



Méthodes à haute résolution

Roland Badeau

roland.badeau@telecom-paristech.fr



Contexte public } sans modifications

Voir page 23

M2 Mathématiques / Vision / Apprentissage - Analyse des signaux audiofréquences



Avant-propos

Dans le cadre du traitement de la parole et des signaux de musique, la partie tonale d'une grande variété de sons trouve une modélisation efficace comme une somme de sinusoides à paramètres lentement variables. Par exemple, les sons qui produisent une sensation de hauteur bien définie possèdent une forme d'onde quasi périodique (sur une durée supérieure à quelques dizaines de millisecondes). L'analyse de Fourier montre que ces signaux sont composés de sinusoides satisfaisant une relation d'*harmonicité*, ce qui signifie que leurs fréquences sont multiples de la fréquence fondamentale, définie comme l'inverse de la période. C'est notamment le cas des signaux de parole dits *voisés*, produits par vibration quasi périodique des cordes vocales, comme les voyelles. De nombreux instruments à vent ou à corde produisent également des sons harmoniques ou quasi harmoniques. Cependant, dans un signal de musique polyphonique, les sons émis simultanément par un ou plusieurs instruments se superposent ; ainsi la relation d'*harmonicité* n'est plus vérifiée, mais le signal reste essentiellement constitué de sinusoides.

L'estimation de sinusoides est un problème classique, vieux de plus de deux cents ans. Dans ce domaine, la transformation de Fourier est un outil privilégié en raison de sa robustesse, de la simplicité de sa mise en oeuvre, et de l'existence d'algorithmes rapides (FFT). Elle présente néanmoins un certain nombre d'inconvénients. Tout d'abord, sa *précision* fréquentielle, c'est-à-dire la précision avec laquelle la fréquence d'une sinusoides peut être estimée, est limitée par le nombre d'échantillons utilisés pour la calculer. Cette première limitation peut être contournée en prolongeant le signal utile par une suite de zéros. Cependant, sa *résolution* fréquentielle, c'est-à-dire sa capacité à distinguer deux sinusoides proches, reste limitée par la durée du signal observé. Malgré ces inconvénients, la transformation de Fourier demeure l'outil le plus utilisé en analyse spectrale. Elle a donné lieu à de nombreuses méthodes d'estimation de fréquences [1].

Les méthodes dites à Haute Résolution (HR), qui trouvent leurs applications en traitement d'antenne comme en analyse spectrale [2], présentent l'avantage de s'affranchir des limitations naturelles de l'analyse de Fourier. En effet, en l'absence de bruit, leur précision et leur résolution fréquentielles sont virtuellement infinies (bien qu'en pratique limitées par la précision finie des machines de calcul). Ceci est rendu possible en s'appuyant fortement sur un modèle paramétrique de signal. Ainsi, contrairement à l'analyse de Fourier qui consiste à représenter le signal dans un domaine transformé, les méthodes HR sont des méthodes d'estimation paramétrique. Dans le cadre du traitement du signal audio, malgré leur supériorité en terme de résolution spectrale (en particulier sur des fenêtres temporelles courtes), elles restent peu utilisées en raison de leur forte complexité algorithmique. Néanmoins, les méthodes HR sont bien adaptées pour estimer les paramètres d'une somme de sinusoides dont l'amplitude varie exponentiellement (modèle ESM). Ce type de modulation permet de décrire l'amortissement naturel des systèmes vibratoires libres, tels que la vibration d'une corde pincée [3]. D'autre part, il a été montré dans [4] que les méthodes HR se révèlent particulièrement performantes dans le cas de signaux fortement atténués. D'une manière plus générale, le modèle ESM permet de décrire des signaux à forte variation d'amplitude [5]. Par ailleurs, les signaux de musique contiennent souvent des paires ou des triplets de fréquences très proches qui engendrent un phénomène de battements. Ces battements contribuent fortement à l'aspect naturel du son. Ils résultent souvent des propriétés particulières des systèmes de vibration. Par exemple, une dissymétrie mineure dans la géométrie d'une cloche conduit à des paires de modes de vibration. Dans le cas d'une guitare, le couplage entre les cordes et le chevalet peut être représenté par une matrice dite de mobilité, dont il est possible de déduire des paires de fréquences [6]. Dans le cas du piano, le couplage des modes de vibration horizontal et vertical de chaque corde et la présence de paires ou de triplets de cordes pour la plupart des notes expliquent la présence de quatre ou six fréquences voisines au niveau de chaque harmonique [7]. L'analyse de Fourier ne permet généralement pas de distinguer toutes ces fréquences. Les études menées dans [4] sur des sons de piano et de guitare ont montré la supériorité des méthodes HR dans ce domaine. La même technique a été utilisée pour estimer des paramètres physiques, comme le facteur de rayonnement d'une guitare [8], et pour étudier la propagation d'ondes mécaniques dans des matériaux solides [9].



Table des matières

1	Estimation de sinusoides	3
1.1	Modèle de signal	4
1.2	Méthode du maximum de vraisemblance	5
1.2.1	Application du principe du maximum de vraisemblance au modèle ESM	5
1.2.2	Maximum de vraisemblance et résolution de Fourier	6
1.3	Méthodes à haute résolution	7
1.3.1	Techniques de prédiction linéaire	8
1.3.1.1	Equations de récurrence linéaires	8
1.3.1.2	Méthode de Prony	9
1.3.1.3	Méthode de Pisarenko	10
1.3.2	Méthodes sous-espace	11
1.3.2.1	Structure singulière de la matrice de données	11
1.3.2.2	Structure singulière de la matrice de corrélation	12
1.3.2.3	Complément : analogie entre le spectre au sens matriciel et le spectre au sens de Fourier	13
1.3.2.4	MULTiple SIGNAL Characterization (MUSIC)	13
1.3.2.5	Estimation of Signal Parameters via Rotational Invariance Techniques	14
1.4	Estimation des autres paramètres	15
1.4.1	Estimation de l'ordre de modélisation	15
1.4.2	Estimation des amplitudes, des phases et de l'écart-type du bruit	16
1.5	Performances des estimateurs	16
1.5.1	Borne de Cramer-Rao	16
1.5.2	Performances des méthodes HR	17
1.6	Conclusion	18
1.7	Annexes	18
1.7.1	Optimisation sous contraintes	18
1.7.2	Matrices de Vandermonde	19
	Licence de droits d'usage	23



Chapitre 1

Estimation de sinusoides

Ce chapitre est consacré à l'estimation des paramètres d'un signal composé d'une somme de sinusoides modulées exponentiellement et perturbées par un bruit additif. Le principe du maximum de vraisemblance ramène alors l'estimation des amplitudes et des phases à un problème de moindres carrés simple, alors que l'estimation des fréquences et des facteurs d'atténuation exige des méthodes plus sophistiquées, dites à *haute résolution*, car elles s'affranchissent des limites de l'analyse de Fourier en terme de résolution spectrale.

L'origine des méthodes HR remonte aux travaux de Prony publiés en 1795, qui visent à estimer une somme d'exponentielles par des techniques de prédiction linéaire [10]. Plus récemment, cette approche a été approfondie par Pisarenko pour estimer des sinusoides d'amplitude constante [11]. En comparaison, les méthodes HR modernes reposent sur les propriétés particulières de la matrice de covariance du signal. Ainsi, l'étude de son rang permet de séparer l'espace des données en deux sous-espaces, l'espace signal engendré par les sinusoides, et l'espace bruit qui est son complémentaire orthogonal. Les méthodes HR issues de cette décomposition en sous-espaces sont connues pour être plus robustes que les techniques de prédiction linéaire. C'est le cas des méthodes MUSIC [12] et root-MUSIC [13] (qui reposent sur l'espace bruit), de l'algorithme TAM [14], ainsi que de l'algorithme ESPRIT [15] et de ses variantes TLS-ESPRIT [16] et PRO-ESPRIT [17] (qui reposent sur l'espace signal). Toutes ces méthodes d'estimation peuvent être appliquées au modèle de signal ESM (Exponential Sinusoidal Model), qui représente le signal comme une somme de sinusoides modulées exponentiellement. Ce modèle est également baptisé EDS (Exponentially Damped Sinusoids) quand la modulation est décroissante [18]. D'autres techniques d'estimation ont été spécifiquement développées pour le modèle ESM, telles que l'algorithme de KT, encore appelé méthode Min-Norm [19], sa version modifiée MKT [20] (de type prédiction linéaire), et la méthode *Matrix Pencil* [21] (de type sous-espace). Une liste plus complète de ces méthodes peut être consultée dans [22].

Ce chapitre n'a pas vocation à présenter les méthodes HR de façon exhaustive, mais plutôt à familiariser le lecteur avec les concepts sur lesquels elles reposent. C'est pourquoi seules certaines d'entre elles sont présentées ici : les méthodes de Prony, de Pisarenko, MUSIC et ESPRIT. Cet exposé débutera par la définition du modèle de signal (section 1.1). Puis la méthode du maximum de vraisemblance, qui permet d'établir un lien avec la transformation de Fourier, sera présentée dans la section 1.2. Ensuite les méthodes à haute résolution pour estimer les pôles complexes seront introduites dans la section 1.3, et des techniques d'estimation des autres paramètres du modèle seront présentées dans la section 1.4. La section 1.5 sera consacrée à l'analyse des performances des méthodes HR. Enfin, les résultats de ce chapitre seront résumés dans la section 1.6.



Notations

\mathbf{v}	vecteur
\mathbf{M}	matrice
$(.)^T$	transposé
$(.)^H$	conjugué hermitien
$\mathcal{R}e(.)$	partie réelle
$\mathcal{I}m(.)$	partie imaginaire
$\text{Im}(.)$	espace image d'une matrice
$\text{Ker}(.)$	noyau d'une matrice
$\text{dim}(.)$	dimension d'un espace vectoriel
$\text{rang}(.)$	rang d'une matrice
$\text{trace}(.)$	trace d'une matrice carrée
$\text{det}(.)$	déterminant d'une matrice carrée
$\text{diag}(.)$	matrice diagonale définie par ses coefficients diagonaux
$\mathbb{E}[.]$	espérance mathématique
$\hat{(.)}$	estimateur d'un paramètre
$\text{CRB}\{.\}$	borne de Cramér-Rao

1.1 Modèle de signal

Considérons le modèle de signal discret (défini pour tout $t \in \mathbb{Z}$)

$$s(t) = \sum_{k=0}^{K-1} \alpha_k z_k^t \quad (1.1)$$

où $K \in \mathbb{N}^*$, $\forall k \in \{0 \dots K-1\}$, $\alpha_k \in \mathbb{C}^*$, et tous les pôles $z_k \in \mathbb{C}^*$ sont distincts. Dans le cas particulier où tous les pôles appartiennent au cercle unité, le signal est représenté comme une somme de sinusoides complexes. Ainsi, chaque pôle z_k s'écrit sous la forme $z_k = e^{i2\pi f_k}$ où $f_k \in \mathbb{R}$ est la fréquence de la sinusoïde. Plus généralement, si les pôles ne se trouvent pas sur le cercle unité, les sinusoides sont modulées exponentiellement (modèle ESM). Dans ce cas, chaque pôle z_k s'écrit sous forme polaire $z_k = e^{\delta_k} e^{i2\pi f_k}$, où $\delta_k \in \mathbb{R}$ est le facteur d'atténuation (ou taux d'amortissement) de la sinusoïde. En particulier, les pôles de même angle polaire et de modules différents sont associés à la même fréquence. Les amplitudes complexes α_k s'écrivent également sous forme polaire $\alpha_k = a_k e^{i\phi_k}$, où $a_k \in \mathbb{R}_+^*$ et $\phi \in \mathbb{R}$.

