

Notes de cours sur l'égalisation aveugle

Joël Le Roux

5 janvier 2004

Table des matières

1	Introduction	3
1.1	Objectif de l'égalisation aveugle	3
1.2	Plan du cours	4
2	Les premières techniques de déconvolution aveugle	4
2.1	Le forçage à zéro des coefficients de la réponse impulsionnelle de Lucky ("zero forcing")	4
2.2	Le critère de Sato	5
2.3	Le forçage à module constant de Godard (CMA "constant modulus algorithm")	6
2.3.1	Analyse du critère de Godard	7
2.4	La formulation de Benveniste <i>et al</i>	10
2.5	L'égalisation n'utilisant que les statistiques d'ordre deux de Tong <i>et al.</i>	12
3	Formulation en termes de statistiques d'ordre supérieur	12
3.1	Formules de base, cumulants	13
3.1.1	Fonctions caractéristiques, moments et cumulants dans le cas d'une variable aléatoire scalaire	13
3.1.2	Fonctions caractéristiques, moments et cumulants dans le cas de vecteurs aléatoires	14
3.1.3	Relation entre cumulants et moments de v.a. réelles	14
3.1.4	Quelques propriétés des moments et des cumulants	15
3.1.5	Cumulant d'ordre quatre pour des variables aléatoires complexes	16
3.2	Triple corrélation et filtrage linéaire, méthode d'identification de Giannakis et Mendel	16
3.2.1	Corrélation d'ordre trois (triple corrélation)	16
3.2.2	Filtrage d'une séquence i.i.d.	17
3.2.3	Support de la corrélation d'ordre trois d'une séquence i.i.d. filtrée par un filtre à réponse impulsionnelle finie	17
3.2.4	Identification de filtre à non minimum de phase (Giannakis et Mendel)	18
3.3	La maximisation du kurtosis de Shalvi et Weinstein	19
3.3.1	Corrélation d'ordre quatre des signaux complexes	19
3.3.2	Maximisation du kurtosis	19
3.3.3	Minimisation d'un critère en respectant une contrainte	25
4	Interprétation spectrale, spectres d'ordre trois (bispectre) et d'ordre quatre (trispectre)	26
4.1	Formules	26
4.1.1	Stationnarité dans le cas des signaux à temps continu	27
4.1.2	Spectre d'ordre trois ou bispectre	27
4.1.3	Spectre d'ordre quatre ou trispectre	28
4.2	Spectres d'ordre supérieur, stationnarité et échantillonnage	30
4.3	Spectres d'ordre supérieur et filtrage	31
4.3.1	Spectre d'ordre trois et filtrage	31
4.3.2	Obtention de la triple corrélation par transformée de Fourier inverse du spectre d'ordre trois	32
4.3.3	Résultat du filtrage sur le spectre d'ordre quatre d'une séquence i.i.d.	33
4.3.4	Validation des hypothèses sur la génération du signal étudié	33
4.4	Effet d'un retard sur les spectres d'ordre supérieur des signaux échantillonnés	33

4.4.1	Statistiques d'ordre quatre et synchronisation	35
4.4.2	Statistiques d'ordre quatre et stationnarité	35
4.5	Reconstruction d'une fonction de transfert dans le domaine des fréquences	35
4.6	Les perturbations en imagerie par interférométrie en astronomie	35
4.7	Clôture de phase	36
4.8	Maximisation du kurtosis dans le domaine des fréquences	39

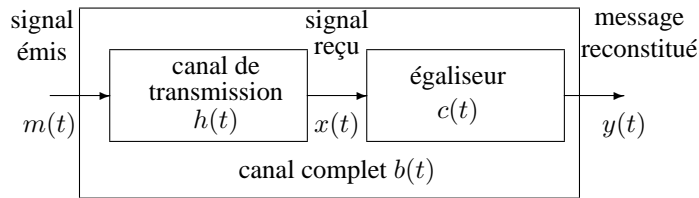


Figure 1: Schéma de l'égalisation

1 Introduction

1.1 Objectif de l'égalisation aveugle

Dans un système de transmission, le milieu où se propage l'onde porteuse du signal informatif déforme cette onde et donc le signal utile. Faute d'éléments de connaissances supplémentaires, on suppose en général que cette déformation est linéaire et que ses effets sont toujours les mêmes, qu'il y a (au moins dans une certaine mesure) stationnarité ou invariance dans le temps. (Nous ne considérons pas le cas de l'évanouissement brusque du signal qui ne vérifie pas ce type d'hypothèses).

Par ailleurs l'antenne qui capte le signal à l'entrée du récepteur capte aussi d'autres signaux parasites, qui seront en général considérés en l'absence d'informations supplémentaires comme des bruits souvent supposés blancs et gaussiens. On suppose que le signal capté se présente sous la forme du résultat d'une convolution à laquelle s'additionne un bruit de mesure $v(t)$, en principe indépendant du signal émis $m(t)$

$$x(t) = \sum_{k=0}^q h(k)m(t-k) + v(t). \quad (1)$$

Les données émises, les caractéristiques du canal, les coefficients des filtres seront des nombres complexes; les problèmes de synchronisation et de génération du signal complexe à partir du signal réel modulé capté sont supposés connus et résolus; on traitera donc le signal en bande de base à valeurs complexes après démodulation.

Un exemple qu'on trouve couramment est celui d'un écho où le signal capté est la somme du signal original retardé et d'une réplique atténuée de ce signal elle aussi retardée:

$$x(t) = am(t-\tau) + a'm(t-\tau'), \quad (2)$$

mais où les retards τ et τ' ne sont pas nécessairement des multiples de la période d'échantillonnage.

Dans ce cas, la réponse du canal de transmission (après filtrage passe-bas) est

$$h(t) = a \frac{\sin \pi(t-\tau)}{\pi(t-\tau)} + a' \frac{\sin \pi(t-\tau')}{\pi(t-\tau')}. \quad (3)$$

Ceci se traduit par une "interférence intersymbole" (ISI "inter symboles interference") et une détérioration du "diagramme de l'œil" qui peut engendrer des erreurs de transmission. Pour améliorer les performances du système de transmission, il faut compenser ces déformations. Cette opération est appelée "égalisation".

On peut remédier à ces défauts en transmettant une séquence connue du récepteur (séquence d'apprentissage). Cette séquence est constituée de données pseudo-aléatoires judicieusement choisies pour que la fonction d'autocorrélation $r(\tau)$ de cette séquence soit nulle lorsque $\tau \neq 0$.

Si on calcule la fonction d'intercorrélation entre le signal capté $x_0(t)$ lorsque cette séquence $m_0(t)$ a été émise, on obtient la réponse impulsionnelle du canal de transmission. Dans le cas où les caractéristiques du canal de transmission évoluent au cours du temps, on peut adapter les coefficients du filtre modélisant le canal en utilisant des algorithmes récursifs comme le gradient stochastique ou les moindres carrés récursifs. J'invite le lecteur à se reporter au polycopié "Estimation récursive des paramètres d'un modèle" sur la page www.essi.fr/~leroux, où on trouvera une présentation des méthodes de ce type.

Dans ces approches où on suppose que l'entrée est connue ou du moins correctement estimée, on cherche à minimiser la variance de l'écart entre le signal après égalisation $y(t)$ et le signal émis $m(t)$

$$E [|y(t) - m(t-t_0)|^2],$$

en acceptant un retard dans la reconstruction du signal transmis. Dans la plupart des formulations, ce retard sera omis, on suppose qu'on admet la réalisation de filtres non causaux, ce qui est équivalent.

Le signal capté par le récepteur est $x(t)$. On suppose que ce signal est le résultat de la convolution de $m(t)$ par la réponse impulsionnelle d'un filtre linéaire inconnu (le canal de transmission) $h(t)$. Le signal transformé par l'égaliseur s'en déduit par la convolution (figure 1)

$$y(t) = \sum_{k=-K}^K c(k)x(t-k). \quad (4)$$

Pour corriger les paramètres de l'égaliseur $c(k)$, on utilise souvent des algorithmes de type "gradient stochastique" où on modifie $c(k)$ à chaque échantillon de manière à réduire $|y(t) - m(t - t_0)|^2$, ce qui se traduit par

$$c_t(k) = c_{t-1}(k) - \Delta [c_{t-1}(k)] x(t-k) [y(t) - m(t - t_0)]. \quad (5)$$

Le pas du gradient $\Delta (c_{t-1}(k))$ peut prendre plusieurs formes plus ou moins complexes et efficaces.

Toutefois il peut être utile de pallier les défauts du système de transmission sans avoir à émettre un signal connu, ou du moins d'améliorer les performances de l'égalisateur en tenant compte de toutes les données reçues, aussi bien les données connues du récepteur que les données utiles. On dispose pour cela d'une information importante: le signal émis appartient à un alphabet fini connu du récepteur.

Si pour atteindre cet objectif on utilise les statistiques ou les informations spectrales d'ordre deux (autocorrélation ou densité spectrale du signal capté $x(t)$), on perd l'information de phase sur la réponse en fréquence du canal de transmission. Par exemple, en utilisant les techniques de prédiction linéaire, on peut "blanchir" le signal $x(t)$ (c'est à dire rendre constant le module de sa densité spectrale) mais on ne peut pas contrôler la compensation apportée sur la phase. Ainsi la prédiction linéaire (cf. "*techniques numériques pour le traitement du signal*" sur la page www.essi.fr/~leroux) permet de trouver un modèle à phase minimum, alors que le canal de transmission n'a aucune raison de présenter cette caractéristique. Toutefois il est utile de commencer l'égalisation par cette étape préalable de blanchiment; ceci permet de traiter indépendamment le module et la phase du modèle de la réponse en fréquence du canal de transmission.

Nous verrons que les méthodes employées pour effectuer l'égalisation se fondent implicitement sur l'utilisation des caractéristiques statistiques et spectrales d'ordre supérieur à deux. Ceci n'est possible que parce que le signal de communication présente des caractéristiques statistiques connues, les signaux émis appartiennent à un ensemble de formes fini (en particulier son moment d'ordre quatre n'est pas celui d'une variable aléatoire gaussienne, son kurtosis n'est pas nul).

Il existe toutefois une approche proposée par Tong, Xu et Kailath, qui permet de réaliser une égalisation en se fondant uniquement sur les statistiques d'ordre deux en utilisant le signal reçu par plusieurs capteurs, à la conditions que ces signaux présentent des différences suffisantes. Nous la décrirons succinctement dans le paragraphe 2.5.

1.2 Plan du cours

Nous étudierons dans le deuxième chapitre les approches les plus souvent citées (Lucky, Sato, Godard, Benvéniste) et considérées comme les travaux créateurs fondamentaux en égalisation aveugle, en détaillant plus particulièrement les travaux de Godard et de Benvéniste. Ensuite nous étudierons la formalisation fondée sur la maximisation du kurtosis proposée par Shalvi et Weinstein ce qui nous amènera à employer dans le troisième chapitre les outils développés dans le cadre de l'analyse statistique d'ordre supérieur à deux (HOS, higher order statistics) et les méthodes d'identification qui s'en déduisent comme celles de Giannakis et Mendel. Enfin, dans un quatrième chapitre nous donnerons l'interprétation spectrale de cette approche fondée sur les statistiques d'ordre supérieur à deux qui est une approche équivalente trouvant des applications en interférométrie astronomique (Lohmann *et al.*)

2 Les premières techniques de déconvolution aveugle

2.1 Le forçage à zéro des coefficients de la réponse impulsionnelle de Lucky ("zero forcing")

Une première proposition a été faite par R. Lucky des Bell Labs en 1966. Il cherche à minimiser la distortion du signal (ce qu'on peut interpréter comme une ouverture de l'œil dans le diagramme du même nom. Si la

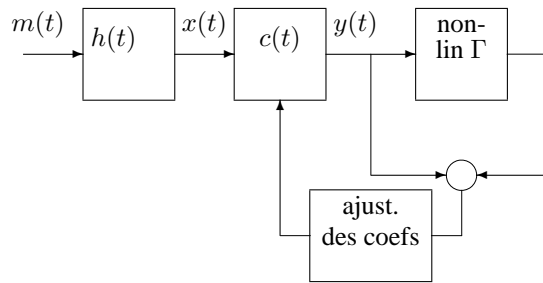


Figure 2: Schéma de l'égaliseur de Sato

réponse impulsionnelle du canal après égalisation est $b(t)$, on cherche à trouver les coefficients de l'égaliseur $c(t)$ pour minimiser

$$\frac{\sum_{t \neq 0} |b(t)|}{b(0)}. \quad (6)$$

Cette approche ne peut être envisagée en pratique car elle est extrêmement sensible au bruit.

2.2 Le critère de Sato

Le filtre égaliseur produit un signal $y(t)$ (figure 2)

$$y(t) = \sum_{k=-K}^K c(k)x(t-k) \quad (7)$$

On applique ensuite une non linéarité sans mémoire $\Gamma[y(t)]$, par exemple une saturation à $y(t)$ et on corrige les coefficients $C(k)$ de manière à réduire l'écart entre $y(t)$ et $\Gamma[y(t)]$. On peut interpréter ce critère de la manière suivante : on connaît les caractéristiques statistiques du signal émis : c'est par exemple un signal quantifié ; on ajuste les coefficients de l'égaliseur de manière à forcer le signal égalisé à prendre des niveaux quantifiés suivant cette loi connue. La non-linéarité a pour fonction implicite de forcer une caractéristique de la densité de probabilité de $y(t)$ de se rapprocher le plus possible de la même caractéristique portant sur $m(t)$. Dans le cas des signaux réels, Sato propose de minimiser

$$E(y^2(t)) - \frac{E(m^2(t))}{E(|m(t)|)} E(|y(t)|)$$

ou plus exactement de minimiser

$$E(y(t) - \gamma \text{signe}(y(t)))$$

où le coefficient γ est ajusté pour que l'amplitude du signal après égalisation soit identique à celle du signal émis. La minimisation se fait en appliquant un méthode du type "gradient stochastique" : on calcule à chaque itération le gradient du critère par rapport aux coefficients c_k de l'égaliseur. Le fait de faire apparaître une non-linéarité dans le calcul de l'erreur à minimiser fait intervenir implicitement les statistiques d'ordre supérieur à deux dans le processus d'égalisation. Ce critère sera ensuite précisé et justifié d'une part par Godard et d'autre part par Benveniste *et al.*

La minimisation d'un critère a en général pour conséquence une bonne robustesse des algorithmes d'estimation, toutefois la forme du critère ne garantit pas nécessairement l'unicité du minimum : on peut donc obtenir des solutions tout-à-fait erronées dans ce cas.

Remarques : Plusieurs articles font référence au travail de Bussgang [2] comme source historique de cette approche. Mais le lien entre l'article de Bussgang, qui traite des signaux gaussiens, et l'égalisation ne me semble pas direct. Plusieurs articles font une référence erronée au travail de Sato, quia publié deux articles différents dans le même numéro d'une revue !)

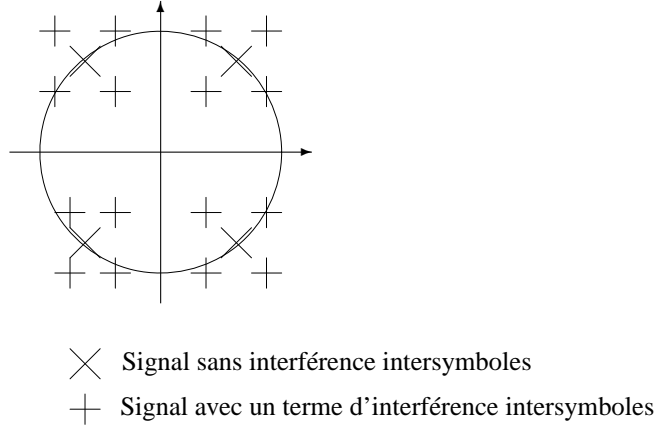


Figure 3: Effet d'une interférence intersymboles sur une constellation: le motif est "dilaté", ce qui se remarque en particulier sur l'amplitude qui est moins concentrée

2.3 Le forçage à module constant de Godard (CMA "constant modulus algorithm")

La séquence émise considérée comme un séquence i.i.d. est $m(t)$. Le signal capté par le récepteur est

$$x(t) = \sum_n m(t-n)h(n)e^{j\varphi(t)} + v(t), \quad (8)$$

où $h(t)$ est la réponse impulsionnelle du canal de transmission supposé linéaire, $e^{j\varphi(t)}$ est un terme aléatoire de fluctuation de phase et $v(t)$ un bruit de mesure. La sortie de l'égaliseur est

$$y(t) = \sum_k c(k)x(t-k), \quad (9)$$

ou bien en fonction de l'entrée $m(t)$ et du filtre complet (canal plus égaliseur, cf figure 1)

$$y(t) = \sum_k b(k)m(t-k). \quad (10)$$

Dans cette approche on suppose que les bruits de mesure sont négligeables; nous verrons, en utilisant l'analyse statistique aux ordres supérieurs, que les bruits de mesure gaussiens indépendants du signal émis n'ont pas d'effet sur les méthodes d'égalisation, du moins en théorie. Par construction, le signal émis $m(t)$ vérifie les deux hypothèses :

-

$$E(m^2(t)) = 0, \quad (11)$$

(le carré qui est complexe de $m(t)$ est une variable aléatoire de moyenne nulle).

