

Interprétation régularisée des périodogrammes et extensions non quadratiques

Jérôme Idier, Jean-François Giovannelli et Philippe Ciuciu

Laboratoire des Signaux et Systèmes,

SUPÉLEC, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France

idier@lss.supelec.fr, giovannelli@lss.supelec.fr, ciuciu@lss.supelec.fr

RÉSUMÉ

L'objectif est, d'une part, de montrer que les méthodes non paramétriques d'analyse spectrale ont une interprétation directe dans le cadre de la théorie de la régularisation des problèmes mal posés par pénalisation quadratique ; d'autre part, par l'introduction de critères convexes non quadratiques, de mettre en évidence que « non paramétrique » et « haute résolution » ne sont pas plus incompatibles en analyse spectrale qu'en imagerie.

ABSTRACT

Our contribution is first to show that non parametric approaches to spectral analysis can be interpreted using quadratic penalization in the theoretical setting of regularization of ill-posed problems. Second, we introduce convex non quadratic criteria, and we show that "non parametric" is not incompatible with "high resolution," neither in the field of spectral analysis nor in imaging problems.

1 Introduction

L'objectif de cette communication¹ est double. D'une part, il s'agit de montrer que certaines techniques de base en analyse spectrale ont une interprétation directe dans le cadre de la régularisation des problèmes mal posés. Nous nous intéressons particulièrement au cas de l'analyse spectrale non paramétrique. La vocation de cette première partie est essentiellement pédagogique, et contient surtout des éléments connus mais peu diffusés. Elle montre que le périodogramme et ses avatars peuvent être considérés comme les restrictions linéaires, donc d'un pouvoir résolvant limité, d'une classe de méthodes d'analyse spectrale non paramétrique beaucoup plus vaste.

Dans une deuxième partie, nous fournissons des résultats originaux, qui permettent d'introduire des estimateurs non linéaires issus de l'analyse convexe. L'accent est porté sur les fonctions de régularisation markoviennes. L'origine de notre étude se trouve précisément dans l'idée qu'une partie des outils markoviens du traitement d'image bas-niveau peuvent s'adapter au problème de l'analyse spectrale.

2 Un problème mal posé

On note $\mathbf{y}_K = [y_1, \dots, y_K]^t$ les observations, supposées complexes, $X(\nu)$ le spectre complexe cherché (en fait le but de l'estimation est plutôt $|X(\nu)|^2$), $x_k = \int_0^1 X(\nu) e^{2i\pi\nu k} d\nu$, et $\mathbf{x}_K = [x_1, \dots, x_K]^t$. On formule ainsi le problème de l'analyse spectrale :

trouver $X(\nu)$, $\nu \in [0, 1]$, tel que $\mathbf{x}_K = \mathbf{y}_K$.

Cette formulation diffère de la formulation statistique du problème, qui consiste à trouver la densité spectrale de puissance d'un processus aléatoire à partir de \mathbf{y}_K considéré comme un fragment de trajectoire. Elle se rapproche plutôt des problèmes de synthèse de Fourier en astronomie [1]. Il existe évidemment une infinité de solutions, si bien que ce problème est *mal posé* au sens de Hadamard. Il s'agit d'ailleurs d'un problème inverse, en considérant que le problème direct associé consiste à trouver \mathbf{y} à partir de $X(\nu)$, problème bien posé.

Plusieurs approches permettent de régulariser un problème initialement mal posé [2]. En particulier, les méthodes de stabilisation par *contrôle de dimensionnalité* s'opposent aux méthodes *pénalisées* [3]. Dans le cadre de l'analyse spectrale, les méthodes paramétriques rentrent naturellement dans la première catégorie, sauf exception. Ici, nous nous intéresserons plutôt à la seconde.

