

Ministère de l'Education Nationale

# Ecole Nationale Polytechnique

Département d'Electronique

المدرسة الوطنية المتعددة التقنيات  
المكثبة — POLYTECHNIQUE  
Ecole Nationale Polytechnique

## THESE DE MAGISTER

Spécialité: Electronique

option: Acquisition et traitement de l'information

Présentée par: Lahcene ABDELOUEL

Ingénieur d'Etat en Electronique

option : Contrôle industriel

### Estimation Spectrale

Basée sur la Modélisation Autorégressive,  
Application à la détection des périodicités  
cachées dans les phénomènes Aléatoires

Soutenue en Mai 1993, devant le Jury d'examen composé de

Mr A. FARAH  
Mr B. DERRAS  
Mr A. GUESSOUM  
Mr B. BOUSSEKSOU  
Mlle M. GUERTI  
Mr F. CHIGARA

Maitre de Conférence  
Ph.D  
Maitre de Conférence  
Chargé de Cours  
Chargé de Cours  
Maitre assistant

Président  
Rapporteur  
Examinateur  
Examinateur  
Examinatrice  
Examinateur

## REMERCIEMENTS

Le travail présenté dans cette thèse a été effectué au laboratoire de traitement de signal au sein du Département d'Electronique à l'Ecole Nationale Polytechnique d'ALGER. Ainsi je tiens à remercier la direction de m'avoir donné la possibilité de le faire.

Il m'est difficile d'exprimer ma gratitude à Monsieur Belkacem DERRAS, qui fut mon Directeur de recherche, et qui au long de ce travail a pu me faire bénéficier de ses connaissances scientifiques et ses qualités humaines. Sincèrement, il a dirigé ma thèse avec brio et modestie, ses suggestions judicieuses et l'aide toujours renouvelée qu'il m'a apportée m'ont été très précieuses.

Je suis reconnaissant à Mr A. FARAH d'avoir accepté la présidence du jury, ainsi qu'à Messieurs A. GUESSOUM, B. BOUSSEKSOU, F. CHIGARA et à Mme M. GUERTI, qui ont bien voulu examiner mon travail et faire partie du jury.

Je ne puis oublier l'accueil chaleureux que m'a manifesté l'équipe du laboratoire "Communication parlé" de l'Institut d'Electronique à l'U.S.T.H.B., et particulièrement Mr TEFFAHI H., qui m'a aidé dans l'utilisation du système d'acquisition et de traitement de la parole.

Je tiens à remercier vivement mes collègues au laboratoire, dont la compagnie aura été à la fois amicale et enrichissante.

Ma reconnaissance et mes amitiés vont à tous les enseignants de l'Institut d'Electronique de l'Université de Setif, particulièrement Mrs BENBAOUCHE L., KHOUNI L., IKNI R., BOUTERFA M. et Mr CHEMMALI H.

Je ne serai oublier dans mes remerciements Mr GUEMIDI Lahcene, Dr GUEMIDI Ali El-Arbi, pour leur collaboration. Qu'ils trouvent ici l'expression de ma très profonde gratitude.

J'adresse mes remerciements les plus sincères à Messieurs HAMZA A., SAIS D., SARI H., DOUZI H., BENLARBI K., ainsi qu'à tous ceux qui de près ou de loin, ont contribué à l'élaboration de cette thèse.

Que tous ceux qui m'ont aidé, soutenu et encouragé trouvent ici le signe de ma gratitude.

Quoique mon envie soit grande, je n'aurai pas l'impertinence de me féliciter, moi même, pour avoir réalisé le logiciel du graphisme ALGRAPH.

## Dédicaces

- Aux êtres qui me sont les plus chers: Mes parents.
- A mes frères et soeurs ...
- A l'âme de la feuë, très chère grande mère **BELMOULOU A.**
- A mon oncle **AOURAN A.Z.**
- A ma seconde famille, la famille **BENMESBAH** Messaoud qui m'ont cordialement accueilli, et m'ont soutenu dans les moments les plus difficiles de ma recherche.
- A tous mes amis et amies ...
- A toute l'équipe de **NETSYS** ...

..... Je dédie ce mémoire en signe d'amitié et de reconnaissance.

**Lahcene ABDELOUEL**

*A l'élue de mon coeur*

O élue de mon coeur,  
O rêve de mes nuits sans sommeil,  
Toi pour qui j'ai délaissé la joie et la gaieté,  
Viens de grâce soulager ma douleur.  
Je suis victime de ta beauté,  
Car je ne vois la vie qu'en ta présence,  
Et ne goûte la joie qu'à ta vue.  
Idéal de mes rêves,  
Sans ton espoir vivrais-je encors ?  
Allah m'unira-t-il avec celle que j'aime ?  
Seule l'espérance me console dans mon impatience.  
Glorie à Allah l'exalté, qui ne retarde le bonheur  
que pour le donner parfait.

**ABDELOUEL L.**

	Page
Introduction	
1. Position du problème	01
2. Domaines d'applications	01
3. Perspective Historique	03
4. But de la recherche	06

المدرسة الوطنية المتعددة التقنيات

المكتبة

المكتبة  
Ecole Nationale Polytechnique

## CHAPITRE I

### Généralités sur l'analyse et l'estimation spectrale

1.1 Pourquoi l'analyse spectrale ?	07
1.2 Signaux et Systemes	07
1.3 Representation des signaux	08
1.4 Series et transformation de FOURIER	10
1.5 Puissance et densité spectrale de puissance (DSP)	11
1.6 Densité Spectrale de puissance des processus aléatoires	13
1.7 Estimation et caractéristiques des estimateurs	14
1.8 Méthodes fondamentales d'estimation spectrale	15

## CHAPITRE II

### Estimateurs Spectraux Classiques

2.1 Introduction	18
2.2 Méthode du Corrélogramme	19
2.2.1 Estimateurs pour la fonction d'autocorrélation	19
2.2.2 Estimateurs pour la densité spectrale de puissance	21
2.3 Méthodes du périodogramme	23
2.3.1 Transformation de Fourier discrète (TFD)	23
2.3.2 Estimateur simple	24
2.3.3 Estimateur Moyenné	25
2.3.4 Estimateur Adouci	26
2.3.5 Estimateur modifié	27
2.4 Méthode hybride	28
2.5 Effet des fenêtres de pondération	30
2.5.1 Caractéristiques des fenêtres	30
2.6 Exemple d'application	31
2.7 Conclusions	33

## CHAPITRE III

### Modélisation paramétrique

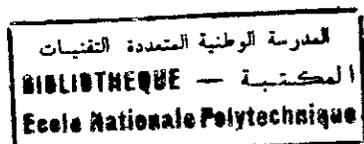
3.1 Introduction	40
3.2 Représentation paramétrique	40
3.3 Description des modèles	41
3.4 Modélisation Autorégressive (AR)	42
3.5 Propriétés des modèles AR	43
3.5.1 Prédiction linéaire	43
3.5.2 Erreurs de prédiction: progressive et rétrograde	44
3.5.3 Variances des erreurs de prédiction	45
3.6 Solution des équations normales	46
3.6.1 Algorithme de Levinson	47
3.7 Coefficients de réflexion	49
3.8 Structure du filtre d'estimation	50
3.8 Conclusion	52

6.1 Introduction.....	108
6.2 Application au traitement de la parole.....	108
6.2.1 Le signal de parole et sa modélisation.....	108
6.2.2 Systèmes de transmission de la parole.....	109
6.2.3 Résultats et interprétations.....	110
6.3 Application au traitement des données météorologiques.....	112
6.3.1 Présentation des données.....	113
6.3.2 Critique des données.....	114
6.3.3 Traitements et interprétations.....	115
-Analyse des séries d'observations Météorologiques.....	115
-Analyse de la série de mesures de l'activité solaire.....	117
-Relation entre température et activité solaire.....	118
-Modélisation AR de la pluviosité.....	119
6.4 Conclusion.....	119
 CONCLUSIONS.....	 133
 BIBLIOGRAPHIE.....	 138
 ANNEXES	
A1 Périodogramme et Transformation de Fourier, Transformation de Fourier Rapide (FFT).....	144
A2 Caractéristiques des fonctions fenêtres.....	147
A3 Algorithme de Levinson.....	150
A4 Algorithme de Burg.....	152
A5 Algorithme de Cholesky.....	154
A6 Phonèmes de l'Arabe standard.....	156
A7 Les variables météorologiques.....	157
A8 L'activité solaire.....	160
A9 Manuel d'utilisation du logiciel.....	162
A10 Unités et bibliothèque de procédures/fonctions.....	181

المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

# Introduction

## INTRODUCTION



### 1. POSITION DU PROBLEME :

Globalement, le problème de l'estimation et de l'analyse spectrale se pose lorsqu' il s'agit d'estimer la densité spectrale de puissance (DSP) d'un signal  $x(t)$  à partir de  $N$  échantillons seulement .

Mathématiquement, la DSP est calculée à partir d'une infinité d'échantillons, or ne disposant que de  $N$  points, il est nécessaire de faire des hypothèses sur le signal en dehors de la fenêtre d'observation.

L'hypothèse la plus classique, consiste à estimer la fonction d'autocorrélation sur les  $N$  points disponibles, les autres sont mis à zéro. Cela conduit aux méthodes mettant en oeuvre la transformée de Fourier.

Le second type d'hypothèse revient à extrapoler la fonction d'autocorrélation en dehors de son support par des valeurs qui satisfassent un certain critère, comme celui de minimiser l'erreur quadratique moyenne de l'erreur de prédiction ou de maximiser l'entropie du processus ...etc. Cela conduit aux méthodes dites modernes ou paramétriques.

La difficulté principale du problème d'estimation de la DSP est due justement au fait que la solution n'est pas unique est dépend des hypothèses faites à priori sur les données non disponibles. Mais, malgré cette difficulté, l'analyse spectrale est un moyen privilégié d'investigation dans les domaines les plus variés de la recherche et de l'industrie.

### 2. DOMAINES D 'APPLICATIONS :

L'analyse spectrale est la discipline de traitement du signal la plus utilisée dans les applications: Décomposer des signaux mesurés en leurs principales composantes fréquentielles, est une pratique extrêmement répandue -et depuis fort longtemps- en Acoustique, Mécanique vibratoire et en Astronomie.

Actuellement l'analyse spectrale est utilisée pour fournir des paramètres pertinents dans des domaines aussi divers que Géophysique, le traitement de la parole, la médecine, les Communications, l'énergie et l'étude des réacteurs nucléaires. Nous nous bornerons à en donner à titre d'exemple seulement quelques applications pratiques parmi tant d'autres.

Dans le domaine des sciences médicales, l'intérêt des méthodes d'estimation est évident: elles apportent une aide précieuse aux médecins dans l'interprétation de l'activité de certains organes humains, (Coeur, cerveau, muscles... ). En effet, dans le cas de l'électro-encéphalographie, une onde EEG est constituée par un grand nombre d'activités élémentaires, et il est important pour les spécialistes de détecter certaines fréquences (rythmes  $\alpha$ ,  $\beta$ ,  $\gamma$ ) noyées dans un bruit caractéristique de l'activité cérébrale pour étudier, par exemple, l'effet de certaines drogues sur le système nerveux. [50].

En ce qui concerne l'étude et la description des comportements vibratoires, le domaine fréquentiel est très largement utilisé, lors de la surveillance des changements des caractéristiques vibratoires d'une structure mécanique complexe, soumise à une excitation non maîtrisée. A titre d'exemple, l'analyse des spectres d'un ensemble machines tournantes / structure, (signature) identifiées sur des signaux multicateurs, permet de déceler la variation du comportement de certains éléments de la machine, et de surveiller la structure par rapport à certains seuils, au delà desquels la fiabilité est compromise et l'instabilité est destructrice [50].

En Géophysique, on fait souvent appel aux méthodes de modélisation AR/ARMA, afin de déterminer les caractéristiques d'une structure géologique [40]. La technique vibrosonique consiste à émettre un signal acoustique dans le sol (ondelette) et d'enregistrer les réflexions se produisant sur l'interface des différentes couches du canal géologique. Une modélisation de cette trace sismique permet de déterminer l'épaisseur et les caractéristiques des couches traversées, ainsi que les coefficients de réflexion correspondants, ce qui renseigne sur la nature des gisements.

Il y a bien d'autres applications, dans tous les domaines des sciences de l'ingénieur. D'une manière générale, on fait appel aux méthodes d'estimation, chaque fois qu'il s'agit de signaux, qui ne se présentent pas du tout comme périodiques et qui ont un caractère aléatoire très prononcé. Et pour toutes ces applications l'analyse spectrale constitue un maillon d'une chaîne globale d'acquisition et traitement des mesures.

Il est bon de rappeler que l'analyse spectrale basée sur un nombre limité de données est devenue possible progressivement grâce aux réflexions et aux réalisations de centaines de savants, mathématiciens et chercheurs, qui travaillaient souvent chacun dans un domaine et avec des objectifs différents.

### 3. Perspective Historique

Historiquement le mot "Spectrum" a été introduit pour la première fois, par NEWTON en 1687 pour caractériser le spectre de la lumière blanche [58]. Mais les premiers travaux concernant les estimateurs spectraux conventionnels sont dus à SHUSTER [60],[61], qui présenta dès 1898 une méthode de calcul du spectre -d'un signal déterministe- à partir d'un nombre entier de périodes [47], d'où son appellation Périodogramme. Shuster a signalé que certaines périodicités supérieures sont dues à la méthode d'estimation elle même, et comme remède à ce problème, il proposa de calculer la moyenne de plusieurs périodogrammes, obtenus en faisant varier l'origine de l'observation.

Plus tard, indépendamment, SLUTSKY (en 1927) et DANIELL (en 1948), ont étudié le problème d'estimation de la DSP d'un bruit blanc [58], ils se sont aperçu qu'une meilleure estimation, peut être obtenue en faisant la moyenne du périodogramme sur une certaine bande de fréquence (filtrage fréquentiel), et que l'estimation est d'autant meilleure (moins de fluctuations indésirables) que la fenêtre est large.

WEINER [71] en 1930 et KHINTCHINE [37] en 1934 se sont intéressés aux processus aléatoires et stationnaires. Ils ont donné une définition statistique de la fonction de corrélation, leurs théorème qui stipule que la fonction de corrélation et la densité spectrale de puissance forment une paire de transformée de FOURIER, est fondamental en analyse Spectrale [38].

En ce qui concerne l'analyse spectrale empirique, on doit signaler que Blackman et Tukey [8] ont présenté en 1949 une méthode d'estimation de la DSP, le corrélogramme en l'occurrence en n'utilisant qu'un nombre fini d'éléments de la fonction d'autocorrélation des mesures, les autres valeurs sont considérées comme étant nulles [47]. Ceci revient à multiplier le signal observé par une fenêtre rectangulaire. Mais cette opération donne naissance à des déformations spectrales particulièrement gênantes. Pour réduire l'effet de cette troncature, de nombreuses autres fenêtres, de caractéristiques fréquentielles meilleures, ont été proposées [49].

On doit probablement à Tukey et Cooley [17] le titre de pionniers de l'analyse spectrale numérique, En effet, le calcul de la transformée de Fourier (T.F) par la méthode directe n'était plus utilisé à cause de la lenteur des calculs, mais le périodogramme repris de l'importance grâce à l'apparition d'un algorithme rapide de calcul de la T.F (Cooley et Tukey en 1965) [17], qui réduit considérablement le nombre d'opérations.

La recherche de la DSP d'un processus aléatoire et stationnaire, est un problème qui peut être résolu par les techniques de modélisation, la première approche a été introduite par le baron de PRONY en 1795 [58]. Le principe de sa théorie consiste à modéliser le processus par une combinaison de sous systèmes, non couplés (développement selon des fonctions exponentielles amorties) [19]. YULE [72] en 1927, afin de déterminer avec précision certaines périodicités cachées dans des séries temporelles (Activité solaire), a considéré le processus comme un modèle de régression ordinaire, de résidu ( $\varepsilon$ ), dont les paramètres peuvent être estimés par la méthode des moindres carrés. Le système d'équations résultant porte les noms d' "équations normales" ou de " YULE-WALKER " [47].

Comme la récurrence Autoregressive (AR) ne caractérise un signal donnée que d'une façon plus ou moins approchée, WOLD en 1938 étudia les modèles à moyenne adaptée (MA), et il proposa une théorie de décomposition pour les séries temporelles stationnaires [58].

Dans l'utilisation pratique de ces théories, on se confronte souvent au problème de l'inversion de la matrice de covariance du signal analysé. Dans le cas stationnaire, cette matrice a une structure bien particulière: elle est de Toeplitz. Alors Levinson [41] en 1947 détermina une procédure récursive vis-à-vis de l'ordre du modèle pour la détermination des paramètres.

En matière de résolution spectrale, l'honneur revient à BURG [10],[12] auteur de la méthode dite maximum d'entropie (MEM) présentée en 1967, bien qu'elle est équivalente aux méthodes auto-régressives [68] (Van den Bos 1971). Ces dernières n'ont connues leurs grand succès qu'avec la MEM.

La modélisation par les méthodes de covariance et d'autocorrélation, a été introduite en 1971 par Atal et Hanauer [58]; il s'agit d'estimer les coefficients du modèle qui minimisent l'erreur de prédiction quadratique.

En utilisant d'autres types de critères, on aboutit à d'autres méthodes d'estimation: méthode selon le maximum de vraisemblance [47] (Maximum Likelihood Method, MLM), méthode dite à variance minimale, méthode de Pisarenko [47], ..etc. Ces méthodes permettent d'améliorer la qualité de l'estimation quand les conditions de travail sont vérifiées .

L'une des principales difficultés des méthodes paramétriques basées sur la modélisation, réside dans le choix de l'ordre du modèle. Plusieurs méthodes ont été proposées afin de le déterminer. En 1975, Makhoul [42] a souligné que l'estimation s'améliore au fur et à mesure que l'on augmente l'ordre, mais Ulrych [67] estiment qu'on ne doit pas dépasser la moitié du nombre d'échantillons dont on dispose.

On doit à Akaike [2],[3] deux critères importants; le FPE (final prediction error, 1974) et le AIC (Akaike information criterion, 1969). D'autres critères ont été proposés par Parzen [56] en 1974 et Hannan [58] en 1979.

En résumé le problème de l'estimation de la DSP reste toujours d'actualité, malgré les nombreux travaux effectués.

#### 4. But de la recherche

Les méthodes d'estimation ont été largement développées, et il nous semble que les chercheurs ont fait le tour des questions essentielles. Cependant ce domaine, à notre avis devrait devenir actif au plan des applications et des systèmes de traitement matériel et logiciel.

C'est dans ce cadre que s'inscrit notre travail. Les objectifs de la présente recherche sont les suivants :

- Etude des méthodes d'estimation de la DSP des processus aléatoires et discrets, en résumant les méthodes classiques et en présentant les méthodes paramétriques basées sur la modélisation autoregressive. En particulier l'étude des problèmes liés à la détermination de l'ordre du modèle, la résolution, le dédoublement des raies, lissage et effet du bruit.
- Etude de l'extension des méthodes d'estimation au cas multicanaux, quand deux ou plusieurs séries de mesures sont disponibles et des densités interspectrales peuvent être estimées. On mettra le point sur les possibilités offertes par les fonctions de cohérence pour l'étude de la similitude entre signaux et la détermination du degré de causalité de certains mesures.
- Développer un logiciel d'estimation de la DSP afin d'évaluer les performances des différentes méthodes, ce qui permet de mettre les techniques d'estimation à la portée de tous ceux concernés par ce problème.
- Application des méthodologies développées à des problèmes typiques du traitement de signal, en procédant à l'analyse spectrale de certains signaux issus de situations réelles, à savoir
  - Les signaux de Parole.
  - Les données Météorologiques.
  - Les données de l'activité solaire (tâches solaires).

# **CHAPITRE I**

## Généralités sur l'analyse et l'estimation spectrale

## CHAPITRE I

### GÉNÉRALITÉS SUR L'ANALYSE ET L'ESTIMATION SPECTRALE

#### 1.1 Pourquoi l'analyse spectrale ?

De tous les outils dont peut disposer le traitement de signal, l'analyse spectrale est certainement l'une des plus importantes. Les raisons de son excellence sont évidemment à chercher dans la relative universalité du concept de la fréquence: que ce soit dans des domaines s'intéressant à des ondes physiques, comme l'acoustique, les vibrations et l'optique, ou reposent sur la périodicité de certains événements; comme c'est le cas en économie, biologie et astronomie. En plus une description fréquentielle est souvent à la base d'une plus grande compréhension des phénomènes mis en jeu, en fournissant un complément d'information indispensable à la seule description temporelle. Si l'on ajoute à cela que l'approche fréquentielle s'accommode aussi bien à une transposition en traitement spatial (imagerie à résonance magnétique nucléaire, imagerie acoustique), on comprend qu'un grand nombre d'études ont été et soient encore consacrées à l'analyse spectrale.

Avant de formaliser du point de vue mathématique le problème, il est indispensable de présenter les principes généraux de l'analyse des signaux et de préciser quelques notions fondamentales de l'estimation.

#### 1.2 Signaux et Systèmes

Les relations de l'homme avec son milieu naturel, ou avec les systèmes techniques qu'il construit, se caractérisent par un échange intense de signaux. Les signaux sont des fonctions d'une ou plusieurs variables qui contiennent des informations sur le comportement et la nature des systèmes, se sont par exemple: un relevé de température, un enregistrement issu d'un sismographe, une statistique de vente, ou un signal de parole capté par un microphone.

Pour certains systèmes et phénomènes physiques, on arrive à exprimer une relation mathématique, généralement approchée, entre certaines variables, c'est souvent une équation différentielle ou aux différences finies. Dans ce cas il s'agit de signaux déterministes. Mais lorsque l'on connaît peu le système, comme en météorologie et en traitement de la parole, les signaux susceptibles d'être produits par ces systèmes sont dites aléatoires. De tels signaux ne possèdent évidemment pas de représentation temporelle analytique, ils peuvent toutefois être caractérisés par leurs propriétés statistiques et fréquentielles [18], [38], [49].

### 1.3 Représentation des signaux

Le fondement de l'analyse des signaux est basé sur le développement en une combinaison linéaire de fonctions [57]:

$$x(t) = \sum_{i=1}^n a_i \psi_i(t). \quad (1.1)$$

où  $a_i$  constituent une représentation discrète du signal  $x(t)$  et dépendent des fonctions  $\psi_i(t)$  choisies. Un choix adéquat des fonctions  $\psi_i$ , peut favoriser la mise en évidence de propriétés particulières du signal (*fréquence, amplitude, ...etc*) et faciliter l'étude des transformations qu'il subit en cours de sa propagation dans un système physique donné.

Si les fonctions  $\psi_i(t)$  forment une base orthogonale, on obtient alors un développement en série de fonctions orthogonales. L'exemple de la figure 1.1 illustre bien la représentation du signal  $x(t)$  par une combinaison linéaire de deux impulsions rectangulaires  $\psi_1$ , et  $\psi_2$ .  
Avec cette représentation, la détection de la périodicité  $T/2$  de l'impulsion  $\psi_2$  est évidente.

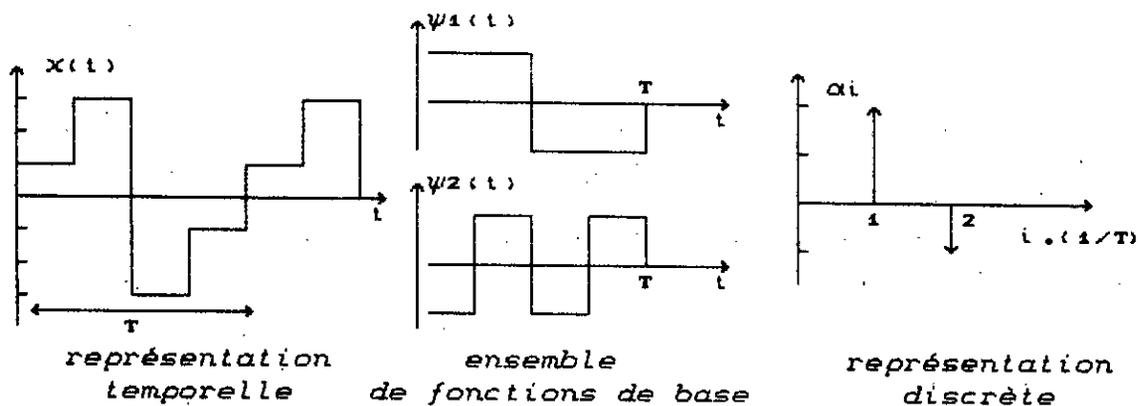


Fig 1.1 :Représentations temporelle /discrète.

Il existe plusieurs ensembles de bases de fonctions orthogonales. Sans être exhaustif on peut mentionner en particulier [18]:

- les fonctions de Bessel, de Radmarcher et de Welch.
- les fonctions de Legendre et de Tchebycheff .
- les fonctions sinusoidales et exponentielles complexes.

Mais de tous les développements orthogonaux d'un signal, le développement en série de FOURIER est le plus adéquat à la description des phénomènes physiques, cela pour plusieurs raisons:

-La solution des équations différentielles qui décrivent plusieurs processus est sinusoidale; le fait que l'environnement des circuits électriques est sinusoidale, donne une place du premier choix à ces fonctions [21].

-La plupart des phénomènes ondulatoires, sont périodiques, de période constante, ce qui n'est pas vérifié pour les fonctions de Bessel, par exemple.

-Les fonctions sinusoidales sont les fonctions propres des systèmes linéaires qui possèdent la propriété de ne pas déformer ou distordre les signaux sinusoidaux [50]. En plus, elles jouissent de la propriété d'invariance par rapport à certaines transformations mathématiques, à titre d'exemple après une intégration ou une différentiation -exempté le changement de phase- une onde sinusoidale demeure toujours sinusoidale.

#### 1.4 Séries et transformation de FOURIER

En 1807, Jean Baptiste FOURIER [24] démontrait que tout signal périodique de période  $T_0$ , peut être représenté sous la forme d'une somme pondérée de fonctions sinusoidales:

$$x(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} (A_n \cdot \cos(2\pi n f_0 \cdot t) + B_n \cdot \sin(2\pi n f_0 \cdot t)) \quad (1.2)$$

avec  $f_0 = 1 / T_0$ , les coefficients  $A_n$  et  $B_n$  sont donnés par

$$\begin{aligned} A_0 &= \frac{1}{T_0} \int_{(T_0)} x(t) \cdot dt \\ A_n &= \frac{1}{T_0} \int_{(T_0)} x(t) \cdot \cos(2\pi n f_0 \cdot t) \cdot dt \\ B_n &= \frac{1}{T_0} \int_{(T_0)} x(t) \cdot \sin(2\pi n f_0 \cdot t) \cdot dt \end{aligned} \quad (1.3)$$

Le développement sous la forme de la somme d'exponentielles est possible et donne

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{\infty} X_n \cdot \exp(j2\pi n f_0 \cdot t) \\ X_n &= \frac{1}{T_0} \int_{(T_0)} x(t) \cdot \exp(-j2\pi n f_0 \cdot t) \cdot dt \end{aligned} \quad (1.4)$$

Le tracé graphique de la fonction discrète  $X_n$  en fonction de la fréquences  $n f_0$  des exponentielles porte le nom de spectre fréquentiel du signal  $x(t)$ .

Pour les signaux non périodiques (extension à l'infini de la période  $T_0$ ), la série de FOURIER converge vers la transformée de Fourier

$$x(t) = \int_{-\infty}^{+\infty} X(f) \cdot \exp(j2\pi f \cdot t) \cdot df \quad (1.5)$$

$$\text{avec } X(f) = \int_{-\infty}^{+\infty} x(t) \exp(-j2\pi f \cdot t) \cdot dt \quad (1.6)$$

Les deux relations (1.4) et (1.5) nous permettent de conclure que le spectre d'un signal périodique est discret, car la distance entre deux raies adjacentes est égale à  $f_0$ . Toutefois pour un signal non périodique le spectre devient une fonction continue de la fréquence  $X(f)$ . Cette fonction analyse le signal  $x(t)$  sous forme d'une infinité de composantes sinusoïdales complexes, elle fournit des informations sur la distribution de l'amplitude (phase, énergie, puissance, ...) en fonction de la fréquence.

### 1.5 Puissance et densité spectrale de puissance

Soit  $i(t)$  le courant électrique à travers une résistance ohmique normalisée de  $1\Omega$ , la puissance moyenne sur un intervalle de temps  $T=[t_1, t_2]$  est

$$P_{\text{moy}} = \frac{1}{T} \int_{(T)} R \cdot i^2 \omega dt = \frac{1}{T} \int_{(T)} \frac{1}{R} v^2 \omega dt = \frac{1}{T} \int_{(T)} i^2 \omega dt = \frac{1}{T} \int_{(T)} v^2 \omega dt$$

En pratique, les quantités physiques sont converties en tensions électriques par les transducteurs, alors le signal  $v(t)$  peut représenter n'importe quel paramètre physique: variation de pression, de niveau, de température, déplacement, contrainte, accélération ... etc. Par analogie, on appelé puissance moyenne d'un signal  $x(t)$ , sur un intervalle  $T=[t_1, t_2]$ , la valeur:

$$P_{\text{moy}} = \frac{1}{T} \int_{(T)} |x(t)|^2 \cdot dt \quad (1.7)$$

Cette puissance peut être considérée dans le domaine fréquentiel. Une mesure de la puissance moyenne contenue dans une bande de fréquence  $\Delta f = [f_1, f_2]$

$$P_{\Delta f} = \int_{(\Delta f)} S(f) df$$

La fonction  $S(f)$  représente la distribution de puissance<sup>(1)</sup> dans le domaine fréquentiel, plus communément appelée: densité spectrale de puissance (DSP).

La puissance moyenne totale du signal est donnée par

$$P = \int_{-\infty}^{\infty} S(f) df \quad (1.8)$$

Historiquement, les densités spectrales, étaient obtenues à partir des fonctions de corrélation. La DSP est la transformée de Fourier de la corrélation. Cette relation est connue sous le nom de théorème de Wiener-Kinchine [47], [71], [37].

$$S(f) = \int_{-\infty}^{+\infty} R(\tau) \cdot \text{Exp}\{ -j2\pi f \cdot \tau \} \cdot d\tau \quad (1.9)$$

$$R(\tau) = \int_{-\infty}^{+\infty} S(f) \text{Exp}\{ j2\pi f \cdot \tau \} \cdot df \quad (1.10)$$

où  $R(\tau)$  est la fonction de corrélation du signal  $x(t)$ . Pour un signal à puissance moyenne finie et ergodique  $R(\tau)$  est définie par:

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) \cdot x(t-\tau) \cdot dt \quad (1.11)$$

Si  $X_T(f)$  est la transformée de Fourier de  $x_T(t)$ , avec

$$x_T(t) = \begin{cases} x(t) & |t| \leq T/2 \\ 0 & \text{ailleurs} \end{cases}, \text{ alors on peut montrer [47] que:}$$

$$S(f) = \lim_{T \rightarrow \infty} \frac{1}{T} | X_T(f) |^2 \quad (1.12)$$

(1): Pour les signaux à énergie finie  $S(f)$  sera une densité spectrale énergétique. Pour les signaux à énergie infinie, mais à puissance moyenne totale finie, ce sera une densité spectrale de puissance.

## 1.6 Densité Spectrale de puissance des processus aléatoires

Les signaux aléatoires ne possèdent pas de représentation temporelle analytique. Leurs lois exactes nous échappent, alors pour en déduire une représentation fréquentielle de ces signaux on doit se contenter de leurs propriétés statistiques.

Pour simplifier le problème de l'estimation de la DSP, on considère que les processus aléatoires en question sont stationnaires, c'est-à-dire que leurs propriétés statistiques sont invariables dans le temps, et érgodiques. Cette dernière hypothèse nous permet d'estimer les caractéristiques du processus par l'analyse des moyennes temporelles d'une seule réalisation sur une durée relativement longue, plutôt que de considérer un grand nombre de réalisations, car l'accès aux moyennes d'ensembles est très difficile à réaliser [51].

La fonction d'autocorrélation d'un signal aléatoire réel et stationnaire  $x(t)$  est définie par la moyenne statistique suivante:

$$R(\tau) = \mathbb{E} \left[ x(t)x(t-\tau) \right] \quad (1.13)$$

$R(\tau)$  doit être calculée sur un nombre infini de réalisations  $x_k(t)$ . Mais si l'on admet que la plus part des signaux aléatoires usuels sont la manifestation de processus érgodiques,  $R(\tau)$  peut être définie comme la limite pour  $T \rightarrow \infty$  de la moyenne temporelle :

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} x(t) \cdot x(t-\tau) \cdot dt \quad (1.14)$$

Malheureusement, l'analyse spectrale expérimentale diffère de ce modèle théorique. La difficulté majeure que l'on rencontre c'est que l'on ne peut jamais calculer les véritables moyennes, compte tenu du fait que l'observation du signal  $x(t)$  se fait sur une durée limitée  $T$ . Dans ce cas, il ne s'agit plus d'un problème de calcul ou de mesure, mais principalement d'un problème d'estimation.

## 1.7 Estimation et caractéristiques des estimateurs

Considérons  $N$  échantillons d'une réalisation d'un processus aléatoire stationnaire et ergodique  $[x_1, x_2, \dots, x_N]$ , l'estimation " $\hat{a}$ " d'un paramètre " $a$ " du processus est donnée par

$$\hat{a} = S [x_1, x_2, \dots, x_N]. \quad (1.15)$$

La fonction  $S$  est appelée estimateur du paramètre " $a$ ". On ne jugera la qualité de l'estimation qu'en fonction des résultats obtenus et des propriétés des estimateurs. Les plus importantes de ces propriétés sont: l'absence de biais, la faible variance, l'efficacité et la consistance. Considérons séparément certaines de ces notions .

-Le biais de l'estimation  $B_{\hat{a}}$ , est défini par la différence [47]:

$$B_{\hat{a}} = a - E(\hat{a}) \quad (1.16)$$

C'est le coût de l'absence d'information sur la moyenne du paramètre estimé ( erreur systématique ). L'estimation est dite non biaisée -absolument correcte- si la valeur moyenne de cette estimation sur l'ensemble des échantillons est exactement égale au paramètre estimé c'est-à-dire :  $B_{\hat{a}}=0$ .

-La variance de l'estimation caractérise l'ampleur de la dispersion des mesures par rapport à la valeur moyenne, c'est une mesure de l'étendu de la probabilité de l'estimation. Elle est définie par :

$$\sigma_{\hat{a}}^2 = E \left[ \hat{a} - E(\hat{a}) \right]^2 = E(\hat{a}^2) - \left[ E(\hat{a}) \right]^2. \quad (1.17)$$

L'estimation est dite consistante si elle converge en probabilité vers le paramètre estimé lorsque le nombre d'échantillons augmente indéfiniment c-à-d :

$$B_{\hat{a}} \longrightarrow 0 \quad \text{et} \quad \sigma_{\hat{a}}^2 \longrightarrow 0 \quad \text{lorsque} \quad N \longrightarrow \infty$$

En général il est impossible d'annuler la variance et le biais en même temps, la diminution de l'un provoque l'augmentation de l'autre, et vice-versa. Mais un bon estimateur doit posséder un biais et une variance aussi petits que possible.

-L'estimation est dite efficace si pour un biais donné, la l'erreur quadratique moyenne (eqm) de l'estimation, n'est pas supérieur à l' e.q.m pour tout autre estimation. Un estimateur non biaisé est dit efficace si sa variance est égale à la borne de Cramer-Rao (borne inférieure de la variance) [47].

L'erreur quadratique moyenne (e.q.m) définie par

$$eqm = E((a - \hat{a})^2) = B^2 + \sigma_a^2 \quad (1.18)$$

est un critère de compromis entre le biais et la variance et caractérise l'efficacité d'un estimateur .

### 1.8 Méthodes fondamentales d'estimation spectrale

Les notions de base étant définies, une question se pose: quelles sont les principales méthodes et techniques utilisées afin d'estimer la DSF ? Avant de répondre à cette question, on doit préciser qu'on s'intéresse à l'estimation numérique de la DSF des phénomènes aléatoires. Ce choix est motivé par le fait que dans de nombreuses recherches scientifiques et applications techniques, les résultats des observations (*expérience, essais, mesures ... etc*) sont fixés par un dispositif analogique sous la forme de réalisations  $x(t)$  de durée finie d'un certain processus aléatoire, soit par un dispositif numérique sous la forme d'un ensemble fini de nombres  $\{x_1, x_2, \dots, x_N\}$ . Avec l'hypothèse que la variable aléatoire  $x(t)$  soit échantillonné aux instants  $\{t_1, t_2, \dots, t_N\}$ , en respectant le critère de Shannon, on obtient un ensemble fini de nombres  $x_i = x(t_i)$ ,  $i \in [1, \dots, N]$  et les résultats se présentent toujours sous forme discrète. Cette représentation est le seul moyen d'aborder le problème par voie numérique, et d'exploiter la puissance de calcul des ordinateurs.

En ce qui concerne les techniques d'estimation de la DSP, on dispose aujourd'hui d'un arsenal de méthodes qu'on peut classer en deux grandes catégories (figure 1.2):

-Méthodes classiques basées sur la transformation de Fourier (périodogramme, corrélogramme ).

-Méthodes paramétriques ou modernes basées sur la modélisation des signaux. Les modèles généralement utilisés sont autorégressifs (AR), à moyenne adaptée (MA) et ARMA. D'autres méthodes sont également utilisées notamment la méthode d'estimation au sens du maximum de vraisemblance, méthode de Pisarenko et méthode de Prony. Ces dernières utilisent d'autres critères d'estimation et sont basées sur des approches différentes de la modélisation autorégressive, néanmoins une relation peut toujours exister [12], [68].

Dans ce mémoire, après un rappel sur les méthodes classiques, on ne considérera que les méthodes paramétriques autoregressives. Une étude sera menée pour établir leurs performances et leurs limitations. Le chapitre suivant sera consacré à un aperçu sur l'estimation spectrale par les méthodes classiques.

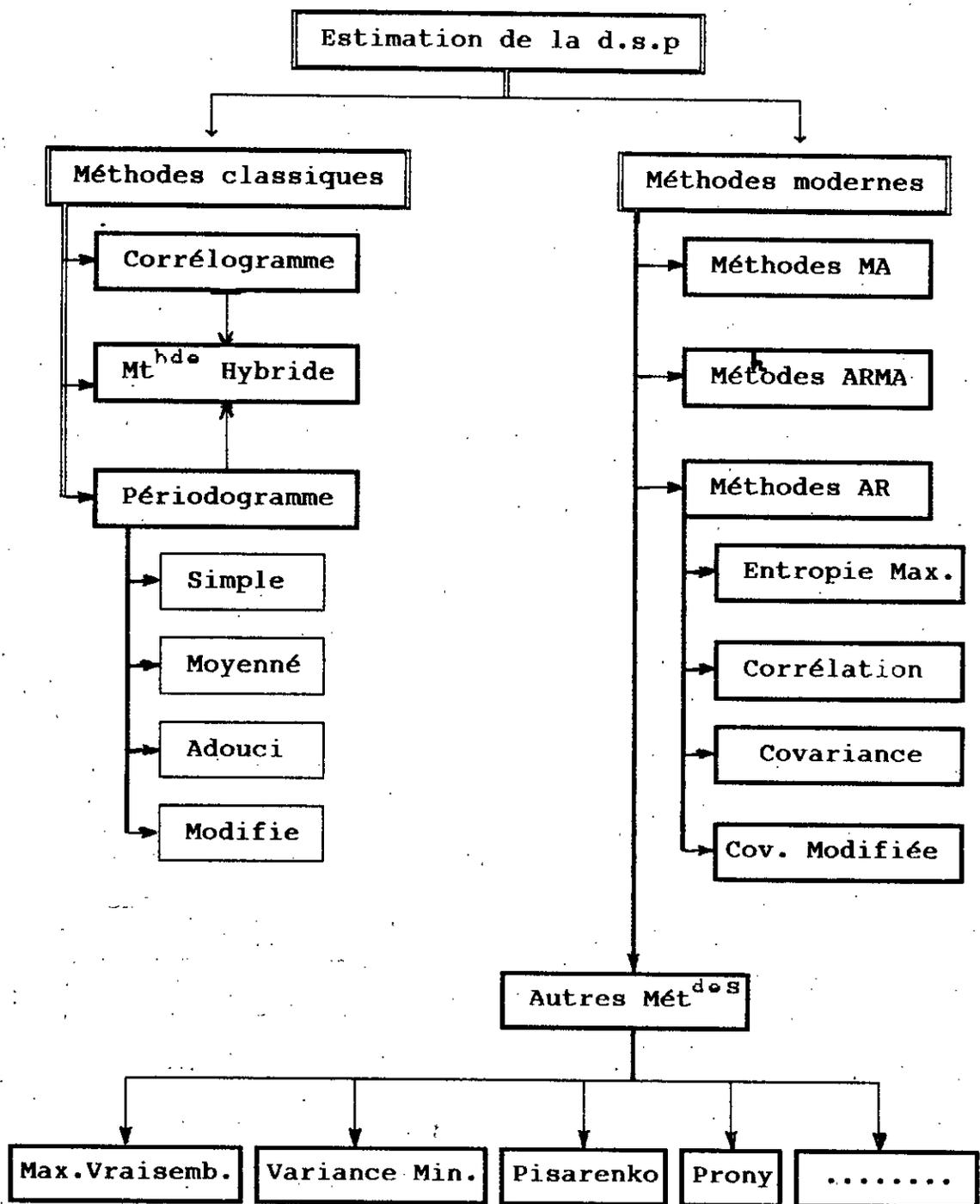


figure 1.2: Les deux grandes familles de méthodes d'estimation de la densité spectrale de puissance

ESTIMATEURS SPECTRAUX CLASSIQUES

2.1 Introduction

L'analyse spectrale peut s'effectuer, soit analogiquement par filtrage électrique, soit numériquement à l'aide d'analyseurs et calculateurs numériques.

La technique analogique, basée sur le filtrage en bande étroite [50], a été pour longtemps la plus utilisée pour la mesure de la DSP. Mais les analyseurs les plus récents sont exclusivement numériques et basés sur l'emploi de la transformation de Fourier discrète [50].

D'après les deux relations (1.9) et (1.12) du chapitre précédent, la DSP peut être estimée de deux manières:

-directement à partir des données disponibles, en élevant au carré la transformée de Fourier du signal. C'est ce que nous appelons méthode du périodogramme. Une fenêtre appropriée  $w(k)$  peut être utilisée (fig. 2.1).

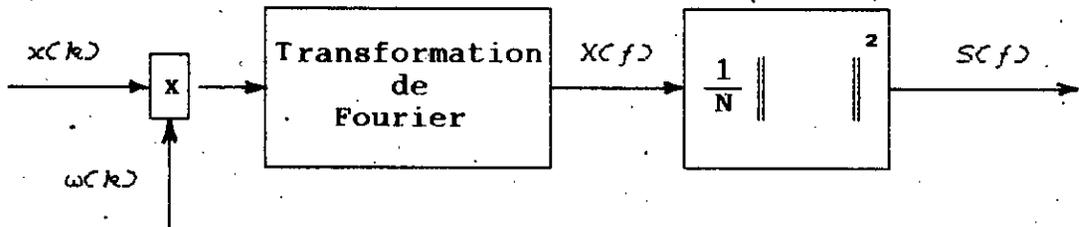


Schéma 2.1: Méthode du périodogramme.

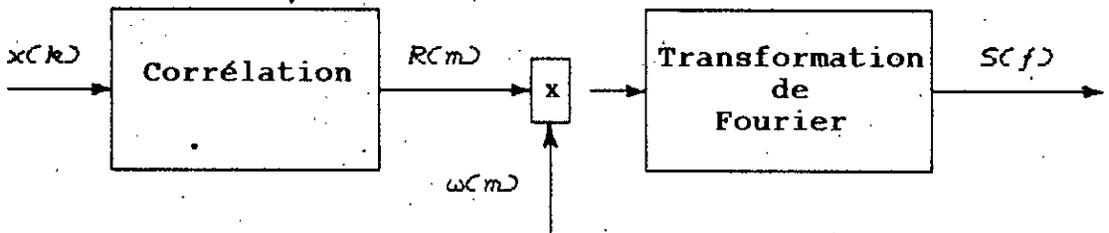


Schéma 2.2: Méthode du Corrélogramme.

-Indirectement, en évaluant successivement, la fonction d'auto-corrélation du signal, avec une éventuelle pondération par une fonction fenêtre  $\omega(m)$ , puis la transformée de Fourier de celle-ci. Cette procédure est connue comme étant la méthode du corrélogramme ou de Blackman-Tukey [8].

## 2.2 Méthode du Corrélogramme

La corrélation est l'opération qui consiste à mettre en évidence le lien qui peut exister entre deux signaux, pris dans des intervalles de temps décalés d'une certaine durée  $\tau$ .

Dans le cas de l'autocorrélation, cette opération permet de voir en quoi un signal, à un instant donné, est influencé par son état un instant  $\tau$  avant. Pour l'intercorrélation, c'est une mesure de la similitude et de l'interaction entre deux signaux différents.

### 2.2.1 Estimateurs pour la fonction d'autocorrélation

Par définition, la fonction d'autocorrélation d'un signal aléatoire, réel, stationnaire et discret, est donnée par:

$$R(m) = E \left[ x(k) \cdot x(k+m) \right] \quad (2.1)$$

qui, si le signal est ergodique équivaut à

$$R(m) = \lim_{N \rightarrow \infty} \left[ \frac{1}{N} \sum_{k=-N/2}^{N/2} x(k) \cdot x(k+m) \right] \quad (2.2)$$

En utilisant un nombre fini  $N$  d'échantillons du signal, la valeur de  $R(m)$ , ne peut être estimée qu'à partir de  $(N-m)$  valeurs du signal:

$$R(m) = \frac{1}{(N-m)} \sum_{k=0}^{N-m-1} x(k) \cdot x(k+m) \quad , \quad m = 0, 1, 2 \dots L. \quad (2.3)$$

$$\text{et } R(-m) = R(m)$$

Certains auteurs [8], pour obtenir une fonction de corrélation avec une bonne précision, ont suggéré une valeur maximale du retard :  $m \leq L = N/10$ . Mais rien nous empêche d'adopter une autre valeur (plus grande), si cela s'avère nécessaire.

La moyenne statistique de cette estimation est :

$$\mathbb{E}[\check{R}(m)] = \frac{1}{(N-m)} \sum_{k=0}^{N-m-1} \mathbb{E}[x(k) \cdot x(k+m)] = R(m) \quad (2.4)$$

compte tenu de la définition (1.16), l'estimation est non biaisée.

Le calcul de la variance s'avère très compliqué et nécessite la connaissance des moments d'ordre supérieurs à deux. Dans le cas d'un processus Gaussien, centré, une approximation de la variance est donnée par [38] :

$$\text{Var}[\check{R}(m)] \simeq \frac{1}{(N-|m|)^2} \sum_{l=-N-|m|-1}^{N-m-1} \binom{N-m-|l|}{l} \cdot \left[ R(m) + R(m+l) \cdot R(l-m) \right]^2$$

et pour des valeurs de  $N$  très supérieur à  $m$ , la variance est inversement proportionnelle à  $N$

$$\text{Var}[\check{R}(m)] \simeq \frac{1}{N^4} \sum_{l=-\infty}^{+\infty} \left[ R(m) + R(m+l) \cdot R(l-m) \right]^2 \quad (2.5)$$

Il s'agit donc d'un estimateur non biaisé et consistant. Toutefois pour des valeurs de  $m$  proches de  $N$ , celui-ci n'est pas très utile : la variance devient trop élevée, et l'estimation n'est plus satisfaisante. Alors on utilise un autre estimateur donné par

$$\hat{R}(m) = \frac{1}{N} \sum_{k=0}^{N-m-1} x(k) \cdot x(k+m) \quad \text{avec } m = 0, 1, 2, \dots, L. \quad (2.6)$$

$$\text{et } \hat{R}(-m) = \hat{R}(m)$$

$$\text{d'où } \hat{R}(m) = \frac{N-|m|}{N} \check{R}(m) \implies \mathbb{E}[\hat{R}(m)] = \frac{N-|m|}{N} R(m) \quad (2.7)$$

L'estimateur est donc biaisé, car

$$B(\hat{R}(m)) = R(m) - E(\hat{R}(m)) = \left(\frac{m}{N}\right) \cdot R(m) \quad (2.8)$$

Lorsque  $N \gg m$ , la variance peut être approximée par la même relation (2.5). On remarque que quand  $N$  tend vers l'infini, la variance ainsi que le biais tendent vers zéro: l'estimateur est asymptotiquement consistant. Pour cet estimateur même si  $N$  est fini, le facteur extérieur de (2.5) ne dépend pas de  $m$ , ce qui rend la variance faible même si  $m$  est proche de  $N$ .

### 2.2.2 Estimateurs pour la densité spectrale de puissance

L'estimateur de la DSP est la transformée de Fourier de l'estimateur de la fonction de corrélation:

$$\hat{S}(f) = \sum_{m=-L+1}^{L-1} \hat{R}(m) \cdot \exp(-2\pi f m T_0) \quad (2.9)$$

où  $\hat{R}(m)$  est l'estimateur biaisé ou non biaisé

La fonction d'autocorrélation  $\hat{R}(m)$  est calculée pour des retards variant de 0 à  $(L-1)T_0$  avec un pas d'échantillonnage  $\Delta\tau = T_0$ .

L'espérance mathématique de cet estimateur est

$$E(\hat{S}(f)) = \sum_{m=-L+1}^{L-1} E[\hat{R}(m)] \cdot \exp(-2\pi f m T_0)$$

avec un estimateur non biaisé pour  $\hat{R}(m)$

$$E(\hat{S}(f)) = \sum_{m=-\infty}^{+\infty} \left[ R(m) \cdot \omega_{2L-1}(m) \right] \cdot \exp(-2\pi f m T_0)$$

où  $\omega_{2L-1}(m)$  est une fenêtre rectangulaire définie sur l'intervalle  $[-(L-1)T_0, \dots, (L-1)T_0]$ , alors:

$$E(\hat{S}(f)) = S(f) * \left[ \frac{\sin(\pi f (2L-1) T_0)}{\sin(\pi f T_0)} \right] \cdot \quad -\frac{1}{2T_0} \leq f \leq \frac{1}{2T_0} \quad (2.10)$$

où  $*$  représente le produit de convolution.

Ainsi le fait d'utiliser un nombre fini d'échantillons pour la fonction d'autocorrélation, et même avec un estimateur non biaisé pour celle-ci, l'estimateur de la DSP est biaisé. La véritable densité spectrale se trouve convoluée par la T.F de la fenêtre rectangulaire.

Tout comme pour la fonction de corrélation, on définit un autre estimateur pour la DSP, en utilisant l'estimateur biaisé pour la corrélation (2.6)

$$\hat{S}(f) = \sum_{m=-L+1}^{L-1} \hat{R}(m) \cdot \exp(-2\pi f m T) \quad (2.11)$$

La valeur moyenne de l'estimation est

$$\mathbb{E}(\hat{S}(f)) = \sum_{m=-(L-1)}^{L-1} \mathbb{E}(\hat{R}(m)) \cdot \exp(-2\pi f m T) = \sum_{m=-\infty}^{+\infty} \omega_{B_M}(m) R(m) \exp(-2\pi f m T) \quad (2.12)$$

avec  $\omega_{B_M}(m)$  fenêtre triangulaire de Bartlett définie par:

$$\omega_{B_M}(m) = \begin{cases} \frac{N-|m|}{N} & \text{si } -(L-1) < m < L-1 \\ 0 & \text{ailleurs} \end{cases}, \quad M=2L-1$$

Alors dans l'intervalle  $\left[-\frac{1}{2T_0}, \frac{1}{2T_0}\right]$  on a:

$$\mathbb{E}(\hat{S}(f)) = S(f) * \frac{1}{L} \left( \frac{\sin(\pi f L T_0)}{\sin(\pi f T_0)} \right)^2 \quad (2.13)$$

dans ce cas, l'erreur du biais est due à la convolution de  $S(f)$  par le spectre de la fonction triangulaire.

Pour améliorer l'estimation et réduire l'effet de la convolution implicite avec le spectre de la fenêtre rectangulaire ou triangulaire, Blackman et Tukey [8] proposèrent un estimateur en utilisant une pondération  $\omega(m)$  pour la corrélation autre que la fenêtre  $\omega_r$  ou  $\omega_B$ :

$$\hat{S}_{BT}(f) = \sum_{m=-(L-1)}^{L-1} \hat{R}(m) \cdot \omega(m) \cdot \exp(-2\pi f m T_0) \quad (2.14)$$

pour l'autocorrélation, on utilisera l'estimateur non biaisé défini par (2.3). Si  $W(f)$  est le spectre de la fenêtre temporelle  $\omega(m)$ , alors la moyenne statistique de la DSP est donnée par:

$$E(S(f)) = S(f) * W(f) \implies E(R(m)) = R(m) \cdot \omega(m) \quad (2.15)$$

La relation 2.15 formalise la convolution de la DSP par le spectre de la fonction fenêtre. Cette convolution va introduire des lobes secondaires. Ces lobes ont une influence non négligeable et peuvent engendrer de oscillations parasites autour des pics de la DSP. Parfois ces lobes latéraux sont négatifs, et vont introduire des parties négatives sur la DSP !

Pour avoir le minimum de déformation, la fenêtre spectrale  $W(f)$ , doit présenter des lobes secondaires de très faibles amplitudes, et si possible pas de lobes négatifs. Enfin, pour que l'estimation soit asymptotiquement non biaisée, et afin de conserver la puissance totale du signal donnée par  $R(0) \cdot \omega(0)$ , il est nécessaire que la fenêtre temporelle  $\omega(m)$  prenne une valeur égale à l'unité pour un retard nul:  $\omega(0) = 1$ . Cette condition est implicitement satisfaite pour la plupart des fonctions fenêtres.

### 2.3 Méthode du périodogramme

L'utilisation de la transformée de Fourier pour l'obtention d'une représentation dans le domaine des fréquences d'un signal, impose qu'il soit considéré comme un signal périodique, d'où la terminologie ancienne "périodogramme" pour les méthodes basées essentiellement sur la transformation de Fourier discrète (TFD).

#### 2.3.1 Transformation de Fourier discrète (TFD)

La version numérique de la transformée de Fourier continue d'un signal discret  $x(kT_0)$  est donnée par

$$X(f) = \sum_{k=-\infty}^{+\infty} x(kT_0) \cdot \exp(-2\pi f k T_0) \quad (2.16)$$

$T_0$  est la période d'échantillonnage

Mais comme on n'a qu'un nombre fini d'échantillons  $x(k)$ , et en même temps on doit échantillonner la transformée  $X(f)$ , la relation (2.16) s'écrit:

$$X(n) = \sum_{k=0}^{N-1} x(k) \cdot \exp\left(-2\pi \frac{n \cdot k}{N}\right) \quad (2.17)$$

L'intérêt de la T.F.D réside dans l'existence d'un algorithme de calcul rapide appelé FFT (Fast Fourier Transform, voir Annexe A1).

### 2.3.2 Estimateur simple

En vertu du théorème de Wiener-Kinchine, appliqué à une estimation biaisée de la fonction d'autocorrélation :

$$\hat{S}(f) = \sum_m \hat{R}(m) \cdot \exp(-2\pi f m T_0) \quad (2.18)$$

on a choisi l'estimateur biaisé  $\hat{R}(m)$ , car il donne une densité spectrale de puissance toujours supérieure à zéro. En utilisant la définition (2.6), on obtient l'estimation numérique du périodogramme simple selon le schéma 2.1 (voir Annexe A1) :

$$\hat{S}(f) = \frac{1}{N} \|X(f)\|^2 \quad \text{où} \quad X(f) = \sum_{k=0}^{N-1} x(k) \cdot \exp(-2\pi \cdot f \cdot k T_0) \quad (2.19)$$

L'estimation est asymptotiquement non biaisé, mais pour un nombre limité d'échantillons, le biais est différent de zéro. Le calcul de la variance est très difficile. Avec l'hypothèse d'un processus Gaussien, elle est donnée par l'approximation suivante [47]:

$$\text{var}(\hat{S}(f)) = \|S(f)\|^2 \left[ 1 + \left( \frac{\text{Sin}(2\pi f N T_0)}{N \cdot \text{Sin}(2\pi f T_0)} \right)^2 \right] \quad (2.20)$$

Soit pour N très grand

$$\begin{aligned} \text{var}(\hat{S}(f)) &\cong \|S(f)\|^2 && \text{si } f \neq 0, f \neq 1/2T_0 \\ \text{var}(\hat{S}(f)) &\cong 2 \cdot \|S(f)\|^2 && \text{pour } f=0 \text{ et } f=1/2T_0 \end{aligned}$$

La variance est proportionnelle au carré de la D.S.P recherché, et ne tend pas vers zéro avec l'augmentation de N, par conséquent le périodogramme simple n'est pas consistant.

Pour remédier à ce problème, il faut établir d'autres estimateurs qui permettent de réduire la variance, en agissant sur les paramètres d'estimation.

### 2.3.3 Estimateur Moyenné

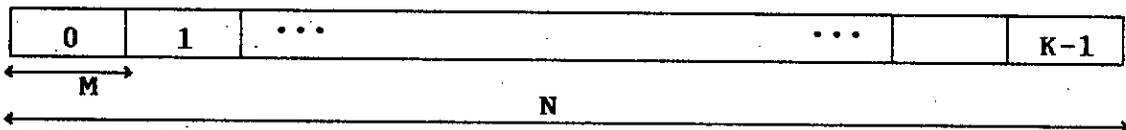
Une méthode directe pour réduire la variance de l'estimateur simple, consiste à segmenter le signal en plusieurs sections, évaluer le périodogramme simple de chaque section puis calculer la moyenne des estimations. Cette procédure est connue comme étant la méthode de Bartlett [15].

Soit

K: Nombre de sections;

M: Nombre d'échantillons par section;

N: Durée d'observation du signal tel que  $N \geq K.M$ .



L'estimateur moyenné est donné par

$$\hat{S}_B(f) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{S}^{(k)}(f) \quad (2.21)$$

$$\text{avec } \begin{cases} \hat{S}^{(k)}(f) = \frac{1}{M} \left| \sum_{n=0}^{M-1} x^{(k)}(nT_0) \cdot \exp(-2\pi n f T_0) \right|^2 \\ x^{(k)}(n) = x(k.M+n) , \quad kM \leq n \leq (k+1)M-1 \end{cases}$$

avec l'hypothèse que les spectres de chaque section du signal sont indépendants, il est bien connu [38] que la variance de la densité spectrale de puissance obtenue par cette méthode est proportionnelle à:

$$\text{var}(\hat{S}_B(f)) \approx \frac{\|S(f)\|^2}{K} \quad (2.22)$$

Ainsi pour diminuer la variance, il faut augmenter le nombre de sections K.

Le biais de l'estimateur moyenné ne dépend que de la longueur des sections M:

$$\mathbb{E} \left[ \widehat{S}_D(f) \right] = \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E} \left[ \widehat{S}^{(k)}(f) \right] = \mathbb{E} \left[ \widehat{S}^{(k)}(f) \right] \quad (2.23)$$

Une comparaison de ces résultats avec ceux de l'estimateur simple  $\widehat{S}(f)$ , montre que pour le périodogramme moyenné, le biais est plus grand, tandis que la variance est K fois plus petite. D'après les expressions (2.22) et (2.23), le biais diminue avec l'augmentation de M, la variance diminue avec l'augmentation de K, alors pour un  $N=K.M$  fixe, ceci conduit à un compromis.

#### 2.3.4 Estimateur Adouci

Daniel [58] proposa une autre alternative pour réduire la variance de l'estimateur simple: filtrer les fluctuations rapides de la densité spectrale en fonction de la fréquence (lissage fréquentiel). Ce qui s'exprime par la convolution:

$$\widehat{S}_D(f) = \widehat{S}(f) * \psi(f) \quad (2.24)$$

où  $\psi(f)$  est une fenêtre spectrale d'adoucissement.

Si les données sont pondérées par  $\omega(k)$ ,  $\psi(f)$  est égale à un facteur  $1/N$  près, au carré de la TF de la fenêtre  $\omega(k)$ , (voir Annexe 1). Par la TF inverse de l'équation (2.24)

$$\widehat{R}_D(m) = \widehat{R}(m) \cdot \phi(m) \quad , \quad \phi(m) = \text{TF}^{-1}[\psi(f)] = \omega(k) * \omega(-k) \quad (2.25)$$

Il est important de rappeler que la fonction  $\phi(m)$  doit être normalisée c'est-à-dire  $\phi(0)=1$ . La valeur moyenne de l'estimation est donnée par

$$\mathbb{E} \left[ \widehat{S}_D(f) \right] = \mathbb{E} \left[ \widehat{S}(f) * \psi(f) \right] \approx S(f) * \psi(f) \quad (2.26)$$

Le biais de l'estimateur adouci, dépend de la forme et de la largeur de la fenêtre.

### 2.2.5 Estimateur modifié

Welch [70] a modifié l'estimateur moyenné de Bartlett de telle sorte que le signal est divisé en K sections, avec la possibilité de recouvrement entre deux sections adjacentes. En plus, chaque section est pondérée par une fonction fenêtre  $\omega(k)$ , avant que ne lui soit appliquée la TFD.

$$\hat{S}_w(f) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{S}^{(k)}(f) \quad (2.27)$$

$$\begin{cases} S^{(k)}(f) = \frac{1}{MU} \left| \sum_n x^{(k)}(n) \cdot \exp(-2\pi \cdot n \cdot \Delta f \cdot T_0) \right|^2 \\ x^{(k)}(n) = x(k \cdot M + n) \cdot \omega(n), \quad kM \leq n \leq (k+1)M-1 \end{cases}$$

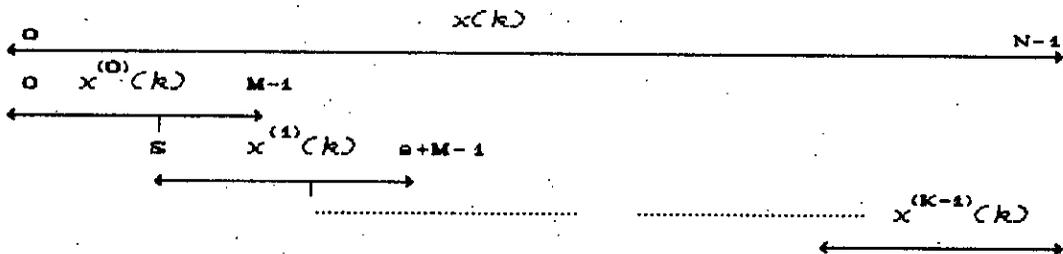
M: Nombre d'échantillons par section

D: Distance entre 2 sections successives.

N: Durée d'observation du signal .

K: Nombre de sections:

$$K = \text{Valeur entière de } \left\lfloor \frac{N-M}{D} \right\rfloor + 1 \quad (2.28)$$



La fenêtre a été introduite afin d'améliorer le biais de l'estimation. U est un facteur de normalisation, afin que l'estimateur soit asymptotiquement non biaisé

$$E(\hat{S}_w(f)) = E(\hat{S}(f)) / U = S(f) * \frac{|W(f)|^2}{U} \quad (2.29)$$

la TF inverse de l'équation (2.29) donne la valeur moyenne de la fonction de corrélation:

$$\mathbb{E}(\hat{R}_V(m)) = R(m) \cdot \frac{\phi(m)}{U} \quad \text{avec} \quad \phi(m) = \sum_{k=0}^{M-m-1} \omega(k+m) \cdot \omega(k) \quad (2.30)$$

$$\hat{R}_V(0) = R(0) \implies U = \phi(0) = \sum_{k=0}^{M-1} \omega^2(k) \quad (2.31)$$

Comme pour l'estimateur moyenné de Bartlett, la variance est inversement proportionnelle au nombre de sections:

$$\text{var}(\hat{S}_w(f)) \approx \frac{S^2(f)}{K} \quad (2.32)$$

Le recouvrement va diminuer la variance, car il y a davantage de spectres partiels pour une même durée du signal. Mais il est difficile de prévoir l'amélioration, vue que l'hypothèse que les spectres de chaque section sont indépendants n'est pas vérifiée dans ce cas.

#### 2.4 Méthode hybride

Nous avons vu que pour assurer une certaine stabilité statistique de l'estimation, on est amené à l'une des deux techniques:

-Moyennage de plusieurs périodogrammes (méthode de Welch).

-Lissage spectral (Méthode de Blackman-Tukey, de Daniel).

Mais il est possible d'utiliser une combinaison de ces deux méthodes pour estimer la DSP, procédure qu'on appelle "Méthode hybride" [47].