- Lorsque $k \neq 0$

$$E(\overline{m(t)}m(t-k)) = 0. \quad (12)$$

$E(|m(t)|^2)$ est connu et indépendant de t ainsi que $E(|m(t)|^4)$. Une forme simplifiée du critère à minimiser proposé par Godard est

$$E\left(\left(|y(t)|^2 - r\right)^2\right). \quad (13)$$

(Godard considère le cas où l'exposant est un entier positif quelconque et non égal à 2).

S'il y a une interférence intersymboles, une constellation sera étalée (à variance constante); ce qui se traduit en termes de probabilité sur la modification des moments d'ordre supérieur à deux.

L'idée proposée par D. Godard est de réduire cet étalement en se fixant un gain, par exemple pour que les constellations émises et reconstituées correspondent à des signaux de même variance. Le choix de

ce gain sera discuté ultérieurement. Dans le cas d'une modulation de phase, on conçoit bien, d'après la figure 3, que le forçage à module constant de la sortie de l'égaliseur annulera nécessairement l'interférence intersymboles. Lorsque la modulation n'est pas une simple modulation de phase, le fait de rapprocher le plus possible le module de la sortie $y(t)$ d'un module constant aura sous certaine hypothèses un effet identique.

2.3.1 Analyse du critère de Godard

Pour étudier le critère (13-26), nous aurons besoin des valeurs de $E(|y(t)|^4)$ et $E(|y(t)|^2|m(t)|^2)$:

$$E(|y(t)|^4) = E\left(\left|\sum_k b(k)m(t-k)\right|^4\right), \quad (14)$$

soit

$$E(|y(t)|^4) = E\left(\left(\sum_k b(k)m(t-k)\overline{\sum_\ell b(\ell)m(t-\ell)}\right)^2\right), \quad (15)$$

ou encore

$$E(|y(t)|^4) = E\left(\left(\sum_k \sum_\ell b(k)\overline{b(\ell)}m(t-k)\overline{m(t-\ell)}\right)^2\right), \quad (16)$$

qui devient

$$E(|y(t)|^4) = E\left(\sum_k \sum_\ell \sum_{k'} \sum_{\ell'} b(k)\overline{b(\ell)}m(t-k)\overline{m(t-\ell)}b(k')\overline{b(\ell')}m(t-k')\overline{m(t-\ell')}\right), \quad (17)$$

et finalement

$$E(|y(t)|^4) = \sum_k \sum_\ell \sum_{k'} \sum_{\ell'} b(k)\overline{b(\ell)}b(k')\overline{b(\ell')}E\left(m(t-k)m(t-k')\overline{m(t-\ell)}\overline{m(t-\ell')}\right). \quad (18)$$

Dans le développement de cette somme, tous les termes qui apparaissent comme des produits de moments d'ordre impair sont nuls par construction de l'ensemble des valeurs possibles de $m(t)$. Il n'y a de termes non nuls que lorsque

- : $k = k', \ell = \ell', \ell \neq k$
- : $k = \ell, k' = \ell', k' \neq k$
- : $k = \ell', k' = \ell, k' \neq k$
- : $k = \ell = k' = \ell'$

Dans le premier cas, d'après l'hypothèse (11)

$$E\left(m(t-k)m(t-k')\overline{m(t-\ell)}\overline{m(t-\ell')}\right) = E(m(t-k)m(t-k'))E\left(\overline{m(t-\ell)}\overline{m(t-\ell')}\right) = 0. \quad (19)$$

Dans le deuxième et le troisième cas

$$E\left(m(t-k)m(t-k')\overline{m(t-\ell)}\overline{m(t-\ell')}\right) = E(|m(t-k)|^2)E(|m(t-k')|^2). \quad (20)$$

Dans le quatrième cas

$$E\left(m(t-k)m(t-k')\overline{m(t-\ell)}\overline{m(t-\ell')}\right) = E(|m(t-k)|^4) \quad (21)$$

On en déduit que

$$\begin{aligned} E(|y(t)|^4) &= E(|m(t)|^4) \sum_k |b(k)|^4 \\ &\quad + 2(E(|m(t)|^2))^2 \sum_k \sum_{k' \neq k} |b(k)|^2 |b(k')|^2, \end{aligned} \quad (22)$$

qu'on réécrit

$$\begin{aligned}
 E(|y(t)|^4) &= E(|m(t)|^4) \sum_k |b(k)|^4 \\
 &\quad + 2(E(|m(t)|^2))^2 \left(\sum_k |b(k)|^2 \right)^2 \\
 &\quad - 2(E(|m(t)|^2))^2 \sum_k |b(k)|^4,
 \end{aligned} \tag{23}$$

soit

$$\begin{aligned}
 E(|y(t)|^4) &= \left(E(|m(t)|^4) - 2(E(|m(t)|^2))^2 \right) \sum_k |b(k)|^4 \\
 &\quad + 2(E(|m(t)|^2))^2 \left(\sum_k |b(k)|^2 \right)^2.
 \end{aligned} \tag{24}$$

Par ailleurs

$$E(|y(t)|^2) = E(|m(t)|^2) \left(\sum_k |b(k)|^2 \right) \tag{25}$$

Le critère à minimiser (13) s'écrit

$$E\left((|y(t)|^2 - r)^2\right) = E(|y(t)|^4) - 2rE(|y(t)|^2) + r^2 \tag{26}$$

ou bien

$$\begin{aligned}
 E\left((|y(t)|^2 - r)^2\right) &= \left(E(|m(t)|^4) - 2(E(|m(t)|^2))^2 \right) \sum_k |b(k)|^4 \\
 &\quad + 2(E(|m(t)|^2))^2 \left(\sum_k |b(k)|^2 \right)^2 \\
 &\quad - 2rE(|m(t)|^2) \left(\sum_k |b(k)|^2 \right) + r^2
 \end{aligned} \tag{27}$$

On veut qu'il y ait un minimum pour

$$b(0) = 1, \tag{28}$$

$$k \neq 0 : b(k) = 0, \tag{29}$$

Ceci nous permet d'évaluer la constante r : En posant

$$\mu_k = |b(k)|^2, \tag{30}$$

on calcule $\frac{\partial E\left((|y(t)|^2 - r)^2\right)}{\partial \mu_k}$:

$$\begin{aligned}
 \frac{\partial E\left((|y(t)|^2 - r)^2\right)}{\partial \mu_k} &= \left(E(|m(t)|^4) - 2(E(|m(t)|^2))^2 \right) 2\mu_k \\
 &\quad + 2(E(|m(t)|^2))^2 2 \left(\sum_k \mu_k \right) \\
 &\quad - 2rE(|m(t)|^2)
 \end{aligned} \tag{31}$$

Une condition nécessaire d'annulation des dérivées partielles est donc que

$$r = \frac{E(|m(t)|^4)}{E(|m(t)|^2)}. \tag{32}$$

Pour vérifier qu'il y a un minimum du critère (13) lorsque l'égalisation est réalisée, c'est à dire pour

$$b(0) = 1, \quad (33)$$

$$k \neq 0 : b(k) = 0, \quad (34)$$

on remplace r par sa valeur (32) et on sépare dans $\sum_k |b(k)|^2$ et $\sum_k |b(k)|^4$ le terme fonction de $b(0)$ des autres

$$\begin{aligned} E \left((|y(t)|^2 - r)^2 \right) &= E (|m(t)|^4) (1 - |b(0)|^2)^2 \\ &\quad + E (|m(t)|^4) \left(\sum_{k \neq 0} |b(k)|^4 \right) \\ &\quad + 2 (E (|m(t)|^2))^2 \left(\left(\sum_{k \neq 0} |b(k)|^2 \right)^2 - \sum_{k \neq 0} |b(k)|^4 \right) \\ &\quad + \left(4 (E (|m(t)|^2))^2 |b(0)|^2 - 2E (|m(t)|^4) \right) \left(\sum_{k \neq 0} |b(k)|^2 \right) \\ &\quad + r^2 - E (|m(t)|^4) \end{aligned} \quad (35)$$

qu'on écrit en fonction des μ_k

$$\begin{aligned} E \left(\left(|y(t)|^2 - \frac{E (|m(t)|^4)}{E (|m(t)|^2)} \right)^2 \right) &= E (|m(t)|^4) (1 - \mu(0))^2 \\ &\quad + E (|m(t)|^4) \left(\sum_{k \neq 0} \mu^2(k) \right) \\ &\quad + 2 (E (|m(t)|^2))^2 \left(\left(\sum_{k \neq 0} \mu(k) \right)^2 - \sum_{k \neq 0} \mu^2(k) \right) \\ &\quad + \left(4 (E (|m(t)|^2))^2 \mu(0) - 2E (|m(t)|^4) \right) \left(\sum_{k \neq 0} \mu(k) \right) \\ &\quad + \left(\frac{E (|m(t)|^4)}{E (|m(t)|^2)} \right)^2 - E (|m(t)|^4) \end{aligned} \quad (36)$$

soit

$$\begin{aligned} E \left(\left(|y(t)|^2 - \frac{E (|m(t)|^4)}{E (|m(t)|^2)} \right)^2 \right) &= E (|m(t)|^4) (1 - \mu(0))^2 \\ &\quad + E (|m(t)|^4) \left(\sum_{k \neq 0} \mu(k)^2 \right) \\ &\quad + 2 (E (|m(t)|^2))^2 \left(\left(\sum_{k \neq 0} \mu(k) \right)^2 - \sum_k \mu(k)^2 \right) \\ &\quad + \left(4 (E (|m(t)|^2))^2 \mu(0) - 2E (|m(t)|^4) \right) \left(\sum_{k \neq 0} \mu(k) \right) \\ &\quad + \left(\frac{E (|m(t)|^4)}{E (|m(t)|^2)} \right)^2 - E (|m(t)|^4) \end{aligned} \quad (37)$$

Le point étudié correspond à

$$\begin{aligned}\mu(0) &= 1, & (38) \\ k \neq 0 : \mu(k) &= 0, & (39)\end{aligned}$$

Si on fait une petite variation $\varepsilon(k)$ (nécessairement positive) des $\mu(k)$, lorsque $k \neq 0$ et $\varepsilon(0)$ à partir de ce point, la variation du critère est, en ne conservant que les termes du premier ordre

$$\Delta = + \left(4 (E(|m(t)|^2))^2 - 2E(|m(t)|^4) \right) \left(\sum_{k \neq 0} \varepsilon(k) \right) \quad (40)$$

Ce terme est nécessairement positif à condition que

$$2 (E(|m(t)|^2))^2 - E(|m(t)|^4) > 0 \quad (41)$$

qui n'est autre qu'une condition sur le signe du "kurtosis" de $m(t)$, point qui sera repris dans l'étude de l'approche de Shalvi et Weinstein.

Nous verrons que le critère du module constant est très semblable au critère de maximisation du kurtosis de Shalvi et Weinstein. (Les calculs de Godard sont d'ailleurs similaires aux développements effectués dans le contexte des statistiques d'ordre supérieur bien qu'il n'y fasse pas référence. L'approche de Shalvi et Weinstein paraît plus directe et nous étudierons le problème fondamental de l'unicité de la solution en nous fondant sur leur travail.

2.4 La formulation de Benveniste *et al*

Benveniste *et al* partent de la remarque suivante, qui se rapproche de l'analyse en composantes indépendantes [4]: on considère deux variables aléatoires indépendantes m_1 et m_2 de même densité de probabilité $p(m)$ et une autre variable aléatoire y combinaison linéaire de m_1 et m_2

$$y = a_1 m_1 + a_2 m_2. \quad (42)$$

telle que

$$a_1^2 + a_2^2 = 1 \quad (43)$$

de sorte que y ait la même variance que x_1 et x_2 .

Dans l'application en égalisation, y sera la sortie de l'égaliseur ($y(t)$); m_1 et m_2 seront deux échantillons successifs du signal émis ($m(t)$ et $m(t-1)$). Le fait que $y(t)$ soit une combinaison linéaire de $m(t)$ et de $m(t-1)$ traduit une interférence intersymboles qu'on cherche à éliminer. L'élimination de l'interférence intersymboles implique que $y(t)$ doit avoir la même densité de probabilité p que m_1 et m_2 . On arrive ainsi à la question:

Dans quelle mesure y peut-elle avoir la même densité de probabilité p_m que m_1 et m_2 ?

La densité de probabilité de y est par construction égale à la convolution de la fonction $\frac{1}{a_1} p(\frac{y}{a_1})$ par $\frac{1}{a_2} p(\frac{y}{a_2})$ car les variables aléatoires $a_1 m_1$ et $a_2 m_2$ sont des variables indépendantes dont les densités de probabilités sont $\frac{1}{a_1} p(\frac{m_1}{a_1})$ et $\frac{1}{a_2} p(\frac{m_2}{a_2})$.

On peut écrire la fonction caractéristique $P_y(u)$ de la variable aléatoire y

$$P_y(u) = P_m(a_1 u) P_m(a_2 u) \quad (44)$$

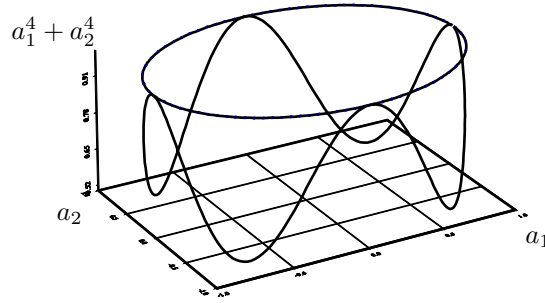
On veut que cette fonction caractéristique soit égale à $P_m(u)$

$$P_m(u) = P_m(a_1 u) P_m(a_2 u) \quad (45)$$

en respectant la contrainte (43).

En développant en série de type Taylor MacLaurin le logarithme de $P_y(u)$, on fait apparaître les moments d'ordre deux et quatre (nous supposons que la densité de probabilité $p(m)$ est symétrique et donc que sa moyenne et son moment d'ordre trois sont nuls)

$$\log(P_y(u)) = \frac{\sigma^2}{2!} (a_1^2 u^2 + a_2^2 u^2) + \frac{K}{4!} (a_1^4 u^4 + a_2^4 u^4) + \dots \quad (46)$$

Figure 4: Evolution de $a_1^4 + a_2^4$ en fonction de a_1 et a_2

Une condition nécessaire d'égalité est donc que

$$K(a_1^4 + a_2^4) = 0 \quad (47)$$

Le "cumulant" d'ordre 4, K est nul si m est une variable aléatoire gaussienne. Dans ce cas il est connu qu'une combinaison linéaire de variables gaussiennes est une variable gaussienne: on ne pourrait pas alors effectuer l'égalisation; ce qui est tout à fait cohérent avec le fait que les statistiques d'ordre deux ne permettent pas de retrouver l'information de phase sur un système linéaire, un signal gaussien étant complètement caractérisé par ses statistiques d'ordre deux. Nous écartons ce cas parce que $m(t)$ est une variable prenant un nombre fini de valeurs et n'est donc pas gaussienne.

La condition pour que l'identité (44) soit vérifiée est

$$a_1^4 + a_2^4 = 1 \quad (48)$$

Or cette condition n'est vérifiée que si

$$a_1^2 = 1, a_2 = 0 \quad (49)$$

ou

$$a_1 = 0, a_2^2 = 1 \quad (50)$$

ce qu'on peut vérifier en écrivant $a_1 = \cos \theta$ et $a_2 = \sin \theta$:

$$a_1^4 + a_2^4 = \frac{3 + \cos 4\theta}{4} \quad (51)$$

Il n'y a donc égalité que lorsque $y(t)$ n'est constitué que d'une seule des deux composantes $m_1(t)$ ou $m_2(t)$, c'est à dire lorsque l'égalisation est réalisée et donc que le filtre $b(k)$ est le filtre identité. On peut généraliser l'approche au cas des cumulants d'ordre quelconque. Toute méthode qui permet d'assurer que la densité de probabilité de la sortie est identique à la densité de probabilité de l'entrée garantira ainsi la reconstruction du signal original. En particulier le critère de Sato force les rapports de moments

$$\frac{E(y^2(t))}{E(|y(t)|)} \text{ et } \frac{E(m^2(t))}{E(|m(t)|)}$$

à être aussi proches que possible. Dans l'approche de Benveniste, Goursat et Ruget, on peut envisager différents critères de similarité entre les densités de probabilité, en particulier la similarité entre les moments d'ordre quatre comme le propose en fait Godard, et que nous retrouverons dans la formulation de Shalvi et Weinstein. Il est aussi envisageable de proposer des fonctions plus appropriées.