3 Périodogramme simple

On montre facilement, en évitant le recours aux principes variationnels, que le périodogramme simple est le carré du module de la solution du problème suivant :

$$\widehat{X}^0 = \arg \min_{X \in L^2_{\mathbb{C}}[0,1]} \int_0^1 |X(\nu)|^2 d\nu \text{ s.c. } \mathbf{x}_K = \mathbf{y}_K.$$

Il s'agit donc de la solution *inverse généralisée* du problème. D'après le théorème de Plancherel-Parseval, il est équivalent de chercher directement la suite infinie des $\{\widehat{x}_k^0\}_{k \in \mathbb{Z}}$ dont $\widehat{X}^0(\nu)$ est la transformée de Fourier, et qui est solution du problème suivant :

$$\widehat{x}^0 = \arg \min_{x \in \ell^2_{\mathbb{C}}} \sum_{k \in \mathbb{Z}} |x_k|^2 \text{ s.c. } \mathbf{x}_K = \mathbf{y}_K.$$

¹Une partie de ce travail a été soutenue par THOMSON CSF AIRSYS dans le cadre de contrats de coopération établis avec le CNRS en 95 et 96 [6,7].

Le résultat est trivial puisque le problème se sépare : $\widehat{x}_k^0 = y_k$ si $k \in \{1, \dots, K\}$; $\widehat{x}_k^0 = 0$ sinon, donc on a bien $\widehat{X}^0(v) = \sum_{k=1}^K y_k e^{-2i\pi vk}$. On peut aussi remarquer que le même problème contraint peut être formulé comme la limite d'un problème pénalisé :

$$\widehat{X}^0 = \lim_{\lambda \rightarrow 0^+} \arg \min_{X \in L^2_{\mathbb{C}}[0,1]} \|x_K - y_K\|^2 + \lambda \int_0^1 |X(v)|^2 dv,$$

en notant d'ailleurs que la solution du problème pénalisé s'écrit $\widehat{X}^\lambda(v) = \widehat{X}^0(v)/(1+\lambda)$, comme transformée de Fourier de la suite $\widehat{x}^\lambda = \widehat{x}^0/(1+\lambda)$, solution du problème équivalent en temps :

$$\widehat{x}^\lambda = \arg \min_{x \in \ell^2_{\mathbb{C}}} \|x_K - y_K\|^2 + \lambda \sum_{k \in \mathbb{Z}} |x_k|^2.$$

4 Périodogrammes fenêtrés

Il s'agit de spectres estimées sous la forme

$$\left| \sum_{k=1}^K \mu_k y_k e^{-2i\pi vk} \right|^2,$$

où $\mu = [\mu_1, \dots, \mu_K]^t$ est une fenêtre temporelle de forme prescrite pour diminuer la variance du périodogramme, au détriment de sa résolution. Pour interpréter ces solutions par minimisation de critère, on est amené cette fois à considérer directement une forme pénalisée. En effet, en introduisant une régularisation adaptative décrite par des paramètres $\lambda_k \in]0, 1]$, le problème pénalisé en temps devient :

$$\arg \min_{x \in \ell^2_{\mathbb{C}}} \|x_K - y_K\|^2 + \sum_{k \in \mathbb{Z}} \lambda_k |x_k|^2. \tag{1}$$

La solution s'obtient sans difficulté sous la forme :

$$\widehat{x}_k^\lambda = y_k/(1 + \lambda_k) \text{ si } k \in \{1, \dots, K\} ; \widehat{x}_k^\lambda = 0 \text{ sinon,}$$

autrement dit, la solution est un périodogramme fenêtré. Coefficients de la fenêtre et paramètres de régularisation $\lambda = [\lambda_1, \dots, \lambda_K]^t$ sont en correspondance terme à terme par la relation $\mu_k = 1/(1 + \lambda_k)$, $k = 1, \dots, K$, tandis que les autres coefficients λ_k n'interviennent pas dans la solution.