La première étape consiste à estimer la fonction de corrélation en utilisant le théorème de Wiener-Kinchine. Pour cela la DSP est estimée par le périodogramme modifié de Welch selon l'équation (2.27). Si les données sont pondérées par une fenêtre  $\omega(k)$ , définie dans l'intervalle  $[0, M-1]$ , ( $M$  = nombre de points dans une section) et dont la TFD est égale à  $W(f)$ , alors

$$\mathbb{E}(\hat{S}_w(f)) = S(f) * |W(f)|^2 \quad (2.33)$$

d'où par transformation de Fourier inverse, et après normalisation:

$$E(\hat{R}_w(m)) = R(m) \cdot \frac{\hat{\phi}(m)}{\hat{\phi}(0)} \quad \text{avec} \quad \hat{\phi}(m) = \text{TF}^{-1} \left[ |W(f)|^2 \right] \quad (2.34)$$

Autrement dit, pour retrouver la fonction de corrélation on effectue la division de la fonction  $\hat{R}_w(m)$  par la fenêtre  $\hat{\phi}(m)$ , et selon ce que nous avons décrit en (2.2.2), on peut multiplier le résultat par une fenêtre (effective)  $\omega_e(m)$  normalisée ( $\omega_e(0)=1$ ) selon le schéma 2.2. Le support de cette fenêtre étant au plus égal à  $2 \cdot L$  avec  $L < M$ .

L'ensemble de ces opérations, revient à pondérer la fonction d'autocorrélation -estimée par la transformée de Fourier inverse du périodogramme moyenné de Daniel- par une fenêtre  $\omega(m)$  tel que (voir figure 2.3):

$$\omega(m) = \frac{\omega_e(m)}{\hat{\phi}(m)/\phi(0)} \quad (2.35)$$

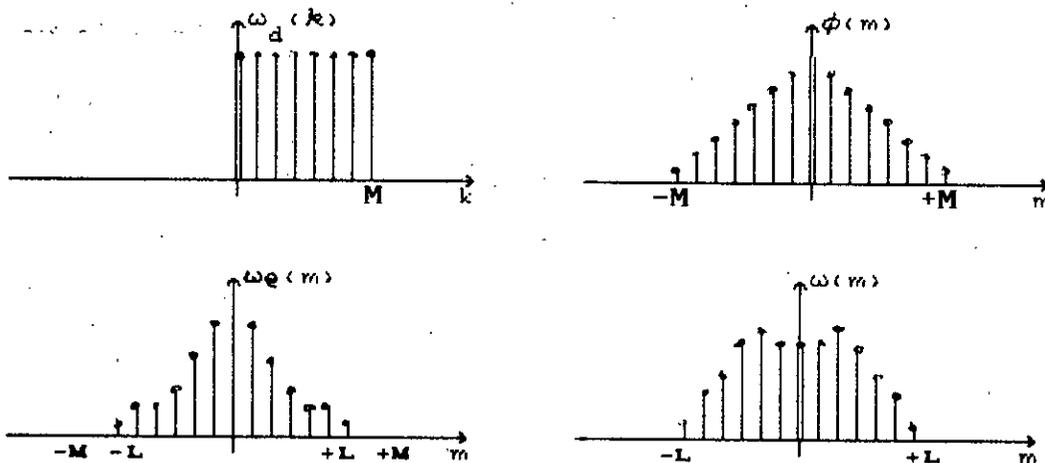


figure 2.3 Fenêtre de correction de la méthode hybride dans le cas où  $\omega_d(k)$  est une fenêtre rectangulaire

La dernière étape ne serait que le calcul de la TFD de  $\hat{R}(m)$ , pour arriver à une estimation de la DSP.

Paradoxalement, cette méthode s'avère numériquement plus efficace, car la pondération directe des différentes sections du signal n'est pas nécessaire, et il suffit de pondérer la fonction d'autocorrélation estimée par la fenêtre  $w(m)$  définie en par l'expression 2.35.

## 2.5 Effet des fenêtres de pondération

Les fenêtres, temporelles ou spectrales, jouent un rôle important en estimation spectrale. Comme on l'a déjà décrit dans (2.2.2) et (2.3.4), la pondération est effectuée pour réduire la variance du corrélogramme et diminuer le biais dans le cas du périodogramme. L'emploi de ces fenêtres se fait selon les schémas blocs (2.1) et (2.2). Elles sont appliquées soit aux données  $x(k)$ , soit à la fonction de corrélation; selon la méthode utilisée.

La DSP estimée est perturbée par la TF de la fenêtre temporelle: une séquence finie de données  $x(k)$ , donne après analyse des lobes secondaires supplémentaires dont le contenu spectral ne provient pas du signal lui même, mais de l'effet du filtrage dû à une fuite d'énergie ("leakage"). Sur ce point, le corrélogramme est vulnérable, car non seulement il introduit des lobes secondaires plus importants que ceux introduits par le périodogramme (figure 2.9), mais en plus, certaines fenêtres spectrales possèdent des lobes secondaires négatifs (voir l'exemple de la figure 2.10).

### 2.5.1 Caractéristiques des fenêtres

La fenêtre la plus simple est la fonction rectangulaire (fenêtre naturelle). Les discontinuités de cette fenêtre font apparaître dans la D.S.P des lobes secondaires relativement importants (phénomène de Gibbs). Ceux-ci peuvent masquer la présence d'une raie de faible puissance située au voisinage d'une raie de forte puissance. Pour palier à cet inconvénient,

d'autres fenêtres à variations plus graduelles ont été proposées. Les fenêtres les plus utilisées sont données dans l'annexe A2.

La meilleure fenêtre doit présenter un pic central le plus étroit possible, peu de lobes secondaires de très faibles amplitudes, et si possible pas de lobes négatifs. L'examen des caractéristiques des fenêtres les plus fréquemment utilisées (Annexe A2) montre que le rapport de l'amplitude du lobe principal au 1<sup>er</sup> lobe secondaire de la fonction  $20.\log|W(f)|$ , est indépendant du nombre d'échantillons  $N$ . La largeur du lobe principal est fonction de  $N$ , la fenêtre est d'autant plus étroite que  $N$  est grand. On remarque aussi que le lobe centrale est d'autant plus large que l'atténuation des lobes secondaires est plus grande.

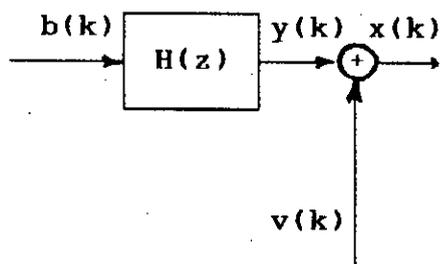
## 2.6 Exemple d'application

Pour fixer les idées, on considère un signal  $y(k)$  obtenu par passage d'un bruit blanc centré  $b(k)$  de variance unité à travers un filtre de fonction de transfert  $H(z)$ :

$$H(z) = (1-z^{-1})(1+z^{-1})^2$$

La DSP exacte du signal de sortie est donnée par l'expression:

$$S(f) = 16.\sin^2(2\pi f) \left[ 1 + \cos(2\pi f) \right]^2$$



Et afin d'étudier l'aptitude des différentes méthodes d'estimation à distinguer entre deux raies spectrales voisines, de même puissance (résolution fréquentielle), nous ajoutons la suite  $v(k)$ :

$$v(k) = 2\cos(k\pi/2) + 2\cos(11k\pi/20).$$

Le signal  $x(k) = y(k) + v(k)$  est analysé par plusieurs corrélogrammes (figures: 2.1, 2.2, 2.3 & 2.4) et périodogrammes (figures: 2.5, 2.6, 2.7 & 2.8) avec les paramètres indiqués en légendes.

De cet exemple il ressort que, pour les méthodes d'estimation de la DSP basées sur la TF, le nombre total d'échantillons ( $N$ ), le nombre d'éléments de la fonction d'autocorrélation ( $L$ ), le nombre de sections ( $K$ ), le nombre de points par section ( $M$ ), ainsi que la fenêtre de pondération  $w(k)$ , sont autant de paramètres à ajuster afin d'améliorer la qualité de l'estimation et arriver à un meilleur compromis entre bonne résolution et faible variance, puisque, si par variation d'un des paramètres cités ci-dessus (segmentation, pondération ...), on projete de réduire les incertitudes statistiques (variance), il peut en résulter une dégradation de la résolution.

Une dernière remarque concernant les méthodes d'évaluation de la D.S.P par TFD ou FFT. Avec la méthode de calcul par TFD directe, on est maître du choix du pas d'échantillonnage, et on peut calculer autant de points qu'on jugera nécessaires, non seulement pour que le spectre soit bien échantillonné, mais pour avoir une bonne visualisation. Les méthodes de calcul de la TFD utilisant l'algorithme FFT, calculent à partir de  $N$  ( $N=2^P$ ) points du signal temporel  $x(k)$ ,  $N$  points de la DSP. La répartition spectrale de ces points ne peut être modifiée par l'utilisateur; l'échantillonnage temporel et fréquentiel sont dépendants l'un de l'autre. Dans certains cas, pour calculer plus de points de la D.S.P qu'il y en a dans le signal, on utilise l'algorithme de Horner [47], comme on peut augmenter le nombre de points en ajoutant des zéros au signal  $x(k)$ .

## 2.7 Conclusions

On a résumé dans ce chapitre, les principales méthodes classiques d'estimation spectrale. La procédure générale de 3 méthodes (Périodogramme, Corrélogramme, méthode hybride) est présentée dans l'algorithme 2.1 ci-après.

On a vu qu'on peut utiliser une multitude de fenêtres, mais le choix d'une fonction de pondération dépend d'une part du but recherché: bonne résolution, précision statistique, et également, bien sûr, de la forme du spectre traité. En pratique, une procédure qui ne serait être recommandée, c'est d'estimer la DSP en utilisant une bonne fenêtre de pondération, afin d'avoir la forme générale du spectre (bonne stabilité statistique), puis sans pondération, pour en étudier les détails (bonne résolution). Toutefois, la résolution des méthodes classiques est limitée par le nombre d'échantillons disponibles: deux raies spectrales associées à deux sinusoides dans le signal ne peuvent être distinguées que si elles sont séparées par un intervalle  $\Delta f$  tel que  $\Delta f \geq 1/N$ .

En résumé, les limitations des méthodes classiques sont des conséquences de l'hypothèse selon laquelle on considère que la fonction de corrélation est nulle en dehors de l'intervalle d'observation. Les méthodes modernes consistent à prolonger la fonction de corrélation au delà du support connu, en s'appuyant sur une connaissance à priori de la nature du signal, et en utilisant certains critères.

\* Acquisition des données

-Echantillonnage, quantification .

\* Choix de la méthode et des paramètres d'estimation

-Choix de la fenêtre de pondération.

-Méthode du périodogramme:

\_ Nombre de Sections (K)

\_ Distance entre deux sections successives (D)

\_ Nombre de points dans chaque section (M)

$K = \text{partie entière de } \left( \frac{N-D}{M} \right) + 1, \text{ et } D \leq M$

\_ Méthode du corrélogramme

\_ Nombre maximal d'éléments de la fonction de corrélation (L):  $L \leq N/10$ .

\* Estimation de la D.S.P

-Périodogramme

-Corrélogramme

-Méthode hybride

\_ Périodogramme de Welch, fenêtre naturelle

\_ TF inverse du périodogramme

\_ Pondération de la fonction de corrélation

fenêtre de correction (Equation 2.35 )

-TF de la corrélation estimée et pondérée.

\* Amélioration de la qualité d'estimation

-Ajuster les paramètres d'estimation K, M, S, L,  $\omega(k)$ , jusqu'à l'obtention d'un compromis acceptable résolution/stabilité.

Algorithme 2.1

Procédure d'estimation de la DSP par les méthodes classiques.

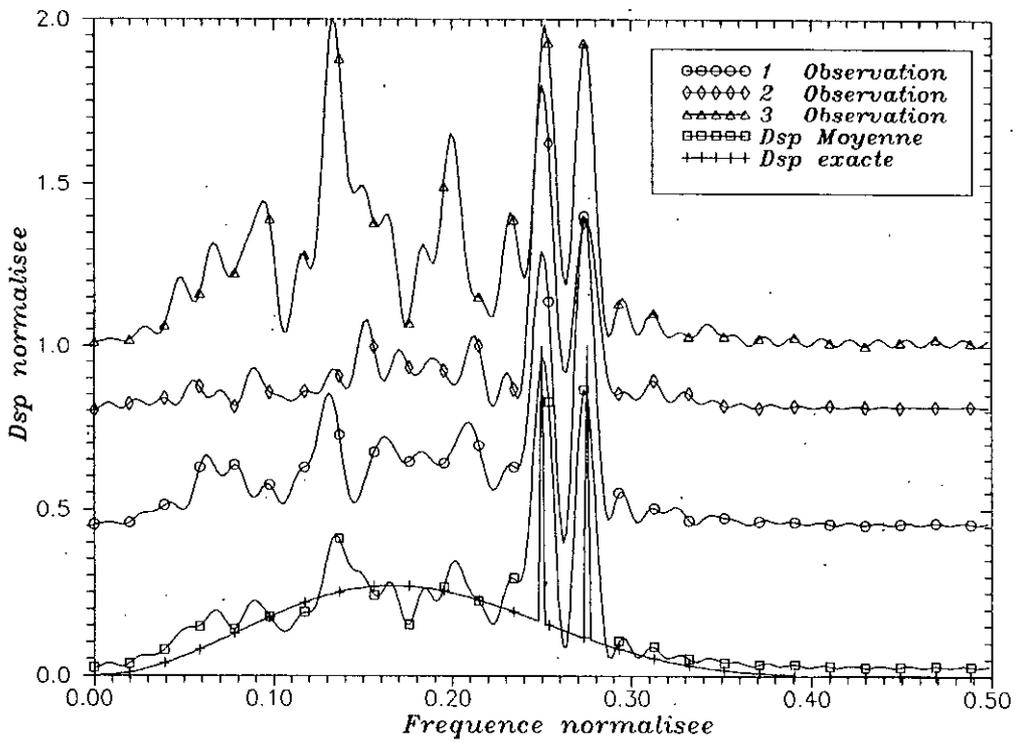


Figure 2.1  
Correlogramme de 3 observations differentes

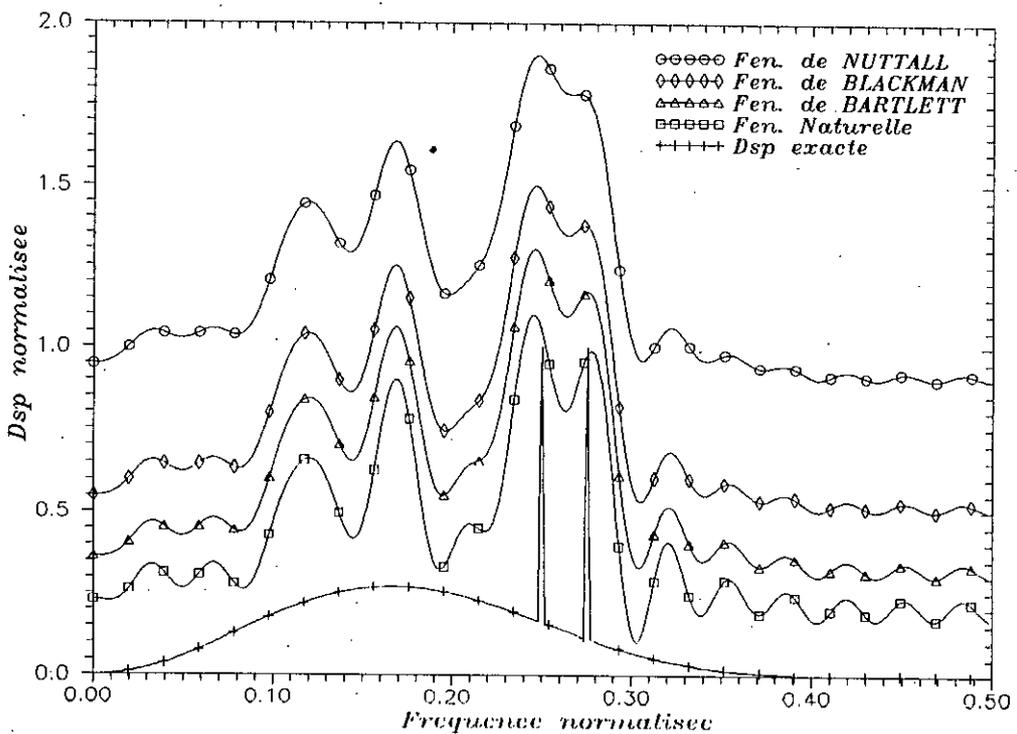


Figure 2.2  
Correlogramme avec differentes fenetres  
Naturelle, Triangulaire, de BLACKMAN, de NUTTALL

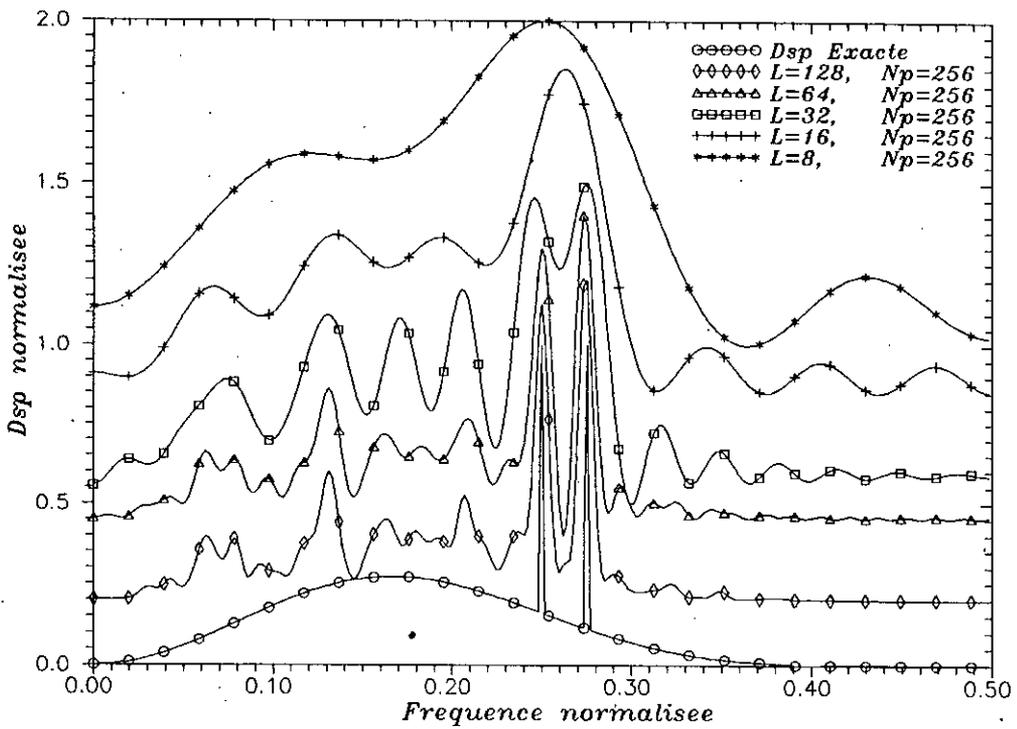


Figure 2.3

Correlogramme avec différentes valeurs de  $L$   
 ( $L$ =Nombre d'éléments d'autocorrelation)

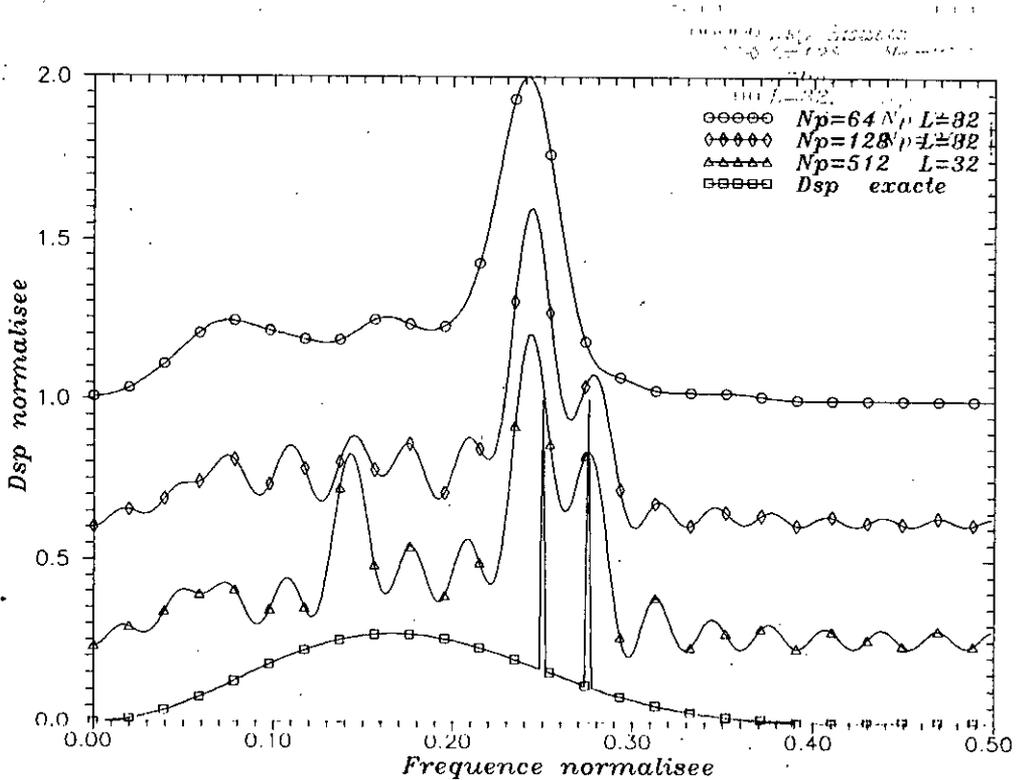


Figure 2.4

Correlogramme en fonction du nombre de points  $N$   
 Ponderation: Fenetre de Hamming

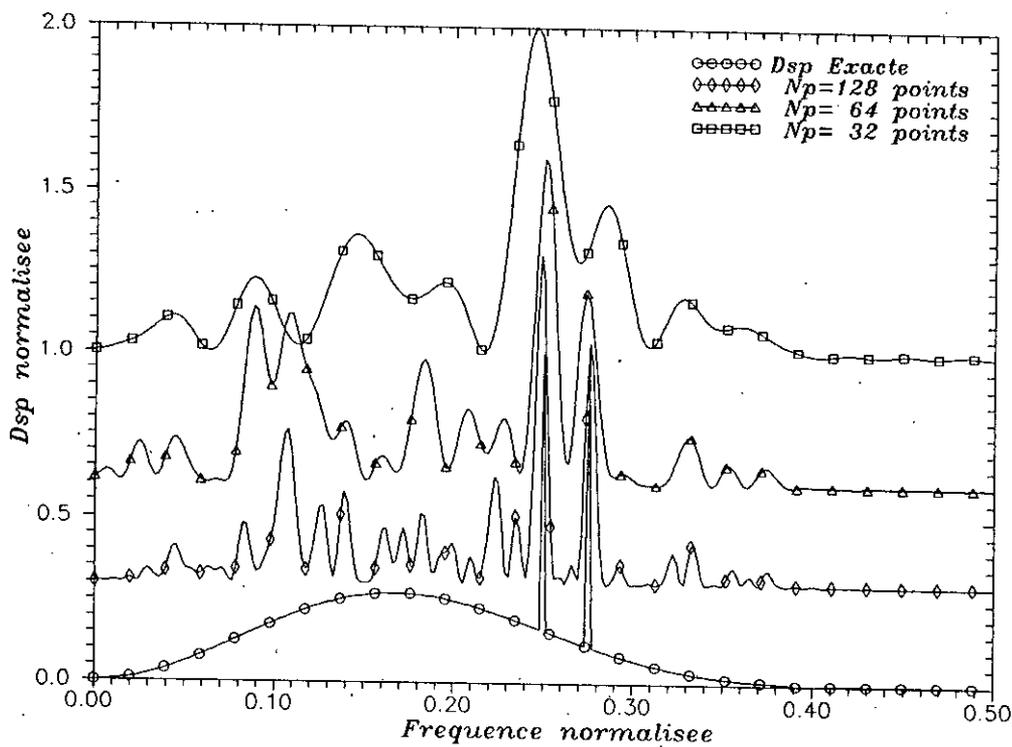


Figure 2.5  
Preriodogramme simple avec différentes valeurs de  $N_p$   
(  $N_p$ =Nombre de points Total )

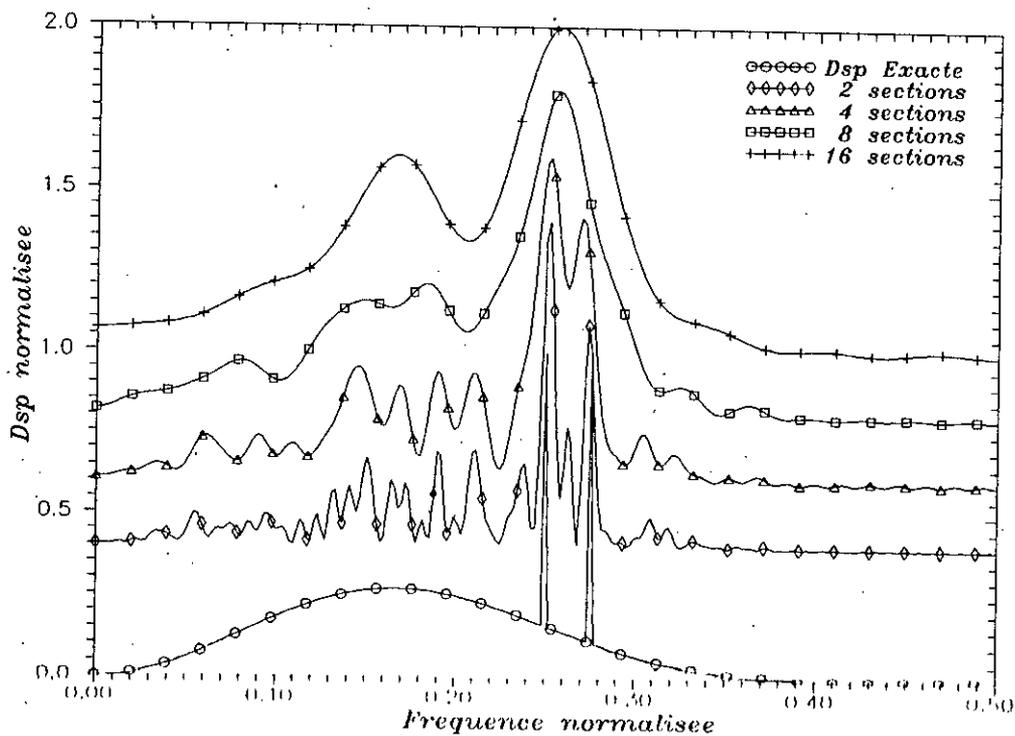


Figure 2.6  
Periodogramme moyenne en fonction du nombre  
de sections  $K$  (avec  $N_p=256$ )

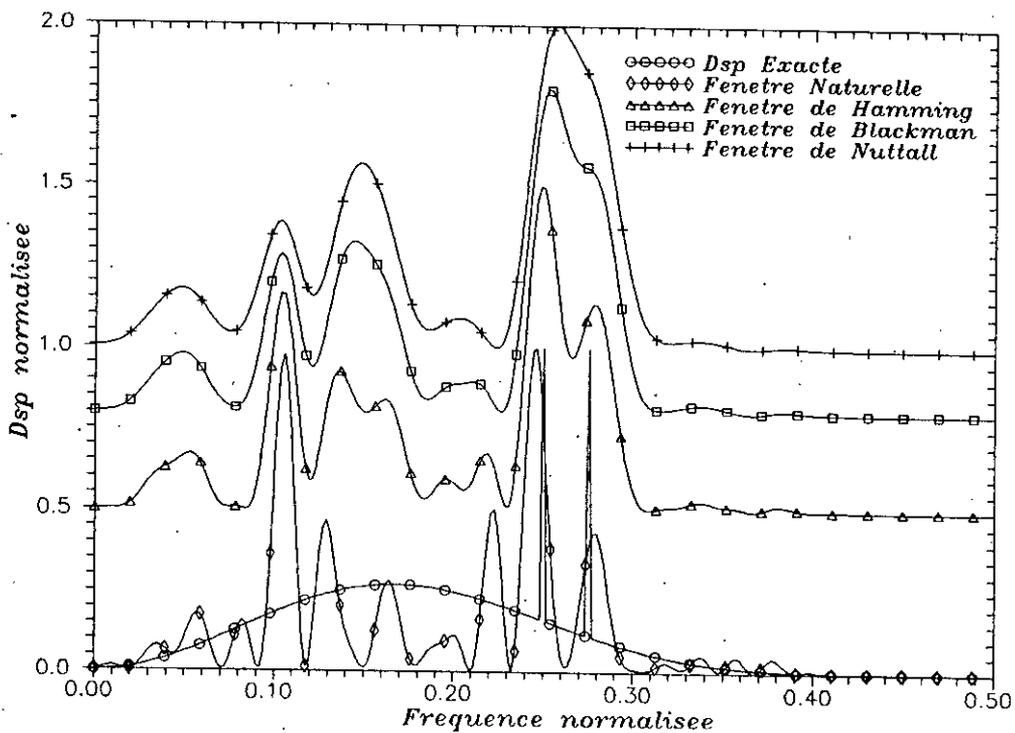


Figure 2.7  
*Periodogramme Adouci avec differentes fenetres*

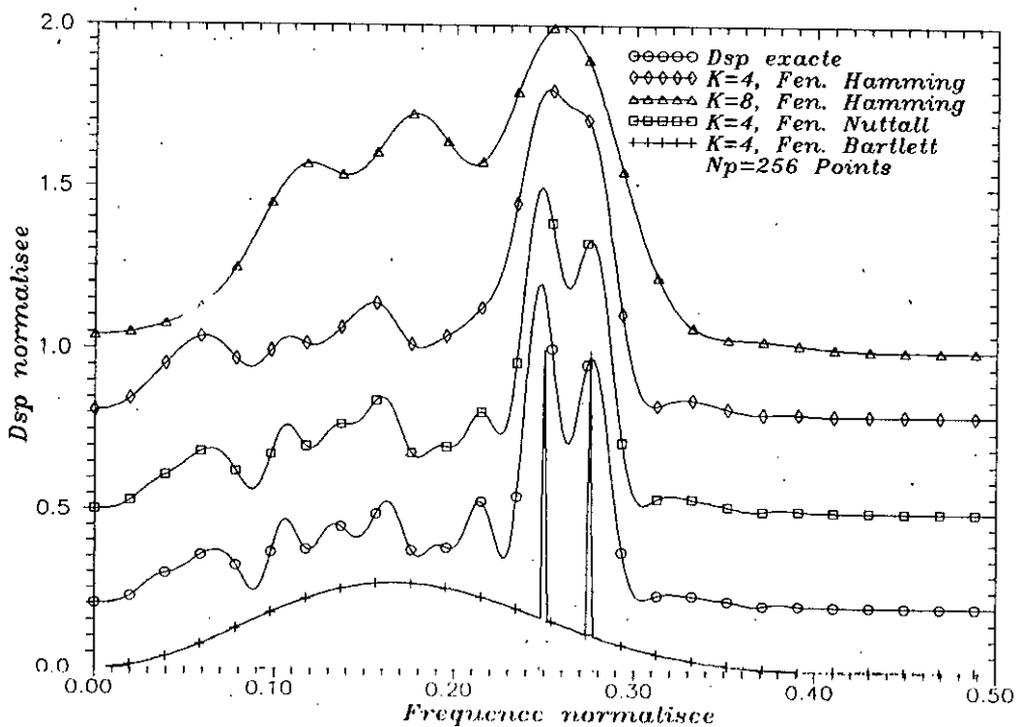


Figure 2.8  
*periodogramme modifie: Segmentation en 4, et 8 sections  
 Ponderation avec differentes fenetres (Hamming, Nuttall, Triangulaire)*

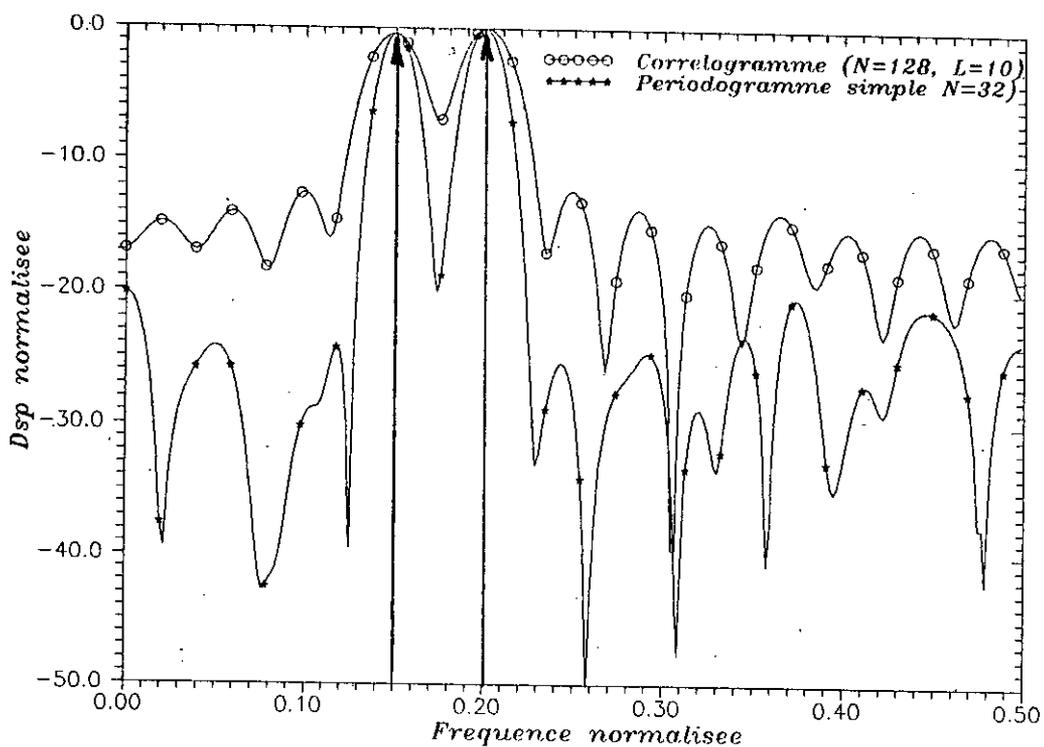


Figure 2.9  
 Comparaison entre les lobes secondaires du Correlogramme  
 et du periodogramme simple

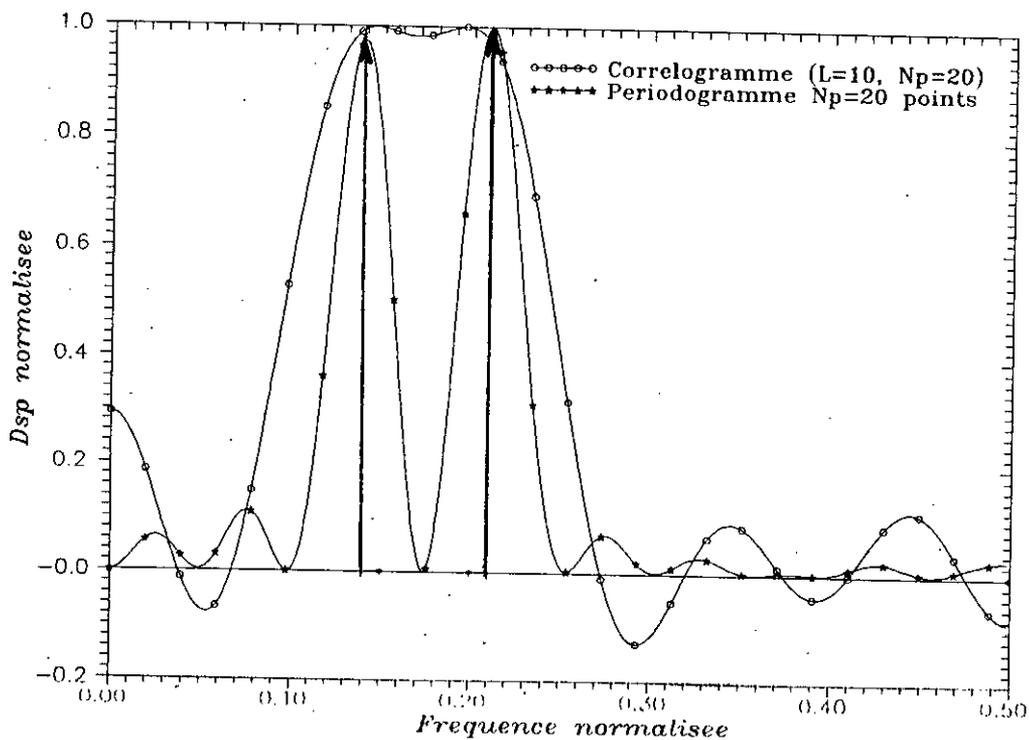


Figure 2.10  
 Lobes secondaires du periodogramme et du correlogramme  
 (apparition de lobes secondaires negatifs)

# **CHAPITRE III**

## **Modélisation paramétrique**

MODÉLISATION PARAMÉTRIQUE

3.1 Introduction

Les méthodes paramétriques d'analyse spectrale, bien qu'anciennes, ont connu un grand essor, surtout depuis quelques décennies. A cela deux raisons:

- Les limites de l'analyse par le classique périodogramme.
- L'évolution des moyens de calcul.

Ces techniques consistent à déterminer un modèle mathématique linéaire pour représenter un processus donné. Ainsi le problème de l'estimation de la DSP se ramène à un problème d'identification, qui conduit à une estimation des paramètres du modèle.

3.2 Représentation paramétrique

Si l'on considère un signal  $y(k)$ , comme étant la sortie d'un filtre générateur, de fonction de transfert  $H(z)$ , et dont l'entrée est un bruit blanc  $e(k)$ , centré de variance finie  $\sigma^2$ . Dans le cas général, on peut montrer que [18]:

$$S_y(z) = S_e(z) |H(z)|^2$$

avec  $S_y(z)$  la DSP du signal  $y(k)$ ,

et  $S_e(z)$  la DSP du signal  $e(k)$ .

Pour un bruit blanc de variance unité:  $S_e(z) = \sigma^2 = 1$ ,  
et avec  $Z = \text{Exp}(-j2\pi f)$  : évaluation sur le cercle unité,

on a 
$$S_y(f) = |H(f)|^2 \tag{3.1}$$

Cette relation est particulièrement importante dans la mesure où elle permet le calcul de la DSP à partir d'un nombre très réduit de paramètres (en général le nombre de paramètres inconnus figurant dans le modèle  $H(z)$  est inférieur au nombre de

données N). L'estimation de la DSP s'effectue selon le schéma de la figure 3.1.

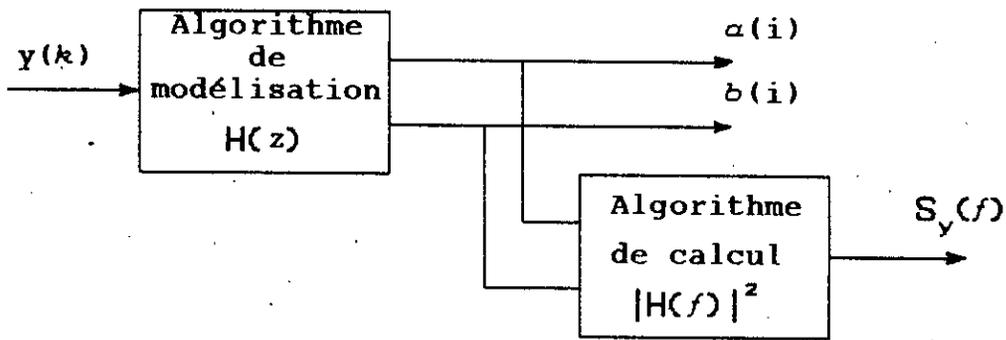


Figure 3.1 Principe de l'estimation de la DSP par les méthodes paramétriques.

### 3.3 Description des modèles

Le plus répandu des modèles paramétriques est le modèle autorégressif à moyenne ajustée (ARMA: Autoregressive Moving Average). Un tel modèle agit sur un signal d'entrée  $e(k)$  et le transforme en un signal de sortie  $y(k)$ . Le processus  $y(k)$  sera décrit par l'équation aux différences finies suivante:

$$y(k) + a_1 y(k-1) + \dots + a_n y(k-n) = b_0 e(k) + b_1 e(k-1) + \dots + b_m e(k-m) \quad (3.2)$$

Le modèle est dit ARMA d'ordre  $(n, m)$ , les coefficients  $a_i$  et  $b_i$  sont appelés paramètres du modèle.

En considérant les conditions initiales nulles, l'équation (3.2) devient après le calcul de la transformée en  $Z$  des deux membres:

$$Y(z) \cdot (1 + a_1 z^{-1} + \dots + a_n z^{-n}) = E(z) \cdot (b_0 + b_1 z^{-1} + \dots + b_m z^{-m})$$

soit

$$H(z) = \frac{Y(z)}{E(z)} = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_m z^{-m}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} \quad (3.3)$$

Deux cas particuliers conduisent à deux autres modèles:

- Si  $B(z) = b_0$ , la fonction de transfert  $H(z)$  ne contient que des pôles, le modèle est dit Autorégressif (AR), d'ordre  $n$ , (modèle tout pôles):

$$y(k) + \sum_{i=1}^n a_i y(k-i) = b_0 e(k) \quad (3.4)$$

-Si  $A(z)=1$ , le modèle est caractérisé par les zéros de  $H(z)$ , et le signal  $y(k)$  s'écrit comme une combinaison linéaire de l'entrée  $e(k)$ , le modèle est dit à moyenne adaptée d'ordre  $m$ ,

$$y(k) = \sum_{i=0}^m b_i e(k-i) \quad (3.5)$$

### 3.4 Modélisation Autorégressive

L'estimation de la DSP se fait selon le schéma de la figure 3.1, néanmoins, plusieurs questions peuvent être soulevées:

- 1-Quel est le type de modèle à utiliser, ARMA, AR, ou MA?
- 2-Comment et selon quel critère doit-on déterminer les coefficients du modèle particulier, qui convient le mieux aux données observées?
- 3-Quel est le nombre minimal de paramètres (ordre du modèle) qui pourra représenter convenablement (selon le critère choisi) le signal?

A la première question, on peut répondre que les modèles AR sont, de loin, les plus utilisés. La synthèse de signaux de parole, et l'étude des signaux sismiques, ne sont que deux exemples d'application, où de tels modèles s'avèrent performants et donnent des résultats satisfaisants [9], [40]. La difficulté avec les modèles ARMA est due à l'aspect non linéaire de la modélisation de la partie MA, car il faudrait dans ce cas déterminer à la fois les paramètres  $b_i$  et l'entrée  $e(k)$  qui est généralement inconnue [54].

Bien que les modèles AR sont insuffisants pour représenter parfaitement les systèmes étudiés, (les modèles ARMA sont plus appropriés) mais d'une façon générale, la complexité des modèles ARMA, ne justifie pas toujours la qualité que l'on peut attendre. De plus, on peut toujours substituer aux modèles ARMA et MA, un modèle AR équivalent dont l'ordre est un peu surestimé [47].

Pour estimer les coefficients du modèle AR, il est nécessaire de définir un critère de coût sur l'erreur commise qu'il faudra minimiser. Le critère usuel est la minimisation de la variance de l'erreur (de prédiction)  $e(k)$ :

$$\sigma_e^2 = E(e^2(k)). \quad (3.6)$$

L'intérêt de ce critère vient du fait qu'il est pourvu de signification physique: c'est une énergie, et qu'il conduit à un problème d'estimation qui est linéaire.

Quant à l'ordre du modèle, il s'agit d'une question particulièrement délicate. Plusieurs critères ont été proposés pour le déterminer, et seront étudiés dans le chapitre suivant.

### 3.5 Propriétés des modèles AR

#### 3.5.1 Prédiction linéaire

La modélisation AR consiste à représenter le signal  $y(k)$ , comme une combinaison linéaire des échantillons qui le précèdent, et de l'entrée actuelle  $e(k)$ :

$$y(k) = - \sum_{i=1}^P a_i y(k-i) + e(k) \quad ; b_0 = 1$$

Si on définit une prédiction de l'échantillon  $y(k)$  en fonction des  $p$  échantillons précédents:

$$\hat{y}^f(k) = - \sum_{i=1}^P a_i y(k-i) \quad (3.7)$$

alors  $e(k)$  apparaît comme une erreur sur le modèle, et elle s'écrit:

$$e^f(k) = y(k) - \hat{y}^f(k) = \sum_{i=0}^P a_i y(k-i) \quad , a_0 = 1. \quad (3.8)$$

Ainsi, quand l'excitation est inaccessible, l'estimation des paramètres du modèle sera basée exclusivement sur l'observation du signal. L'erreur  $e(k)$  peut être considérée comme source de bruit agissant sur le système.

### 3.5.2 Erreurs de prédiction: progressive et rétrograde

Il y a deux manières différentes pour obtenir l'erreur de prédiction, correspondant l'une à la prédiction progressive (directe), l'autre à la prédiction rétrograde (inverse) (voir fig. 3.2). L'erreur progressive de la prédiction directe (forward error) est donnée par

$$e^f(k) = y(k) - \hat{y}(k) = \sum_{i=0}^p a_i^f y(k-i) \quad ; a_0^f = 1 \quad (3.9)$$

c'est l'erreur de prédiction de l'échantillon  $y(k)$  à partir des  $p$  échantillons passés. L'erreur rétrograde (backward error) de la prédiction arrière, est l'erreur commise lors de l'estimation de l'échantillon  $y(k-p)$  à partir de l'ensemble des échantillons  $[y(k-p+1) \dots y(k-1), y(k)]$ . Bien que l'erreur rétrograde  $e^b$  concerne l'échantillon  $y(k-p)$ , mais elle est noté  $e^b(k)$ :

Avec

$$\hat{y}^b(k) = - \sum_{i=1}^p a_i^b y(k+i)$$

$$e^b(k) = y(k-p) - \hat{y}^b(k-p) \quad (3.10)$$

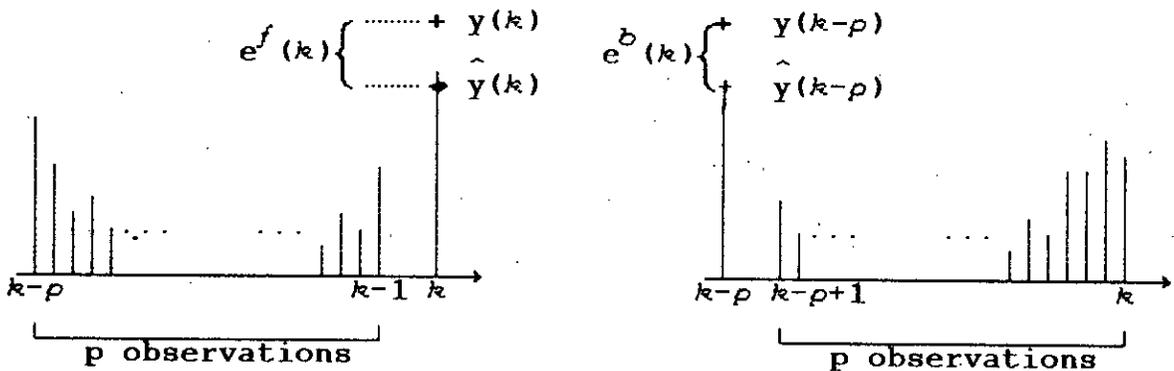


Figure 3.2: Prédiction directe et rétrograde d'ordre  $p$

### 3.5.3 Variances des erreurs de prédiction

La variance de l'erreur de prédiction directe d'ordre  $p$ , notée  $\alpha_p$ , vaut d'après (3.6) et (3.10):

$$\alpha_p = \sum_{i=0}^p \sum_{j=0}^p a_i^f a_j^f E(y(k-i)y(k-j))$$

Avec l'hypothèse de la stationnarité du signal  $y(k)$

$$E(y(k-i)y(k-j)) = r(i-j)$$

Alors

$$\alpha_p = \sum_{i=0}^p \sum_{j=0}^p a_i^f a_j^f r(i-j) \quad (3.11)$$

La minimisation de  $\alpha_p(k)$ , par rapport aux coefficients  $a_i^f$ , conduit au système :

$$\frac{\partial \alpha_p^f}{\partial a_i^f} = 2 \sum_{m=0}^p a_m^f r(i-m) = 0 \quad (3.12)$$

Après minimisation, la variance de l'erreur de prédiction est donnée par

$$(\alpha_p)_{\min} = \alpha_p^f = \sum_{i=0}^p a_i^f r(i) \quad (3.13)$$

Les deux relations (3.12) et (3.13) représentent un système de  $p$  équations linéaires à  $p$  inconnues. Système qu'on peut écrire sous la forme matricielle suivante<sup>(1)</sup>:

$$\begin{bmatrix} r(0) & r(1) & \dots & r(p) \\ r(1) & r(0) & \dots & r(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(p) & \dots & r(1) & r(0) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ a_1^f \\ \vdots \\ a_p^f \end{bmatrix} = \begin{bmatrix} \alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.14)$$

(1): Dans ces définitions, nous avons admis la stationnarité au second ordre du signal  $y(k)$ , et on a utilisé la propriété de symétrie de la fonction de corrélation:

$$r(-k) = r(k).$$

Soit sous la forme condensée:  
tel que

$$R_p \cdot A_p = E_p$$

$$R_p = \begin{bmatrix} r(0) & r(1) & \dots & r(p) \\ r(1) & r(0) & \dots & r(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(p) & \dots & r(1) & r(0) \end{bmatrix}, \quad A_p = \begin{bmatrix} 1 \\ a_1^f \\ \vdots \\ a_p^f \end{bmatrix}, \quad E_p = \begin{bmatrix} \alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.15)$$

Le système représenté par cette équation matricielle porte souvent le nom d'équations normales de Yule-Walker.

Un développement similaire peut être effectué pour la variance de l'erreur de prédiction arrière  $\beta_p$ , on arrive au résultat suivant:

$$\begin{bmatrix} r(0) & r(1) & \dots & r(p) \\ r(1) & r(0) & \dots & r(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(p) & \dots & r(1) & r(0) \end{bmatrix} \cdot \begin{bmatrix} a_p^b \\ \vdots \\ a_1^b \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \beta_p \end{bmatrix} \quad (3.16)$$

De la comparaison avec la relation (3.14), il vient immédiatement  $\alpha_p = \beta_p$ , c'est à dire que la variance de l'erreur de prédiction avant est égale à l'erreur de prédiction arrière. D'autre part, les coefficients de prédiction progressive /rétrograde sont liés par la relation suivante:

$$a_i^f = a_{p-i}^b \quad \dots \quad 1 \leq i \leq p$$

### 3.6 Solution des équations normales

On obtiendra les coefficients  $a_i$  en résolvant le système d'équations linéaire (3.14) ou (3.16). Ainsi le problème de modélisation est donc, essentiellement, un problème d'inversion matricielle. L'utilisation de la méthode générale d'élimination de Gauss nécessite un nombre de multiplications et d'additions de l'ordre de  $(p)^3$ , où  $p$  est l'ordre de la matrice. Certains algorithmes (Levinson, Cholesky), tiennent compte de la

structure particulière de la matrice de corrélation  $R_p$  (connue sous le nom de structure de Toeplitz) pour réduire le nombre d'opérations .

### 3.6.1 Algorithme de Levinson

Dans le cas où les éléments de chaque diagonal de la matrice de corrélation sont identiques; structure particulièrement intéressante, dite de Toeplitz, l'algorithme de Levinson permet de résoudre le système d'équations normales avec un coût de calcul de l'ordre de  $P^2$  [15]. Le secret de la récursion de Levinson réside dans le fait qu'une matrice de Toeplitz d'ordre  $m+1$ , contient comme sous matrices principales, toutes les matrices d'ordres inférieurs:

$$R_m = \begin{bmatrix} r(0) & r(1) & \dots & r(m) \\ r(1) & r(0) & \dots & r(m-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(m-1) & \dots & r(0) & r(1) \\ r(m) & \dots & r(1) & r(0) \end{bmatrix} = \begin{bmatrix} R_{m-1} & J \cdot \underline{r}_m \\ \underline{r}_m^* J & r(0) \end{bmatrix} \quad (3.17)$$

$$R_m = \begin{bmatrix} r(0) & r(1) & \dots & r(m) \\ r(1) & r(0) & \dots & r(m-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(m) & r(m-1) & \dots & r(0) \end{bmatrix} = \begin{bmatrix} r(0) & \underline{r}_m^T \\ \underline{r}_m^* & R_{m-1} \end{bmatrix} \quad (3.18)$$

Où  $J$  est une matrice de réflexion d'ordre  $m$  définie par

$$J = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}, \quad J^2 = J J^T = I \quad (3.19)$$

et nous avons adopté les notations suivantes:

$$\begin{aligned} \underline{r}_m &= (r(1), r(2), \dots, r(m))^T \\ \underline{a}_m &= (a_m(1), a_m(2), \dots, a_m(m))^T \\ A_m &= (1, a_m(1), a_m(2), \dots, a_m(m))^T \end{aligned} \quad (3.20)$$

Supposons que le vecteur  $A_{m-1}$  est connu, le problème est de déterminer le vecteur d'ordre  $m$ ,  $A_m$ . Si l'on admet que ce dernier peut être obtenu à partir de la combinaison linéaire suivante:

$$A_m = \begin{pmatrix} 1 \\ \underline{a}_m \end{pmatrix} = \begin{pmatrix} 1 \\ \underline{a}_{m-1} \\ 0 \end{pmatrix} + \rho_m \begin{pmatrix} 0 \\ J \underline{a}_{m-1} \\ 1 \end{pmatrix} \quad (3.21)$$

Où  $\rho_m$  est un coefficient de réflexion à déterminer:

1-Il est évident que  $\rho_m$  est exactement égale au coefficient  $a_m(m)$ .

2-Après multiplication des deux membres de l'équation (3.21), par  $R_m$  exploitée sous ses deux formes (3.17) et (3.18), et compte tenu de la relation (3.14), on obtient le résultat suivant:

$$\begin{pmatrix} \alpha_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_{m-1} \\ 0 \\ \vdots \\ \Delta_m \end{pmatrix} + \rho_m \begin{pmatrix} \Delta_m \\ 0 \\ \vdots \\ \alpha_{m-1} \end{pmatrix} \quad (3.22)$$

$\Delta_m$  est donné par

$$\Delta_m = \underline{r}_m^T J \begin{pmatrix} 1 \\ \underline{a}_{m-1} \end{pmatrix} = \sum_{k=0}^{m-1} \alpha_{m-1}(m-k) \cdot r(k) + r(m) \quad (3.23)$$

Alors

$$\rho_m = a_m(m) = -\Delta_m / \alpha_{m-1} \quad (3.24)$$

et la variance de l'erreur de prédiction peut être calculée de manière récursive [47]:

$$\alpha_m = \alpha_{m-1} (1 - |\rho_m|^2) \quad (3.25)$$

Donc, les coefficients  $a_m(i)$  du modèle peuvent être calculés pour les ordres successifs  $m=1,2, \dots, p$ , à partir des relations:

$$a_m(i) = \begin{cases} 1 & \text{si } i=0 \\ -\Delta_m/a_{m-1} & \text{si } i=m \\ a_{m-1}(i) + a_m^{(m)} \cdot a_{m-1}(m-i) & , 1 \leq i \leq m-1 \end{cases} \quad (3.26)$$

Il suffit seulement d'initialiser l'algorithme par  $a_0=r(0)$ . (voir Annexe A4).

### 3.6.2 Coefficients de réflexion

Dans l'algorithme de Levinson, les paramètres  $\rho_m$ , semblent jouer un rôle auxiliaire; en réalité, on verra dans le chapitre suivant, qu'ils ont un rôle essentiel. D'ailleurs, on peut représenter un modèle AR indifféremment soit par [47]:

- Une séquence de  $p+1$  éléments de corrélation
- L'ensemble des paramètres du modèle  $a(i) \ i=1 \dots p$  et  $a_p$
- L'ensemble des coefficients de réflexion  $\rho_m, \ m=1 \dots p$  et  $r(0)$ .

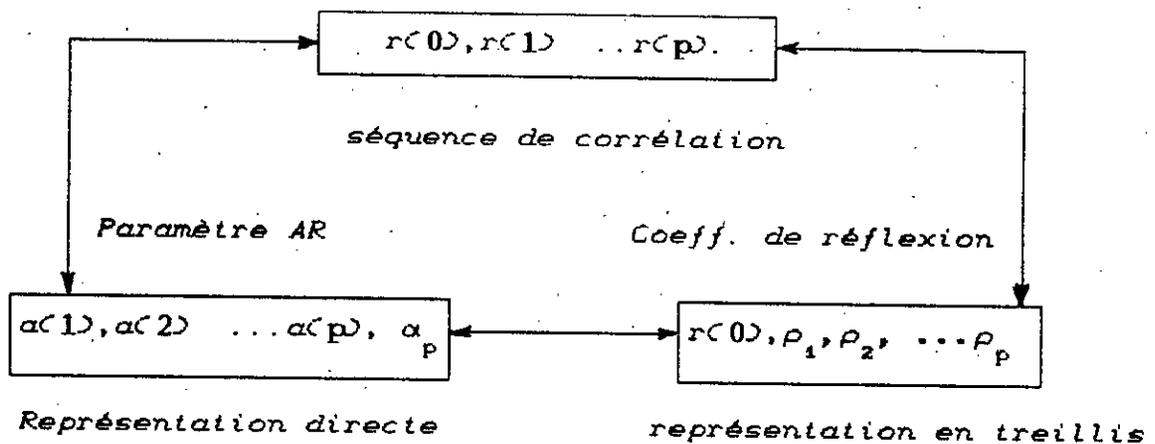


Figure 3.3: Les trois représentations des modèles AR.

La dénomination coefficient de réflexion est consacrée par l'usage dans la représentation des signaux sismiques [40] et la synthèse des signaux de la parole [9]. En effet, la trace

sismique peut être représentée par la sortie d'un système linéaire, excité par une ondelette, et ayant pour réponse impulsionnelle la série des coefficients de réflexion dûs aux discontinuités résultant du passage de l'onde sismique d'une couche vers une autre [54]. De même le conduit vocale peut être considéré comme une succession de tubes ou cavités acoustiques, de sections diverses [9]. Ainsi le signal de parole est fortement dépendant du couplage entre ces différentes cavités.

### 3.7 Structure du filtre d'estimation

La structure standard d'un filtre de prédiction AR(p), est décrite à l'aide des équations (3.9) et (3.10). Ceci conduit à la réalisation transversale (directe), illustrée dans la figure (3.4). Mais comme le filtre AR est complètement défini par les coefficients de réflexion  $\rho_m$ , on peut montrer que les erreurs de prédiction aux différents ordres peuvent être obtenus à l'aide des équations suivantes [47]:

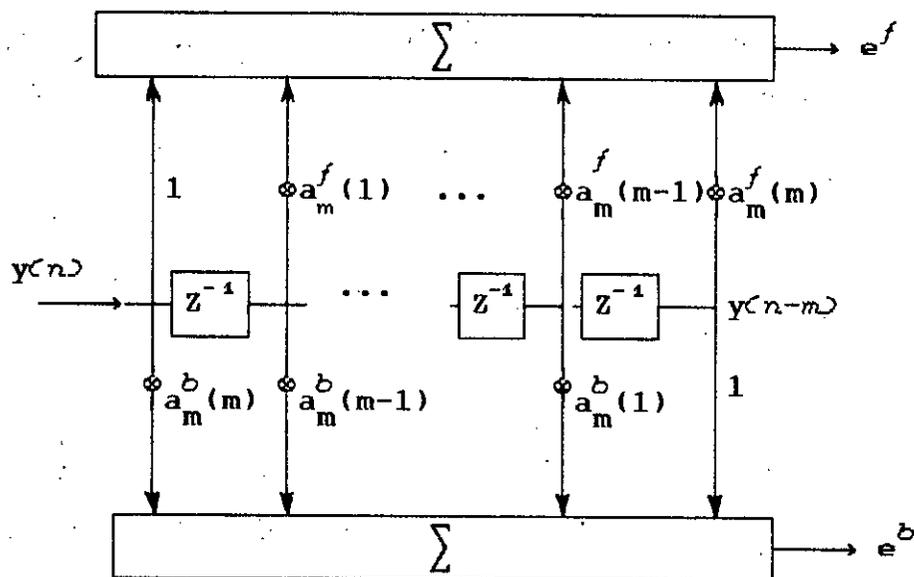


Figure 3.4: structure transversale du filtre de prédiction.

$$\begin{cases} e_m^f(k) = e_m^b(k) + \rho_m e_m^b(k-1) \\ e_m^b(k) = e_m^b(k-1) + \rho_m e_m^f(k) \end{cases} \quad (3.27)$$

avec  $k=1.. \dots p.. \dots N$

et  $e_0^f(k) = e_0^b(k) = y(k)$ .

Les paramètres  $\rho_m$  sont également appelés coefficients de corrélation partielle (PAR-CORR: Partial correlation).  $\rho_m$  est le coefficient de corrélation normalisé entre l'erreur directe  $e_m^f(k)$  et l'erreur rétrograde  $e_m^b(k-1)$  [47]:

$$\rho_m = \frac{E(e_m^f(n) e_m^b(n-1))}{\sqrt{E(|e_m^f(n)|^2) E(|e_m^b(n-1)|^2)}} \quad (3.28)$$

La récurrence (3.27) permet de définir une structure intéressante -appelée structure en treillis- constituée à l'aide de cellules en treillis disposées en cascade. Entre chaque <sup>deux</sup> cellules, il est nécessaire d'insérer un élément de retard élémentaire (illustration sur la fig. 3.5).

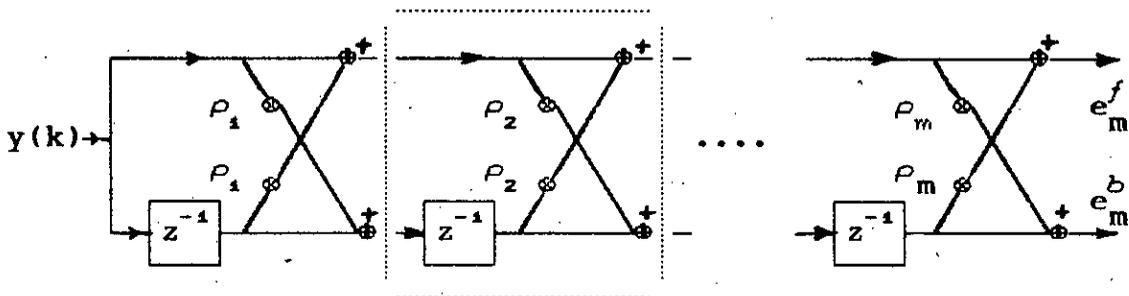


Figure 3.5 Structure en treillis du filtre de prédiction.

La structure en treillis -plus récente que la structure directe- a l'avantage d'être modulaire, car elle est faite d'une cascade d'étages identiques, et l'on montre qu'on

allongeant le treillis par l'ajout d'un dernier étage, les étages antérieures ne sont pas modifiés. Le modèle AR résultant est toujours stable, du fait que la valeur absolue de chacun des coefficients  $|\rho_i|$  du treillis ne dépasse pas l'unité. Enfin, cette structure possède de bonnes propriétés numériques ce que n'est pas le cas de la réalisation transversale qui utilise les coefficients de prédiction  $\alpha_i$ , et qui est beaucoup trop sensible aux arrondis de ces coefficients [22].

### 3.8 Conclusion

Nous avons vu dans ce chapitre, que les modèles AR ont dans beaucoup de cas une excellente justification physique, et sont plus faciles à utiliser. La résolution du problème de prédiction aboutit aux équations normales de Yule-Walker, et l'utilisation de l'algorithme de Levinson conduit à la structure en treillis. Dans le chapitre suivant, seront étudiées, les différentes méthodes AR d'estimation de la DSP. On montrera la supériorité de ces méthodes sur les méthodes classiques. Les limitations des algorithmes AR seront examinées en détail, notamment, le biais dans l'estimation des composantes sinusoidales, le dédoublement des raies spectrales, et le problème de sélection de l'ordre du modèle. Enfin on présentera d'autres méthodes modernes, en mettant le point sur leur relation avec les méthodes AR.

# **CHAPITRE IV**

## Estimation spectrale autoregressive

## CHAPITRE IV

### ESTIMATION SPECTRALE AUTORÉGRESSIVE

#### 4.1 Introduction

Le critère usuel pour l'estimation des paramètres du modèle AR reste la minimisation de la variance de l'erreur de prédiction  $e(k)$  définie en (3.6). Critère mathématiquement et physiquement justifié:

$$J(e) = E(e^2(k)) \quad (4.1.a)$$

Avec l'hypothèse de la stationnarité de l'erreur  $e(k)$ , on peut invoquer à son propos l'érgodicité et minimiser une somme des carrés de  $e(k)$ , d'où le critère des moindres carrés:

$$\alpha(e) = \sum_{k=k_1}^{k_2} e^2(k) \quad (4.1.b)$$

Selon l'intervalle de sommation  $K=[k_1..k_2]$  de cette expression, et selon qu'on s'intéresse à l'erreur directe  $e^f(k)$  ou rétrograde  $e^b(k)$  ou bien une combinaison des deux, on aboutira à des différentes méthodes semblables mais pas identiques.

