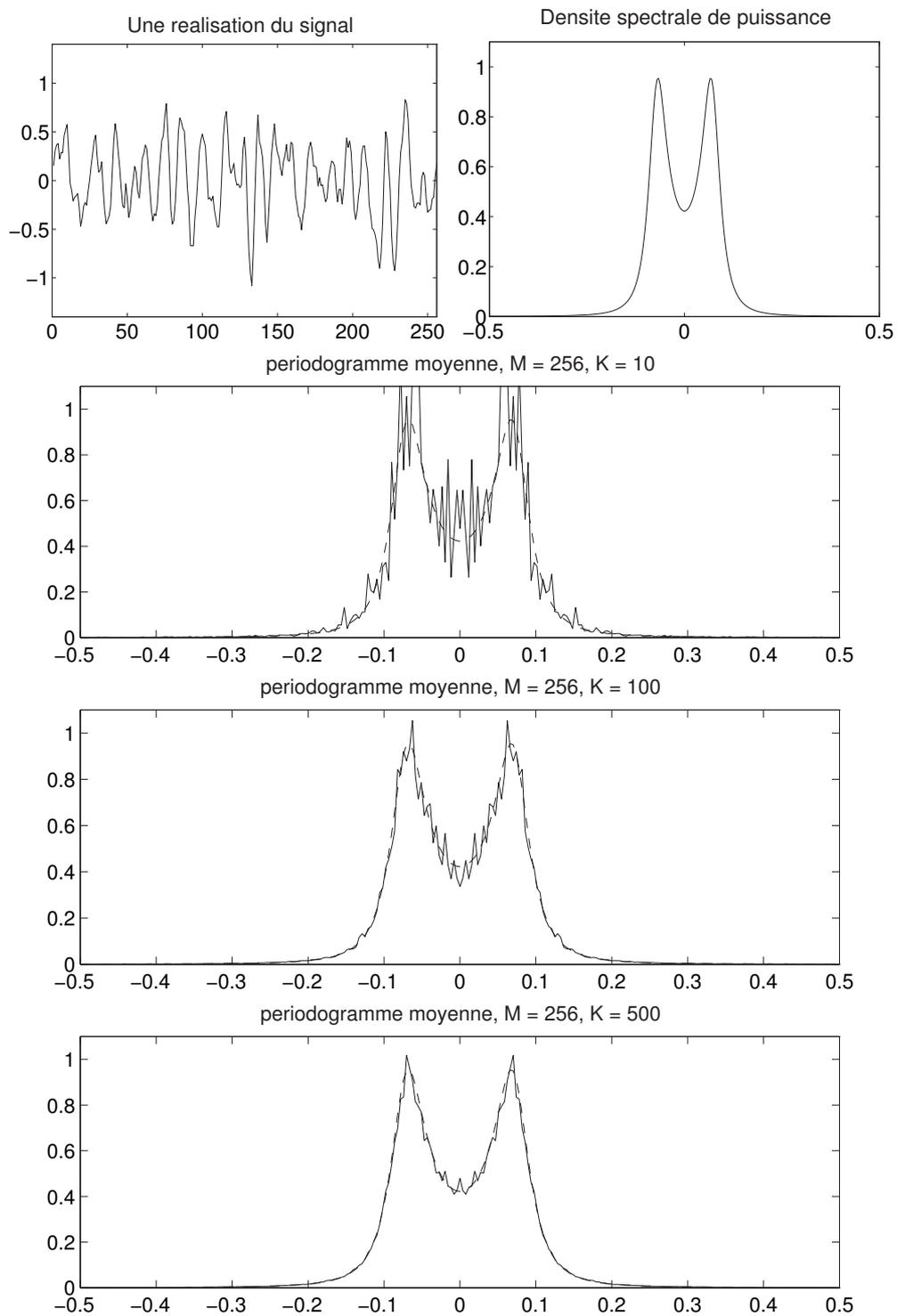


Fig. VI.2. Périodogramme moyenné

Les signaux ARMA jouent un rôle très important en traitement du signal, d'une part parce qu'ils permettent de représenter de façon réaliste un certain nombre de signaux « physiques » (parole humaine par exemple) et d'autre part parce que leur structure spécifique peut être directement exploitée comme nous le verrons dans le domaine de l'analyse spectrale et de l'estimation. Nous allons à présent rappeler brièvement leur définition et étudier leurs caractéristiques en terme de densité spectrale.