Analysons plus en détail le développement du logarithme de la fonction caractéristique $P_y''(u)$ (encore appelé seconde fonction caractéristique). Si on écrit $P_y''(u)$ dans le cas où il a seulement interférence entre deux symboles et lorsque la variance de $y(t)$ est égale à celle de $m(t)$ (ce qui se traduit par $a_1^2 + a_2^2 = 1$)

$$P_y''(u) = P_m''(a_1 u) + P_m''(a_2 u) \quad (52)$$

Quand l'égalisation est réalisée on a donc

$$P_m'(u) = P_m'(a_1 u) + P_m'(a_2 u) \quad (53)$$

En développant en fonction des cumulants (que nous définirons au paragraphe 3.1

$$\sigma_m^2 u^2 = \sigma_m^2 (a_1^2 + a_2^2) u^2 \quad (54)$$

$$cum_4 u^4 = cum_4 (a_1^4 + a_2^4) u^4 \quad (55)$$

$$cum_6 u^6 = cum_6 (a_1^6 + a_2^6) u^6 \quad (56)$$

$$\dots \quad (57)$$

identités qui ne peut être vérifiées dans le cas non gaussien que si un seul des coefficients a_1 et a_2 est non nul.

Si on trouve une méthode pour forcer l'un des cumulants de $y(t)$ à être égal à un des cumulants de $m(t)$, on aura bien réalisé l'égalisation. On peut aussi envisager de forcer l'égalité entre une fonction de ces différents cumulants.

Du point de vue de la robustesse des algorithmes (dans le cas des données bruitées) et la facilité d'implémentation (fondée en général sur une méthode de type gradient stochastique), il est préférable de définir un critère qui sera minimum lorsque l'égalité est atteinte. Une fois ce critère défini, et après s'être assuré qu'il ne présente pas de minima locaux. On peut chercher ainsi à minimiser une fonction de la forme

$$E(f(y)) = \int f(y) p_x(y) dy \quad (58)$$

la fonction $f(y)$ est normée et choisie pour atteindre son minimum lorsque la densité de probabilité de y est celle de x . Une fois qu'on s'est assuré de la validité de ce minimum, on cherche à chaque échantillon à réduire

$$\sum_t f(y(t)).$$

2.5 L'égalisation n'utilisant que les statistiques d'ordre deux de Tong *et al.*

Nous en donnons l'interprétation dans le domaine des fréquences: On suppose qu'on mesure les sorties $x_1(t)$ et $x_2(t)$ de deux canaux $H_1(z)$ et $H_2(z)$ soumis au même signal émis $m(t)$. On calcule les corrélations $R_{11}(z)$, $R_{12}(z)$ et $R_{22}(z)$

$$R_{11}(z) = E(X_1(z)X_1(z^{-1})) \quad (59)$$

$$R_{12}(z) = E(X_1(z)X_2(z^{-1})) \quad (60)$$

$$R_{22}(z) = E(X_2(z)X_2(z^{-1})) \quad (61)$$

On suppose que les deux canaux $H_1(z)$ et $H_2(z)$ sont des filtres à réponse impulsionnelle finie *qui n'ont pas de zéros communs*; si la séquence émise $m(t)$ est une séquence i.i.d., $R_{11}(z)$, $R_{12}(z)$ et $R_{22}(z)$ seront des polynômes (après une multiplication par un facteur de la forme z^{-p}).

On calcule les racines de $R_{12}(z)$. Soit z_0 une de ces racines. On cherche à savoir si z_0 est racine de $H_1(z)$ ou de $H_2(z^{-1})$. Dans le premier cas z_0 est nécessairement racine de $R_{11}(z)$. Dans le deuxième cas z_0 est nécessairement racine de $R_{22}(z)$. Le seul cas où on ne peut pas décider est celui où z_0 est à la fois racine de $R_{11}(z)$ et de $R_{22}(z)$. Il faut donc écarter ce cas correspondant à la situation où $H_1(z)$ et $H_2(z)$ ont des racines communes.

La limitation principale de cette approche comme de celle de Giannakis et Mendel que vous verrez au paragraphe 3.2 est la nécessité de la validité du modèle linéaire et la connaissance de l'ordre du modèle du canal. Tandis que les méthodes robustes se fondent sur la minimisation d'un critère.

3 Formulation en termes de statistiques d'ordre supérieur

Pour développer le critère qui est fondamental dans toutes ces approches, la maximisation du kurtosis sous contrainte, il est nécessaire de définir les outils développés pour l'étude des moments d'ordre supérieur à deux.

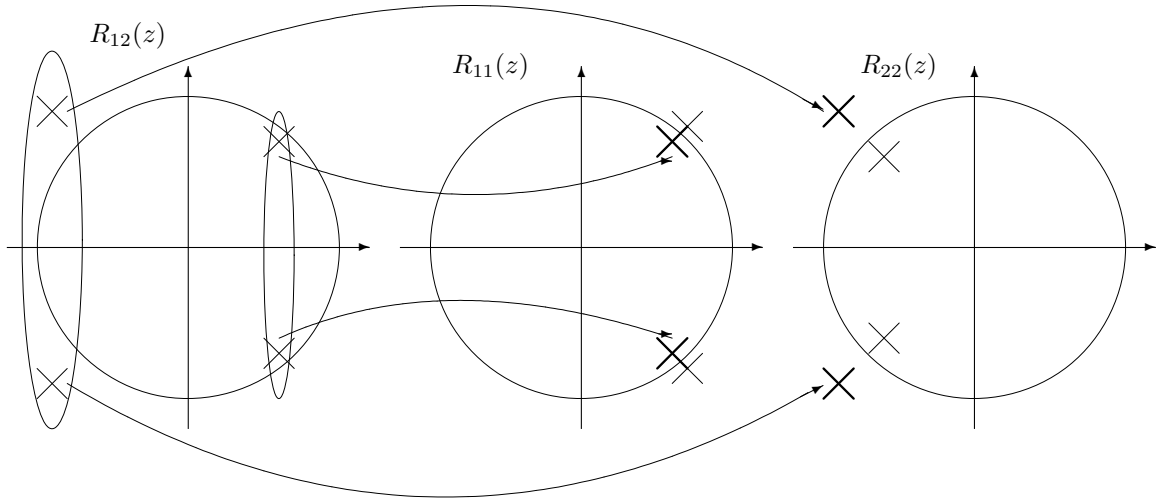


Figure 5: Choix des regroupements de racines dans l'identification conjointe de deux canaux de transmission qui ne présentent pas de racines communes

3.1 Formules de base, cumulants

3.1.1 Fonctions caractéristiques, moments et cumulants dans le cas d'une variable aléatoire scalaire

La (première) fonction caractéristique d'une variable aléatoire est la transformée de Fourier de sa densité de probabilité

$$P_x(u) = E(e^{jux}) = \int_{-\infty}^{\infty} p_x(x)e^{jux} dx \quad (62)$$

En calculant le développement en fonction des puissances de u

$$P_x(u) = E(e^{jux}) = \int_{-\infty}^{\infty} p_x(x) \left(1 + jux + \frac{j^2 u^2 x^2}{2!} + \frac{j^3 u^3 x^3}{3!} + \frac{j^4 u^4 x^4}{4!} + \dots \right) dx \quad (63)$$

$$P_x(u) = E(e^{jux}) = \int_{-\infty}^{\infty} p_x(x) dx + ju \int_{-\infty}^{\infty} p_x(x)x dx + \frac{j^2 u^2}{2!} \int_{-\infty}^{\infty} p_x(x)x^2 dx + \frac{j^3 u^3}{3!} \int_{-\infty}^{\infty} p_x(x)x^3 dx + \frac{j^4 u^4}{4!} \int_{-\infty}^{\infty} p_x(x)x^4 dx + \dots \quad (64)$$

$$P_x(u) = 1 + juE(x) + \frac{j^2 u^2}{2!} E(x^2) + \frac{j^3 u^3}{3!} E(x^3) + \frac{j^4 u^4}{4!} E(x^4) + \dots \quad (65)$$

Le moment d'ordre k se calcule à partir de la fonction caractéristique

$$E(x^k) = \frac{1}{j^k} \left[\frac{d^k P_x(u)}{du^k} \right]_{u=0} \quad (66)$$

On appelle seconde fonction caractéristique le logarithme (néperien) de la première fonction caractéristique

$$P_x''(u) = \log P_x(u) \quad (67)$$

Son développement en série de MacLaurin donne des coefficients qui ne sont plus les moments de x mais ses "cumulants" que nous noterons "cum" dans les formules.

$$cum_k(x) = \frac{1}{j^k} \frac{d^k P_x''(u)}{du^k} \quad (68)$$

Le cumulants d'ordre un est la moyenne; le cumulants d'ordre deux est la variance (moment d'ordre deux après soustraction du carré de la moyenne). Le cumulants d'ordre trois (skewness en anglais, dissymétrie (?) en français) est nul lorsque la variable aléatoire est centrée et sa densité de probabilité est symétrique. Nous nous intéresserons surtout au cumulants d'ordre quatre appelé "kurtosis" à partir du mot grec signifiant "épaule".

3.1.2 Fonctions caractéristiques, moments et cumulants dans le cas de vecteurs aléatoires

Il est nécessaire d'utiliser ces définitions assez lourdes pour traiter le cas des fonctions aléatoires.

Soit un vecteur \vec{x} de composantes (x_1, \dots, x_N) . La fonction caractéristique de ce vecteur sera une fonction de N variables, $\vec{u} = (u_1, \dots, u_N)^T$ calculée comme la transformée de Fourier multidimensionnelle de la densité de probabilité $p_{\vec{x}}(x_1, \dots, x_N)$

$$P_{\vec{x}}(u_1, \dots, u_N) = \int_{R^N} e^{j\vec{x}^T \vec{u}} p_{\vec{x}}(x_1, \dots, x_N) dx_1 \dots dx_N \quad (69)$$

Le développement de cette fonction fera apparaître les moments joints d'ordre k quelconque en particulier les moments joints d'ordre trois et quatre.

$$P_{\vec{x}}(u_1, \dots, u_N) = \sum_{k \geq 0} \frac{j^k}{n_1! \dots n_N!} E(x_1^{n_1} \dots x_N^{n_N}) u_1^{n_1} \dots u_N^{n_N} \quad (70)$$

où les n_1, \dots, n_N vérifient

$$n_1 + \dots + n_N = k \quad (71)$$

Le développement de la seconde fonction caractéristique (logarithme de la première)

$$P_{\vec{x}}''(u_1, \dots, u_N) = \log P_{\vec{x}}(u_1, \dots, u_N) \quad (72)$$

donne encore les cumulants joints

3.1.3 Relation entre cumulants et moments de v.a. réelles

En identifiant les développements de la seconde fonction caractéristique et du logarithme de la première, on obtient les relations entre cumulants et moments.

Ainsi le cumulants d'ordre un est égal au moment d'ordre un.

$$E(x(t_1)) = cum_1(x(t_1)) \quad (73)$$

$$E(x(t_1)x(t_2)) = cum_1(x(t_1))cum_1(x(t_2)) + cum_2(x(t_1)x(t_2)) \quad (74)$$

ou encore

$$cum_2(x(t_1)x(t_2)) = E(x(t_1)x(t_2)) - E(x(t_1))E(x(t_2)) \quad (75)$$

Le cumulants d'ordre deux est égal à la variance (après soustraction du carré de la moyenne au moment d'ordre deux.) Pour donner les formules à l'ordre trois nous supposons que le signal est à moyenne nulle. Dans ce cas

$$cum_3(x(t_1)x(t_2)x(t_3)) = E(x(t_1)x(t_2)x(t_3)) \quad (76)$$

Si $x(t)$ a une moyenne nulle, il y a identité entre moment et cumulants d'ordre trois, exactement comme à l'ordre deux; mais ce n'est pas le cas si $x(t)$ n'est pas de moyenne nulle. A l'ordre quatre nous donnerons la relation en supposant que les densités de probabilité sont centrées et symétriques, $p_{\vec{x}}(-\vec{x}) = p_{\vec{x}}(\vec{x})$ et donc que les moments et cumulants d'ordre trois sont nuls ce qui est le cas la plupart du temps dans les applications.

On a alors la relation

$$\begin{aligned} E(x(t_1)x(t_2)x(t_3)x(t_4)) &= cum_4(x(t_1)x(t_2)x(t_3)x(t_4)) \\ &+ cum_2(x(t_1)x(t_2))cum_2(x(t_3)x(t_4)) \\ &+ cum_2(x(t_1)x(t_3))cum_2(x(t_2)x(t_4)) \\ &+ cum_2(x(t_1)x(t_4))cum_2(x(t_2)x(t_3)) \end{aligned} \quad (77)$$

$$\begin{aligned}
E(x(t_1)x(t_2)x(t_3)x(t_4)) &= cum_4(x(t_1)x(t_2)x(t_3)x(t_4)) \\
&+ E(x(t_1)x(t_2)) E(x(t_3)x(t_4)) \\
&+ E(x(t_1)x(t_3)) E(x(t_2)x(t_4)) \\
&+ E(x(t_1)x(t_4)) E(x(t_2)x(t_3))
\end{aligned} \tag{78}$$

Un moment d'ordre k se déduit des moments d'ordre inférieur par addition d'un terme qui n'en dépend pas et n'est autre que le cumulants d'ordre k . En particulier, dans le cas des signaux réels

$$cum_4(x) = E(x^4) - 3 [E(x^2)]^2. \tag{79}$$

3.1.4 Quelques propriétés des moments et des cumulants

- Multiplication des variables par des constantes

$$E(\alpha_1 x_1, \dots, \alpha_N x_N) = \left(\prod_{i=1}^N \alpha_i \right) E(x_1, \dots, x_N) \tag{80}$$

(relation similaire pour les cumulants)

- Permutation des variables

Si on modifie l'ordre des variables (x_1, \dots, x_N) les moments et cumulants sont inchangés.

- Addition des variables

$$E(x_1 + y_1, \dots, x_N) = E(x_1, \dots, x_N) + E(y_1, \dots, x_N) \tag{81}$$

(relation similaire pour les cumulants)

- Invariance par translation d'une variable pour les cumulants d'ordre supérieur à un

$$cum_k(x_1 + c, \dots, x_N) = cum_k(x_1, \dots, x_N) \tag{82}$$

(cette relation n'est pas vérifiée pour les moments)

- Sommes de variables indépendantes

Si x_i et y_i sont deux variables aléatoires indépendantes

$$cum_k(x_1 + y_1, \dots, x_N + y_N) = cum_k(x_1, \dots, x_N) + cum_k(y_1, \dots, y_N) \tag{83}$$

(cette relation n'est pas vérifiée pour les moments)

- Indépendance des composantes

Si les premières composantes x_1, \dots, x_p sont indépendantes des dernières

$$cum_k(x_1, \dots, x_N) = cum_k(x_1, \dots, x_p) + cum_k(x_{p+1}, \dots, x_N) \tag{84}$$

(cette relation n'est pas vérifiée pour les moments)

Par conséquent, les cumulants des échantillons d'une séquence d'échantillons indépendants de moyenne nulle

$$x(t_1), \dots, x(t_N)$$

sont tous nuls sauf si

$$t_1 = t_2 = \dots = t_N \tag{85}$$

Un des intérêts majeurs des cumulants est la conséquence de l'indépendance qui simplifie considérablement les formules et les traitements.

En particulier, si deux signaux sont indépendants et que l'un d'entre eux est gaussien à moyenne nulle, les cumulants d'ordre supérieur à deux seront ceux du deuxième signal car les cumulants du signal gaussien sont nuls; on peut ainsi étudier les caractéristiques des signaux malgré la présence d'un bruit additif gaussien.

3.1.5 Cumulant d'ordre quatre pour des variables aléatoires complexes

Pour définir les moments présentant un intérêt, on fait intervenir les variables ainsi que leurs complexes conjuguées. Nous supposons, comme c'est souvent le cas des constellations complexes utilisées en transmission numérique que la variable aléatoire complexe z présente les symétries telles que

$$E(z) = 0 \quad (86)$$

$$E(z^2) = 0 \quad (87)$$

(cette deuxième hypothèse n'est pas toujours supposée vérifiée; bien sûr, elle n'implique pas que $E(|z|^2) = 0$.)

$$E(z^3) = 0 \quad (88)$$

$$E(z^2 \bar{z}) = 0 \quad (89)$$

Nous nous intéresserons plus particulièrement au cumulants

$$\begin{aligned} cum_4 \left(z(t_1)z(t_2)\overline{z(t_3)z(t_4)} \right) &= E \left(z(t_1)z(t_2)\overline{z(t_3)z(t_4)} \right) \\ &\quad - E(z(t_1)z(t_2)) E(\overline{z(t_3)z(t_4)}) \\ &\quad - E(z(t_1)\overline{z(t_3)}) E(z(t_2)\overline{z(t_4)}) \\ &\quad - E(z(t_1)\overline{z(t_4)}) E(z(t_2)\overline{z(t_3)}) \end{aligned} \quad (90)$$

les autres étant égaux à zéro pour plusieurs configurations intéressantes des constellations en transmission numérique. Notons que les moments du deuxième ordre apparaissant dans la formule (90) seront nuls la plupart du temps du fait des hypothèses sur z . Dans le cas particulier où les instants de mesure sont identiques

$$t_1 = t_2 = t_3 = t_4 = t \quad (91)$$

$$cum_4(|z(t)|^4) = E(|z(t)|^4) - 2E(|z(t)|^2)^2 - (|E(z^2)|)^2 \quad (92)$$

(Le dernier terme s'annule dans le cas de l'hypothèse (87); noter la différence avec le cas des signaux réels.)