En exploitant à nouveau le théorème de Plancherel-Parseval, et en introduisant $\Lambda(v) = \sum_{k=1}^K \lambda_k e^{-2i\pi vk}$, on obtient la formulation équivalente :

$$\widehat{X}^\lambda = \arg \min_{X \in L^2_{\mathbb{C}}[0,1]} \|x_K - y_K\|^2 + \iint_{[0,1]^2} \Lambda(v - \omega) X(\omega) X^*(v) dv d\omega.$$

Exemple 4.1 (Pénalisation de la dérivée première). On recherche le fenêtrage associé à :

$$\arg \min_{X \in L^2_{\mathbb{C}}[0,1]} \|x_K - y_K\|^2 + \lambda \int_0^1 |X'(v)|^2 dv.$$

D'après le théorème de Plancherel-Parseval appliqué à $(-2i\pi kx_k)$, dont $X'(v)$ est la transformée de Fourier, on obtient $\lambda_k = \lambda |-2i\pi k|^2$ par identification. Par conséquent $\mu_k = 1/(1 + 4\pi^2 k^2 \lambda)$, c'est-à-dire une fenêtre de Cauchy.

Il existe enfin une troisième formulation, adaptée à un calcul direct de la solution \widehat{X}^λ sur une grille fréquentielle régulière, possédant un nombre P de points multiple de K . Soit \widehat{X}^λ le vecteur ainsi constitué. Par transformée de Fourier discrète (TFD) du problème (1) sur $\{1, \dots, P\}$, on obtient directement le résultat suivant :

$$\widehat{X}_P^\lambda = \arg \min_{X \in \mathbb{R}^P} \|x_K - y_K\|^2 + X^s \Lambda X,$$

où X est la TFD de x_P , et Λ est une matrice circulante dont la première ligne est la TFD du vecteur λ étendu sur P points par des zéros. Dans cette interprétation, les périodogrammes moyennés sont issus d'un principe des moindres carrés régularisés. Pour $P = K$, la fenêtre de pondération μ est en relation bijective avec la matrice circulante Λ .

Exemple 4.2 (Pénalisation de la différence première). Soit

$$X^s \Lambda X = \lambda P^2 \sum_{k=1}^P |X_k - X_{k-1}|^2$$

(on note $X_0 = X_K$). Sous une forme centrée autour de 0, les paramètres de régularisation sont donnés par $\lambda_k = 2\lambda P^2(1 - \cos 2\pi k/P)$. Pour une discrétisation de plus en plus fine, $\lim_{P \rightarrow \infty} \lambda_k = 4\pi^2 k^2 \lambda$, et on retrouve pour μ la fenêtre de Cauchy de l'Exemple 4.1.

5 Périodogrammes « non quadratiques »

5.1 Pénalisation séparable

Le périodogramme simple s'interprète comme une solution régularisée par pénalisation quadratique. Or le caractère « non résolvant » de la pénalisation quadratique est bien connu en restauration d'image. Comme dans ce domaine, il peut sembler intéressant de remplacer la pénalisation quadratique par une fonctionnelle

$$\mathcal{F}(|X|) = \lambda_0 \int_0^1 J_0(|X(v)|) dv,$$

où $J_0(x)$ est une fonction croissant moins vite à l'infini que x^2 , afin d'augmenter le pouvoir de restitution de spectres « piqués ».

La convexité de $J_0(|\cdot|)$ implique celle de $\mathcal{F}(|\cdot|)$, ce qui assure l'existence d'une solution unique dans $L^2_{\mathbb{C}}$, continue en fonction des données, donc le caractère *bien posé* du problème ainsi régularisé [2]. En revanche la convexité de J_0 n'implique pas celle de $J_0(|\cdot|)$. Appliqué à $g(X) = |X|$ et à $f(\rho) = J_0(\rho)$, le théorème suivant assure la convexité pour toutes les fonctions J_0 convexes croissantes sur \mathbb{R}_+ .

Théorème 5.1 (Composée de fonct. convexes I) — Soit g et f deux fonctions, respectivement de \mathbb{R}^n dans \mathbb{R}_+ et de \mathbb{R}_+ dans \mathbb{R}_+ . Si g est convexe et f convexe et croissante, alors $f \circ g$ est convexe.