### VI.5.1 Définition

Un signal aléatoire à temps discret stationnaire du second ordre est dit *ARMA d'ordre  $p, q$* , ou *ARMA( $p, q$ )* si il existe deux ensembles de coefficients réels  $\{a_i\}$  et  $\{b_j\}$  de dimensions respectives  $p$  et  $q$  et un bruit blanc échantillonné  $\{B_n\}$  de variance  $\sigma^2$  tels que

$$X_n + \sum_{i=1}^p a_i X_{n-i} = B_n + \sum_{j=1}^q b_j B_{n-j}, \quad \forall n \in \mathbb{Z}. \quad (\text{VI.14})$$

— Lorsque  $q$  est égal à 0, le signal  $\{X_n\}$  est dit AR d'ordre  $p$ , et on a

$$X_n + \sum_{i=1}^p a_i X_{n-i} = B_n.$$

— Lorsque  $p$  est égal à 0, le signal  $\{X_n\}$  est dit à moyenne mobile (MA de l'anglais *Moving Average*) d'ordre  $q$  :

$$X_n = B_n + \sum_{j=1}^q b_j B_{n-j}.$$

*Remarque :*

En automatique, ces représentations sont fréquentes mais avec l'adjonction de la prise en compte de la commande  $u_k$  supposée connue sous la forme

$$X_n + \sum_{i=1}^p a_i X_{n-i} = B_n + \sum_{j=1}^q b_j B_{n-j} + \sum_{k=1}^r c_k u_{n-k}, \quad \forall n \in \mathbb{Z}.$$

Pour tenir compte de la présence de  $u_k$ , le signal est dénommé ARMAX où le X indique la présence d'une *entrée eXogène*. On définit de façon similaire les ARX.

### VI.5.2 ARMA réguliers

Le signal  $\{X_n\}$  peut s'interpréter comme le signal de sortie d'un filtre de fonction de transfert

$$\frac{1 + \sum_{j=1}^q b_j z^{-j}}{1 + \sum_{i=1}^p a_i z^{-i}} = \frac{P(z)}{Q(z)}$$

dont l'entrée est le bruit blanc  $\{B_n\}$ . Le signal ARMA est dit *régulier* si la fonction de transfert est stable, *i.e.*, si le polynôme  $Q(z)$  n'a pas de racine sur le cercle unité. Dans ce cas, on montre (VI.7) que la densité spectrale de puissance du signal  $\{X_n\}$  s'écrit

$$\phi_X(\nu) = \sigma^2 \left| \frac{P(e^{2i\pi\nu})}{Q(e^{2i\pi\nu})} \right|^2.$$

De manière unique, on peut trouver  $P'(z)$  et  $Q'(z)$ , deux polynômes à coefficients réels *sans racine hors du disque unité*  $\{z, |z| \leq 1\}$ , tels que  $P'(0) = Q'(0) = 1$ , et tels que

$$\phi_X(\nu) = \sigma^2 \left| \frac{P'(e^{2i\pi\nu})}{Q'(e^{2i\pi\nu})} \right|^2.$$

Dans ce cas, (VI.15) correspond au filtrage linéaire *stable et causal* du bruit blanc  $\{B_n\}$  par le filtre rationnel de fonction de transfert  $P'(z)/Q'(z)$ . La représentation associée à  $P'(z)$  et  $Q'(z)$  est la *représentation canonique* de  $\{X_n\}$ . Le bruit blanc associé  $B_n$  apparaissant dans (VI.14) est alors non corrélé à  $X_{n-k}$ ,  $\forall k > 0$ .

### VI.5.3 Représentation spectrale des signaux autorégressifs

Soit  $\{X_n\}$  un signal centré stationnaire du second ordre autorégressif régulier d'ordre  $p$ . Dans ce cas, la densité spectrale de puissance du signal  $X$  est définie par

$$\phi_X(\nu) = \frac{\sigma^2}{|Q(e^{2i\pi\nu})|^2}$$

où  $\sigma^2$  est la variance du bruit blanc générateur. En utilisant la représentation du signal AR sous la forme

$$X_n + \sum_{k=1}^p a_k X_{n-k} = B_n,$$

l'équation précédente peut s'écrire

$$\phi_X(\nu) = \frac{\sigma^2}{|1 + \sum_{k=1}^p a_k e^{-2i\pi k\nu}|^2} \quad (\text{VI.15})$$

qui correspond au module au carré de la FFT du signal de sortie du filtre. Il suffit donc de connaître les paramètres  $a_k$  du modèle autorégressif pour obtenir la valeur de la densité spectrale de puissance.

### VI.5.4 Équations de Yule-Walker

On considère un signal AR dont on ne connaît pas la représentation canonique en termes des coefficients  $a_k$ ,  $k = 1, \dots, p$  mais pour lequel on dispose d'une suite de valeurs de la fonction d'autocorrélation  $C_X(k)$ . On cherche à calculer sa densité spectrale de puissance  $\phi_X$  et donc d'après (VI.15) dans un premier temps à déterminer les coefficients  $a_k$  du modèle AR.