En outre, le signal observé $x(t)$ peut être modélisé comme la somme du signal déterministe $s(t)$ défini ci-dessus et d'un bruit blanc gaussien complexe centré $b(t)$ de variance σ^2 . Rappelons qu'un bruit blanc gaussien complexe centré est une suite de variables aléatoires *i.i.d* à valeurs complexes, de densité de probabilité $p(b) = \frac{1}{\pi\sigma^2} e^{-\frac{|b|^2}{\sigma^2}}$. On obtient ainsi la relation

$$x(t) = s(t) + b(t). \quad (1.2)$$

Le signal est observé sur des fenêtres temporelles de longueur $N \geq K$. Ainsi, pour tout $t \in \mathbb{Z}$, on considère la fenêtre temporelle $\{t-l+1 \dots t+n-1\}$, où les entiers n et l sont tels que $N = n+l-1$, et on définit le vecteur $\mathbf{s}(t) = [s(t-l+1), \dots, s(t+n-1)]^T$, de dimension N . Pour tout $z \in \mathbb{C}$, posons $\mathbf{v}(z) = [1, z, \dots, z^{N-1}]^T$. Or $\mathbf{s}(t) = \sum_{k=0}^{K-1} \alpha_k z_k^{t-l+1} \mathbf{v}(z_k)$. Cette égalité peut être réécrite sous la forme d'un produit : $\mathbf{s}(t) = \mathbf{V}^N \mathbf{D}^{t-l+1} \boldsymbol{\alpha}$, où $\boldsymbol{\alpha} = [\alpha_0, \dots, \alpha_{K-1}]^T$ est un vecteur de dimension K , $\mathbf{D} = \text{diag}(z_0, \dots, z_{K-1})$ est une matrice diagonale de dimension $K \times K$, et $\mathbf{V}^N = [\mathbf{v}(z_0), \dots, \mathbf{v}(z_{K-1})]$ est une matrice de Vandermonde de dimensions $N \times K$: (cf. définition 2 de l'annexe 1.7.2 page 19)

$$\mathbf{V}^N = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_0 & z_1 & \dots & z_{K-1} \\ \vdots & \vdots & \vdots & \vdots \\ z_0^{N-1} & z_1^{N-1} & \dots & z_{K-1}^{N-1} \end{bmatrix}.$$



Définissons alors le vecteur d'amplitudes à l'instant t , $\alpha(t) = \mathbf{D}^{t-l+1}\alpha$, de sorte que $s(t) = \mathbf{V}^N \alpha(t)$. Il est connu que la matrice de Vandermonde carrée \mathbf{V}^K extraite des K premières lignes de \mathbf{V}^N (rappelons que $N \geq K$) a pour déterminant (cf. proposition 9 de l'annexe 1.7.2 page 19)

$$\det(\mathbf{V}^K) = \prod_{0 \leq k_1 < k_2 \leq K-1} (z_{k_2} - z_{k_1}). \quad (1.3)$$

Ainsi, la matrice \mathbf{V}^N est de rang plein si et seulement si tous les pôles sont distincts. La relation $s(t) = \mathbf{V}^N \alpha(t)$ montre donc que pour chaque instant t le vecteur $s(t)$ vit dans l'espace image de la matrice \mathbf{V}^N , de dimension inférieure ou égale à K dans le cas général, et égale à K si tous les pôles sont distincts.

Soit $\mathbf{b}(t) = [b(t-l+1), \dots, b(t+n-1)]^T$ le vecteur contenant les échantillons du bruit additif. Il s'agit d'un vecteur aléatoire gaussien centré, dont la matrice de covariance est $\mathbf{R}_{bb} = \sigma^2 \mathbf{I}_N$. Notons enfin $\mathbf{x}(t) = [x(t-l+1), \dots, x(t+n-1)]^T$ le vecteur de données observées. Ce vecteur vérifie donc $\mathbf{x}(t) = s(t) + \mathbf{b}(t)$. Le modèle étant posé, l'analyse du signal $s(t)$ va consister à estimer les paramètres σ^2 , z_0, \dots, z_{K-1} et $\alpha(t)$. Une technique classique d'estimation paramétrique, la méthode du maximum de vraisemblance, est appliquée à ce modèle dans la prochaine section.

1.2 Méthode du maximum de vraisemblance

Le principe du maximum de vraisemblance est une méthode générale d'estimation de paramètres. Il fournit des estimateurs asymptotiquement efficaces et sans biais. C'est pourquoi il est souvent préféré aux autres techniques d'estimation lorsqu'il possède une solution analytique simple.

1.2.1 Application du principe du maximum de vraisemblance au modèle ESM

Le principe du maximum de vraisemblance consiste à maximiser la probabilité conditionnelle d'observer le signal x sur l'intervalle $\{t-l+1, \dots, t+n-1\}$, connaissant les paramètres σ^2 , z_0, \dots, z_{K-1} et $\alpha(t)$ (ou le logarithme népérien de cette probabilité, appelé *log-vraisemblance* des observations). Puisque $\mathbf{x}(t) = s(t) + \mathbf{b}(t)$, où $s(t) = \mathbf{V}^N \alpha(t)$ est un vecteur déterministe et $\mathbf{b}(t)$ est un vecteur aléatoire gaussien complexe centré de matrice de covariance $\mathbf{R}_{bb} = \sigma^2 \mathbf{I}_N$, $\mathbf{x}(t)$ est lui-même un vecteur aléatoire gaussien complexe d'espérance $s(t)$ et de matrice de covariance \mathbf{R}_{bb} . Rappelons que la densité de probabilité d'un tel vecteur aléatoire est

$$p(\mathbf{x}(t)) = \frac{1}{\pi^N \det(\mathbf{R}_{bb})} e^{-(\mathbf{x}(t)-s(t))^H \mathbf{R}_{bb}^{-1} (\mathbf{x}(t)-s(t))}.$$

Ainsi, la log-vraisemblance des observations est

$$L(\sigma^2, z_0 \dots z_{K-1}, \alpha(t)) = -N \ln(\pi \sigma^2) - \frac{1}{\sigma^2} g(z_0 \dots z_{K-1}, \alpha(t))$$

où

$$g(z_0 \dots z_{K-1}, \alpha(t)) = (\mathbf{x}(t) - \mathbf{V}^N \alpha(t))^H (\mathbf{x}(t) - \mathbf{V}^N \alpha(t)).$$

La maximisation de cette log-vraisemblance par rapport aux paramètres $(\sigma^2, z_0 \dots z_{K-1}, \alpha(t))$ peut être effectuée en minimisant d'abord g par rapport au couple $(z_0 \dots z_{K-1}, \alpha(t))$, puis en maximisant L par rapport à σ . On obtient ainsi $\sigma^2 = \frac{1}{N} g(z_0 \dots z_{K-1}, \alpha(t))$, ou encore

$$\sigma^2 = \frac{1}{N} \|\mathbf{x}(t) - \mathbf{V}^N \alpha(t)\|^2. \quad (1.4)$$

Il apparaît que σ^2 est estimé en calculant la puissance du résiduel obtenu en soustrayant les exponentielles du signal observé.

La matrice \mathbf{V}^N est de rang plein, puisqu'il a été supposé dans la section 1.1 que les pôles sont distincts deux à deux. Ainsi, la matrice $\mathbf{V}^{NH}\mathbf{V}^N$ est inversible. Pour minimiser g par rapport au couple $(z_0 \dots z_{K-1}, \alpha(t))$, il suffit d'utiliser la décomposition

$$g(z_0 \dots z_{K-1}, \alpha(t)) = \mathbf{x}(t)^H \mathbf{x}(t) - \mathbf{x}(t)^H \mathbf{V}^N (\mathbf{V}^{NH} \mathbf{V}^N)^{-1} \mathbf{V}^{NH} \mathbf{x}(t) + (\alpha(t) - (\mathbf{V}^{NH} \mathbf{V}^N)^{-1} \mathbf{V}^{NH} \mathbf{x}(t))^H (\mathbf{V}^{NH} \mathbf{V}^N) (\alpha(t) - (\mathbf{V}^{NH} \mathbf{V}^N)^{-1} \mathbf{V}^{NH} \mathbf{x}(t)).$$

Le dernier terme de cette équation est toujours positif, et peut être rendu nul en posant

$$\alpha(t) = (\mathbf{V}^{NH} \mathbf{V}^N)^{-1} \mathbf{V}^{NH} \mathbf{x}(t). \quad (1.5)$$

Il apparaît que le vecteur des amplitudes complexes $\alpha(t)$ est estimé de la même façon qu'en utilisant la méthode des moindres carrés ordinaire.

La fonction g est donc minimale quand le K -uplet $(z_0 \dots z_{K-1})$ maximise la fonction \mathcal{J} définie par

$$\mathcal{J}(z_0, \dots, z_{K-1}) = \mathbf{x}(t)^H \mathbf{V}^N (\mathbf{V}^{NH} \mathbf{V}^N)^{-1} \mathbf{V}^{NH} \mathbf{x}(t). \quad (1.6)$$

Comme ce problème d'optimisation ne possède pas de solution analytique dans le cas général, il doit être résolu numériquement. En résumé, le principe du maximum de vraisemblance conduit à estimer les paramètres du modèle en trois étapes :

les pôles complexes sont obtenus en effectuant la maximisation de la fonction \mathcal{J} (équation (1.6)),

les amplitudes complexes sont obtenues en calculant le membre de droite de l'équation (1.5),

l'écart-type est alors donné par l'équation (1.4).

Malheureusement, il se trouve que la première étape de cette méthode d'estimation, qui requiert l'optimisation d'une fonction de K variables complexes, est difficile à implémenter, car la fonction à maximiser possède de nombreux maxima locaux. De plus, elle s'avère extrêmement coûteuse en temps de calcul. C'est pourquoi on utilise généralement des méthodes plus fiables et plus rapides pour estimer les pôles complexes. Néanmoins, une fois les pôles estimés, le principe du maximum de vraisemblance peut être utilisé pour déterminer les amplitudes complexes et l'écart-type du bruit.

1.2.2 Maximum de vraisemblance et résolution de Fourier

Intéressons-nous maintenant au cas particulier où tous les pôles sont sur le cercle unité ($\forall k, \delta_k = 0$). Les résultats de la section 1.2.1 ont montré que le principe du maximum de vraisemblance conduit à un problème d'optimisation qui ne possède pas de solution analytique simple dans le cas général. Cependant, une telle solution existe dans le cas particulier où $K = 1$, ainsi qu'une solution approchée si $K > 1$.