#### 4.2 Méthodes d'estimation basées sur le critère des moindres carrés

Les observations disponibles sont  $y\{0..N-1\}$ . Une alternance se présente dans laquelle on pourra soit utiliser au mieux les données sans déborder hors de l'intervalle connu, soit déborder au delà en remplaçant par zéro les données manquantes. Les intervalles susceptibles d'être choisis pour calculer l'erreur de prédiction d'ordre  $p$  sont explicités dans l'équation matricielle suivante:

$$E_p = Y_p \begin{bmatrix} 1 \\ \underline{A}_p \end{bmatrix} \quad (4.2)$$

Avec  $E_p = (e(0), e(1), \dots, e(N+p-1))^T$

$$\underline{A}_p = (a(1), \dots, a(p))^T \quad \text{et} \quad A_p = (1, a(1), \dots, a(p))^T$$



$O_p$  est le vecteur formé de  $p$  zéros :  $O_p = (0, 0 \dots 0)^T$ .

L'élément  $r(i, j)$  de la matrice  $R_p = Y_p^T Y_p$  est le produit scalaire de deux colonnes de  $Y$ :

$$r(i, j) := \sum_{k=0}^{N+p} y(k-i)y(k-j) = r(i-j) = \sum_{k=0}^{N-(i-j)} y(k+(i-j))y(k) \quad (4.5)$$

Il est évident que ceci n'est autre qu'une estimation de la fonction d'autocorrélation (AC) à un facteur près  $1/N$ , en considérant l'estimateur biaisé. De ce fait la matrice  $R_p$  est une matrice de Toeplitz symétrique; comme dans le cas asymptotique (3.5.3).

Ainsi avec cette méthode, pour avoir une estimation de la DSP, on calcule la séquence  $r[0, \dots, p]$  à partir des données  $y[0 \dots N-1]$  à l'aide de l'équation (2.6). On résoud le système (4.4) en utilisant l'algorithme de Levinson, et finalement on peut se servir des coefficients du modèle  $a[1, \dots, p]$ , afin d'obtenir une estimation de la DSP sous la forme

$$S(f) = \frac{\alpha_p}{A_p(e^{-j2\pi f})A_p(e^{j2\pi f})} \quad (4.6)$$

$$\text{où } A_p(z) = 1 + \sum_{i=1}^p a(i)z^{-i}$$

#### 4.2.2 Méthodes du pré/Post-fenêtrage

Il s'agit dans ce cas de ne prendre en compte dans la sommation (4.1) que  $N$  échantillons du signal  $e(k)$ . Ceci revient à considérer que l'on tronque l'erreur de prédiction à travers une fenêtre de largeur  $N$ :

$$\text{Méthode de préfenêtrage : } \alpha_p = \sum_{k=0}^{N-1} e^2(k) \quad (4.7)$$

$$\text{Méthode de postfenêtrage : } \alpha_p = \sum_{k=p}^{N+p-1} e^2(k) \quad (4.8)$$

Pour mettre en évidence les équations normales qui interviennent dans ces deux cas, la matrice  $Y_p$  sera notée:

$$Y_p = \begin{pmatrix} L_p \\ Y_{1p} \\ U_p \end{pmatrix} \quad (4.9)$$

où la matrice triangulaire inférieure  $L_p$ , la matrice rectangulaire  $Y_{1p}$  et la matrice triangulaire supérieure  $U_p$  sont définies par:

$$L_p = \begin{pmatrix} Y_0 & 0 & \dots & 0 \\ Y_1 & Y_0 & \dots & 0 \\ \cdot & & & \\ Y_{p-1} & \dots & Y_0 & 0 \end{pmatrix} \quad U_p = \begin{pmatrix} 0 & Y_{N-1} & Y_{N-p} \\ & \cdot & \cdot \\ & & \cdot \\ 0 & \dots & Y_{N-1} \end{pmatrix}$$

$$Y_{1p} = \begin{pmatrix} Y_p & Y_{p-1} & \dots & Y_0 \\ \vdots & \vdots & & \\ \vdots & \vdots & & \\ Y_{N-1} & Y_{N-2} & \dots & Y_{N-p+1} \end{pmatrix}$$

Avec ces notations l'erreur de prédiction dans le cas du préfenêtrage est définie par l'équation suivante:

$$E_p = \begin{pmatrix} L_p \\ Y_{1p} \end{pmatrix} \begin{pmatrix} 1 \\ \underline{A}_p \end{pmatrix} \quad (4.10)$$

et après minimisation de la variance  $\alpha_p = E_p^T E_p$ , on obtient le système de Yule-Walker:

$$R_p \begin{pmatrix} 1 \\ \underline{A}_p \end{pmatrix} = \begin{pmatrix} \alpha_p \\ \underline{O}_p \end{pmatrix} \quad (4.11)$$

la matrice  $R_p$  est donnée par:

$$R_p = L_p^T L_p + Y_{1p}^T Y_{1p} \quad (4.12)$$

de même, pour la méthode du postfenêtrage on a:

$$E_p = \begin{pmatrix} Y_{1p} \\ U_p \end{pmatrix} \begin{pmatrix} 1 \\ A_p \end{pmatrix}, \quad R_p = U_p^T U_p + Y_{1p}^T Y_{1p} \quad (4.13)$$

La résolution des systèmes d'équations associés donne les coefficients du modèle AR, à partir desquels une estimation de la DSP peut être calculée en utilisant la formule (4.6).

A part l'utilisation de la méthode de préfenêtrage dans les techniques d'estimation récursives de la DSP, ces deux méthodes sont très peu utilisées. Ceci est certainement dû au fenêtrage implicite des données d'une part, et d'autre part, la matrice  $R_p$  définie par (4.11) ou (4.13) n'est pas de Toeplitz (mais elle est proche Toeplitz), l'algorithme de Levinson ne peut être utilisé pour la résolution des systèmes d'équations normales.

#### 4.2.3 Méthode de la covariance

Avec le choix  $K4$  [p...N-1], les coefficients du filtre AR produisant l'innovation  $e(k)$ , sont solution des équations de Yule-Walker suivantes:

$$Y_{1p}^T Y_{1p} \begin{pmatrix} 1 \\ A_p \end{pmatrix} = \begin{pmatrix} \alpha_p \\ 0_p \end{pmatrix} \quad (4.14)$$

$\alpha_p$  est la valeur minimale de la variance de l'erreur:

$$\alpha_p = \sum_{k=p}^{N-1} e^2(k) \quad (4.15)$$

La matrice  $R_p$  dans ce cas est donnée par  $R_p = Y_{1p}^T Y_{1p}$  et l'expression des éléments  $r(i, j)$  devient:

$$r(i, j) = \sum_{k=p}^{N-1} y(k-i)y(k-j) \quad (4.16)$$

où l'on observe que:  $r(i,j) \neq r(i-j)$ .

La matrice  $R_p$  est seulement symétrique, la résolution du système (4.14) peut se faire à l'aide de l'algorithme de Cholesky (Annexe A5), néanmoins Morf, Dickinson & Kailath [46] ont exploité le fait que  $R_p$  est un produit de deux matrices de Toeplitz, et ils ont développé un algorithme rapide et efficace pour la covariance autant que l'algorithme de Levinson pour l'autocorrélation.

Une variante de la méthode de covariance [47], consiste à pondérer l'erreur quadratique suivant l'expression:

$$\alpha_p = \sum_{k=p}^{N-1} w(k) \cdot e^2(k) \quad (4.17)$$

où  $w(k)$  représente l'énergie des données utilisées pour la prédiction

$$w(k) = \sum_{l=p+1}^{N-1} Y^2(l) \quad (4.18)$$

la résolution des équations de Yule-Walker associées à la méthode de covariance pondérée se fait à l'aide de méthodes de résolution générales de systèmes d'équations, étant donné qu'il n'y a pas -du moins à notre connaissance- d'algorithmes rapides pour cette méthode.

#### 4.2.4 Méthode de la covariance modifiée

Il faut noter que tout ce qu'a été développé dans les paragraphes précédents s'applique aussi bien à la minimisation de la puissance de l'erreur directe que rétrograde.

Pour la méthode de l'autocorrélation (AC), les puissances des erreurs de prédiction directe/rétrograde sont égales (voir 3.5.3), et les coefficients de réflexion estimés sont les mêmes. Cependant, avec la méthode de la covariance une différence apparaît comme on le montrera dans le paragraphe 4.3.

Les meilleures performances sont atteintes en combinant les deux approches, c'est-à-dire, en minimisant une combinaison linéaire de la puissance des erreurs directe et rétrograde.

Ainsi pour la méthode dite de la covariance modifiée, les paramètres du modèle AR d'ordre  $p$  s'obtiennent par minimisation de la moyenne des erreurs de prédiction directe et rétrograde par rapport aux coefficients  $a_i$ ,  $i=1, \dots, p$

$$\alpha_p = \frac{1}{2} \sum_{k=p}^{N-1} \left[ [e^f(k)]^2 + [e^b(k)]^2 \right] \quad (4.19)$$

$$\alpha_p = \frac{1}{2} \left( E_p^{fT} E_p^f + E_p^{bT} E_p^b \right) = \frac{1}{2} E_p^T E_p \quad (4.20)$$

où

$$E_p = \begin{pmatrix} E_p^f \\ E_p^b \end{pmatrix} = \begin{pmatrix} Y_{1P} \\ Y_{1P} J \end{pmatrix} \begin{pmatrix} 1 \\ A_p \end{pmatrix} \quad (4.21)$$

$J$  est la matrice coidentité définie par (3.19).

Les équations de Yule-Walker dans ce cas sont données par:

$$R_p \begin{pmatrix} 1 \\ A_p \end{pmatrix} = \begin{pmatrix} 2\alpha \\ 0_p \end{pmatrix} \quad (4.22)$$

$$\text{Avec} \quad R_p = Y_{1P}^T Y_{1P} + J Y_{1P}^T Y_{1P} J \quad (4.23)$$

Comme pour la covariance, la résolution du système (4.22) se fait à l'aide de l'algorithme de Cholesky, soit à l'aide d'un algorithme plus rapide (mais plus complexe) dû à Marple [47].

#### 4.2.5 Méthode de l'entropie maximum

L'entropie caractérise le degré du désordre d'un système [68], c'est une mesure de l'information -quoique incomplète- concernant un système donné. L'analyse spectrale basée sur le principe du maximum d'entropie consiste à choisir parmi l'infinité des processus compatibles avec les données, le processus le plus désordonné (aléatoire) possible en dehors de l'intervalle des d'observation.

Si l'on définit l'entropie  $\Delta H$  associée à la DSP  $S(f)$  par la fonction [10]:

$$\Delta H(S) = \int_{-1/2}^{1/2} \text{Ln}(S(f)) df \quad (4.24)$$

$S(f)$  est la DSP définie par la relation de Wiener-Kinchine (1.9).

Maximiser  $\Delta H(S)$ , sous la contrainte:

$$\int_{-1/2}^{1/2} S(f) e^{j2\pi k f} df = r(k) \quad k=0 \dots p \quad (4.25)$$

revient à résoudre l'équation:

$$\frac{\delta(\Delta H)}{\delta r(k)} = \int_{-1/2}^{1/2} \frac{1}{S(f)} e^{j2\pi k f} df = 0 \quad \text{pour } |k| > p \quad (4.26)$$

c'est-à-dire, les coefficients de la série de Fourier  $C_k$  de la fonction  $1/S(f)$  sont nuls pour  $|k| > p$ :

$$\frac{1}{S(f)} = \sum_{k=-p}^p C_k \cdot e^{j2\pi k f} \quad (4.27)$$

Comme  $S(f) \geq 0$ , alors l'expression (4.27) conduit à la DSP d'un filtre AR [66]:

$$S(f) = \left[ \frac{\sigma^2}{A(z) \cdot A(z^{-1})} \right]_{z=e^{j2\pi f}} \quad (4.28)$$

où  $A(z)$ ,  $\sigma^2$  sont donnés par la factorisation de Fejer [58]:

$$\sum_{k=-p}^p C_k \cdot z^{-k} = \frac{A_p(z) A_p(z^{-1})}{\sigma^2} \quad (4.29)$$

$$A_p(z) = 1 + \sum_{i=1}^p a(i) z^{-i}$$

Avec cette méthode, dite aussi méthode de Burg [10],[11], l'estimation de la fonction d'AC n'est plus nécessaire.

Les coefficients du modèle AR s'obtiennent à partir de la récursion de Levinson (équation 3.21). Le coefficient de réflexion  $\rho_m$  est estimé en minimisant le critère

$$\alpha_m = \sum_{k=m}^{N-1} \left( |e^f(k)|^2 + |e^b(k)|^2 \right) \quad (4.30)$$

par rapport au dernier coefficient  $a_m(m)$ . Dans ce cas les erreurs de prédiction directe et rétrograde vérifient la récurrence (3.27).

#### 4.3 Estimation des coefficients de corrélation partielle

Une approche efficace pour résoudre les équations normales consiste à utiliser la structure en treillis, qui repose essentiellement sur l'estimation des coefficients de réflexion. Ces derniers peuvent être caractérisés d'une manière statistique: ils représentent l'intercorrélacion normalisée des erreurs de prédiction directe et rétrograde [47].

A la prédiction d'ordre  $m+1$ , le coefficient  $\rho_{m+1}$  peut être estimé de plusieurs manières. Avec les notations suivantes

$$\alpha_m(n) = \mathbb{E} \left[ |e_m^f(n)|^2 \right] \quad \beta_m(n) = \mathbb{E} \left[ |e_m^b(n)|^2 \right]$$

$$\gamma_m(n) = \mathbb{E} \left[ e_m^f(n) e_m^b(n-1) \right]$$

les estimateurs du tableau (4.1) conduisent à des versions différentes d'algorithmes en treillis [43]. Les quatre derniers estimateurs, du N°3 au N°6, ont l'agréable propriété de garantir la condition de stabilité  $|\rho_{m+1}| < 1$ . En plus l'estimateur N°5 (moyenne harmonique), résulte de la minimisation d'un critère qu'est la somme  $[\alpha(m) + \beta(m)]$  [10], ce qui prouve que la méthode de Burg est une forme particulière des méthodes en treillis.

tableau 4.1: Estimateurs des coefficients de réflexion

N°	Méthode	Définition statistique	Observation
1	Directe	$\rho_{m+1}^f = -\gamma_m(n) / \beta_m(n-1)$	si $ \rho_{m+1}^f  > 1$
2	Inverse	$\rho_{m+1}^b = -\gamma_m(n) / \alpha_m(n)$	Alors $ \rho_{m+1}^b  < 1$
3	Coéff. Min.	$\rho_{m+1}^M = S \cdot \min \left[  \rho_{m+1}^f ,  \rho_{m+1}^b  \right]$ $S = \text{signe}(k_{m+1}^f) = \text{signe}(k_{m+1}^b)$	si $ \rho_{m+1}^b  > 1$ Alors $ \rho_{m+1}^f  < 1$
4	Itakura	$\rho_{m+1}^I = \frac{-\gamma_m(n)}{\sqrt{\alpha_m(n)\beta_m(n-1)}}$ $\rho_{m+1}^I = S \sqrt{\rho_{m+1}^f \rho_{m+1}^b}$	Moyenne géométrique
5	Burg	$\rho_{m+1}^H = \frac{-2\gamma_m(n)}{\alpha_m(n) + \beta_m(n-1)}$ $\rho_{m+1}^H = \frac{2 \rho_{m+1}^f \rho_{m+1}^b}{\rho_{m+1}^f + \rho_{m+1}^b}$	Moyenne harmonique
6	Moyenne générale d'ordre r r ≤ 0	$\rho_{m+1}^r = S \left[ \frac{1}{2} ( \rho_{m+1}^f ^r +  \rho_{m+1}^b ^r) \right]^{1/r}$	si r → 0 $\rho_{m+1}^r \rightarrow \rho_{m+1}^I$ si r → ∞ $\rho_{m+1}^r \rightarrow \rho_{m+1}^M$
Remarque:		$ \rho_{m+1}^M  \leq  \rho_{m+1}^H  \leq  \rho_{m+1}^I $	

#### 4.4 Estimation AR de la fonction d'autocorrélation

Une relation directe entre les coefficients  $a_i$  du modèle AR d'ordre p et les éléments de la fonction d'AC  $r(k)$ , peut être obtenue. En effet, en multipliant les deux membres de l'équation (4.31 a) par  $y(k-m)$ , et en prenant l'espérance mathématique, on arrive au résultat (4.31 b)

$$y(k) = - \sum_{i=1}^p a(i)y(k-i) \quad (4.31 a)$$

$$r(0) = \sigma^2 - \sum_{i=1}^p a(i)r(i)$$

$$r(m) = - \sum_{i=1}^p a(i)r(m-i) \quad (4.31 b)$$

Il est important de bien noter que pour  $k \geq p$ , la suite  $r(k)$  est engendrée par une relation de récurrence faisant intervenir les  $p$  termes précédents. Ce n'est pas le cas des méthodes classiques étudiées au Chapitre II, où il n'y a aucune extrapolation de la fonction d'AC; les valeurs non disponibles sont considérées simplement comme nulles.

Un autre point à retenir: si la matrice de corrélation  $R_p$  est une matrice de Toeplitz, la décomposition triangulaire dite de Cholesky de  $R_p^{-1}$  est donnée par:

$$R_p^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ a_m(1) & 1 & 0 & & \\ a_m(2) & a_m(1) & 1 & & \\ \cdot & & & 0 & \\ a_m(m) & a_m(m-1) & \dots & 1 & \end{pmatrix} \begin{pmatrix} 1/\alpha_p & 0 & 0 \\ 0 & 1/\alpha_{p-1} & \\ \cdot & \cdot & \\ \cdot & \cdot & \\ 0 & 0 & 1/\alpha_0 \end{pmatrix} \begin{pmatrix} 1 & a_m(1) & a_m(m) \\ 0 & 1 & a_m(m-1) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

$$R_p^{-1} = A_p^T P_p^{-1} A_p \quad (4.32)$$

Ainsi avec les méthodes en treillis, l'algorithme de résolution n'a pas à manipuler directement l'inverse de la matrice de corrélation  $R_p^{-1}$ , mais seulement à estimer à chaque étape  $m$ , la variance de l'erreur de prédiction  $\alpha_m$ , ainsi que les coefficients du modèle.

## 4.5 Simulations et résultats

Après avoir décrit les différents algorithmes de modélisation, nous allons étudier leurs comportements sur des signaux synthétiques particuliers. Ceci afin d'évaluer leurs performances en comparant le modèle exact avec le modèle estimé.

Les signaux analysés sont de deux types:

1-Signaux obtenus par passage d'un bruit blanc à travers un modèle paramétrique (AR,ARMA). Ces signaux sont éventuellement bruités. Le signal considéré est généré par le processus AR d'ordre 4 caractérisé par l'équation de récurrence suivante:

$$x(k) = \sum_{i=1}^4 a_i \cdot x(k-i) + e(k), \quad y(k) = x(k) + b(k).$$

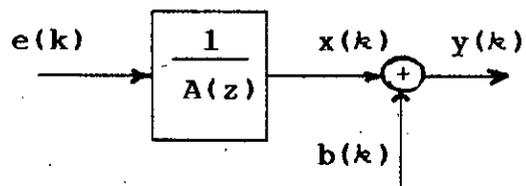
Avec

$$a_1 = 2.7607$$

$$a_2 = -3.8106$$

$$a_3 = 2.6535$$

$$a_4 = -0.9238$$



2-Signaux pseudo-déterministes formés par plusieurs sinusoides noyées dans du bruit blanc.

$$y(k) = \sum_i A_i \cdot \cos(2 \cdot \pi \cdot f_i \cdot k + \phi_i) + b(k)$$

Les paramètres d'estimation sont indiqués en légendes.

Comme le spectre d'un signal réel est symétrique, seule la bande  $0, \dots, f_0/2$  est représentée, où  $f_0$  désignant la fréquence d'échantillonnage.

### 4.5.1 Variance

Notre but n'est pas de comparer les différents algorithmes sur le plan statistique, mais d'étudier leur comportement sur un petit nombre d'épreuves. Puisqu'en pratique, il est

souvent demandé de conclure à partir d'un très petit nombre de réalisations. Ainsi nous avons effectué pour le modèle AR(4) un traitement statistique portant sur 10 observations obtenues en utilisant plusieurs réalisations du signal d'excitation (bruit blanc).

Il ressort de l'examen des résultats présentés aux figures 4.1-4.4 que du point de vue stabilité statistique, la méthode d'autocorrélation est la plus mauvaise. La méthode est moins performante dans le cas d'une estimation non biaisée de la fonction de corrélation. Ceci est attribué au fait que la condition: coefficients de réflexion  $|\rho_i| < 1$  n'est pas toujours satisfaite. Les autres méthodes (Burg, Covariance ... etc) ont une variance relativement faible et acceptable.

Il est à remarquer que pour l'ensemble des méthodes AR, l'instabilité de l'estimation se manifeste clairement au voisinage des pics. La situation devient critique dans le cas d'estimation de la DSP d'un modèle ARMA caractérisé par des vallées très prononcées (zéros proches du cercle unité), ce qui forme des zones spectrales difficiles à estimer par les méthodes Autorégressives (Figure 4.5).

#### 4.5.2 Résolution et Biais

Détecter des périodicités cachées dans un signal stochastique est une opération difficile, souvent compliquée, et pourtant de plus en plus indispensable. En effet, il faut arriver à détecter toutes les composantes fréquentielles (bonne résolution), et à les identifier exactement (sans biais).

En ce qui concerne l'aptitude à séparer deux sinusoides très proches et noyées dans du bruit, les estimateurs AR sont beaucoup plus performants que les estimateurs classiques basés sur la transformée de FOURIER [10], [52]. Néanmoins, la résolution des méthodes AR dépend à la fois du nombre de points utilisé, de l'ordre du modèle, ainsi que du niveau signal/bruit (RSB). Pour illustrer ces points, on considère le signal test habituel de plusieurs sinusoides bruitées. soit le signal composé de trois sinusoides et d'un bruit blanc Gaussien  $b(k)$ ,

le rapport signal/bruit est de 20 dB.

$$y_2(k) := \sum_{i=1}^3 A_i \cdot \cos(2 \cdot \pi \cdot f_i \cdot k + \phi_i) + b(k)$$

$f_1 = 0.15$	$A_1 = 10$	$\phi_1 = 0$
$f_2 = 0.19$	$A_2 = 05$	$\phi_2 = \pi / 4$
$f_3 = 0.23$	$A_3 = 10$	$\phi_3 = 0$

On identifie un modèle AR(6) à ce signal (3 sinusoides\*2). Les graphes des figures 4.6 et 4.7, montrent qu'il est nécessaire de surestimer cet ordre afin de détecter toutes les composantes fréquentielles. En effet, un ordre  $p=6$  ne nous permet pas de détecter la fréquence 0.19, même en utilisant 200 échantillons du signal. Tandis qu'avec un ordre  $p=12$ , on obtient avec, la méthode de Burg, trois pics distincts sur un bloc de 40 points seulement. Mais avec la méthode de corrélation, il faut tripler l'ordre ( $p=18$ ) et utiliser un nombre de points plus élevé ( $N_p=200$ ) pour arriver au même résultat. D'après les graphes des figures 4.8 et 4.9, on constate que, si le RSB est de 40 dB, les trois fréquences sont détectées par toutes les méthodes, toutefois la finesse des pics est meilleure dans le cas de la méthode de Burg que celle de la méthode de corrélation. Mais quand le rapport RSB décroît de 40 db à 10db, le nombre de composantes fréquentielles détectées décroît de trois à deux.

Il est clair que la présence de biais ou de variance fréquentielle est très gênante lors de la détection des périodicités cachées. C'est pourquoi avoir un estimateur sans biais est une qualité très recherchée. Malheureusement, la méthode de Burg, qui est avantageuse du point de vue variance et résolution, présente deux faiblesses pour les signaux à fréquences pures:

1-Le biais est fonction de la phase initiale du signal. Ceci est illustré dans l'exemple des figures 4.10, 4.11, 4.12, 4.15.a et 4.16.a, où on observe une estimation par les différentes méthodes (MEM, corrélation et covariance) de la DSP d'une sinusoïde pure d'amplitude unité et de fréquence 1hz,

échantillonnée à une cadence de 0.05s. On constate un biais considérable pour la MEM (décalage de l'ordre de 20% par rapport à la fréquence exacte) dans le cas où la phase initiale est un multiple impaire de  $45^\circ$ , et le nombre de points utilisés dans l'estimation correspond à un nombre impaire de quarts de périodes. Une étude plus détaillée du problème est présentée dans [14].

2-Dédoublage des raies, où l'occurrence de deux pics, parfois plus, lorsque seulement un pic doit être présent. D'après Fugore et Al [47], ce phénomène se présente lorsque:

- Le rapport RSB est élevé
- La phase initiale est un nombre impaire de  $45^\circ$
- L'ordre du modèle AR est considérable devant le nombre total d'échantillons utilisés pour l'estimation.

Les figures 4.13, 4.14, 4.15.b et 4.16.b donnent l'estimation (par les différentes méthodes) de la DSP d'une sinusoïde de fréquence 1.5 Hz, avec  $45^\circ$  comme phase initiale. La fréquence d'échantillonnage est de 100 Hz. le rapport RSB est de 61 dB.

Le tableau suivant résume les performances des différentes méthodes. La meilleure occupe la première place, tandis que la moins bonne est classée N°4.

Tableau 4.1: Performances des différentes méthodes

Méthode	Résolution	Variance	Biais	Efficacité Numérique
BURG	1	1	4	1
Covariance	2	3	1	3
Cov. Modifiée	3	2	1	4
Corrélation	4	4	1	2

Enfin de compte on peut dire que les méthodes AR peuvent s'accorder à des signaux relativement brefs, et offrent une capacité de détection supérieure. Les méthodes utilisant la structure en treillis ont en générale une bonne stabilité statistique.

#### 4.6 Sélection de l'ordre du modèle AR

Comment déterminer l'ordre optimal  $M$  ? C'est un problème qui se pose lors de toute modélisation. Tout d'abord nous allons examiner les questions suivantes:

1-Que se passe-t-il lorsqu'on surestime l'ordre ( $p > M$ ) ?

2-Que se passe-t-il lorsqu'on sousestime l'ordre ( $p < M$ ) ?

Les courbes de la figure (4.17) montrent que pour un ordre inférieur à l'ordre exact ( $p < M$ ) l'estimation reproduit l'enveloppe de la DSP réelle mais la résolution est très mauvaise. Par contre, si l'ordre est surestimé ( $p > M$ ), les principales composantes fréquentielles se retrouvent dans la DSP estimée, mais on voit apparaître des pics parasites arbitraires qui n'ont rien à avoir avec la DSP exacte.

Ainsi le choix du nombre optimum de paramètres du modèle est l'un des points les plus importants en analyse spectrale par les méthodes paramétriques.

##### 4.6.1 Critères de sélection

L'algorithme de Levinson possède la propriété de fournir la variance de l'erreur à chaque étape de la récursion. Ce qui incite par intuition à choisir l'ordre  $p$  qui minimise la puissance de l'erreur (fig 4.18). Cependant on remarque que cette dernière décroît de façon monotone avec l'augmentation de l'ordre  $p$  du modèle: cette méthode n'est pas toujours exploitable. Ainsi, plusieurs critères ont été introduits pour la sélection de l'ordre optimal du modèle, les critères les plus importants sont:  $FPE^{(1)}$  [2],  $AIC^{(2)}$  [3],  $MDL^{(3)}$  [15], et le  $CAT^{(4)}$  [56].

---

(1)  $FPE$  = Final prediction error

(2)  $AIC$  = Akaike Information Criterion

(3)  $MDL$  = Minimum description length

(4)  $CAT$  = Criterion Autoregressive Transfer function.

Le principe de ces critères est de chercher le modèle le plus vraisemblable, ce qui conduirait à un ordre infini, mais d'ajouter un coût croissant pour limiter le nombre de paramètres. La plupart de ces critères peuvent être exprimés sous la forme [26]

$$C(p) = -2L(p) + pQ(p) \quad (4.33)$$

$C(p)$  est le critère en fonction de l'ordre  $p$

$L(p)$  est la vraisemblance du modèle estimé d'ordre  $p$

$Q(p)$  est un facteur de pondération dépendant du nombre d'échantillons.

Dans le cas gaussien, si on note par  $e_p$  la puissance de l'erreur de prédiction, le critère devient:

$$C(p) = N \text{Log}(e_p) + pQ(p) \quad (4.34)$$

où  $N$  est le nombre de données.

tableau 4.2 Critères de sélection de l'ordre du modèle AR

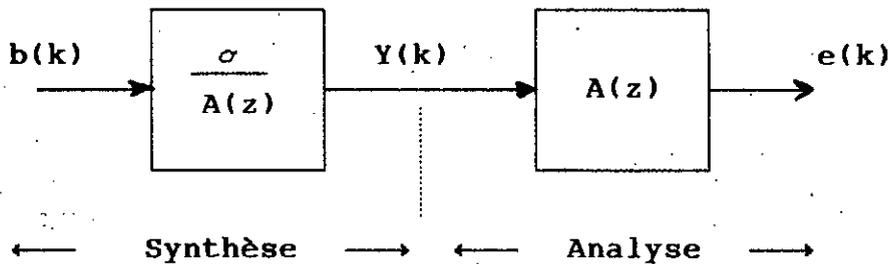
Critère	Auteur
$FPE(p) = \left[ \frac{N + p + 1}{N - p - 1} \right] e_p$	AKAIKE [2], 1969
$AIC(p) = N \cdot \text{Ln}(e_p) + 2p$	AKAIKE [3], 1974
$HAIC(p) = N \cdot \text{Ln}(e_p) + 2C \cdot \text{Ln}(\ln(n))$	HANNAN [15], 1979
$MDL(p) = N \cdot \text{Ln}(e_p) + p \cdot \text{Ln}(N)$	RISSANEN [47] 83
$CAT(p) = \left[ \frac{1}{N} \sum_{j=1}^p \underline{e}_j^{-1} \right] - \underline{e}_p^{-1}$ $\underline{e}_i = \left[ \frac{N}{N - i} \right] e_i$	PARZEN [56] 1974

## 4.6.2 Un nouveau critère de sélection de l'ordre du modèle AR

### 4.6.2.1 Analyse et synthèse

Comme on l'a déjà vu au paragraphe 3.5, il existe deux approches nous permettant de résoudre le problème de la modélisation autorégressive:

- L'approche synthèse où le signal  $y(k)$  est considéré comme étant la sortie d'un filtre AR linéaire alimenté par un bruit blanc  $b(k)$  de moyenne nulle et de variance  $\sigma^2$ .
- L'approche analyse où on considère le filtre inverse, qui, excité par le signal  $y(k)$ , engendre à sa sortie l'erreur de prédiction  $e(k)$ .



Analyse et synthèse

Dans le cas où le signal a été effectivement produit par un système AR d'ordre  $p$ , l'erreur  $e(k)$  reproduit l'excitation  $b(k)$  à un facteur près. Comme cela est illustré par le schéma synoptique ci-dessus. Il est bon de rappeler que les deux approches (Analyse/ Synthèse) conduisent aux mêmes ensembles d'équations, respectivement équations normales et équations de Yule-Walker. L'estimation de la DSP se ramène à la résolution de ces équations donc au calcul des paramètres  $(a_1, a_2, \dots, a_p)$ .

### 4.6.2.2 Formulation du critère

Supposons que les données sont engendrées par un modèle AR d'ordre  $p$ , et soit  $a_1, a_2, \dots, a_p$  les coefficients de ce modèle.

Du point de vue synthèse, on a :

$$Y(z) = \sigma \frac{B(z)}{A_p(z)} \quad (4.35)$$

où  $A_p(z) = 1 + \sum_{i=1}^p a_i z^{-i}$  et  $B(z) = 1$

du point de vue Analyse, l'erreur de prédiction de l'estimateur d'ordre  $m$  est donnée par :

$$E(z) = Y(z) \hat{A}_m(z) \Rightarrow Y(z) = \frac{E(z)}{\hat{A}_m(z)} \quad (4.36)$$

où  $\hat{A}_m(z) = 1 + \sum_{i=1}^m \hat{a}_i z^{-i}$

et  $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_m$  les  $m$  coefficients du modèle estimé d'ordre  $m$ .  
Comme les deux approches sont identiques alors :

$$Y(z) = \frac{E(z)}{\hat{A}_m(z)} = \sigma \frac{B(z)}{A_p(z)} \quad (4.36)$$

$$E(z) = \sigma \frac{\hat{A}_m(z)}{A_p(z)} B(z).$$

Ainsi, l'erreur de prédiction est donnée par un modèle ARMA(p,m). Si l'estimateur des coefficients  $a_i$  est consistant, alors à  $m=p$   $\hat{A}_m(z) = A_p(z)$  et on aura  $E(z) = 1 = B(z)$ , et l'erreur reproduit exactement l'excitation  $b(k)$ . A ce stade deux constatations nous ont conduit à formuler un nouveau critère :

- 1-Il est bien connu que la fonction d'autocorrélation (AC) d'un bruit blanc discret  $b(k)$  est une impulsion de Kronecker  $\delta(k)$ .
- 2-Toutes les méthodes d'estimation s'accordent à prendre comme valeur initiale de l'erreur, le signal lui même  $e_0(k) = y(k)$ . Alors l'AC de l'erreur  $R_{e_0}(k) = R_y(k)$ . Dans le cas général  $R_y(k) \neq \delta(k)$  à moins que le signal est complètement noyé dans du bruit.

Notre critère est basé sur l'idée de blanchiment de l'erreur de prédiction (Critère BEP). Le principe consiste à choisir l'ordre qui correspond à la valeur minimale de  $\eta(m)$  où  $m = \{1, \dots, \text{Ordre\_Max}\}$  et  $\eta(m)$  est définie comme étant le rapport de la moyenne arithmétique des éléments de l'AC de l'erreur de prédiction à l'ordre  $m$ :  $R_{e_m}(k) \ k \geq 1$  à la valeur  $R_{e_m}(0)$ :

$$\eta(m) = \frac{\frac{1}{L} \sum_{i=1}^L |R_{e_m}(i)|}{R_{e_m}(0)} \quad (4.37)$$

En effet, le fait de vouloir minimiser  $\eta(m)$  revient à chercher l'ordre pour lequel la fonction d'autocorrélation de l'erreur de prédiction soit concentrée à l'origine, et par conséquent,  $e(k)$  soit proche d'un bruit blanc. Le minimum de  $\eta(m)$  correspondra à l'ordre  $p$  du modèle AR le mieux adapté à la modélisation du signal.

#### 4.8.3 Optimisation du temps machine

Le temps machine nécessaire au calcul d'une valeur  $\eta(m)$  est extrêmement long. La majeure partie de ce temps est consommé par le calcul des éléments de l'AC  $R_e(k)$ . Alors on s'y pose la question suivante:

Est-il possible d'exploiter la structure en treillis pour calculer  $R_e(k)$  d'une manière récursive, et par conséquent plus rapidement que la méthode directe ?

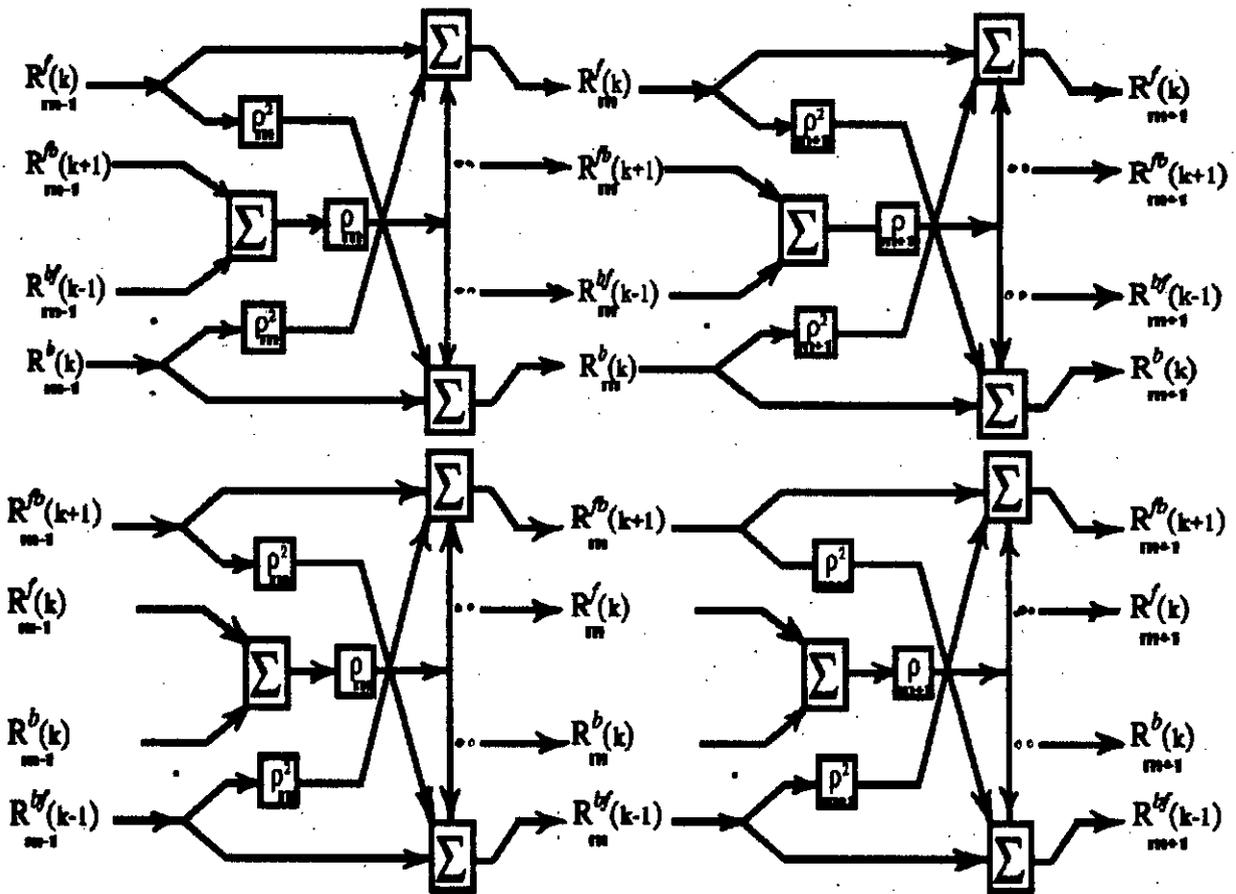
En effet, les fonctions d'inter- et d'autocorrélation, des erreurs de prédiction directe et inverse d'ordre  $m$ , sont définies par les expressions suivantes.

$$\begin{aligned} R_{e_m}^f(k) &= \mathbb{E}(E_m^f(n+k)E_m^f(n)) & R_{e_m}^b(k) &= \mathbb{E}(E_m^b(n+k)E_m^b(n)) \\ R_{e_m}^{fb}(k) &= \mathbb{E}(E_m^f(n+k)E_m^b(n)) & R_{e_m}^{bf}(k) &= \mathbb{E}(E_m^b(n+k)E_m^f(n)) \end{aligned} \quad (4.37)$$

En reportant dans (4.37) les expressions récursives des erreurs  $E_m^f$  et  $E_m^b$  de (3.27), les expressions (4.37) prennent la forme:

$$\begin{aligned}
 R_{m-1}^f(k) &= R_{m-1}^f(k) + \rho_m^2 R_{m-1}^b(k) + \rho_m (R_{m-1}^{fb}(k+1) + R_{m-1}^{bf}(k-1)) \\
 R_{m-1}^b(k) &= R_{m-1}^b(k) + \rho_m^2 R_{m-1}^f(k) + \rho_m (R_{m-1}^{fb}(k+1) + R_{m-1}^{bf}(k-1)) \\
 R_{m-1}^{fb}(k) &= R_{m-1}^{fb}(k+1) + \rho_m^2 R_{m-1}^{bf}(k-1) + \rho_m (R_{m-1}^f(k) + R_{m-1}^b(k)) \\
 R_{m-1}^{bf}(k) &= R_{m-1}^{bf}(k-1) + \rho_m^2 R_{m-1}^{fb}(k+1) + \rho_m (R_{m-1}^f(k) + R_{m-1}^b(k))
 \end{aligned}
 \tag{4.38}$$

ce qui correspond aux structures en treillis suivantes:



Structure en treillis  
des fonctions d'auto- et d'intercorrélation des erreurs  
de prédiction directe et rétrograde

Cette méthode conduit pratiquement au même résultat que la méthode directe (figure 4.20), mais en un nombre limité d'opérations élémentaires qui est de l'ordre de  $7L$  multiplications par itération, où  $L$  est le nombre d'éléments de la fonction d'autocorrelation à estimer. Tandis que le calcul par la méthode directe nécessite  $N.L$  multiplications. Le calcul par la structure en treillis s'avère beaucoup plus rapide.

#### 4.6.2.4 Filtrage du critère

L'examen de quelques courbes nous a permis de constater que, dans certains cas, la courbe représentant l'évolution du critère, présente des variations plus ou moins irrégulières (voir figures 4.20, 4.21). La courbe en question contient plusieurs minimas locaux, ce qui rend le choix automatique de la valeur minimale du critère une tâche difficile et délicate. Néanmoins, il est possible d'y remédier à ce problème par une opération de filtrage.

Des simulations nous ont permis de conclure qu'un filtre passe bas du premier ordre élimine complètement les fluctuations rapides. La nouvelle courbe obtenue après filtrage possède un seul minimum, mais elle est déphasée par rapport à la courbe initiale. Ce qui va se traduire par une légère sur-estimation de l'ordre. Ceci n'est pas très gênant dans la mesure où il faut augmenter l'ordre afin de détecter les composantes fréquentielles. Toutefois, un choix judicieux des paramètres du filtre est nécessaire.

Dans le plan  $Z$ , la fonction de transfert du filtre est donnée par:

$$H(z) = \frac{1}{1 - b.z^{-1}}$$

où  $b$  est une constante réelle inférieure à l'unité.

Il résulte des simulations que la valeur  $b=0.707$  est un bon choix, et permet d'atteindre le meilleur compromis entre filtrage et déphasage.

Des résultats meilleurs ont été obtenus en considérant le logarithme du critère BEP pondéré par la variance de l'erreur de prédiction:  $\text{Ln}(\eta(p).e_p)$  (voir figures 4.22).

#### 4.6.2.5 Comparaison entre les différents critères

Afin d'évaluer les performances des différents critères, on a procédé à l'estimation de l'ordre du modèle AR donné à la section 4.5, par différents critères de sélection (FPE, AIC, MDL, CAT et BEP). L'estimation est basée sur un nombre de points très limité (seulement 50 échantillons du signal). On rappelle que l'ordre exacte du modèle est égale à 4.

Sur 100 réalisations, nous reportons dans le tableau 4.3 le nombre de fois où l'ordre a été sélectionné.

Tableau 4.3: Comparaison entre les différents critères.

Ordre	FPE	AIC	MDL	CAT	BEP
1	01	01	01	01	00
2	15	14	29	13	00
3	28	18	18	16	18
4	34	40	48	42	10
5	18	20	03	22	30
6	02	05	01	04	24
7	01	02	00	02	22
8	00	00	00	00	04
9	00	00	00	00	02

L'examen des résultats de ce tableau, montre que les critères MDL et FPE ont tendance à sous-estimer l'ordre, tandis que le BEP donne un ordre légèrement surestimé.

#### 4.7 Autres méthodes d'estimation de la DSP

Cette étude serait incomplète sans aborder -au moins d'une manière succincte- certaines techniques d'estimation de la DSP liées à des signaux particuliers (méthodes de Pisarenko et de Prony), ainsi que la méthode du maximum de vraisemblance. Enfin, il est intéressant de mettre en évidence la relation entre ces méthodes et les méthodes Autorégressives.

##### 4.7.1 Méthode du maximum de la vraisemblance

la méthode du maximum de vraisemblance dite de Capon [47], est basée sur la mise en oeuvre d'un banc de filtres dont les coefficients sont ajustés pour chacune des fréquences en vue de minimiser la variance.

Afin de préserver l'amplitude d'une sinusoïde à une fréquence  $f_0$ , les coefficients  $h_i$  du filtre sont soumis à la contrainte:

$$\sum_{i=0}^P h_i \cdot \exp(-j2\pi f_0 k T) = 1 \quad T=1/f_0 \quad (4.39)$$

Le maximum de vraisemblance est obtenu pour une variance minimale donnée par [47]:

$$\sigma_{mv}^2 = \frac{1}{\mathbf{H}^{-1} \cdot \mathbf{W}_p(f_0) \cdot \mathbf{R}_p^{-1} \cdot \mathbf{W}_p(f_0)} \quad (4.40)$$

$\mathbf{W}(f)$  est le vecteur des exponentielles complexes défini par:

$$\mathbf{W}_p(f) = \left[ 1, \exp(j2\pi f T), \dots, \exp(j2\pi f p T) \right]^T$$

$\mathbf{W}_p^H$  est la transposée hermitienne de  $\mathbf{W}_p$ .  $\mathbf{R}_p^{-1}$  est l'inverse de la matrice d'AC qui doit être estimée à partir des données disponibles.

En supposant que ce calcul est effectué pour chaque valeur de la fréquence, on obtient une estimation de la DSP

$$S_{mv}(f) = \frac{T}{W_P^H(f) \cdot R_P^{-1} \cdot W_P(f)} \quad (4.41)$$

La décomposition triangulaire (4.32) de la matrice  $R_P^{-1}$  donne

$$\frac{1}{S_{mv}(f)} = \frac{1}{T} W_P^H(f) \cdot R_P^{-1} \cdot W_P(f) \quad (4.42)$$

$$\frac{1}{S_{mv}(f)} = \frac{1}{T} W_P^H(f) \cdot A_P \cdot P_P^{-1} \cdot A_P \cdot W_P(f) \quad (4.43)$$

$$= \frac{1}{T} \sum_{k=0}^P \frac{1}{\alpha_k} \left| \sum_{m=0}^k a_k(m) \exp(-j2\pi f m T) \right|^2 \quad (4.44)$$

$$\frac{1}{S_{mv}(f)} = \sum_{k=0}^P \frac{1}{S_k^{AR}(f)} \quad (4.45)$$

Ainsi  $S_{mv}(f)$  apparaît comme la moyenne harmonique des estimateurs AR pour les ordres 0, ..., p, et par suite, elle offre une résolution moins bonne que l'estimation AR d'ordre le plus élevé. Car dans cette approche, l'accent est mis sur la minimisation de la variance de l'erreur de prédiction [12].

#### 4.7.2 Méthode de Pisarenko

Pisarenko a montré que pour un ensemble fini d'éléments de la fonction d'AC  $r[0 \dots p]$ , une estimation de la DSP peut être obtenue par l'expression [50]

$$S(f) = \sigma^2 + \sum_{k=0}^P \frac{A_k^2}{2} \delta(f - f_k) \quad (4.46)$$

C'est-à-dire que le signal  $y(k)$  peut être considéré comme un ensemble de sinusoides d'amplitudes  $A_k$  dans un bruit de variance  $\sigma^2$ . Le modèle de Pisarenko (équation 4.46) est un cas limite d'une classe de méthodes: bruit blanc + modèle AR d'ordre  $p$ .

En effet, la matrice  $R_p$  peut se mettre sous la forme suivante:

$$R_p = \underbrace{\sigma^2 I_p}_{\text{Bruit blanc}} + \underbrace{\left( R_p - \sigma^2 I_p \right)}_{\text{Modèle AR}(p)} \quad (4.47)$$

La seule condition est que le deuxième terme  $\left( R_p - \sigma^2 I_p \right)$  soit défini positif (toutes les valeurs propres  $\lambda_k$  sont positives). Donc au maximum la variance du bruit est la plus petite valeur propre  $\lambda_{\min}$  de la matrice  $R_p$ , c'est-à-dire:

$$0 \leq \sigma^2 \leq \lambda_{\min} \quad (4.48)$$

Les coefficients du modèle AR sont donc les éléments du vecteur propre associé.

#### 4.7.3 Méthode de Prony

Parmi les méthodes d'analyse spectrale fondées sur une modélisation, la méthode de Prony [19], est particulièrement adaptée aux signaux de vibration. Elle s'avère comme un outil puissant pour l'étude des signaux transitoires [26].

La méthode de Prony consiste à décomposer un signal  $y(k)$  en une somme pondérée d'exponentielles (modes):

$$y(k) = \sum_{m=1}^P b_m w_m^k \quad k=0, \dots, N-1 \quad (4.49)$$

$$\text{Avec } \begin{cases} b_m = A_m \cdot \text{Exp}(j\theta_m) \\ W_m = \text{Exp}[(\delta_m + j2\pi f_m)T] \end{cases} \quad (4.50)$$

Toute la difficulté consiste à trouver les coefficients  $A_m$ ,  $\theta_m$ ,  $\delta_m$  et  $f_m$  qui rendent l'erreur quadratique moyenne minimale. La démarche à suivre pour résoudre ce problème peut se résumer en trois étapes [47],[19]:

1-Détermination des paramètres  $(a(0), \dots, a(p))$  de la prédiction linéaire compatibles avec les données  $y(k)$   $k \in [0 \dots N-1]$  :

$$e(k) = \sum_{i=0}^p a(i) \cdot y(k-i) \quad (4.51)$$

2-Recherche des racines du polynôme  $P(z)$  défini par

$$P(z) = \sum_{i=0}^p a(i) \cdot z^{p-i} = \prod_{m=1}^p (z - W_m) \quad (4.52)$$

afin de déterminer les modes  $W_m$  (fréquence  $f_m$  et atténuation  $\delta_m$ ).

3-Résolution d'un ensemble d'équations linéaires donné par (4.49), afin de déterminer les coefficients  $b_m$ , et par suite, les amplitudes  $A_m$  et les phases initiales  $\theta_m$  à partir des relations suivantes [19]:

$$\begin{aligned} A_m &= |b_m| \\ \theta_m &= \text{tg}^{-1} \left[ \text{Im}(b_m) / \text{Re}(b_m) \right] \\ \delta_m &= \text{Ln}(|Z_m|) \end{aligned} \quad (4.53)$$

Enfin il est à remarquer que l'astuce de la méthode de Prony consiste à ramener un problème de modélisation très non linéaire (4.49) à un problème linéaire (4.51) (modélisation AR) et à un problème de racines (4.52).

#### 4.8 Conclusions

Les méthodes paramétriques autorégressives sont particulièrement intéressantes: elles apportent d'incontestables progrès en matière de résolution spectrale et analyse des signaux de courte durée.

Il en résulte de cette étude, que chaque méthode met en relief une caractéristique bien déterminée de l'estimateur (variance, résolution, rapidité). C'est ainsi que du point de vue résolution la méthode de Pisarenko semble être la meilleure [47], tandis que pour une bonne stabilité statistique, il n'y a pas mieux que la méthode du maximum de vraisemblance. La méthode de Burg offre aussi une bonne résolution et s'apprête bien à une implementation avec la structure en treillis, mais malheureusement elle souffre de deux inconvénients: l'apparition de pics parasites et le biais des composantes sinusoidales en fonction de la phase initiale [14]. Ce n'est pas le cas des autres méthodes AR (Corrélation, Covariance, et covariance modifiée) qui donnent des résultats satisfaisants. Néanmoins elles requièrent un temps de calcul plus important.

Par ailleurs, les méthodes AR sont appropriées pour les signaux générés par des processus autorégressifs, la méthode de Capon aux signaux gaussiens stationnaires et la méthode de Prony aux signaux transitoires. Ainsi, une information sur la nature du signal analysé serait très utile et l'estimation de la DSP ne peut qu'en tirer bénéfice.

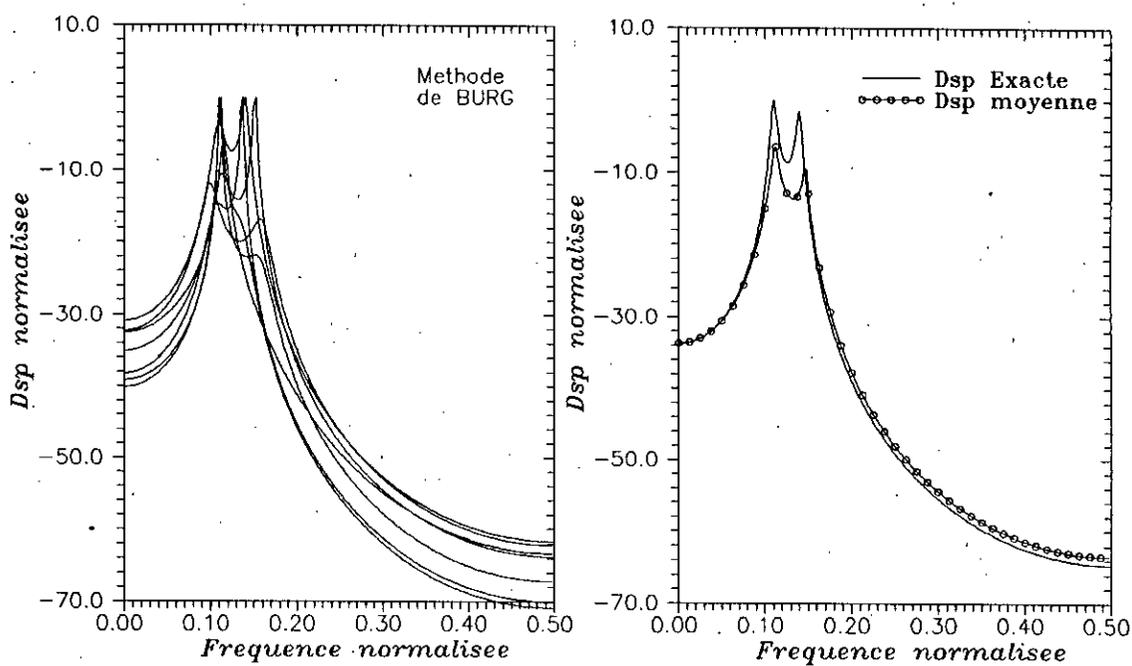


Figure 4.1.  
Dsp de plusieurs observations d'un processus AR(4),  $N_p=40$  points.

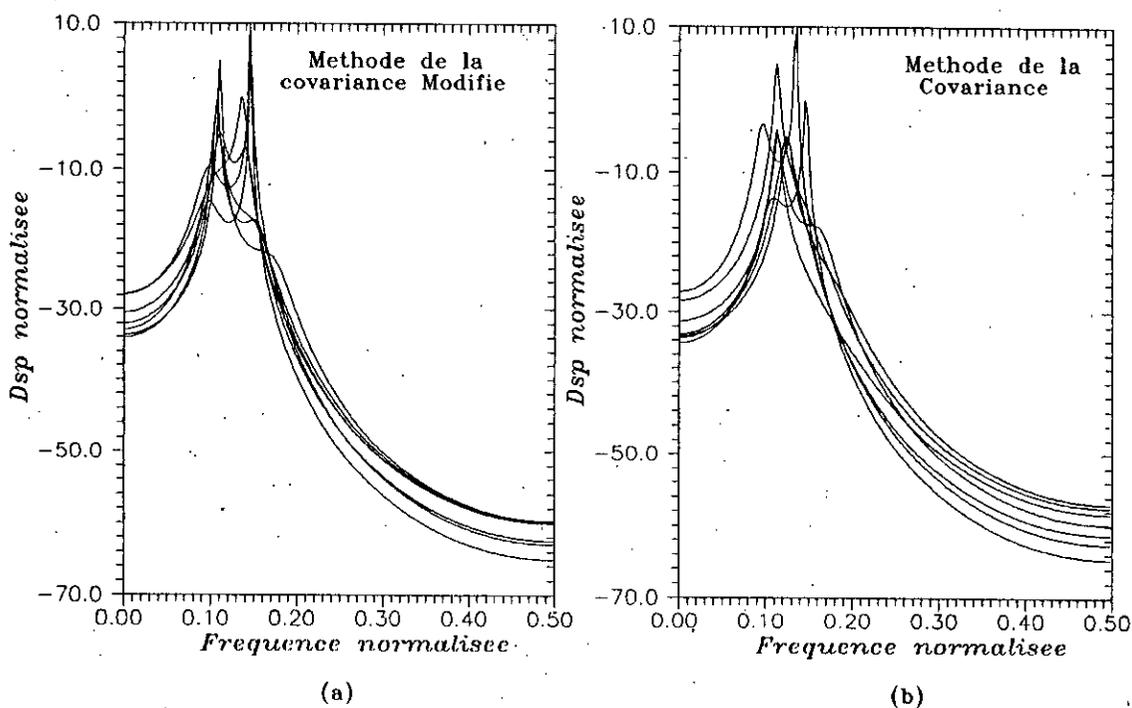


Figure 4.2.  
Dsp de plusieurs realisation d'un processus AR(4) avec  $N_p=40$  Points.

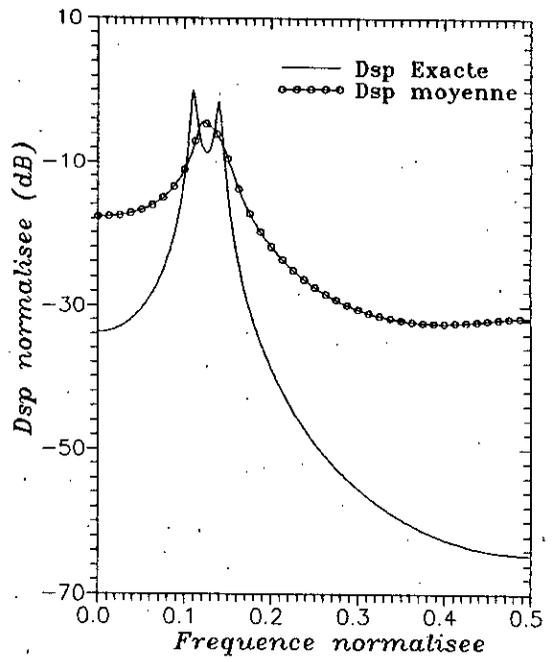
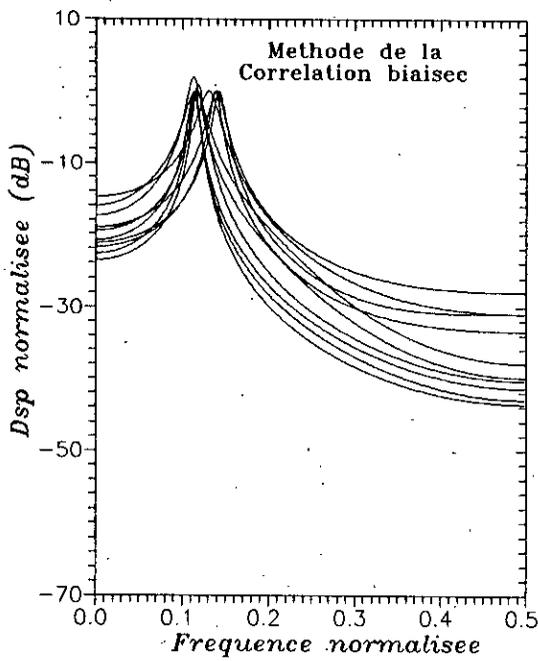


Figure 4.3.  
Dsp de plusieurs observations d'un processus AR(4),  $N_p=40$  points.

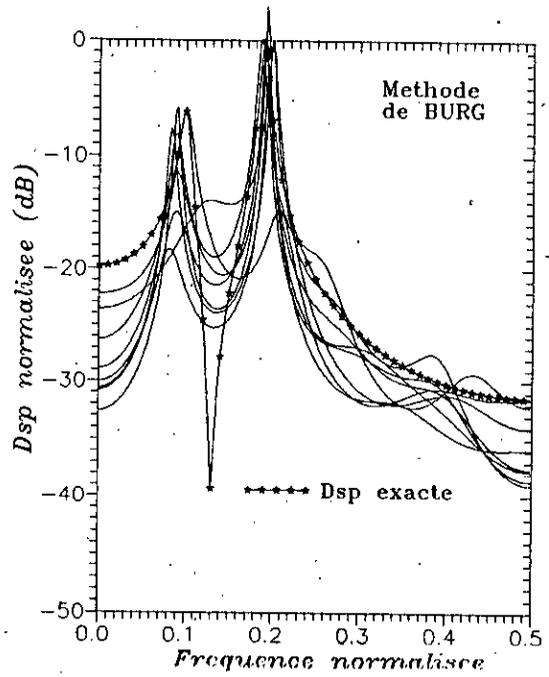
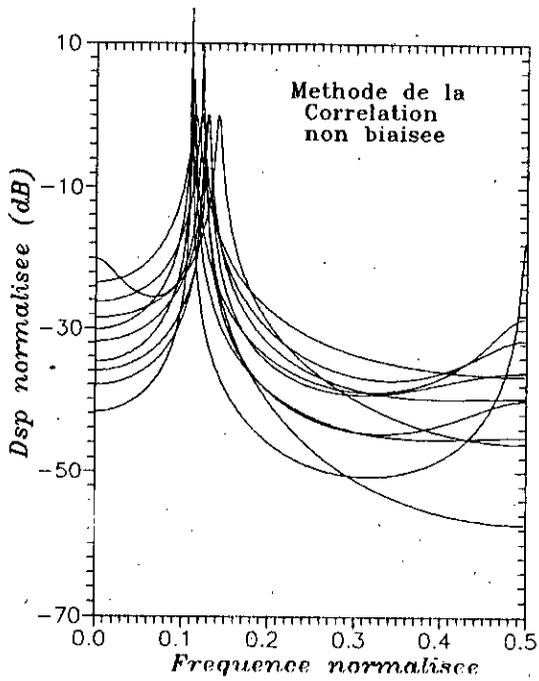


Figure 4.4  
Dsp de plusieurs realisations du modele AR(4) de l'exemple N 1, avec  $N_p=40$

Figure 4.5.  
Dsp de Plusieurs realisations d'un modele ARMA(4,2), avec  $N_p=40$  points

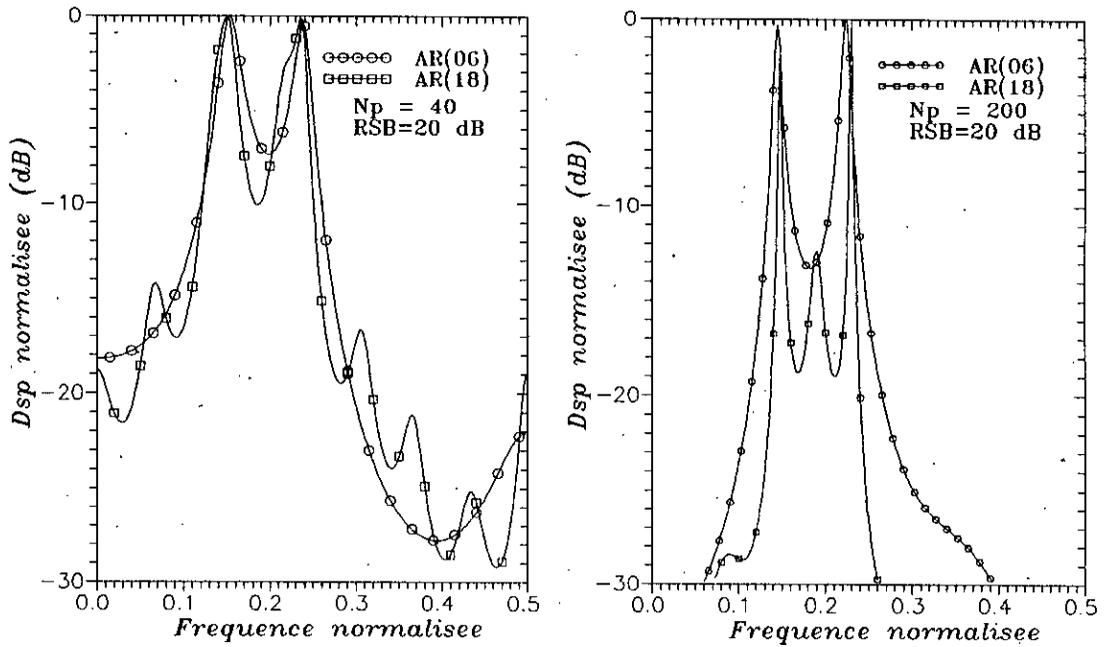


Figure 4.6. Resolution de la methode de correlation en fonction du nombre de points et de l'ordre du modele AR

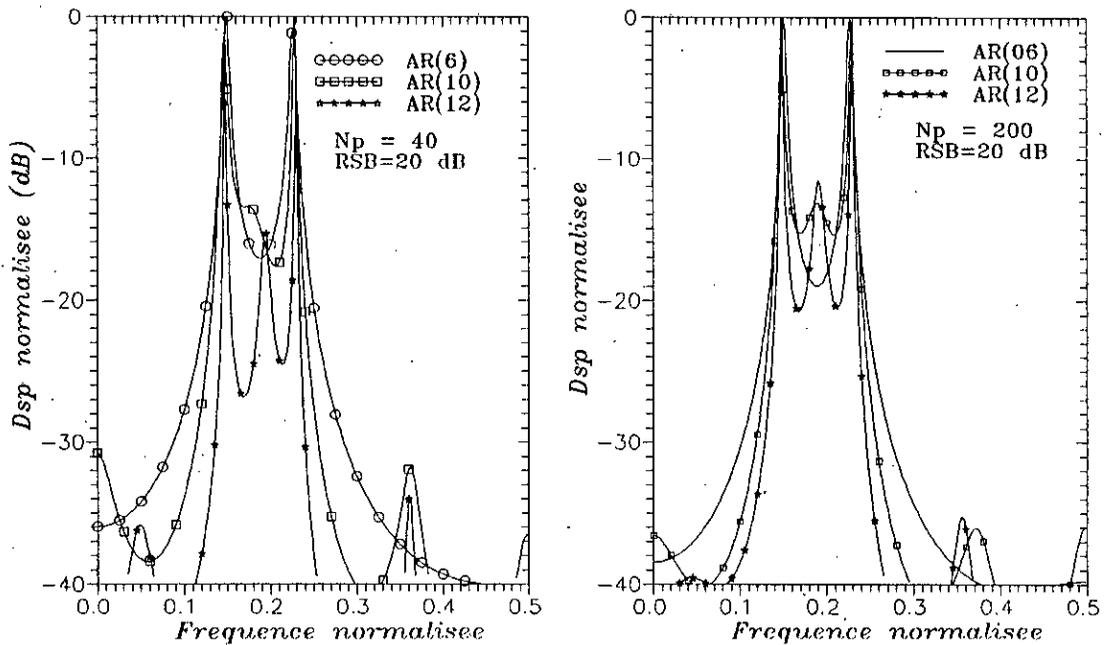


Figure 4.7. Resolution de la MEM en fonction du nombre de points et de l'ordre du modele AR

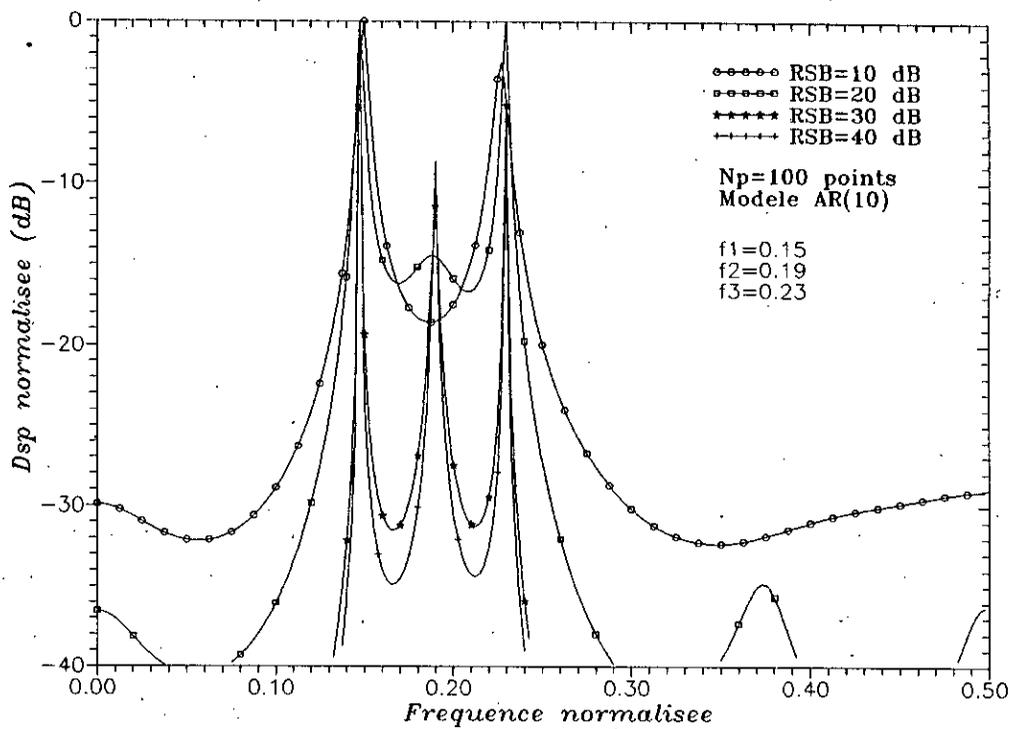


Figure 4.8. Resolution de la methode de BURG en fonction du RSB.