On retrouve des formules similaires à celles du critère proposé par Godard; ou pourra ainsi établir un lien entre les algorithmes d'égalisation fondés sur le "module constant" et ceux fondés sur la maximisation du kurtosis que nous verrons dans le paragraphe 3.3.

3.2 Triple corrélation et filtrage linéaire, méthode d'identification de Giannakis et Mendel

A partir de ces cumulants, on peut définir les corrélations d'ordre supérieur pour une fonction du temps.

3.2.1 Corrélation d'ordre trois (triple corrélation)

Si le moment d'ordre un est nul, elle s'écrit dans le cas des signaux réels (elle n'est guère utilisée dans le cas complexe car les symétries des constellations impliquent souvent sa nullité)

$$E(x(t_1)x(t_2)x(t_3)) \quad (93)$$

Si $x(t)$ est stationnaire à l'ordre trois

$$r_x^{(3)}(\tau_1, \tau_2) = E(x(t)x(t+\tau_1)x(t+\tau_2)) \quad (94)$$

(la stationnarité à l'ordre trois est plus contraignante que la stationnarité à l'ordre deux: un signal peut être stationnaire à l'ordre deux sans l'être à l'ordre trois) C'est une fonction de deux variables. Nous allons voir que cette fonction appliquée à la sortie d'un filtre linéaire inconnu permet d'en retrouver les paramètres lorsque l'entrée du filtre est un bruit blanc ou une séquence i.i.d. sans faire référence au signal d'entrée à condition que la corrélation d'ordre trois existe. Nous verrons aussi dans le chapitre 4 que cette identification aveugle peut aussi se réaliser dans le domaine des fréquences.

3.2.2 Filtrage d'une séquence i.i.d.

Un signal i.i.d. $x(t)$ a une triple corrélation qui est nulle partout sauf à l'origine du plan où elle est égale au facteur d'asymétrie ("skewness") $cum_3(x)$.

Si $x(t)$ est filtré par une filtre numérique non récursif de réponse impulsionnelle $b(t)$, produisant la sortie $y(t)$,

$$y(t) = \sum_k b(k)x(t-k) \quad (95)$$

la corrélation d'ordre trois de $y(t)$ est

$$r_3^{(y)}(\tau_1, \tau_2) = E \left(\sum_{k, k', k''} x(t-k)b(k)x(t-k'+\tau_1)b(k')x(t-k''+\tau_2)b(k'') \right) \quad (96)$$

en réarrangeant les termes

$$r_3^{(y)}(\tau_1, \tau_2) = \sum_{k, k', k''} b(k)b(k')b(k'')E(x(t-k)x(t-k'+\tau_1)x(t-k''+\tau_2)) \quad (97)$$

$x(t)$ étant une séquence i.i.d., les seuls termes de la somme qui ne sont pas nuls vérifient

$$t-k = t-k'+\tau_1 = t-k''+\tau_2 \quad (98)$$

soit

$$k' = k - \tau_1 \quad (99)$$

$$k'' = k - \tau_2 \quad (100)$$

ce qui se traduit par

$$r_3^{(y)}(\tau_1, \tau_2) = cum_3(x) \sum_k b(k)b(k-\tau_1)b(k-\tau_2) \quad (101)$$

La sommation est une génération d'un calcul de corrélation classique : ici on fait intervenir un double produit et trois valeurs de la fonction, au lieu du produit de deux valeurs dans la corrélation d'ordre deux.

3.2.3 Support de la corrélation d'ordre trois d'une séquence i.i.d. filtrée par un filtre à réponse impulsionnelle finie

La fonction de transfert est

$$B(z) = \sum_{k=0}^K b(k)z^{-k} \quad (102)$$

D'après l'équation (101)

$$r_3^{(y)}(\tau_1, \tau_2) = cum_3(x) \sum_k b(k)b(k-\tau_1)b(k-\tau_2) \quad (103)$$

Lorsque k est fixé, le domaine de variation du couple de variables (τ_1, τ_2) est le carré (figure 6)

$$k-K \leq \tau_1 \leq k \quad (104)$$

$$k-K \leq \tau_2 \leq k \quad (105)$$

Lorsque k varie de 0 à K chaque point de ce carré se translate le long d'un segment dont les extrémités ont pour coordonnées

$$\begin{aligned} &(\tau_1(0), \tau_2(0)) \text{ dans le carré } (-K, 0) \times (-K, 0) \\ &(\tau_1(0) + K, \tau_2(0) + K) \text{ dans le carré } (0, K) \times (0, K) \end{aligned}$$

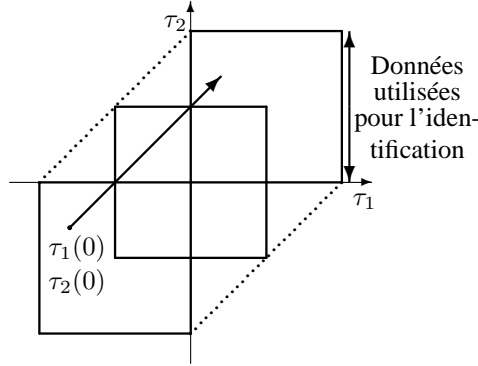


Figure 6: Support hexagonal de la corrélation d'ordre trois d'un signal iid filtré par un filtre à réponse impulsionnelle finie. Les données utilisées dans la méthode de Giannakis sont sur la lisière du domaine.

Tableau 1: Valeur des coefficients de la corrélation d'ordre trois du résultat du filtrage d'une séquence i.i.d. par le filtre $b(z) = 1 + bz^{-1}$

0	0	0	0	0
0	0	b	b^2	0
0	b^2	$b^3 + 1$	b	0
0	b	b^2	0	0
0	0	0	0	0

Ainsi lorsque k varie le carré glisse le long de la première diagonale et le support de la triple corrélation est un hexagone représenté sur la figure 6. La corrélation d'ordre trois présente plusieurs symétries en particulier le changement de signe simultané sur les deux variables τ_1 et τ_2 , l'échange de ces variables τ_1 et τ_2 , ainsi que la transformation

$$\nu_1 = -\tau_1 \quad (106)$$

$$\nu_2 = \tau_2 - \tau_1 \quad (107)$$

Un exemple de corrélation d'ordre trois pour le filtre non récursif

$$B(z) = 1 + bz^{-1} \quad (108)$$

est donné dans le tableau 1

3.2.4 Identification de filtre à non minimum de phase (Giannakis et Mendel)

La corrélation d'ordre deux permet de retrouver les caractéristiques d'un filtre à minimum de phase par factorisation spectrale (filtre à réponse impulsionnelle finie) ou d'un filtre récursif stable par résolution des équations de Yule-Walker (algorithme de Levinson). On ne peut pas à partir de ces informations séparer les racines d'un filtre qui seraient soit à l'intérieur soit à l'extérieur du cercle de rayon un. On voit que cette identification est possible lorsqu'on se base sur la corrélation d'ordre trois.

Si on fixe $\tau_2 = K$ l'équation (101) reliant la corrélation d'ordre trois aux coefficients du filtre devient

$$r_3^{(y)}(\tau_1, K) = cum_3(x) \sum_k b(k)b(k - \tau_1)b(k - K) \quad (109)$$

Comme $(k - K)$ ne doit pas être négatif pour que $b(k - K)$ ne soit pas nul, la somme se réduit à un seul terme non nul ($k = K$)

$$r_3^{(y)}(\tau_1, K) = cum_3(x)b(0)b(K)b(K - \tau_1) \quad (110)$$

Les échantillons situés sur la lisière du support de la triple corrélation (figure 6) permettent à un coefficient multiplicatif près de retrouver la réponse impulsionnelle du filtre.

Toutefois cette méthode est peu robuste (il faut une excellente estimation de la triple corrélation, être assuré de l'ordre du modèle, etc. . . . Il en existe de nombreuses variantes.

On remarque que si la densité de probabilité de $x(t)$ est symétrique, le moment d'ordre trois est nul et cette identification n'est pas possible; il faudra alors étendre l'approche à l'ordre quatre, à condition que le cumulants d'ordre quatre de $x(t)$ ne soit pas nul. Dans le cas des signaux d'entrée gaussiens tous les cumulants d'ordre supérieur à deux sont nuls et il n'est pas possible d'effectuer cette identification de filtres non minimum de phase.

3.3 La maximisation du kurtosis de Shalvi et Weinstein

Kurtosis positif et kurtosis négatif Dans les applications en égalisation on s'intéresse surtout au cumulants d'ordre quatre, et il peut être important (par exemple dans les hypothèses de Godard et Benveniste) de connaître le signe du kurtosis du signal étudié (le signal émis et le signal reçu ont des kurtosis de même signe s'ils ne sont pas perturbés par des bruits non gaussiens).

On parle parfois de variable "surgaussienne" quand le kurtosis est positif (loi de Cauchy par exemple) et de variable "sousgaussienne" quand le kurtosis est négatif (variable équirépartie).

Si une densité de probabilité est bornée dans $[-A, A]$ centrée et de variance fixée, le kurtosis le plus petit (négatif) est obtenu lorsque

$$p(x) = \frac{1}{2}\delta(x - A) + \frac{1}{2}\delta(x + A) \quad (111)$$

pour une variance égale à A^2

Si la densité de probabilité est bornée dans $[-A, A]$ centrée et de variance fixée égale à σ^2 (on suppose que $A^2 > 3\sigma^2$), le kurtosis le plus grand(positif) est obtenu lorsque

$$p(x) = (1 - \alpha)\delta(x) + \frac{\alpha}{2}\delta(x - A) + \frac{\alpha}{2}\delta(x + A) \quad (112)$$

pour une variance fixée σ^2 et où α est égal à $\frac{\sigma^2}{2A^2}$

3.3.1 Corrélation d'ordre quatre des signaux complexes

Nous supposons que les moments d'ordre un et trois sont nuls. Le cumulants d'ordre quatre s'écrit dans le cas des signaux stationnaires

$$\begin{aligned} cum_4^{(x)}(\tau_1, \tau_2, \tau_3) &= E \left(x(t)x(t + \tau_1)\overline{x(t + \tau_2)x(t + \tau_3)} \right) \\ &\quad - E(x(t)x(t + \tau_1)) E(\overline{x(t + \tau_2)x(t + \tau_3)}) \\ &\quad - E(x(t)\overline{x(t + \tau_2)}) E(x(t + \tau_1)\overline{x(t + \tau_3)}) \\ &\quad - E(x(t)\overline{x(t + \tau_3)}) E(x(t + \tau_1)\overline{x(t + \tau_2)}) \end{aligned} \quad (113)$$

en particulier, si $\tau_1 = \tau_2 = \tau_3 = 0$

$$cum_4^{(x)}(0, 0, 0) = E(|x(t)|^4) - |E(x^2(t))|^2 - 2E(|x(t)|^2)^2 \quad (114)$$

Si, pour des raisons de symétrie de la densité de probabilité de $x(t)$, le terme $E(x^2(t))$ s'annule, on retrouve les données utilisées par Godard dans le critère du module constant.

3.3.2 Maximisation du kurtosis

Le point de départ de l'approche de Benveniste est d'essayer de filtrer le signal mesuré $x(t)$ pour que certaines caractéristiques de la densité de probabilité de $y(t)$ se rapprochent le plus possibles des mêmes caractéristiques concernant le signal émis $m(t)$; ce que font aussi Sato et Godard. Shalvi et Weinstein vont un peu plus loin; ils montrent qu'il suffit de maximiser (en valeur absolue) le kurtosis de la sortie pour trouver les paramètres de l'égaliseur à un retard et éventuellement à un déphasage près.

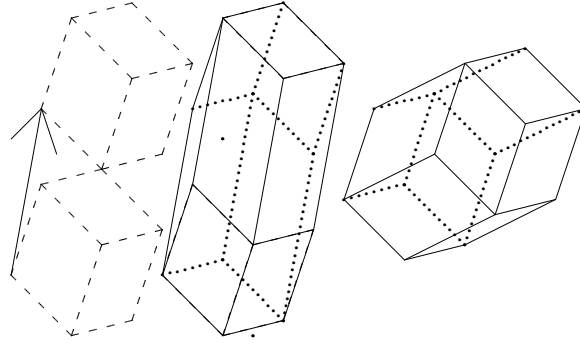


Figure 7: Support dodécaédrique de la corrélation d'ordre quatre d'un signal iid filtré par un filtre à réponse impulsionnelle finie. Ce support est le volume obtenu en faisant glisser un cube dont les côtés de longueur $(K + 1)$ sont parallèles aux axes et dont un des sommets est situé à l'origine de l'espace. Le glissement se fait suivant le vecteur $((K + 1), (K + 1), (K + 1))$ de sorte qu'à la fin du glissement le sommet opposé au premier se retrouve à l'origine.

Quelques formules préalables Nous reprenons pour préparer les calculs quelques formules semblables à celles de Godard (en supposant toujours que $m(t)$ est une séquence i.i.d.)

$$E(|y(t)|^2) = E\left(\left|\sum_k m(t-k)b(k)\right|^2\right) \quad (115)$$

$$E(|y(t)|^2) = E(|m(t)|^2) \sum_k |b(k)|^2 \quad (116)$$

Si on ne suppose pas *a priori* que $E(m^2(t))$ est nul, on a aussi

$$E(y(t)^2) = E(m(t)^2) \sum_k b(k)^2 \quad (117)$$

D'après les propriétés (90)

$$\begin{aligned} E(|y(t)|^4) &= E(|m(t)|^4) \left(\sum_k |b(k)|^4\right) \\ &+ 2(E(|m(t)|^2))^2 \left(\left(\sum_k |b(k)|^2\right)^2 - \sum_k |b(k)|^4\right) \\ &+ |E(|m(t)|^2)|^2 \left(\left|\sum_k |b(k)|^2\right|^2 - \sum_k |b(k)|^4\right) \end{aligned} \quad (118)$$

Le kurtosis $cum_4(y)$ s'écrivant

$$K(y) = cum_4(y) = E(|y(t)|^4) - 2(E(|y(t)|^2))^2 - |E(y^2(t))|^2 \quad (119)$$

on fait disparaître les termes d'ordre deux et on obtient d'après les équations (116) et (117)

$$K(y) = K(m) \left(\sum_k |b(k)|^4\right) \quad (120)$$

Ce résultat n'est valable que sous l'hypothèse selon laquelle $m(t)$ est une séquence i.i.d. Cette formule implique que le kurtosis ne change pas de signe lorsqu'un signal est filtré par un signal linéaire. Cette formule permet de mettre en évidence un résultat semblable à celui de Benveniste:

Si les modules des signaux $m(t)$ et $y(t)$ ont la même variance

$$E(|y(t)|^2) = E(|m(t)|^2) \quad (121)$$

on a le résultat fondamental:

- (a) -

$$|K(y)| \leq |K(m)| \quad (122)$$

$K(y)$ ayant toujours le même signe que $K(m)$

- (b) -

$$|K(y)| = |K(m)| \quad (123)$$

seulement si $B(z)$ se réduit à un retard d'un nombre entier d'échantillons à en déphasage près.

En effet

$$\left(\sum_k |b(k)|^2 \right)^2 = \left(\sum_k |b(k)|^4 \right) + \left(\sum_k \sum_{\ell \neq k} |b(k)|^2 |b(\ell)|^2 \right) \quad (124)$$

Le deuxième terme de la somme est toujours positif ou nul. Il n'est nul que lorsqu'un seul des $|b(k)|^2$ est différent de zéro. D'où le résultat.

Tout filtrage d'un signal i.i.d. tel que le signal de sortie ait la même variance que le signal d'entrée réduira donc le module du kurtosis du signal de sortie; il le conservera uniquement s'il réalise un simple décalage (d'un nombre entier d'échantillons).

Retrouver le signal original revient donc à trouver un filtre tel que la séquence de sortie ait la même variance et le kurtosis maximum en valeur absolue.

Il s'agit donc de retrouver les coefficients de l'égaliseur maximisant $|K(y)|$ sous la contrainte

$$\sum_k |b(k)|^2 = 1 \quad (125)$$

(imposer la contrainte reviendra en pratique à appliquer un gain à $y(t)$ de manière à ce que la variance de $|y(t)|$ soit égale à celle de $|m(t)|$.)

Il faut toutefois s'assurer que cette recherche de maximum n'aboutit pas sur des maxima locaux (problème traité de manière moins simple dans les approches de Godard et Benveniste).