Examinons d'abord le cas d'une seule exponentielle complexe ($K = 1$). Alors l'équation (1.6) se simplifie en $\mathcal{J}(z_0) = \widehat{R}_x(z_0)$, où \widehat{R}_x est le périodogramme du signal $x(t)$ observé sur la fenêtre temporelle $\{t - l + 1 \dots t + n - 1\}$:

$$\widehat{R}_x(e^{i2\pi f_0}) = \frac{1}{N} |X(e^{i2\pi f_0})|^2$$

où $X(e^{i2\pi f_0}) = \mathbf{v}(e^{i2\pi f_0})^H \mathbf{x}(t) = \sum_{\tau=0}^{N-1} x(t - l + 1 + \tau) e^{-i2\pi f_0 \tau}$. De même, l'équation (1.5) se simplifie en $\alpha_0(t) = \frac{1}{N} X(e^{i2\pi f_0})$. Enfin, l'équation (1.4) se simplifie en $\sigma^2 = \frac{1}{N} (\|\mathbf{x}(t)\|^2 - \widehat{R}_x(e^{i2\pi f_0}))$.

Ces résultats induisent la conclusion suivante :

Le principe du maximum de vraisemblance conduit dans le cas d'une sinusoïde complexe à détecter la fréquence pour laquelle le périodogramme atteint son maximum. L'amplitude complexe correspondante est proportionnelle à la valeur de la TFD du signal à cette fréquence. La variance du bruit est estimée comme la puissance du signal après soustraction de la sinusoïde.

Abordons maintenant le cas général $K \geq 1$, pour lequel la maximisation de la fonction $\mathcal{J}(z)$ ne possède plus de solution analytique exacte. On introduit alors l'hypothèse suivante :

$$N \gg \frac{1}{\min_{k_1 \neq k_2} |f_{k_2} - f_{k_1}|}.$$

La matrice $\mathbf{V}^{NH} \mathbf{V}^N$ est une matrice hermitienne définie positive de dimension $K \times K$, dont les coefficients peuvent être calculés analytiquement : $\{\mathbf{V}^{NH} \mathbf{V}^N\}_{(k_1, k_2)} = \sum_{\tau=0}^{N-1} (z_{k_1}^* z_{k_2})^\tau$. On obtient alors

$$\begin{aligned} \frac{1}{N} \{\mathbf{V}^{NH} \mathbf{V}^N\}_{(k_1, k_2)} &= e^{i\pi(N-1)(f_{k_2} - f_{k_1})} \frac{\sin(\pi N(f_{k_2} - f_{k_1}))}{N \sin(\pi(f_{k_2} - f_{k_1}))} & \text{si } k_1 \neq k_2 \\ \frac{1}{N} \{\mathbf{V}^{NH} \mathbf{V}^N\}_{(k, k)} &= 1 & \text{si } k_1 = k_2 = k \end{aligned}$$

Ainsi, quand $N \gg \frac{1}{\min_{k_1 \neq k_2} |f_{k_2} - f_{k_1}|}$, $\frac{1}{N} \mathbf{V}^{NH} \mathbf{V}^N = \mathbf{I}_K + O\left(\frac{1}{N}\right)$, donc

$$(\mathbf{V}^{NH} \mathbf{V}^N)^{-1} = \frac{1}{N} \mathbf{I}_K + O\left(\frac{1}{N^2}\right).$$

Alors l'équation (1.6) se simplifie en

$$\mathcal{J}(z_0, \dots, z_{K-1}) = \frac{1}{N} \|\mathbf{V}^{NH} \mathbf{x}(t)\|^2 + O\left(\frac{1}{N^2}\right) = \sum_{k=0}^{K-1} \widehat{R}(z_k) + O\left(\frac{1}{N^2}\right).$$

De même, l'équation (1.5) se simplifie en $\alpha(t) = \frac{1}{N} \mathbf{V}^{NH} \mathbf{x}(t) + O\left(\frac{1}{N^2}\right)$, d'où

$$\alpha_k(t) = \frac{1}{N} X(e^{i2\pi f_k}) + O\left(\frac{1}{N^2}\right).$$

Enfin, l'équation (1.4) se simplifie en $\sigma^2 = \frac{1}{N} \left(\|\mathbf{x}(t)\|^2 - \sum_{k=0}^{K-1} \widehat{R}(e^{i2\pi f_k}) \right) + O\left(\frac{1}{N^2}\right)$.

Ainsi, la maximisation conjointe de \mathcal{J} par rapport à z_0, \dots, z_{K-1} conduit à déterminer les K fréquences associées aux K plus grandes valeurs du périodogramme. Les amplitudes complexes correspondantes sont proportionnelles à la valeur de la TFD du signal à ces fréquences. Rappelons que ces résultats ne sont valables que si tous les pôles sont sur le cercle unité et reposent sur l'hypothèse $N \gg \frac{1}{\min_{k_1 \neq k_2} |f_{k_2} - f_{k_1}|}$.

On observe ainsi la limite de l'analyse de Fourier en terme de résolution spectrale : les paramètres sont estimés correctement à condition que la longueur de la fenêtre observée soit suffisamment grande devant l'inverse du plus petit écart fréquentiel entre deux pôles voisins. C'est de cette limite que les méthodes HR présentées dans la section 1.3 permettent de s'affranchir. Ainsi, les méthodes HR sont capables de distinguer deux sinusoides proches, que l'analyse de Fourier ne permet pas de discerner. Dans les applications, les méthodes HR pourront être utilisées avec des fenêtres plus courtes que celles que l'on utilise habituellement avec l'analyse de Fourier.

1.3 Méthodes à haute résolution

On commence par introduire ici les méthodes à haute résolution les plus anciennes, qui reposent sur des techniques de prédiction linéaire (section 1.3.1), avant d'aborder dans la section 1.3.2 les méthodes de type sous-espace, plus récentes.



FIGURE 1.1 – Jean Baptiste Joseph FOURIER (1768-1830)

1.3.1 Techniques de prédiction linéaire

Les deux premières méthodes à haute résolution présentées dans ce chapitre reposent sur un résultat fondamental portant sur les équations de récurrence linéaires, présenté dans la section 1.3.1.1.

1.3.1.1 Equations de récurrence linéaires

Soient $p_0 \in \mathbb{C}^*$, $K \in \mathbb{N}^*$ et $\{z_0, \dots, z_{K-1}\}$ K nombres complexes distincts et non nuls. On définit le polynôme de degré K dont le coefficient dominant est p_0 et dont les racines sont les z_k :

$$P[z] = p_0 \prod_{k=0}^{K-1} (z - z_k) = \sum_{\tau=0}^K p_{K-\tau} z^\tau.$$

Le théorème suivant caractérise le modèle de signal.

Theorème 1 *Un signal discret complexe $\{s(t)\}_{t \in \mathbb{Z}}$ satisfait l'équation de récurrence*

$$\sum_{\tau=0}^K p_\tau s(t - \tau) = 0 \quad (1.7)$$

pour tout $t \in \mathbb{Z}$ si et seulement si il existe des scalaires $\alpha_0, \dots, \alpha_{K-1} \in \mathbb{C}$ tels que $s(t) = \sum_{k=0}^{K-1} \alpha_k z_k^t$.

Démonstration Tout d'abord, il est immédiat de vérifier que l'ensemble des signaux qui satisfont la relation (1.7) constitue un espace vectoriel E sur \mathbb{C} . Ensuite, nous allons démontrer que cet espace vectoriel est de dimension inférieure ou égale à K . Considérons l'application

$$\begin{aligned} f: E &\rightarrow \mathbb{C}^K \\ s[t] &\mapsto [s[0], \dots, s[K-1]]^T \end{aligned}$$

On peut remarquer que f est une application linéaire. Soit un signal $s \in E$ tel que $f(s) = \mathbf{0}$. Alors s est nul sur l'intervalle $[0, K-1]$. En utilisant la récurrence (1.7), on en déduit que s est également nul sur l'intervalle $[K, +\infty[$. Enfin, en utilisant la récurrence (1.7) et le fait que $p_K \neq 0$, on montre que s est également nul sur l'intervalle $]-\infty, -1]$. Par conséquent, $s \equiv 0$, donc l'application linéaire f est injective. On en conclut que l'espace vectoriel E est au plus de dimension K .

A présent, nous allons montrer que tout signal de la forme $s[t] = z_k^t$ où $k \in \{0, \dots, K-1\}$ appartient à l'espace vectoriel E . En effet, si $s[t] = z_k^t$, alors $\forall t \in \mathbb{Z}$, $\sum_{k=0}^K p_k s[t-k] = z_k^{t-K} \sum_{k=0}^K p_k z_k^{K-k} = z_k^{t-K} P[z_k] = 0$, donc $s[t]$ satisfait la relation (1.7).

Finalement, considérons la famille des vecteurs $\{z_k^t\}_{k \in \{0, \dots, K-1\}}$. La matrice carrée dont les colonnes sont extraites de ces vecteurs et dont les lignes correspondent aux instants $\{0 \dots K-1\}$ est une matrice de Vandermonde (cf. définition 2 de l'annexe 1.7.2 page 19). D'après la proposition 9 de l'annexe 1.7.2, elle est inversible, puisque les pôles z_k sont distincts deux à deux. Par conséquent, la famille $\{z_k^t\}_{k \in \{0, \dots, K-1\}}$ est libre. Or elle contient précisément K vecteurs de E . Cet espace est donc exactement de dimension K , et cette famille en constitue une base. Ainsi, un signal $s[t]$ appartient à E si et seulement si il est de la forme (1.2). \square

1.3.1.2 Méthode de Prony



FIGURE 1.2 – Gaspard-Marie RICHE de PRONY (1755-1839)

Les travaux du baron de Prony sont à l'origine du développement des méthodes à haute résolution. Celui-ci a proposé une méthode d'estimation s'inspirant du résultat précédent sur les équations de récurrence linéaires [10]. Cette méthode était originellement destinée à estimer des exponentielles réelles non bruitées ; cependant nous l'appliquons ici à l'estimation d'exponentielles complexes bruitées. La méthode de Prony consiste à déterminer dans un premier temps le polynôme $P[z]$ à l'aide de techniques de prédiction linéaire, puis à extraire les racines de ce polynôme. On définit l'erreur de prédiction

$$\varepsilon(t) \triangleq \sum_{\tau=0}^K p_{\tau} x(t-\tau). \quad (1.8)$$

En particulier, en substituant les équations (1.2) et (1.7) dans l'équation (1.8), on obtient $\varepsilon(t) = \sum_{\tau=0}^K p_{\tau} b(t - \tau)$. L'erreur de prédiction caractérise donc uniquement le bruit qui se superpose au signal. Plaçons-nous dans le cas particulier $n = K + 1$, et supposons que $l \geq K + 1$. Ainsi, le signal est observé sur la fenêtre $\{t - l + 1 \dots t + K\}$. En appliquant l'équation (1.8) aux instants $\{t - l + K + 1, t - l + K + 2, \dots, t + K\}$, on obtient le système d'équations

$$\begin{cases} p_0 x(t - l + K + 1) + p_1 x(t - l + K) + \dots + p_K x(t - l + 1) = \varepsilon(t - l + K + 1) \\ p_0 x(t - l + K + 2) + p_1 x(t - l + K + 1) + \dots + p_K x(t - l + 2) = \varepsilon(t - l + K + 2) \\ \vdots + \vdots + \dots + \vdots = \vdots \\ p_0 x(t + K) + p_1 x(t + K - 1) + \dots + p_K x(t) = \varepsilon(t + K) \end{cases} \quad (1.9)$$

Posons alors $\mathbf{p} = [p_K, p_{(K-1)}, \dots, p_0]^H$, $\boldsymbol{\varepsilon}(t) = [\varepsilon(t - l + K + 1), \varepsilon(t - l + K + 2), \dots, \varepsilon(t + K)]^H$ et

$$\mathbf{X}(t) = \begin{bmatrix} x(t - l + 1) & \dots & x(t - 1) & x(t) \\ x(t - l + 2) & \dots & x(t) & x(t + 1) \\ \vdots & \dots & \vdots & \vdots \\ x(t - l + K + 1) & \dots & x(t + K - 1) & x(t + K) \end{bmatrix} \quad (1.10)$$

de sorte que le système d'équations (1.9) peut être condensé sous la forme $\mathbf{p}^H \mathbf{X}(t) = \boldsymbol{\varepsilon}(t)^H$.