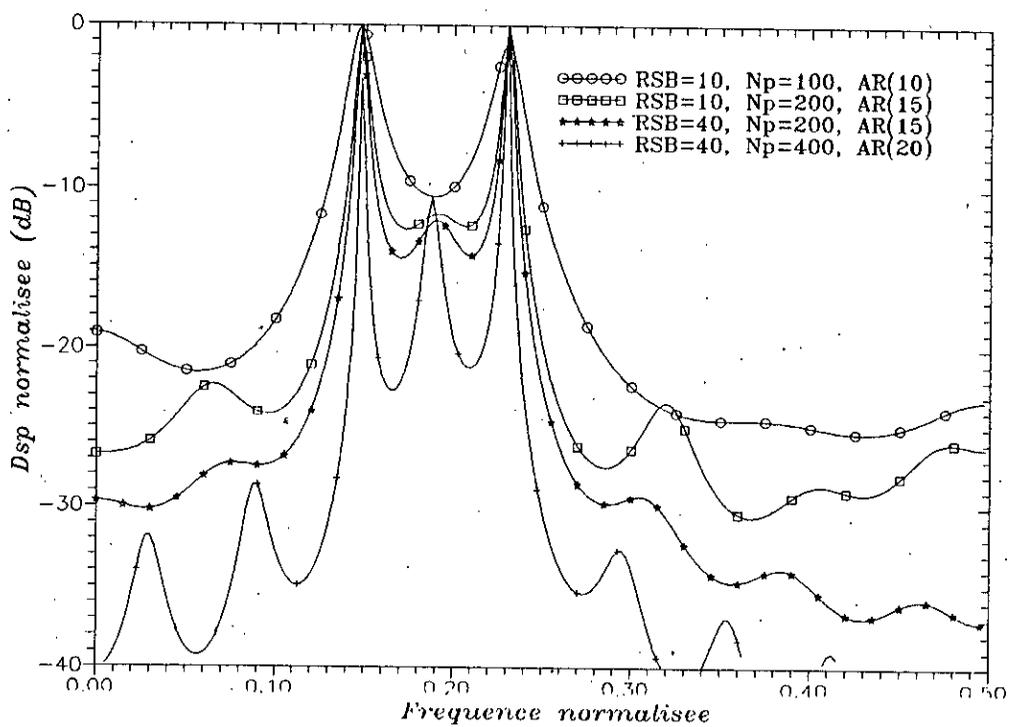


Figure 4.9. Resolution de la methode de Correlation en fonction du RSB

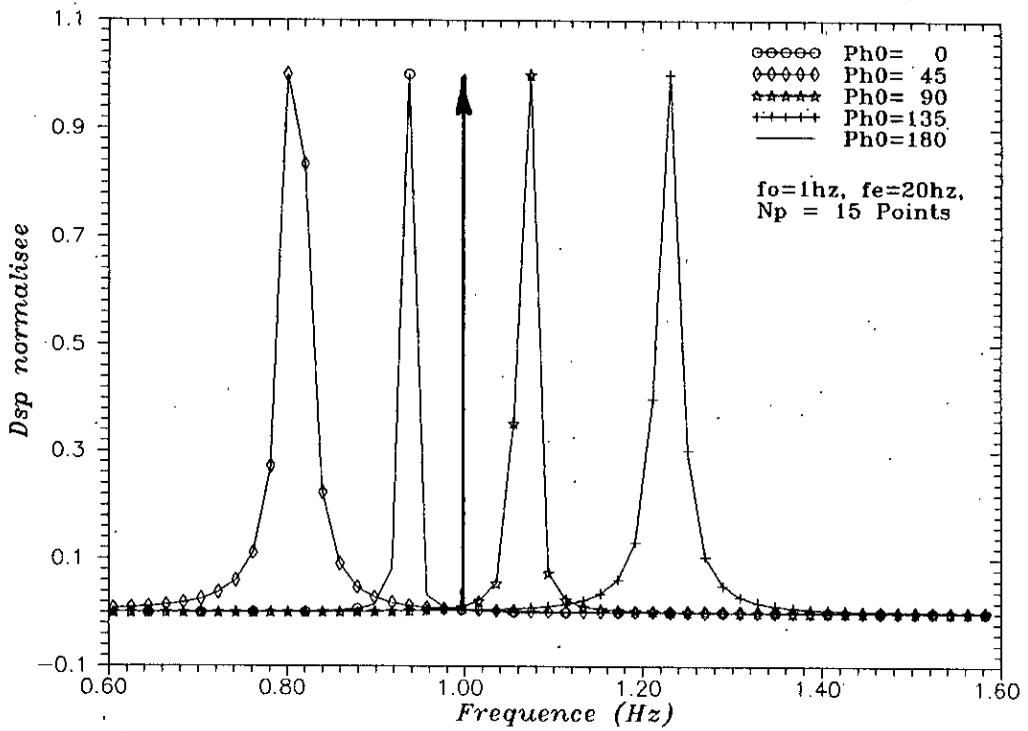


Figure 4.10.

Biais de la methode de BURG (MEM) en fonction de la phase initiale de la sinusoid.

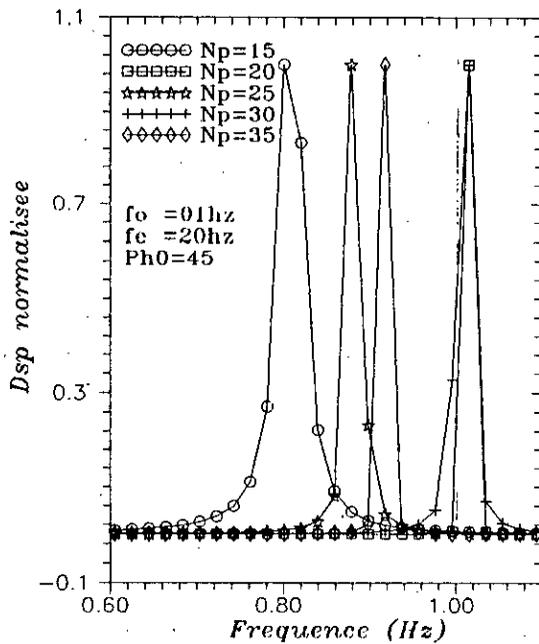


Figure 4.11.

Biais de la MEM en fonction du nombre de points  $N_p$

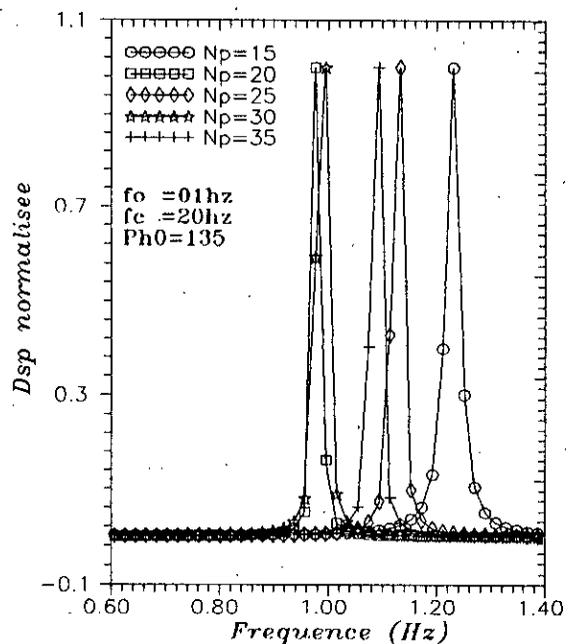


Figure 4.12.

Biais de la MEM en fonction du nombre de points  $N_p$

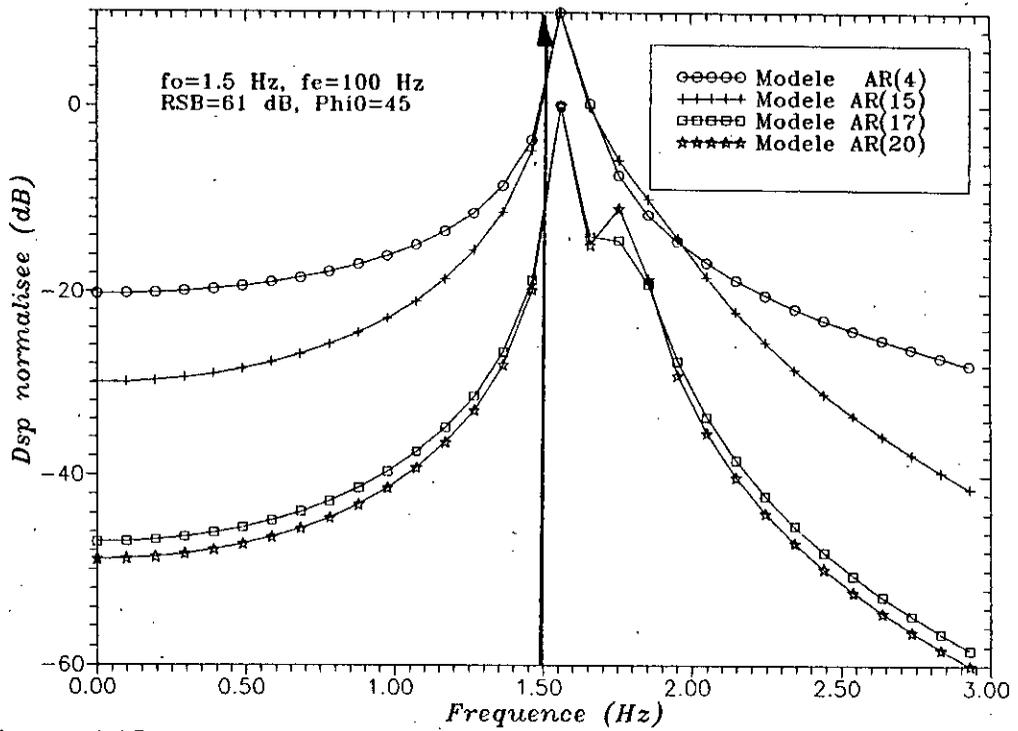


Figure 4.13.  
Le dedoublement des raies spectrales avec la MEM  
en fonction de l'ordre du modele

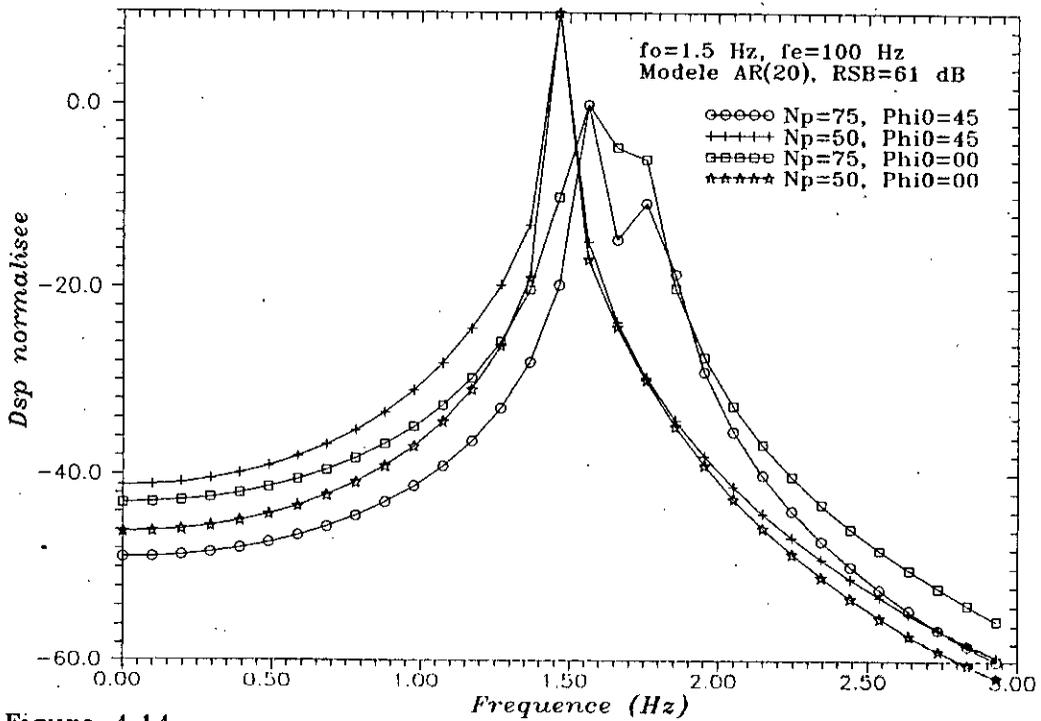


Figure 4.14.  
Le dedoublement des raies spectrales avec la MEM depend  
de la phase initiale de la sinusoide, et du nombre de  
points utilises dans l'estimation

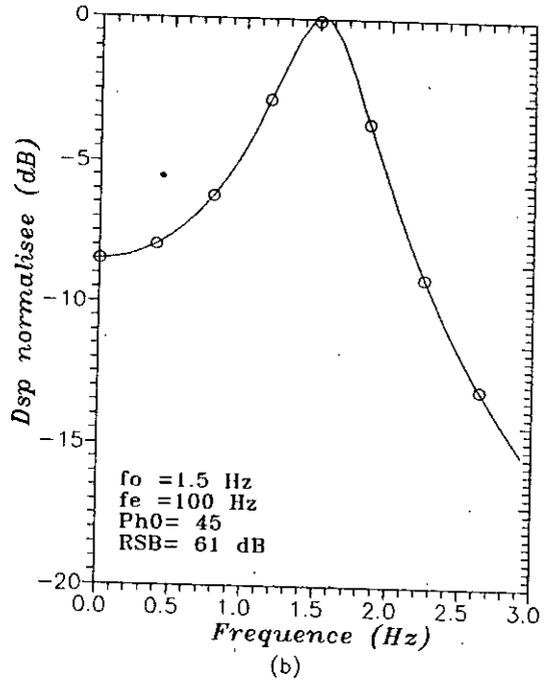
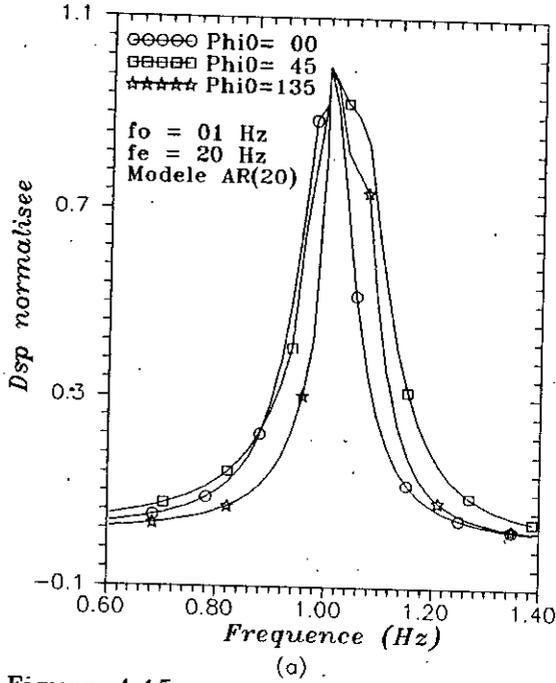


Figure 4.15. (a) Biais, (b) dedoubling des raies spectrales par la methode de correlation

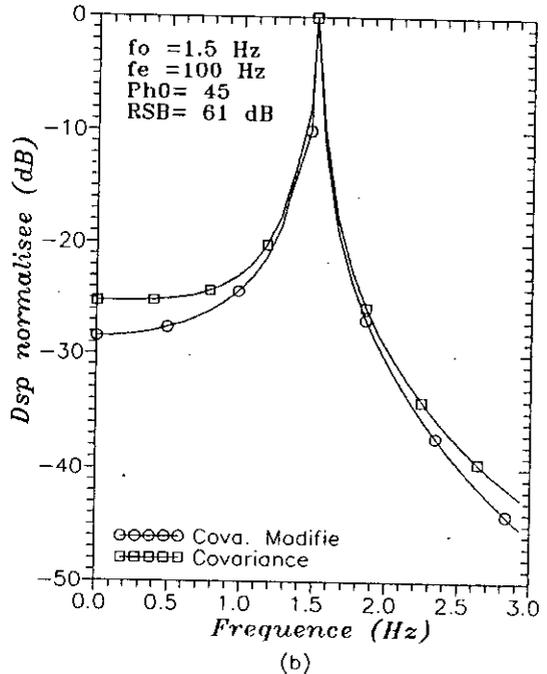
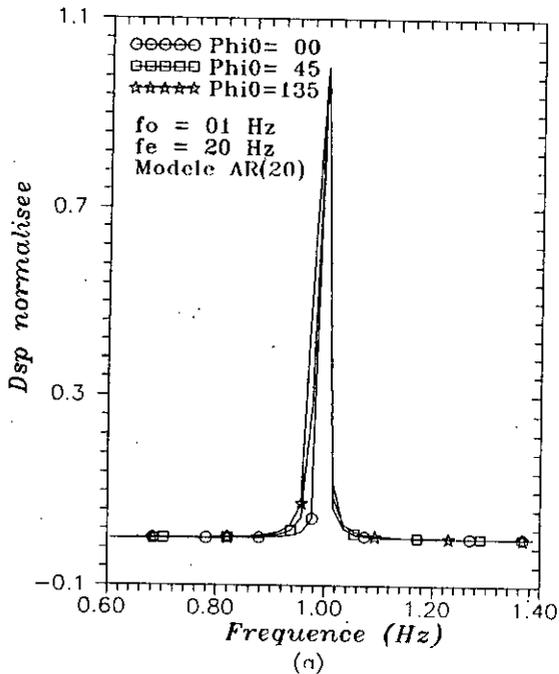


Figure 4.16. (a) Biais, (b) dedoubling des raies spectrales de la methode de covariance

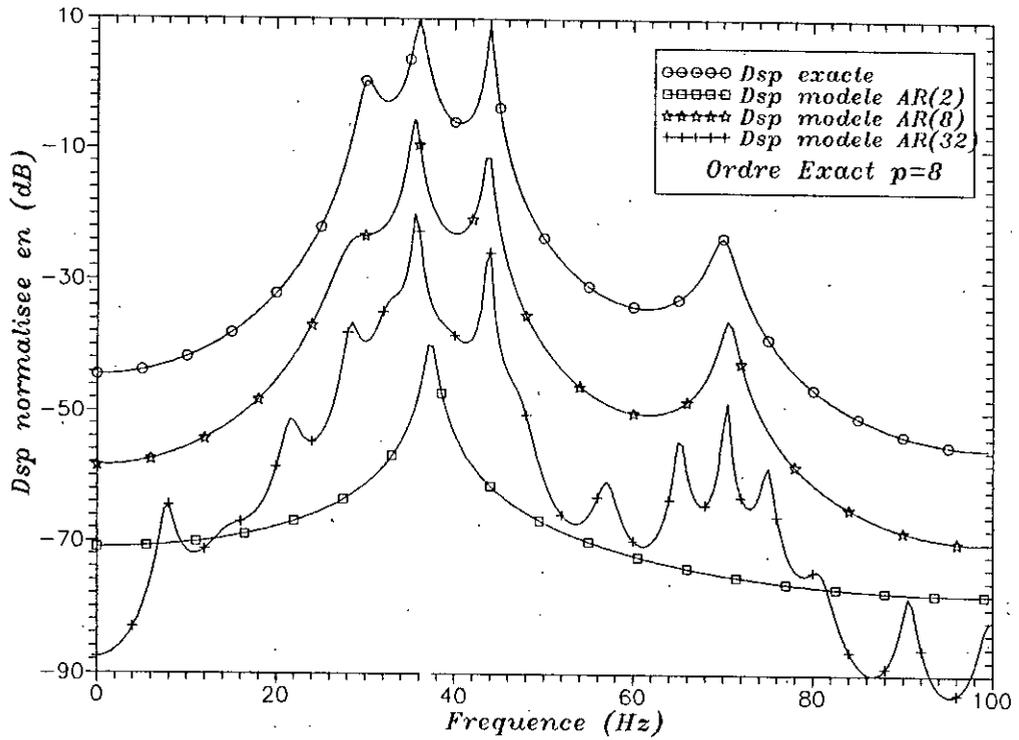


Figure 4.17  
 Probleme du choix de l'ordre du modele AR

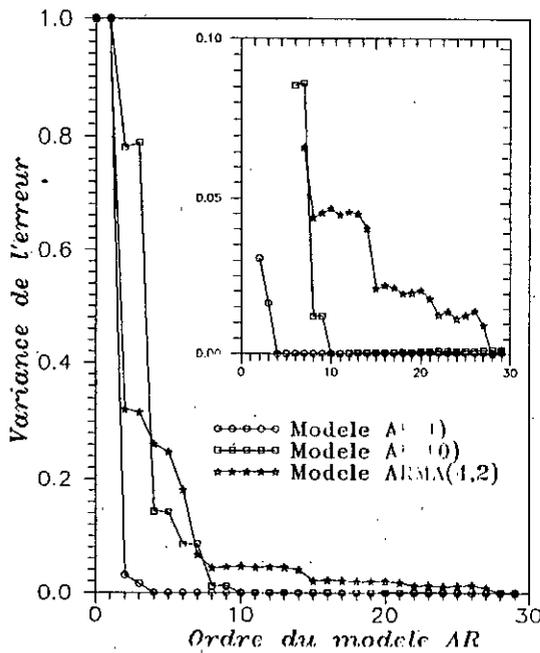


Figure 4.18  
 Variance de l'erreur de prediction de plusieurs modeles

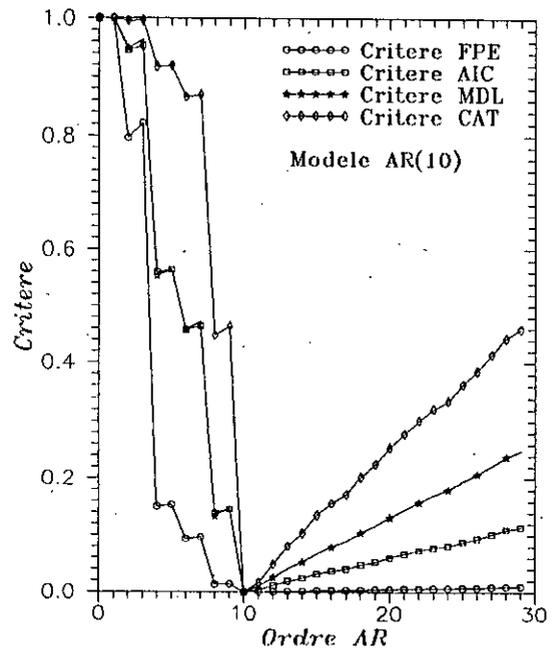


Figure 4.19.  
 Evolution des differents Criteres pour un modele AR(10)

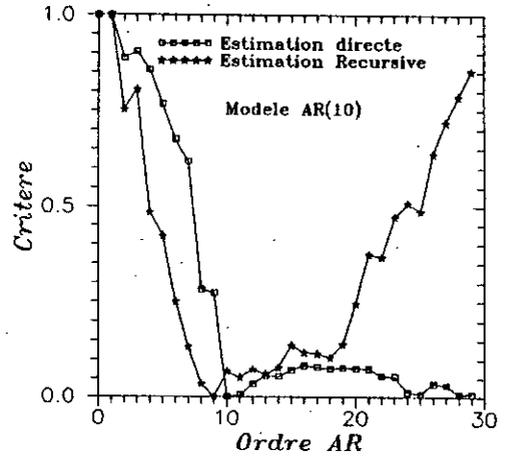
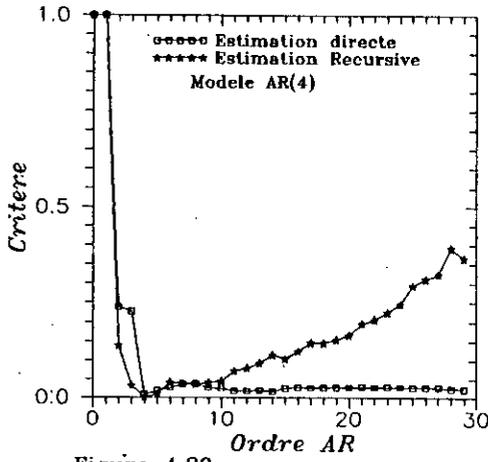


Figure 4.20. Estimation directe et Estimation recursive du critere de selection de l'ordre AR

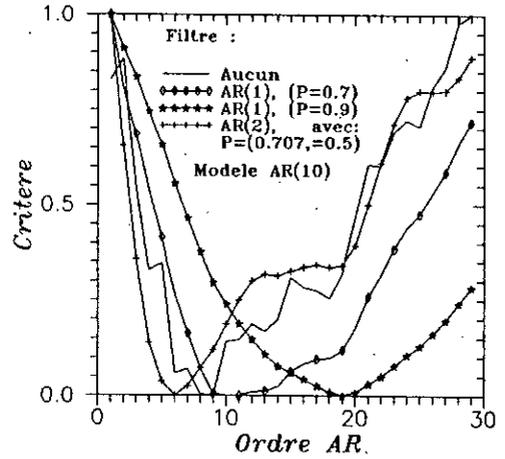
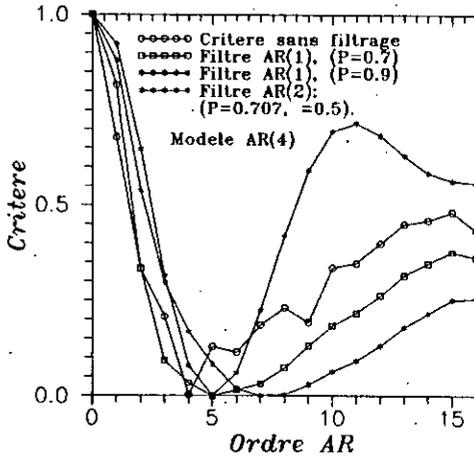


Figure 4.21. Filtrage du critere de selection de l'ordre

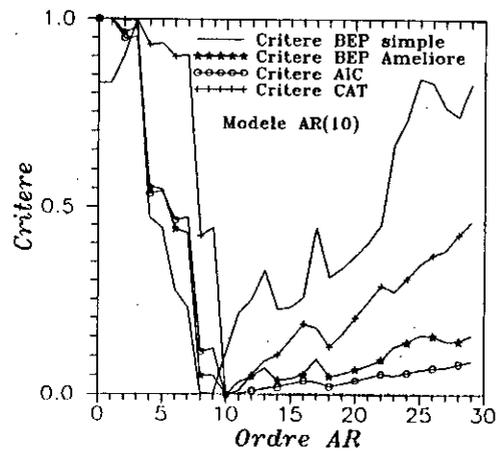
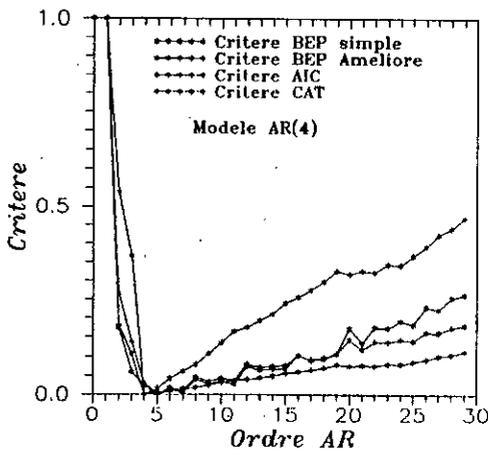


Figure 4.22. Amelioration du critere de selection de l'ordre AR

# **CHAPITRE V**

## **Extention aux signaux Multicanaux**

## CHAPITRE V

### EXTENSION AUX SIGNAUX MULTICANAUX

#### 5.1 Introduction

La plupart des algorithmes et techniques d'analyse spectrale, présentés dans les chapitres précédents, sont conçus pour les signaux -réels ou complexes- qu'on peut qualifier de signaux à simple canal. Ces signaux correspondent à un très large domaine d'applications.

Cependant, on rencontre dans de nombreuses applications, notamment dans le domaine de la géophysique, l'acoustique sous-marine, et les systèmes de transmission à antennes multiples, des signaux multicanaux<sup>(1)</sup>. Un signal multicanaux est formé de plusieurs signaux indépendants, mais supposés corrélés, sans quoi, ils pourraient être traités comme autant de signaux scalaires. Chaque signal représente une information partielle sur le système et son utilisation optimale nécessite un traitement simultané avec tous les autres signaux.

Pour l'étude de ce type de signaux, on est conduit logiquement à étendre les méthodes déjà étudiées (dans le cas d'un signal simple canal) au cas des signaux multicanaux.

#### 5.2 Matrice des DSP et fonction de cohérence

##### 5.2.1 DSP d'un signal multicanaux

Soit un signal multicanaux défini comme un vecteur à  $m$  éléments (canaux):

$$Y_m[k] = \left[ x_1(k), x_2(k), \dots, x_m(k) \right]^T, \quad k=0 \dots N-1$$

---

(1) Il faut marquer la différence entre les signaux multicanaux et les signaux multidimensionnels:

Un signal à plusieurs canaux est une fonction vectorielle d'une variable scalaire, exemple:  $Climat(t) = f(T(t), P(t), \dots)$ .

Un signal multidimensionnel est une fonction scalaire de plusieurs variables, exemple:  $Image(x, y, z, t)$ .

Dans la pratique, il s'agit le plus souvent soit de signaux issus d'un nombre fini ( $m$ ) de capteurs, soit de plusieurs enregistrements représentant différents paramètres d'un processus complexe (par exemple le climat (température, pression, pluviométrie, humidité ..etc) ). La DSP recherchée est une matrice  $S(f)$  de dimension  $(m,m)$ :

$$S(f) = \begin{bmatrix} S_{11}(f) & S_{12}(f) & \dots & S_{1m}(f) \\ S_{21}(f) & S_{22}(f) & \dots & \\ \vdots & & & \\ S_{m1}(f) & & \dots & S_{mm}(f) \end{bmatrix} \quad (5.1)$$

Les termes diagonaux  $S_{ii}(f)$ , sont les densités autospectrales de puissance des signaux élémentaires  $x_i(k)$ ,  $i=1..m$ . Le terme  $S_{ij}(f)$ ,  $i \neq j$ , représente la densité interspectrale de puissance (DIP) entre les deux canaux  $x_i(k)$  et  $x_j(k)$ ;  $i, j=1 .. m$ ;  $i \neq j$ . Si les signaux élémentaires  $x_1(k) .. x_m(k)$  sont indépendants ou simplement non corrélés, alors  $S(f)$  est une matrice diagonale.

### 5.2.2 Fonction de cohérence

Soit  $x_i(k)$  et  $x_j(k)$  deux canaux différents. Pour mesurer la similitude entre les deux signaux pour chaque fréquence, on utilise le rapport de l'interspectre à la racine carrée des autospectres; ce qu'on appelle fonction de cohérence<sup>(2)</sup>.

$$\phi_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f) S_{jj}(f)}} \quad (5.2)$$

(2) Certains auteurs définissent la fonction de cohérence par

$$\phi_{ij}(f) = \frac{|S_{ij}(f)|^2}{S_{ii}(f) S_{jj}(f)}$$

La DIP  $S_{ij}(f)$  peut être complexe, même si les deux signaux  $x_i(k)$  et  $x_j(k)$  sont réels, alors la cohérence  $\phi_{ij}(f)$  est en général une fonction complexe.

On définit deux autres fonctions:

-le carré du module de la fonction de cohérence (MSC=*Magnitude Squared Coherence*)

$$MSC(f) = |\phi_{ij}(f)|^2 = \frac{|S_{ij}(f)|^2}{S_{ii}(f) S_{jj}(f)} \quad (5.3)$$

-La phase de la fonction de cohérence:

$$\theta_{ij}(f) = \text{tg}^{-1} \left[ \text{Img}[\phi_{ij}(f)] / \text{Re}[\phi_{ij}(f)] \right] \quad (5.4)$$

La fonction de cohérence peut être vue comme étant une densité interspectrale normalisée. Normalisée du faite que le MSC est un coefficient variant entre 0 et 1:

$$0 \leq MSC(f) = |\phi_{ij}(f)|^2 \leq 1 \quad \forall f \quad (5.5)$$

La cohérence, particulièrement le MSC, est une mesure de la linéarité entre deux processus: une valeur nulle indique que les deux processus sont soit indépendants soit caractérisés par une relation non linéaire [50].

Le MSC trouve son utilisation dans de nombreux champs d'application: mesure des fonctions de transfert, du rapport signal sur bruit des systèmes linéaires, et du décalage temporel entre deux signaux [50] ... etc.

### 5.2.3 Application de la fonction de Cohérence au test d'homogénéité des séries chronologiques.

L'étude des séries d'observations met parfois en évidence des changements de valeurs qui surviennent brusquement et se maintiennent ultérieurement, on dit alors que ces séries ne

sont pas homogènes. La question d'homogénéité ou de non homogénéité des observations, est un des problèmes les plus importants, compte tenu des conséquences qu'entraîne l'utilisation des séries non cohérentes: les conclusions peuvent se trouver faussées ou même inversées.

Nous avons imaginé une méthodologie de détection et de correction de la non homogénéité des séries chronologiques. Le principe est très simple: on considère la série comme étant un signal multicanaux et constituée de la juxtaposition de plusieurs enregistrements représentant plusieurs signaux. Si l'homogénéité d'un enregistrement est remise en cause alors on calcule la fonction de cohérence entre deux enregistrements différents: une valeur maximale proche de l'unité du MSC indique une parfaite cohérence des mesures. Tandis qu'une valeur nulle ou très faible nous informe que, soit le processus observé n'est pas stationnaire, ou bien les observations ne sont pas homogènes. Dans ce cas, pour identifier l'enregistrement qui présente un défaut d'homogénéité, on compare les deux enregistrements précédents à un troisième présumé homogène.

Une fois les mesures incohérentes sont identifiées, il est possible d'apporter une correction. A partir des données cohérentes; la série est modélisée par un modèle autorégressif par exemple. Les coefficients de ce modèle seront utilisés ensuite pour générer des données plus cohérentes et homogènes avec les mesures.

A titre d'exemple simple montrant les possibilités de cette méthodologie, nous l'avons appliqué sur les données de la figure 5.1.a. L'examen des graphes des figures 5.1.b et 5.1.d permet de déceler la non homogénéité de l'enregistrement N°2. Après correction de ce dernier par prédiction linéaire, les données sont plus cohérentes (figure 5.1.e).

### 5.3 Estimation AR de la DSP

Considérons le modèle AR multicanaux:

$$Y_m[n] + \sum_{k=1}^p A_m[k] Y_m[n-k] = E_m^f[n] \quad (5.6)$$

où  $Y_m[n]$  est l'échantillon du signal vectoriel stationnaire de dimension  $m$  à l'instant  $n$ . Les matrices  $A_m[k]$  de dimension  $(m \times m)$  sont les paramètres du modèle AR d'ordre  $p$ .

Avec l'hypothèse selon laquelle l'entrée  $E_m^f[n]$  est un bruit blanc vectoriel constitué de  $m$  bruits indépendants, de matrice de covariance diagonale et constante notée  $P_m$ , la DSP peut être exprimée par [47]:

$$S_{AR}(f) = T \underline{A}^{-1}(f) P_m \underline{A}^{-H}(f) \quad (5.7)$$

$$\text{où } \underline{A}(f) = I_m + \sum_{k=1}^p A_m[k] \text{Exp}(-j2\pi kf) \quad (5.8)$$

$T$  est la période d'échantillonnage

et  $\underline{A}_m^H$  est la transposée hermitienne de la matrice  $A_m$ .

### 5.4 Equations de Yule-Walker

Les méthodes couramment utilisées minimisent une fonction de l'erreur de prédiction directe  $E_m^f$  par rapport aux paramètres du modèle AR. Ceci conduit aux équations de Yule-Walker suivantes [47]:

$$\left[ I_m \ A_m[1] \ \dots \ A_m[p] \right] R_p = \begin{bmatrix} P_m^f & 0 & \dots & 0 \end{bmatrix} \quad (5.9)$$

De même en considérant le modèle AR d'ordre  $p$  correspondant à l'estimation rétrograde:

$$Y_m[n-p-1] + \sum_{k=1}^p B_m[k] Y_m[n-p-1+k] = E_m^b[n-1] \quad (5.10)$$

on aboutit aux équations (5.11)

$$\left( B_m[p] \ B_m[p-1] \ \dots \ B_m[1] \ I_m \right) R_p = \left[ 0 \ \dots \ 0 \ P_p^b \right] \quad (5.11)$$

Les matrices  $P_p^f$  et  $P_p^b$  représentent les variances des erreurs de prédiction et sont définies par:

$$P_p^f = E \left\{ E_p^f[n] (E_p^f[n])^T \right\} \quad P_p^b = E \left\{ E_p^b[n-1] (E_p^b[n-1])^T \right\} \quad (5.12)$$

la matrice de covariance est définie par

$$P_p^{fb} = E \left\{ E_p^f[n] (E_p^b[n-1])^T \right\} \quad (5.13)$$

Dans le cas général, les variances  $P^f$ ,  $P^b$ , ainsi que la covariance  $P^{fb}$  sont estimées respectivement par  $\hat{P}_m^f$ ,  $\hat{P}_m^b$  et  $\hat{P}_m^{fb}$  données par les relations [47]:

$$\hat{P}_m^f = \frac{1}{N} \sum_{n=p+2}^N E_p^f[n] (E_p^f[n])^T \quad (5.14)$$

$$\hat{P}_m^b = \frac{1}{N} \sum_{n=p+2}^N E_p^b[n-1] (E_p^b[n-1])^T \quad (5.15)$$

$$\hat{P}_m^{fb} = \frac{1}{N} \sum_{n=p+2}^N E_p^f[n] (E_p^b[n-1])^T \quad (5.16)$$

La matrice de corrélation  $R_p$  utilisée en (5.9) et (5.11) est Toeplitz par bloc, elle est formée à partir des  $(p+1) \times (p+1)$  matrices  $R_{YY}(k)$ :

$$R_{YY}(k) = E \left\{ Y_m[n] Y_m^T[n-k] \right\} \quad (5.17)$$

$$R_p = \begin{bmatrix} E \left\{ Y_m[n] Y_m^T[n] \right\} & \dots & E \left\{ Y_m[n] Y_m^T[n-p] \right\} \\ \vdots & & \vdots \\ E \left\{ Y_m[n-p] Y_m^T[n] \right\} & \dots & E \left\{ Y_m[n-p] Y_m^T[n-p] \right\} \end{bmatrix}$$

$$R_p = \begin{pmatrix} R_{YY}^{(0)} & R_{YY}^{(1)} & & R_{YY}^{(p)} \\ R_{YY}^{(-1)} & & & \\ & & & \\ R_{YY}^{(-p)} & & & R_{YY}^{(0)} \end{pmatrix} \quad (5.18)$$

Les matrices  $R_{YY}^{(k)}$  possèdent une certaine propriété de symétrie:

$$R_{YY}^{(-k)} = R_{YY}^{(k)T}.$$

En ce qui concerne l'estimation des matrices  $R_{YY}^{(k)}$  à partir des données  $Y_m[0]$ ,  $Y_m[1]$ , ...,  $Y_m[N-1]$ , on préfère l'estimateur biaisé [47]

$$R_{YY}^{(k)} = \frac{1}{N} \sum_{n=0}^{N-k} Y_m[n+k] Y_m^T[n] \quad (5.19)$$

ceci afin de garantir que la matrice  $R_{YY}^{(k)}$  soit définie positive.

### 5.5 Algorithme de Levinson-Wiggins-Robinson (LWR)

Une solution récursive sur l'ordre du modèle pour les deux systèmes (5.9) et (5.11) s'obtient par l'algorithme de Levinson-Wiggins-Robinson [15]: c'est une extension au cas multicanaux de l'algorithme de Levinson. En effet, en procédant d'une façon analogue à la démarche du paragraphe (3.6), on peut vérifier facilement que les paramètres des modèles autorégressifs direct et rétrograde sont donnés respectivement par les récursions suivantes:

$$A_{p+1}[k] = A_p[k] + A_{p+1}[p+1] B_p[p+1-k] \quad (5.20)$$

$$B_{p+1}[k] = B_p[k] + B_{p+1}[p+1] A_p[p+1-k] \quad (5.21)$$

Les matrices de réflexion  $A_{p+1}[p+1]$  et  $B_{p+1}[p+1]$  sont définies par les expressions (5.22) et (5.23)

$$A_{p+1}[p+1] = -\Delta_{p+1} (P_p^b)^{-1} \quad (5.22)$$

$$B_{p+1}[p+1] = -\nabla_{p+1} (P_p^f)^{-1} \quad (5.23)$$

où  $\Delta_{p+1}$  est l'intercorrélation entre l'erreur de prédiction directe  $E_p^f$  et l'erreur de prédiction rétrograde  $E_p^b$ .

$$\Delta_{p+1} = E \left\{ E_p^f[n] (E_p^b[n-1])^T \right\} = P_p^{fb}, \quad \nabla_{p+1} = \Delta_{p+1}^T \quad (5.24)$$

L'intérêt des équations (5.20) et (5.21) vient du fait que l'on obtient en même temps les récurrences sur les erreurs de prédiction:

$$E_{p+1}^f[k] = E_p^f[k] + A_{p+1}[p+1] E_p^b[n-1] \quad (5.25)$$

$$E_{p+1}^b[k] = E_p^b[k-1] + B_{p+1}[p+1] E_p^f[k] \quad (5.26)$$

Ainsi la généralisation de l'algorithme de Levinson au cas vectoriel se fait sans difficultés. Toutefois, il faut remarquer que dans le cas multicanaux, les deux matrices de réflexion définies par (5.22) et (5.23), et la matrice de corrélation partielle donnée par l'expression <sup>(3)</sup> (5.27):

$$\Lambda_{p+1} = (P_p^{f1/2})^{-1} (P_p^{fb}) (P_p^{b1/2})^{-T} \quad (5.27)$$

sont trois matrices différentes. Tandis que dans le cas simple canal, le coefficient de réflexion est identique au coefficient de corrélation partielle [47].

(3) L'exposant 1/2 d'une matrice définie positive dénote la racine carrée de cette dernière, et elle est donnée par la matrice triangulaire inférieure de la décomposition de Cholesky.

Résumé de l'algorithme LWR

Entrées:

Ordre du modèle AR (P)

$R_{yy}(k)$   $k=1 \dots P$ , Matrices de corrélation du signal  $Y(k)$   
(équations (5.17) et (5.19)).

Sorties:

Paramètres du modèle AR:  $A[1], A[2], \dots, A[n]$  :

Variances des erreurs de prédiction:  $P_p^f, P_p^b$ .

INITIALISATION

$$P_o^f = P_o^b = R_{yy}(0), \quad A_o[0] = B_o[0] = I_m, \quad \Delta_o = R_{yy}(1), \quad \nabla_o = \Delta_o^T$$

Récursion

For  $p=0$  to  $P-1$  do

\*Estimation des matrices de réflexion  $A_{p+1}[p+1], B_{p+1}[p+1]$

$$A_{p+1}[p+1] = -\Delta_p (P_p^b)^{-1}, \quad B_{p+1}[p+1] = -\nabla_p (P_p^f)^{-1}$$

\*Actualisation des variances des erreurs de prédiction  $P_p^f, P_p^b$ :

$$P_{p+1}^f = (I_m - A_{p+1}[p+1] B_{p+1}[p+1]) P_p^f$$

$$P_{p+1}^b = (I_m - B_{p+1}[p+1] A_{p+1}[p+1]) P_p^b$$

\*Actualisation des paramètres du modèle AR:  $A_p[k], B_p[k]$

For  $i = 0$  to  $p+1$  do

$$A_{p+1}[i] = A_p[i] + A_{p+1}[p+1] B_p[p+1-i]$$

$$B_{p+1}[i] = B_p[i] + B_{p+1}[p+1] A_p[p+1-i]$$

End loop on  $i$

$$\Delta_{p+1} = \sum_{k=0}^p A_p[k] R_{yy}[p+1-k], \quad \nabla_{p+1} = \Delta_{p+1}^T$$

End loop on  $p$

## 5.6 Algorithme de NUTTALL-STRAND [47]

La version multicanaux de l'algorithme de BURG a été proposée indépendamment par Nuttall en 1976 et Strand en 1977. L'idée de base est d'estimer la matrice de corrélation partielle  $\hat{\Delta}_p$  directement à partir des données. Le procédé proposé consiste à estimer la matrice  $\Delta_p$ , qui minimise la fonction  $f(\Delta_p)$  donnée par la trace de la moyenne (arithmétique) pondérée des matrices de covariance des erreurs de prédiction directe  $P_p^f$  et rétrograde  $P_p^b$ .

$$f(\Delta_p) = \text{tr} \left\{ V_{p-1} \hat{P}_p^f + W_{p-1} \hat{P}_p^b \right\} \quad (5.28)$$

où  $V_p$ , et  $W_p$  sont des matrices définies positives quelconques.

et  $\hat{P}_p^f$ ,  $\hat{P}_p^b$  sont définies par les relations (5.14) et (5.15)

Nuttall, en choisissant  $V_p = (P_p^f)^{-1}$  et  $W_p = (P_p^b)^{-1}$  et en minimisant l'expression (5.28) sous les trois contraintes suivantes:

$$E_{p+1}^f[n] = E_p^f[n] - \Delta_p (P_p^b)^{-1} E_p^b[n-1] \quad (5.29)$$

$$E_{p+1}^b[n] = E_p^b[n-1] - (\Delta_p)^H (P_p^f)^{-1} E_p^f[n] \quad (5.30)$$

$$A_p[p] P_p^f = B_p[p] P_p^b \quad (5.31)$$

a aboutit au système

$$\hat{P}_p^f (P_p^f)^{-1} \hat{\Delta}_{p+1} + \hat{\Delta}_{p+1} (P_p^b)^{-1} \hat{P}_p^b = -2 \hat{P}_p^{fb} \quad (5.32)$$

Dans le cas d'un seul canal:  $P_p^f = P_p^b$  et  $A_{p+1}[p+1] = -\Delta_{p+1} P_p^f$ . Alors l'estimation  $A_{p+1}[p+1]$  est donnée par:

$$A_{p+1}[p+1] = \frac{-2 \hat{P}_p^{fb}}{\hat{P}_p^f + \hat{P}_p^b} \quad (5.34)$$

Pour les signaux multicanaux, l'expression (5.32) est un cas particulier du système d'équations bilinéaire de type de Lyapunov:

$$AX + XB = C \quad (5.35)$$

où A, B, C, et X sont des matrices d'ordre (m x m).

Plusieurs auteurs se sont intéressés à la résolution de ce système. Pour Nuttall, la solution est de la forme:

$$X = P \cdot Q^{-1}$$

$$P = \sum_{k=0}^{m-1} (-1)^k A_k C B^{m-1-k}, \quad Q = \sum_{k=0}^m (-1)^k \alpha_k B^{m-k} \quad (5.36)$$

$$\alpha_k = -\frac{1}{k} \text{tr}(A A_{k-1}), \quad A_k = A A_{k-1} + \alpha_k I_k; \quad 1 \leq k \leq m$$

$$\alpha_0 = 1 \text{ et } A_0 = I$$

Dans le cas le plus fréquent, -donc le plus intéressant- le nombre de canaux est égale à deux (m=2), les matrices P et Q peuvent être données par les expressions suivantes:

$$P = CB - (A - \text{tr}(A)I)C$$

$$Q = (\text{tr}(A) + \text{tr}(B))B + (\det(A) - \det(B))I$$

Résumé de l'algorithme <sup>de</sup> BURG multicanaux

Entrées: Ordre du modèle AR (P)

$Y_m(k)$   $k=1 \dots N$ , échantillons d'un signal à m canaux.

Sorties:

Matrices de reflexion:  $C(1), C(2), \dots, C(P)$ :

Variances des erreurs de prédiction:  $P_p^f, P_p^b$ .

INITIALISATION

For  $k=1$  to  $N$  do  $E_0^b[k] = E_0^b[k] = Y_m(k)$

$$P_0^f = P_0^b = \frac{1}{N} \sum_{k=1}^N Y_m[k] Y_m[k],$$

Récursion

For  $p=0$  to  $P-1$  do

\*Estimation des variances des erreurs de prédiction:

$P_p^f, P_p^b$  et  $P_p^{fb}$ , équations (5.14), (5.15) et (5.16).

\*Estimation de la matrice de corrélation partielle  $\Delta_{p+1}$

Résolution de l'équation (5.32),  $\nabla_{p+1} = \Delta_{p+1}^T$

\*Estimation des matrices de reflexion  $A_{p+1}[p+1], B_{p+1}[p+1]$

$$C(p+1) = -\Delta_{p+1} (P_p^b)^{-1}, \quad D(p+1) = -\nabla_{p+1} (P_p^f)^{-1}$$

\*Actualisation des erreurs de prédiction  $E_p^f[k], E_p^b[k]$ :

For  $k = 0$  to  $p+1$  do

$$E_{p+1}^f[k] = E_p^f[k] + C(p+1)E_p^b[k-1]$$

$$E_{p+1}^b[k] = E_p^b[k-1] + D(p+1)E_p^f[k]$$

End loop on  $k$

End loop on  $p$

L'estimation des paramètres du modèle AR:  $A_p[k], B_p[k]$  se fait en utilisant la recursion de Levinson:

For  $p=0$  to  $P$  do

For  $k = 0$  to  $p+1$  do

$$A_{p+1}[k] = A_p[k] + C(p+1)B_p[p+1-k]$$

$$B_{p+1}[k] = B_p[k] + D(p+1)A_p[p+1-k]$$

End loop on  $k$

End loop on  $p$

### 5.7 Application: Cas d'un signal à deux canaux

Afin de mettre en évidence les différentes étapes de l'algorithme LWR, et à titre d'exemple, nous allons nous limiter à estimer la DSP d'un signal à deux canaux, car pratiquement, le nombre de canaux n'est pas assez élevé.

Soit  $Y(k) = (x_1(k), x_2(k))^T$ , la sortie d'un modèle AR d'ordre 1 de coefficients:

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad A_1 = \begin{bmatrix} -0.85 & +0.75 \\ -0.65 & -0.55 \end{bmatrix}$$

et excité par un bruit blanc de variance  $P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$   
Si l'on dispose de N échantillons

$$(y(1), y(2), \dots, y(N)) = ((x_1(1), x_2(1))^T, \dots, (x_1(N), x_2(N))^T)$$

la matrice de corrélation d'ordre 1 (p=1) est :

$$R_1 = \begin{bmatrix} R_{YY}(0) & R_{YY}(1) \\ R_{YY}(-1) & R_{YY}(0) \end{bmatrix}$$

$$\text{où } R_{YY}(k) = \frac{1}{N} \sum_{n=1}^{N-k} Y[n+k] Y^T[n]$$

$$R_{yy}(k) = \frac{1}{N} \begin{bmatrix} \sum_{n=1}^{N-k} x_1(n+k)x_1(n) & \sum_{n=1}^{N-k} x_1(n+k)x_2(n) \\ \sum_{n=1}^{N-k} x_2(n+k)x_1(n) & \sum_{n=1}^{N-k} x_2(n+k)x_2(n) \end{bmatrix}$$

Remarquons que les éléments de la matrice  $R_1$  sont des matrices est non plus des scalaires.

Pour mener les calculs, il est nécessaire d'initialiser l'algorithme par:  $P_o^b = P_o^f = R_{yy}(0)$ ,  $A_o(0) = B_o(0) = I_2$ .

$$\Delta_o = R_{yy}(1), \quad \nabla_o = \Delta_o^T$$

Les matrices de réflexion  $A_1[1]$  et  $B_1[1]$  sont données par:

$$A_1[1] = -\Delta_0(P_0^b)^{-1} \quad B_1[1] = -(\Delta_0)^T(P_0^f)^{-1}$$

A ce stade d'estimation, il faut actualiser les variances des erreurs de prédiction  $P_0^f, P_0^b$

$$P_1^f = (I_2 - A_1[1]B_1[1])P_0^f, \quad P_1^b = (I_2 - B_1[1]A_1[1])P_0^b$$

ainsi que la matrice  $\Delta_1$ ,  $\Delta_1 = A_0[0]R_{yy}[1]$

Les paramètres du modèle d'ordre 2 s'obtiennent en procédant comme suit:

-Calcul des matrices  $A_2[2]$  et  $B_2[2]$ :

$$A_2[2] = -\Delta_1(P_1^b)^{-1}, \quad B_2[2] = -(\Delta_1)^T(P_1^f)^{-1}$$

-Actualisation des variances  $P_0^f, P_0^b$

$$P_2^f = (I_2 - A_2[2]B_2[2])P_1^f$$

$$P_2^b = (I_2 - B_2[2]A_2[2])P_1^b$$

-Actualisation des paramètres du modèle:  $A_2[i]$  et  $B_2[i]$  en utilisant la recursion de Levinson:

i = 0 to 2 do

$$A_2[i] = A_1[i] + A_2[2]B_1[2-i]$$

$$B_2[i] = B_1[i] + B_2[2]A_1[2-i]$$

Fin de la boucle de i.

-Enfin, la nouvelle valeur de la matrice de corrélation partielle est donnée par:

$$\Delta_2 = \sum_{k=0}^1 A_1[k]R_{yy}[2-k]$$

De la même manière on peut estimer les paramètres du modèle d'ordre supérieur.

Pour un modèle AR d'ordre 1, les DSP obtenues par l'expression analytique (5.7) ainsi que le module et la phase de la fonction de cohérence données par (5.3) et (5.4), respectivement, sont représentées à la figure 5.2 sont également représentées sur cette figure, la cohérence et les DSP théoriques.

### 5.8 Calcul des DSP d'un signal multicanaux

La DSP d'un signal unidimensionnel modélisé par un modèle AR d'ordre  $p$  est aussi autorégressive caractérisée par  $2p$  pôles dans le plan complexe  $z$ . Dans le cas d'un signal à  $m$  canaux, les densités auto/inter-spectrales d'un modèle AR( $p$ ) sont représentées par des modèles ARMA avec  $m^2 p$  pôles et  $(m^2 - 1)p$  zéros dont  $p$  zéros à l'origine [47]. En effet, dans le cas simple d'un processus à deux canaux du premier ordre, si la matrice  $A[1]$  est donnée par

$$A[1] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

la matrice des DSP, dans le plan  $z$ , est donnée par [47]:

$$S(z) = \frac{1}{D(z) \cdot D(1/z)} \begin{bmatrix} 1 + b^2 + d^2 + d \cdot z^{-1} + d \cdot z & -(ab + cd + b \cdot z^{-1} + c \cdot z) \\ -(ab + cd + c \cdot z^{-1} + b \cdot z) & 1 + a^2 + c^2 + a \cdot z^{-1} + a \cdot z \end{bmatrix}$$

où  $D(z)$  est un polynôme d'ordre 2 donné par l'expression:

$$D(z) = (1 - (a+d)z^{-1} + (ad-bc)z^{-2})$$

Ainsi la DSP de chaque canal est donnée par la DSP d'un modèle ARMA, malgré que le processus vectoriel est considéré seulement comme autorégressif (AR). A partir des DSP données sous cette forme, on peut conclure que, même si on observe une cohérence élevée entre certains canaux, ceci ne traduit pas forcément une corrélation entre ces derniers. Il se peut que cette bonne cohérence n'est due qu'à un effet d'interférence entre les dits canaux.

Par ailleurs, l'évaluation de la matrice des DSP  $S(f)$  à partir des paramètres  $A_p(1) \dots A_p(p)$ , constitue une complication numérique: elle nécessite l'inversion d'une matrice complexe en chaque point de l'axe des fréquences (équation (5.7)).

## 5.9 Conclusion

Les signaux multicanaux peuvent conduire aux mêmes algorithmes et structures que les signaux à simple canal. Certes, la mise en oeuvre pratique des algorithmes multicanaux pose des problèmes numériques non négligeables. Néanmoins la complexité du modèle, le volume des données à traiter, et les limitations matérielles ne constituent pas un obstacle infranchissable, mais elles réduisent nos possibilités de choix aux méthodes les plus simples. L'étude du cas particulier de deux canaux est plus simple et s'avère plus fréquent dans la pratique. La fonction de cohérence entre deux canaux nous permet de définir dans quelles parties du spectre les deux signaux sont cohérents. Nous l'avons utilisé pour le teste d'homogénéité des séries chronologiques. Dans le chapitre suivant, on essaiera de mettre en application l'ensemble de ces techniques sur des signaux réels et concrets.

## **CHAPITRE VI**

# Applications

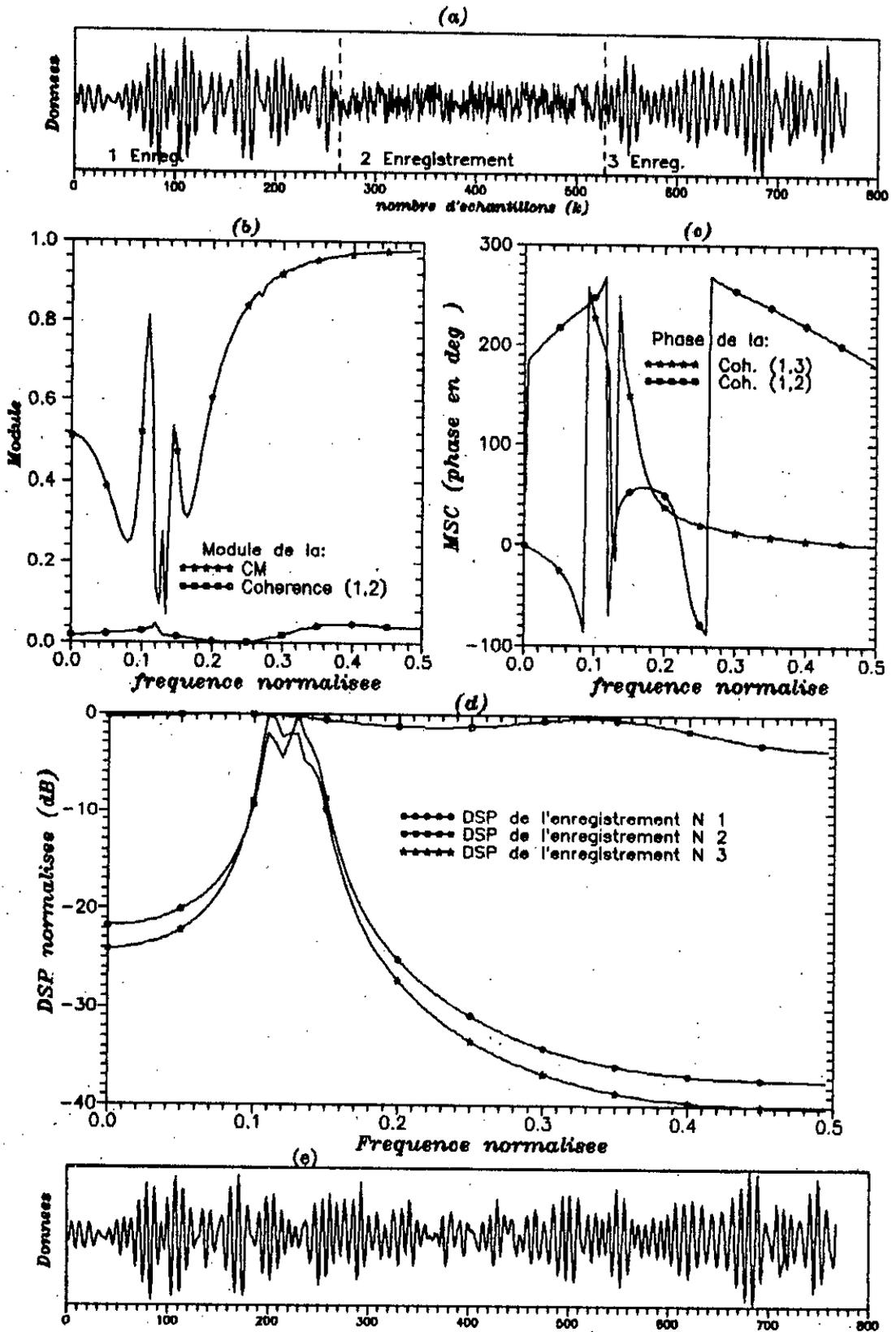


Figure 5.1 Exemple de test de la cohérence des mesures  
 (a) Données presumes non homogenes (b),(c) modules et phase  
 de la fct de cohérence, (d) DSP des 3 enregistrements 1,2 & 3  
 (e) correction des données (modélisation AR + prédiction)

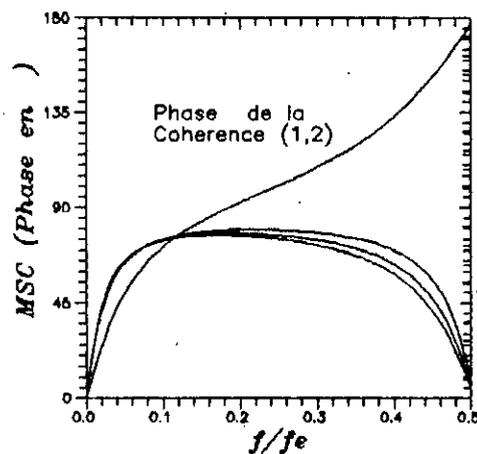
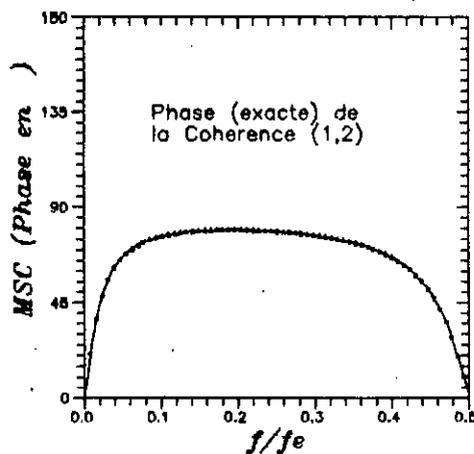
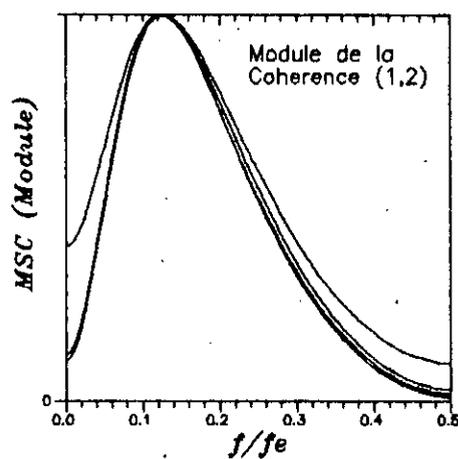
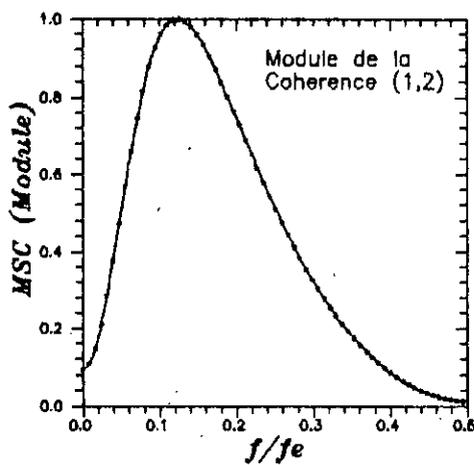
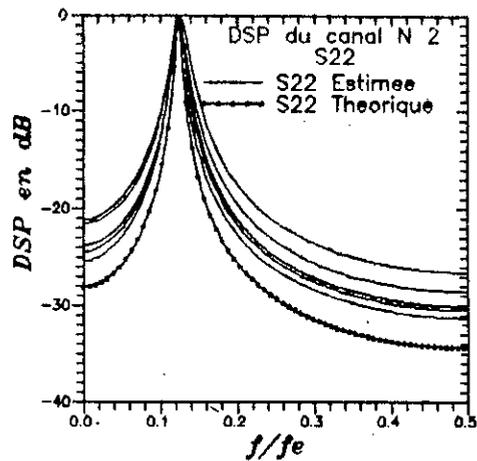
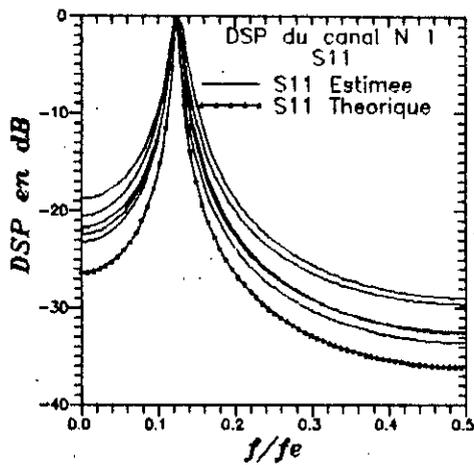


Figure 5.2 Exemple d'un modele AR(1) a deux canaux  
(a) DSP S11, (b) DSP S22, (c) & (d) Module de la coherence  
(e) & (f) Phase de la coherence.