Démonstration : g est convexe et f croissante, donc, pour tout $x, y \in \mathbb{R}^n$ et $\alpha \in [0, 1]$,

$$f(g(\alpha x + (1 - \alpha)y)) \leq f(\alpha g(x) + (1 - \alpha)g(y)),$$

$$\leq \alpha f(g(x)) + (1 - \alpha)f(g(y)),$$

car f est convexe. ■

Exemple 5.1. Les fonctions « entropiques » $\rho \ln \rho$ et $-\ln \rho$ introduites en synthèse de Fourier [1], et plus généralement en imagerie pour restaurer des objets X réels positifs, sont convexes mais ne croissent pas sur \mathbb{R}_+ . Elles ne sont donc pas utilisables dans le présent contexte.

Exemple 5.2. Les fonctions « L_p » ρ^p ($p \geq 1$) et les fonctions « L_{21} » telles que $\sqrt{\beta + \rho^2}$ satisfont les conditions du théorème et croissent moins vite que la fonction « L_2 » ρ^2 . La Figure 1b représente un spectre obtenu pour $J_0(\rho) = \sqrt{\beta + \rho^2}$, qui représente un gain en résolution très sensible par rapport au périodogramme représenté Figure 1a.

En pratique, la résolution du problème résulte de la discrétisation du critère fonctionnel sur une grille fréquentielle régulière sous la forme

$$\widehat{X}_P = \arg \min_{\mathbf{x} \in \mathbb{R}^P} \|\mathbf{x}_K - \mathbf{y}_K\|^2 + \frac{\lambda_0}{P} \sum_{k=1}^P J_0(|X_k|),$$

puis de la mise en œuvre d'un algorithme de descente exploitant la convexité du critère.

5.2 Pénalisation gibbsienne

En favorisant l'apparition de pics, la pénalisation séparable manque de stabilité pour restituer les parties les plus régulières du spectre recherché. Il est donc naturel de lui adjoindre une contribution gibbsienne [4, 5]; une extension directe du périodogramme fenêtré de l'Exemple 4.1 suggère le terme

$$\lambda_1 \int_0^1 J_1(|X'(v)|) dv,$$

en choisissant pour J_1 une fonction non quadratique, auquel le Théorème 5.1 s'applique comme à J_0 . La Figure 1c représente un spectre obtenu pour $J_1(\rho) = \sqrt{\beta + \rho^2}$ ($\lambda_0 = 0$). La régularité du résultat a été obtenue au détriment du pouvoir de séparation des pics proches. Il faut également remarquer que la pénalisation porte simultanément sur le module $|X(v)|$ et sur la phase, tandis qu'il peut sembler plus pertinent de contrôler la régularité du module seulement grâce à une fonctionnelle de pénalisation du type :

$$\lambda_1 \int_0^1 J_1\left(\frac{d}{dv}|X(v)|\right) dv.$$

Sous forme discrétisée, la fonction de pénalisation devient

$$J(|X|) = \frac{\lambda_0}{P} \sum_{k=1}^P J_0(|X_k|) + \frac{\lambda_1}{P} \sum_{k=1}^P J_1(P(|X_k| - |X_{k-1}|)).$$

L'étude de convexité est plus délicate. Elle nécessite une extension multivariée du théorème précédent :

Théorème 5.2 (Composée de fonct. convexes II) — Soit g et f deux fonctions, respectivement de \mathbb{R}^n dans \mathbb{R}_+^m et de \mathbb{R}_+^m dans \mathbb{R}_+ . Si chaque composante g_k de g est convexe et si f est convexe et croissante coordonnée par coordonnée, en tout point de \mathbb{R}_+^m , alors $f \circ g$ est convexe.