La fonction d'autocorrélation de  $\{X_n\}$  est définie par

$$\begin{aligned} C_X(j) = E(X_n X_{n+j}) &= E\left(X_n \left(B_{n+j} - \sum_{k=1}^p a_k X_{n-k+j}\right)\right) \\ &= E(X_n B_{n+j}) - \sum_{k=1}^p a_k C_X(j-k). \end{aligned}$$

Le deuxième terme  $E(B_{n+j} X_n)$  se simplifie puisque la canonicité impose l'indépendance entre les réalisations du bruit blanc à l'instant  $j$  et les valeurs passées du signal, on a donc

$$E(B_{n+j} X_n) = \begin{cases} 0 & \text{si } j > 0, \\ \sigma^2 & \text{si } j = 0. \end{cases}$$

L'expression générale de  $C_X(j)$  est donc

$$C_X(j) = \begin{cases} -\sum_{k=1}^p a_k C_X(j-k) & \text{si } j > 0, \\ -\sum_{k=1}^p a_k C_X(-k) + \sigma^2 & \text{si } j = 0. \end{cases} \quad (\text{VI.16})$$

Ces expressions définissent les équations de Yule-Walker. Pour déterminer les coefficients  $a_k$  lorsqu'on dispose des valeurs de  $C_X$ , il suffit donc de résoudre  $p$  équations pour  $j > 0$  et de déterminer  $\sigma^2$  pour  $j = 0$ . Ces équations peuvent être exprimées sous forme matricielle par

$$\begin{bmatrix} C(0) & C(-1) & \dots & C(-(p-1)) \\ C(1) & C(0) & \dots & C(-(p-2)) \\ \vdots & \vdots & \ddots & \vdots \\ C(p-1) & C(p-2) & \dots & C(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} C(1) \\ C(2) \\ \vdots \\ C(p) \end{bmatrix}.$$

Le signal étant réel, l'équation précédente peut s'écrire sous la forme plus simple

$$\begin{bmatrix} C(0) & C(1) & \dots & C(p-1) \\ C(1) & C(0) & \dots & C(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ C(p-1) & C(p-2) & \dots & C(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} C(1) \\ C(2) \\ \vdots \\ C(p) \end{bmatrix}. \quad (\text{VI.17})$$

La matrice d'autocorrélation est symétrique et de Toeplitz. Il suffit finalement de connaître  $p$  coefficients d'autocorrélation pour déterminer la densité spectrale.

La détermination des coefficients  $a_k$  nécessite donc de résoudre le système précédent qui admet une solution unique si la matrice d'autocorrélation est inversible. L'inversion de cette matrice peut s'effectuer par les procédures classiques de type Gauss-Jordan mais on verra par la suite qu'en exploitant la structure spécifique de cette matrice on peut définir des algorithmes plus efficaces (Algorithme de Levinson) qui permettent de réduire le nombre d'opérations à effectuer de  $O(p^3)$  à  $O(p^2)$ .

### VI.5.5 Estimation spectrale autorégressive

Dans la pratique, disposant de  $N$  échantillons d'un signal inconnu, l'analyse spectrale autorégressive consiste à assimiler ce signal à un signal autorégressif d'ordre  $p$ , puis à estimer sa densité spectrale de puissance en calculant la fonction d'autocorrélation empirique biaisée (VI.10)

$$\hat{C}_N(k) = \frac{1}{N} \sum_{j=0}^{N-|k|-1} X_j X_{j+|k|}.$$

En utilisant ces coefficients dans les équations de Yule-Walker on peut alors en déduire les coefficients  $\hat{a}_1, \dots, \hat{a}_p$  par inversion du système VI.17 dans lequel on a remplacé  $C(k)$  par  $\hat{C}(k)$ . La densité spectrale correspondante est alors obtenue par

$$\hat{\Phi}_X(\nu) = \frac{\hat{\sigma}^2}{|1 + \sum_{k=1}^p \hat{a}_k e^{-2i\pi k\nu}|^2}.$$

où

$$\hat{\sigma}^2 = \hat{C}_N(0) + \sum_{j=1}^p \hat{C}_N(j) \hat{a}_j$$

obtenu à partir de l'équation (VI.16) pour  $k = 0$ . Cette technique d'analyse spectrale présente dans un certain nombre de cas l'avantage d'obtenir une meilleure discrimination entre raies et une variance inférieure à celle obtenue par l'analyse spectrale non paramétrique.

# Bibliographie

- [1] P. BRÉMAUD, 1993, *Signaux aléatoires pour le traitement du signal et les communications*, Ellipses, Paris.
- [2] E. JURY, 1964, *Theory and application of the z transform method*, John Wiley, New York.
- [3] M. LABARRÈRE, J.-P. KRIEF et B. GIMONET, 1993, *Le filtrage et ses applications*, Cepadues Edition.
- [4] M. J. LIGHTHILL, 1959, *An introduction to Fourier analysis and generalized functions*, Cambridge University Press.
- [5] M. LOÈVE, 1977, *Probability theory I*, Springer Verlag, New York.
- [6] L. SCHWARTZ, 1965, *Méthodes mathématiques pour les sciences physiques*, Hermann, Paris.
- [7] C. SOIZE, 1993, *Éléments mathématiques de la théorie déterministe et aléatoire du signal*, cours de l'ENSTA, Masson, Paris.