## CHAPITRE VI

### APPLICATIONS

#### 6.1 Introduction

Les chapitres précédents de la thèse ont été consacrés à l'étude des différentes méthodes classiques et autorégressives d'estimation de la densité spectrale. Il s'agit maintenant de les appliquer à des problèmes typiques du traitement de signal. La première application concerne l'analyse des enregistrements de signaux de parole de quelques phonèmes de la langue ARABE. Pour la seconde application, on s'intéresse à l'analyse et au traitement de quelques séries chronologiques, comme les séries d'observations Météorologiques (température, précipitations, durée d'insolation, humidité ...etc), ainsi que la série d'observations du nombre de taches solaires à la surface visible du soleil [Annexe AB].

#### 6.2 Application au traitement de la parole

##### 6.2.1 Le signal de parole et sa Modélisation

Le signal vocal est l'ensemble des fluctuations de la pression de l'air engendrées et émises par l'appareil phonatoire. Il est très difficile de décrire précisément le fonctionnement acoustique de l'appareil vocal, mais généralement, on considère que la parole peut être générée par un signal source, périodique ou non, fourni par le larynx et modulé par le conduit vocal qui est formé par les cavités orales et nasales [9].

En effet, l'analyse d'un signal de parole montre qu'il est la succession de segments de bruits (sons non voisés) et d'un signal pseudo-périodique (sons voisés) (voir figure 6.1), et c'est ainsi qu'on modélise le système phonatoire par

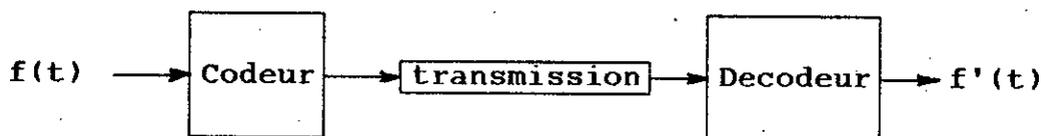
un modèle paramétrique excité, soit par un bruit blanc, soit par une suite d'impulsions périodiques. La périodicité de ces impulsions correspond à la fréquence  $F_0$  du fondamentale des ondes vocales appelée "Pitch" [9].

Le spectre de l'onde à la sortie du conduit vocal peut être décrit comme le produit du signal source par la fonction de transfert du conduit vocal. Cette dernière est le plus souvent assimilée à un filtre présentant plusieurs résonances. Ce qui favorise certaines zones fréquentielles nommées formants.

Avec cette modélisation du mécanisme de la production de la parole, le traitement de la parole (Analyse, Transmission, Synthèse, Reconnaissance, ...) est réduit à l'estimation des paramètres du modèle, du mode d'excitation (bruit ou impulsions), de la fréquence du pitch et de la position des formants. Cependant, il est essentiel de rappeler que le signal vocal n'est pas un signal stationnaire: le conduit vocal se déforme d'une façon continue. Toutefois, le signal de parole peut être considéré comme quasi-stationnaire sur des intervalles de temps de l'ordre 20 à 25 ms [9],[16]. Donc les paramètres doivent être rafraîchis toutes les 20ms environ.

### 6.2.2 Systèmes de transmission de la parole

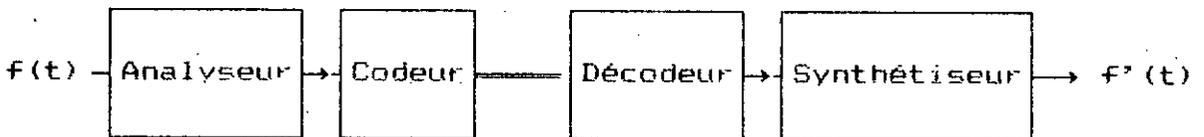
Il existe divers systèmes d'analyse et de synthèse de la parole en vue de sa transmission. On peut distinguer: -Système de transmission de l'information physique: Dans ce cas on s'intéresse à la transmission du signal de parole sous sa forme temporelle.



Système de transmission de l'information physique

L'idéale est alors de restituer un signal  $f'(t)$ , identique au signal initial  $f(t)$ . Ceci au moyen de codage le plus économique possible. Actuellement, le codage prédictif (LPC) est l'une des techniques les plus dominantes, car connaissant le mode d'excitation (bruit/impulsion) et la période du pitch, il suffit d'avoir les coefficients d'un filtre récursif (RII) d'ordre très réduit (10 à 15), généralement de type AR, pour synthétiser le signal original.

-Système de transmission de l'information perceptible: Il s'agit là de transmettre le spectre instantané d'amplitude du signal puisqu'il transporte l'essentiel de l'information perceptible. Et afin de réduire encore le débit, on ne transmet que les paramètres caractéristiques des premiers formants (la fréquence et la bande passante). Pour l'intelligibilité de la parole, les trois premiers formants sont suffisants, mais l'oreille peut percevoir une demi-douzaine dont le rôle est important dans toutes les nuances qu'il est possible d'apporter au message vocal [9].



Système de transmission de l'information perceptible

Dans ce type de système, il importe peu que le signal final  $f'(t)$  soit semblable au signal original  $f(t)$ , l'essentiel est qu'un auditeur humain ne doit pas percevoir une différence sensible entre le signal émis et le signal reçu.

### 6.2.3 Résultats et Interprétations

Il ressort du paragraphe précédent que pour la transmission de la parole, il faut passer par une analyse acoustique du signal vocal, permettant d'extraire les caractéristiques jugées les meilleures.

Nous avons effectué de nombreuses analyses spectrales sur des signaux naturels de quelques segments de phonèmes de la langue ARABE ( $/a/$ ,  $/i/$ ,  $/u/$ ,  $/b/$ ,  $/s/$ ,  $/z/$ ). Les

signaux sont convenablement échantillonnés à une fréquence de 10 khz, qui est une moyenne fréquemment utilisée en pratique. L'analyse porte sur des fenêtres de 25.6 ms prélevées dans des zones stables. Les figures 6.2 à 6.8 nous montrent, pour chaque phonème:

- (a): L'allure temporelle des 256 échantillons du signal de données,
- (b): Une estimation biaisée de la fonction d'autocorrélation,
- (c): L'évolution des différents critères de sélection de l'ordre AR,
- (d): Ainsi que la densité spectrale de puissance calculée par la méthode de BURG, et en utilisant l'ordre adéquat donné par les critères de sélection.

L'estimation du pitch est liée à la localisation des tranches voisées. La période du fondamentale,  $T_0=1/F_0$ , peut être estimée par des méthodes basées sur la fonction d'autocorrélation (recherche du 2<sup>ème</sup> maximum) [38].

Pour la recherche de l'ordre optimal du modèle AR, quatre critères ont été employés (AIC, FPE, MDL, BEP), ils nous donnent le nombre de coefficients du filtre prédicteur à transmettre. Cet même ordre est utilisé pour le calcul de la DSP du signal.

Les courbes des figures 6.2, 6.3 et 6.4 nous permettent d'observer les caractéristiques des voyelles /a/, [ , et /i/ , ainsi que celles de la semi voyelle /w/, ɔ . Après examen de ces courbes, nous remarquons que ces sons voisés présentent une structure quasi-périodique dont la fréquence fondamentale correspondant à  $F_0$  varie entre 80hz et 100hz (voix masculine). Le spectre type de ces signaux se compose de quelques formants (4 à 5), avec des amplitudes décroissantes régulièrement du graves vers l'aiguë.

Certaines consonnes, soit occlusives comme /b/, /p /, ou fricatives comme /z̃/, /ç /, présentent des caractéristiques analogues, toutefois avec une différence dans la distribution fréquentielle des formants, (voir figures 6.5 et 6.6). Pour les phonèmes /i/ / ɔ /, /w/, / ɔ /, le phénomène d'antirésonance apparaît pour certaines fréquences,

et simultanément les formants supérieurs s'affaiblissent et leur largeur augmente. Ceci est certainement dû à un couplage conduit oral/conduit nasal. Les signaux possédant de tels spectres doivent être modélisés par un modèle ARMA [6]. Les pôles correspondent aux résonances, c'est-à-dire aux formants du conduit vocal (pics du spectre); tandis que les zéros correspondent aux antiformants (vallées du spectre). Néanmoins, on voit bien que les consonnes et voyelles nasales peuvent être représentées avantageusement par un modèle AR dont l'ordre est un peu élevé (de 15 à 18).

Un autre type de consonnes dites fricatives non voisées, est présentée à la figure 6.7. Il s'agit de la consonne ( $\text{ش}$ ) /  $\text{ش}$  /. En ce qui est des données, on n'observe aucune structure périodique. Quant au spectre on remarque une nette accentuation vers les hautes fréquences.

L'ensemble des résultats obtenus sont regroupés dans le tableau 6.1 ci-après. Pour chaque phonème, on donne les différents formants (fréquence, et amplitude normalisée), ainsi que l'ordre du modèle AR adéquat.

### 6.3 Application au traitement des données météorologiques

La 2<sup>ème</sup> application envisagée concerne le traitement de séries d'observations de quelques variables météorologiques, étant donné que la météo est devenue le partenaire privilégié de plusieurs secteurs d'activité, notamment pour l'agriculture.

Les variables atmosphériques et climatiques ne sont manifestement pas distribuées au hasard. L'atmosphère est structurée dans l'espace et dans le temps; la stratification des couches atmosphériques et les périodicités dues au cycle journalier et au cycle annuel en témoignent. Cependant, en dépit de cet ordre global, on constate une variabilité très prononcée de l'atmosphère au cours du temps.

Tableau 6.1: Caractéristiques des Formants de quelques phonèmes de la langue ARABE

Phonème	Ordre AR	Formants					
		F & A	F1	F2	F3	F4	F5
/a/ ا	12	F (hz)	750	1244	2510	3030	3906
		Ampl.	40 dB	37 dB	24 dB	18 dB	07 dB
/i/ اِ	18	F (hz)	250	2000	2900	3400	3800
		Ampl.	35 dB	23 dB	32 dB	36 dB	12 dB
/u/ و	16	F (hz)	145	755	2100	3000	4100
		Ampl.	38 dB	32 dB	13 dB	11 dB	04 dB
/b/ ب	11	F (hz)	100	1000	1500	2944	4531
		Ampl.	40 dB	25 dB	12 dB	10 dB	40 dB
/s/ ش	14	F (hz)	350	2000	2800	3200	4215
		Ampl.	00 dB	04 dB	20 dB	25 dB	20 dB
/c/ ج	17	F (hz)	200	1756	2050	3000	4334
		Ampl.	10 dB	26 dB	30 dB	32 dB	10 dB

### 6.3.1 Présentation des données

Les données utilisées sont le résultat des mesures et observations météorologiques effectuées par l'Office National de la Météo (ONM) en deux stations (ALGER et ANNABA), depuis presque trois décennies. Les séries chronologiques s'échelonnent de 1967 à 1986 pour la station de ANNABA, et de 1958 à 1986 pour la station Dar-El-Beida d'ALGER. Ainsi nous disposons de 20 à 30 années d'observations journalières de plusieurs variables météorologiques (Annexe A7), à savoir: la température (°C), La pluviométrie (mm), l'insolation, et l'humidité relative. En ce qui concerne l'activité solaire, nous avons à notre disposition la série chronologique représentant l'évolution des moyennes mensuelles du nombre de taches solaires (Annexe A8) durant la période 1800-1984, ce qui correspond à 2208 échantillons (12\*184). Les fichiers de mesures disponibles sont très volumineux (356\*28=7120 échantillons, pour chaque variable), pour réduire le nombre de données à manipuler, nous avons effectuer nos traitements

sur les valeurs moyennes annuelles de l'activité solaire, moyennes mensuelles des températures et sur les cumuls mensuels des durées d'insolation et les quantités de pluies. Nous disposons ainsi de 336 échantillons pour la station d'Alger (28\*12), et de 228 échantillons (19\*12) pour la station de Annaba.

### 6.3.2 Critique des données

En général, une longue série de mesures est rarement homogène. Elle présente souvent sur certaines périodes des sur-estimations ou des sous-estimations du phénomène mesuré. En observant les différentes séries chronologiques (voir Annexe A7), les mesures de la durée d'insolation à la station d'Alger a attiré notre attention. Déjà à l'oeil nu on remarque, vers le début des années 80, une transition abrupte entre un régime d'ensoleillement quasi normal à un régime de faible durée d'insolation. Pour tester l'homogénéité de ces mesures, nous avons appliqué la méthodologie proposée en 5.2.3 sur les observations de la période 67-86 en prenant comme référence la période 73-80 (2° Enregistrement). A partir des courbes des fonctions de cohérence fig. 6.8. b, on constate une faible cohérence entre l'enregistrement N°3 (période 80-87) et l'enregistrement de référence. Par contre la cohérence entre l'enregistrement N°1 (période 67-73) et la référence est assez bonne. Ce qui confirme que les mesures de la période 80-86 sont effectivement erronées. Ceci est d'ailleurs prévisible à partir des courbes des DSP des trois enregistrements, où on observe l'absence de la composante cyclique annuelle de la DSP de l'enregistrement N°3 (voir fig. 6.8. d). Afin de rendre la série des mesures de la durée d'insolation plus proche de la réalité, on a remplacé les mauvaises observations de la période en question par les données de la prédiction linéaire à partir du modèle AR correspondant aux observations correctes 67-80. Une fois les corrections établies, les données semblent correctes et homogènes (fig. 6.8.e).

### 6.3.3 Traitements et Interprétations

#### 6.3.3.1 Analyse des séries d'observations Météorologiques

A partir des données ainsi disponibles nous avons estimé les densités spectrales de puissance des trois variables météorologiques (précipitations, température, et durée d'insolation) relatives aux deux stations, Alger et Annaba (Voir fig. 6.10 et 6.11). Mais auparavant, nous avons essayé de calculer l'ordre AR optimal en utilisant les critères usuels du choix de l'ordre AR. Les graphes de la figure 6.9 montrent bien que les variations de la température ainsi que la durée d'insolation peuvent être modélisées par un modèle AR dont l'ordre peut varier de 8 à 13. Mais en ce qui concerne les précipitations, les différents critères ne s'accordent pas sur un ordre bien fixe. Ce qui nous amène à penser que les quantités de pluies mensuelles ne peuvent être représentées par un modèle AR, et peut être un modèle ARMA conviendrait mieux. Les résultats de la modélisation sont résumés dans le tableau 6.2, où on peut trouver, l'ordre du modèle, les coefficients  $a_i$ , les coefficients de corrélation partielle  $\rho_i$ , ainsi que la variance de l'erreur de prédiction  $\sigma^2$ .

Tableau 6.2.a: Ordre optimal du modèle AR

Variable Météo.	Station d'ALGER	Station d'ANNABA
Température	08	13
Durée d'Insolation	12	13

Les DSP des phénomènes météorologiques étudiées ne paraissent pas présenter -exception faites des périodicités annuelle (1 cycle par année) et saisonnières (2 & 4 cycles/année)- des périodicités inférieures à un cycle/année, à la fois nettes et importantes. Cette constatation de fait n'implique pas obligatoirement la non existence de composantes cycliques dont la période est de quelques années, et peut être si on disposait de séries chronologiques suffisamment longues (observations sur un siècle par exemple), la détection de telles périodicités serait plus aisée.

Tableau 6.2.b: Modélisation de quelques variables météo.

Station d'ALGER

S t a t i o n	i	Température			Durée d'insolation		
		$\frac{n}{al}, n=8$	$\rho_i$	$\sigma^2$	$\frac{n}{al}, n=12$	$\rho_i$	$\sigma^2$
	0	1	-----	322.954	1	-----	66.325
	1	-1.049	-0.985	9.790	-0.344	-0.977	3.057
	2	0.060	0.543	6.905	-0.123	0.143	2.994
	3	0.160	0.032	6.898	-0.090	0.183	2.894
	4	0.044	-0.258	6.439	0.218	0.050	2.887
	5	0.055	-0.324	5.765	0.017	-0.352	2.530
	6	-0.171	-0.346	5.074	-0.029	-0.326	2.261
	7	0.214	-0.196	4.879	-0.046	-0.305	2.051
	8	0.031	-0.370	4.213	0.036	-0.262	1.910
	9	-----	-----	-----	0.021	-0.324	1.710
	10	-----	-----	-----	-0.130	-0.390	1.450
	11	-----	-----	-----	-0.163	-0.315	1.306
	12	-----	-----	-----	-0.216	-0.274	1.200

Tableau 6.2.c: Modélisation de quelques variables météo.

Station de ANNABA

S t a t i o n	i	Température			Durée d'insolation		
		$\frac{n}{al}, n=13$	$\rho_i$	$\sigma^2$	$\frac{n}{al}, n=13$	$\rho_i$	$\sigma^2$
	0	1	-----	240.055	1	-----	60.359
	1	-0.9042	-0.984	7.718	-0.5075	-0.976	2.902
	2	0.2407	0.665	4.307	-0.0096	0.270	2.691
	3	-0.0191	-0.153	4.207	-0.0020	0.120	2.652
	4	0.0127	-0.278	3.882	0.1354	-0.054	2.644
	5	-0.0042	-0.313	3.501	-0.0354	-0.298	2.410
	6	-0.0406	-0.290	3.206	-0.0202	-0.262	2.244
	7	0.0796	-0.293	2.931	0.0275	-0.249	2.105
	8	-0.0477	-0.375	2.519	0.0266	-0.310	1.904
	9	0.1288	-0.361	2.191	-0.0981	-0.367	1.648
	10	-0.3084	-0.454	1.740	0.0005	-0.290	1.510
	11	0.0394	-0.182	1.683	-0.1732	-0.368	1.305
	12	-0.2215	-0.229	1.595	-0.2297	-0.295	1.191
	13	-0.0908	0.034	1.593	-0.1306	-0.122	1.173

### 6.3.3.2 Analyse de la série de mesures de l'activité solaire

L'étude de l'activité du soleil qui se manifeste par le changement du nombre de taches solaires à sa surface visible, a certainement de l'importance, puisque c'est à la chaleur de cet astre que nous devons toutes les manifestations de la vie terrestre.

La densité spectrale de l'activité solaire (A.S.) estimée en se servant des moyennes annuelles des observations durant la période 1800-1984, présente deux pics importants à 0.0925 et 0.175 cycle/année, correspondant, respectivement, à des périodes de 10.81 et 5.7 années. Une estimation de la DSP de l' A.S. à partir d'une série d'observations relativement courte, en l'occurrence des moyennes mensuelles de la période 58-84 -qui ne couvre que deux cycles de 11 ans (312 échantillons)-, met en évidence la difficulté d'estimation des deux composantes cycliques citées auparavant. La période de 10.8 années correspondant à 0.0077 cycle/mois, se cache certainement dans la forte composante continue (voir figure 6.12). Ce résultat consolide l'hypothèse d'existence de composantes cycliques dont la période est de quelques années pour les variables météorologiques, notamment la température et la durée d'insolation.

Les différents critères de sélection de l'ordre nous indique que l'ordre optimal du modèle AR représentant l'activité solaire est égale à 10. Le tableau 6.3 nous donne les différents paramètres de la modélisation de l'activité solaire.

L'utilisation de ce modèle a permis d'établir une bonne prédiction des cycles 20 & 21 de l'activités solaire, voir fig 6.14. A partir de ce même modèle, on prévoit que le cycles 22 et 23 seront aussi intenses que le cycle 21, toutefois sans surpasser le cycle 19, le plus actif jamais observé.

Tableau 6.3 Modélisation de l'activité solaire

i	Coeff. $a_i^n$ , $n=10$	Coeff: $\rho_i$	Variance $\sigma^2$
0	1	-----	3770.045
1	-1.1399968	-0.9292001	514.940
2	0.2934803	0.59182	334.582
3	0.2379534	-0.0722647	332.834
4	-0.1746125	-0.2912226	304.606
5	0.0640767	-0.1443268	298.261
6	-0.033992	-0.2581893	278.379
7	-0.0131508	-0.2803497	256.500
8	0.0678879	-0.3048863	232.660
9	-0.1728486	-0.3125753	209.925
10	-0.1614434	-0.1072225	207.512

### 6.3.3.3 Relation entre température et activité solaire

Afin d'étudier l'effet de l'activité solaire sur la température à Alger, on a construit un signal a deux canaux (Température & Activité Solaire), et on est parti à la recherche d'éventuel corrélation entre les deux variables (fig. 6.13). Une grande cohérence (0.78) est observée à la fréquence de 0.015 cycle/mois (Période de 5.55 années). Ainsi la demi-période de 5.55 du cycle solaire semble avoir des effets sur les variations de la température. Le module de la cohérence(T,A) présente aussi un pic à la fréquence de 0.086 cycle/mois, ce qui revient à dire q'une composante cyclique annuelle est présente dans les deux densités spectrales S11 et S22. Or une analyse simple canal de l'activité solaire nous n'indique en aucun cas la présence de cette composante. La cohérence observée à cette fréquence ne s'explique que par l'effet d'interférence<sup>(1)</sup> entre les deux canaux "feed-across artifact" [47].

(1) Interférence entre canaux: Cet effet se manifeste par la présence d'une composante fréquentielle dans deux canaux, tandis qu'en réalité, elle n'est présente que dans un seul.

(Formants, Composantes cycliques) cachées dans les enregistrements de données, ainsi que la procédure à suivre afin d'homogénéiser et corriger certaines anomalies des données, et prédire si nécessaire le comportement du phénomène. Cette méthodologie peut être résumée en 4 étapes:

- 1-Prétraitement des données (filtrage, calcul des valeurs moyennes, teste d'homogénéité et de cohérence ...etc).
- 2-Détermination de l'ordre optimal du modèle & modélisation.
- 3-Estimation des densités spectrales de puissance.
- 4-Interprétations et utilisation des résultats (pour la prédiction par exemple).

Enfin, on tient à signaler qu'en principe, rien ne s'oppose à appliquer des traitements analogues aux signaux provenant d'autres systèmes, tels que les signaux ECG, EEG, signaux et observations sismiques ...etc. Nous étions prêts à effectuer ces traitements si seulement les enregistrements et les données nécessaires étaient disponibles.

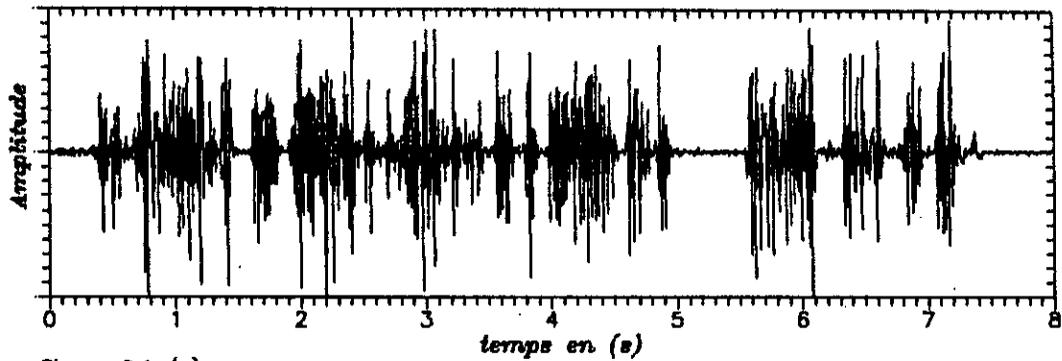


Figure 6.1 (a) Signal de parole representant l'expression suivante:  
 'كيفما وجدت إلا مسور فعندنا بساطة ولا تهاش فإن الهاس من الشيطان'

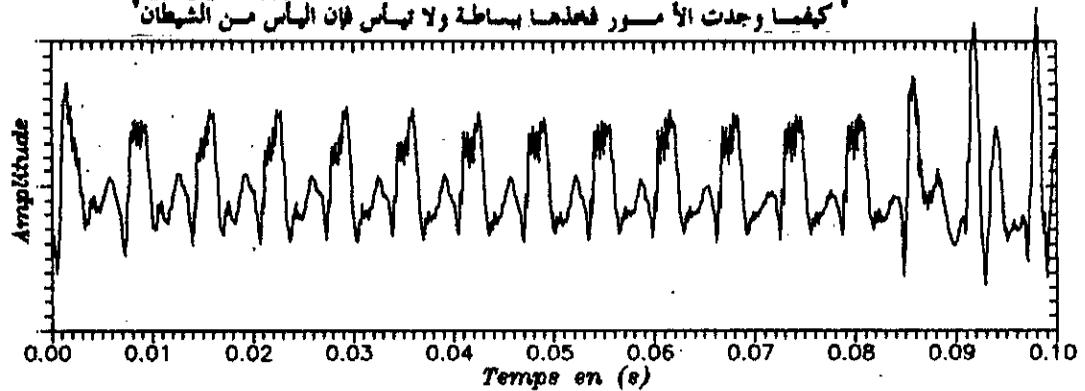


Figure 6.1 (b) Signal de parole: Exemple de tranche voisée

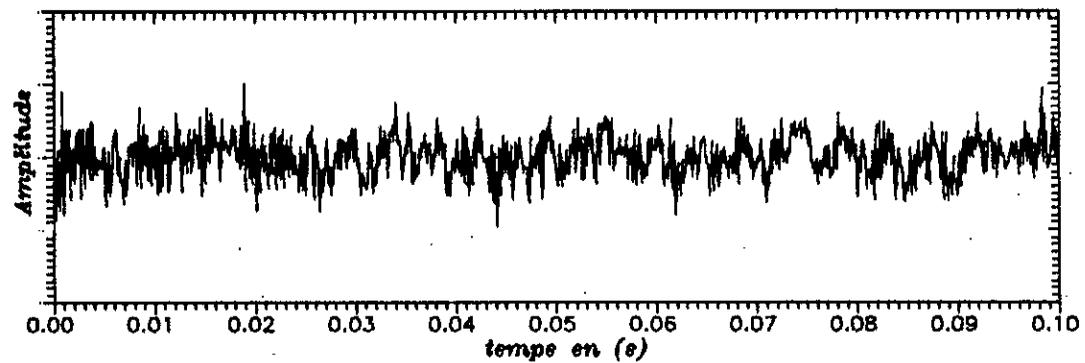


Figure 6.1 (c) Signal de parole: Exemple de tranche non voisée

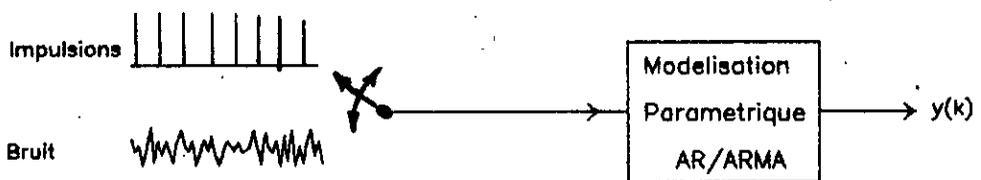
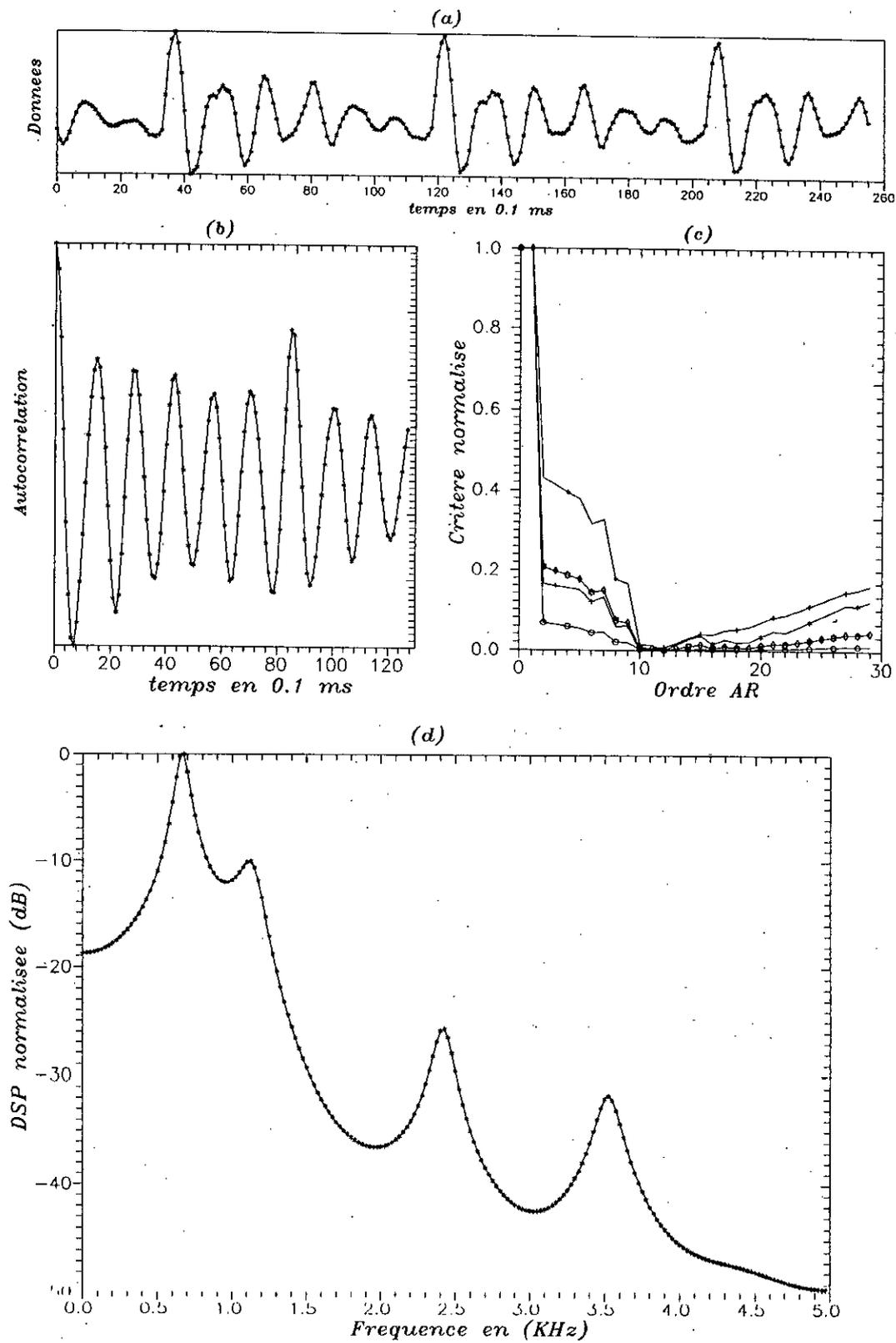
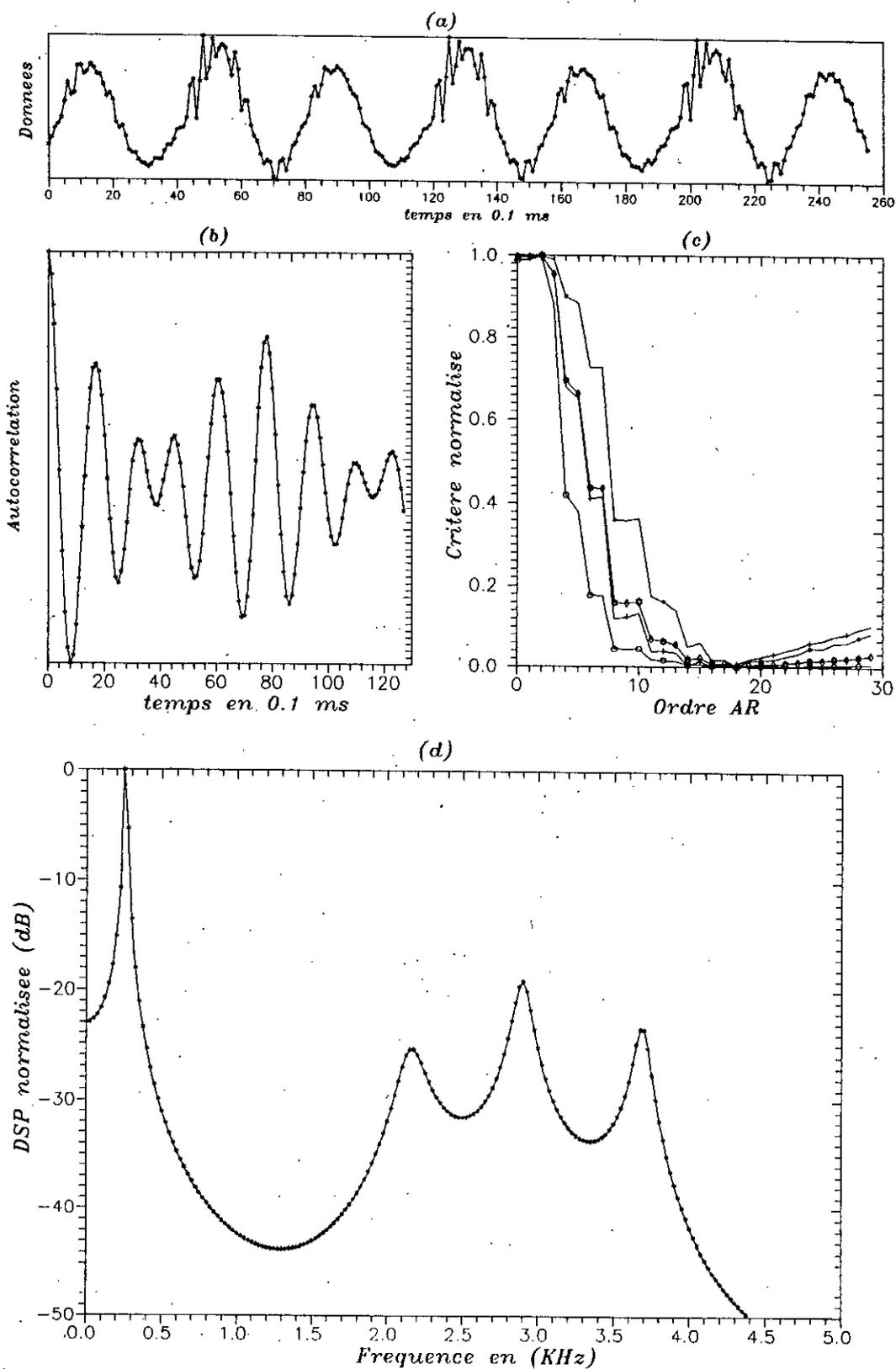


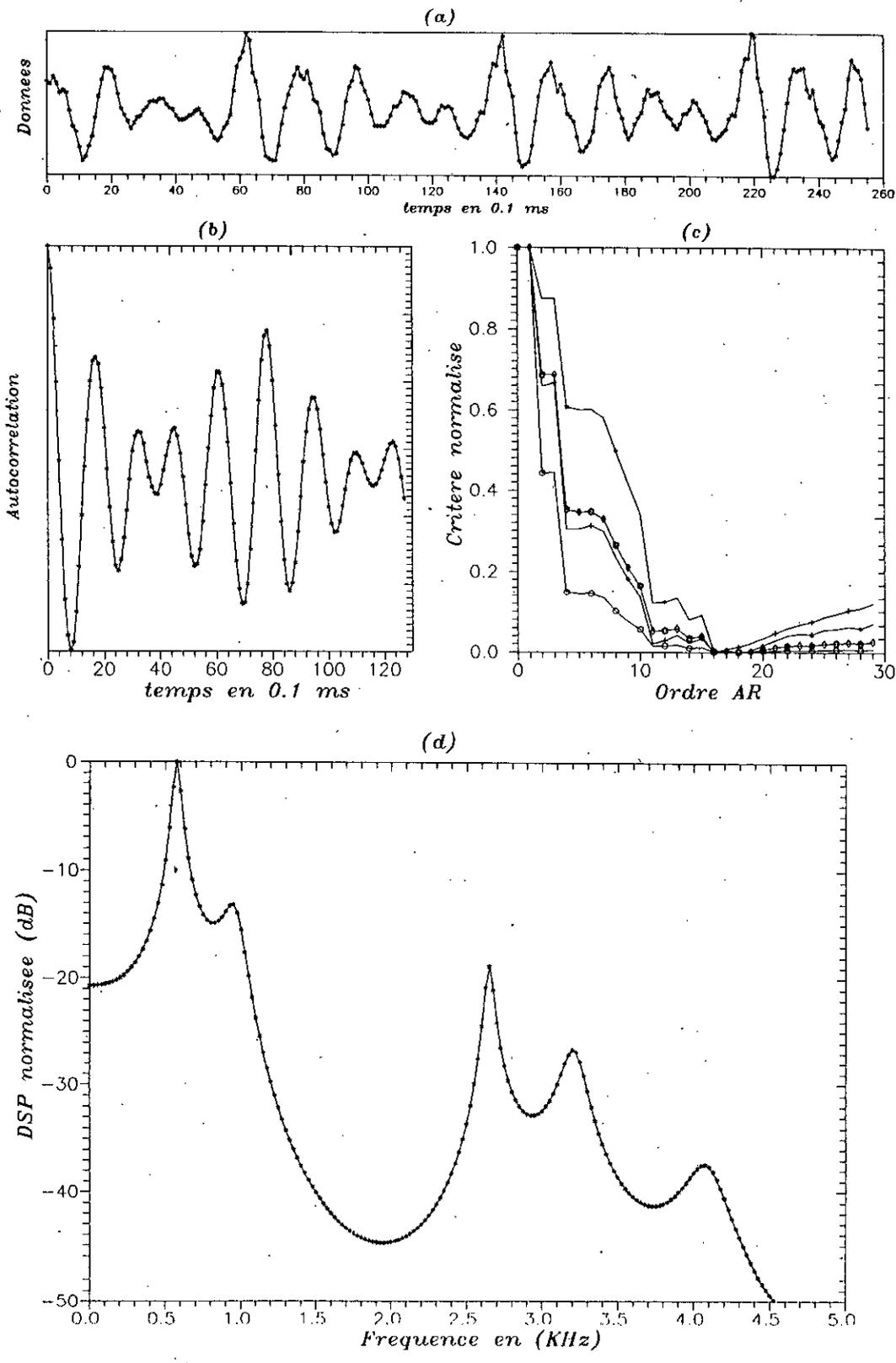
Figure 6.1 (d) Modélisation paramétrique du signal de parole



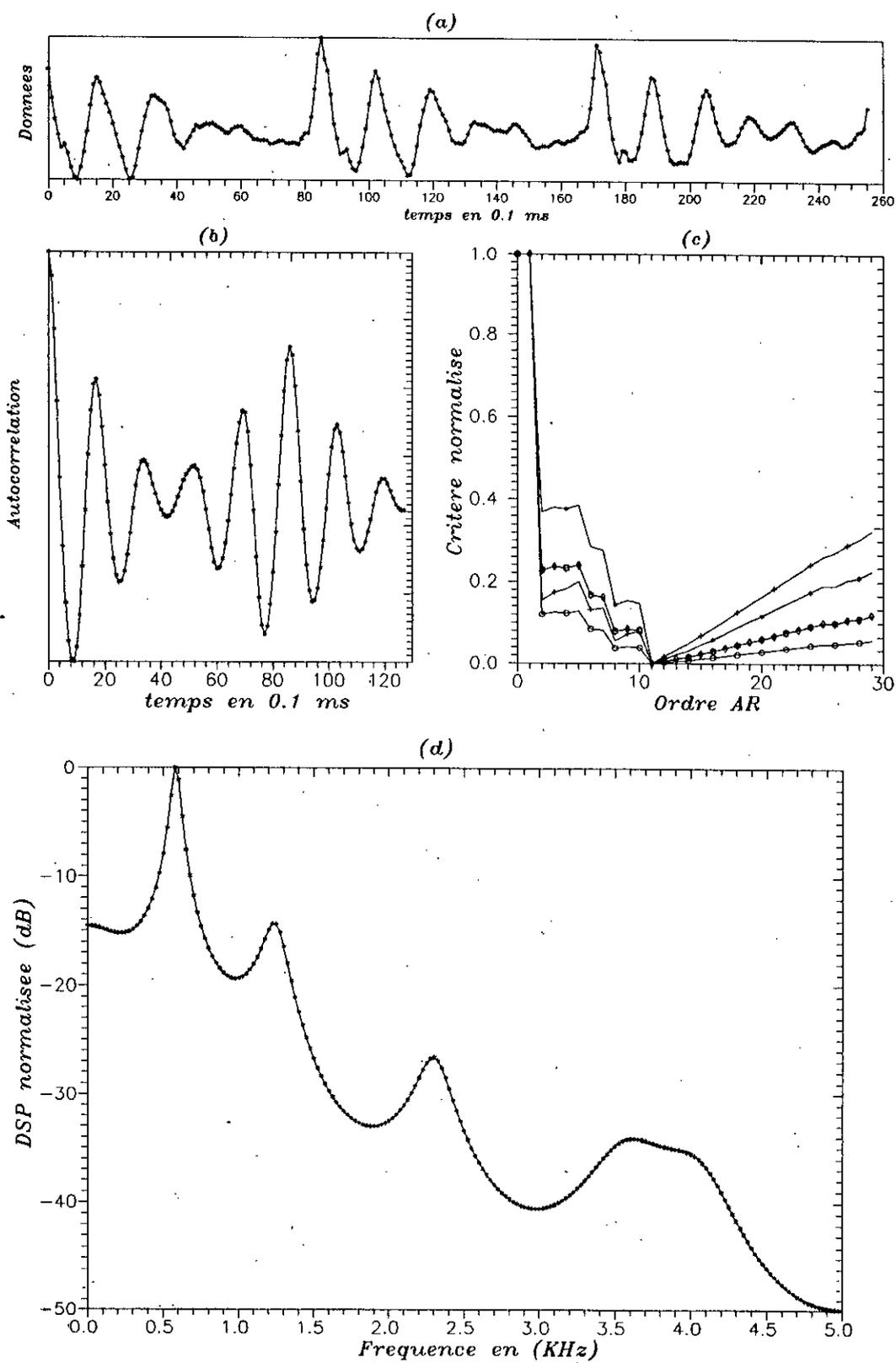
**Figure 6.2** Phoneme < A > < ʔ > de la langue ARABE  
 (a) Donnees, (b) Autocorrelation, (c) evolution des criteres AR  
 (d) densite spectrale de puissance



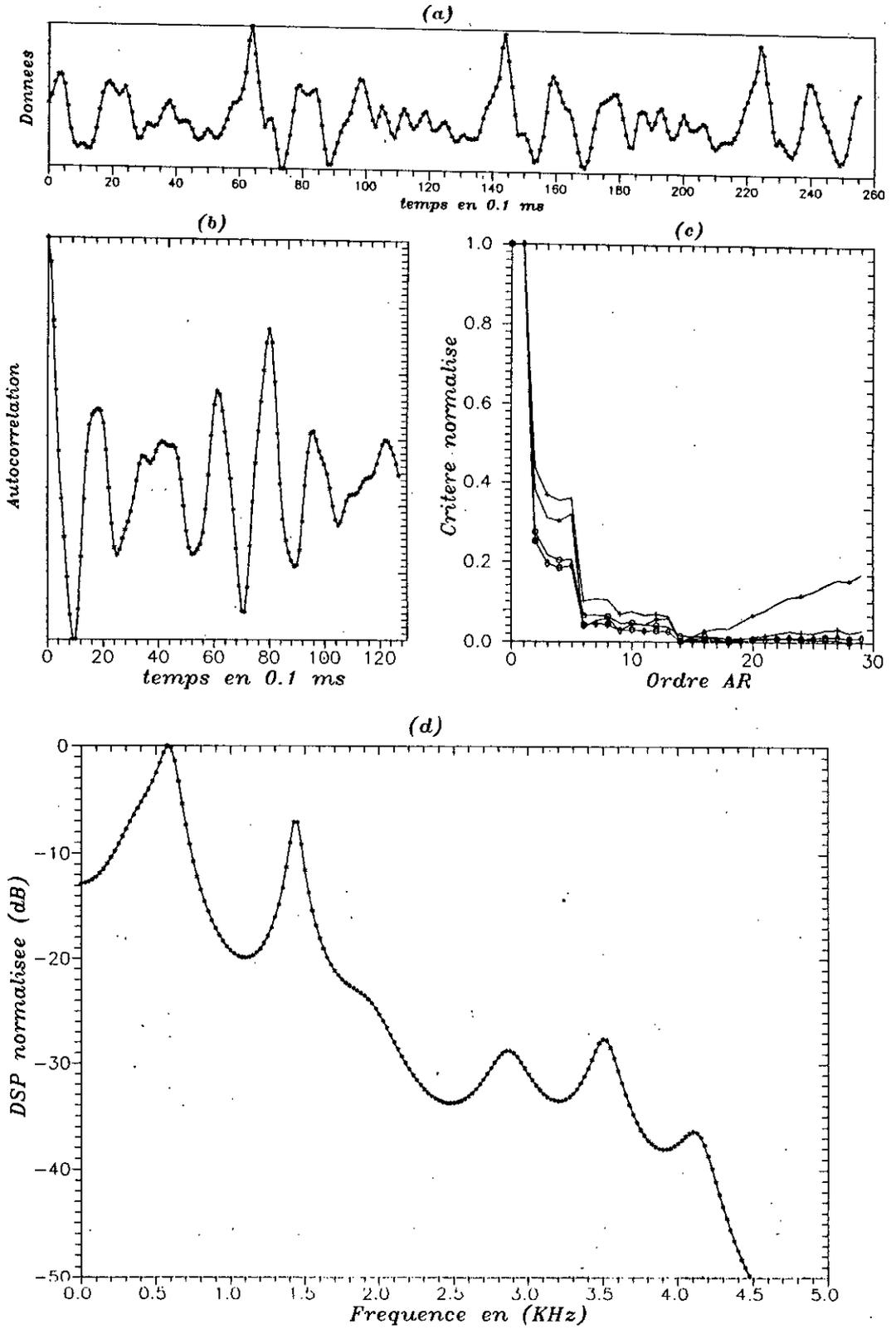
**Figure 6.3** Phoneme [l] de la langue ARABE  
 (a) Donnees, (b) Autocorrelation, (c) evolution des criteres AR  
 (d) densite spectrale de puissance



**Figure 6.4** Phoneme { WA } { و } de la langue ARABE  
 (a) Donnees, (b) Autocorrelation, (c) evolution des criteres AR  
 (d) densite spectrale de puissance



**Figure 6.5** Phoneme  $\langle BA \rangle$  (ب) de la langue ARABE  
 (a) Données, (b) Autocorrelation, (c) evolution des criteres AR  
 (d) densite spectrale de puissance



**Figure 6.6** Phoneme  $\langle \text{Djim} \rangle$   $\langle \text{ج} \rangle$  de la langue ARABE  
 (a) Donnees, (b) Autocorrelation, (c) evolution des criteres AR  
 (d) densite spectrale de puissance

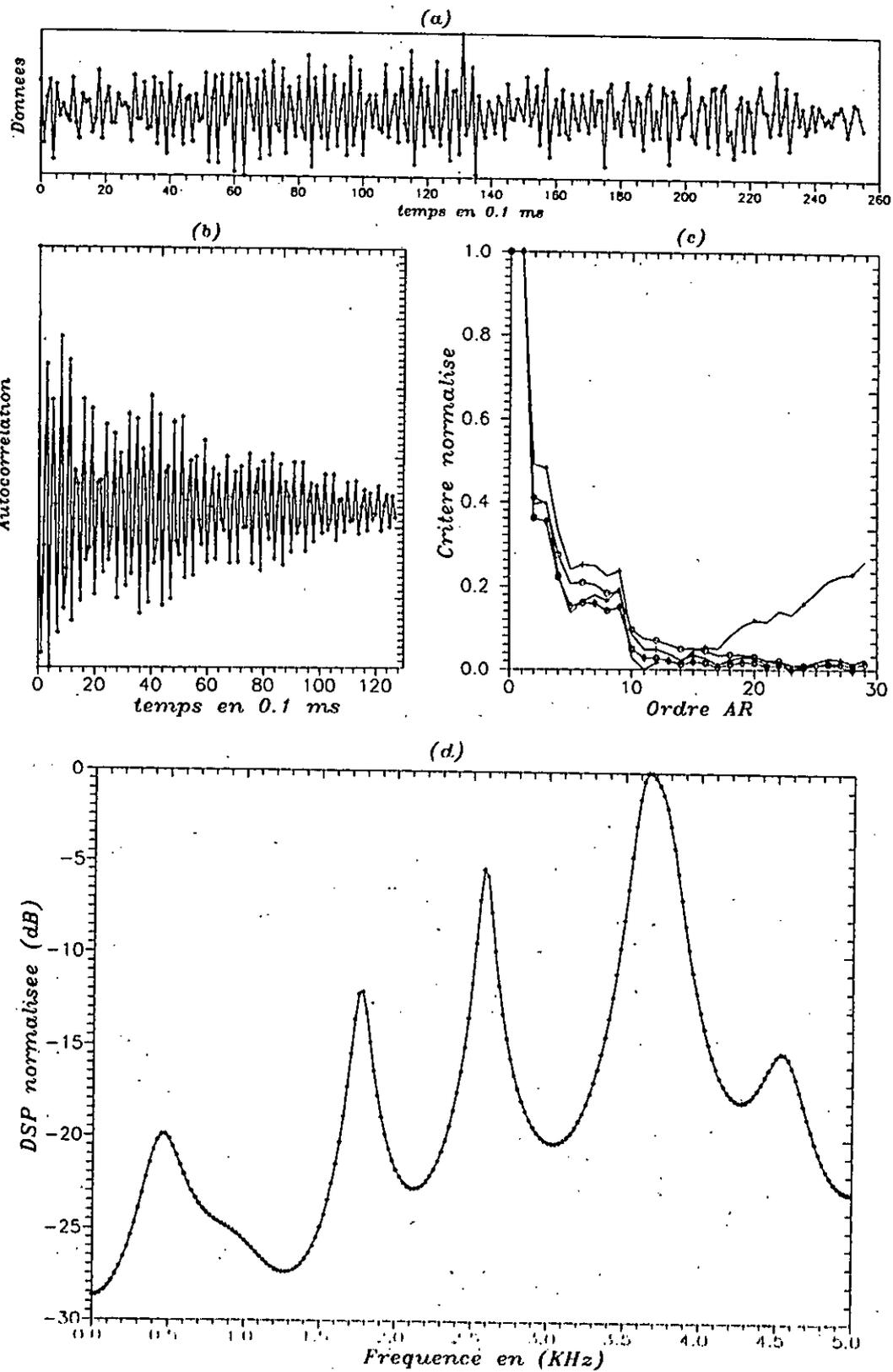


Figure 6.7 Phoneme  $\langle \text{ش} \rangle$  de la langue ARABE  
 (a) Donnees, (b) Autocorrelation, (c) evolution des criteres AR  
 (d) densite spectrale de puissance

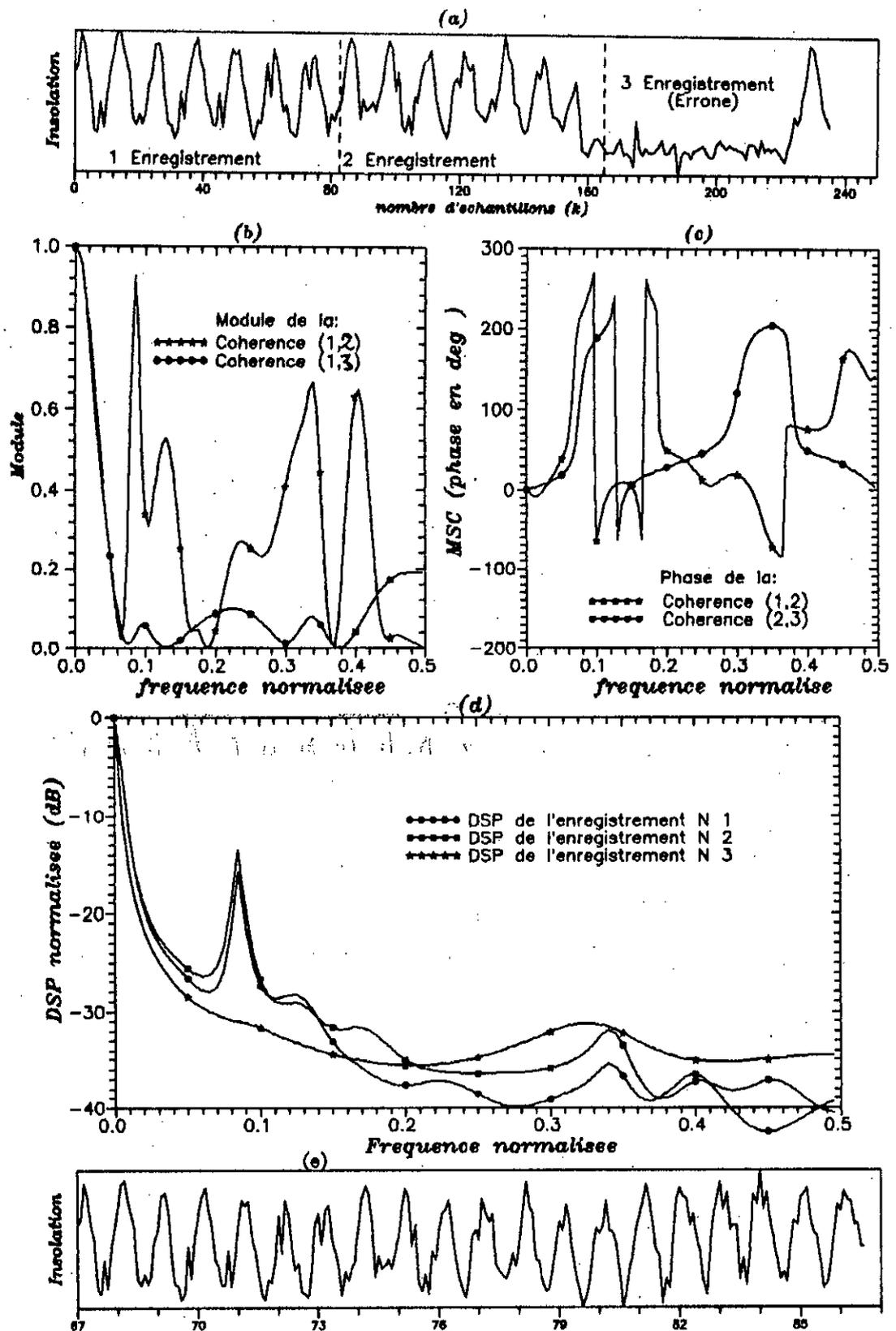
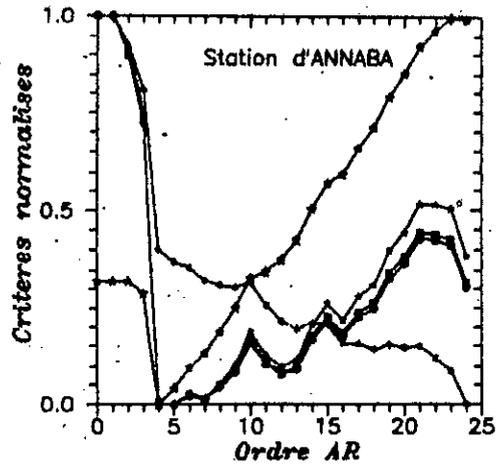
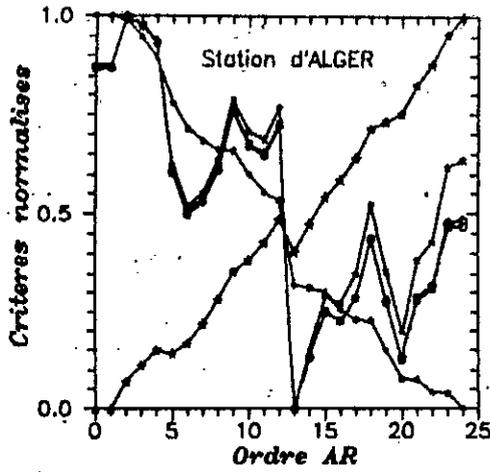
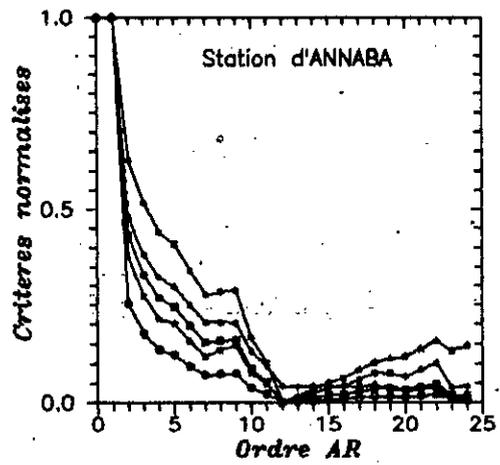
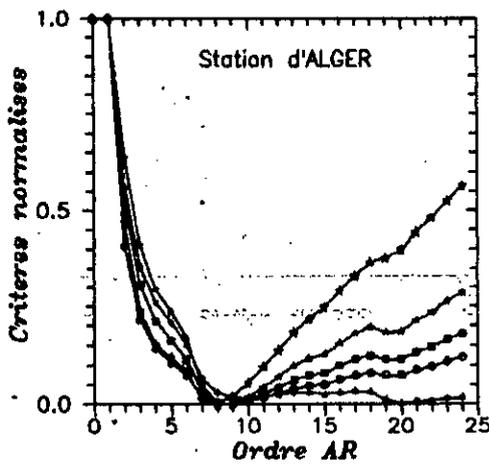


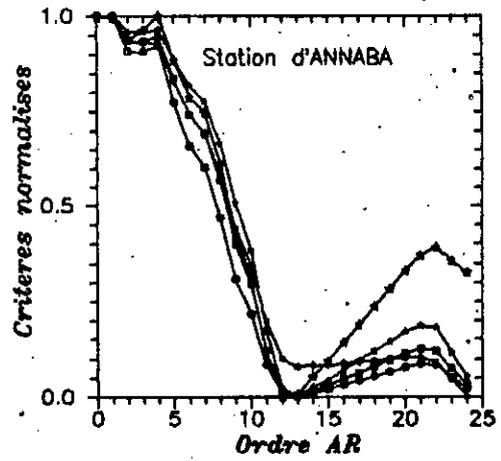
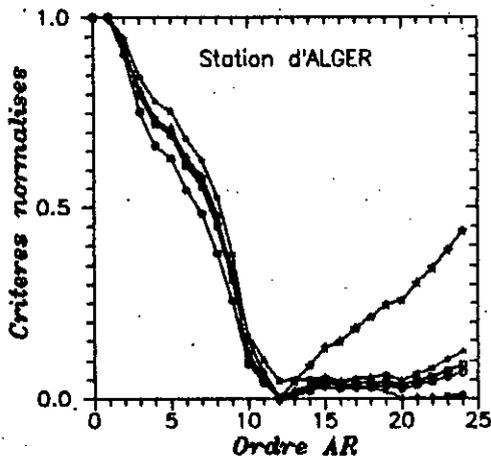
Figure 6.8 Test d'homogeneite des mesures d'insolation a ALGER  
 (a) Observations mensuelles (67-86) (b),(c) modules et phase de la fct de coherence, (d) DSP des 3 enregistrements 1,2 & 3  
 (e) correction des mesures (modelisation AR + prediction)



Choix de l'ordre AR optimal pour les variations des precepitation dans les stations d'ALGER et d' ANNABA

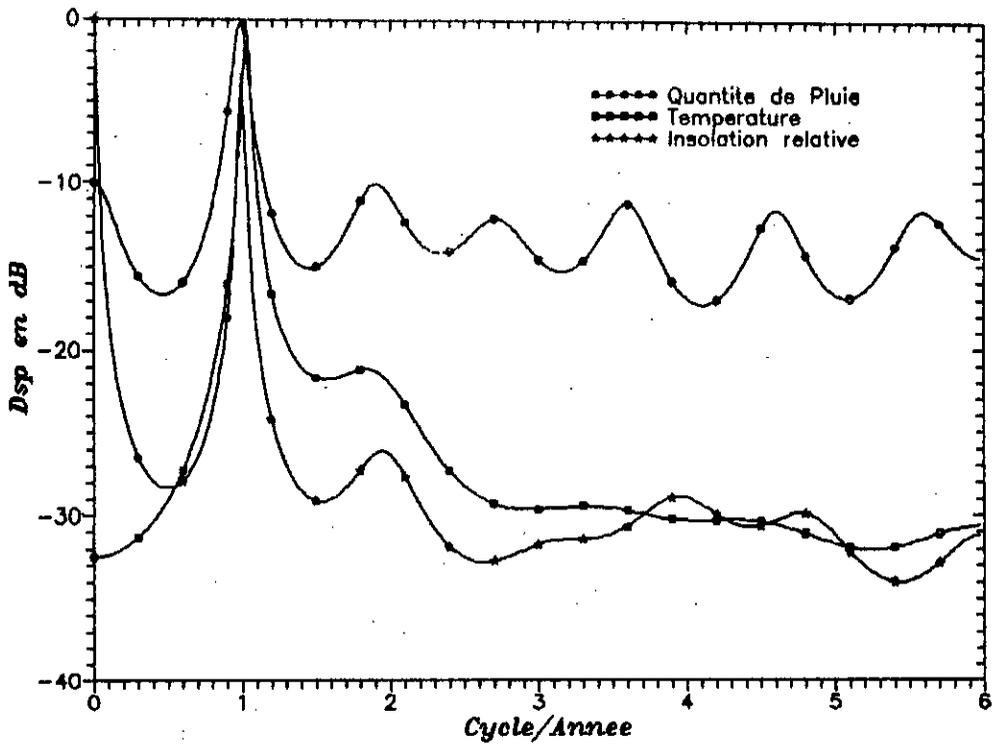


Choix de l'ordre AR optimal pour les variations de la Temperatur dans les stations d'ALGER et d' ANNABA