La méthode de Prony consiste à minimiser la puissance de l'erreur de prédiction $\frac{1}{l} \|\boldsymbol{\varepsilon}\|^2$ par rapport à \mathbf{p} , sous la contrainte $p_0 = 1$. Or il est possible d'écrire $\frac{1}{l} \|\boldsymbol{\varepsilon}\|^2 = \mathbf{p}^H \widehat{\mathbf{R}}_{xx}(t) \mathbf{p}$, où la matrice $\widehat{\mathbf{R}}_{xx}(t) = \frac{1}{l} \mathbf{X}(t) \mathbf{X}(t)^H$ est de dimension $(K + 1) \times (K + 1)$. Comme la matrice $\mathbf{X}(t)$ possède $K + 1$ lignes et $l \geq K + 1$ colonnes, on peut supposer que la matrice $\widehat{\mathbf{R}}_{xx}(t)$ est inversible.

Le théorème 8 de l'annexe 1.7.1 page 18 permet de démontrer¹ que la solution de ce problème d'optimisation est²

$$\mathbf{p} = \frac{1}{\mathbf{e}_1^H \widehat{\mathbf{R}}_{xx}(t)^{-1} \mathbf{e}_1} \widehat{\mathbf{R}}_{xx}(t)^{-1} \mathbf{e}_1$$

où $\mathbf{e}_1 \triangleq [1, 0 \dots 0]^T$ est un vecteur de dimension $K + 1$. Ainsi, la méthode d'estimation de Prony comprend les étapes suivantes :

- Construire la matrice $\mathbf{X}(t)$ et calculer $\widehat{\mathbf{R}}_{xx}(t)$;
- Calculer $\mathbf{p} = \frac{1}{\mathbf{e}_1^H \widehat{\mathbf{R}}_{xx}(t)^{-1} \mathbf{e}_1} \widehat{\mathbf{R}}_{xx}(t)^{-1} \mathbf{e}_1$;
- Déterminer les pôles $\{z_0, \dots, z_{K-1}\}$ en tant que racines du polynôme $P[z] = \sum_{k=0}^K p_k z^{K-k}$.

1.3.1.3 Méthode de Pisarenko

La méthode de Pisarenko est une variante de la méthode de Prony. Elle consiste à minimiser la puissance de l'erreur de prédiction $\frac{1}{l} \|\boldsymbol{\varepsilon}\|^2 = \mathbf{p}^H \widehat{\mathbf{R}}_{xx}(t) \mathbf{p}$ sous la contrainte que le vecteur \mathbf{p} soit de norme 1. Le théorème 8 de l'annexe 1.7.1 page 18 permet de démontrer que la solution de ce problème d'optimisation est le vecteur propre de la matrice $\widehat{\mathbf{R}}_{xx}(t)$ associé à la plus petite valeur propre.

Ainsi la méthode de Pisarenko [11] consiste à

- calculer et diagonaliser $\widehat{\mathbf{R}}_{xx}(t)$;
- déterminer \mathbf{p} comme le vecteur propre associé à la plus petite valeur propre ;
- extraire les racines du polynôme $P[z]$.

Les méthodes de Prony et de Pisarenko constituent les plus anciennes méthodes HR. Comme nous le montrons dans la section 1.5.2, elles s'avèrent peu robustes en pratique, c'est pourquoi les méthodes de type sous-espace, proposées plus récemment, leur sont généralement préférées.

1. Comme les données sont complexes, il est nécessaire de décomposer préalablement le vecteur \mathbf{p} en sa partie réelle et sa partie imaginaire pour pouvoir appliquer le théorème 8, qui traite exclusivement le cas de données réelles.

2. Le scalaire $\mathbf{e}_1^H \widehat{\mathbf{R}}_{xx}(t)^{-1} \mathbf{e}_1$ est non nul, puisque le vecteur \mathbf{e}_1 est unitaire et par hypothèse la matrice $\widehat{\mathbf{R}}_{xx}(t)$ est définie positive.

1.3.2 Méthodes sous-espace

Dans le même esprit que la méthode de Pisarenko, les méthodes HR modernes (par exemple [12, 15, 21]) reposent sur une décomposition de la matrice $\widehat{\mathbf{R}}_{xx}(t)$.

1.3.2.1 Structure singulière de la matrice de données

Supposons désormais que $n \geq K + 1$ et $l \geq K + 1$, et construisons la matrice de données du signal non bruité $s(t)$ sur le même modèle que la matrice $\mathbf{X}(t)$ dans l'équation (1.10), selon une structure de Hankel :

$$\mathbf{S}(t) = \begin{bmatrix} s(t-l+1) & \cdots & s(t-1) & s(t) \\ s(t-l+2) & \cdots & s(t) & s(t+1) \\ \vdots & \cdots & \vdots & \vdots \\ s(t-l+n) & \cdots & s(t+n-2) & s(t+n-1) \end{bmatrix}. \quad (1.11)$$

La proposition suivante caractérise le modèle de signal.

Proposition 2 (Factorisation de la matrice de données) *Les assertions suivantes sont équivalentes :*

1. Le signal $s(t)$ satisfait le modèle défini dans l'équation (1.1) sur l'intervalle $\{t-l+1, \dots, t+n-1\}$;
2. La matrice $\mathbf{S}(t)$ définie dans l'équation (1.11) peut être factorisée sous la forme

$$\mathbf{S}(t) = \mathbf{V}^n \mathbf{A}(t) \mathbf{V}^{lT} \quad (1.12)$$

où la matrice diagonale $\mathbf{A}(t) = \text{diag}(z_0^{t-l+1}\alpha_0, \dots, z_{(K-1)}^{t-l+1}\alpha_{(K-1)})$ est de dimension $K \times K$, \mathbf{V}^n est de dimensions $n \times K$, et \mathbf{V}^l est de dimensions $l \times K$.

Démonstration de la proposition 2 : Démontrons chacune des deux implications.

Démonstration de 1. \Rightarrow 2. Le signal $s(t)$ est défini comme une somme d'exponentielles complexes : $s(t) = \sum_{k=0}^{K-1} s_k(t)$, où $s_k(t) = \alpha_k z_k^t$. Par conséquent, la matrice $\mathbf{S}(t)$ définie dans l'équation (1.11) peut être décomposée de façon analogue : $\mathbf{S}(t) = \sum_{k=0}^{K-1} \mathbf{S}_k(t)$, où les matrices $\mathbf{S}_k(t)$ sont construites de la même façon que $\mathbf{S}(t)$ dans l'équation (1.11), à partir des signaux $s_k(t)$. Or il est immédiat de vérifier que $\mathbf{S}_k(t) = \alpha_k z_k^{t-l+1} \mathbf{v}^n(z_k) \mathbf{v}^l(z_k)^T$, où $\mathbf{v}^n(z) = [1, z, \dots, z^{n-1}]^T$ et $\mathbf{v}^l(z) = [1, z, \dots, z^{l-1}]^T$. L'équation (1.12) en découle directement.

Démonstration de 2. \Rightarrow 1. Si la matrice $\mathbf{S}(t)$ est définie par l'équation (1.12), le cheminement inverse permet de montrer que $\mathbf{S}(t)$ peut s'écrire sous la forme (1.11), où $s(t)$ est défini dans l'équation (1.1). \square

La proposition 2 induit un résultat sur lequel s'appuient toutes les méthodes sous-espace :

Corollaire 3 *La matrice $\mathbf{S}(t)$ définie dans l'équation (1.11) est de rang inférieur ou égal à K . Plus précisément, elle est de rang K si et seulement si $n \geq K$, $l \geq K$, tous les pôles z_k sont distincts et non nuls, et toutes les amplitudes α_k sont non nulles. Dans ce cas, son espace image est engendré par la matrice \mathbf{V}^n .*

Démonstration du corollaire 3 : Il se trouve que les matrices \mathbf{V}^n et \mathbf{V}^l sont de rang plein égal à K . En effet, si \mathbf{V}_0 est la matrice de Vandermonde constituée des K premières lignes de \mathbf{V}^n ou de \mathbf{V}^l , la proposition 9 montre que \mathbf{V}_0 est inversible, puisque les pôles z_k sont distincts deux à deux. Par conséquent, $\text{rang}(\mathbf{V}^n) \geq K$ et $\text{rang}(\mathbf{V}^l) \geq K$. Or $\dim(\mathbf{V}^n) = n \times K$ et $\dim(\mathbf{V}^l) = l \times K$, donc $\text{rang}(\mathbf{V}^n) = \text{rang}(\mathbf{V}^l) = K$. Par ailleurs, la matrice $\mathbf{A}(t)$, de dimensions $K \times K$, est inversible, donc de rang K .

On peut déduire de cette remarque que la matrice $\mathbf{S}(t)$ est également de rang K . Pour cela, montrons dans un premier temps que $\text{Ker}(\mathbf{S}(t)) = \text{Ker}(\mathbf{V}^{lT})$. En effet,

- $\forall \mathbf{y} \in \mathbb{C}^n, \mathbf{V}^{lT} \mathbf{y} = \mathbf{0} \Rightarrow \mathbf{V}^n \mathbf{A}(t) \mathbf{V}^{lT} \mathbf{y} = \mathbf{0}$;
- $\forall \mathbf{y} \in \mathbb{C}^n, \mathbf{V}^n \mathbf{A}(t) \mathbf{V}^{lT} \mathbf{y} = \mathbf{0} \Rightarrow (\mathbf{V}^{nH} \mathbf{V}^n) \mathbf{A}(t) \mathbf{V}^{lT} \mathbf{y} = \mathbf{0}$. Or la matrice $(\mathbf{V}^{nH} \mathbf{V}^n) \mathbf{A}(t)$ est inversible, donc $\mathbf{V}^{lT} \mathbf{y} = \mathbf{0}$.

Le théorème du rang induit alors : $\text{rang}(S(t)) = n - \dim(\text{Ker}(S(t))) = n - \dim(\text{Ker}(V^{lT})) = \text{rang}(V^{lT}) = K$. \square

La structure singulière de la matrice de données induit une structure équivalente pour la matrice de corrélation, définie ci-dessous.