Pour recherche ce maximum on utilise la méthode des multiplicateurs de Lagrange

$$L(B, \lambda) = \sum_k |b(k)|^4 - \lambda \left(\sum_k |b(k)|^2 - 1 \right) \quad (126)$$

Les points stationnaires de $L(B, \lambda)$ vérifient

$$\frac{\partial L(B, \lambda)}{\partial b(\ell)} = 4b(\ell) \left(|b(\ell)|^2 - \frac{\lambda}{2} \right) \quad (127)$$

pour tout ℓ et

$$\frac{\partial L(B, \lambda)}{\partial \lambda} = - \sum_k |b(k)|^2 + 1 = 0 \quad (128)$$

D'après l'équation (127), il faut que

soit

$$b(\ell) = 0 \quad (129)$$

soit

$$|b(\ell)|^2 - \frac{\lambda}{2} = 0. \quad (130)$$

Supposons donc qu'il y ait N coefficients $b(k)$ non nuls (et donc tous égaux en valeur absolue à $\sqrt{\lambda/2}$). D'après la contrainte (128), il faut que

soit

$$|b(k)|^2 = 1/N \quad (131)$$

soit

$$b(k) = 0 \quad (132)$$

On est amené à envisager trois cas :

- (a) Le cas où tous les $|b(k)|$ sont égaux et valent donc $\frac{1}{K}$ (en supposant qu'on ne considère que le cas où la chaîne complète : filtre + égaliseur est un filtre à réponse impulsionnelle finie, ce qui n'est pas très réaliste en théorie et n'est d'ailleurs pas envisagé dans l'article de Shalvi et Weinstein;

- (b) Le cas général où $N \neq K$ et $N \neq 1$

- (c) Le cas où un seul des $b(k)$ est différent de zéro et vaut donc un;

Shalvi et Weinstein montrent que les solutions sont des points selles (ou cols) dans le cas (b) et que le maximum n'est obtenu que dans le cas (c), le cas (a) donnant un minimum.

Pour cela on étudie les petites variations réelles $\varepsilon(\ell)$ des composantes $|b(k)|^2$ telles que

$$\sum_{\ell} \varepsilon(\ell) = 0 \quad (133)$$

- **Cas (a) Il y a un minimum lorsque tous les $b(k)$ sont différents de zéro**

Dans ce cas, $N = K$. Lorsqu'on ajoute une petite variation aux K composantes $b(k)$ le critère (126) devient

$$L(B + \varepsilon, \lambda) = \sum_k |b(k)|^4 + \sum_{\ell} \varepsilon^2(\ell) \quad (134)$$

La valeur du critère au point d'équilibre ($\varepsilon(\ell) = 0$) augmente quelquesoit la variation vérifiant la contrainte (133), et le point ne peut être qu'un minimum.

- **Cas (b) Il y a un col (point-selle) lorsque le nombre N de $b(k)$ différents de zéro est inférieur à K et supérieur à un**

- (b1) Lorsqu'on ajoute au critère (126) une petite variation non nulle seulement pour les N composantes non nulles $b(k)$ et nulle pour les $(K - N)$ composantes nulles, on obtient

$$L(B + \varepsilon, \lambda) = \sum_k |b(k)|^4 + \sum_{\ell} \varepsilon^2(\ell) \quad (135)$$

La valeur du critère au point d'équilibre ($\varepsilon(\ell) = 0$) est dépassée, et le point ne peut être qu'un minimum ou un point selle.

- (b2) Considérons le cas où au moins une des composantes $b(k)$, soit $b(0)$, est nulle.

Les petites variations $\varepsilon(k)$ sont nécessairement positives lorsque $B(k) = 0$ car un carré de module ne peut qu'augmenter.

Le critère est dans ce cas pour une petite variation vérifiant la contrainte (133)

$$L(B + \varepsilon, \lambda) = \varepsilon^2(0) + \sum_{k \neq 0} (|b(k)|^2 + \varepsilon(k))^2 \quad (136)$$

En remplaçant les $b(k)$ par leurs valeurs

$$L(B + \varepsilon, \lambda) = \varepsilon^2(0) + \sum_{k \neq 0} \left(\frac{1}{N} + \varepsilon(k) \right)^2 \quad (137)$$

En développant le carré

$$L(B + \varepsilon, \lambda) = \varepsilon^2(0) + \sum_{k \neq 0} \left(\frac{1}{N^2} + 2\frac{1}{N}\varepsilon(k) + \varepsilon^2(k) \right)^2 \quad (138)$$

D'après la contrainte (133)

$$\sum_{k \neq 0} \varepsilon(k) = -\varepsilon(0) \quad (139)$$

et le critère devient

$$L(B + \varepsilon, \lambda) = \varepsilon^2(0) + \frac{1}{N} - 2\frac{1}{N}\varepsilon(0) + \sum_{k \neq 0} \varepsilon^2(k) \quad (140)$$

$$L(B + \varepsilon, \lambda) = \frac{1}{N} + \varepsilon^2(0) + \sum_{k \neq 0} \varepsilon^2(k) - 2\frac{1}{N}\varepsilon(0) \quad (141)$$

En négligeant les termes du deuxième ordre

$$L(B + \varepsilon, \lambda) = \frac{1}{N} - 2\frac{1}{N}\varepsilon(0) \quad (142)$$

une variation positive $\varepsilon(0)$ entraîne dans ce cas une diminution du critère.

• (b3) Ainsi, en effectuant une petite variation dans une direction (premier cas considéré), on augmente le critère (135); en effectuant une petite variation dans une autre direction on diminue ce critère (142); par conséquent, le point analysé est un point selle et non un point d'équilibre pour la recherche d'un maximum.

• **Cas (c) : Il y a un maximum lorsque un seul des $b(k)$ différent de zéro** Dans ce cas ce $b(k)$ vaut un. On suppose que c'est le premier ($k = 0$), mais on peut appliquer l'approche à n'importe lequel des $b(k)$ (on rappelle que la solution n'est recherchée qu'à une translation temporelle près).

Les petites variations $\varepsilon(k)$ sont nécessairement positives lorsque $b(k) = 0$ c'est à dire lorsque $k \neq 0$ et donc la variation $\varepsilon(0)$ est négative d'après la contrainte (133).

Le critère est dans ce cas pour une petite variation vérifiant la contrainte (133)

$$L(B + \varepsilon, \lambda) = \sum_{k \neq 0} \varepsilon(k) + (|b(0)|^2 + \varepsilon(0))^2 \quad (143)$$

D'après la contrainte (133)

$$\sum_{k \neq 0} \varepsilon(k) = -\varepsilon(0) \quad (144)$$

Tous les $\varepsilon(\ell)$ sont positifs lorsque $\ell \neq 0$, par conséquent, d'après la contrainte (133)

$$\varepsilon(0) < 0 \quad (145)$$

Le critère devient

$$L(B + \varepsilon, \lambda) = -\varepsilon(0) + 1 + 2\varepsilon(0) + \varepsilon^2(0) \quad (146)$$

$$L(B + \varepsilon, \lambda) = 1 + \varepsilon(0) + \varepsilon^2(0) \quad (147)$$

En négligeant les termes du deuxième ordre

$$L(B + \varepsilon, \lambda) = 1 + \varepsilon(0) \quad (148)$$

une variation (nécessairement négative) $\varepsilon(0)$ associée à n'importe quelle variation (nécessairement positive) des autres $\varepsilon(\ell)$ entraîne donc une diminution du critère : il s'agit bien d'un maximum.

Remarques : Le critère ne présente qu'un seul minimum, au signe près ou dans le cas complexe à un déphasage près: on peut multiplier $B(z)$ un facteur $e^{j\varphi}$ dans le développement précédent sans rien changer au résultat; la constellation du signal après égalisation peut subir une rotation quelconque qu'il faudra compenser en introduisant une connaissance connue *a priori* du récepteur sur le message émis.

Shalvi et Weinstein supposent que le filtre complet (canal + égaliseur) a une longueur infinie, et ne considèrent donc pas le cas du minimum absolu, obtenu lorsque tous les $|b(k)|^2$ sont identiques et égaux à $1/K$; pour qu'il y ait un point selle il faut qu'au moins une des composantes $b(k)$ soit nulle. Sinon on se trouve toujours dans le cas (135): il y a augmentation du critère quelquesoit la variation $\varepsilon(k)$: si toutes les composantes sont non nulles et donc égales à $1/K$ le kurtosis est minimum (on trouve le filtre de longueur k qui mélange le plus les $m(t)$).

Les considérations sur la recherche du maximum ont été faites en référence au canal complet ($B(z)$) et non de l'égaliseur $C(z)$. Le passage de $B(z)$ à $C(z)$ est immédiat: on passe de l'un à l'autre par une transformation linéaire inversible

$$B(z) = H(z)C(z) \quad (149)$$

et on peut effectuer la recherche du minimum en faisant varier les coefficients $c(k)$ et non les coefficients $b(k)$.

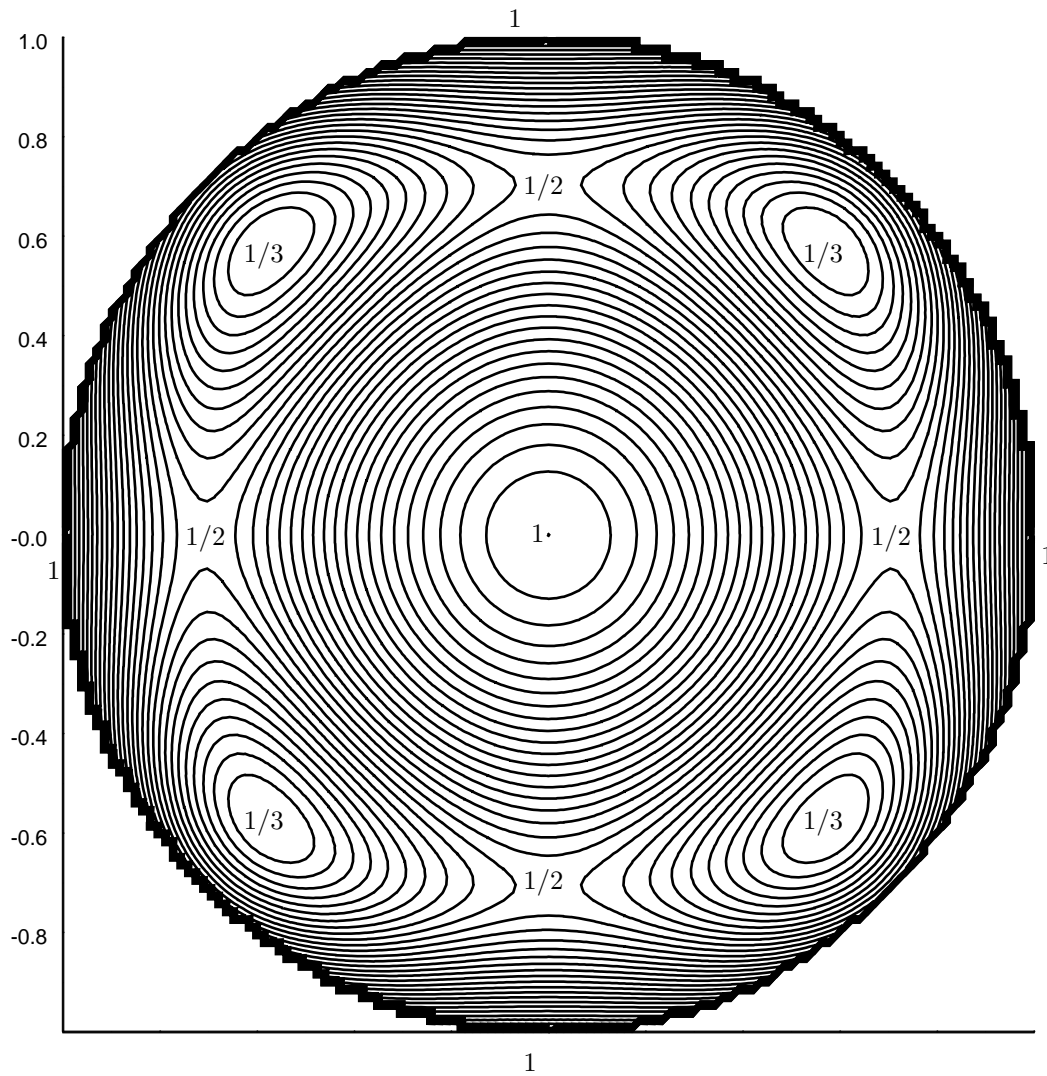


Figure 8: Courbes de niveau de la fonction $x^4 + y^4 + z^4$ sur la sphère $x^2 + y^2 + z^2 = 1$; Les coordonnées des six maxima sont $(\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)$, celles des huit minima sont $\frac{1}{\sqrt{3}}(\pm 1, \pm 1, \pm 1)$ celles des douze points selle sont $\frac{1}{\sqrt{2}}(0, \pm 1, \pm 1), \frac{1}{\sqrt{2}}(\pm 1, 0, \pm 1), \frac{1}{\sqrt{2}}(\pm 1, \pm 1, 0)$

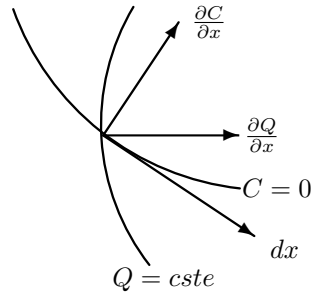


Figure 9: Projection du gradient sur la tangente à la contrainte

3.3.3 Minimisation d'un critère en respectant une contrainte

Si on cherche à minimiser un critère Q par la méthode du gradient en respectant une contrainte $C = 0$, on peut procéder de la manière suivante :

On cherche à diminuer $Q + \lambda C$. Le gradient est $\frac{\partial Q}{\partial x} + \lambda \frac{\partial C}{\partial x}$. Dans cette expression, l'écriture condensée $\frac{\partial Q}{\partial x}$ signifie dérivation par rapport à chacune des variables x_i composant le vecteur x . On calcule une première correction qui est tangente à la contrainte $C = 0$ à partir d'un point courant qui vérifie aussi cette contrainte (voir figure 9)

$$C(x(n)) = 0 \quad (150)$$

$$C(x(n) + dx) = 0 \quad (151)$$

en première approximation

$$C(x(n)) \frac{\partial C}{\partial x} = 0 \quad (152)$$

On veut que la variation dx soit de la forme

$$dx = \frac{\partial Q}{\partial x} + \lambda \frac{\partial C}{\partial x} \quad (153)$$

et vérifie la condition d'orthogonalité

$$\langle dx, \frac{\partial C}{\partial x} \rangle = 0 \quad (154)$$

$$\langle \frac{\partial Q}{\partial x} + \lambda \frac{\partial C}{\partial x}, \frac{\partial C}{\partial x} \rangle = 0 \quad (155)$$

$$\lambda = - \frac{\langle \frac{\partial Q}{\partial x}, \frac{\partial C}{\partial x} \rangle}{\langle \frac{\partial C}{\partial x}, \frac{\partial C}{\partial x} \rangle} \quad (156)$$

La correction suivant la direction du gradient est donc dans la direction

$$dx = \frac{\partial Q}{\partial x} - \frac{\partial C}{\partial x} \frac{\langle \frac{\partial Q}{\partial x}, \frac{\partial C}{\partial x} \rangle}{\langle \frac{\partial C}{\partial x}, \frac{\partial C}{\partial x} \rangle} \quad (157)$$

Il faut cependant que le résultat de la correction vérifie la contrainte

$$C(x(n+1)) = 0 \quad (158)$$

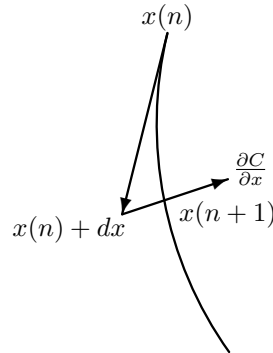


Figure 10: Deuxième correction forçant la vérification de la contrainte

Si la contrainte est assez régulière, on peut effectuer une approximation de projection, c'est à dire trouver l'amplitude (scalaire) μ telle que la correction prenne la forme (voir figure 10).

$$x(n+1) = x(n) + dx + \mu \frac{\partial C}{\partial x} \quad (159)$$

et vérifie

$$C(x(n) + dx + \mu \frac{\partial C}{\partial x}) = 0 \quad (160)$$

A chaque étape, il est prudent de vérifier *a posteriori* que le critère a effectivement été réduit par cette correction.

4 Interprétation spectrale, spectres d'ordre trois (bispectre) et d'ordre quatre (trispectre)

L'interprétation dans le domaine des fréquences des statistiques d'ordre supérieur est très utile: beaucoup de résultats s'énoncent simplement dans ce domaine et on rejoint un champ d'application important de l'identification (et de la déconvolution) aveugles en traitement d'images, la suppression d'effets de miroitement en observation astronomique.

4.1 Formules

Nous éludons les problèmes théoriques liés à l'existence et la formalisation correcte du problème de la représentation fréquentielle et nous invitons le lecteur intéressé par les développements théoriques à se référer aux ouvrages de A. Blanc-Lapierre et R. Fortet, "Théorie des fonctions aléatoires", publié aux éditions Masson en 1953 et de B. Picinbono, "Signaux aléatoires", publié aux éditions Dunod en 1994 ou encore les présentations de Brillinger, *An introduction to polyspectra*, in the annals of mathematical statistics, vol. 36, 1965, pp 1351-1374, et de Brillinger et Rosenblatt, *computation and interpretation of k-th order spectra*, Wiley, 1967, pp 189-232. En ce qui concerne le développement présenté ici nous supposons que les corrélations d'ordre trois et quatre des signaux étudiés possèdent des transformées de Fourier et en particulier, nous supposons la stationnarité à l'ordre trois et quatre des signaux échantillonnés. Cette hypothèse est, nous le verrons plus contraignante que la stationnarité à l'ordre deux.