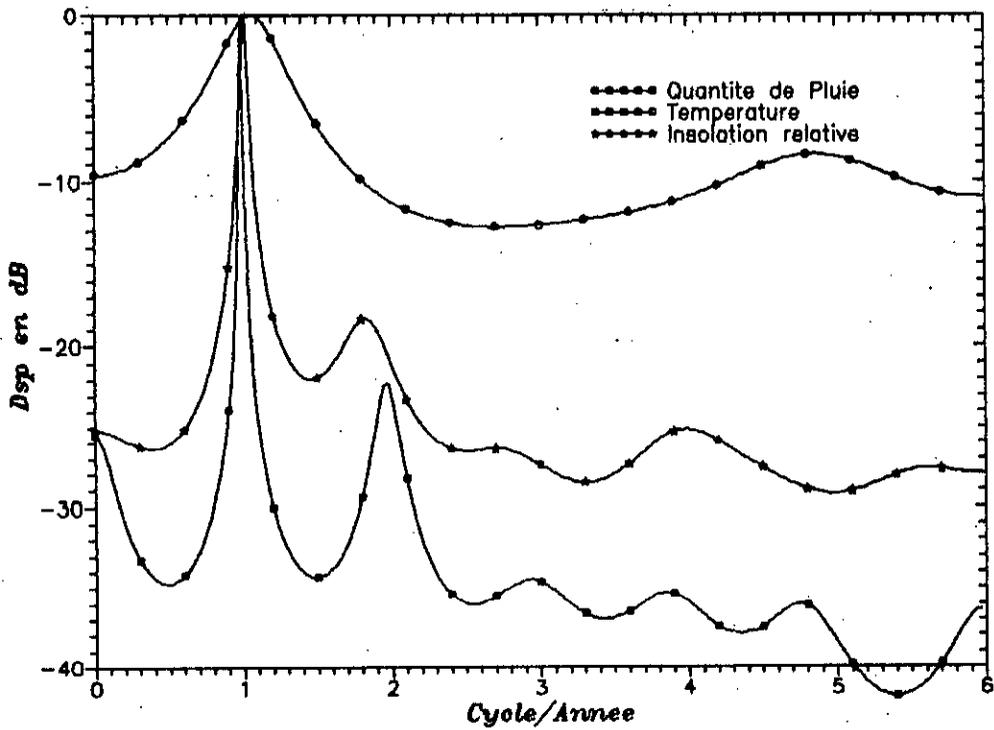


Choix de l'ordre AR optimal pour les variations de l'insolation dans les stations d'ALGER et d' ANNABA

**Figure 6.9 : Determination de l'ordre AR**



**Figure 6.10**  
*Densité spectrale de puissance de quelques variables climatiques observées à la station d'ALGER.*



**Figure 6.11**  
*Densité spectrale de puissance de quelques variables climatiques observées à la station d'ANNABA.*

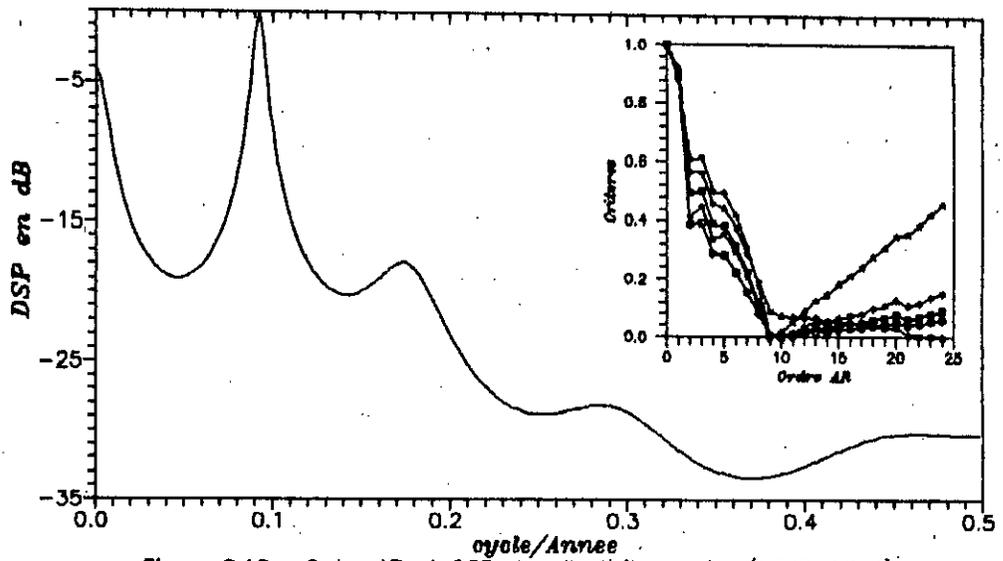
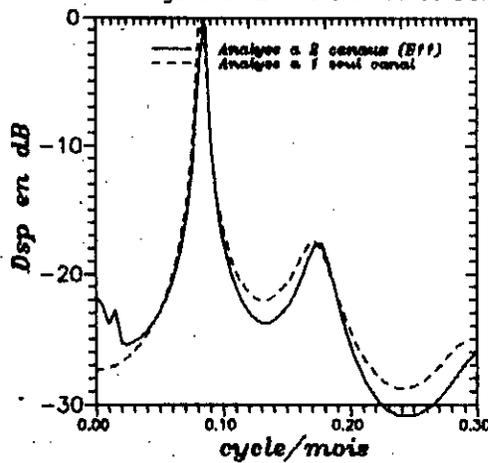
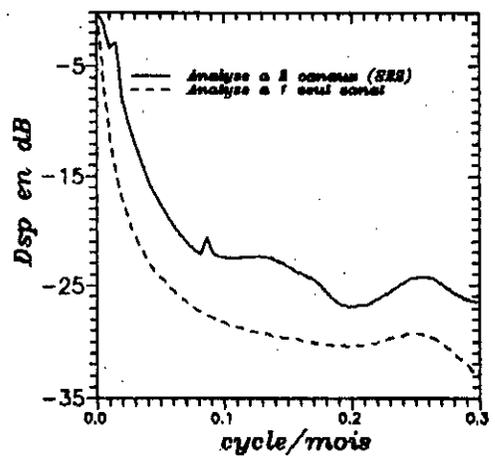


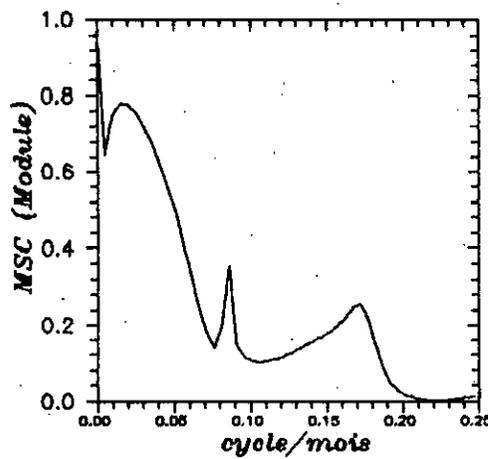
Figure 6.12 : Ordre AR et DSP de l'activité solaire (1800-1984)



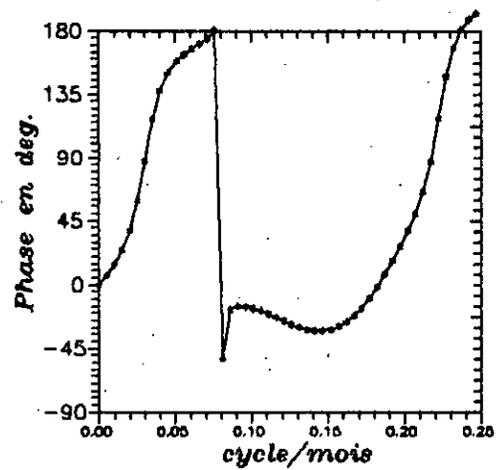
(a): DSP S11 (Temperature a ALGER)



(b): DSP S22 de l'activité solaire



(c): Module de la cohérence (T,A)



(d): Phase de la cohérence (T,A)

Figure 6.13 : Analyse spectrale a deux canaux (Temperature & Activite solaire)

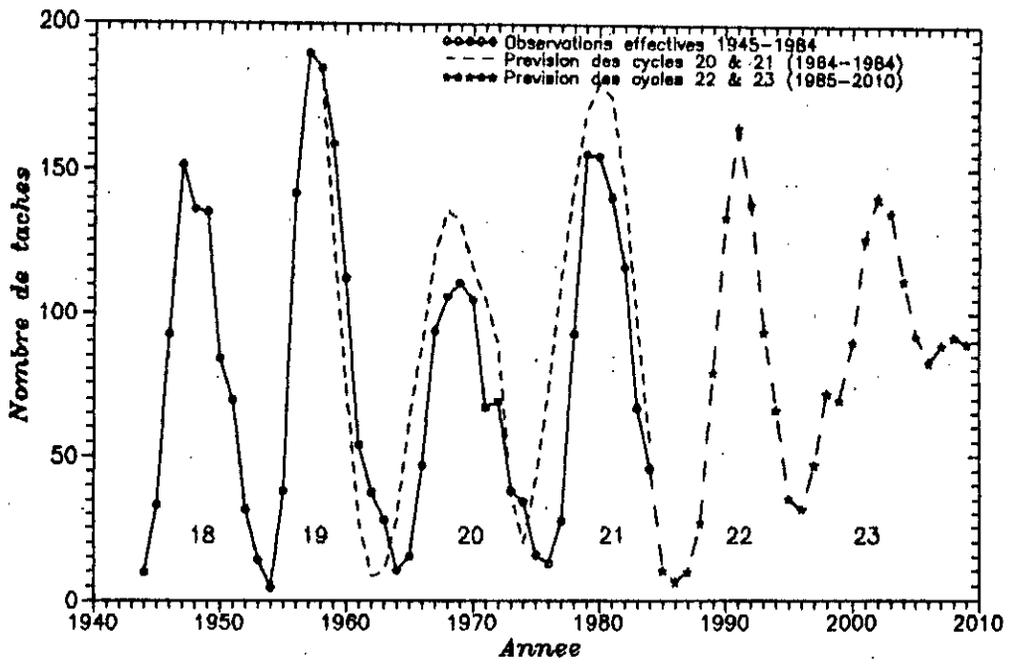


Figure 6.14 Prevision de quelques cycles de l'activite solaire.

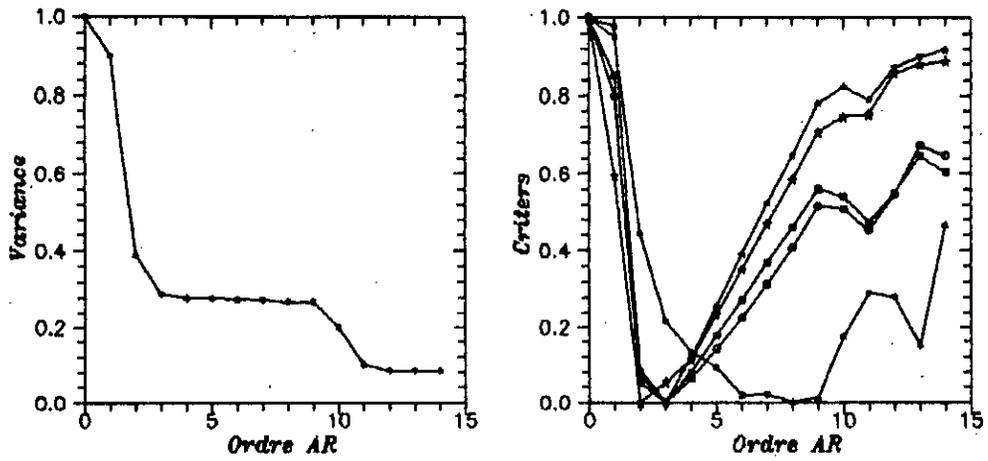


Figure 6.15 Evolution de la Variance de l'erreur, et des differents criteres de selection de l'ordre AR pour les Quantites de pluies cumulees (Sep. Dec) a la station d'ALGER

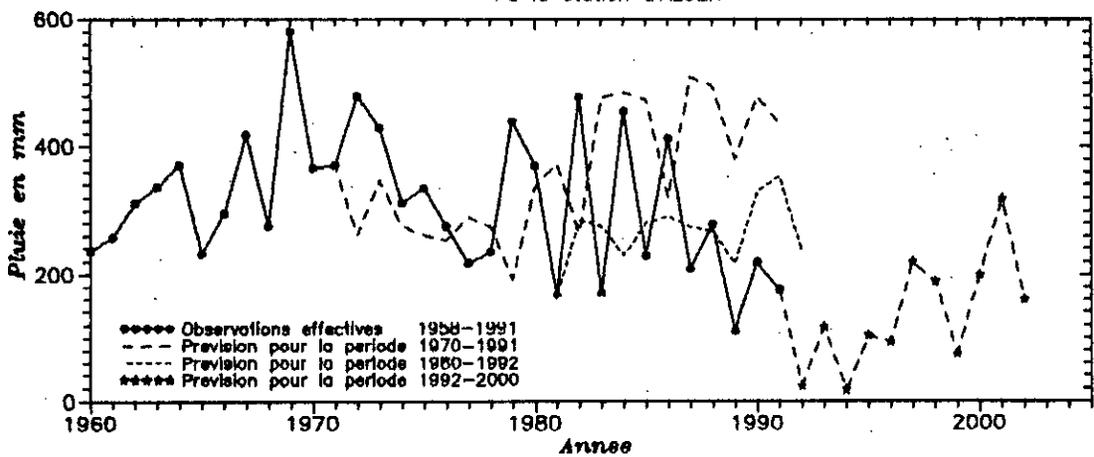


Figure 6.16 Quantites de pluies cumulees de septembre a decembre (Station Dar El-Beida a ALGER)

# CONCLUSIONS

## CONCLUSIONS

Dans ce travail de recherche, les fondements des algorithmes d'estimation de la DSP ont été étudiés de manière détaillée. Ce qui nous a permis de prendre conscience des atouts et faiblesses des différentes méthodes en tant que moyen de traitement de l'information.

Ainsi, en ce qui concerne la méthode du périodogramme, le plus simple des estimateurs classiques possède une variance qui ne peut être diminuée en agissant sur les paramètres de l'estimation. La segmentation des données en plusieurs sections (périodogramme moyenné) permet de réduire la variance, mais comme la résolution fréquentielle dépend du nombre de points dans chaque section, si on ne peut plus augmenter le nombre total d'échantillons  $N$ , la diminution de la variance ne peut être qu'au prix de l'augmentation du biais. L'estimateur adouci fait intervenir une fonction fenêtre, dont les paramètres et la forme influencent considérablement l'estimation. C'est ainsi qu'on peut améliorer la variance en choisissant une fenêtre qui présente une bonne atténuation des lobes secondaires. Mais ceci affecte bien sûr la résolution.

Avec l'estimateur modifié: une combinaison des deux estimateurs moyenné et adouci, on dispose de deux moyens pour améliorer les résultats de l'estimation: jouer sur le nombre de sections du signal, et utiliser d'autres fenêtres de pondération plus performantes que la fenêtre rectangulaire, pour cette raison, cet estimateur est généralement, le plus utilisé.

Par la méthode du corrélogramme (méthode de Blackman-Tukey), on obtient la véritable DSP convoluée par une fonction fenêtre spectrale. Cet effet de filtrage est d'autant plus faible que la durée de corrélation est plus grande. Cependant cette dernière est limitée à une fraction du temps d'observation du signal. La convolution avec la fenêtre spectrale va

introduire des lobes secondaires -parfois négatifs<sup>(1)</sup> - plus importants que ceux introduits dans le cas du périodogramme.

Les imperfections des méthodes classiques (Périodogramme, Corrélogramme) sont liées essentiellement à l'hypothèse faite sur le signal ou sa fonction d'autocorrélation en considérant que ce(tte) dernier(ère) est nul(le) en dehors de la période d'observation. De ceci résulte une mauvaise résolution, surtout lorsque la durée d'observation est faible. Les méthodes dites paramétriques essaient de palier à cette difficulté en postulant un modèle particulier du signal: Modèle AR pour la méthode de BURG, modèle d'un bruit blanc additionné à des fréquences pures pour le modèle de PISARENKO, somme de sinus-oïdes amorties pour la méthode de PRONY...etc. Conséquence de cette approche: la résolution spectrale est meilleure comparée à la résolution des méthodes classiques.

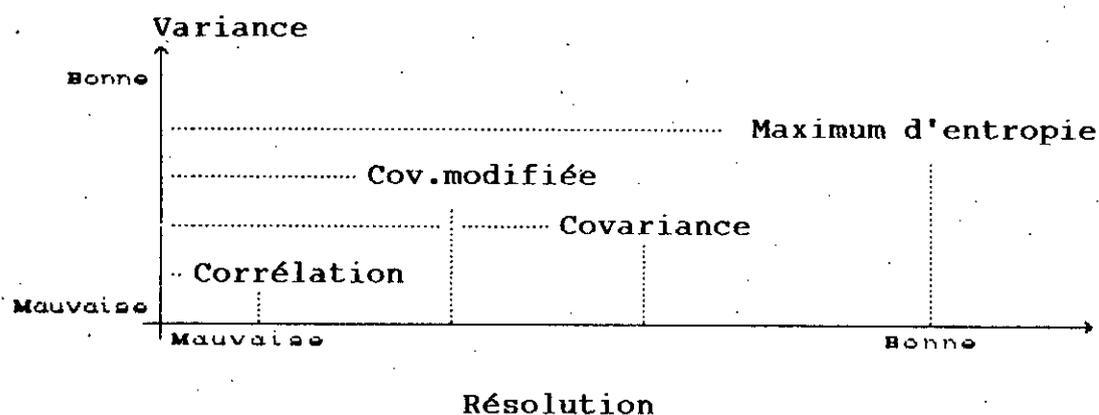
L'ensemble de ces méthodes paramétriques a été présenté, et on s'est particulièrement attaché à étudier les méthodes basées sur une modélisation Autorégressives (AR). Les coefficients du modèle AR s'obtiennent par minimisation de l'erreur quadratique totale de prédiction (directe et/ou rétrograde). Plusieurs méthodes ont été proposées pour l'estimation des paramètres du modèle. La méthode de l'autocorrélation conduit aux équations de YULE-WALKER, pour lesquelles un algorithme dû à LEVINSON et DURBIN apporte une solution efficace.

Les méthodes de covariance, covariance modifiée et fenêtrage sont basées sur le même principe, et donnent des résultats meilleurs du point de vue variance. Mais le problème principal de ces méthodes est la stabilité du modèle (non garantie). Avec la méthode du maximum d'entropie les coefficients de la modélisation AR sont estimés directement à partir des données et la stabilité du modèle est assurée du

---

(1) Seules quelques fenêtres de pondération temporelles utilisées avec la fonction de corrélation ne présentent pas de lobes secondaires négatifs (BARTLETT, PARZEN ...) [49].

fait que le module des coefficients de réflexion est inférieur à l'unité. Mais malheureusement pour des signaux à fréquences pures cette méthode présente deux faiblesses: le dédoublement des raies et la sensibilité à la phase initiale du signal, ce qui est très gênant pour la détection des composantes fréquentielles noyées dans du bruit. Du point de vue variance et résolution spectrale, les différentes méthodes étudiées se répartissent de la façon suivante, de la moins bonne à la meilleure:



Un problème important est celui du choix de l'ordre du modèle. Les critères usuels ( FPE, AIC, MDL) ont tendances à sous-estimer l'ordre. Nous avons proposé un nouveau critère basé sur le test de la blancheur de l'erreur de prédiction en vérifiant la similitude de la fonction d'autocorrélation de cette erreur avec l'impulsion de Dirac, et afin de présenter un intérêt pratique les éléments de la corrélation sont estimés de manière récursive en utilisant la structure en treillis. Les résultats obtenus par simulation se révèlent tout à fait satisfaisants.

Après l'analyse spectrale scalaire, on a essayé d'aborder le problème de l'analyse spectrale multicanaux. Il s'agit dans la plupart des cas d'extensions -pas toujours évidentes- des méthodes conçues pour un seul canal aux signaux multicanaux. L'algorithme de LEVINSON-WIGGINS-ROBINSON (LWR),

(extension de l'algorithme de LEVINSON) et l'algorithme de NUTTALL-STRAND (version multicanaux de l'algorithme de BURG) ont été étudiés. Un intérêt particulier a été accordé à la fonction de cohérence et une méthodologie de test de la cohérence des enregistrements d'un signal stationnaire a été proposée.

A l'issue de l'étude des différentes méthodes, un logiciel interactif, supporté par les IBM PC/XT/AT (et compatibles) a été développé avec l'objectif de mettre les différentes techniques à la portée de tous ceux concernés par le problème de l'estimation de la densité spectrale de puissance des signaux aléatoires et stationnaires.

Deux exemples d'applications ont été considérés:

- 1) L'analyse des enregistrements de quelques segments de phonèmes de la langue ARABE. Les résultats confirment que le modèle AR est efficace pour la représentation des voyelles et semi-voyelles. Mais il semble limité pour la modélisation des consonnes nasales et non voisées. Néanmoins, ces dernières peuvent, avantageusement être représentées par un modèle AR dont l'ordre est relativement élevé (de 14 à 16).
- 2) Analyse et traitement de certaines séries chronologiques, représentant l'évolution de quelques variables météorologiques (précipitations, température, durée d'insolation) en deux stations: ALGER et ANNABA, ainsi que la série des observations universelles de l'activité solaire (nombre de taches). La méthodologie de test de cohérence a été utilisée sur les observations de la durée d'insolation à ALGER, ce qui nous a permis de prouver la non homogénéité des observations de la période 80-86, et d'apporter par la suite des corrections aux mesures incohérentes. Les résultats de l'analyse spectrale montrent que les variations de la température ainsi que la durée d'insolation peuvent être représentées par un modèle AR. Mais en ce qui concerne les précipitations, à notre avis, une modélisation ARMA serait plus convenable.

Quant aux périodicités, nous n'avons pas pu détecter dans les différentes séries, des périodicités autres que les périodicités annuelles et saisonnières. Ceci est certainement dû au nombre insuffisant d'années d'observation. En effet, pour l'activité solaire, la détection des cycles de 10.8 et 5.7 années était très facile avec l'utilisation des moyennes annuelles de la période 1800-1984. Mais ceci devient impossible lorsqu'on utilise les moyennes mensuelles de la période 1958-1984. Enfin les résultats de l'analyse multicanaux, montrent qu'il est probable que la demi-période de l'activité solaire (5.7 années) ait un effet sur la température.

Nous espérons que ce travail va ouvrir la porte à un grand nombre d'activités dans ce sens. Ainsi il serait intéressant d'étudier la version adaptative des différents algorithmes AR ainsi que les méthodes et les algorithmes pour la modélisation ARMA. De plus, il serait intéressant d'étudier les méthodes d'estimation des spectres évolutifs: méthodes adaptées aux signaux non stationnaires et basées sur des représentations temps-fréquence comme la distribution de WIGNER-VILLE par exemple [50]. Comme il est souhaitable de voir l'apport des méthodes d'estimation utilisant des normes ( $L_1$  (Valeur absolue),  $L_\infty$  (minmax) ...) autre que la norme  $L_2$  des moindres carrées. Enfin, étant donné que l'analyse spectrale dépend à la fois des connaissances à priori (nature et type des signaux) et de la finalité de l'analyse (bonne Résolution, faible variance, stabilité du modèle, ...etc), l'unification des méthodes d'estimation (classiques et haute résolution) en vue de la réalisation d'un système expert d'aide à l'analyse serait d'un grand apport au problème de l'estimation spectrale.

# BIBLIOGRAPHIE

## BIBLIOGRAPHIE

- [1] J. G. Ables, "Maximum Entropy Spectral Analysis," Astronomy and Astrophysics Supplement Series, vol.15, pp. 383-393, June 1974.
- [2] H. Akaike, "Power Spectrum Estimation through Autoregressive Model Fitting," Ann. Inst. Statist. Math., Vol 21, pp. 407-419, 1969.
- [3] H. Akaike, "A New Look at the Statistical Model Identification," IEEE Trans. Autom. Control, vol AC-19 pp. 716-723, Dec. 1974.
- [4] N. Andersen "On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis," Geophysics, vol. 39, pp. 69-72, February 1974.
- [5] M. Bellanger, Analyse des Signaux et Filtrage Numérique Adaptatif, Ed. Masson 1989.
- [6] A. Bensidi-Slimane, Analyse spectrale A.R.M.A et application au traitement de la parole, Thèse de Magister en Electronique, USTHB, Avril 1992.
- [7] C. Bingham, M. D. Godfrey, and J.W. Tukey, "Modern Techniques of Power Spectrum Estimation", IEEE Transaction on Audio and Electroacoustics, Vol. AU-15, pp. 56-66, June 1967.
- [8] R. B. Blackman, J. W. Tukey, The Measurement of Power Spectra, From the Point of View of Communications Engineering, Dover Publication Inc., New York, 1958.
- [9] R. Boité et M. Kunt, Traitement de la Parole, Presses polytechnique romandes, Lausanne, 1987.
- [10] J. P. Burg, "Maximum Entropy Spectral Analysis," Proceedings of the 37th Meeting of the Society of Exploration Geophysicists, 1967.
- [11] J. P. Burg, "A New Analysis Technique For Time Series Data," NATO Advanced Study Institute on Signal Processing with Emphasis on UnderWater Acoustics, August 12-23 1968.
- [12] J. P. Burg, "The Relationship between Maximum Entropy Spectra and Maximum Likelihood Spectra," Geophysics., vol. 37, pp. 375-376, April 1972.

- [13] D. Burshtein and W. Ehud, "Confidence Intervals for the Maximum Entropy Spectrum," IEEE Trans., Acoust., Speech, Signal Processing, vol. ASSP- 35, No. 4, pp. 504-510, April 1987.
- [14] W. Y. Chen and G. R. Stegen, "Experiments with maximum entropy power spectra of sinusoids," J. Geophysics, vol. 79, pp. 3019-3022, July 1974.
- [15] D. G. Childers, (Ed.), Modern Spectrum Analysis, IEEE Press, New York, 1978.
- [16] D. Cochran, D. Sinno and H. Gish, "A Comparison of Nonparametric Statistics for Multiple-Channel Detection of Signals in Ocean Noise," IEEE International Conference on ASSP, Vol. V CH2977 7/91/0000-3685, pp. 3685-3688, Jul. 1991.
- [17] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Fourier Series," Math Comput., vol 19, pp 297-301, 1965.
- [18] F. de Coulon, Theorie et Traitement des Signaux, Editoin Georgi, Lausanne 1984.
- [19] B. Derras, "Prony's Method and its Application to Power Spectral Estimation", (Course Project), University of Colorado, Boulder, Colorado, Spring 1985.
- [20] R. E. DuBroff, "The Effective Autocorrelation Function Of Maximum Entropy Spectra," IEEE Trans., Acoust., Speech, Signal Processing, vol. ASSP-63, pp. 1622-1623, Nov. 1975.
- [21] M. Engelson, Modern Spectrum Analyzer: Theory and Applications, Artech House, Inc., 1984.
- [22] F. Fabre, "Algorithmes Transversaux Rapides d'Identification Recursives et Méthodes de Modélisation des Réponses Impulsionnelles Complexes Associées. Application au Signal de Parole," Thèse de Docteur Ingenieur E.N.S.T, Paris, 1986.
- [23] A. Ferrari and L. Alengrin, "Estimation of the Frequencies of a Complex Sinusoidal Noisy Signal Using Fourth Order Statistics," IEEE International Conference on ASSP, Vol. V CH2977 7/91/0000-3457, pp. 3457-3460, Jul. 1991.
- [24] J. P. Fourier , Théorie Analytique de la Chaleur, Paris, France: Didot, 1822.
- [25] B. Friedlander, "Lattice methods for Spectral Estimation," Proceedings of IEEE, Vol.70, pp. 990-1017, Sep. 1982.

- [26] D. Garreau et B. Georgel, "La Méthode de Prony en Analyse de Vibrations," *Traitement du signal*, Vol. 3 N°4-5, 1986.
- [27] G. B. Giannakis and M. K. Tsatsanis, "HOS or SOS for Parametric Modelling?", *IEEE International Conference on ASSP*, Vol. V CH2977 7/91/0000-3097, PP. 3097-3101, Jul. 1991.
- [28] D. F. Gingras, E. Masry, "Autoregressive Spectrale Estimation in Additive Noise," *IEEE Trans., Acoust., Speech, Signal Processing*, vol. ASSP-36, No. 4, April 1988.
- [29] P. R. Gutowski, E. A. Robinson and S. Treitel, "Spectral Estimation: Fact or Fiction," *IEEE Transactions on Geoscience electronics*, vol. GE-16, pp. 80-84, April 1978.
- [30] B. I. Helme and C. L. Nikias, "Improved Spectrum Performance via a Data-Adaptive Weighted Burg Technique," *IEEE Trans., Acoust., Speech, Signal Processing*, vol. ASSP-33, No. 4, August 1985.
- [31] R. Hippenstiel, "Spectral Estimation of Short Duration Signals Using Time Domain Extrapolation," *IEEE International Conference on ASSP*, Vol. V CH2977 7/91/0000-3133, PP. 3133-3136, Jul. 1991.
- [32] R. H. Jones, "Autoregressive Ordre Selection," *Geophysics.*, vol 41, pp. 771-773, Aug. 1976.
- [33] R. H. Jones, "Identification and Autoregressive Spectrum Estimation," *IEEE Trans. Autom. Control*, vol. AC-19, pp. 894-897, Dec. 1974.
- [34] M. Kaveh and G. R. Cooper, "An Empirical Investigation of the Properties of the Autoregressive Spectral Estimator," *IEEE Trans. On Info. Theory*, Vol. IT-22, pp.313-323, May 1976.
- [35] S. M. Kay and S. L. Marple jr, "Spectral Analysis, a Modern Perspective," *Proceedings of the IEEE*, Vol. 69, N° 11, pp 1379-1419, Nov.1981.
- [36] S.M. Kay and V. Negesha "Spectral Analysis Based on the Canonical Autoregressive Decomposition," *IEEE International Conference on ASSP*, Vol. V CH2977 7/91/0000-3137, PP. 3137-3140, jul. 1991.
- [37] A. Y. Khintchine, "Korrelations Theorie der Stationaren Stochastischen Prozesse," *Math Ann*, vol 109, pp. 604-615, 1934.

- [38] M. Kunt, *Traitement Numérique des Signaux*, Editon Dunod, Presses Polytechniques Rommandes, Lausanne 1984.
- [39] R. T. Lacoss "Data Adaptive Spectral Analysis Methods" *Geophysics*, vol. 36, pp. 661-675, Aug. 1971.
- [40] T. E. Landers and R. T. Lacoss, "Some Geophysical Applications of Autoregressive Spectral Estimates," *IEEE Trans. on Geoscience Electronics*, Vol. GE-15, pp. 26-32, Jan. 1977.
- [41] N. Levinson N. "A Heuristic Exposition of Wiener's Mathematical Theory of Prediction and Filtering," *Journal of Math. and Physics*, Vol 26, PP 110-119, 1947.
- [42] J. Makhoul "Linear Prediction: A Tutorial Review," *IEEE Trans., Acoust., Speech, Signal Processing*, vol. ASSP-63, Nov. 1975; vol ASSP-64, pp.561-580, Feb. 1976.
- [43] J. Makhoul, "Stable and Efficient Lattice Methodes for Linear Prediction," *IEEE Trans., Acoust., Speech, Signal Processing*, vol. ASSP-25, pp.423-428, Nov. 1977.
- [44] J. Makhoul and A Steinhardt, "Where's the Peak ?," *IEEE International Conference on ASSP*, Vol. V CH2977 7/91/0000-3361, PP. 3361-3464, Jul. 1991.
- [45] J. D. Markel and A. Gray, "On Autocorrelation Equation as Applied to Speech Analysis," *IEEE transaction on Audio and Electroacoustics* Vol. Au-21, N°02, pp. 61-80, April 1973.
- [46] L. Marple, "A New Autoregressive Spectrum Analysis Algorithm," *IEEE Trans., Acoust., Speech, Signal Processing*, vol. ASSP-28, No. 4, pp. 441-454, Aug. 1980.
- [47] L. Marple, *Digital Spectral Analysis with Applications*, Prentice-Hall Inc, New Jersy, 1987.
- [48] K. W. Martin and M. Padmanabhan, " The Spectral Analysis of Short Data-Segments Using an IIR adaptive Filtre," *IEEE International Conference on ASSP*, Vol. V CH2977 7/91/0000-3237, PP. 3237-3240, Jul. 1991.
- [49] J. Max, *Méthodes et Techniques de Traitement du Signal et Applications aux Mesures Physiques*, 5° Edition Masson, Tomes I, Paris 1985.
- [50] J. Max, *Méthodes et Techniques de Traitement du Signal et Applications aux Mesures Physiques*, 5° Edition Masson, Tomes II, Paris 1985.
- [51] J. Max, M. Diot et R. Bigret "Les Analyseurs de Spectre à FFT et les Analyseurs de Spectre à Corrélation," *Traitement du signal*, Vol. 3, N°4-5, 1986.

- [52] S. Mayrargue and T. Blu, "Relationship Between High-Resolution Methods and Discret Fourier Transform," IEEE International Conference on ASSP, Vol. V CH2977 7/91/0000-3321, pp. 3321-3324, Jul. 1991.
- [53] R. L. Moses and J. A. Cadzow, "A Recursive Procedure for ARMA Modeling," IEEE Trans., Acoust., Speech, Signal Processing, vol. ASSP-33, No. 4, October 1985.
- [54] M. Najim, Modélisation et identification en traitement du signal, Edition Masson, Paris, 1988.
- [55] A. Papoulis, "Maximum Entropy and Spectral Estimation: a Review," IEEE Trans., Acoust., Speech, Signal Processing, vol. ASSP-29 No. 6, pp. 1176-1186, Dec. 1981.
- [56] E. Parzen, "Some Recent Advance in Time Series Modeling," IEEE Trans. Autom. Control, Vol. AC-19, pp. 723-730, Dec. 1974.
- [57] A. V. Oppenheim and R. W. Schaffer, Digital Signal processing, Printice-Hall Inc, New Jersey, 1975.
- [58] E. A. Robinson, "An Historical Perspective of Spectrum Estimation," Proceedings of the IEEE, Vol. 70, No. 9, pp. 885-907, Sep. 1982
- [59] C. Serviere, D Baudois and A Silvent, "Estimation of Correlated Frequencies in Noise Concelling Using Higher order moments," IEEE International Conference on ASSP, Vol. V CH2977 7/91/0000-3477, pp. 3477-3480, Jul. 1991.
- [60] A. Schuster "On the Investigation of Hidden Periodicities with application to supposed 26-day period of Meteorological Phenomena," Terr. Magnet. Vol 3, pp. 13-41, Mar. 1898.
- [61] A. Schuster, "On the Periodicities of Sun Spots," Philos Trans, Roy. Soc. London, Ser. A, Vol 206, 1906.
- [62] P. Stolca, "On Spectral and Root Forms of Sinusoidal Frequency Estimator," IEEE International Conference on ASSP, Vol. V CH2977 7/91/0000-3257, pp. 3257-3260, Jul. 1991.
- [63] H. Tong, "Autoregressive Model Fitting with Noisy Data by Akaike's Information Criterion," IEEE Trans. On Information Theory, Vol.IT-21, pp. 476-480, July 1975.
- [64] H. Tong, "More on Autoregressive Model Fitting with Noisy Data by Akaike's Information Criterion," IEEE Trans. On Information Theory, Vol. IT-23, pp. 409-410, May 1977.

- [65] J. W. Tukey, "The Sampling Theory of Power Spectrum Estimates," Proceeding Symposium on Applied Auto-correlation Analyses of Physical Problemes, U.S. Office of Naval Research 1949.
- [66] J. W. Tukey, "An Introduction to the Calculation of Numerical Spectrum Analysis," in Spectral Analysis of Time Series, B. Harris, (Ed.), New York: Wiley, 1967.
- [67] T. J. Ulrych and T. N. Bishop, "Maximum Entropy Spectral Analysis and Autoregressive Decomposition," Reviews of geophysics and space physics, Vol.13, pp. 183-200, Feb. 1975.
- [68] A. VanDen Bos, "Alternative Interpretation of Maximum Entropy Spectral Analysis," IEEE Trans. On Information Theory, Vol. IT-17, pp. 493-494, Jul. 1971.
- [69] M. Wax, "Ordre Selection for AR Models by Predictive Least Squares," IEEE Trans., Acoust., Speech, Signal Processing, Vol. ASSP-36, No. 4, pp. 581-588, April 1988.
- [70] P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short Modified Periodograms," IEEE Transaction on Audio and Electroacoustics, Vol AU-15, pp 70-73, Jun. 1967.
- [71] N. Wiener, Generalized Harmonic Analysis, Acta Math., Vol. 55, pp. 117-158, 1930.
- [72] G. U. Yule, "On a Methode of Investigation Periodicities in Distrubed Series, with special reference to Wolfer's Sunspot Numbers," Phil. Trans. Roy. Soc. London, A, vol 226, pp. 267-298, 1927.

# ANNEXES

## ANNEXE A1

### (a): Périodogramme et Transformation de FOURIER discrète

Selon le théorème de Wiener-Kinchine, la densité spectrale de puissance  $S(f)$  d'un signal  $x(k)$ , est la transformée de FOURIER de la fonction d'autocorrélation  $r(m)$ .

$$S(f) = \sum_{m=-\infty}^{+\infty} r(m) \cdot \exp(-j2\pi m f T) \quad (\text{A1.1})$$

$T$  est la période d'échantillonnage du signal  $x(k)$ .

Si on note par  $x_N(k)$  le signal défini par

$$x_N(k) = \begin{cases} x(k) & \text{si } 0 \leq k \leq N-1 \\ 0 & \text{Ailleurs} \end{cases} \quad (\text{A1.2})$$

Alors la transformation de FOURIER discrète de  $x_N(k)$  est donnée par l'expression suivante:

$$X_N(f) = \sum_{k=-\infty}^{+\infty} x_N(k) \cdot \exp(-j2\pi k f T) = \sum_{k=0}^{N-1} x(k) \cdot \exp(-j2\pi k f T) \quad (\text{A1.3})$$

En ce qui concerne la fonction d'autocorrélation, en utilisant  $N$  échantillons du signal  $x(k)$ , l'estimation biaisée de  $r(m)$  est définie par

$$\hat{r}(m) = \frac{1}{N} \sum_{k=0}^{N-m-1} x^*(k) \cdot x(k+m) \quad m=0, 1, \dots, N-1$$

Compte tenu de la relation (A1.2) on peut écrire:

$$\hat{r}(m) = \frac{1}{N} \sum_{k=-\infty}^{+\infty} x_N^*(k) \cdot x_N(k+m) \quad m=0, \pm 1, \dots \quad (\text{A1.4})$$

En remplaçant  $r(m)$  par  $\hat{r}(m)$  dans l'expression (A1.1), on obtient

$$S(f) = \frac{1}{N} \sum_{m=-\infty}^{+\infty} \left[ \sum_{k=-\infty}^{+\infty} x_N^*(k) \cdot x_N(k+m) \right] \cdot \exp\{ -j2\pi m f T \} \quad (A1.5)$$

$$S(f) = \frac{1}{N} \left[ \sum_{k=-\infty}^{+\infty} x_N^*(k) \cdot \exp\{ j2\pi k f T \} \right] \left[ \sum_{k=-\infty}^{+\infty} x_N(k+m) \cdot \exp\{ -j2\pi (k+m) f T \} \right]$$

d'après la relation (A1.3) :

$$S(f) = \frac{1}{N} X_N^*(f) X_N(f) = \frac{1}{N} |X_N(f)|^2 \quad (A1.6)$$

c'est la méthode du périodogramme simple.

(b): Transformation de Fourier discrète

La version numérique de la Transformée de Fourier continue d'un signal discret  $x(kT_0)$  est donnée par

$$X(f) = \sum_{k=-\infty}^{+\infty} x(kT_0) \cdot \exp\{ -2\pi f k T_0 \} \quad (A1.7)$$

$T_0$  est la période d'échantillonnage

Mais, pratiquement il est impossible d'évaluer l'expression (A1.7) sur un nombre infini d'échantillons. On doit se limiter à la durée du signal ( $N \cdot T_0$ ). De plus, avec un système de traitement numérique, on ne peut évaluer  $X(f)$  en fonction de la variable continue  $f$ . Il faut procéder à une discréditation de la fréquence. Comme  $X(f)$  est périodique, de période  $1/T_0$ , en prélevant  $N$  échantillons avec un pas d'échantillonnage  $\Delta f = 1/NT_0$  on aura:

$$X(n \cdot \Delta f) = \sum_{k=0}^{N-1} x(k) \cdot \exp\{ -2\pi \cdot n \cdot \Delta f \cdot k T_0 \} \quad (A1.8)$$

ou simplement

$$X(n) = \sum_{k=0}^{N-1} x(k) \cdot \exp\{ -2\pi \frac{n \cdot k}{N} \} \quad (A1.9)$$

L'intérêt de la T.F.D réside dans l'existence d'un algorithme de calcul rapide [14] dénomé "Transformée de Fourier Rapide" T.F.R, plus couramment utilisé sous la dénomination anglo-saxone de F.F.T (Fast Fourier Transform). La F.F.T consiste à décomposer le calcul de la T.F.D d'une suite numérique de longueur N, en plusieurs T.F.D de longueur plus petites, de façon à exploiter la périodicité et la symétrie des exponentielles complexes [30].

En effet, la somme finie (A1.9) s'écrit:

$$X(k) = x(0) + x(1)W^k + \dots + x(N-1)W^{(N-1)k}; \quad W = \exp(-j2\pi/N)$$

Si l'on regroupe les termes d'ordre paire et impaire on obtient:

$$X(k) = P(k) + I(k) \text{ où}$$

$$P(k) = x(0) + x(2)W^{2k} + x(4)W^{4k} + \dots + x(N-2)W^{(N-2)k}$$

$$I(k) = x(1) + x(3)W^{2k} + x(5)W^{4k} + \dots + x(N-1)W^{(N-2)k}$$

$P(k)$  et  $I(k)$  sont aussi des TFD d'ordre moitié M ( $M=N/2$ ), indépendantes l'une de l'autre. Ceci nous permet de réitérer le procédé sur des intervalles d'ordre M/2.

Le coût de calcul par le procédé de la TFD normale est de l'ordre de  $N^2$  opérations complexes. Si N est une puissance de deux ( $N=2^p$ ), avec la FFT on peut faire beaucoup mieux, et passer à une complexité de l'ordre de  $N \cdot \log_2(N)$ .

## ANNEXE A2

### Caracteristiques des fonctions fenêtres

Tableau A2.1: définition temporelle de quelques fenêtres

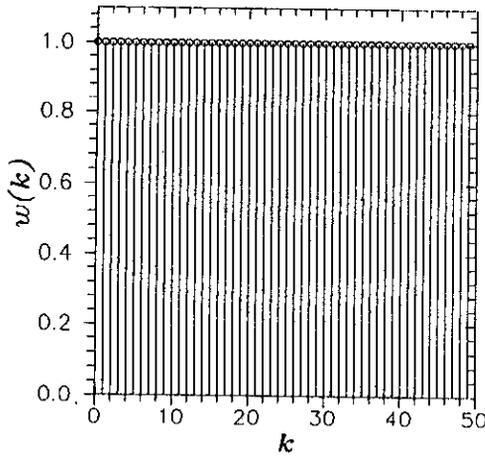
Fenêtre	$\omega(k)$ ou $\omega(m)$
Naturelle	1
Bartlett	$1 - 2 \cdot  t[n] $
Hanning	$0.50 + 0.50 \cdot \cos(2 \cdot \pi \cdot t[n])$
Hamming	$0.54 + 0.46 \cdot \cos(2 \cdot \pi \cdot t[n])$
Blackman <sup>(1)</sup>	$0.42 + 0.5 \cdot \cos(2 \cdot \pi \cdot t[n]) + 0.008 \cdot \cos(4 \cdot \pi \cdot t[n])$
Nuttall	$\sum_{r=0}^3 \alpha_r \cdot \cos(2 \cdot \pi \cdot r \cdot t[n])$ $\alpha_0 = 0.3635819$ $\alpha_1 = 0.4891755, \quad \alpha_2 = 0.1365995, \quad \alpha_3 = 0.0106411$
$t[k] = (k - (N-1)/2) / (N-1) \quad 0 \leq k \leq N-1$	
$t[m] = m/N \quad -(N-1)/2 \leq m \leq (N-1)/2, \quad N \text{ impair}$	

tableau A2.2 Caracteristiques de quelques fenêtres

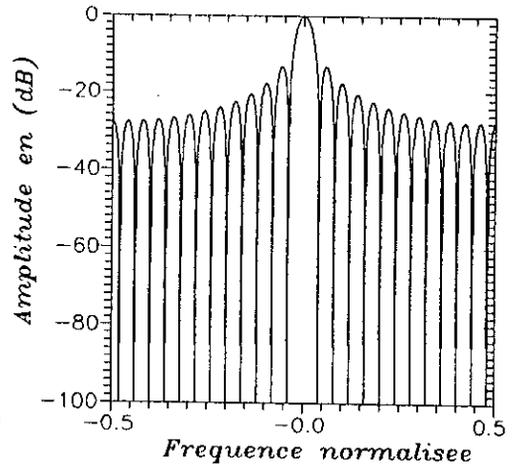
Fenêtre	Atténuation (db) 1° lobe secon. / pic central	Largeur de la base du pic central en $(1/T_0)$
Naturelle	-13	$2/N$
Bartlett	-26	$4/N$
Hanning	-32	$4/N$
Hamming	-43	$4/N$
Blackman	-57	$6/N$
Nuttall	-98	$8/N$

(1) Plusieurs fonctions fenêtres peuvent être générées par la relation de Blackman généralisée:

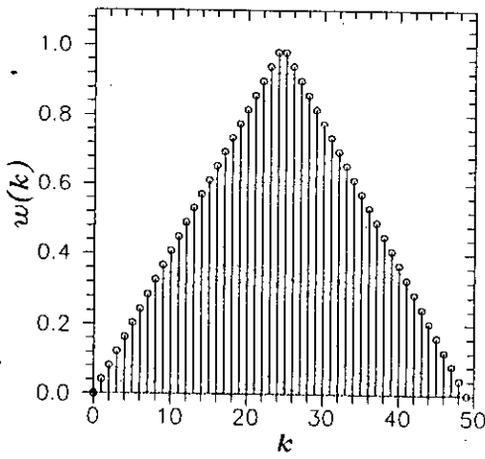
$$w(n) = \sum_{r=0}^R \alpha_r \cdot \cos(2\pi r \cdot t[n]) \quad \text{avec} \quad \sum_{r=0}^R \alpha_r = 1$$



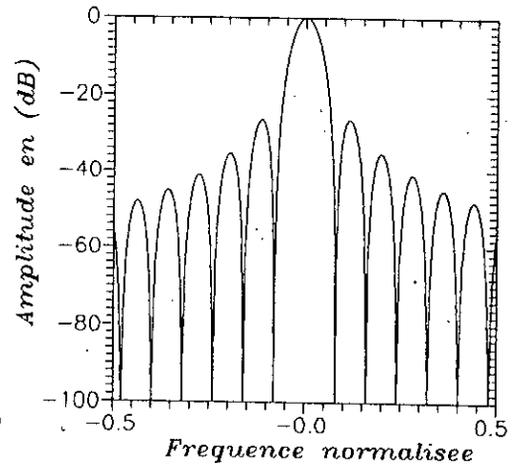
Fenêtre naturelle



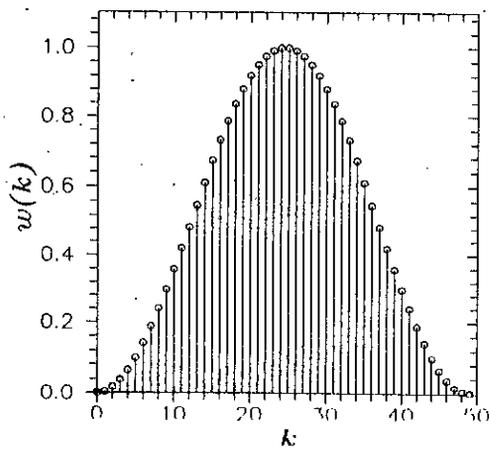
Fenêtre naturelle



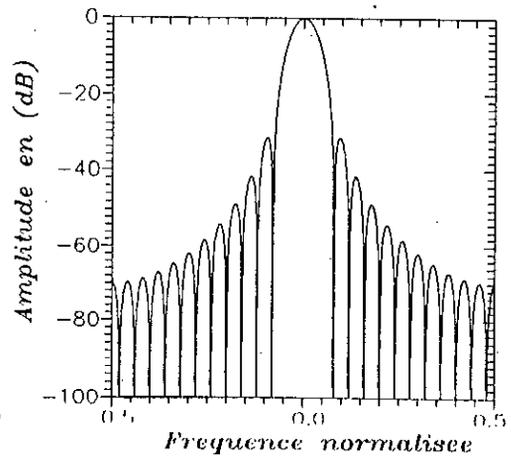
Fenêtre de Bartlett



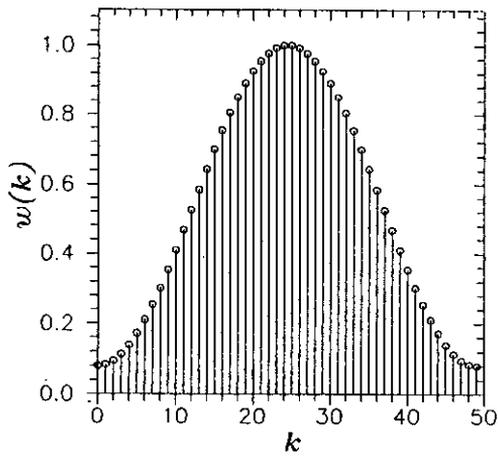
Fenêtre de Bartlett



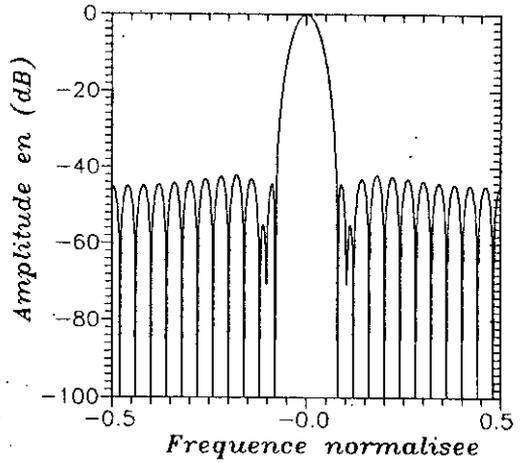
Fenêtre de Hanning



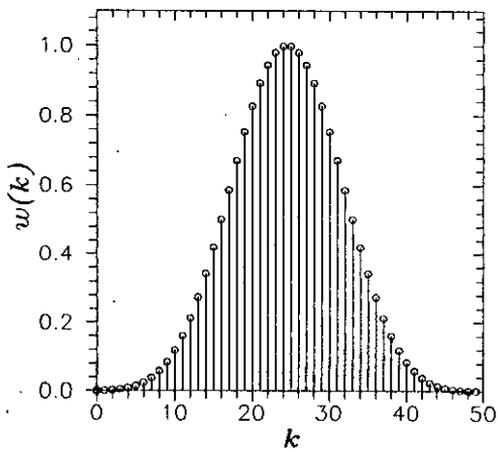
Fenêtre de Hanning



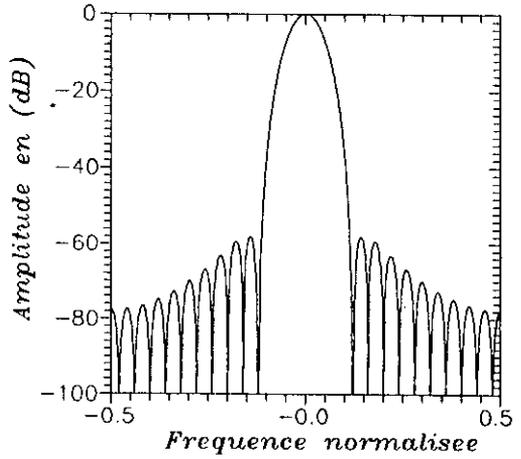
Fenêtre de Hamming



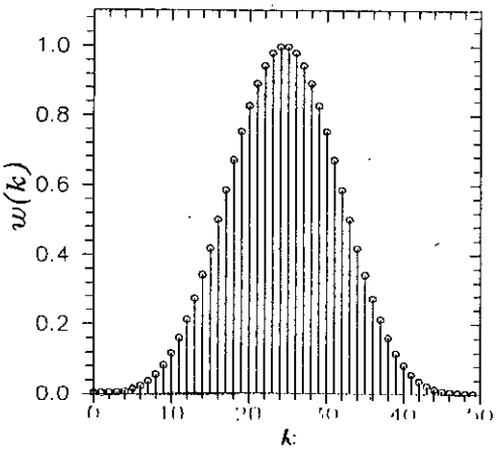
Fenêtre de Hamming



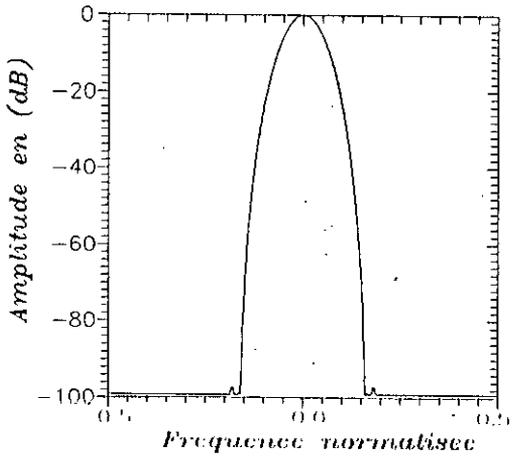
Fenêtre de Blackman



Fenêtre de Blackman



Fenêtre de Nuttall



Fenêtre de Nuttall

## ANNEXE A3

### Algorithme de LEVINSON

L'algorithme de LEVINSON permet de résoudre le système d'équations normales (A3.1), par une récursion sur l'ordre de prédiction

$$R_p \cdot A_p = E_p \quad (A3.1)$$

tel que

$$R_p = \begin{bmatrix} r(0) & r^*(1) & \dots & r^*(p) \\ r(1) & r(0) & \dots & r^*(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(p) & \dots & r(1) & r(0) \end{bmatrix}, \quad A_p = \begin{bmatrix} 1 \\ a_1^f \\ \vdots \\ a_p^f \end{bmatrix}, \quad E_p = \begin{bmatrix} \alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

la matrice d'autocorrélation  $R_p$  est supposé connue.

L'algorithme de Levinson peut être considéré comme étant un algorithme rapide pour la décomposition triangulaire dite de CHOLESKY de la matrice de corrélation  $R_p$ .

### ALGORITHME

#### Entrées:

$P$  : ordre du modèle

$r(0), r(1), \dots, r(p)$ : éléments de la fonction d'autocorrélation du signal  $x(k)$ .

#### Sorties:

$a_m(1), a_m(2), \dots, a_m(m)$ : Coefficients du modèle AR.

$C_m \quad 1 \leq m \leq P$ : coefficients de réflexion.

$\alpha_p$ : Variance de l'erreur de prédiction

### INITIALISATION

$$\alpha_0 = r(0)$$

$$a_{0,0} = 1$$

## Récursion

For m=0 to P-1 do

$$\gamma_m = \sum_{k=0}^m r(m-(k-1)) a_{k,m}$$

$$C_{m+1} = \frac{-\gamma_m}{\alpha_m}$$

$$\alpha_{m+1} = \alpha_m \left( 1 - C_{m+1}^2 \right)$$

$$a_{m,m+1} = 0$$

For k = 0 to m+1 do

$$a_{m+1,k} = a_{m,k} + C_{m+1} a_{m,m+1-k}$$

End loop on k

End loop on m

Exemple:

$$x(k) \left\{ x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5 \right\}; N=5, P=3;$$

$$\text{Matrice } R_3 = \left\{ r_0 = 55, r_1 = 40, r_2 = 26, r_3 = 14 \right\}$$

$$\alpha_0 = 55$$

$$m=0 \quad C_1 = -40/55 = -0.72727 \quad \alpha_1 = 25,909$$

$$\alpha_{11} = -0.72727, \alpha_{10} = 1;$$

$$m=1 \quad C_2 = 0.11929 \quad \alpha_2 = 25,5403$$

$$\alpha_{22} = 0.11929, \alpha_{21} = -0.81403, \alpha_{20} = 1;$$

$$m=2 \quad C_3 = 0.093702 \quad \alpha_3 = 25,3160$$

$$\alpha_{33} = 0.093702, \alpha_{32} = 0.04301, \alpha_{31} = -0.80285, \alpha_{30} = 1;$$

## ANNEXE A4

### Algorithme de BURG

La méthode générale pour le calcul de  $p$  coefficients de prédiction d'un signal stationnaire  $x(k)$ , consiste à résoudre le système d'équations normales où les  $p$  premières valeurs de la fonction d'autocorrélation doivent être estimées.

BURG a montré que ce calcul peut être fait directement, et sans faire appel aux éléments de la fonction d'autocorrélation. Il suffit pour cela de minimiser la somme des puissances des deux erreurs de prédiction: directe et rétrograde sans considérer l'hypothèse de fenêtrage.

L'algorithme de Burg donne les coefficients de réflexion. Les paramètres du modèle Autorégressif s'obtiennent à l'aide de la récurrence de LEVINSON.

### ALGORITHME

Entrées:

$P$  : ordre du modèle

$x(0), x(1), \dots, x(N-1)$ :  $N$  échantillons du signal  $x(k)$

Sorties:

$C_m$   $1 \leq m \leq P$  : coefficients de réflexion.

$\sigma_p$  : Variance de l'erreur de prédiction

$E_f^m, E_b^m$ ,  $0 \leq m \leq p-1$  : Erreurs de prédiction

### INITIALISATION

$$d_0 = 2 \sum_{k=0}^{N-1} y^2(k)$$

$C_0 = 0$

For  $k=0$  to  $N-1$  do

$E_f^0(k) = y(k)$

$E_b^0(k) = y(k)$

End Loop on  $k$

## RécurSION

For m=1 to P do

$$d_m = (1 - C_m^2) \cdot d_{m-1} - E_{m-1}^f(m-1) - E_{m-1}^b(N-1)$$

$$f_m = \sum_{k=m}^{N-1} E_{m-1}^f(k) E_{m-1}^b(k-1)$$

$$C_m = \frac{-2 f_m}{d_m}$$

For k = m to N-1 do

$$E_m^f(k) = E_{m-1}^f(k) + C_m E_{m-1}^b(k-1)$$

$$E_m^b(k) = E_{m-1}^b(k-1) + C_m E_{m-1}^f(k)$$

End Loop on k

$$a_{m-1, m} = 0$$

For k = 0 to m do

$$a_{m, k} = a_{m-1, k} + C_m a_{m-1, m-k}$$

End loop on k

End Loop on m

$$\alpha_P = \frac{d_P}{2(N-p)}$$

Exemple:

$$x(k) \cdot \left\{ x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5 \right\}; N=5, P=2$$

$$d_0 = 84;$$

$$E_{12}^f = 1.0476$$

$$E_{12}^b = -0.9048$$

$$E_{13}^f = 1.0952$$

$$E_{13}^b = -0.8972$$

$$C_1 = 80/84 = 0.9524$$

$$E_{14}^f = 1.1428$$

$$E_{14}^b = -0.8096$$

$$a_{11} = 0.9524, \alpha_{10} = 1;$$

$$E_{15}^f = 1.1908$$

$$E_{15}^b = -0.7620$$

$$d_1 = 6.13 \quad C_2 = 0.9575$$

$$a_{22} = 0.9575, \alpha_{21} = -1.8637, \alpha_{20} = 1;$$

$$E_{23}^f = 0.2289$$

$$E_{23}^b = 0.3306$$

$$d_2 = 0.3483$$

## ANNEXE A5

### Algorithme de CHOLESKY

Les opérations de factorisation et de triangulation de matrices constituent les opérations de base des nouveaux algorithmes de traitement du signal. Ainsi le procédé de décomposition d'une matrice  $A(n,n)$  en LU, avec L triangulaire inférieure et U triangulaire supérieure, permet de ramener la résolution du système d'équations linéaires:

$$AX=B$$

où A est une matrice carrée d'ordre N,  $X, B \in \mathbb{R}^n$ ,

à la résolution de deux systèmes particuliers:

-Système triangulaire inférieur  $LY=B$

-Système triangulaire supérieur  $UX=Y$

Si A est une matrice symétrique et définie positive<sup>(1)</sup>, alors CHOLESKY nous affirme que A peut se mettre sous la forme:

$$A = S S^T$$

où S est une matrice triangulaire inférieure (ou supérieure).

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & \dots & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} s_{11} & 0 & \dots & 0 \\ s_{21} & s_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & \dots & \dots & s_{nn} \end{pmatrix} \begin{pmatrix} s_{11} & s_{21} & \dots & s_{n1} \\ 0 & s_{22} & \dots & s_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & s_{nn} \end{pmatrix}$$

Certains auteurs considèrent S comme étant la racine carrée de A, du fait qu'une telle décomposition ( $A=SS^T$ ) est unique.

L'algorithme de décomposition de CHOLESKY permet le calcul des coefficients  $s_{l,j}$  de la matrice S par récurrence.

#### ALGORITHME

Entrées: A :Matrice carrée d'ordre N,

B :Vecteur à N éléments.

Sorties: S : Matrice triangulaire inférieure d'ordre N.

X : Vecteur solution du système  $AX=B$ .

(1): A Symétrique est dite définie positive si  $\forall X \in \mathbb{R}^n, X \neq 0$   
 $X^T A X > 0$ . Si  $A=SS^T \Rightarrow X^T S S^T X = \|S^T X\|^2 > 0$

Décomposition de CHOLESKY :  $A=SS^T$

For  $i=1$  to  $n$  Do

$$S_{ii} = \left( a_{ii} - \sum_{k=1}^{i-1} S_{ik}^2 \right)^{1/2}$$

if  $S_{ii}=0$  then Halt (A est non définie positive)

While  $j > i$  and  $j \leq n$  do

$$S_{ij} = \frac{1}{S_{ii}} \left( a_{ij} - \sum_{k=1}^{i-1} S_{ik} S_{jk} \right)$$

End Loop on  $i$

Resolution du système  $SY=B$

For  $j=1$  to  $N$  do

$$Y_j = \frac{1}{S_{jj}} \left( B_j - \sum_{i=1}^{j-1} Y_i \cdot S_{ji} \right)$$

Resolution du système  $S^T X=Y$

For  $j=N$  down to 1 do

$$X_j = \frac{1}{S_{jj}} \left( Y_j - \sum_{i=j+1}^N S_{ij} \cdot X_i \right)$$

Exemples:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 1 & 10 \\ 3 & 1 & 35 & 5 \\ 4 & 10 & 5 & 45 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 9 \\ 35 \\ -45 \\ -89 \end{pmatrix} \implies S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & -5 & 1 & 0 \\ 4 & 2 & 3 & 4 \end{pmatrix}, X = \begin{pmatrix} 1 \\ -3 \\ 2 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} -9 & -6 & 15 \\ -6 & 12 & 14 \\ 15 & 14 & -25 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 93 \\ -2 \\ -179 \end{pmatrix}$$

A est non définie positive.  
Donc la décomposition  $A=SS^T$   
n'est pas possible.



## ANNEXE A7

### LES VARIABLES MÉTÉOROLOGIQUES.

L'Office National de la Météo (ONM), procède, d'une façon permanente, à la mesure et l'observation de plusieurs variables météorologiques, parmi lesquelles on peut citer:

- 1) La Température ou l'état thermodynamique de l'aire, mesurée en degré Celsius ( $^{\circ}\text{C}$ ).
- 2) La pluviosité ou la quantité de pluie mesurée à l'aide d'instruments pluviométriques. Elle est donnée en mm/Jour. Il est bon de rappeler que 1mm correspond à 1 Litre/m<sup>2</sup>.
- 3) La durée d'insolation ou la période durant laquelle le soleil a brillé sur le terrain d'observation. Elle est enregistrée à l'aide d'héliographes à un dixième (1/10) d'heure près (6 min).
- 4) L'insolation relative qui n'est autre chose que le rapport de l'insolation effective à sa durée astronomique possible.
- 5) Le vent, qui représente le mouvement de l'aire. Il est caractérisé principalement par sa vitesse, à 2.5 m du sol, exprimée en m/s.
- 6) L'humidité ou la teneur de l'atmosphère en vapeur d'eau. Les instruments de mesure usuels, mesurent l'humidité relative (en %): c'est-à-dire le rapport de la tension de vapeur réellement observée à la tension de vapeur maximale possible.