Démonstration : Pour tout $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ et $\alpha \in [0, 1]$,

$$\forall k, \quad g_k(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha g_k(\mathbf{x}) + (1 - \alpha)g_k(\mathbf{y})$$

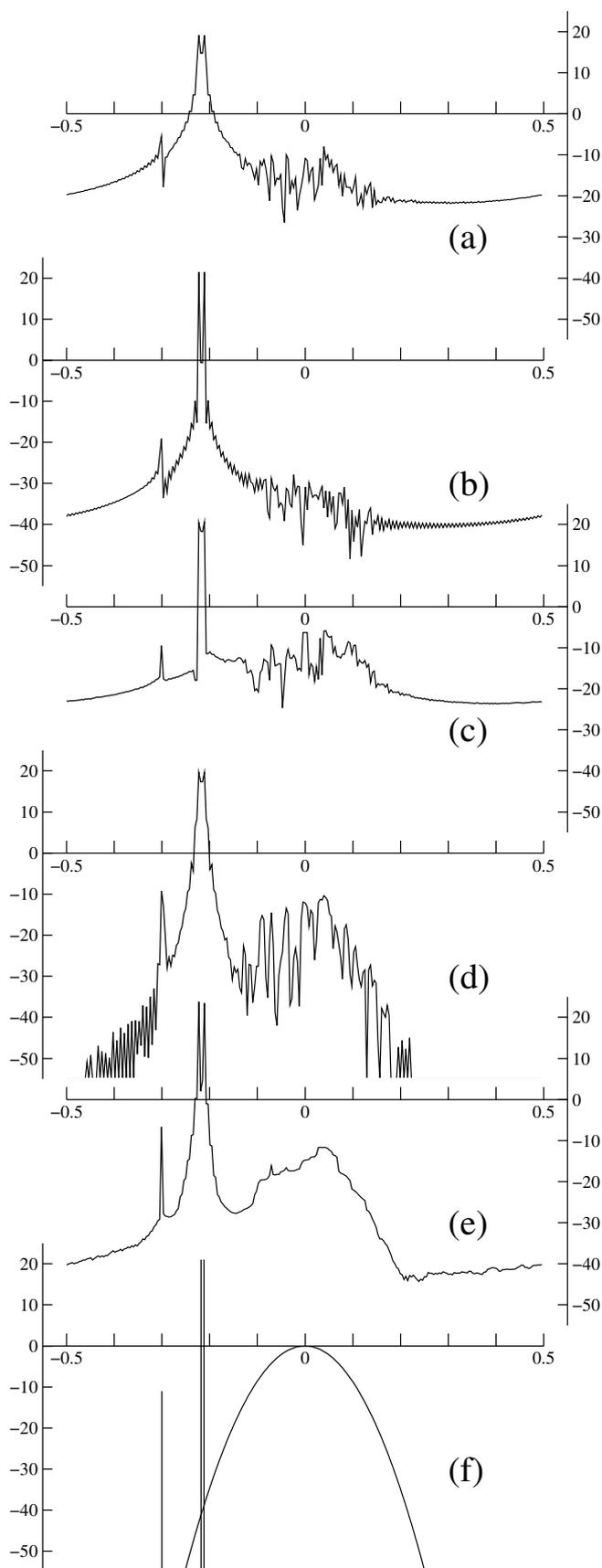


FIG. 1 — (a) Périodogramme simple et plusieurs variantes non quadratiques ; (b) Pénalisation séparable ; (c) Pénalisation gibbsienne sur les amplitudes complexes ; (d) Pénalisation gibbsienne sur les modules ; (e) Variante de (d) ; (f) Spectre de référence

d'après la convexité composante par composante de g . En prenant l'image par f , et en utilisant successivement la croissance de f pour ses m coordonnées,

$$\begin{aligned} f(g(\alpha x + (1 - \alpha)y)) &\leq f(\alpha g(x) + (1 - \alpha)g(y)), \\ &\leq \alpha f(g(x)) + (1 - \alpha)f(g(y)), \end{aligned}$$

car f est convexe. ■

Pour appliquer ce résultat à $J(|X|)$, posons $g_k(X_1, X_2) = |X_k|$ ($k \in \{1, 2\}$) et $f(\rho_1, \rho_2) = J_0(\rho_1) + J_0(\rho_2) + \mu J_1(P(\rho_1 - \rho_2))$, avec $\mu = 2\lambda_1/\lambda_0$. L'application du Théorème 5.2 conduit à la condition