1.3.2.2 Structure singulière de la matrice de corrélation

Les méthodes sous-espace reposent sur la structure particulière de la matrice de corrélation du signal $C_{ss}(t) = S(t)S(t)^H$, et en particulier sur ses sous-espaces propres, que nous allons maintenant étudier. Définissons $R_{ss}(t) = \frac{1}{l} C_{ss}(t)$. L'équation (1.12) montre que

$$R_{ss}(t) = V^n P(t) V^{nH} \quad (1.13)$$

où

$$P(t) = \frac{1}{l} A(t) V^{lT} V^{l*} A(t)^H \quad (1.14)$$

est une matrice symétrique définie positive. Ainsi, l'équation (1.13) montre que sous les mêmes hypothèses que pour $S(t)$, la matrice $R_{ss}(t)$ est de rang K .

L'espace image de la matrice $R_{ss}(t)$, de dimension K , est engendré par la matrice V^n . Cet espace est appelé *espace signal* dans la littérature.

Soit alors $\{w_m\}_{m=0..n-1}$ une base orthonormée de vecteurs propres de la matrice $R_{ss}(t)$, associés aux valeurs propres $\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{n-1} \geq 0$. Comme la matrice $R_{ss}(t)$ est seulement de rang K , on a en fait $\lambda_m = 0 \forall m \geq K$. On note $W(t)$ la matrice $[w_0 \dots w_{K-1}]$, et $W_{\perp}(t)$ la matrice $[w_K \dots w_{n-1}]$. On peut alors vérifier que $\text{Im}(W(t)) = \text{Im}(V^n)$. En effet, les vecteurs $\{w_m\}_{m=0..K-1}$ sont vecteurs propres de la matrice $R_{ss}(t) = V^n P(t) V^{nH}$ associés à des valeurs propres non nulles. Donc $\forall k \in \{0 \dots K-1\}$, $w_m \in \text{Im}(V^n)$. Donc $\text{Im}(W(t)) \subset \text{Im}(V^n)$. Or les matrices $W(t)$ et V^n sont de même rang K , donc $\text{Im}(W(t)) = \text{Im}(V^n)$.

La matrice $W(t)$ est une autre base de l'espace signal, généralement distincte de V^n .

On définit ensuite la matrice $X(t)$ à partir des échantillons du signal bruité $x(t)$, de la même manière que la matrice $S(t)$ dans l'équation (1.11), et on considère la matrice de corrélation

$$C_{xx}(t) = X(t) X(t)^H. \quad (1.15)$$

Posons alors $\widehat{R}_{xx}(t) = \frac{1}{l} C_{xx}(t)$ (comme dans la section 1.3.1.2). Puisque le bruit additif $b(t)$ est blanc et centré, de variance σ^2 , la matrice $R_{xx}(t) = \mathbb{E}[\widehat{R}_{xx}(t)]$ vérifie

$$R_{xx}(t) = R_{ss}(t) + \sigma^2 I_n. \quad (1.16)$$

En utilisant l'équation (1.16), on montre que la famille $\{w_m\}_{m=0..n-1}$ définie ci-dessus est également une base orthonormée de vecteurs propres de la matrice $R_{ss}(t)$, associés aux valeurs propres

$$\bar{\lambda}_m = \begin{cases} \lambda_m + \sigma^2 & \forall m \in \{0, \dots, K-1\} \\ \sigma^2 & \forall m \in \{K, \dots, n-1\} \end{cases}.$$

Ainsi, tous les vecteurs propres de la matrice $R_{ss}(t)$ sont également vecteurs propres de $R_{xx}(t)$, et les valeurs propres correspondantes de $R_{xx}(t)$ sont égales à celles de $R_{ss}(t)$ plus σ^2 . Par conséquent, l'espace signal, défini comme l'espace image de la matrice $R_{ss}(t)$, est également l'espace principal de dimension K de la matrice $R_{xx}(t)$, c'est-à-dire l'espace propre de $R_{xx}(t)$ associé aux K plus grandes valeurs propres, toutes strictement supérieures à σ^2 . Les $n - K$ valeurs propres associées au complémentaire orthogonal de l'espace signal, appelé *espace bruit*, sont toutes égales à σ^2 . Ainsi, il est possible d'estimer l'espace signal et l'espace bruit en calculant l'EVD de la matrice $\widehat{R}_{xx}(t)$, ou encore la SVD de $X(t)$. En juxtaposant les K principaux vecteurs propres ou singuliers de l'une de ces matrices, on obtient ainsi la matrice $W(t) = [w_0 \dots w_{K-1}]$ de dimensions $n \times K$ engendrant l'espace signal, et en juxtaposant les $n - K$ autres vecteurs, on obtient la matrice $W_{\perp}(t) = [w_K \dots w_{n-1}]$ de dimensions $n \times (n - K)$ engendrant l'espace bruit.

L'idée qui consiste à décomposer l'espace des données en deux sous-espaces (signal et bruit) est à l'origine de plusieurs méthodes à haute résolution, parmi lesquelles la méthode MUSIC, présentée dans la section 1.3.2.4, et la méthode ESPRIT, présentée dans la section 1.3.2.5.

1.3.2.3 Complément : analogie entre le spectre au sens matriciel et le spectre au sens de Fourier

On examine ici le cas particulier où tous les pôles sont sur le cercle unité ($\forall k, \delta_k = 0$) et où toutes les fréquences f_k sont à la fois multiples de $\frac{1}{n}$ et de $\frac{1}{l}$ (on pourra considérer les résultats obtenus comme des résultats asymptotiques). On note $X(e^{i2\pi\frac{y}{n}})$ la TFD du signal observé sur la fenêtre $\{t, \dots, t+n-1\}$:

$$X(e^{i2\pi\frac{y}{n}}) = \sum_{\tau=0}^{n-1} x(t+\tau) e^{-i2\pi\frac{y}{n}\tau}.$$

On définit alors le périodogramme du signal de la façon suivante

$$\widehat{R}_x(e^{i2\pi\frac{y}{n}}) = \frac{1}{n} \left| X(e^{i2\pi\frac{y}{n}}) \right|^2.$$

Comme toutes les fréquences f_k sont multiples de $\frac{1}{n}$, le spectre discret $R_x(e^{i2\pi\frac{y}{n}}) \triangleq \mathbb{E}[\widehat{R}_x(e^{i2\pi\frac{y}{n}})]$ vérifie

$$R_x(e^{i2\pi\frac{y}{n}}) = \sigma^2 + \sum_{k=0}^{K-1} a_k^2 \mathbf{1}_{\{e^{i2\pi\frac{y}{n}} = z_k\}}. \quad (1.17)$$

Par ailleurs, comme toutes les fréquences f_k sont multiples de $\frac{1}{l}$, on peut vérifier que

$$\mathbf{P}(t) = \text{diag}(a_0^2, \dots, a_{(K-1)}^2).$$

Par conséquent, $\mathbf{R}_{xx}(t) = \sigma^2 \mathbf{I}_n + \sum_{k=0}^{K-1} a_k^2 \mathbf{v}(z_k) \mathbf{v}(z_k)^H$. On suppose alors sans perte de généralité que les a_k sont rangés par ordre décroissant. Or la famille $\{\mathbf{v}(z_k)\}_{k=0 \dots K-1}$ est orthogonale. Par conséquent $\forall m \in \{0 \dots K-1\}$, $\mathbf{R}_{xx}(t) \mathbf{v}(z_k) = (a_k^2 + \sigma^2) \mathbf{v}(z_k)$. Ainsi, $\{\mathbf{v}(z_k)\}_{k=0 \dots K-1}$ constitue une famille de vecteurs propres de la matrice $\mathbf{R}_{xx}(t)$, associés aux valeurs propres $\{a_k^2 + \sigma^2\}_{k=0 \dots K-1}$. En complétant cette famille avec les autres vecteurs de la base de Fourier des signaux de longueur n , on obtient une base de vecteurs propres de la matrice $\mathbf{R}_{xx}(t)$ (les autres vecteurs propres étant tous associés à la même valeur propre σ^2).

Par conséquent, le spectre des valeurs propres de la matrice $\mathbf{R}_{xx}(t)$ coïncide avec le spectre discret donné dans l'équation (1.17) (dont le périodogramme constitue un estimateur).

1.3.2.4 Multiple Signal Characterization (MUSIC)

La méthode MUSIC, développée par R. O. Schmidt [23], repose sur la remarque suivante : les pôles $\{z_k\}_{k=0 \dots K-1}$ sont les uniques solutions de l'équation

$$\|\mathbf{W}_\perp(t)^H \mathbf{v}(z)\|^2 = 0 \quad (1.18)$$

où $\mathbf{v}(z) = [1, z, \dots, z^{n-1}]^T$. En effet, z est solution si et seulement si $\mathbf{v}(z) \in \text{Im}(\mathbf{W}(t)) = \text{Im}(\mathbf{V}^n)$. Donc tout pôle z_k est solution, et il ne peut y en avoir d'autre car dans le cas contraire l'espace signal serait de dimension strictement plus grande que K . Ainsi, la méthode *root-MUSIC* [13] consiste à

- calculer et diagonaliser la matrice $\widehat{\mathbf{R}}_{xx}(t)$;
- en déduire une base de l'espace bruit $\widehat{\mathbf{W}}_\perp(t)$;
- extraire les racines de l'équation (1.18).

Dans le cas particulier où l'espace bruit est de dimension 1, elle est équivalente à la méthode de Pisarenko présentée dans la section 1.3.1.3.

Dans la pratique, les signaux réels ne correspondent pas rigoureusement au modèle, et l'équation (1.18) n'est pas rigoureusement vérifiée. C'est pourquoi la méthode *spectral-MUSIC* [12] consiste plutôt à rechercher les K pics les plus élevés de la fonction $\widehat{S}(z) = \frac{1}{\|\widehat{\mathbf{W}}_\perp^H \mathbf{v}(z)\|^2}$. La méthode *spectral-MUSIC* est illustrée sur la figure I.3, où elle est appliquée à une note de piano.

La méthode ESPRIT, présentée ci-dessous, permet d'éviter l'optimisation de la fonction $\widehat{S}(z)$, ou la résolution de l'équation (1.18), et fournit les valeurs des pôles complexes d'une manière plus directe.