Dans les figures qui suivent, on a représenté en fonction du temps les variations de l'ensemble de ces variables atmosphériques relatives aux deux stations: ALGER et ANNABA

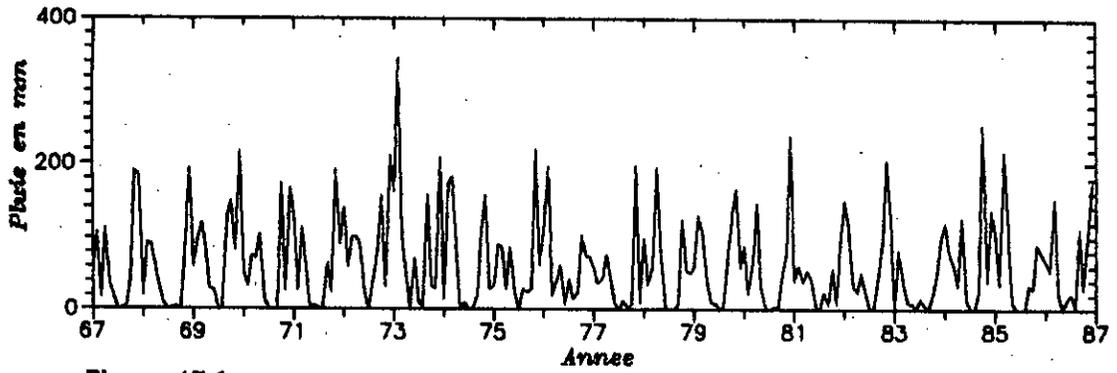


Figure A7.1

Quantité de pluie mensuelle  
Période 1967-1986, Station d'ALGER

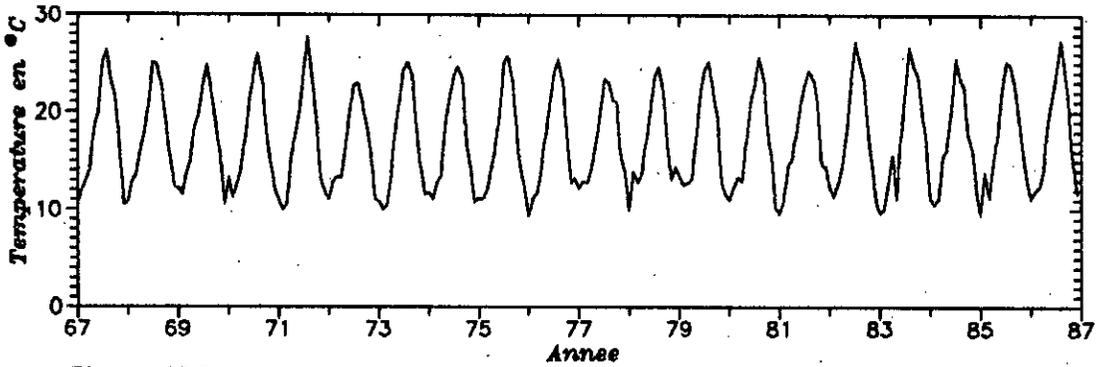


Figure A7.2

Variation de la température moyenne mensuelle  
Période 1967-1986, Station d'ALGER

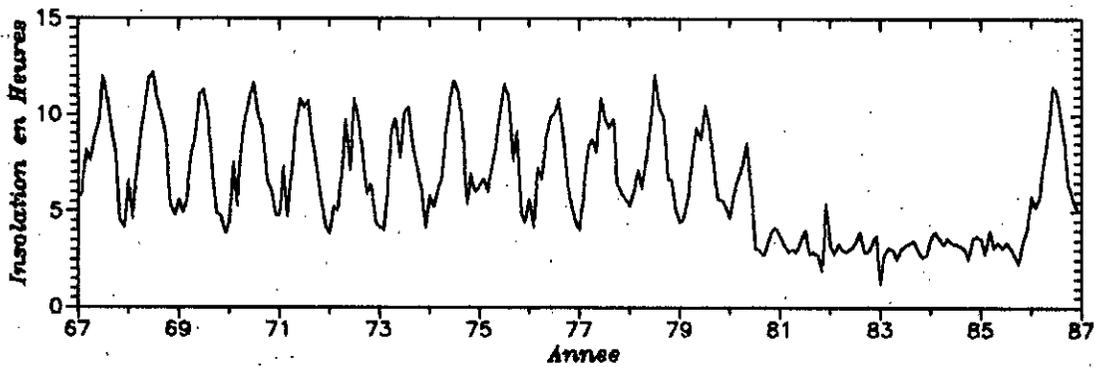


Figure A7.3

Variations mensuelles de la durée d'insolation  
Période 1967-1986, Station d'ALGER

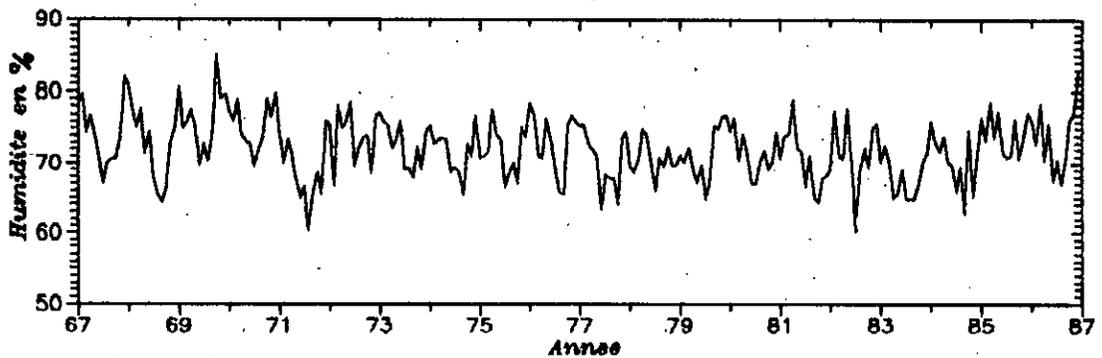
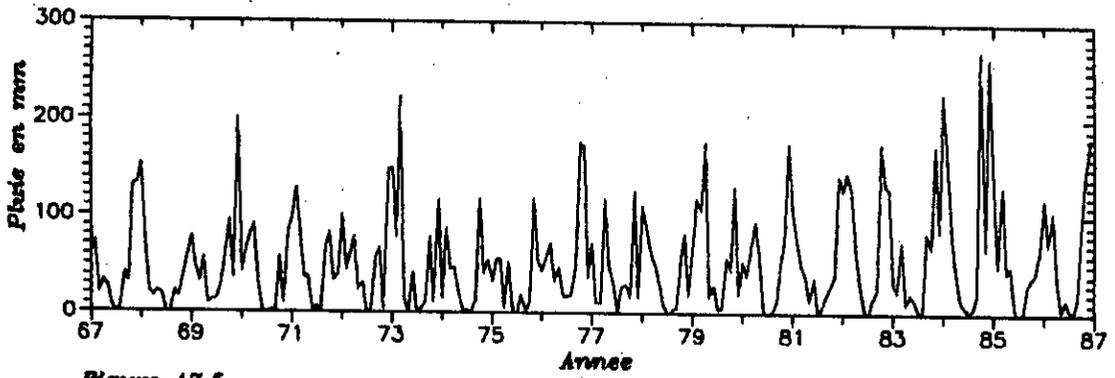
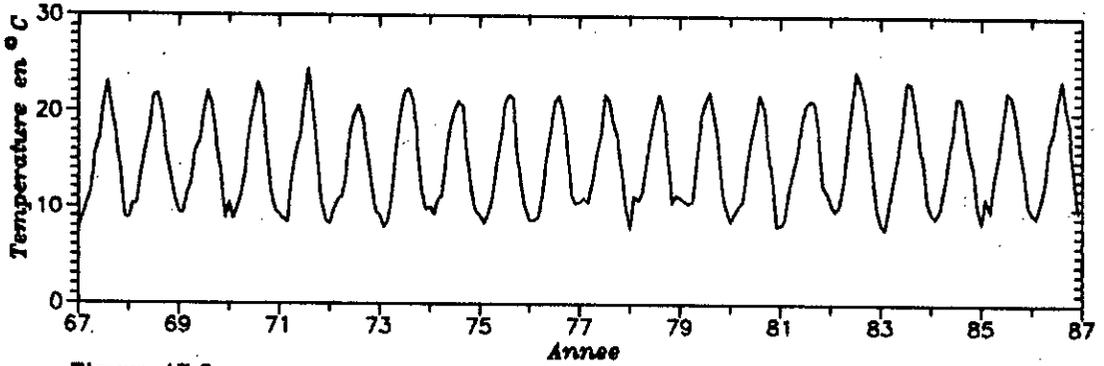


Figure A7.4

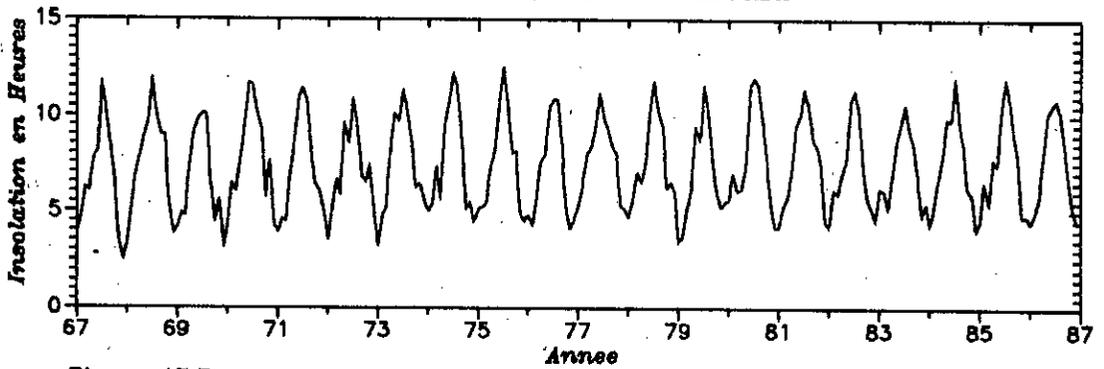
Variation de la moyenne mensuelle de l'humidité  
Période 1967-1986, Station d'ALGER



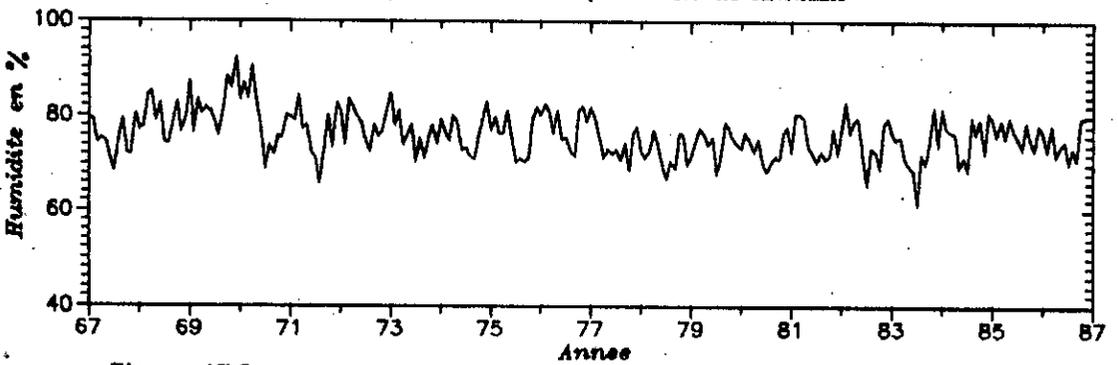
Quantité de pluie mensuelle  
Période 1967-1986, Station de ANNABA



Variation de la température moyenne mensuelle  
Période 1967-1986, Station de ANNABA



Variations mensuelles de la durée d'insolation  
Période 1967-1986, Station de ANNABA



Variation de la moyenne mensuelle de l'humidité  
Période 1967-1986, Station de ANNABA

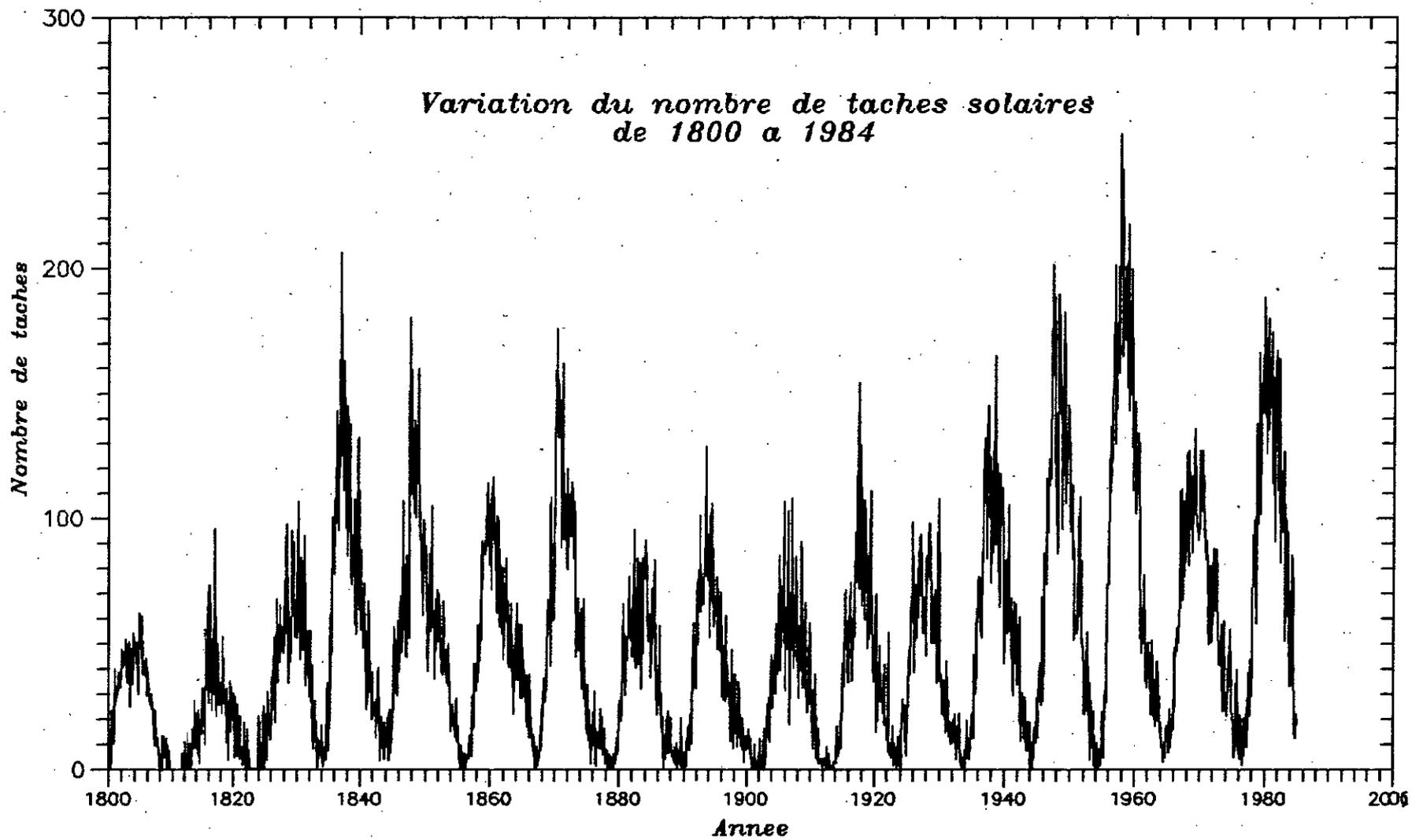
## ANNEXE A8

### L'ACTIVITE SOLAIRE (47)

Le soleil nous semble d'une ardeur constante. Pourtant si l'on regarde de près, on s'aperçoit que c'est une étoile variable. En effet, des phénomènes complexes commandent l'activité de notre astre du jour: réactions de fusion thermonucléaires, rotation des couches externes ...etc), mais, d'une façon générale, des observations systématiques révèlent que les régions d'activité intense sont plus froides que les zones environnantes; elles apparaissent comme des zones plus sombres sur un fond plus chaud et donc plus brillant du disque solaire. Ces régions d'activité solaire sont ce que l'on nomme des taches solaires (SunSpots).

Le cycle solaire se manifeste notamment par un changement du nombre de taches à la surface visible du soleil. Depuis 1715 ce nombre a oscillé sans interruption (voir figure ci après). En 1843 Henrich SCHWABE remarqua que le cycle a une périodicité de 10 ans. En 1848 le Directeur de l'observatoire de Zurich, Rudolf WOLF entreprit une recension quotidienne des taches solaires. Simultanément il rechercha l'évolution du nombre de ces taches durant les 150 dernières années, et conclut que la période du cycle étant égale à 11.2 ans. Depuis 1848, la période a varié entre 10 et 12 ans, malgré d'importantes irrégularités.

L'étude de ces variations est importante car notre planète s'adapte à l'activité solaire. Les climatologues cherchent aujourd'hui à préciser les relations entre les phénomènes à l'intérieur du soleil et leurs conséquences sur la terre.



## MANUEL D'UTILISATION DU LOGICIEL ESTIMA

## Présentation

ESTIMA est un logiciel pour l'estimation de la densité spectrale des signaux réels, que ce soit simulés par ordinateur ou introduits par l'utilisateur. Ces signaux peuvent être simples (1 seul canal) ou formé de plusieurs canaux. Les méthodes d'estimation implémentées sont basées essentiellement sur la transformation de FOURIER pour les méthodes classiques, et sur la modélisation Autorégressive en ce qui concerne les méthodes paramétriques.

## Méthodes classiques

Périodogrammes  
Corrélogramme  
Méthode hybride

## Méthodes Paramétriques

Méthode de l'Autocorrélation  
Méthode de la covariance  
Méthode de la cov. modifiée  
Méthode de Burg (MEM)

La particularité du logiciel ESTIMA est qu'il intègre un grapheur, ce qui permet de représenter sous forme graphique, les données saisies et les résultats de l'estimation.

Les programmes du logiciel sont modulaires, ils correspondent chacun d'entre eux à une opération spécifique relative aux deux types de signaux (simple canal et multicanaux):

DATASIM : Simulation et génération de signaux (1 canal).  
ESCALE : Estimation de la DSP par les méthodes classiques.  
ESPACE : Estimation de DSP par les méthodes Paramétriques  
MCSIMULA: Simulation et génération de signaux multicanaux.  
MCDSP : Estimation de la DSP des signaux multicanaux.

La convivialité de ces programmes les rendent de telle sorte que même ceux qui ont des connaissances très faibles en informatique peuvent les utiliser sans faire appel à ce manuel. Néanmoins dans ce qui suit, nous allons donner les principales commandes.

## INSTALLATION

Les utilitaires ESTIMA fonctionnent sur tout système PC-XT/AT IBM et compatible avec:

- Une mémoire minimale de 512 Ko
- Un Moniteur VGA, EGA, CGA, ou Hercule
- Une Imprimante compatible avec l'IBM Proprinter ou EPSON.
- Un disque dur ou un système avec deux lecteurs de disquettes
- Un coprocesseur 80X87 sera le bienvenu, mais son absence n'entrave pas le bon déroulement des programmes (Emulation du Coprocesseur).

En plus des utilitaires ESTIMA et de quelques fichiers de données, ESTIMA contient divers programmes pilotes pour pouvoir s'adapter aux différentes cartes graphiques. La liste complète des fichiers ESTIMA est comme suit:

- DATASIM.EXE : Simulation et génération de signaux (1 canal).  
ESCALE.EXE : Estimation de la densité Spectrale de puissance  
ESCALE.OVR par les méthodes classiques.  
ESPACE.EXE : Estimation de la densité Spectrale de puissance  
ESPACE.OVR par les méthodes Paramétriques.  
MCSIMULA.EXE : Simulation et génération de signaux  
multicanaux.  
MCDSP.EXE : Estimation de la densité Spectrale de puissance  
MCDSP.OVR des signaux multicanaux.  
ALGRAPH.EXE Module graphique

### Fichiers pilotes de cartes graphiques:

ATT	BGI	CGA	BGI	EGAVGA	BGI
HERC	BGI	IBM8514	BGI	PC3270	BGI

Fichiers Polices de caractères:

GOTH.CHR

SANS.CHR

LITT.CHR

TRIP.CHR

Unités et bibliothèques de fonctions:

ALCRT.TPU	ALEGAVGA.TPU	ALKEYS.TPU	ALPRINT.TPU
ALSMALL.TPU	ARMAPSD.TPU	ESCALELB.TPU	ESCOVR.TPU
ESPACELB.TPU	ESPOIRLB.TPU	ESPOVR.TPU	GRAPH.TPU
GRAPHNET.TPU	GRAPHSYS.TPU	IODSPLIB.TPU	MCALMAT.TPU
MCKERNEL.TPU	MCLULEV.TPU	MCMATDEF.TPU	MCNDATA.TPU
MCOVR.TPU	MCRSEDIT.TPU	NCOMPLEX.TPU	SIMOVR.TPU

Fichiers de données:

D44ALAM DAT Modèle ARMA(4,4);  
DAR10AD DAT Modèle AR(10);  
DAR4KAY DAT Modèle AR(4);  
DMCAR1 DAT Modèle AR(1) à 2 canaux;  
DATACON DAT Phonèmes de la langue ARABE: Consonnes.  
DATAVOY DAT Phonèmes de la langue ARABE: Voyelles.  
MALG5288 DAT Observations Météo. Station d'ALGER (1952-1986)  
MANN6788 DAT Observations Météo. Station d'ANNABA(1967-1986)  
SUNSPOTS DAT Observations du nombre de taches solaires de  
1850 à 1984.

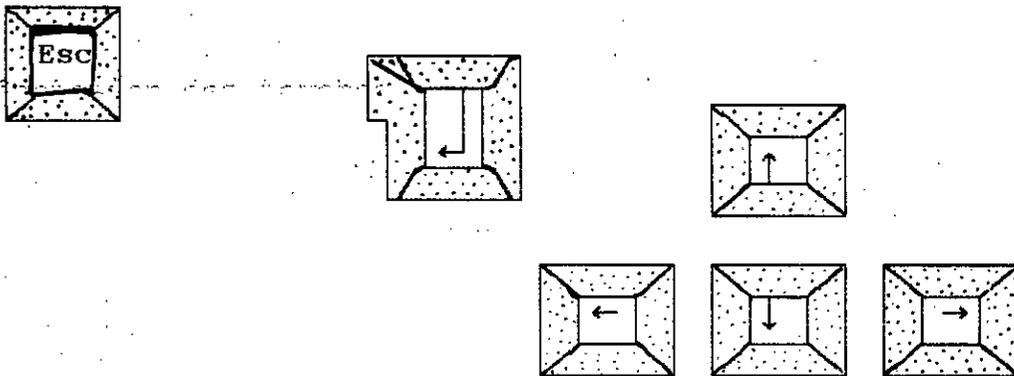
Fichier Modèle ARMA + Sinusoides:

MAR10AD.PZP	MAR4ALAM.PZP	MAR8AD.PZP
M04MA.ABC	M44ALAM.ABC	MAR10AD.ABC
MAR4KAY.ABC	P2S10DBA.SIN	P3S30DB.SIN
MCMARPLE.ABC		

## Affectation des touches

Le logiciel est écrit de manière conversationnelle; les principales commandes sont accessibles à travers les menus déroulants. Le passage d'une commande à l'autre se fait à l'aide du curseur (touches ← et →), ou encore en frappant la lettre en surbrillance du menu choisi. Le menu se déroule, et si nécessaire vous demande les valeurs des paramètres à entrer sous forme claire et compréhensible.

La validation de votre choix ou la prise en considération par le calculateur des paramètres se fait à l'aide du Retour charriot (↵) ou la touche (↓); curseur vers le bas. L'annulation d'une commande se fait au moyen de la touche Echapement (ESC) ou bien la touche (↑) curseur vers le haut.

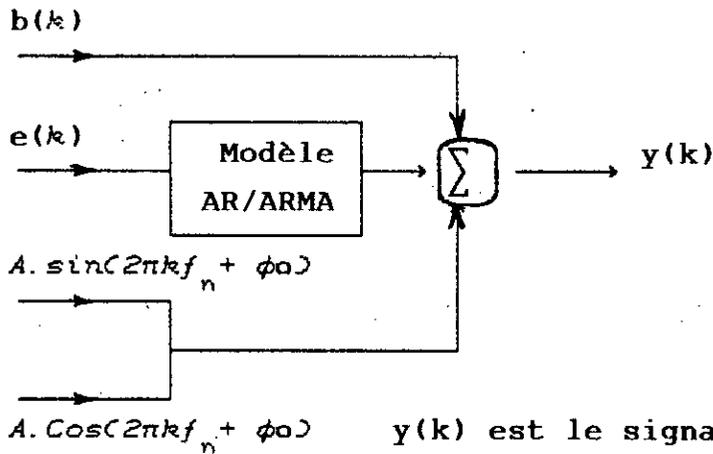


## Les Menus

Nous allons voir sommairement les différents menus relatifs à chaque programme, en donnant quelques précisions quand il est nécessaire.

### Simulation de processus et Génération de données

Le Programme DATASIM permet de générer des données selon le schéma synoptique suivant:



$y(k)$  est le signal généré  
 $e(k)$  est un bruit blanc  
d'excitation  
 $b(k)$  est un bruit additif

Le signal  $y(k)$  peut être le résultat d'une combinaison des trois type de signaux:

- Sinusoïdes (fréquences pures).
- Bruit blanc (Gaussien ou Uniforme)
- Signaux recueillis à la sortie d'un processus AR/ARMA excité par un bruit blanc.

Les sinusoïdes sont générées par l'une des deux expressions:

$$y(k) = A \cdot \sin(2\pi k f_n + \phi_0); \quad y(k) = A \cdot \cos(2\pi k f_n + \phi_0)$$

où  $A$  est l'amplitude de la sinusoïde,

$f_n$  est la fréquence normalisée variant dans l'intervalle  $-0.5, \dots, 0.5$  et définie par  $f_n = f/f_0$ ,  $f_0$  est la fréquence d'échantillonnage.

$\phi_0$  est la phase initiale en radians.

Ces paramètres peuvent être lus à partir d'un fichier conforme au format de l'exemple suivant:

N° Sinusoïde	Fréquence	Amplitude	Phase	Validation
1	0.195313	1.41	0.0	1
2	0.333399	1.41	0.0	1
0	0.1	Variance du bruit		

Quand  $y(k)$  est généré par un processus AR ou ARMA excité par un bruit blanc, les paramètres à spécifier sont les suivants:  
 1-Type du bruit blanc d'excitation : Gaussien ou Uniforme. La variance étant toujours égale à l'unité ( $\sigma_e^2=1$ ).

2-Les paramètres du modèle AR/ARMA

-Pour un processus AR d'ordre  $p$ ,  $y(k)$  est donné par l'équation:  $y(k) = -\alpha_1 y(k-1) - \alpha_2 y(k-2) - \dots - \alpha_p y(k-p) + e(k)$

les paramètres requis sont : l'ordre  $P$  et les coefficients  $\alpha_1, \alpha_2 \dots \alpha_p$  du modèle.

-Pour un processus ARMA  $y(k)$  est donné par l'équation:

$$y(k) = -\sum_{i=1}^P \alpha_i y(k-i) + \sum_{i=0}^Q b_i e(k-i)$$

les paramètres convenables sont :

l'ordre  $P$  et les coefficients  $\alpha_i$  de la partie AR,

l'ordre  $Q$  et les coefficients  $b_i$  de la partie MA.

Le programme nous offre la possibilité de caractériser le modèle par ses pôles et zéros. Ainsi il suffit de spécifier les  $P$  pôles /  $Q$  zéros du modèle, et le programme se charge du calcul des coefficients  $\alpha_i$  et  $b_i$ .

Exemples de Fichiers de paramètres ARMA:

Modèle AR(8):

8 Pôles

0 Aucun Zéro

r	$\theta(\pi)$
0.95	0.20
0.95	-0.20
0.97	0.30
0.97	-0.30
0.94	0.40
0.94	-0.40
0.97	0.60
0.97	-0.60

Modèle AR ( 8 )

8 Ordre AR

0 Ordre MA

Ai	Bi
1	1
-2.6588946591	0
5.0227781629	0
-6.3781332962	0
6.9886991296	0
-5.9252768189	0
4.2852054074	0
-2.0723614818	0
0.705975866	0

**ARMA(4,4): Exemple d'ALAM**

4 Ordre AR

4 Ordre MA

Ai	Bi
1	1
-3.3781	-2.3790
4.7367	2.9968
-3.2882	-2.2770
0.9509	0.9273

**AR(4): Exemple d'ALAM.**

4 Pôles

0 Aucun Zéro

Re	Im
0.7155	+0.6718
0.7155	-0.6718
0.9736	+0.1983
0.9736	-0.1983

Enfin un bruit blanc (Gaussien ou Uniforme) de variance  $\sigma^2 > 0$  peut se superposer au signal  $y(k)$  généré auparavant (quelques sinusoides et/ou signal AR/ARMA).

En ce qui concerne les signaux multicanaux, le schéma synoptique reste toujours valable. Néanmoins, les paramètres du modèle AR sont des matrices carrées d'ordre  $m$ , où  $m$  est le nombre de canaux. Le bruit blanc est un bruit vectoriel formé à partir de  $m$  bruits indépendants.

Pour les signaux AR/ARMA, une fois les coefficients sont connus, on peut facilement calculer la DSP idéale du processus en utilisant l'algorithme de Horner [47].

**Méthodes d'estimation:**

Comme les méthodes d'estimation ont été l'objet d'études faites aux chapitres II et IV, nous nous bornerons à présenter les différents menus et à résumer les différentes options.

Simulation de Processus Aléatoires

Ecran  
Principal  
du module  
DATASIM

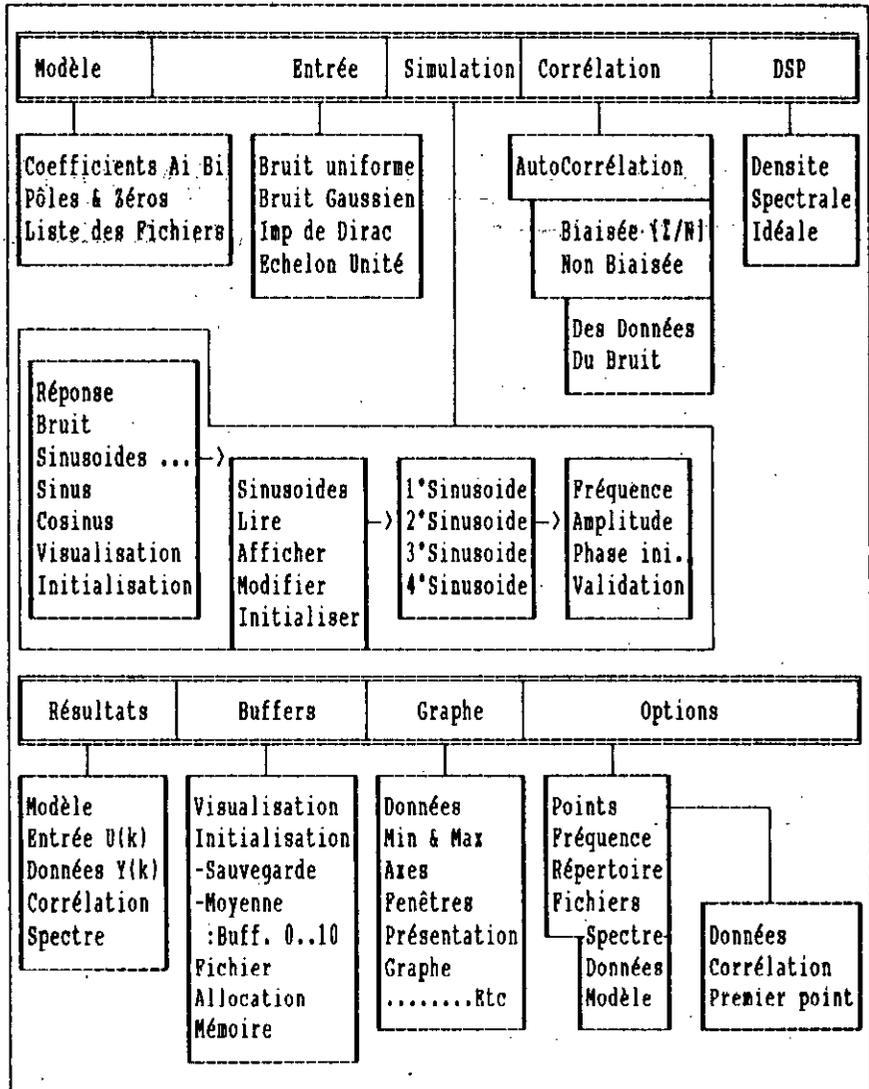
Modèle ARMA [ 1 , 0]. Entrée Bruit Gaussien Bruit Blanc ( $\sigma^2 = 1$ ) Signal/Bruit : 30 dB. f/échantillonnage 1 Hz Nombre de Points Données : 512 P.Point : 64 Corrélation: 128. Spectre : 200	Répertoire : C:\LAHCENE <hr/> Modèle : ARMA.ABC Données : DATA.DAT Spectre : DSPI.DAT Fichier : BUFFER.DAT
--	--

Modèle ARMA :  $H(z) = B(z)/A(z)$

$$H(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_m z^{-q}}{1 + a_1 z^{-1} + \dots + a_n z^{-p}}$$

Modèle Entrée Simulation Corrélation Dsp Résultats Buffers Graphe Options

Menu principal  
et sous menus  
du module  
DATASIM



Ecran  
Principal  
du module  
ESCALE

**ESTIMATION SPECTRALE PAR LES METHODES CLASSIQUES**

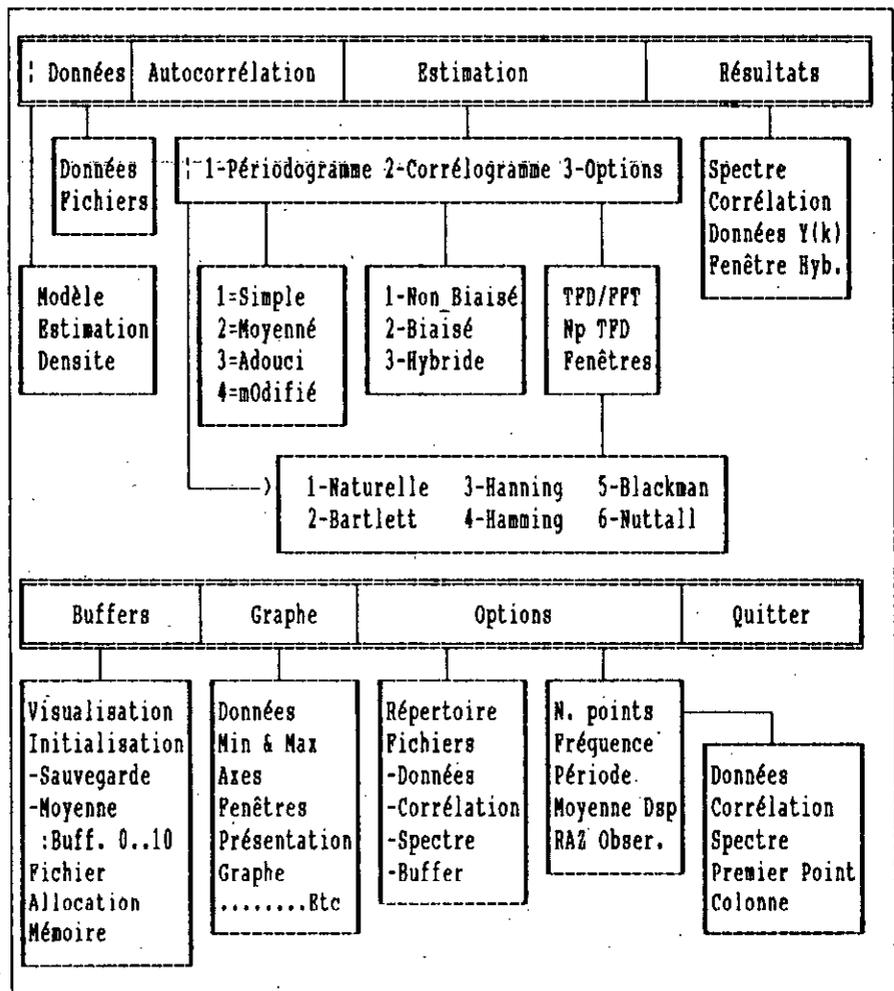
---

1-Périodogramme Simple (.....)    2-Périodogramme Moyenné (BARTELETT)  
 3-Périodogramme Adouci (DANIELLE)    4-Périodogramme Modifié (WELCH)  
 5-Corrélogramme (BLACKMAN & TUKEY)    6-Méthode Hybride

Trans. de FOURIER Rapide (PPT) Fenêtre Naturelle Nombre de sections : 1 Nombre de points Données : 512 Spectre : 256 Autocorrélation : 128 Observations : 0 P/échantillonnage : 1	Répertoire : C:\LAHCENE  Fichiers Données : DATA.ADT Colonne N° : 2 Premier point: 0 Spectre : DSP.DAT AutoCorrélation : AC.DAT Fichier : BUFFER.DAT
---	--

Données Autocorrélation Estimation Résultats Buffers Graphe Options

Menu principal  
et sous menus  
du module  
ESCALE



ESTIMATION SPECTRALE PAR LES METHODES PARAMETRIQUES

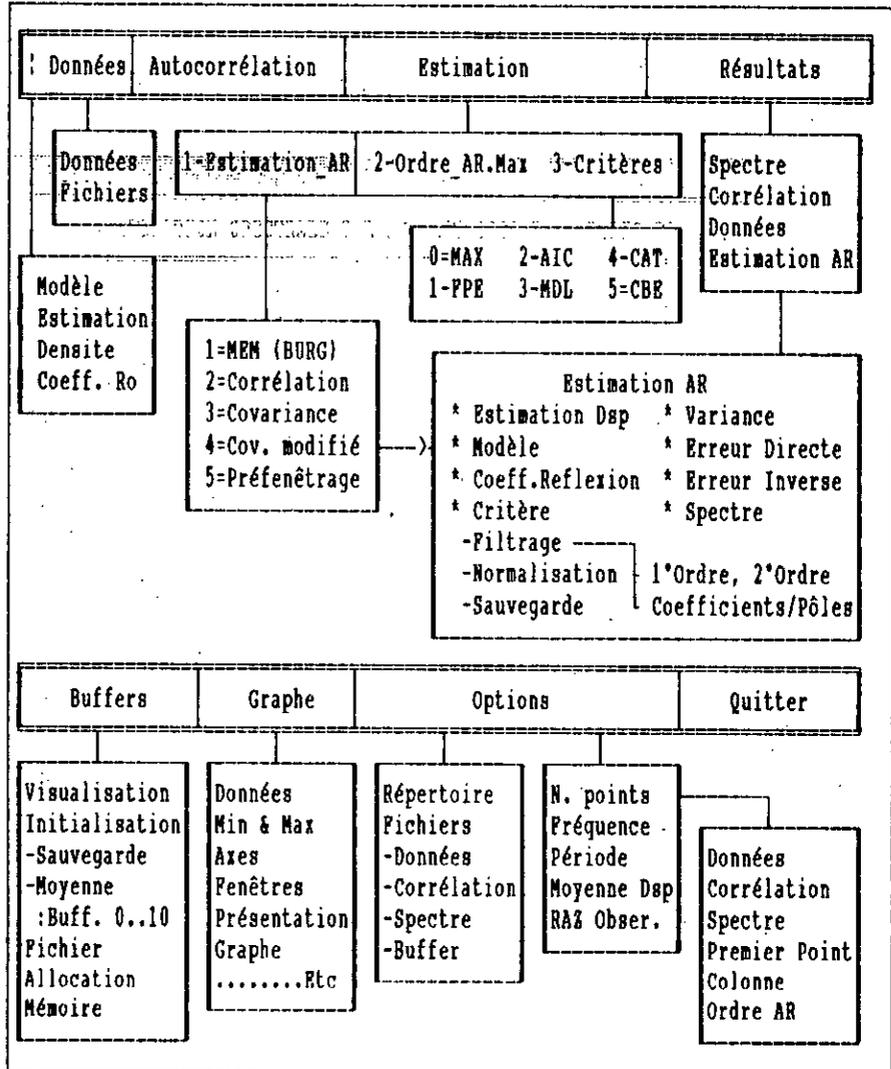
- 1-Maximum d'entropie (Burg)
- 2-Corrélation (Durbin)
- 3-Covariance /Covariance modifiée
- 4-Préfenêtrage

Ecran  
Principal  
du module  
ESPACE

Ordre du Modèle AR Ordre AR : 1 Ordre Max : 64 Critère N° : 5 Nombre de points Données : 252 Spectre : 200 Autocorrélation : 126 Observations : 0 P/échantillonnage : 1	Répertoire : C:\LANCENE Fichiers Données : E:DATA.DAT Colonne N° : 2 Premier point: 0 Spectre : DSP.DAT AutoCorrélation: AC.DAT Critère_AR : AR.DAT Buffer : BUFFER.DAT
--	---

Données Autocorrélation Estimation Résultats Buffers Graphe Options

Menu principal  
et sous menus  
du module  
ESPACE



Simulation & Génération de Signaux multicanaux

Ecran principal du module MCSIMULA

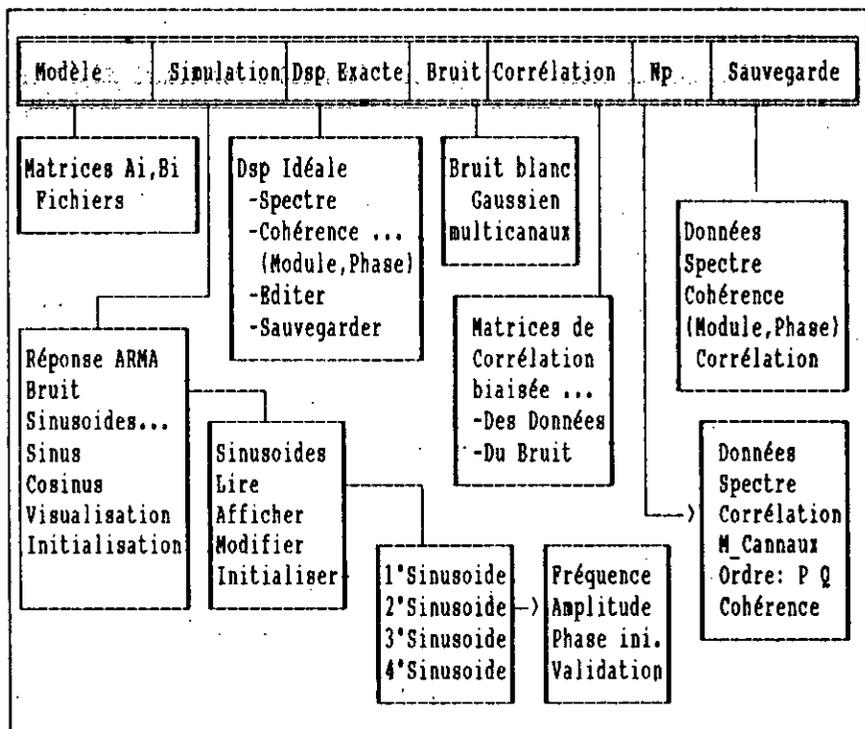
Bruit Blanc ( $\sigma^2 = 1$ )	Nombre de points
Modèle ARMA [ 10, 0].	Données : 800
Nombre de Canaux 4	Spectre : 256
Simulation ...	Répertoire C:\LANCENE
Densité spectrale Ideale	Modèle : MCARMA
Rapport (Signal/Bruit) 20 dB.	Données : MCdata
f/échantillonnage 1	Spectre : MCpsd

Modèle ARMA :  $H(z) = B(z)/A(z)$

$$H(z) = \frac{B_0 + B_1 z^{-1} + \dots + B_m z^{-m}}{1 + A_1 z^{-1} + \dots + A_n z^{-n}}$$

Modèle Dsp\_Exacte Bruit Simulation Corrélation (S)auvegarde Quitter

Menu principal et sous menus du module MCSIMULA



Analyse Spectrale des signaux multicanaux

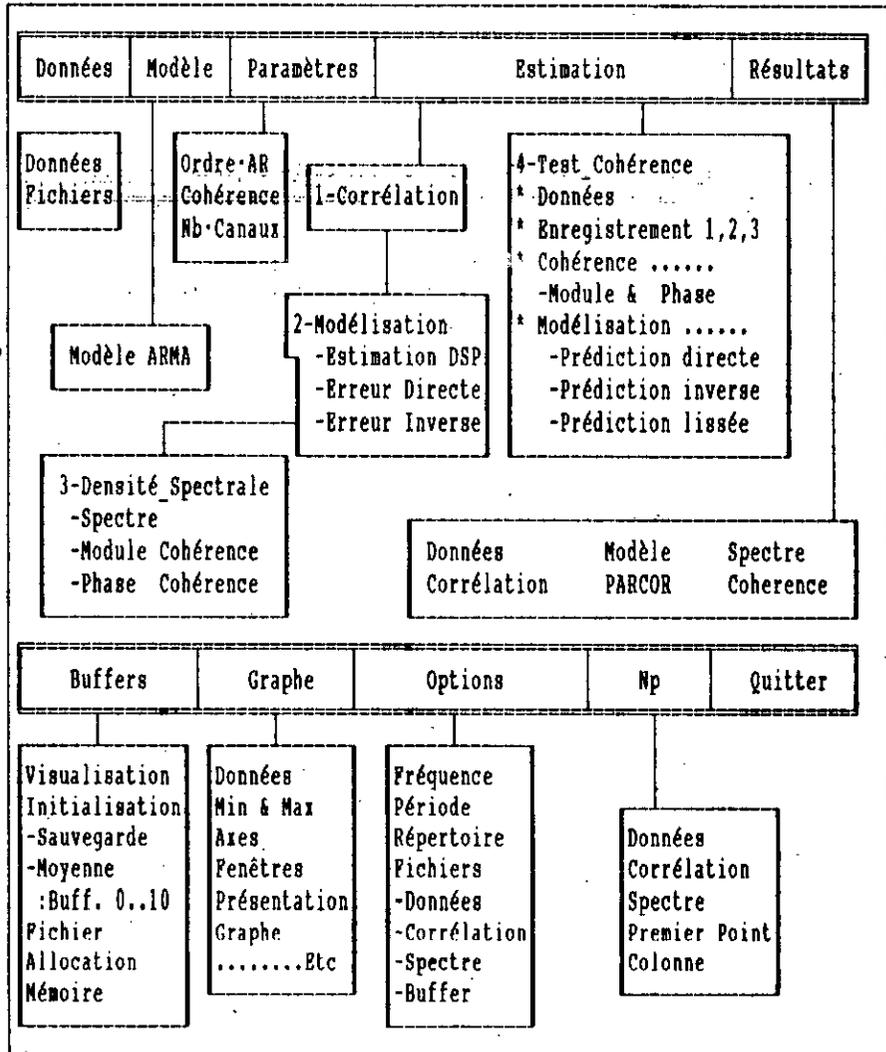
Ecran  
Principal  
du module  
MCDSP

Nombre de Canaux : 4	^Répertoire : C:\LABCENE
Modèle AR d'ordre : 1	^Fichiers
Premier point : 50	Données : MCDATA.DAT
Nombre de points	Spectre : MCDSP.DAT
Données : 256	Modèle : NCAR.DAT
Spectre : 100	Résultats : BUFFER.DAT
Corrélation : 100	
f/échantillonnage : 1	

0	1° Enregistrement	100
100	2° Enregistrement	200
200	3° Enregistrement	300

Données Modèle Paramètres Estimation Résultats Buffers Graphe Np Quitter

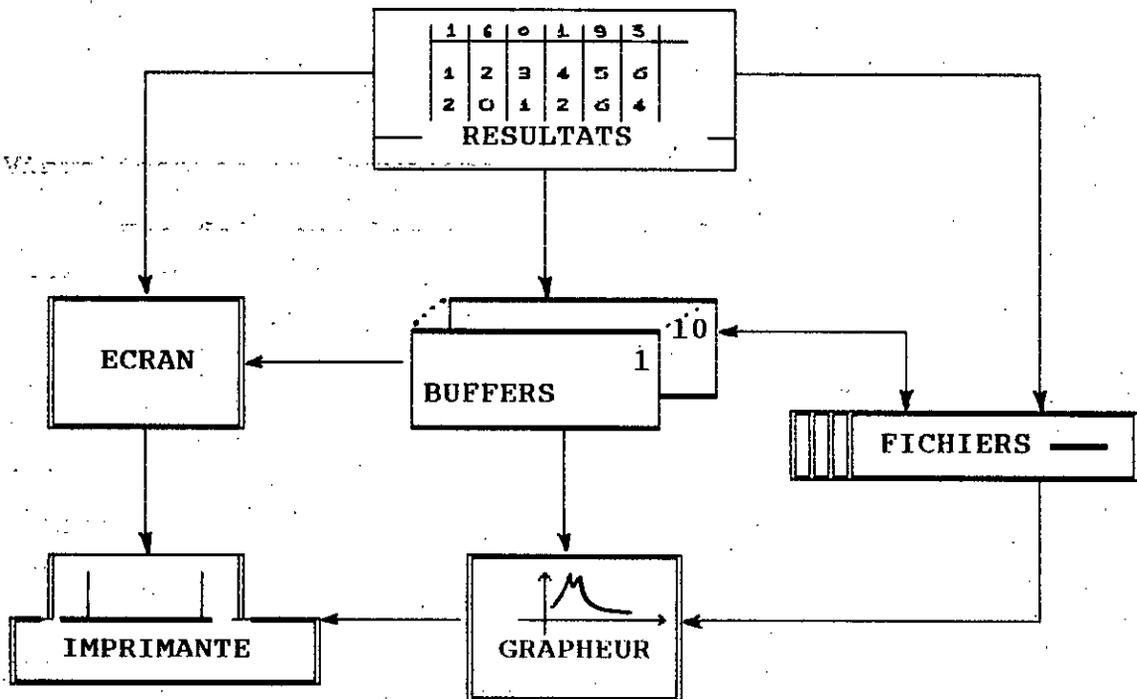
Menu principal  
et sous menus  
du module MCDSP



## Visualisation et Sauvegarde

Une fois que les calculs sont faits et les résultats sont prêts, trois possibilités nous sont offertes:

- Visualiser les résultats sur écran sous forme de tableaux.
- Sauvegarder les tableaux de valeurs dans la RAM du calculateur (Buffer). Cette opération peut s'avérer nécessaire soit pour visualiser les résultats de plusieurs estimation sur un même écran. Soit comme étape intermédiaire entre la partie calcul et la partie graphique du logiciel.
- Enregistrer les résultats sur un support magnétique (disque dur, disquettes ...). On doit dans ce cas introduire le nom du fichier, et éventuellement le chemin complet s'il s'agit de l'enregistrement dans un autre répertoire.



## Présentation graphique des résultats

On dit souvent qu'un graphe est nettement plus "parlant" qu'une série de valeurs. Alors il nous a semblé nécessaire que le logiciel ESTIMA dispose de son propre grapheur, et c'est ainsi que nous avons développé ALGRAPH.

ALGRAPH permet de représenter sous forme graphique les données saisies, les DSP estimées, et toute autre série de valeurs ( Erreurs de prédiction, Critère de sélection de l'ordre ... etc).

Pour créer un graphe, on doit seulement spécifier la source des données, soit un fichier disque ou un buffer RAM. ALGRAPH dessine automatiquement la courbe à l'écran. Néanmoins, vous pouvez personnaliser les graphiques en choisissant vous même la configuration: les colonnes à représenter, les valeurs minimale et maximale des axes, l'épaisseur des traits, les couleurs des courbes, la police de caractères, la fenêtre du graphe, ...etc.

Le graphe tel qu'il est proposé par ALGRAPH est déjà très représentatif des valeurs qu'il représente. Avec quelques ajouts, tels qu'un beau titre, un petit commentaire, et votre graphe peut être qualifié de "professionnel". Il y a tellement de variations possibles dans la représentation des graphes que l'on est seulement limiter par son imagination. Il ne peut pas être question de donner des exemples de tous les cas possibles. Nous Vous laissons le soin de découvrir le formidable potentiel qui se cache dans cette partie du logiciel.

#### Résumé des principales commandes ALGRAPH

**DONNEES:** Source des données: Fichier disque ou Buffers

**COLONNES:** Colonnes à représenter :

\* Colonne des Y : de la 1<sup>o</sup> Colonne à la 7<sup>o</sup> colonne.

\* Colonne des X : 0 => Aucune Colonne (données en y(k)).

**AXES:** Choix du format des Axes

valeurs Min & Max, Echelle linéaire ou logarithmique, Grille, Titres des axes...etc.

**FENETRES:** Position du graphe à l'écran: On a le choix entre la fenêtre par défaut, la fenêtre utilisateur et l'une des fenêtres prédéfinies (1, 2, 4, ou 6 par écran )

ALGRAPH permet de représenter sous forme graphique les données saisies, les DSP estimées, et toute autre série de valeurs ( Erreurs de prédiction, Critère de sélection de l'ordre ... etc).

Pour créer un graphe, on doit seulement spécifier la source des données, soit un fichier disque ou un buffer RAM. ALGRAPH dessine automatiquement la courbe à l'écran. Néanmoins, vous pouvez personnaliser les graphiques en choisissant vous même la configuration: les colonnes à représenter, les valeurs minimale et maximale des axes, l'épaisseur des traits, les couleurs des courbes, la police de caractères, la fenêtre du graphe, ...etc.

Le graphe tel qu'il est proposé par ALGRAPH est déjà très représentatif des valeurs qu'il représente. Avec quelques ajouts, tels qu'un beau titre, un petit commentaire, et votre graphe peut être qualifié de "professionnel". Il y a tellement de variations possibles dans la représentation des graphes que l'on est seulement limité par son imagination. Il ne peut pas être question de donner des exemples de tous les cas possibles. Nous Vous laissons le soin de découvrir le formidable potentiel qui se cache dans cette partie du logiciel.

#### Résumé des principales commandes ALGRAPH

**DONNEES:** Source des données: Fichier disque ou Buffers

**COLONNES:** Colonnes à représenter :

\* Colonne des Y : de la 1<sup>o</sup> Colonne à la 7<sup>o</sup> colonne.

\* Colonne des X : 0 => Aucune Colonne (données en y(k)).

**AXES:** Choix du format des Axes

valeurs Min & Max, Echelle linéaire ou logarithmique, Grille, Titres des axes...etc.

**FENETRES:** Position du graphe à l'écran: On a le choix entre la fenêtre par défaut, la fenêtre utilisateur et l'une des fenêtres prédéfinies (1, 2, 4 , ou 6 par écran )

F1	F2	F3	F4	F5
<i>Données</i>	<i>Colonnes</i>	<i>Axes</i>	<i>Fenêtres</i>	<i>Graphe</i>
1-Fichier	1-Col .des X	1-Axe des X	1-optimale	1-Graphe
2-Buffer	2-Col .des Y	2-Axe des Y	2-1,2,4,6 /Ecran	2-Couleur
3-Coordonnées		4-Epaisseur	5-Lignes	6-Symboles
F6	F7	F8	F9	F10
<i>Effacer</i>	<i>Textes</i>	<i>Imprimer</i>	<i>Options</i>	<i>Quitter</i>
1-l'Ecran	1-Saisie	1-Imp-Ecran	1-Répertoire	1-Exit
2-Une Fenêtre	2-Afficher	2-Imprimer	2-Dos Shell	
3-Le Graphe	3-Effacer	3-Options	3-Couleurs	2-Stop

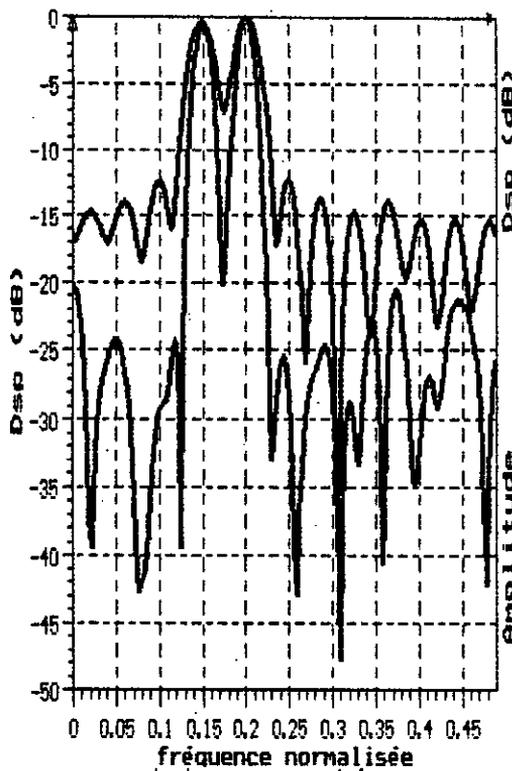
**GRAPHE** : Plusieurs options nous sont offertes:

- \* Choix de la couleur de la courbe [0 .. 15].
- \* Style et épaisseur du tracé des lignes.
- \* Symboles: Motifs à utiliser pour représenter les points sur écran (14 symboles sont disponibles \*,+,.,□,-)..etc.
- \* Afficher les coordonnées réelles des points.

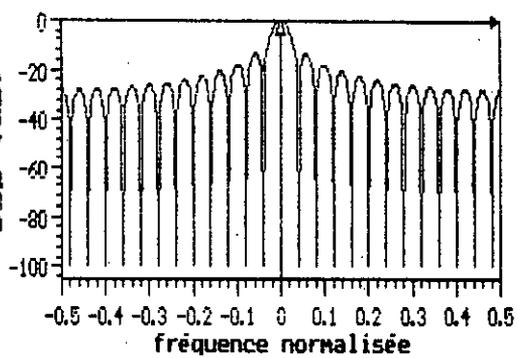
**EFFACER**: Tout l'écran, Seulement le graphe, ou une fenêtre bien déterminée (coordonnées de deux points diagonaux ).

**TEXTES**: Mêmes la police de caractères, la taille, la direction, la justification du texte peuvent être modifiées:

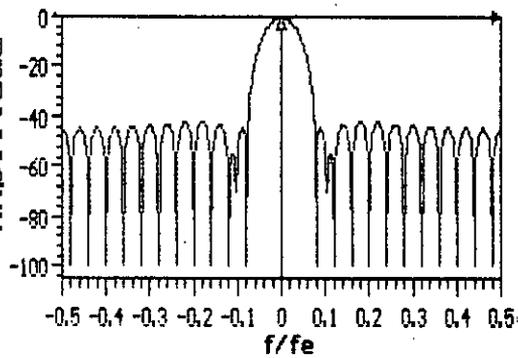
- \* Fonts: Défaut, Triplex, Petite, Bâtonnée, et Gothique.
- \* Direction: Horizontale /Verticale
- \* Taille: Facteur d'agrandissement de 1 à 9.
- \* Justification: A-gauche/droite, Vers le bas/haut, Centré et la fenêtre définie par l'utilisateur



lobes secondaires  
des méthodes classiques

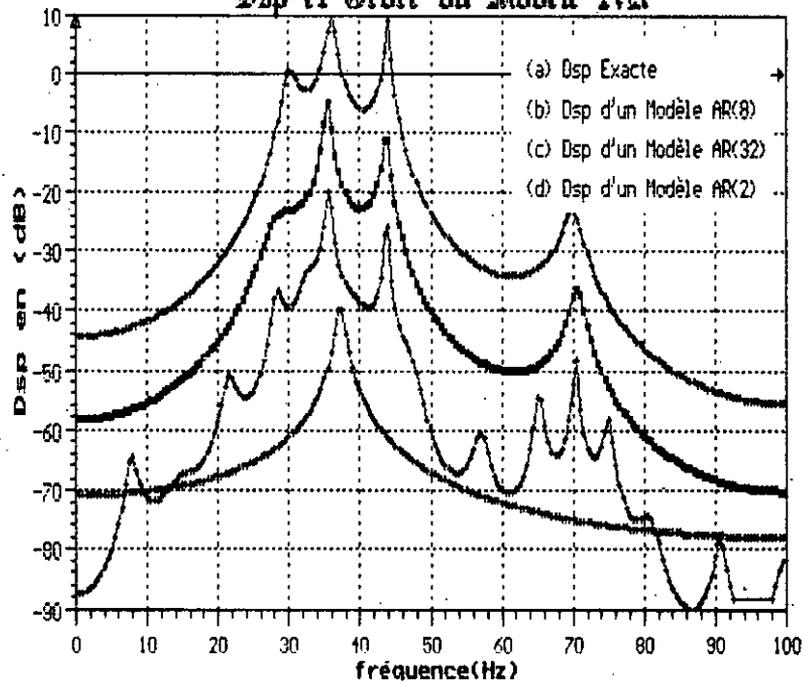


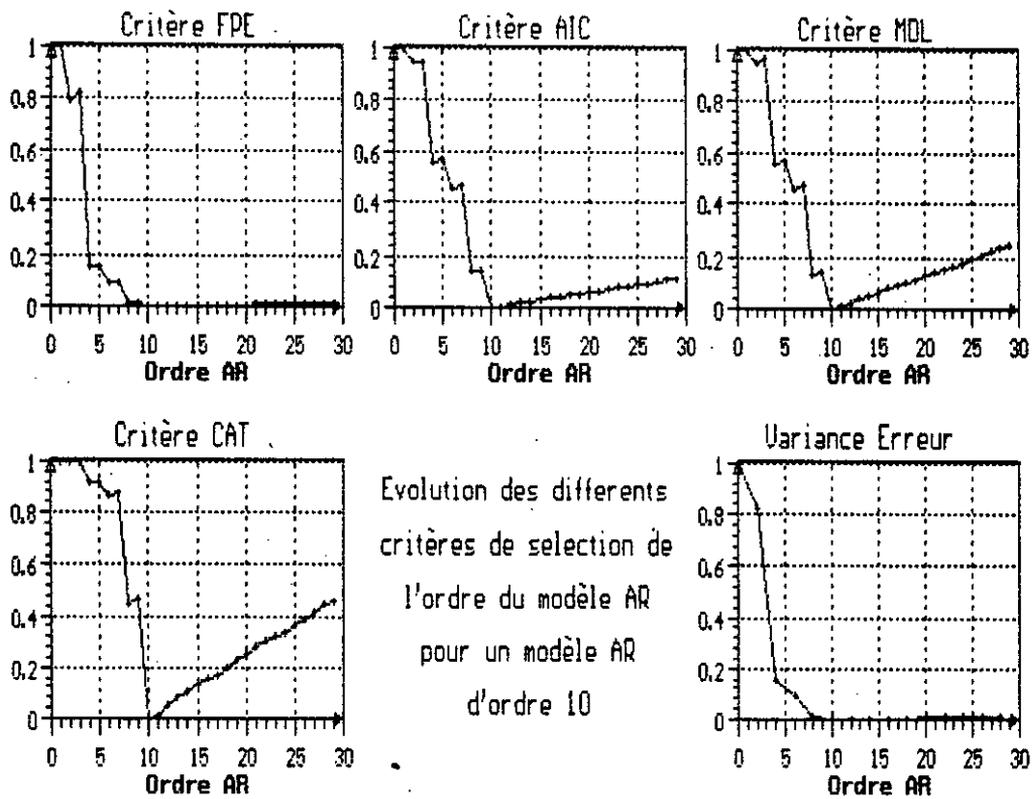
Spectre de la fenêtre Rectangulaire



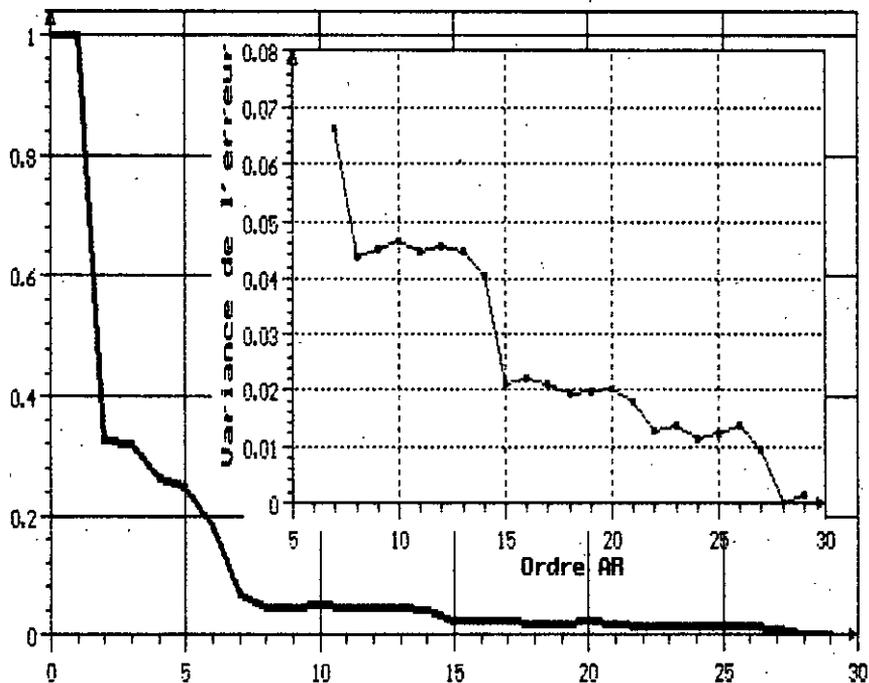
Spectre de la fenêtre de HANNING

### Dsp et Ordre du Modèle AR





### Variance de l'erreur de prédiction



Variance de l'erreur de prédiction d'un modèle ARMA(4, 4) en fonction de l'ordre AR

**IMPRIMER:** On a le choix entre l'impression écran avec DOS, [GRAPHICS+Touche Imp-Ecran] ou avec l'unité ALPRINT. Dans ce dernier cas on agit sur la commande marge on peut déplacer la feuille écran par rapport au bord de la feuille papier. Comme on peut modifier les dimensions du graphe à l'aide des commandes Echelle et Mode d'impression (densité).