4.1.1 Stationnarité dans le cas des signaux à temps continu

La stationnarité se traduit par une invariance par translation

$$\text{cum}_k(x(t_1), \dots, x(t_k)) = \text{cum}_k(x(t_1 + \tau), \dots, x(t_k + \tau)) \quad (161)$$

Ceci implique que la transformée de Fourier $\text{Cum}_k(X(\omega_1, \dots, \omega_k))$ est nulle en dehors des sous-espace

$$\omega_1 + \dots + \omega_k = 0 \quad (162)$$

dans le cas des signaux à temps continu et

$$\omega_1 + \dots + \omega_k = 0 \pmod{2\pi} \quad (163)$$

dans le cas des signaux à temps discret.

En général, on traduit cette condition de support en réduisant l'espace des variables à un espace de dimension inférieure :

dimension deux à l'ordre trois en posant

$$\omega_3 = -\omega_1 - \omega_2 \quad (164)$$

et dimension trois à l'ordre quatre en posant

$$\omega_4 = -\omega_1 - \omega_2 - \omega_3 \quad (165)$$

Toutefois il peut être préférable de revenir à la formulation en fonction de toutes les variables pour mieux étudier certains problèmes comme les symétries des spectres d'ordre supérieur comme dans la thèse de C. Huet [10].

4.1.2 Spectre d'ordre trois ou bispectre

La transformée de Fourier d'une corrélation d'ordre trois est souvent appelée bispectre (mais pour éviter les confusions des terminologies dans le domaine temporel et fréquentiel il est peut-être préférable d'utiliser les termes de spectre d'ordre trois).

Si la corrélation d'ordre trois est notée

$$r_3^{(x)}(\tau_1, \tau_2) = E(x(t)x(t + \tau_1)x(t + \tau_2)) \quad (166)$$

le spectre d'ordre trois est, dans le cas des signaux échantillonnés

$$R_3^{(x)}(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} r_3^{(x)}(\tau_1, \tau_2) \exp -j(\omega_1\tau_1 + \omega_2\tau_2) \quad (167)$$

La transformation inverse est

$$r_3^{(x)} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} R_3^{(x)}(\omega_1, \omega_2) \exp -j(\omega_1\tau_1 + \omega_2\tau_2) d\omega_1 d\omega_2 \quad (168)$$

On peut calculer directement ce spectre d'ordre trois à partir des transformées de Fourier de différentes réalisations

$$R_3^{(x)}(\omega_1, \omega_2) = E(X(\omega_1)X(\omega_2)X(-\omega_1 - \omega_2)) \quad (169)$$

Si on suppose que le support de $X(\omega)$ est borné entre $-\pi$ et π , le support est un hexagone représenté sur la figure 11 : le facteur $X(\omega_1)$ a un support borné, ce qui réduit le support de $R_3^{(x)}(\omega_1, \omega_2)$ à une bande verticale

$$-\pi < \omega_1 < \pi; \quad (170)$$

de même le facteur $X(\omega_2)$ a un support borné

$$-\pi < \omega_2 < \pi; \quad (171)$$

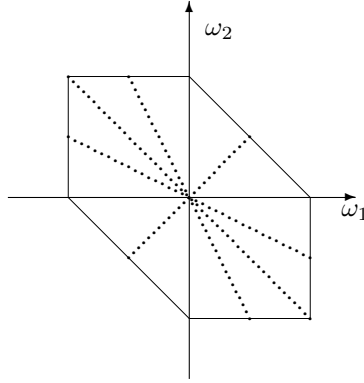


Figure 11: Support hexagonal du spectre d'ordre trois d'un signal dont la bande de fréquence est limitée; les axes de symétrie sont les lignes pointillées

ce qui réduit le support de $R_3^{(x)}(\omega_1, \omega_2)$ à l'intersection de cette bande verticale avec une bande horizontale; finalement le facteur $X(-\omega_1 - \omega_2)$ a pour support une bande diagonale,

$$-\pi < \omega_2 + \omega_1 < \pi; \quad (172)$$

Le support de $R_3^{(x)}(\omega_1, \omega_2)$ est l'hexagone intersection de ces trois supports, représenté sur la figure 11.

D'après la formule (169), dans le cas des signaux réels, ce spectre d'ordre trois présente six axes de symétrie. On peut aussi interpréter ce support à partir de la formule (188)

$$R_3^{(x)}(\omega_1, \omega_2) = E(X(\omega_1)X(\omega_2)X(\omega_3)) \quad (173)$$

où les trois fréquences ω_1, ω_2 et ω_3 vérifient

$$\omega_1 + \omega_2 + \omega_3 = 0 \quad (174)$$

On voit ainsi que le support du spectre d'ordre trois est l'intersection d'un cube centré à l'origine et de côté 2π avec le plan (174), ce qui permet de voir plus facilement les symétries.

On trouvera plusieurs résultats concernant les propriétés parfois peu connues des spectres d'ordre trois dans la thèse de Delphine Rossille [19].

4.1.3 Spectre d'ordre quatre ou trispectre

Dans le cas du spectre d'ordre quatre nous nous intéressons aux signaux complexes aussi bien que réels. Dans le cas des signaux complexes, il faut choisir les données pour lesquelles on utilise la complexe conjugué. La formulation la plus utile de la corrélation d'ordre quatre est celle qui permettra de calculer $E(|x|^4)$ donnée dans la définition (114)

$$\begin{aligned} cum_4^{(x)}(\tau_1, \tau_2, \tau_3, \tau_4) &= E\left(x(t + \tau_1)x(t + \tau_2)\overline{x(t + \tau_3)x(t + \tau_4)}\right) \\ &\quad - E\left(x(t + \tau_1)x(t + \tau_2)\right) E\left(\overline{x(t + \tau_3)x(t + \tau_4)}\right) \\ &\quad - E\left(x(t + \tau_1)\overline{x(t + \tau_3)}\right) E\left(x(t + \tau_2)\overline{x(t + \tau_4)}\right) \\ &\quad - E\left(x(t + \tau_1)\overline{x(t + \tau_4)}\right) E\left(x(t + \tau_2)\overline{x(t + \tau_3)}\right) \end{aligned} \quad (175)$$

invariante par translation de la forme (t, t, t, t) La transformée de Fourier appartient donc à

$$\omega_1 + \omega_2 + \omega_3 + \omega_4 = 0 \quad (176)$$

Elle s'écrit sous la forme

$$\begin{aligned} Cum_4^{(x)}(\omega_1, \omega_2, \omega_3, \omega_4) = & E \left(X(\omega_1) X(\omega_2) \overline{X(-\omega_3)} \overline{X(-\omega_4)} \right) \\ & - E \left(X(\omega_1) X(\omega_2) \right) E \left(\overline{X(-\omega_3)} \overline{X(-\omega_4)} \right) \\ & - E \left(X(\omega_1) \overline{X(-\omega_3)} \right) E \left(X(\omega_2) \overline{X(-\omega_4)} \right) \\ & - E \left(X(\omega_1) \overline{X(-\omega_4)} \right) E \left(X(\omega_2) \overline{X(-\omega_3)} \right) \end{aligned} \quad (177)$$

Les termes du deuxième ordre vérifient la stationnarité à l'ordre deux (qui est impliquée par la stationnarité à l'ordre quatre); donc le premier terme soustrait est nul en dehors du support

$$\omega_1 + \omega_2 = 0 \text{ et } \omega_3 + \omega_4 = 0 \quad (178)$$

le deuxième terme est nul en dehors du support

$$\omega_1 + \omega_3 = 0 \text{ et } \omega_2 + \omega_4 = 0 \quad (179)$$

et le troisième en dehors du support

$$\omega_1 + \omega_4 = 0 \text{ et } \omega_2 + \omega_3 = 0 \quad (180)$$

Comme ω_4 peut être éliminé, le spectre d'ordre 4 (178) s'écrit

$$\begin{aligned} Cum_4^{(x)}(\omega_1, \omega_2, \omega_3, -\omega_1 - \omega_2 - \omega_3) = & E \left(X(\omega_1) X(\omega_2) \overline{X(-\omega_3)} \overline{X(\omega_1 + \omega_2 + \omega_3)} \right) \\ & - E \left(X^2(\omega_1) \right) E \left(\overline{X^2(-\omega_3)} \right) \delta(\omega_1 + \omega_2) \\ & - E \left(|X(\omega_1)|^2 \right) E \left(|X(\omega_2)|^2 \right) \delta(\omega_1 + \omega_3) \\ & - E \left(|X(\omega_1)|^2 \right) E \left(|X(\omega_2)|^2 \right) \delta(\omega_2 + \omega_3) \end{aligned} \quad (181)$$

Soustraction des termes spectraux d'ordre deux En pratique on commence par calculer des moments pour en déduire des cumulants. Les termes spectraux du deuxième ordre apparaissent ainsi comme des données qu'il faut retrancher du spectre des moments d'ordre quatre pour obtenir les spectres de cumulants d'ordre quatre. Ces termes d'ordre deux sont non nuls sur les plans

$$\omega_1 + \omega_2 = 0 \quad (182)$$

$$\omega_2 + \omega_3 = 0 \quad (183)$$

$$\omega_3 + \omega_1 = 0 \quad (184)$$

Le premier des trois termes à retrancher est différent dans le cas réel et dans le cas complexe: dans ce dernier cas ce terme

$$E \left(X^2(\omega_1) \right) E \left(\overline{X^2(-\omega_3)} \right) \delta(\omega_1 + \omega_2)$$

est un nombre complexe qui ne se ramène pas à un calcul de densité spectrale. Ce terme peut être nul pour certaines symétries du de la densité de probabilité du signal d'entrée.

La soustraction de ces termes du deuxième ordre, dont le support est formé de sous espaces de dimension deux est délicate: on doit soustraire des fonction qui ont sur ce sous espace une amplitude bien plus élevée que le spectre d'ordre quatre et le résultat peut être biaisé. De plus les conséquences implicites de l'utilisation de fenêtres d'analyse spectrale peut conduire à ce que les termes d'ordre deux soient aussi visibles dans le voisinage des plans (184). Une analyse détaillée de ce problème se trouve dans la thèse de Cécile Huet [10].

Symétries du spectre d'ordre quatre Dans le cas des signaux à bande limitée (à temps continu), le spectre d'ordre quatre est une fonction de trois variable non nulle à l'intérieur d'un support octaédrique représenté sur la figure 12. Exactement comme le support du spectre d'ordre trois est obtenu en coupant deux des quatre sommets d'un carré, le support du spectre d'ordre quatre est obtenu en coupant deux

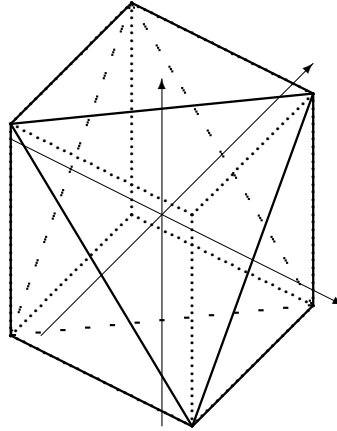


Figure 12: Support octaédrique du spectre d'ordre quatre d'un signal dont la bande de fréquence est limitée

sommets d'un cube, ce qui forme un octaèdre. L'analyse des symétries du spectre d'ordre quatre se fait plus simplement si on considère l'expression

$$R_4^{(x)}(\omega_1, \omega_2, \omega_3, \omega_4) = E \left(X(\omega_1)X(\omega_2)\overline{X(\omega_3)}\overline{X(\omega_4)} \right) - E \left(X(\omega_1)\overline{X(\omega_3)} \right) E \left(X(\omega_2)\overline{X(\omega_4)} \right) \quad (185)$$

$$- E \left(X(\omega_1)\overline{X(\omega_4)} \right) E \left(X(\omega_2)\overline{X(\omega_3)} \right) \quad (186)$$

$$- E \left(X(\omega_1)X(\omega_2) \right) E \left(\overline{X(\omega_3)}\overline{X(\omega_4)} \right) \quad (187)$$

$$(188)$$

en la restreignant à l'hyperplan

$$\omega_1 + \omega_2 + \omega_3 + \omega_4 = 0 \quad (189)$$

Une analyse détaillée de ces symétries se trouve aussi dans la thèse de Cécile Huet [10].

4.2 Spectres d'ordre supérieur, stationnarité et échantillonnage

Lorsqu'on traite des signaux échantillonnés, la notion de stationnarité n'est pas en général la même que celle utilisés pour le temps continu: on fait l'hypothèse qu'il y a stationnarité uniquement aux instants d'échantillonnage et non entre les instants d'échantillonnage. Un signal échantillonné n'est pas un signal stationnaire car c'est une suite d'impulsions régulièrement espacées et modulées en amplitude, le signal étant nul entre les instants d'échantillonnage. Le support des spectres d'ordre deux et trois des signaux échantillonnés ne se réduisent pas à l'hexagone de la figure ni à l'octaèdre de la figure. Ce sont un carré contenant l'hexagone et un cube contenant l'octaèdre. Il n'y a identité entre les spectres d'ordre trois du signal à temps continu et du même signal après échantillonnage que si la fréquence maximale du signal est égale au tiers de la fréquence d'échantillonnage. On peut montrer ce résultat graphiquement en rappelant que la transformée de Fourier du signal échantillonné est la périodisation (modulo 2π) de la transformée de Fourier du signal à temps continu. Dans ce cas, le support de

$$E \left(X(\omega_1)X(\omega_2)X(-\omega_1 - \omega_2) \right)$$

est obtenu par intersection de bandes horizontales, verticales et obliques correspondant aux supports périodiques des spectres des signaux échantillonnés (figure 13): L'hexagone est complété par deux triangles qui disparaissent si le support spectral est dans la bande $] - 2\pi/3, 2\pi/3[$; si la bande de fréquence du signal à temps continu est $] - \pi, \pi[$, les deux triangles et l'hexagone se rejoignent et le support est le carré complet.

De même pour qu'il y ait identité entre les spectres d'ordre quatre du signal à temps continu et du signal échantillonné, il faut que la fréquence maximale du signal soit inférieure au quart de la fréquence d'échantillonnage.

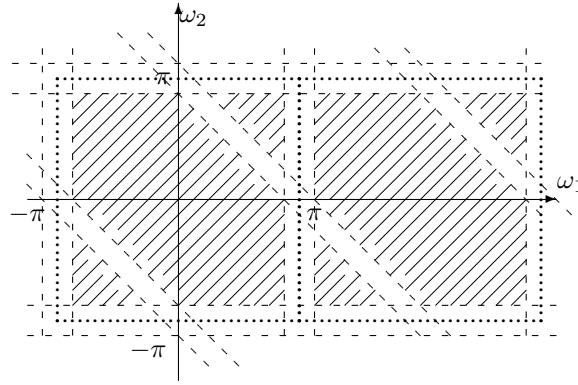


Figure 13: Support du spectre d'ordre trois d'un signal dont la bande de fréquence est limitée

Une analyse détaillée des conséquences de cette remarque sur l'échantillonnage se trouve dans la thèse de Delphine Rossille [19].

4.3 Spectres d'ordre supérieur et filtrage

Nous considérerons essentiellement le cas du spectre d'ordre trois dont nous étendrons les résultats au cas des spectres d'ordre quatre réels ou complexes.

Pour que ces résultats s'appliquent, il est essentiel que la séquence filtrée $y(t)$ soit un bruit blanc dans le cas du temps continu ou une séquence i.i.d. dans le cas du temps discret.

4.3.1 Spectre d'ordre trois et filtrage

Soit une séquence i.i.d. $x(t)$ stationnaire et centrée filtrée par un filtre $B(z)$

$$y(t) = \sum_n b(n)x(t-n) \quad (190)$$

La corrélation d'ordre trois de $y(t)$ est (voir l'équation (101))

$$r_3^{(y)}(\tau_1, \tau_2) = cum_3(x) \sum_k b(k)b(k-\tau_1)b(k-\tau_2) \quad (191)$$

Sa transformée de Fourier bidimensionnelle est

$$R_3(\omega_1, \omega_2) = \sum_{\tau_1} \sum_{\tau_2} cum_3(x) \sum_k b(k)b(k-\tau_1)b(k-\tau_2) \exp -j(\omega_1\tau_1 + \omega_2\tau_2) \quad (192)$$

En effectuant le changement de variable

$$\tau_1 - k = \tau'_1 \quad (193)$$

$$\tau_2 - k = \tau'_2 \quad (194)$$

L'équation (192) devient

$$R_3(\omega_1, \omega_2) = \sum_{\tau'_1} \sum_{\tau'_2} cum_3(x) \sum_k b(k)b(-\tau'_1)b(-\tau'_2) \exp -j(\omega_1\tau'_1 + \omega_2\tau'_2 + (\omega_1 + \omega_2)k) \quad (195)$$

$$R_3(\omega_1, \omega_2) = cum_3(x) \left(\sum_k b(k) \exp -j(\omega_1 + \omega_2)k \right) \left(\sum_{\tau'_1} b(-\tau'_1) \exp -j\omega_1\tau'_1 \right) \left(\sum_{\tau'_2} b(-\tau'_2) \exp -j\omega_2\tau'_2 \right) \quad (196)$$

$$R_3(\omega_1, \omega_2) = \text{cum}_3(x) B(\exp j\omega_1) B(\exp j\omega_2) B(\exp -j(\omega_1 + \omega_2)) \quad (197)$$

Le spectre d'ordre trois est obtenu en multipliant l'asymétrie de l'entrée par trois facteurs qui sont la réponse en fréquence du filtre calculée pour les fréquences ω_1 , ω_2 et $-\omega_1 - \omega_2$.