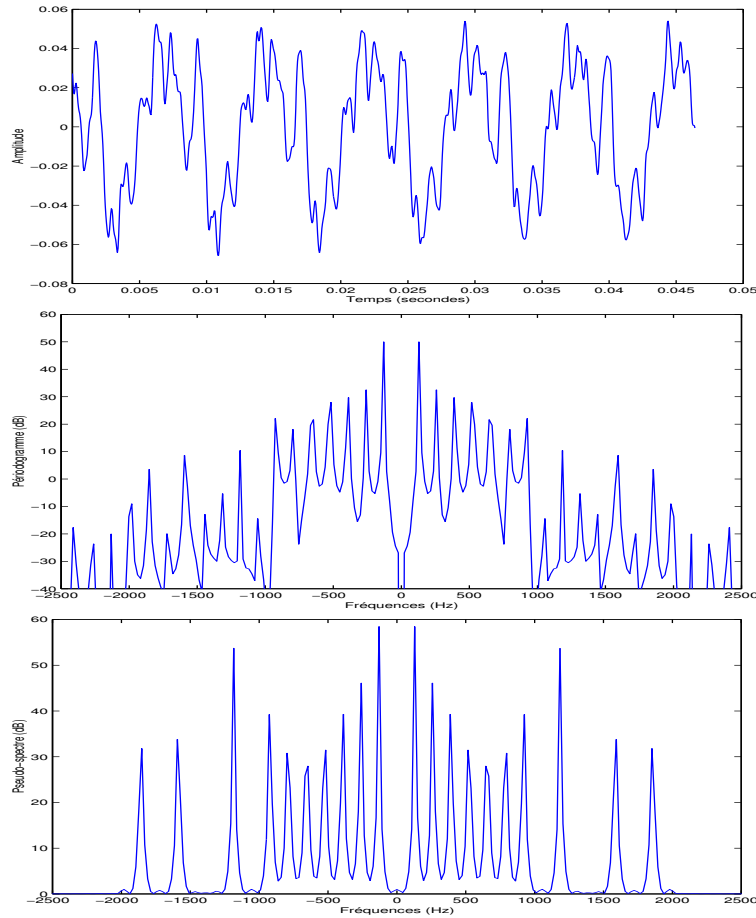


FIGURE 1.3 – Forme d’onde ($t \mapsto x(t)$), périodogramme ($f \mapsto 20 \log_{10} |S(e^{i2\pi f})|$) et pseudo-spectre ($f \mapsto 20 \log_{10} |\widehat{S}(e^{i2\pi f})|$) avec $K = 20$ et $n = 256$ d’une note de piano

1.3.2.5 Estimation of Signal Parameters via Rotational Invariance Techniques

La méthode ESPRIT [15] s’appuie sur une propriété particulière de l’espace signal : l’invariance rotationnelle. Soit \mathbf{V}_\downarrow^n la matrice de dimensions $(n-1) \times K$ qui contient les $n-1$ premières lignes de \mathbf{V}^n , et \mathbf{V}_\uparrow^n la matrice de dimensions $(n-1) \times K$ qui contient les $n-1$ dernières lignes de \mathbf{V}^n . De même, soit $\mathbf{W}(t)_\downarrow$ la matrice de dimensions $(n-1) \times K$ qui contient les $n-1$ premières lignes de $\mathbf{W}(t)$, et $\mathbf{W}(t)_\uparrow$ la matrice de dimensions $(n-1) \times K$ qui contient les $n-1$ dernières lignes de $\mathbf{W}(t)$. Alors on vérifie que

$$\boxed{\mathbf{V}_\uparrow^n = \mathbf{V}_\downarrow^n \mathbf{D}} \quad (1.19)$$

où $\mathbf{D} = \text{diag}(z_0, \dots, z_{K-1})$. Or les colonnes de \mathbf{V}^n et celles de $\mathbf{W}(t)$ constituent deux bases d’un même espace vectoriel de dimension K . Ainsi, il existe une matrice inversible $\mathbf{G}(t)$ de dimension $K \times K$ telle que

$$\mathbf{V}^n = \mathbf{W}(t) \mathbf{G}(t) \quad (1.20)$$

où $\mathbf{G}(t)$ est définie comme la matrice de passage de la première base à la seconde. En substituant l’équation (1.20) dans l’équation (1.19), on montre que

$$\boxed{\mathbf{W}(t)_\uparrow = \mathbf{W}(t)_\downarrow \Phi(t)} \quad (1.21)$$

où $\Phi(t)$, appelée *matrice spectrale*, est définie par son EVD :

$$\Phi(t) = \mathbf{G}(t) \mathbf{D} \mathbf{G}(t)^{-1}. \quad (1.22)$$

En particulier, les valeurs propres de $\Phi(t)$ sont les pôles $\{z_k\}_{k=0\dots K-1}$.
En multipliant l'équation (1.21) à gauche par $\mathbf{W}(t)_\downarrow^H$, on obtient

$$\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\uparrow = \mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow \Phi(t). \quad (1.23)$$

Or, si $\text{rang}(\mathbf{W}(t)_\downarrow) = K$, la matrice $\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow$ est inversible. En effet, on remarque trivialement que $\forall \mathbf{x} \in \mathbb{C}^n$, $\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow \mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{W}(t)_\downarrow \mathbf{x} = \mathbf{0}$. Donc $\dim(\ker(\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow)) = \dim(\ker(\mathbf{W}(t)_\downarrow))$. Le théorème du rang permet de conclure que $\text{rang}(\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow) = \text{rang}(\mathbf{W}(t)_\downarrow) = K$, donc $\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow$ est inversible. Ainsi, l'équation (1.23) implique $\Phi(t) = (\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow)^{-1} \mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\uparrow$.

Finalemment, l'algorithme ESPRIT se décompose en quatre étapes :

- calculer et diagonaliser la matrice $\widehat{\mathbf{R}}_{xx}(t)$;
- en déduire une base de l'espace signal $\mathbf{W}(t)$;
- extraire de $\mathbf{W}(t)$ les matrices $\mathbf{W}(t)_\downarrow$ et $\mathbf{W}(t)_\uparrow$;
- calculer la matrice spectrale $\Phi(t) = (\mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\downarrow)^{-1} \mathbf{W}(t)_\downarrow^H \mathbf{W}(t)_\uparrow$
- diagonaliser $\Phi(t)$ et en déduire les pôles estimés.

Des études théoriques et expérimentales ont montré que la méthode ESPRIT est la plus performante des méthodes HR présentées ci-dessus (cf. section 1.5.2).

1.4 Estimation des autres paramètres

Les méthodes à haute résolution exposées dans les sections précédentes estiment uniquement les pôles z_k . On s'intéresse maintenant à l'estimation des autres paramètres du modèle.

1.4.1 Estimation de l'ordre de modélisation

Jusqu'à présent, l'ordre du modèle ESM était supposé connu, ce qui n'est généralement pas le cas dans la pratique. De nombreuses méthodes ont été proposées dans la littérature pour estimer le nombre de sinusoides présentes dans un bruit blanc. Les plus classiques sont la méthode du maximum de vraisemblance [24] et les critères issus de la théorie de l'information, dits ITC [25], parmi lesquels les critères AIC [26] et MDL par Schwartz [27] et Rissanen [28]. Une autre technique dans le cadre des ITC est le critère EDC [29], qui s'avère également robuste à un bruit blanc multiplicatif [30]. Ces divers critères ITC reposent sur la similarité des valeurs propres dans l'espace bruit, et non sur l'existence d'une cassure entre les espaces signal et bruit [31]. Un critère de sélection de l'ordre de modélisation fondé sur cette cassure, formulée en terme de décomposition *maximalement stable*, a été développé dans [32]. D'autres approches reposent sur les matrices de Wishart [33] et sur la méthode de validation croisée [34].

Cependant, dans le cas où le bruit est coloré, toutes ces méthodes tendent à surestimer l'ordre du modèle. Ainsi, des méthodes spécifiques ont été conçues pour traiter le cas d'un bruit coloré, parmi lesquelles de nouveaux critères ITC [35, 36], une technique basée sur un modèle de fonction d'autocovariance du bruit à support fini [37], et un critère de maximum a posteriori [38].

Parmi toutes ces méthodes, nous présentons ici les plus classiques, à savoir les trois principaux critères ITC : AIC, MDL et EDC (qui est une généralisation robuste de AIC et MDL). Ces méthodes consistent à minimiser une fonction de coût composée d'un premier terme commun et d'un second terme qui constitue un facteur de pénalisation :

$$\text{ITC}(p) = -(n-p)l \ln \left(\frac{\left(\prod_{q=p+1}^n \sigma_q^2 \right)^{\frac{1}{n-p}}}{\frac{1}{n-p} \sum_{q=p+1}^n \sigma_q^2} \right) + p(2n-p)C(l)$$

où les scalaires σ_q^2 sont les valeurs propres de la matrice $\widehat{\mathbf{R}}_{xx}(t)$ classées par ordre décroissant, et $C(l)$ est une fonction de la variable l . Le critère AIC est défini en posant $C(l) = 1$, et le critère MDL est défini en posant $C(l) = \frac{1}{2} \ln(l)$. Les critères EDC sont obtenus pour toutes les fonctions $l \mapsto C(l)$ telles que $\lim_{l \rightarrow +\infty} \frac{C(l)}{l} = 0$ et $\lim_{l \rightarrow +\infty} \frac{C(l)}{\ln(\ln(l))} = +\infty$. Ces critères conduisent à maximiser le rapport de la moyenne géométrique des valeurs propres de l'espace bruit sur leur moyenne arithmétique. Or ce rapport est maximal et égal à 1 lorsque toutes ces valeurs propres sont égales ; il mesure donc la blancheur du bruit (en théorie les valeurs propres sont toutes égales à σ^2). Le terme de pénalisation $C(l)$ permet d'éviter de surestimer p . Dans la pratique, ces méthodes sont relativement satisfaisantes pour traiter des signaux qui vérifient bien le modèle de signal, mais leurs performances s'effondrent quand ce modèle est moins bien vérifié, en particulier quand le bruit est coloré.

1.4.2 Estimation des amplitudes, des phases et de l'écart-type du bruit

Le principe du maximum de vraisemblance développé dans la section 1.2.1 suggère d'utiliser la méthode des moindres carrés pour estimer les amplitudes complexes (cf. équation (1.5)) :

$$\alpha(t) = \mathbf{V}^{N\dagger} \mathbf{x}(t),$$

dont se déduisent $a_k = |\alpha_k|$ et $\phi_k = \arg(\alpha_k)$. Rappelons que d'après le théorème de Gauss-Markov, l'estimateur des moindres carrés est un estimateur linéaire sans biais, de variance minimale parmi tous les estimateurs linéaires sans biais, dans la mesure où le bruit additif est blanc. Dans le cas où le bruit additif est coloré, l'estimateur optimal est obtenu par la méthode des moindres carrés pondérés (on pourra consulter [39] pour des informations détaillées sur l'estimation des amplitudes par la méthode des moindres carrés pondérés).

Enfin, le principe du maximum de vraisemblance suggère d'estimer l'écart-type en calculant la puissance du résiduel (cf. équation (1.4)) :

$$\sigma^2 = \frac{1}{N} \|\mathbf{x}(t) - \mathbf{V}^N \alpha(t)\|^2.$$

1.5 Performances des estimateurs

1.5.1 Borne de Cramer-Rao

La borne de Cramér-Rao est un outil fondamental en théorie des probabilités, car elle permet d'analyser les performances d'un estimateur, en rapportant la variance de celui-ci à une valeur optimale, qui tient en quelque sorte lieu de repère de qualité. Dans le cas particulier du modèle de signal ESM, une étude de la borne de Cramér-Rao a été proposée dans [21]. Le théorème général de la borne de Cramér-Rao est rappelé ci-dessous (cf. [40]). Il repose sur l'hypothèse d'un modèle statistique régulier.