**OPTIONS:** Enfin on trouve d'autres commandes telles

- \* Répertoire: Chemin d'accès au répertoire de données.
- \* DOS-Shell: Exécuter une commande Dos sans quitter ALGRAPH.
- \* Couleur: Choix des couleurs du graphe (Axes, Cadre, Grille), et des menus ( Barre, Commande, ... etc). Les goûts et les couleurs ne ...

### Les FICHIERS

Les fichiers constituent un élément important du logiciel ESTIMA. Lors de l'enregistrement des résultats, il y a création de fichiers, et il y a lecture de fichier lors de chaque saisie de données d'estimation. Dans ce qui suit, nous donnons des exemples de format des fichiers lus et créés par ESTIMA.

#### Fichiers de données

D o n n é e s:Exemple de KAY	
k	.....
Y(k)	.....
Modèle AR d'ordre 4	.....
0	0.9911174859
....	.....
1023	-2.9189127260

Les quatres premières lignes sont réservées. Debut des données à la 5 <sup>ème</sup> ligne.
---



II) Fichiers \*.PZP (Position des pôles & des zéros):

Modèle ARMA(3,1)	
3 Pôles	
2 Zéros	
r	$\theta(\pi)$
0.9	0.3
0.9	-0.3
0.95	0.7
0.7	0.5

AR(4): Exemple d' ALAM	
4 Pôles	
0 Aucun Zéro	
Re	Im
0.7155	+0.6718
0.7155	-0.6718
0.9736	+0.1983
0.9736	-0.1983

III) Fichiers \*.ABC (Matrices A B ):

Modèle A(1) avec 2 canaux	
2	1 0
Matrice A0	
1	0
0	1
Matrice A1	
-0.85	0.75
-0.65	-0.55
Matrice B1	
1	0
0	1

nombre de canaux = 2  
 Ordre AR = 1  
 Ordre MA = 0

Fichiers Résultats:

Densité Spectrale de puissance			
fréquence normalisée			
Dsp			
fn=f/fe	S(fn)/Max	S(fn) en db	S(fn)
0.0025125628	0.0004190116	-33.777739444	9.7705435821
0.0050251256	0.000421223	-33.754878881	9.8221097506
0.0075376884	0.000424941	-33.716714035	9.9088046674
.....	.....	.....	.....
0.108040201	0.5536754511	-2.5674473184	12910.644977
0.1105527638	1	0	23318.073704
0.1130653266	0.698120878	-1.5606937382	16278.834087
.....	.....	.....	.....
0.4949748744	0.0000003459	-64.609996586	0.0080666463
0.4974874372	0.0000003454	-64.616295202	0.0080549556
0.5	0.0000003451	64.620073096	0.0080479502

## ANNEXE A10

### UNITÉS ET BIBLIOTHÈQUES DE FONCTIONS DU LOGICIEL

Unité ALCRT  
Interface avec DOS et Ecran

```

Unit ALCRT;
Interface
Uses Crt,Dos,alkeys;
CONST
  HardAdd : Word = $B800; { hardware screen }
  SaveTxtAtt:Word=White; Color24:Word=Yellow;
  ErrColor:Word=16*Red+White+Blink;
  BgHcolor25:Word=7;selectcolor:Word=black;
  Bgcolor25:Word=Blue; Color25:Word=White;
  ComColor:Word=Yellow; MenuColor:Word=Cyan;
  NoCursor = $2000; DelCursor = $0A0A;
  InsCursor = $0A11; XlWind:Integer=48;
  DimMax = 512 ;
Type RealTab = Array[0..DimMax] of Real;
Procedure setcolor(Color:Word);
Function GetDostime:pathstr;
Procedure SetCursor(NewCursor : Word);
Function GetCursor : Word;
Function Blanks(Num : Integer) : PathStr;
Function StrOfChar(Long:Integer;car:Char):PathStr;
Function UpperCase(S : PathStr) : PathStr;
Procedure WriteXY(S : PathStr; Col, Row : Word);
Procedure Write24(MsgStr : PathStr);
Procedure Write25(MsgStr : PathStr);
Procedure Effacer24; Procedure Effacer25;
Procedure MsgErreur(S : PathStr);
Function OKLecture(Var S:PathStr;Legal:PathStr;
                  MaxLength:Word):Boolean;
Function OkLEFichier(RW:Boolean ;Msg:pathstr;Fext:ExtStr;
                    Var Fichier:PathStr):Boolean;
Function Ufilename(Fname:pathstr):pathstr;
Function Filename(PFname:pathstr):namestr;
Function BarMenu(MsgStr,ComStr :PathStr;Var counter:Integer):Char;
Function OKOrNO(S:pathstr):Boolean;
Procedure WriteMenu(S:String);
Function OkEntier(Var Number :Integer; Low,High:Integer):Boolean;
Function OkReal(Var Number : Real) : Boolean;
Function Read2Numbers(msg1,msg2,msg3:dirstr;Var Rp,Ip:Real):Boolean;
Function ReadReal(Var A:Real):Boolean;
Function RealString(Num : Real) : PathStr;
Function FRealString(Num : Real;L,D:Integer):PathStr;
Function Waitdelay(Var dt:Integer):Boolean;

```

```

Procedure Cadre(x1,y1,x2,y2:byte;Cadrefgnd,Cadrebkgnd:Byte);
Procedure Fcadre(x1,y1,x2,y2:byte;CadreBkgnd,Cadrefgnd:Byte);
Procedure FTcadre(x1,y1,x2,y2:byte;CadreBkgnd,
                  Cadrefgnd:Byte;Msg:PathStr);
Procedure ALcadre(Msg:PathStr;Bkgnd,fgnd:Byte;C17X:Char);
Procedure Files_in(Var Path:PathStr;X1,Y1,X2,Y2:Integer);
Procedure STxtScreen; Procedure PcPrintScreen;
Var      CrtCursor:Word; Selected:Integer; OldDirectory:Pathstr;
implementation

```

Unité IODSPLIB Fonctions d'Edition et de Sauvegarde
--

```
Unit IODSPLib;
```

```
Interface
```

```
Uses CRT,Dos,ALkeys,ALcrt;
```

```
Const Ordmax:Integer=100;
```

```
TabMax=128;
```

```
GNmax=512;
```

```
NBuff:Integer=1;
```

```
Bdelay:Integer=50;
```

```
DspShift:Real=0;
```

```
Type TypeCoeff =array[0 .. TabMax] of Real;
```

```
IntegerTab =array[0 .. GNmax ] of Integer;
```

```
TypeES =array[0 .. GNmax ] of Real;
```

```
PtrTypeES =TypeES;
```

```
Procedure T10RBuffers=array[1..10] of pathStr;
```

```
Procedure NP10RBuffers=array[1..10] of Integer;
```

```
RamBuffers =record
```

```
  XRAM,YRam :array[1..10] of pointer;
```

```
  TRam :T10RBuffers;
```

```
  NPRam :NP10RBuffers;
```

```
end;
```

```
Var
```

```
  Np,npds,NpAC,i,j,k :Integer;
```

```
  kb,NpMemRam,NmemRam :Integer;
```

```
  YK,Rk,Mdensites,Densite :TypeES;
```

```
  w,dw,fe,df :Real;
```

```
  DSmax,DSmin :Real;
```

```
  MyBuffers :RamBuffers;
```

```
  XYBuffer :RealTab ;
```

```
  TmemRam, BuffFile :Pathstr;
```

```
  EndBrepeat :Boolean;
```

```
  OkNewDsp,Okcorr :Boolean;
```

```
Function Horner(A:TypeCoeff;N:Integer;w:Real):Real ;
```

```
Procedure XCorrelation(Var X,Y,R:TypeES;N,M,Mode:Integer;msg:PathStr);
```

```
Procedure Correlation(Var X,Y,R:TypeES;N,M,Mode:Integer);
```

```
Procedure REad_AR(Var A,B:typecoeff;Var N,M:Integer;
```

```
                  Var MFICHIER:PathStr);
```

```
Procedure Save_AR(Var A:typeCoeff;N:Integer;
```

```
  titreMsg:pathstr;Var Mfichier:PathStr);
```

```
Procedure AR_Ro( a:typecoeff;Var ro:typecoeff;N:Integer);
```

```
Procedure Read_Data( Dfichier:PathStr ;Var Yk:typeES;
```

```
  Var np:Integer;pp,Colonne:Integer);
```

```
Procedure MinMaxDensite(DSPPTR:Pointer;npds:Integer;
```

```
  Var Dsmin,Dsmax:Real);
```

```

Procedure DbMinMaxDensite(DSPPTR:Pointer;npds:Integer;
                          Var Dsmin,Dsmax:Real);
Procedure Densite_Estimee(DspPtr,NdspPtr:Pointer;npds:Integer);
Procedure densite_en_db(DspPtr,DbPtr:Pointer; Var npds:Integer);
Procedure Densite_Normalisee(DspPtr,NdspPtr:Pointer;npds:Integer);
Procedure Dsp_Normalisee(DspPtr,NdspPtr:Pointer;npds:Integer);
Procedure Normalisation(DspPtr,NdspPtr:Pointer;npds:Integer);
Procedure Save_Spectre(Var Sfichier:PathStr;DspPtr:Pointer;
                      NPds:Integer; Xtitre,Ytitre,Titre:pathstr);
Procedure FreeMemory;
Procedure Save1TabInBuff(Aptr:Pointer;Np:Integer;
                        Tstr:PathStr;D:Integer);

Procedure edit1Tab(Aptr:Pointer;N,NF:Integer;
                  tablstr,TAB2str,titreMsg:pathstr;Okzero:Integer);
Procedure edit2Tab(Aptr,Bptr:Pointer;N,NF:Integer;
                  tablstr,TAB2str,titreMsg:pathstr;Okzero:Integer);

Procedure Save1Tab(Var Ffichier:PathStr;Aptr:Pointer;N,NF:Integer;
                  tablstr,TAB2str,titreMsg:pathstr;Okzero:Integer);
Procedure Save2Tab(Var Ffichier:PathStr;Aptr,Bptr:Pointer;
                  N,NF:Integer; tablstr,TAB2str,titreMsg:pathstr;Okzero:Integer);
Procedure fxfe2TAB(fe:Real;Npmax:Integer;fxfePtr:pointer);
Function SetXYbuff(Np,D:Integer;Var Buffptr:pointer):Boolean;
Procedure ResetXYbuff(Var Buffptr:pointer;Var Npb:Integer);
Procedure AllocationProc(Var MyRAMBuff:RAMBuffers);
Function BuffAllocation(Var MyBuff:RAMBuffers;Var nb:Integer;
                       Np:Integer;Qb:char;msg:pathstr):Boolean;
Function out2SBF(msg:pathstr;buffok:Boolean):Char;
Procedure saveInBuffer(XYptr:pointer; Var MyBuff:RAMBuffers;
                      NB,Np:Integer;Qb:char;msg:pathstr);
Procedure EditOrSave1Tab(Var Ffichier:PathStr;Aptr:Pointer;
                        Np,NF:Integer; Astr,Bstr,Tstr:PathStr;OK0:Integer);
Procedure EditOrSave2Tab(Var Ffichier:PathStr;Aptr,Bptr:Pointer;
                        Np,NF:Integer; Astr,Bstr,Tstr:PathStr;OK0:Integer);
Procedure BuffAverage(Var mybuffers:RAMbuffers);
Procedure BufferMenu(Var mybuffers:RAMbuffers);

```

Implementation

Unité ESCALELB  
Estimation Spectrale Classique (FFT,TFD)

```

unit EscaleLb;
Interface
Uses      CRT,Dos,ALkeys,ALcrt,IODspLib;
Const    TFstr:string[15]='Rapide (FFT)';
Var
OkFFT      :Boolean;
Type      Tab10=array[0..10] of Integer;
          TWind kOrm (Wdek,Wdem);
Const
Puissance2 :Tab10=(1,2,4,8,16,32,64,128,256,512,1024);
Nuttall:array[0..3] of Real=(0.3635819,0.4891775,
0.1365995,0.0106411);
FFtdelay   :Integer=0;DIstr:Char='D';

```

```

Var
  TFDFFT :Procedure(Var Y,Densite:typeES;T:Real;
  Nsections,Np,Fenetre :Integer;Var Npds:Integer);
  SFFT,P2,Nwindow :Integer;
  Key :Char;
  FFTAuto :Boolean;
  SomInv :TypeCoeff;
  Xr,Xi,Hwinm,Hwink :TypeES;
  OrdInv :IntegerTab;
Procedure SinCosTab(Np,SFFT:Integer;T:Real);
Function InitFFT(SFFT,N:Integer;T:Real;Var P2:Integer):Boolean;
Procedure FFT(N,P2,SFFT:Integer;T:Real;Var Xr,XI,densite:typeES);

Procedure Spectre_FFT(Var Y,Densite:typeES; T:Real;
  Nsection,Np,Fenetre:Integer;Var Npds:Integer);
Procedure Spectre_TFD(Var Y,Densite:typeES; T:Real;
  Nsections,Nptsection,Fenetre:Integer;Var Npds:Integer);
Procedure PeriodoGramme(Var Npoints:Integer;NSections,Fenetre:Integer;
  T:Real; Var Npds:Integer; Var Mdensite:TypeES);
Procedure PeriodoGramme_Simple(Var N,Npds:Integer;T:Real;
  Var Densite:TypeES);
Procedure PeriodoGramme_Moyenne(Var N:Integer;Nsections:Integer ;
  Var NpDs:Integer;T:Real;Var Mdensite:TypeES);
Procedure PeriodoGramme_Adouci(Var N:Integer;Fenetre:Integer ;
  Var Npds:Integer;T:Real;Var Mdensite:TypeES);
Procedure PeriodoGramme_Modifie(Var N:Integer;NSections,Fenetre:
  Integer;Var Npds:Integer; T:Real;Var Mdensite:TypeES);
Procedure Correlogramme(Var RV,Densite:typeES;T:Real;L,Fenetre:Integer;
  Var Npds:Integer);
Procedure Fenetre_Hybride(Var M:Integer;Fenetre1,Fenetre2:Integer);

```

#### Implementation

```

Var sinTab,CosTab :TypeES;
Procedure PeriodoGramme_Simple(Var N,Npds:Integer;T:Real;
  Var Densite:TypeES);
Var Fenetre:Integer;
  begin Fenetre:=1; PeriodoGramme(N,1,Fenetre,T,Npds,densite); end;
Procedure PeriodoGramme_Moyenne(Var N:Integer;Nsections:Integer ;
  Var NpDs:Integer;T:Real;Var Mdensite:TypeES);
  Var fenetre:Integer;
  begin fenetre:=1;
  PeriodoGramme(N,Nsections,Fenetre,T,Npds,Mdensite); end;
Procedure PeriodoGramme_Adouci(Var N:Integer;Fenetre:Integer ;
  Var Npds:Integer;T:Real; Var Mdensite:TypeES);
  begin PeriodoGramme(N,1,Fenetre,T,Npds,Mdensite); end;
Procedure PeriodoGramme_Modifie(Var N:Integer;NSections,Fenetre:Intege
  Var Npds:Integer; T:Real;Var Mdensite:TypeES);
  begin PeriodoGramme(N,NSections,Fenetre,T,Npds,Mdensite); end;

```

Unité ESPOIRLB  
Estimation Spectrale Parametrique  
Méthodes Autoregressives

```

unit EspoirLb;
Interface
Uses CRT,Dos,ALkeys,ALcrt,IODspLib;
Type AllCritTab=Array [1..10] of TypeCoeff;
Var
    Efm,Ebm :TypeES;
    CAi,CBi,CRO,R,Alpha,CSordre:TypeCoeff;
    Akaike,LFAC :Integer;
    OkNewRFRb :Boolean;
    CritereN :AllCritTab;
    CS1:Real;

    Coefffiltre:TypeCoeff;
    OrdFiltre:Integer;
Procedure Reponse(A,B:TypeCoeff;N,M:Integer;
    Var TabUk,TabYk:typeES;GN:Integer);
Procedure bruit_blanc(Var BB:typeES;GN:Integer;Var Varbb:Real);
Procedure Reponse_Bruit(A,B:TypeCoeff;N,M:Integer;
    Var Y:typeES;GN:Integer);
Function Ordre_AKAIKE(Akaike,m,n:Integer; alpha:Real):Real;
Function Burg(Var X,Fm,Bm:typeES;N,Mmax:Integer;
    Var ro,alpha,A:typecoeff;Var-ordAR:Integer):PathStr;
Procedure Durbin(Var R,Alpha,Ro,A:TypeCoeff;N,P:Integer);
Function Cholesky(Var A,B:TypeCoeff;N:Integer;EPS:Real):Boolean;
Function InitFACriterion(Var Np,L:Integer):Boolean;

```

Implementation

Extention aux signaux multicanaux  
Unité MCMatDef: Definition des types Matrices  
Unité MCKernel: Structure des données

```

Unit MCMatdef;
Interface
Const
    MchannelMax=4;
Type
    MCDim =1..MchannelMax;
    MCmatrice =array[MCDim,MCDim] of Real;
    DataMC =array[MCDim] of Real;

```

implementation

Unité MCKernel: Structure des données

```

Unit MCKernel;
Interface
uses MCmatDef,Dos;
{$IFDEF SIMULA} Const NdataMax =800;

```

```

{$ELSE} Const NdataMax =256;
{$ENDIF}
Const
  Mchannel2Max=8;
  NPsdMax =256;
  AROMax =20;   Rkmax= 100;
  I0:Integer =50;   Npds:Integer=200;

```

```

Type
  MCCDim      =1..Mchannel2Max;
  DataDim     =0..NdataMax;
  DspDim      =0..NPsdMax;
  OrdDim      =0..AROMax;
  Sdata       =array[DataDim] of Real;
  ITab        =array[OrdDim]  of Integer;
  MCVecteur   =Array[OrdDim]   of MCmatrice;
  MCRkVecteur=Array[0..Rkmax]  of MCmatrice;
  MCSpectrum  =Array[dspDim]   of MCmatrice;
  MCdata      =array[DataDim]  of DataMC;
  MMCdata     =array[Datadim, Mcdim] of Real;
  MCCphase    =record
    Msc, HPhase: Sdata;
    Cx, Cy      : Integer;
  end;

```

implémentation

```

{$ELSE} Const NdataMax =256;
{$ENDIF}

```

Extention aux signaux multicanaux  
Bibliothèque de fonctions mathématiques

Unité MCALMAT: Opérations sur des matrices

```

unit McAlmat;
interface
uses crt, dos, McMatdef;
Function trMat(Mata:Pointer;la:Integer):Real;
Procedure ZeroMat( Mata:Pointer;N,M:Integer);
Procedure ScaleMat( Mata,MatB:Pointer;N,M:Integer;S:Real);
Procedure AssignMat( Mata,MatB:Pointer; N,M:Integer);
Procedure IdentMat( Mata:Pointer; N:Integer);
Procedure XposeMat(Mata:Pointer;N,M:Integer; MatB:Pointer);
Procedure OutProd(Vx,Vy:Pointer; MatM:Pointer; N:Integer);
Procedure RReadMat(MatStr:namestr; Mat1:Pointer;lgn1, coll:Integer);
Procedure CReadMat(MatStr:namestr; Mat1:Pointer;lgn1, coll:Integer);
Procedure WriteMat(Mat1:Pointer;lgn1, coll:Integer);
Procedure SumMAT(Mat1,Mat2:Pointer; Mat3:Pointer;lgn1, coll:Integer);
Procedure SubMAT(Mat1,Mat2:Pointer; Mat3:Pointer;lgn1, coll:Integer);
Procedure MulMATVect(Mat1, Vect2:Pointer; Vect3:Pointer;
  lgn1, coll:Integer);
Procedure MulMAT(Mat1,Mat2:Pointer;Mat3:Pointer;
  lgn1, coll, lgn2, col2:Integer);
Procedure InvertMat(Mata:Pointer;N:Integer; MatB:Pointer);
implementation

```

Unité NCOMPLEX: Opérations sur des nombres complexes

Unité NCOMPLEX:

UNIT NComplex;

interface

Uses Dos,crt,McMatDef;

type Type\_Complex = record

Re,Im :Real;end;

ReIm=(Re,Im);

Type\_CP = record

rayon,phase:Real;

end;

VecComplex=array[0..Mchannelmax] of Type\_Complex;

MatComplex=array [MCDim,MCDim] of Type\_Complex;

Var Zsum,Ztemp:Type\_Complex;

Procedure Complex (Var Z:Type\_Complex;x,y:Real);

Procedure CAssign (Var Z :Type\_Complex;C:Type\_Complex);

Procedure Cconj (Var C :Type\_Complex;Z:Type\_Complex);

Procedure CSum (Var S:Type\_Complex; C1,C2:Type\_Complex);

Procedure CSub (Var S:Type\_Complex; C1,C2:Type\_Complex);

Procedure Cscale (Var S:Type\_Complex;A:Real;Z:Type\_Complex );

Procedure CMul (Var P:Type\_Complex; C1,C2:Type\_Complex);

Procedure CDiv (Var Zd:Type\_Complex; Z1,Z2:Type\_Complex);

Procedure CMdiv (Var Zd:Type\_Complex; V1,V2:Type\_CP);

Function CAbs (Z:Type\_Complex):Real;

Function CPhase (Z:Type\_Complex):Real;

Procedure PComplex (Var comp\_Z: Type\_Complex;mod\_phase:Type\_CP);

Procedure Csqrt (Var Z:Type\_Complex;C:Type\_Complex);

Procedure CPower (Var ZN:Type\_Complex;Z1:Type\_Complex;n:Integer);

Procedure ScaleCMat( MatA,MatB:Pointer;N,M:Integer;S:Real);

Procedure AssignCMat( MatA,MatB:Pointer; N,M:Integer);

Procedure IdentCMat( MatA:Pointer; N:Integer);

Procedure ZeroCMat( MatA:Pointer; N:Integer);

Procedure XposeCMat(MatA:Pointer;N,M:Integer; MatB:Pointer);

Procedure SumCMat(Mat1,Mat2:Pointer;Mat3:Pointer;lgn1,col1:Integer);

Procedure SubCMat(Mat1,Mat2:Pointer;Mat3:Pointer;lgn1,col1:Integer);

Procedure MulCMat(Mat1,Mat2:Pointer;Mat3:Pointer;

lgn1,col1,lgn2,col2:Integer);

Procedure ReadCMat(MatStr:namestr; Mat1:Pointer;lgn1,col1:Integer);

Procedure WriteCMat(Mat1:Pointer;lgn1,col1:Integer);

implementation

Unité MCLULEV Factorisation LU & ALgorithme LWR
--

Unit MCLULEV;

interface

Uses crt, dos,

alkeys,alert,

McMatDef,mckernel,mcalmat,Ncomplex;

Function LUfact(MatA:pointer; Var P:Itab; n:Integer):boolean;

Function CLUfact(MatA:pointer; Var P:Itab; n:Integer):boolean;

Procedure Subst(MatA,CB,CY:pointer;P:Itab;N:Integer);

Procedure CSubst(MatA,CB,CY:pointer;P:Itab;N:Integer);

```

Function Solve(MatA,matX,MatB:pointer;m:Integer):boolean;
Function Csolve(MatA,matX,MatB:pointer;m:Integer):boolean;
Function InvCmat(MatA,MatX:pointer;m:Integer):boolean;
Procedure MCCorrelation
  (Var X:MCdata;N,M:Integer;Var R:MCRkvecteur;Kmax:Integer);
Procedure Delta(Var MCR:McrkVecteur;Var Mca:MCvecteur;
  matDD,MatDI:pointer;m,P:Integer);
Function MatReflexion(PtrMCCA,PtrMCCB,PtrMCA,PtrMCB:Pointer;
  Var Pf,Pb,Dd,Di:mcmatrice ;m,p:Integer):boolean;
Procedure MatErreur( Var MCA,MCB:MCvecteur;
  Var Pf,Pb,Dd,Di:mcmatrice ;m,p:Integer);
Procedure MClevinson(Var MCA,MCB:MCvecteur; m,p:Integer);
Implementation

```

Unité ARMAPSD Evaluation de la DSP d'un signal multicanaux
---

```

unit ARMAPSD;
interface
uses dos,Crt,
  Alkeys,
  McMatDef,
  McKernel;
Function McALMatolve(MatA,matX,MatB:pointer;m:Integer):boolean;
Function Ncomplex;
Procedure Horner(A:MCvecteur;Var MAlpha,mBeta:MCmatrice;
  m,p:Integer;w:Real);
Procedure MCMat2CMat(Var CM:MatComplex;Var MCM1,MCM2:MCmatrice;
  N,M:Integer);
Procedure DSARMA(Var A,B:MCvecteur;vpf:mcmatrice;fe : Real ;
  Var mcdsp:McSpectrum;Var Mscphase:MCCPhase;
  m,p,q,Npdsi:Integer);
implementation

```

Unité MCNDATA Bruit blanc/Gaussien Multicanaux
---

```

Unit MCNdata;
interface
uses Crt,Dos,McMatDef,McKernel, McALMat;
Procedure GenIseed(Var K:Real);
Procedure WUniforme(N:Integer;Var Iseed,WgnVar:Real;PtrBB:pointer);
Procedure WGaussien(N:Integer; Var Iseed,WgnVar:Real;ptrBB:pointer);
Procedure BruitMC(m,Gn:Integer;Var bbMC:MCdata);
Procedure Reponse(A,B:MCvecteur;p,q:Integer;Var UK,MCY:MCdata;
  m,GN:Integer);
Procedure MCCorrelation(Var X:MCdata;N,M:Integer;
  Var R:MCRkvecteur;Kmax:Integer);
Implementation

```

Unité MCRSEDIT Fonctions d'Edition et de Sauvegarde
--

Unit MCRSEDIT;

Interface

```

Uses Dos , Crt ,
      Alkeys,Alcrt,
      McMATDef,McKernel;
Const XlWind:Integer=40;
Var   McDspi,McDspj:Integer;
Procedure EditSdata(Tstr:pathstr;Msg:NameStr;Pdata:pointer;
                   Np:Integer);
Procedure SaveSData(Var Dfichier:pathstr; Msg1,Titre,Xstr,
                   Ystr:pathstr;Var data:Sdata;fe:Real;Npmax,I0:Integer);
Procedure EditMcddata(Tstr:pathStr;Msg:NameStr;Var MCDmat:MCdata;
                    m,Np:Integer);
Procedure Read_MCdata(Var Dfichier:pathstr;
                    Var MCDmat:MCdata;m,Dcolonne,pp:Integer;Var Np:Integer);
Procedure Edit_MC_Coeff(MCAptr:pointer;m,P:Integer;Msg,Tstr:pathStr);
Procedure Read_MC_Coeff(Var Fichier:pathstr ;
                    Var MCA,McB:MCvecteur;Var m,P,q:Integer);
Procedure Save_MC_Coeff(Tmsg,Imsg:pathstr;Var Fichier:pathstr ;
                    PtrMCR:pointer;m,P:Integer);
Procedure Write_in_file(Var Dfichier:pathstr;-Msg1,Titre,
                    Xstr,Ystr:pathstr; Var MCDmat:MCdata;m,Npmax,I0:Integer);
Procedure MaxMCDsp(Var psd:MCspectrum ; Var maxpsd:MCMatrice;
                  m,npds:Integer);
Procedure fxfe2Tab(fe:Real;Npmax:Integer;fxfeptr:pointer);
Procedure MCMatrice2Tab(Psdptr,SDptr:pointer;Nx,Ny,Np:Integer);
Procedure Save_MCPSD(Var Fichier:pathstr;Msg1,Titre,Xstr,Ystr:PathStr;
                    fe:Real; Var Mcdsp:MCspectrum;maxPsd:mcmatrice; m,Npmax:Integer);
Procedure Edit_MCPSD(fe:Real;Var Mcdsp:MCspectrum;
                    Var maxpsd:mcmatrice; m,Npmax:Integer);
Procedure SaveMAT1Buffer(MatPtr,XYptr:pointer;m,np:Integer;
                        Var Tmsg:pathstr);
Procedure Savedata1Buffer(DPtr,XYptr:pointer;m,np:Integer;
                        Var Tmsg:pathstr);

```

Implementation

Implementation de la procédure de génération d'un bruit blanc Uniforme/ Gaussien
---

```

Procedure GenIseed(Var K:Real);
(* 12357 <= K <= 69069 *)

```

```

  Var KK,Ran:LongInt;
  begin
    Ran:=random(30000);
    KK:=Random(Ran);
    K:=K+KK;
    if k <12357 then K:=12357+Random(Random(30000));
    if k >69069 then K:=12357+Random(Random(30000));
    kk:=Trunc(K);K:=Kk;
    if (Kk Mod 2 ) = 0 then K:=K+1;
  end;

```

```

Procedure bruit blanc(Var BB:typeES;GN:Integer;Var Varbb:Real);
const KBB=12357;

```

```

  Var som2,Cfbb,kkbb      :Real;
      N10, i,j,k         :Integer;
  function white(Var K:Real) :Real;
  Var k1 :longint;
  white01 :Real;
  begin
    k:=2045*k+1;
    k1:=trunc(K);
    k:=k -(k1 div 1048576)*1048576;
    white01:=(k+1)/1048577;
    white:=(white01 - 0.5)/cfbb;
  end;

```

```

begin
  if Varbb=0 then VarBB:=1;
  som2:=0;
  kkBB:=kBB;
  Cfbb:=sqrt(1/12)*sqrt(1/VarBB);
  for i:= 0 to GN do
    begin
      BB[i]:=white(kkbb);
      som2:=som2+sqr(BB[i]);
    end;
  Varbb:=som2/GN;
end;

```

```

Procedure Bruit_Uniforme(Var BBU:RealTab;Var BBVar,iseed:Real;
  Mbb,Np:Integer);

```

```

  Var som2,Cfbb :Real;
      i         :Integer;
  function BBlanc(Var Iseed:Real):Real;
  Var k1 :Longint;
  white01 :Real;

```

```

begin
  Iseed:=2045*Iseed+1;
  k1:=trunc(ISeed);
  Iseed:=Iseed -(k1 div 1048576)*1048576;
  white01:=(Iseed+1)/1048577;
  if Mbb=0 then Bblanc:=(white01 - 0.5)/cfbb
  else Bblanc:= white01;
end;

begin
  FillChar(BBu,Sizeof(BBU),0);
  if bbVar=0 then Exit;
  som2:=0;
  Cfbb:=sqrt(1/12)*sqrt(1/BBVar);
  for i:= 1 to Np do
    begin
      BBU[i]:=bblanc(Iseed);
      som2:=som2+sqr(BBU[i]);
    end;
  bbVar:=som2/Np;
end;

Procedure BBUniforme(Var BBU:RealTab;Var BBVar,iseed:Real;N:Integer);
begin
  Bruit_Uniforme(BBU,BBVar,iseed,0,N);
end;

procedure BBGaussien(Var BBG:RealTab ;Var BBVar,Iseed:Real;
                    Np:Integer);
  Var N1,L           :Integer;
      cfgbb,som2,bbgVar:Real;
  Procedure Box_Mueller;
  Var i              :Integer;
      DPi,U1,U2,Temp:Real;
  begin
    dPi:=2*Pi;
    for i:=1 to L do
      begin
        U1:=BBG[2*I-1];
        U2:=BBG[2*I];
        TEMP:=SQRT(-2*Ln(U1));
        BBG[2*I-1]:=TEMP*COS(DPi*U2)*SQRT(BBVar);
        BBG[2*I]:=TEMP*SIN(DPi*U2)*SQRT(BBVar);
      end;
    End;
begin
  N1:=Np ;bbgVar:=1;
  IF (Np MOD 2) <> 0 then N1:=Np+1 ;
  Bruit_uniforme(BBG,BBgVar,iseed,1,N1);
  L:=N1 div 2; Box_Mueller; som2:=0;
  if bbVar=0 then exit;
  Cfgbg:=1;{sqrt(1/12)*sqrt(1/BBVar);}
  for i:= 1 to Np do
    begin
      BBG[i]:=-bbg[i]/cfgbg;
      som2:=som2+sqr(BBg[i]);
    end;
  bbVar:=som2/Np;
end;

```

# **Publication**

# CHOIX RECURSIF DE L'ORDRE AR BASE SUR LE CRITERE DE BLANCHISSEMENT DE L'ERREUR DE PREDICTION

Lahcene ABDELOUEL & Belkacem DERRAS.  
Laboratoire de Traitement du Signal

Département d'Electronique Ecole Nationale Polytechnique  
10 Avenue Hassen BADI El-Harrach 16200, Alger.

## Resumé:

Cet article présente un nouveau critère de sélection de l'ordre d'un modèle AR, basé sur l'idée de blanchissement de l'erreur (directe ou inverse) de prédiction linéaire. Le test se fait sur la fonction d'autocorrélation de cette erreur, calculée d'une manière récursive en exploitant la structure en treillis des erreurs de prédiction. Les résultats obtenus à partir des simulations se révèlent encourageants.

Mots clés: Analyse spectrale, Modélisation AR, Sélection de l'Ordre AR, Erreurs de prédiction linéaire, fonction de Corrélation, Algorithme de Levinson

## Abstract:

This paper presents a new criterion for autoregressive (AR) model order selection based on the idea of whitening of the forward/backward linear prediction error by testing the autocorrelation function of this error. The computational efficiency is obtained by exploiting the lattice structure of prediction errors. Simulation results are presented and found to be very encouraging.

Keywords: Spectral analysis, AR modelling, AR model order selection, Linear prediction errors, Correlation function, Levinson recursion.

## 1. Introduction

La modélisation d'un processus donné consiste à déterminer un modèle mathématique basé sur les données d'observation pour représenter le système considéré. La plupart des signaux discrétisés en traitement du signal peuvent être représentés par des modèles ARMA(X), ARMA ou AR [11].

Les méthodes de modélisation paramétriques n'ont plus à faire leurs preuves. Leurs bonnes performances font qu'elles soient de plus en plus utilisées dans des domaines aussi variés que le traitement de la parole, la géophysique, les communications ... etc. En effet, la connaissance du modèle d'un signal permettrait de réaliser le traitement désiré à savoir: l'analyse spectrale [6], le filtrage adaptatif [12], le lissage, la prédiction [9], la déconvolution, l'égalisation adaptative, la suppression d'interférence, le codage [16]...etc.

## 2. Problème du choix de l'ordre AR.

Si l'on considère un signal  $y(k)$ , réel, comme étant la sortie d'un filtre générateur de type ARMA alors sa fonction de transfert  $H(z)$  est donnée par:

$$H(z) = \frac{Y(z)}{E(z)} = \frac{B(z)}{A(z)} \quad (1)$$

où l'entrée est un bruit blanc  $E(k)$ , centré de variance  $\sigma^2$ .  
Si  $B(z)=b=Cte$ , la fonction de transfert  $H(z)$  contient  $n$  pôles le modèle est dit Autorégressif (AR), d'ordre  $n$ , (modèle tout pôles) et donné par :

$$y(k) + \sum_{i=1}^n a_i y(k-i) = b_0 E(k). \quad (2)$$

Le problème du choix de l'ordre du modèle se pose comme suit:

- Quel est le nombre minimal (et optimal  $p$ ) de paramètres  $a_i$  qui pourra représenter convenablement les données  $y(k)$  ?

le choix de l'ordre du modèle a une grande importance, car il permet de juger l'aptitude d'un modèle à représenter un certain type de signal. pour illustrer ce point nous allons examiner l'effet de sur/sous-estimation de l'ordre  $p$  sur l'estimation de la densité spectrale de puissance (DSP).

Les courbes de la figure 4 montrent que pour un ordre inférieur à l'ordre exact ( $n < p$ ), l'estimation reproduit l'enveloppe de la DSP réelle mais la résolution est très mauvaise. Par contre, si l'ordre est surestimé ( $n > p$ ), les principales composantes fréquentielles se retrouvent dans la DSP estimée, mais on voit apparaître des pics parasites arbitraires qui n'ont rien à avoir avec la DSP exacte et qui sont dûs aux pôles ajoutés au modèle.

### Critères de sélection

Les critères de sélection portent essentiellement sur deux aspects:

1- Test de la variance de l'erreur de prédiction: la règle de base consiste à augmenter l'ordre d'une unité tant que la diminution de la variance  $\sigma^2$  est significative. (voir figure 5). Plusieurs critères basés sur ce principe ont été introduit. Les plus importants sont:

FPE<sup>(1)</sup> [1], AIC<sup>(2)</sup> [1], MDL<sup>(3)</sup> [14], et le CAT<sup>(4)</sup> [4]. (Voir figure 6) pour comparaison ).

(1) FPE= Final prediction error

(2) AIC= Akaike Information Criterion

(3) MDL= Minimum description length

(4) CAT= Criterion Autoregressive Transfer function.

- 2- Test des propriétés statistiques de l'erreur de prédiction:
- a-Test de la décorrélation entre l'erreur de prédiction et le signal prédit par le modèle, par calcul de leurs fonctions d'intercorrélations [11].
  - b-Test de la blancheur des erreurs de prédiction par calcul de leurs fonctions d'autocorrélations [4],[11].
- Le critère proposé se classe dans cette deuxième catégorie.

### 3. Estimation AR et Erreurs de prédiction

La modélisation AR consiste à représenter le signal  $y(k)$ , comme une combinaison linéaire des échantillons qui le précèdent, et de l'entrée actuelle  $E(k)$ :

$$y(k) = - \sum_{i=1}^p a_i y(k-i) + E(k) \quad ; b_0 = 1 \quad (3)$$

Si on définit une prédiction de l'échantillon  $y(k)$  en fonction de  $p$  échantillons, alors il y a deux manières différentes pour obtenir l'erreur de prédiction: l'une correspond à la prédiction progressive (directe), l'autre à la prédiction rétrograde (inverse), (voir figure 1). L'erreur progressive de la prédiction directe (forward error) est donnée par [9]:

$$E^f(k) = y(k) - \hat{y}^f(k) \quad (4)$$

L'erreur rétrograde de la prédiction inverse (backward error) est donnée par [9]:

$$E^b(k) = y(k-p) - \hat{y}^b(k-p) \quad (5)$$

où  $\hat{y}^f(k)$  et  $\hat{y}^b(k)$  sont définies à la figure (1).

Comme le filtre AR est complètement défini par ses coefficients de réflexion  $\rho_m$ , on peut montrer que les erreurs de prédiction aux différents ordres peuvent être obtenus à l'aide des équations suivantes [17]

$$\begin{cases} E_m^f(k) = E_{m-1}^f(k) + \rho_m^b E_{m-1}^b(k-1) \\ E_m^b(k) = E_{m-1}^b(k-1) + \rho_m^f E_{m-1}^f(k) \end{cases} \quad (6)$$

avec  $k = 1 \dots p \dots N$ , où  $N$  est le nombre d'observations, et  $E_0^f(k) = E_0^b(k) = y(k)$ .

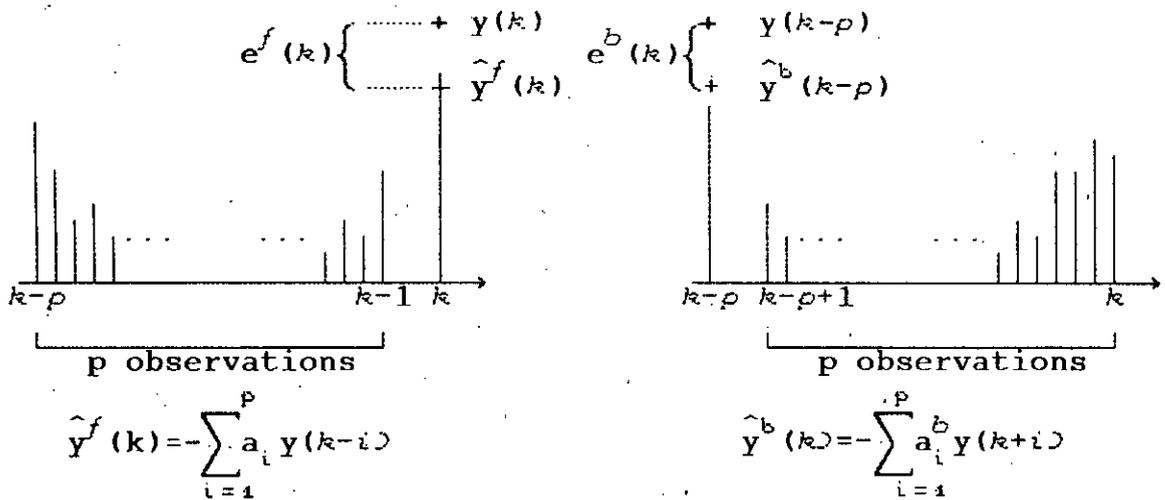
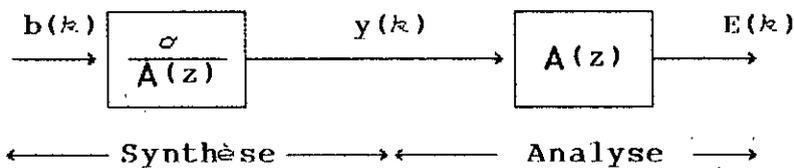


Figure 1: Prédiction directe et rétrograde d'ordre  $p$

#### 4. Choix récursif de l'ordre AR

Le problème de la modélisation AR peut être abordé par deux approches: Analyse et synthèse. L'approche synthèse considère que le signal  $y(k)$  est la sortie d'un filtre AR excité par un bruit blanc  $b(k)$ . L'approche analyse considère que l'erreur de prédiction  $E(k)$  est engendrée par le filtre inverse alimenté par le signal  $y(k)$ .



Dans le cas où le signal est effectivement produit par un modèle AR d'ordre  $p$ , l'erreur  $E(k)$  reproduit l'excitation  $b(k)$ . Le principe du critère BEP (Blanchissement de l'Erreur de Prédiction) consiste à choisir l'ordre qui correspond à la valeur minimale de  $\eta(m)$   $m=|1..Ordre\ Max|$  où  $\eta(m)$  est définie comme étant le rapport de la moyenne arithmétique des éléments de l'autocorrélation (AC) de l'erreur de prédiction à l'ordre  $m$   $Re_m(k)$  avec  $k \geq 1$ , à la valeur  $Re_m(0)$ :

$$\eta(m) = \frac{1}{L} \sum_{i=1}^L \frac{|Re_m(i)|}{Re_m(0)} \quad (7)$$

En effet, le fait de vouloir minimiser  $\eta(m)$ , revient à chercher l'ordre pour lequel la fonction d'autocorrélation de l'erreur de prédiction soit concentrée à l'origine; et par conséquent  $E(k)$  soit proche d'un bruit blanc.

Le temps machine nécessaire au calcul d'une valeur  $\eta(m)$  est extrêmement long. La majeure partie de ce temps est consommé par le calcul des éléments d'autoCorrélation  $Re_m(k)$ . Mais en exploitant la structure en treillis on peut calculer  $Re_m(k)$  d'une manière récursive, et par conséquent d'une façon plus rapide que la méthode directe. En effet, les fonctions d'Inter/Autocorrélation des erreurs de prédiction directe et inverse d'ordre  $m$  sont définies par

$$Re_m^f(k) = E(E_m^f(n+k)E_m^f(n)) \quad (8) \quad Re_m^{fb}(k) = E(E_m^f(n+k)E_m^b(n)) \quad (9)$$

$$Re_m^b(k) = E(E_m^b(n+k)E_m^b(n)) \quad Re_m^{bf}(k) = E(E_m^b(n+k)E_m^f(n))$$

Avec l'hypothèse de l'ergodicité et de la stationnarité des signaux  $E_m^f(k)$ ,  $E_m^b(k)$ , la méthode de calcul directe des fonctions d'autocorrélation  $Re_m^f$ ,  $Re_m^b$  est basée sur un estimateur (généralement biaisé) qui nécessite  $NL$  multiplications par iteration. où  $N$  est le nombre d'observations et  $L$  le nombre d'éléments de  $Re_m^f(k)$  ou  $Re_m^b(k)$  ( $L$  est généralement pris inférieur ou égal à  $N/5$  [11]).

En reportant les expressions récursives (6) des erreurs  $E^f(k)$  et  $E^b(k)$ , dans les expressions (8) et (9), les éléments des fonctions d'autocorrélation prennent la forme:

$$Re_m^f(k) = Re_{m-1}^f(k) + \rho_m^2 Re_{m-1}^b(k) + \rho_m (Re_{m-1}^{fb}(k+1) + Re_{m-1}^{bf}(k-1)) \quad (10)$$

$$Re_m^b(k) = Re_{m-1}^b(k) + \rho_m^2 Re_{m-1}^f(k) + \rho_m (Re_{m-1}^{fb}(k+1) + Re_{m-1}^{bf}(k-1))$$

ce qui correspond à la structure en treillis de la figure 2.

Quant aux expressions (9) des éléments des fonctions d'inter-corrélation, elles prennent la forme:

$$Re_m^{fb}(k) = Re_{m-1}^{fb}(k+1) + \rho_m^2 Re_{m-1}^{bf}(k-1) + \rho_m (Re_{m-1}^f(k) + Re_{m-1}^b(k)) \quad (11)$$

$$Re_m^{bf}(k) = Re_{m-1}^{bf}(k+1) + \rho_m^2 Re_{m-1}^{fb}(k-1) + \rho_m (Re_{m-1}^f(k) + Re_{m-1}^b(k))$$

ce qui correspond à la structure en treillis de la figure 3.

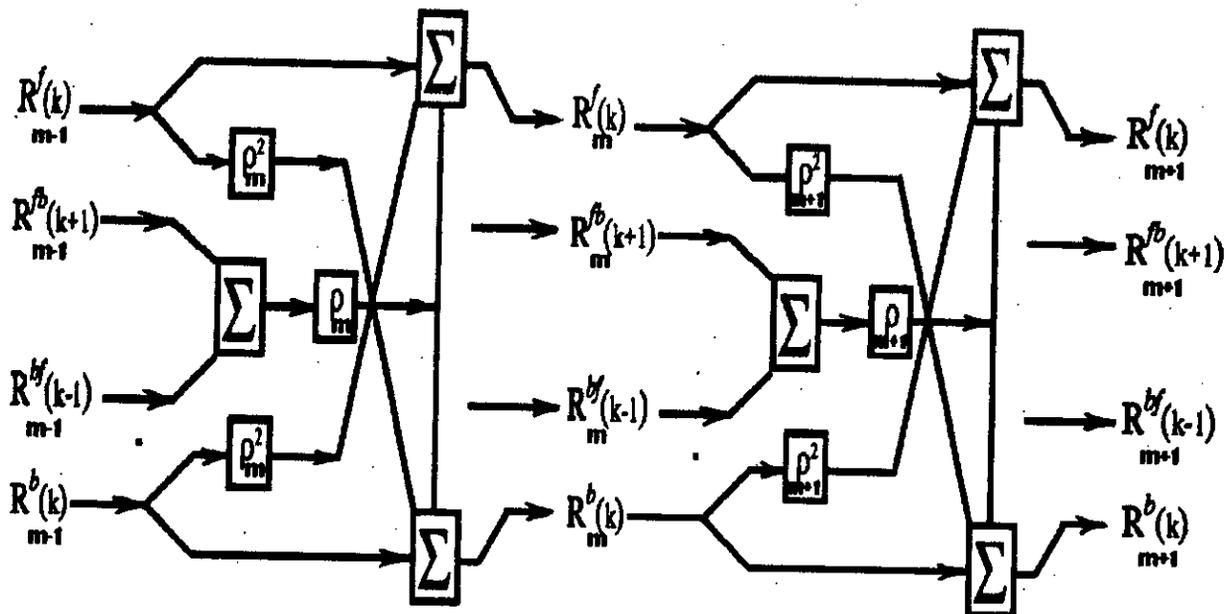


Figure 2: Estimation récursive des fonctions d'autocorrélation des erreurs de prédiction

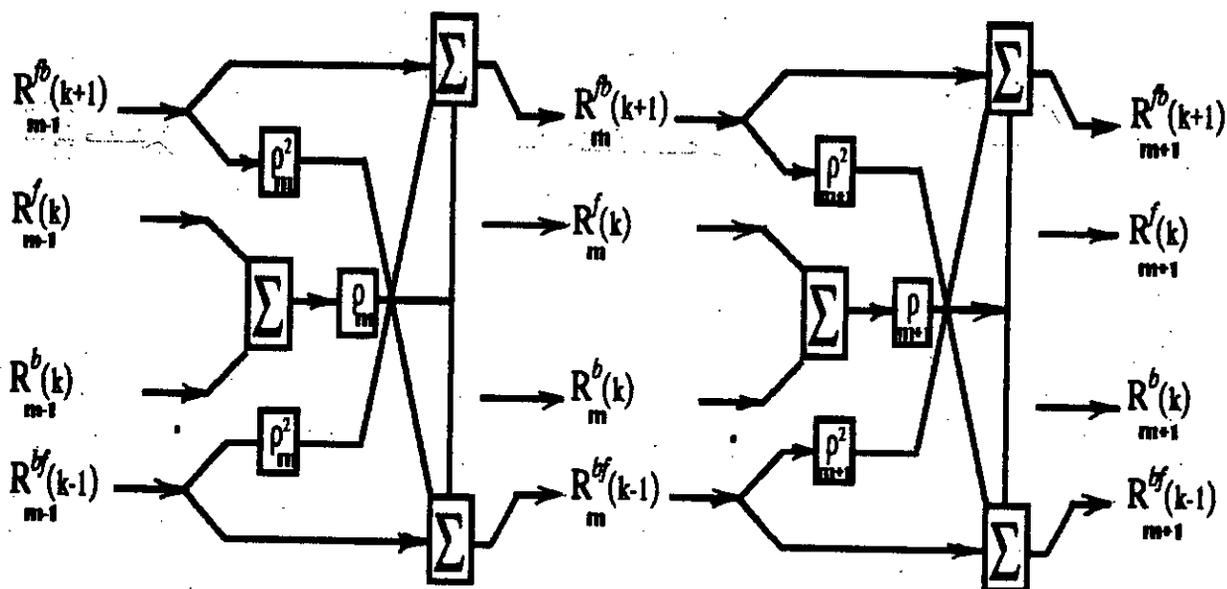


Figure 3: Estimation récursive des fonctions d'intercorrélation des erreurs de prédiction

Cette méthode conduit pratiquement au même résultat en un nombre limité d'opérations élémentaires de l'ordre de  $7L$  multiplications par itération. Donc le calcul par la structure en treillis est beaucoup plus rapide que par un calcul direct.

## 5. Simulations et Résultats

L'examen de quelques courbes nous a permis de constater que, dans certains cas, la courbe représentant l'évolution du critère (BEP), présente des variations irrégulières (figure 7). La courbe en question contient plusieurs minima locaux, ce qui rend le choix automatique de la valeur minimale du critère une tâche difficile et délicate.

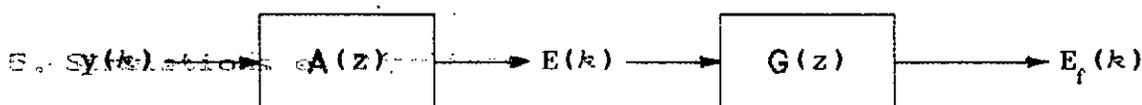
Néanmoins, il est possible d'y remédier à ce problème par une opération de filtrage. Des simulations nous ont permis de conclure qu'un filtre passe-bas du premier ordre élimine complètement les fluctuations rapides.

Dans le plan  $Z$ , la fonction de transfert du filtre du premier ordre est donnée par :

$$G(z) = \frac{1}{1 - a.z^{-1}} \quad (12)$$

où  $a$  est une constante réelle dont la valeur absolue est inférieure à l'unité.

La nouvelle courbe obtenue après filtrage possède un seul minimum, mais elle est déphasée par rapport à la courbe initiale, ceci est certainement dû au fait que le modèle inverse global se compose de la mise en cascade des deux fonctions de transfert  $A(z)$  et  $G(z)$ :



et par conséquent le signal  $y(k)$  est la sortie d'un filtre ARMA de fonction de transfert

$$\frac{Y(z)}{E(z)} = \frac{1 - a.z^{-1}}{A(z)} \quad (13)$$

Le déphasage peut être évalué a priori, car il dépend de la position du pôle du filtre du premier ordre (zéro de la nouvelle fonction de transfert), il est très important lorsque le pôle est très proche du cercle unité. Mais lorsque le pôle est très proche de l'origine, le filtrage n'a pratiquement aucun effet.

Il résulte des simulations que la valeur  $a = 0.707$  est un bon choix et permet d'atteindre le meilleur des compromis entre filtrage et déphasage (figure 8).

## 6. Conclusion

Un problème important dans la modélisation et de l'estimation spectrale paramétrique est celui du choix de l'ordre du modèle. Les critères usuels (FPE, AIC, MDL) ont tendance à sous-estimer l'ordre. Nous avons proposé un nouveau critère basé sur le test de la blancheur de l'erreur de

prédiction d'une manière récursive en vérifiant la similitude de la fonction d'autocorrélation avec une impulsion à l'origine, et afin de présenter un intérêt pratique, les éléments de cette fonction sont estimés d'une manière récursive en utilisant la structure en treillis déduite de celle du modèle AR. Les résultats obtenus par simulation se révèlent tout à fait satisfaisants.

#### Références bibliographiques

- [1] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans. Autom. Control*, vol. AC-19 pp. 716-723, Dec. 1974.
- [2] N. Andersen "On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis," *Geophysics*, vol. 39, pp. 69-72, Feb. 1974.
- [3] R. Boite et M. Kunt, *Traitement de la Parole*, Presses polytechnique romandes, Lausanne, 1987.
- [4] D. G. Childers, (Ed.), *Modern Spectrum Analysis*, IEEE Press, New York (USA), 1978.
- [5] R. H. Jones, "Autoregressive Order Selection," *Geophysics*, vol. 41, pp. 771-773, Aug. 1976.
- [6] S. M. Kay and S. L. Marple, jr. "Spectral Analysis, a Modern Perspective," *Proceeding of the IEEE*, Vol. 69, N° 11, pp. 1379-1419, Nov. 1981.
- [7] R. T. Lacoss "Data Adaptive Spectral Analysis Methods" *Geophysics*, vol. 36, pp. 661-675 Aug. 1971.
- [8] N. Levinson N. "A Heuristic Exposition of Wiener's Mathematical Theory of Prediction and Filtering," *Journal of Math. and Physics*, Vol 26, pp. 110-119, 1947.
- [9] J. Makhoul "Linear Prediction: A Tutorial Review," *Proceeding of IEE*, vol.63, pp.561-580, April 1975.
- [10] J. D. Markel and A. Gray, "On Autocorrelation Equation as Applied to Speech Analysis," *IEEE transaction on Audio and Electroacoustics*, Vol. AU-21, N°02, April 1973.
- [11] L. Marple, *Digital Spectral Analysis with Applications*, Prentice-Hall Inc., New Jersey (USA), 1987.
- [12] K. W. Martin and M. Padmanabhan, " The Spectral Analysis of Short Data-Segments Using an IIR adaptive Filter," *IEEE International Conference on ASSP*, pp.3237-3240, July 1991.
- [13] J. Max, *Méthodes et Techniques de Traitement du Signal et Applications aux Mesures Physiques*, 5° Edition Masson, Tomes I, II, Paris 1985.

- [14] E. Parzen, "Some Recent Advance in Time Series Modeling," *IEEE Trans. Autom. Control*, Vol. AC-19, pp. 723-730, Dec. 1974.
- [15] E. A. Robinson " A Historical Perspective of Spectrum Estimation, " *Proceeding of the IEEE*, Vol.70, N°.9, pp. 885-907, Sep. 1982.
- [16] L. H. Sibul (Editor), *Adaptive signal processing*, IEEE Press, New York (USA), 1987.
- [17] M. Wax, "Ordre Selection for AR Models by Predictive Least Squares," *IEEE Trans., Acoust., Speech, Signal Processing*, vol. ASSP-36, No. 4, pp 581-588, April 1988.

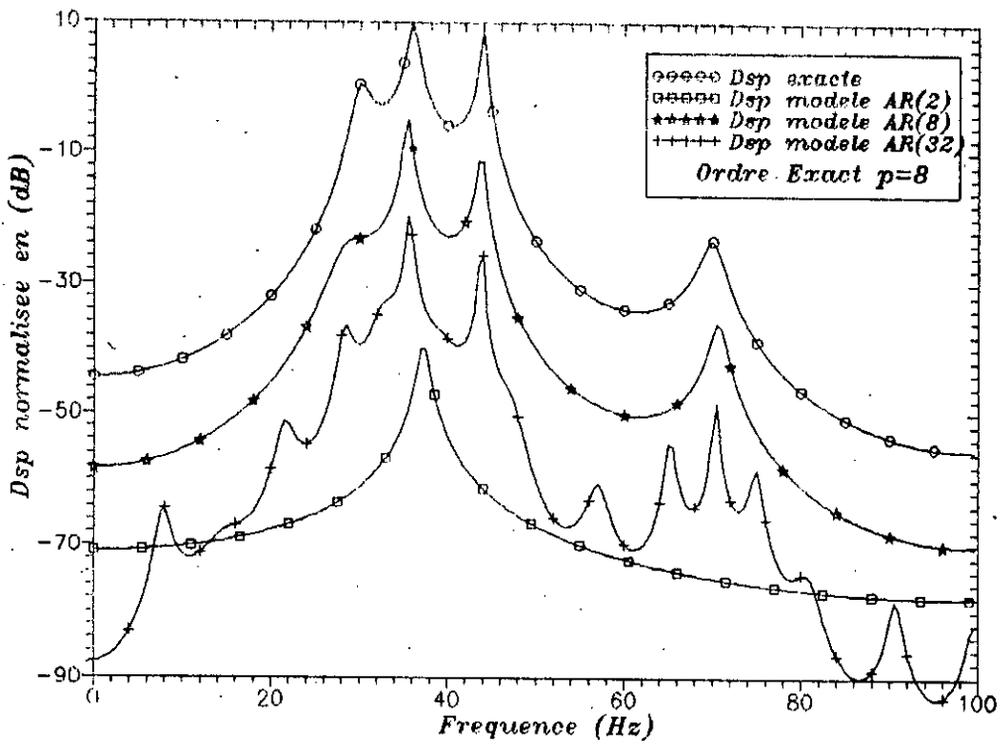


Figure 4:  
Probleme du choix de l'ordre du modele AR

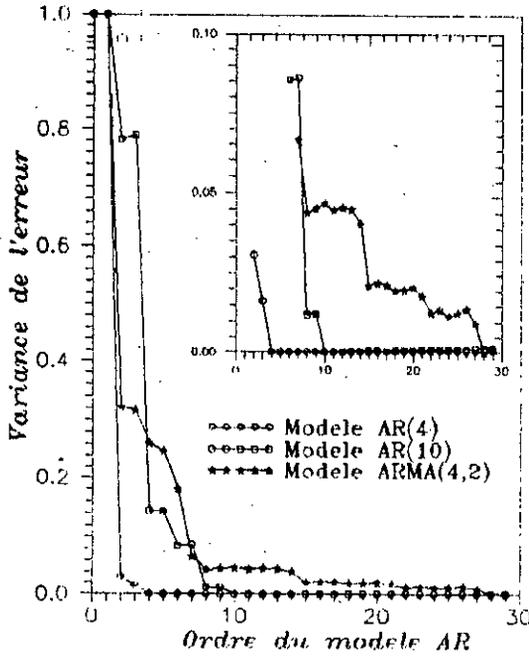


Figure 5:  
Variance de l'erreur de prediction  
de plusieurs modeles

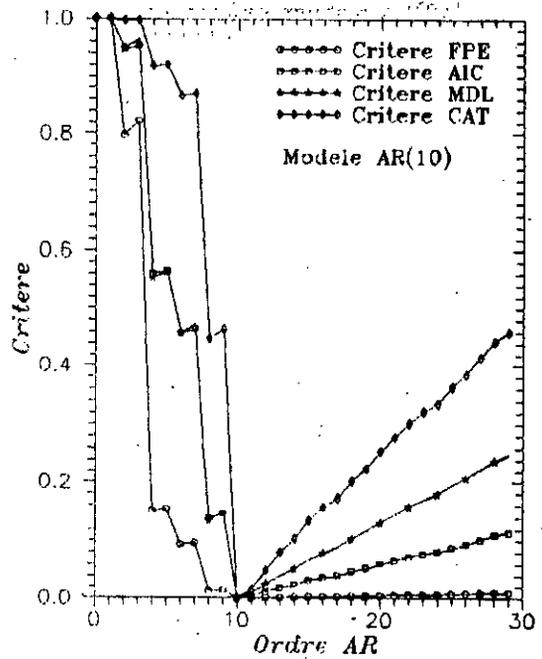


Figure 6:  
Evolution des differents  
Criteres pour un modele AR(10)

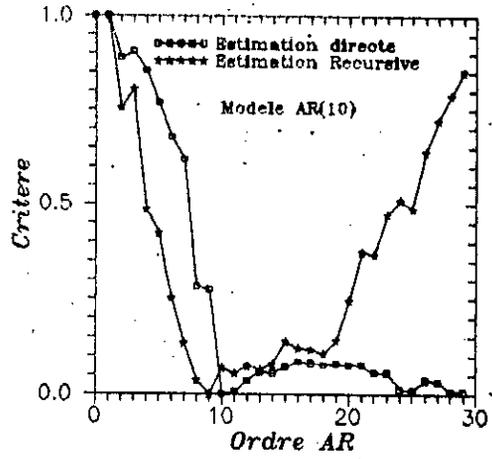
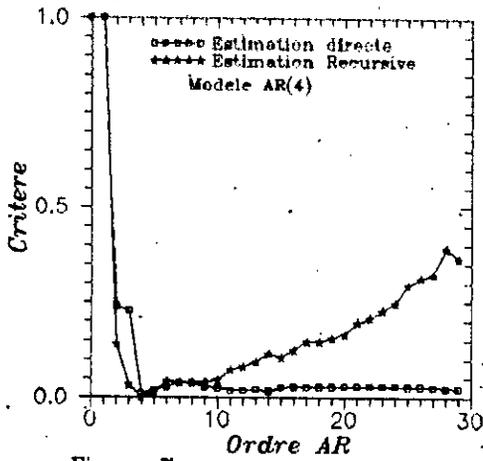


Figure 7: Estimation directe et Estimation recursive du critere de selection de l'ordre AR

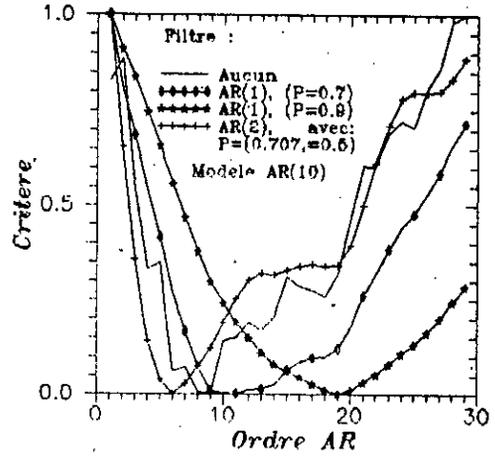
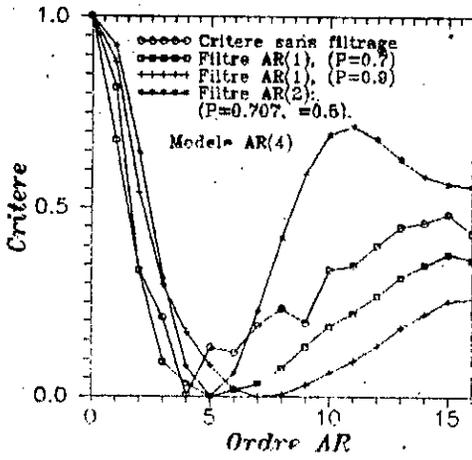


Figure 8: Filtrage du critere de selection de l'ordre