4.3.2 Obtention de la triple corrélation par transformée de Fourier inverse du spectre d'ordre trois

Nous avons vu que le spectre d'ordre trois est la transformée de Fourier de la triple corrélation. Inversement, on peut retrouver la triple corrélation à partir du spectre d'ordre trois.

Nous nous placerons dans le cas des signaux déterministes qu'on peut généraliser, sous les hypothèses convenables au cas des signaux aléatoires (cas des séquences i.i.d. transformées par un filtre linéaire invariant dans le temps).

Soit un bispectre vérifiant

$$R(\omega_1, \omega_2) = F(\omega_1)F(\omega_2)F(-\omega_1 - \omega_2). \quad (198)$$

C'est une fonction de deux variables. Sa transformée de Fourier inverse est

$$r(\tau_1, \tau_2) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F(\omega_1)F(\omega_2)F(-\omega_1 - \omega_2) \exp(j\omega_1\tau_1 + j\omega_2\tau_2) d\omega_1 d\omega_2 \quad (199)$$

ou encore

$$r(\tau_1, \tau_2) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega_1) \exp(j\omega_1\tau_1) \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega_2)F(-\omega_1 - \omega_2) \exp(j\omega_2\tau_2) d\omega_2 \right) d\omega_1 \quad (200)$$

La transformée de Fourier inverse du produit de deux fonctions $F(\omega)G(\omega)$ est une convolution

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega)G(\omega) \exp(j\omega t) d\omega = \sum_{\tau=-\infty}^{\infty} f(\tau)g(t - \tau) \quad (201)$$

Si on choisit pour variable ω_2 à la place de ω et la fonction

$$G(\omega_2) = F(-\omega_1 - \omega_2), \quad (202)$$

on obtient

$$g(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(-\omega_1 - \omega_2) \exp(j\omega_2 t) d\omega_2 \quad (203)$$

soit, en effectuant le changement de variable

$$\omega_2 + \omega_1 = \omega \quad (204)$$

$$g(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(-\omega) \exp[j(\omega - \omega_1)t] d\omega \quad (205)$$

ce qui donne une translation ω_1 dans le domaine des fréquences (ω) ou une modulation dans le domaine temporel (t)

$$g(t) = f(-t) \exp(-j\omega_1 t) \quad (206)$$

On peut reporter cette valeur de $g(t)$ dans la convolution transformée de Fourier inverse du produit (201)

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega_2)F(-\omega_1 - \omega_2) \exp(j\omega_2\tau_2) d\omega_2 = \sum_{\tau=-\infty}^{\infty} f(\tau)f(\tau - \tau_2) \exp j(\omega_1\tau - \omega_1\tau_2) \quad (207)$$

$r(\tau_1, \tau_2)$ s'écrit alors

$$r(\tau_1, \tau_2) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega_1) \exp(j\omega_1\tau_1) \left(\sum_{\tau=-\infty}^{\infty} f(\tau)f(\tau - \tau_2) \exp j(\omega_1\tau - \omega_1\tau_2) \right) d\omega_1 \quad (208)$$

En commutation les opérateurs de sommation

$$r(\tau_1, \tau_2) = \sum_{\tau=-\infty}^{\infty} f(\tau)f(\tau - \tau_2) \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega_1) \exp j(\omega_1\tau_1 + \omega_1\tau - \omega_1\tau_2) d\omega_1 \right) \quad (209)$$

c'est à dire, en exprimant la transformée de Fourier inverse

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(\omega_1) \exp j(\omega_1 \tau_1 + \omega_1 \tau - \omega_1 \tau_2) d\omega_1 = f(\tau_1 + \tau - \tau_2) \quad (210)$$

$$r(\tau_1, \tau_2) = \sum_{\tau=-\infty}^{\infty} f(\tau) f(\tau_1 + \tau - \tau_2) f(\tau - \tau_2) d\tau \quad (211)$$

En posant

$$\tau - \tau_2 = \tau' \quad (212)$$

cette expression devient

$$r(\tau_1, \tau_2) = \sum_{\tau=-\infty}^{\infty} f(\tau') f(\tau' + \tau_1) f(\tau' + \tau_2) d\tau \quad (213)$$

On retrouve bien la triple corrélation d'un signal déterministe.

4.3.3 Résultat du filtrage sur le spectre d'ordre quatre d'une séquence i.i.d.

A l'ordre quatre, dans le cas des signaux réels, on obtient l'expression

$$R_4(\omega_1, \omega_2, \omega_3) = cum_4(x) B(\exp j\omega_1) B(\exp j\omega_1) B(\exp j\omega_1) B(\exp -j(\omega_1 + \omega_2 + \omega_3)) \quad (214)$$

Dans le cas des signaux complexes, si on calcule le cumulante $cum_4(x(t)\overline{x(t+\tau_1)}\overline{x(t+\tau_2)}\overline{x(t+\tau_3)})$ le spectre d'ordre quatre sera

$$R_4(\omega_1, \omega_2, \omega_3) = cum_4(x) B(\exp j\omega_1) \overline{B(\exp j\omega_2)} \overline{B(\exp j\omega_3)} \overline{B(\exp -j(\omega_1 + \omega_2 + \omega_3))} \quad (215)$$

4.3.4 Validation des hypothèses sur la génération du signal étudié

Nous avons insisté à plusieurs reprises sur l'importance de l'hypothèse selon laquelle le signal étudié est obtenu en appliquant un filtre linéaire invariant dans le temps à une séquence i.i.d. stationnaire; on peut vérifier que le spectre d'ordre supérieur est compatible avec cette hypothèse en s'assurant que l'on a pour tout ω_1 et tout ω_2 et pour un ω non nul :

$$\frac{R_3(\omega_1, \omega) R_3(\omega_1 + \omega, \omega_2)}{R_2(\omega_1 + \omega)} = \frac{R_3(\omega_1, \omega_2) R_3(\omega_1 + \omega_2, \omega)}{R_2(\omega_1 + \omega_2)}. \quad (216)$$

où $R_2(\omega)$ est le spectre d'ordre deux; (il faut aussi que les symétries du spectre d'ordre trois soient vérifiées, ce qui est en général le cas par construction.) Cette formule est donnée ici à l'ordre trois. On peut généraliser à un ordre quelconque. Ce point est détaillé dans l'article [12].

4.4 Effet d'un retard sur les spectres d'ordre supérieur des signaux échantillonnés

Nous avons vu que le repliement spectral dû à l'échantillonnage fait apparaître des informations supplémentaires dans le spectre d'ordre trois ou d'ordre quatre du signal échantillonné. Cette remarque peut avoir des conséquences intéressantes pour traiter les problèmes de retard d'une durée inférieure à la période d'échantillonnage

Invariance par translation des spectres d'ordre supérieur des signaux à temps continu Exactement comme le spectre d'ordre deux d'un signal déterministe ou aléatoire stationnaire est invariant par translation, les spectres d'ordre supérieur sont invariants par translation du signal dans le domaine temporel ou par un terme de déphasage linéaire dans le domaine des fréquences. Cette formulation simple permet de développer des algorithmes d'identification et d'égalisation dans le domaine des fréquences.

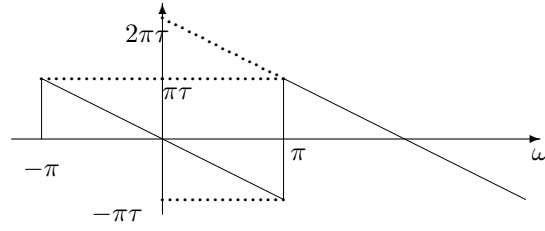


Figure 14: Effet de déphasage linéaire correspondant à un retard inférieur à la période d'échantillonnage

Invariance par translation d'un nombre entier d'échantillons dans le cas du temps discret Le même résultat d'invariance par translation s'applique si la translation est un nombre entier d'échantillons. Cette invariance se voit directement sur le spectre d'ordre trois, le résultat se généralise directement aux spectres d'ordre supérieur à trois:

Si on ajoute un retard $\exp -j\omega\tau$ à la fonction de transfert $B(z)$, τ étant un nombre entier d'échantillons, l'équation (197) devient

$$R_3(\omega_1, \omega_2) = \quad (217)$$

$$\text{cum}_3(x)B(\exp j\omega_1) \exp -j\omega_1\tau B(\exp j\omega_2) \exp -j\omega_1\tau B(\exp -j(\omega_1 + \omega_2)) \exp j(\omega_1 + \omega_2)\tau$$

en regroupant les déphasages linéaires

$$R_3(\omega_1, \omega_2) = \quad (218)$$

$$\text{cum}_3(x)B(\exp j\omega_1) B(\exp j\omega_2) B(\exp -j(\omega_1 + \omega_2)) \exp -j\omega_1\tau \exp -j\omega_2\tau \exp j(\omega_1 + \omega_2)\tau$$

$$R_3(\omega_1, \omega_2) = \text{cum}_3(x)B(\exp j\omega_1) B(\exp j\omega_2) B(\exp -j(\omega_1 + \omega_2)) \quad (219)$$

Le spectre d'ordre trois est inchangé.

Effet d'un retard inférieur à la période d'échantillonnage Dans le cas où le retard τ est inférieur à la période d'échantillonnage, le déphasage linéaire traduisant ce retard dans le domaine des fréquences est

$$\exp -j\omega\tau \quad (220)$$

pour $-\pi \leq \tau < \pi$. Il est périodique en ω de période 2π et présente en $\pm\pi$ un saut qui est égal à $2\tau\pi$ et non 2π comme dans le cas du retard entier où $\tau = 1$.

Pour $\pi \leq \omega < 3\pi$, on a

$$B(e^{j\omega}) = e^{j(2\pi - \omega)\tau} \quad (221)$$

Lorsqu'on calcule le bispectre par la formule (197) pour

$$B(e^{j\omega}) = e^{-j\omega\tau} \quad (222)$$

on voit que l'effet à l'intérieur du support carré $(-\pi, \pi) \times (-\pi, \pi)$ est différent à l'intérieur de l'hexagone et dans les deux triangles qui le complètent: à l'intérieur de l'hexagone,

$$R_3(\omega_1, \omega_2) = 1 \quad (223)$$

et dans les deux triangles

$$R_3(\omega_1, \omega_2) = e^{-2\pi\tau} \quad (224)$$

Il est intéressant de noter que l'information contenue dans ces portions triangulaires du support peut être utilisée pour caractériser des retards inférieurs à une période d'échantillonnage: par exemple le facteur d'asymétrie d'une séquence i.i.d. filtrée passe-bas (à la moitié de la fréquence d'échantillonnage) puis retardée de τ et rééchantillonnée voit son facteur d'asymétrie cum_3 multiplié par

$$\frac{3}{4} + \frac{1}{4} \cos(2\pi\tau).$$

4.4.1 Statistiques d'ordre quatre et synchronisation

On aboutit à une constatation identique dans le cas des spectres d'ordre quatre, mais l'effet est plus sensible car les pyramides à base triangulaires qui complètent l'octaèdre occupent un sous-espace proportionnellement plus important que les triangles du cas du spectre d'ordre trois. Dans le cas du spectre d'ordre quatre le kurtosis d'un signal i.i.d. filtré à la moitié de la fréquence d'échantillonnage et retardé de τ sera

$$\frac{2}{3} + \frac{1}{3} \cos(2\pi\tau).$$

La synchronisation sur une source pourrait être fondée directement sur la mesure de cette fonction.

D'ailleurs, tout comme l'égalisation utilise implicitement des statistiques d'ordre quatre, les techniques de synchronisation y font aussi appel par l'utilisation de fonctions non linéaires.

4.4.2 Statistiques d'ordre quatre et stationnarité

Notons que cet effet disparaît si l'information contenue dans les triangles extérieurs à l'hexagone du spectre d'ordre trois (ou les pyramides extérieures à l'octaèdre du spectre d'ordre quatre) est nulle ; c'est à dire si la bande passante du signal est inférieure au tiers de la fréquence d'échantillonnage dans le premier cas et au quart de la fréquence d'échantillonnage dans le second cas, c'est à dire quand les signaux à temps continu (avant échantillonnage) sont stationnaires respectivement à l'ordre trois et à l'ordre quatre.

Pour une analyse détaillée de ces résultats, on se reportera à la thèse de Delphine Rossille [19].

4.5 Reconstruction d'une fonction de transfert dans le domaine des fréquences

Si on se donne $R_3(\omega_1, \omega_2)$ on peut retrouver $B(e^{j\omega})$ à un déphasage linéaire (retard près). Nous traiterons le cas d'un signal réel pour lequel on a calculé le spectre d'ordre trois pour N valeurs discrètes des fréquences (ω_1 et ω_2 sont des multiples de $2\pi/N$)

On a, si le spectre d'ordre trois est parfaitement estimé,

$$R_3\left(\omega_1, \frac{2\pi}{N}\right) = cum_3(x)B\left(\exp \frac{2j\pi}{N}\right)B(\exp j\omega_1)B\left(\exp -j\left(\omega_1 + \frac{2\pi}{N}\right)\right) \quad (225)$$

En général on s'intéresse à la reconstruction de la phase $\varphi(\omega)$ de $B(\exp j\omega)$ dans la mesure où les caractéristiques des modules de fonctions de transfert peuvent se déduire des statistiques d'ordre deux. On peut généraliser aux modules l'approche donnée ici pour les phases. En termes de phases l'équation (225) devient

$$\psi\left(\omega_1, \frac{2\pi}{N}\right) = \varphi\left(\frac{2\pi}{N}\right) + \varphi(\omega_1) - \varphi\left(\omega_1 + \frac{2\pi}{N}\right) \quad (226)$$

On peut forcer $\varphi(0)$ à zéro (on peut choisir le signe de la fonction à reconstituer et se donner arbitrairement $\varphi\left(\frac{2\pi}{N}\right)$ qui caractérisera le déphasage linéaire qu'on doit choisir pour reconstituer la fonction de transfert. L'équation (226) est une équation récurrente qui permet de retrouver toutes les valeurs de $\varphi(\omega)$: si on pose

$$\omega_1 = \frac{2\pi}{N}n \quad (227)$$

$$\varphi\left(\frac{2\pi}{N}n + \frac{2\pi}{N}\right) = \varphi\left(\frac{2\pi}{N}\right) + \varphi\left(\frac{2\pi}{N}n\right) - \psi\left(\frac{2\pi}{N}n, \frac{2\pi}{N}\right) \quad (228)$$

Cette proposition est due à Lohmann, Weigelt et Wirtzner [14]. Il existe des variantes de cette formule et on peut aussi en déduire des techniques plus élaborées. De nombreuses références bibliographiques et des algorithmes originaux sont décrits dans les thèses [19][10].

4.6 Les perturbations en imagerie par interférométrie en astronomie

On peut montrer, d'après la théorie de Maxwell sur la propagation des ondes électromagnétiques, que l'information qui se propage après émission par une source lumineuse (supposée bidimensionnelle) peut s'interpréter comme la transformée de Fourier bidimensionnelle de cette source. Par exemple dans les expériences élémentaires d'interférométrie, si on perce un écran vertical de deux ouvertures ponctuelles situées sur la même horizontale et qu'on éclaire par une source de lumière cohérente, la lumière qui se propage horizontalement à partir de ces deux ouvertures va interférer, et si on l'observe sur un écran situé à une

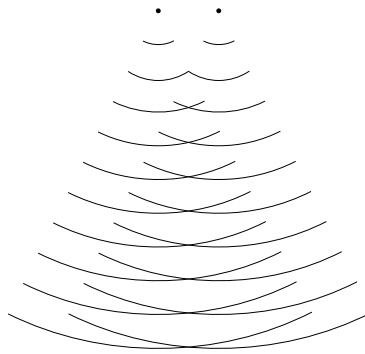


Figure 15: Les interférences entre deux sources lumineuses ponctuelles cohérentes produisent un signal sinusoïdal dont l'amplitude et la phase caractérisent la géométrie de ces sources

grande distance des deux sources et perpendiculaire à la direction de propagation des ondes, on y verra un signal sinusoïdal bidimensionnel dont les courbes de niveau sont perpendiculaires au vecteur reliant les deux points et dont la fréquence est inversement proportionnelle à la distance entre les deux ouvertures. C'est bien la transformée de Fourier de l'image de l'écran qu'on interprète comme deux impulsions. Cette propriété est utilisée en observation astronomique : comme il est pratiquement impossible de réaliser des télescopes de très grand diamètre, on capte les ondes en provenance d'une direction par deux télescopes et on fait interférer ces ondes. En mesurant l'amplitude et la phase des franges d'interférence pour une position des télescopes on obtient l'amplitude et la phase de l'objet pour une fréquence spatiale caractérisée par le vecteur donnant la différence de position entre les deux télescopes; on peut ainsi arriver à distinguer la distance angulaire séparant deux composantes d'une étoile double. En dehors des problèmes techniques posés par la mise au point des franges d'interférence, il existe une difficulté supplémentaire parce que la traversée de l'atmosphère par les ondes lumineuses les affecte d'une perturbation: les ondes issues de l'objet étudié considérées comme planes sont déformées par les fluctuations de pression de l'atmosphère, ce qui se traduit par un effet de miroitement ou de scintillement (tavelures ou "speckles" en anglais) qui peut rendre l'image indéchiffrable. On peut considérer que cette perturbation est un déphasage aléatoire dans le temps et dans l'espace, de moyenne nulle.