Définition 1 (Modèle statistique régulier) Soit un modèle statistique dominé par une mesure μ et paramétré par $\theta \in \Theta$, où Θ est une partie ouverte de \mathbb{R}^q . Notons \mathbf{x} la variable vectorielle de dimension N . La paramétrisation est dite régulière si les conditions suivantes sont vérifiées :

1. la densité de probabilité $p(\mathbf{x}; \theta)$ est continûment dérivable, μ -presque partout, par rapport à θ .
2. la matrice d'information de Fisher

$$\mathbf{F}(\theta) \triangleq \int_H \mathbf{l}(\mathbf{x}; \theta) \mathbf{l}(\mathbf{x}; \theta)^T p(\mathbf{x}; \theta) d\mathbf{x}$$

définie à partir de la fonction de score $\mathbf{l}(\mathbf{x}; \theta) \triangleq \nabla_{\theta} \ln p(\mathbf{x}; \theta) \mathbf{1}_{p(\mathbf{x}; \theta) > 0}$ est définie positive pour toute valeur du paramètre θ et continue par rapport à θ .

Théorème 4 (Borne de Cramér-Rao) Soit un modèle statistique régulier paramétré par $\theta \in \Theta$. Soit $\widehat{\theta}$ un estimateur non biaisé de θ ($\forall \theta \in \Theta, \mathbb{E}_{\theta}[\widehat{\theta}] = \theta$). Alors la matrice de dispersion $\mathbf{D}(\theta, \widehat{\theta}) \triangleq \mathbb{E}_{\theta} \left[(\widehat{\theta} - \theta)(\widehat{\theta} - \theta)^T \right]$ est telle que la matrice $\mathbf{D}(\theta, \widehat{\theta}) - \mathbf{F}(\theta)^{-1}$ est positive.

En particulier, les éléments diagonaux de la matrice $\mathbf{D}(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}) - \mathbf{F}(\boldsymbol{\theta})^{-1}$ sont positifs. Par conséquent, les variances des coefficients de $\widehat{\boldsymbol{\theta}}$ sont supérieures aux éléments diagonaux de la matrice $\mathbf{F}(\boldsymbol{\theta})^{-1}$. Ainsi les bornes d'estimation de Cramér-Rao pour l'ensemble des paramètres scalaires s'obtiennent en trois étapes :

- calcul de la matrice d'information de Fisher ;
- inversion de cette matrice ;
- extraction des éléments diagonaux.

Comme cela a été mentionné dans la section 1.2.1, le vecteur $\mathbf{x}(t)$ contenant les N échantillons du signal observé est un vecteur aléatoire gaussien d'espérance $\mathbf{s}(t)$ et de matrice de covariance \mathbf{R}_{bb} . Ci-dessous, la dépendance de $\mathbf{s}(t)$ et de \mathbf{R}_{bb} par rapport aux paramètres du modèle sera mentionnée explicitement. En revanche, pour simplifier les notations, nous omettrons la dépendance de $\mathbf{s}(t)$ par rapport au temps. Il est connu que la matrice d'information de Fisher d'un vecteur aléatoire gaussien s'exprime simplement en fonction des paramètres du modèle, comme le montre la proposition suivante [40, pp. 525].

Proposition 5 (Matrice d'information de Fisher pour une densité gaussienne) *Pour une famille de lois de probabilité gaussiennes complexes de matrice de covariance $\mathbf{R}_{bb}(\boldsymbol{\theta})$ et de moyenne $\mathbf{s}(\boldsymbol{\theta})$, où $\mathbf{R}_{bb} \in \mathcal{C}^1(\Theta, \mathbb{C}^{N \times N})$ et $\mathbf{s} \in \mathcal{C}^1(\Theta, \mathbb{C}^N)$, les coefficients de la matrice d'information de Fisher $\{\mathbf{F}_{(i,j)}(\boldsymbol{\theta})\}_{1 \leq i, j \leq k}$ sont donnés par la formule de Bangs-Slepian étendue :*

$$\mathbf{F}_{(i,j)}(\boldsymbol{\theta}) = \text{trace} \left(\mathbf{R}_{bb}^{-1} \frac{\partial \mathbf{R}_{bb}(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{R}_{bb}^{-1} \frac{\partial \mathbf{R}_{bb}(\boldsymbol{\theta})}{\partial \theta_j} \right) + 2 \mathcal{R}e \left(\frac{\partial \mathbf{s}(\boldsymbol{\theta})^H}{\partial \theta_i} \mathbf{R}_{bb}^{-1} \frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_j} \right). \quad (1.24)$$

En appliquant la formule (1.24) au modèle ESM, on obtient une expression analytique de la matrice d'information de Fisher. On en déduit le théorème suivant, démontré dans [21] :

Proposition 6 *Les bornes de Cramér-Rao pour les paramètres (ϕ_k, δ_k, f_k) sont indépendantes de a_k pour tout $k' \neq k$, mais proportionnelles à $\frac{1}{a_k^2}$. La borne pour le paramètre a_k est indépendante de tous les a_k . Enfin, les bornes pour tous les paramètres sont indépendantes de toutes les phases ϕ_k , et sont inchangées par une translation de l'ensemble des fréquences f_k .*

En outre, les bornes de Cramér-Rao peuvent être calculées analytiquement sous certaines hypothèses, comme cela a été fait dans [41].

Proposition 7 *Supposons que tous les facteurs d'atténuation sont nuls, et faisons tendre N vers $+\infty$. Alors les bornes de Cramér-Rao pour les paramètres du modèle ESM admettent les développements limités au premier ordre suivants :*

- $\text{CRB}\{\sigma\} = \frac{\sigma^2}{4N} + \mathcal{O}\left(\frac{1}{N^2}\right)$;
- $\text{CRB}\{f_k\} = \frac{6\sigma^2}{4\pi^2 N^3 a_k^2} + \mathcal{O}\left(\frac{1}{N^4}\right)$;
- $\text{CRB}\{a_k\} = \frac{2\sigma^2}{N} + \mathcal{O}\left(\frac{1}{N^2}\right)$;
- $\text{CRB}\{\phi_k\} = \frac{2\sigma^2}{N a_k^2} + \mathcal{O}\left(\frac{1}{N^2}\right)$.

On remarque en particulier que les bornes de Cramér-Rao relatives aux fréquences f_k sont de l'ordre de $\frac{1}{N^3}$, ce qui est peu courant en estimation paramétrique. Par ailleurs, il est connu que le principe du maximum de vraisemblance fournit des estimateurs asymptotiquement efficaces [40]. Ainsi, les variances des estimateurs donnés dans la section 1.2.1 sont asymptotiquement équivalentes aux bornes de Cramér-Rao données dans la proposition 7. Le cas des méthodes HR est abordé ci-dessous.

1.5.2 Performances des méthodes HR

Les performances d'un estimateur sont généralement exprimées en terme de biais et de variance. Il est également possible de mesurer son efficacité, définie comme le rapport de sa variance sur la borne de Cramér-Rao. En particulier, un estimateur est dit efficace si son efficacité est égale à 1.



Dans le cas des méthodes HR, il s'avère malheureusement impossible de calculer analytiquement biais et variance, car l'extraction des racines d'un polynôme, ou des valeurs propres d'une matrice, induit une relation complexe entre les statistiques du signal et celles des estimateurs. Cependant, des résultats asymptotiques ont pu être obtenus grâce à la théorie des perturbations. Ces résultats reposent soit sur l'hypothèse $N \rightarrow +\infty$ (dans le cas où tous les pôles sont sur le cercle unité), soit sur l'hypothèse d'un fort Rapport Signal à Bruit ($RSB \rightarrow +\infty$). Sous chacune de ces deux hypothèses, il a été démontré que toutes les méthodes HR présentées dans ce chapitre sont non biaisées. De plus, sous l'hypothèse $N \rightarrow +\infty$, les variances des méthodes de Prony et Pisarenko ont été calculées dans [42], et celles de MUSIC et ESPRIT dans [43]. Sous l'hypothèse $RSB \rightarrow +\infty$, la variance de la méthode de Prony a été calculée dans [44], celle de MUSIC dans [45], et celle de ESPRIT dans [45, 46].

Les développements mathématiques proposés dans tous ces articles sont assez complexes, et sont fortement liés à la méthode d'estimation considérée, c'est pourquoi ils ne sont pas reproduits dans le cadre de ce document. Seuls les principaux résultats sont résumés ici. Tout d'abord, il a été démontré dans [42, 44] que les méthodes de Prony et de Pisarenko sont très inefficaces, au sens statistique du terme : leurs variances sont largement supérieures aux bornes de Cramér-Rao. De plus, elles augmentent plus vite que les bornes de Cramér-Rao quand le RSB décroît. En revanche les méthodes MUSIC et ESPRIT ont une efficacité asymptotique proche de 1. Plus précisément, il a été démontré dans [43, 45] (dans le cadre de sinusoïdes non modulées) que ces deux méthodes atteignent des performances presque identiques, mais que ESPRIT est légèrement meilleur que MUSIC. L'étude menée dans [46] (dans le cas plus général de sinusoïdes modulées exponentiellement) va dans le même sens : ESPRIT s'avère moins sensible au bruit que MUSIC.

1.6 Conclusion

Dans ce chapitre, nous avons montré que l'estimation des fréquences et des facteurs d'atténuation par la méthode du maximum de vraisemblance conduit à un problème d'optimisation difficile. Quand tous les pôles du signal sont sur le cercle unité, elle peut être approchée par la détection des K principaux pics du périodogramme. Ce résultat n'est valable que lorsque la longueur de la fenêtre d'observation est suffisamment grande devant l'inverse du plus petit écart fréquentiel entre pôles voisins. L'intérêt principal des méthodes HR est qu'elles s'affranchissent de cette limite de l'analyse de Fourier en terme de résolution spectrale. Les premières méthodes de cette famille, proposées par Prony et Pisarenko, reposent sur les équations de récurrence linéaire qui caractérisent le modèle de signal. En revanche les techniques plus modernes, parmi lesquelles les méthodes MUSIC et ESPRIT, s'appuient sur la décomposition de l'espace des données en deux sous-espaces propres de la matrice de covariance, appelés espace signal et espace bruit. L'étude statistique de ces diverses techniques d'estimation a montré que la méthode ESPRIT est la plus performante. Les amplitudes et les phases des exponentielles complexes peuvent ensuite être estimées par la méthode des moindres carrés.

1.7 Annexes

1.7.1 Optimisation sous contraintes

Soit V un espace de Hilbert et F un sous-ensemble fermé de V défini par $F = \{\mathbf{y} \in V / f_1(\mathbf{y}) = 0 \dots f_M(\mathbf{y}) = 0\}$, où les fonctions $\{f_m\}_{m=1\dots M}$ sont continues de V dans \mathbb{R} . Soit J une fonction à valeurs réelles sur V et \mathbf{p} un minimum local de J sur F . On suppose de plus que les fonctions J et $\{f_m\}_{m=1\dots M}$ sont dérivables en \mathbf{p} .