$$\forall \rho_0, \rho_1 > 0, \quad J'_0(\rho_0) > \mu P J'_1(\rho_1),$$

où J_0 et J_1 sont des fonctions convexes symétriques supposées dérivables. Par exemple, cette condition est satisfaite par la fonction

$$f(\rho_1, \rho_2) = \rho_1 + \rho_2 + \mu \sqrt{\alpha + P |\rho_1 - \rho_2|} \quad \text{si } \mu P < 1.$$

La Figure 1d est une application directe de ce résultat. Même si la forme du spectre est assez fidèlement respectée, l'introduction du terme gibbsien pénalise à nouveau la séparation des pics rapprochés. En revanche, une combinaison plus raffinée des deux types de pénalisation, exploitant les mêmes condition de convexité, permet d'obtenir un excellent compromis (Figure 1e) [7].

6 Paramètre de régularisation

Les spectres représentés Figure 1 résultent de la minimisation de critères pénalisés. Dans tous les cas, le paramètre de régularisation a été choisi en sorte que la puissance empirique soit respectée : $\|\widehat{X}\|^2 = \|y\|^2$ (les autres paramètres résultent d'un réglage empirique). Le résultat suivant montre qu'un tel choix est toujours possible.

Théorème 6.1 — Soit J une fonction convexe de \mathbb{R}^P dans \mathbb{R}_+ , dont le minimum unique est atteint en $\mathbf{0}$, et soit $\widehat{X}(\lambda)$ l'argument du minimum de $\|x_K - y_K\|^2 + \lambda J(X)$, supposée strictement convexe (P est un multiple de K , et on note $x_k = \sum_{p=1}^P X_p e^{2i\pi p k/P}$). Alors il existe au moins un λ tel que $\|\widehat{X}(\lambda)\|^2 = \|y_K\|^2$.

Démonstration : Le résultat est la conséquence des trois remarques suivantes :

- $\|\widehat{X}(0^+)\|^2 \geq \|y_K\|^2$ car $\widehat{X}(0^+)$ vérifie $x_K = y_K$ (et minimise $J(X)$ sous cette contrainte); or le minimum de $\|\widehat{X}(0^+)\|^2$ sous la contrainte $x_K = y_K$ est atteint par le périodogramme et vaut précisément $\|y_K\|^2$.
- $\|\widehat{X}(\infty)\|^2 = 0$, car $\widehat{X}(\infty)$ minimise $J(X)$, donc vaut $\mathbf{0}$.
- $\widehat{X}(\lambda)$ est une fonction continue de λ [5]. ■

Références

[1] S. F. Gull et G. J. Daniell, « Image reconstruction from incomplete and noisy data », *Nature*, vol. 272, pp. 686–690, 1978.

- [2] A. Tikhonov et V. Arsenin, *Solutions of Ill-Posed Problems*, Winston, Washington DC, 1977.
- [3] M. Nashed, « Operator-theoretic and computational approaches to ill-posed problems with applications to antenna theory », *IEEE Trans. Ant. Propag.*, vol. 29, pp. 220–231, 1981.
- [4] S. Geman et D. Geman, « Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images », *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, 6, pp. 721–741, Nov. 1984.
- [5] C. Bouman et K. Sauer, « A generalized Gaussian image model for edge-preserving MAP estimation », *IEEE Trans. Image Processing*, vol. IP-2, 3, pp. 296–310, July 1993.
- [6] J.-F. Giovannelli et J. Idier, « Une nouvelle approche non-paramétrique de l'imagerie radar Doppler », Rapport de contrat CNRS–Société THOMSON, GPI–LSS, 1995.
- [7] P. Ciuciu, J.-F. Giovannelli, et J. Idier, « Analyse spectrale post-moderne. application aux signaux radars. », Rapport de contrat CNRS–Société THOMSON, GPI–LSS, 1996.