4.7 Clôture de phase

Le signal d'interférence obtenu à partir de deux signaux issus de capteurs situés en deux points A et B est égal à un signal sinusoïdal bidimensionnel dont l'amplitude et la phase sont celles de la transformée de Fourier $F(u, v)$ du signal analysé (la répartition spatiale $f(x, y)$ de la luminosité de l'étoile) calculée pour la fréquence spatiale (u_α, v_α) où u_α et v_α sont les composantes du vecteur AB . La phase de $F(u_\alpha, v_\alpha)$ est ainsi la différence de phase entre les signaux reçus par le capteur A et le capteur B , soit

$$(\varphi_B - \varphi_A)$$

qui ne dépend que de l'image analysée et non pas du temps. Lors de mesures à travers l'atmosphère présentant une agitation thermique, une perturbation aléatoire

$$\Delta\varphi_B(t) - \Delta\varphi_A(t)$$

est ajoutée à cette différence de phase. Elle ne peut donc pas être mesurée directement.

Cependant, il est possible d'éliminer ces déphasages aléatoires de la manière suivante:

On considère trois points A , B et C , et on nomme $\vec{\alpha}$, $\vec{\beta}$ et $\vec{\gamma}$ les vecteurs \vec{AB} , \vec{BC} et \vec{CA} . Leurs composantes sont respectivement (u_α, v_α) , (u_β, v_β) et (u_γ, v_γ) .

On a

$$\vec{\alpha} + \vec{\beta} + \vec{\gamma} = 0 \quad (229)$$

si on mesure les amplitudes et les phases des trois composantes, on obtient

$$F(u_\alpha, v_\alpha) = \rho(u_\alpha, v_\alpha) \exp[j(\varphi(u_\alpha, v_\alpha) + \Delta\varphi_B(t) - \Delta\varphi_A(t))] \quad (230)$$

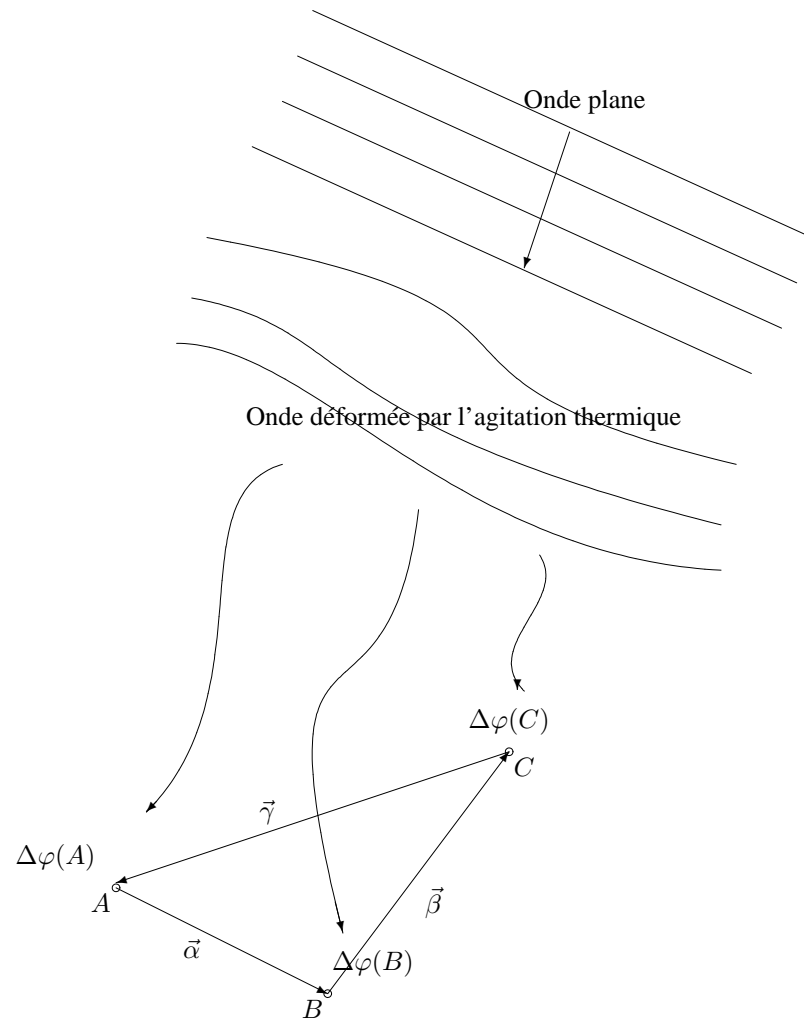


Figure 16: Schéma représentant les perturbations sur les phases des signaux captés par trois télescopes

$$F(u_\beta, v_\beta) = \rho(u_\beta, v_\beta) \exp[j(\varphi(u_\beta, v_\beta) + \Delta\varphi_C(t) - \Delta\varphi_B(t))] \quad (231)$$

$$F(u_\gamma, v_\gamma) = \rho(u_\gamma, v_\gamma) \exp[j(\varphi(u_\gamma, v_\gamma) + \Delta\varphi_A(t) - \Delta\varphi_C(t))] \quad (232)$$

On effectue le produit de ces trois valeurs, on obtient de cette façon le *bispectre* que nous nommerons

$$R(u_\alpha, v_\alpha, u_\beta, v_\beta)$$

. Il ne fait intervenir que les variables

$$u_\alpha, v_\alpha, u_\beta, v_\beta$$

car les deux autres variables se déduisent des quatre premières d'après l'éq. (229), les trois termes de déphasage aléatoires dûs aux perturbations atmosphériques se compensent. Cette propriété est connue sous le nom de cloture (parfois de fermeture) de phase. D'après l'éq. (229)

$$u_\gamma = -(u_\alpha + u_\beta) \quad (233)$$

$$v_\gamma = -(v_\alpha + v_\beta) \quad (234)$$

$$(235)$$

et le bispectre s'écrit

$$R(u_\alpha, v_\alpha, u_\beta, v_\beta) = \rho(u_\alpha, v_\alpha) \rho(u_\beta, v_\beta) \rho(-u_\alpha - u_\beta, -v_\alpha - v_\beta) \exp[j(\varphi(u_\alpha, v_\alpha) + \varphi(u_\beta, v_\beta) + \varphi(-u_\alpha - u_\beta, -v_\alpha - v_\beta))] \quad (236)$$

car la somme des trois perturbations est nulle:

$$\Delta\varphi_B(t) - \Delta\varphi_A(t) + \Delta\varphi_C(t) - \Delta\varphi_B(t) - \Delta\varphi_A(t) - \Delta\varphi_C(t) = 0 \quad (237)$$

On a ainsi éliminé les perturbations aléatoires. En effectuant cette opération on a aussi éliminé les éventuels composantes linéaires de la forme $(\lambda\vec{u} + \mu\vec{v})$ présentes dans les phases $\varphi(u_\alpha, v_\alpha)$, $\varphi(u_\beta, v_\beta)$ et $\varphi(u_\gamma, v_\gamma)$. Cette élimination n'a pas de conséquences majeures car un déphasage linéaire se traduit dans le domaine spatial par une translation mais ne déforme pas l'image analysée. Si on excepte cette translation, le bispectre contient l'information $F(u, v)$. Il faut bien sûr trouver des moyens de déduire $F(u, v)$ de $R(u_\alpha, v_\alpha, u_\beta, v_\beta)$

En pratique, il existe des techniques indépendantes de la mesure du bispectre pour calculer l'amplitude de la transformée de Fourier de l'image ; le bispectre est surtout utile pour la reconstitution de la phase de cette transformée et on sait que cette de phase est très informative dans la représentation de la forme de l'image cherchée.

En ce qui concerne la recherche de la phase de $F(u, v)$ à partir de la phase $\psi(u_\alpha, v_\alpha, u_\beta, v_\beta)$ de $R(u_\alpha, v_\alpha, u_\beta, v_\beta)$, on peut utiliser l'algorithme décrit dans le cas monodimensionnel on peut initialiser un algorithme en se fixant a priori et arbitrairement (par exemple à 0) deux premières valeurs de la phase $\varphi(u_\alpha, v_\alpha)$, $\varphi(u_\beta, v_\beta)$. Ce choix arbitraire revient simplement à se fixer la translation que subira l'image, et cette translation n'a pas d'importance dans le problème étudié où on s'intéresse à la forme d'un objet . Si on a mesuré $\psi(u_\alpha, v_\alpha, u_\beta, v_\beta)$, on peut donc en déduire la phase φ en un troisième point du plan des fréquences

$$\varphi(-u_\alpha - u_\beta, -v_\alpha - v_\beta) = -\psi(u_\alpha, v_\alpha, u_\beta, v_\beta) + \varphi(u_\alpha, v_\alpha) + \varphi(u_\beta, v_\beta) \quad (238)$$

L'image analysée est réelle et sa phase est une fonction impaire

$$\varphi(-u, -v) = -\varphi(u, v) \quad (239)$$

On connaît alors la phase en six points du plan des fréquences : $\alpha, -\alpha, \beta, -\beta, \gamma, -\gamma$. Il y a cloture de phase pour le triplet (α, β, γ) et pour le triplet $(-\alpha, -\beta, -\gamma)$. Si maintenant on mesure simultanément la phase pour les fréquences $\alpha, -\gamma$ et $(-\alpha + \gamma)$ et si on calcule le bispectre associé à ce triplet on peut donc en déduire la phase non perturbée $\varphi(u, v)$ à la fréquence $(-\alpha + \gamma)$. Par un choix judicieux du triple de fréquences pour lequel on calcule le bispectre on peut donc augmenter le nombre d'échantillons du plan des fréquences pour lesquels la phase est connue. Si de plus on tient compte d'informations connues a priori sur l'image cherchée ou d'hypothèses permettant de compléter la transformée de Fourier dans le plan des fréquences, on peut estimer cette transformée de Fourier de l'image dans la région utile du domaine bidimensionnel des fréquences et donc par transformée inverse, reconstruire l'image de l'objet étudié.

4.8 Maximisation du kurtosis dans le domaine des fréquences

On peut interpréter la maximisation du kurtosis dans le domaine des fréquences, ce qui conduit à des résultats assez simples.

Nous nous intéressons au cas de l'identification d'une fonction de transfert de module constant égal à un et qui agit donc comme un déphaseur pur sur un signal i.i.d. de kurtosis non nul. Le filtre $B(z)$ a pour forme

$$B(e^{j\omega}) = e^{j\varphi(\omega)} \quad (240)$$

La contrainte imposant que la variance de l'entrée est égale à la variance de la sortie est donc nécessairement respectée. La maximisation du kurtosis, connaissant le spectre d'ordre trois de la sortie du filtre inconnu revient donc à la maximisation de

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j\psi(\omega_1, \omega_2, \omega_3)} \left(e^{j\varphi(\omega_1)} e^{-j\varphi(\omega_2)} e^{j\varphi(\omega_3)} e^{-j\varphi(\omega_1 + \omega_2 + \omega_3)} \right)^{-1} d\omega_1 d\omega_2 d\omega_3 \quad (241)$$

(par application du théorème de Parseval). Ce qui donne par symétrie

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} 2 \cos(\psi(\omega_1, \omega_2, \omega_3) - \varphi(\omega_1) + \varphi(\omega_2) - \varphi(\omega_3) + \varphi(\omega_1 + \omega_2 + \omega_3)) d\omega_1 d\omega_2 d\omega_3 \quad (242)$$

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} 2 \sin^2 \frac{1}{2} (\psi(\omega_1, \omega_2, \omega_3) - \varphi(\omega_1) + \varphi(\omega_2) - \varphi(\omega_3) + \varphi(\omega_1 + \omega_2 + \omega_3)) d\omega_1 d\omega_2 d\omega_3 \quad (243)$$

C'est l'expression d'une distance euclidienne dans le plan complexe entre deux points de module un.

Cette expression dans le domaine des fréquences permet une interprétation simple de l'efficacité de la recherche du maximum du kurtosis par la méthode du point fixe proposée par Shalvi et Weinstein [21]: en l'absence de bruit, pour une estimation parfaite des caractéristiques statistiques d'ordre quatre, la fonction étudiée passe par un maximum au point d'équilibre de l'algorithme du point fixe, ce qui garantit une convergence très rapide (voir [13]):

A l'itération p , la nouvelle estimation de $\varphi(\omega)$ est ainsi donnée par

$$\varphi_p(\omega) = \arg \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \exp j (\psi(\omega, \omega_2, \omega_3) + \varphi_{p-1}(\omega_2) - \varphi(\omega_3) + \varphi_{p-1}(\omega + \omega_2 + \omega_3)) d\omega_2 d\omega_3 \quad (244)$$

Références

- [1] A. Benveniste, M. Goursat and G. Ruget, "Robust identification of a nonminimum phase system: blind adjustment of a linear equalizer in data communications," *IEEE trans. automatic control*, vol. AC-25, no. 3, June 1980, pp. 1740-1748.
- [2] J. J. Bussgang, "Cross correlation functions of amplitude-distributed gaussian signals", *technical report 216*, MIT Research laboratory of electronics, Cambridge, 1952.
- [3] J. A. Cadzow, "Blind deconvolution via cumulant extrema," *IEEE sig. proc. magazine*, 1996.
- [4] P. Comon, "Independent Component Analysis, a new concept?," *Signal Processing, Special issue on Higher-Order Statistics*, vol. 36, no 3, April 1994 pp 287-314.
- [5] Z. Ding and Y. Li, "Blind equalization and identification", Marcel Dekker ed. , 2000
- [6] D. Donoho, "On minimum entropy deconvolution", *Applied time series analysis II*, Academic Press, 1981, pp. 565-609.
- [7] G. Giannakis and J. M. Mendel, "Identification of non-minimum phase systems using higher-order statistics," *IEEE trans. ASSP*, vol. 37, 1989, pp. 360-377.
- [8] D. Godard, "Self recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE trans. on Communications*, vol. COM-28 no. 11, nov. 1980, pp. 1867-1875.
- [9] S. Haykin, ed., "Unsupervised adaptive filtering, vol. 2: Blind deconvolution", Wiley, 2000.

- [10] C. Huet, "Factorisation des spectres d'ordre supérieur de signaux complexes, application en identification aveugle de systèmes et reconstruction d'images", *Thèse de l'université de Nice*, 1997.
- [11] J. L. Lacoume, P.O. Amblard et P. Comomn, "Statistiques d'ordre supérieur pour le traitement du signal", Ed. Masson, 1997
- [12] J. Le Roux and C. Huet, "On a condition of higher order spectra factorization", *IEEE trans. on Signal Processing*, vol. 47, n° 7, July 1999, pp. 1903-1909.
- [13] J. Le Roux and S. Icart, "System frequency response phase estimation based on skewness or kurtosis maximization", *ICASSP 2000 in Istanbul, Turkey*, 2000, June 5-9.
- [14] A. W. Lohmann, G. Weigelt and B. Winitzer, "Speckle masking in astronomy: triple correlation theory and application", *Applied physics*, vol. 22, pp 4028-4037.
- [15] R. W. Lucky, "Techniques for adaptive equalization of digital communication systems", *Bell system tech. J.*, vol. 45, 1966, pp. 255-286.
- [16] J. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications", *Proceedings of the IEEE*, vol. 79, no. 3, March 1991, pp. 278-305
- [17] S. U. H. Qureshi, "Adaptive equalization", *Proc. IEEE*, vol. 73, no. 9, Sept. 1985, pp. 1349-1387.
- [18] Y. Sato, "A method of self recovering equalization for multilevel amplitude modulation", *IEEE trans. communications*, vol. COM-23, june 1975, pp. 679-682.
- [19] D. Rossille, "Reconstruction de signaux à partir du bispectre, application à l'astronomie : effet de l'échantillonnage et de la stationnarité sur les spectres d'ordre supérieur", *Thèse de Doctorat, Université de Nice*, Juin 1995.
- [20] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of non-minimum phase systems (channels)," *IEEE trans. information theory*, vol. 36, March 1990, pp. 312-321.
- [21] O. Shalvi and E. Weinstein, "Super exponential methods for blind deconvolution," *Proc. SPIE 1991, Adaptive signal processing*, vol. 1565, Jan. 1991, pp. 143-151 ; aussi dans *IEEE Trans. Information Theory* vol. 39, no. 2, March 1993, pp 504-519.
- [22] L. Tong, G. Xu and T. Kailath, "Blind identification and equalization based on on second-order statistics, : A time domain approach", *IEEE trans. Information theory*, vol. IT-40, no 2, March 1994, pp. 340-349.