Theorème 8 (Multipliateurs de Lagrange) *Il existe $\mu_0, \dots, \mu_M \in \mathbb{R}$ non tous nuls tels que $\mu_0 J'(\mathbf{p}) + \sum_{m=1}^M \mu_m f'_m(\mathbf{p}) = 0$. Si de plus les vecteurs $\{f'_m(\mathbf{p})\}_{m=1\dots M}$ sont linéairement indépendants, alors il existe des coefficients $\lambda_1 \dots \lambda_M \in \mathbb{R}$ appelés multipliateurs de Lagrange tels que $J'(\mathbf{p}) + \sum_{m=1}^M \lambda_m f'_m(\mathbf{p}) = 0$.*



FIGURE 1.4 – Joseph-Louis LAGRANGE (1736-1813)

1.7.2 Matrices de Vandermonde

Définition 2 (Matrice de Vandermonde) Soient $K \in \mathbb{N}^*$ et $z_0, \dots, z_{K-1} \in \mathbb{C}$. On appelle matrice de Vandermonde une matrice V de dimension $K \times K$ de la forme

$$V = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_0 & z_1 & \dots & z_{K-1} \\ \vdots & \vdots & \ddots & \vdots \\ z_0^{K-1} & z_1^{K-1} & \dots & z_{K-1}^{K-1} \end{bmatrix}.$$

Proposition 9 (Déterminant de Vandermonde) Le déterminant de la matrice de Vandermonde est

$$\det(V) = \prod_{0 \leq k_1 < k_2 \leq K-1} (z_{k_2} - z_{k_1}).$$

Bibliographie

- [1] F. Keiler and S. Marchand, “Survey on extraction of sinusoids in stationary sounds,” in *Proc. of DAFx-02*, Hamburg, Allemagne, Sept. 2002, pp. 51–58.
- [2] S. Marcos, J. Sanchez-Araujo, N. Bertaux, P. Larzabal, and P. Forster, *Les Méthodes à haute résolution : traitement d’antenne et analyse spectrale. Chapitres 4 et 5*, Hermès, Paris, France, 1998, Ouvrage collectif sous la direction de S. Marcos.
- [3] J. Jensen, R. Heusdens, and S. H. Jensen, “A perceptual subspace approach for modeling of speech and audio signals with damped sinusoids,” *IEEE Trans. Speech Audio Processing*, vol. 12, no. 2, pp. 121–132, Mar. 2004.
- [4] J. Laroche, “The use of the Matrix Pencil method for the spectrum analysis of musical signals,” *Journal of the Acoustical Society of America*, vol. 94, no. 4, pp. 1958–1965, Oct. 1993.
- [5] K. Hermus, W. Verhelst, and P. Wambacq, “Psychoacoustic modeling of audio with exponentially damped sinusoids,” in *Proc. of ICASSP’02*. 2002, vol. 2, pp. 1821–1824, IEEE.
- [6] C. Lambourg and A. Chaigne, “Measurements and modeling of the admittance matrix at bridge in guitars,” in *Proc. of SMAC’93*, Stockholm, Suède, July 1993, pp. 449–453.
- [7] G. Weinreich, “Coupled piano strings,” *Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1474–1484, 1977.
- [8] B. David, *Caractérisations acoustiques de structures vibrantes par mise en atmosphère raréfiée*, Ph.D. thesis, University of Paris VI, 1999.
- [9] M. Jeanneau, P. Mouyon, and C. Pendaries, “Sintrack analysis, application to detection and estimation of flutter for flexible structures,” in *Proc. of EUSIPCO*, Ile de Rhodes, Grèce, Sept. 1998, pp. 789–792.
- [10] G. M. Riche de Prony, “Essai expérimental et analytique : sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de l’eau et de la vapeur de l’alcool à différentes températures,” *Journal de l’école polytechnique*, vol. 1, no. 22, pp. 24–76, 1795.
- [11] V. F. Pisarenko, “The retrieval of harmonics from a covariance function,” *Geophysical J. Royal Astron. Soc.*, vol. 33, pp. 347–366, 1973.
- [12] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [13] A. J. Barabell, “Improving the resolution performance of eigenstructure-based direction-finding algorithms,” in *Proc. of ICASSP’83*, Boston, MA, USA, 1983, pp. 336–339, IEEE.
- [14] S. Y. Kung, K. S. Arun, and D. B. Rao, “State-space and singular value decomposition based approximation methods for harmonic retrieval problem,” *J. of Opt. Soc. of America*, vol. 73, pp. 1799–1811, Dec. 1983.
- [15] R. Roy, A. Paulraj, and T. Kailath, “ESPRIT—A subspace rotation approach to estimation of parameters of cisoids in noise,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 5, pp. 1340–1342, Oct. 1986.
- [16] R. Roy and T. Kailath, “Total least squares ESPRIT,” in *Proc. of 21st Asilomar Conference on Signals, Systems, and Computers*, Nov. 1987, pp. 297–301.
- [17] M. Zoltawski and D. Stavrinos, “Sensor array signal processing via a Procrustes rotations based eigen-analysis of the ESPRIT data pencil,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 6, pp. 832–861, June 1989.



- [18] J. Nieuwenhuijse, R. Heusens, and Ed. F. Deprettere, “Robust exponential modeling of audio signals,” in *Proc. of ICASSP’98*. May 1998, vol. 6, pp. 3581–3584, IEEE.
- [19] R. Kumaresan and D. W. Tufts, “Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, no. 6, pp. 833–840, Dec. 1982.
- [20] Y. Li, K. Liu, and J. Razavilar, “A parameter estimation scheme for damped sinusoidal signals based on low-rank Hankel approximation,” *IEEE Trans. Signal Processing*, vol. 45, pp. 481–486, Feb. 1997.
- [21] Y. Hua and T. K. Sarkar, “Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 5, pp. 814–824, May 1990.
- [22] A.-J. Van der Veen, Ed. F. Deprettere, and A. L. Swindlehurst, “Subspace based signal analysis using singular value decomposition,” *Proc. of IEEE*, vol. 81, no. 9, pp. 1277–1308, Sept. 1993.
- [23] R. O. Schmidt, *A signal subspace approach to multiple emitter location and spectral estimation*, Ph.D. thesis, Stanford University, Stanford, California, USA, Nov. 1981.
- [24] G. Bienvenu and L. Kopp, “Optimality of high-resolution array processing using the eigensystem method,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, no. 5, pp. 1235–1245, Oct. 1983.
- [25] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 387–392, Apr. 1985.
- [26] H. Akaike, “Information theory and an extension of the maximum likelihood principle,” in *Proc. of the 2nd International Symposium on Information Theory*, B. N. Petrov and F. Csaki, Eds., Budapest, Hongrie, 1973, pp. 267–281, Akademia Kiado.
- [27] G. Schwarz, “Estimating the dimension of a model,” *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.
- [28] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, pp. 465–471, 1978.
- [29] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai, “On detection of the number of signals in presence of white noise,” *Journal of Multivariate Analysis*, vol. 20, no. 1, pp. 1–25, 1986.
- [30] F. Gini and F. Bordononi, “On the behavior of information theoretic criteria for model order selection of InSAR signals corrupted by multiplicative noise,” *Signal Processing*, vol. 83, pp. 1047–1063, 2003.
- [31] A. P. Liavas and P. A. Regalia, “On the behavior of Information Theoretic Criteria for model order selection,” *IEEE Trans. Signal Processing*, vol. 49, no. 8, pp. 1689–1695, Aug. 2001.
- [32] A. P. Liavas, P. A. Regalia, and J.-P. Delmas, “Blind channel approximation : effective channel order determination,” *IEEE Trans. Signal Processing*, vol. 47, no. 12, pp. 3336–3344, Dec. 1999.
- [33] J. Grouffaud, P. Larzabal, and H. Clergeot, “Some properties of ordered eigenvalues of a Wishart matrix : application in detection test and model order selection,” in *Proc. of ICASSP’96*. 1996, vol. 5, pp. 2465–2468, IEEE.
- [34] D. Kundu and A. Mitra, “Detecting the number of signals for an undamped exponential model using cross-validation approach,” *Signal Processing*, vol. 80, no. 3, pp. 525–534, 2000.
- [35] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai, “On detection of the number of signals when the noise covariance matrix is arbitrary,” *Journal of Multivariate Analysis*, vol. 20, no. 1, pp. 26–49, 1986.
- [36] Q. T. Zhang and K. M. Wong, “Information theoretic criteria for the determination of the number of signals in spatially correlated noise,” *IEEE Trans. Signal Processing*, vol. 41, no. 4, pp. 1652–1663, Apr. 1993.
- [37] J. J. Fuchs, “Estimation of the number of signals in the presence of unknown correlated sensor noise,” *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1053–1061, May 1992.
- [38] W. B. Bishop and P. M. Djuric, “Model order selection of damped sinusoids in noise by predictive densities,” *IEEE Trans. Signal Processing*, vol. 44, no. 3, pp. 611–619, Mar. 1996.
- [39] P. Stoica, H. Li, and J. Li, “Amplitude estimation of sinusoidal signals : survey, new results, and an application,” *IEEE Trans. Signal Processing*, vol. 48, no. 2, pp. 338–352, 2000.



- [40] S. M. Kay, *Fundamentals of Statistical Signal Processing : Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
- [41] C. R. Rao and L. C. Zhao, “Asymptotic behavior of maximum likelihood estimates of superimposed exponential signals,” *IEEE Trans. Signal Processing*, vol. 41, no. 3, pp. 1461–1464, Mar. 1993.
- [42] P. Stoica and A. Nehorai, “Study of the statistical performance of the Pisarenko harmonic decomposition method,” *IEE Proceedings Radar and Signal Processing*, vol. 135, no. 2, pp. 161–168, Apr. 1988.
- [43] P. Stoica and T. Söderström, “Statistical Analysis of MUSIC and Subspace Rotation Estimates of Sinusoidal Frequencies,” *IEEE Trans. Signal Processing*, vol. 39, pp. 1836–1847, Aug. 1991.
- [44] A. Kot, S. Parthasarathy, D. Tufts, and R. Vaccaro, “The statistical performance of state-variable balancing and Prony’s method in parameter estimation,” in *Proc. of ICASSP’87*, Apr. 1987, vol. 12, pp. 1549–1552.
- [45] A. Eriksson, P. Stoica, and T. Soderstrom, “Second-order properties of MUSIC and ESPRIT estimates of sinusoidal frequencies in high SNR scenarios,” *IEE Proceedings on Radar, Sonar and Navigation*, vol. 140, no. 4, pp. 266–272, Aug. 1993.
- [46] Y. Hua and T. K. Sarkar, “On SVD for estimating generalized eigenvalues of singular matrix pencil in noise,” *IEEE Trans. Signal Processing*, vol. 39, no. 4, pp. 892–900, Apr. 1991.





Contexte public } sans modifications

Par le téléchargement ou la consultation de ce document, l'utilisateur accepte la licence d'utilisation qui y est attachée, telle que détaillée dans les dispositions suivantes, et s'engage à la respecter intégralement.

La licence confère à l'utilisateur un droit d'usage sur le document consulté ou téléchargé, totalement ou en partie, dans les conditions définies ci-après et à l'exclusion expresse de toute utilisation commerciale.

Le droit d'usage défini par la licence autorise un usage à destination de tout public qui comprend :

- Le droit de reproduire tout ou partie du document sur support informatique ou papier,
- Le droit de diffuser tout ou partie du document au public sur support papier ou informatique, y compris par la mise à la disposition du public sur un réseau numérique.

Aucune modification du document dans son contenu, sa forme ou sa présentation n'est autorisée.

Les mentions relatives à la source du document et/ou à son auteur doivent être conservées dans leur intégralité.

Le droit d'usage défini par la licence est personnel, non exclusif et non transmissible.

Tout autre usage que ceux prévus par la licence est soumis à autorisation préalable et expresse de l'auteur : sitepedago@telecom-paristech.fr