

# Modèles ARMA à coefficients dépendant du temps : estimateurs et applications

ARMA models with time-dependent coefficients: estimators and applications



Yves GRENIER

ENST, Systèmes et Communications, 46, rue Barrault,  
75634 PARIS CEDEX 13

Yves Grenier est né à Ham, Somme, France, en 1950. Ingénieur de l'École Centrale de Paris, en 1972, et Docteur-Ingénieur de l'École Nationale Supérieure des Télécommunications, Paris, en 1977, il est aussi titulaire du Doctorat d'État ès Sciences Physiques, depuis 1984. Depuis 1974, il fait partie du Département Systèmes et Communications de l'École Nationale Supérieure des Télécommunications, Paris, où il a effectué des recherches sur la reconnaissance de la parole, l'identification du locuteur, et l'adaptation au locuteur des systèmes de reconnaissance de parole. Depuis 1977, il travaille sur la modélisation des signaux, l'analyse spectrale de signaux bruités, avec des applications à la reconnaissance et la synthèse de parole, ainsi que sur l'estimation de modèles non stationnaires, et de représentations temps-fréquence, appliqués à la parole, et aux signaux sismiques. Il étudie également la modélisation de signaux vectoriels, et le filtrage d'antennes à large bande. Il a été co-organisateur d'un séminaire GALF sur l'analyse de signaux de parole, en décembre 1983, et a organisé un séminaire CNRS sur la modélisation non stationnaire, à Nice, en juin 1984. Il est membre du GALF (Groupement des Acousticiens de Langue Française), et de IEEE.

## RÉSUMÉ

Ce texte, effectuant une synthèse de divers travaux de l'auteur, présente une méthodologie de modélisation des signaux non stationnaires, au moyen d'une classe de modèles autorégressifs à moyenne ajustée (ARMA). La non-stationnarité du signal y est caractérisée par l'évolution temporelle des coefficients du modèle : ceux-ci s'expriment comme combinaisons linéaires d'une famille de fonctions connues, suivant en cela les idées dues à Rao, Mendel, puis Liporace. Il sera montré que cette hypothèse conduit à un ensemble d'estimateurs dont on détaillera plusieurs variantes, pour les modèles autorégressifs, puis les modèles à moyenne ajustée, enfin pour les filtres en treillis, et leur paramétrisation en terme de fonctions d'aire logarithmiques. La synthèse de parole est une des applications où de tels modèles s'avèrent performants et une description de cette application conclura l'article.

## MOTS CLÉS

Modèles ARMA, signaux non stationnaires, filtres en treillis, fonctions d'aire logarithmiques, codage de parole

## SUMMARY

*This paper describes a framework for time-dependent modelling of nonstationary signals, based upon autoregressive or ARMA models with time-varying coefficients. Previous works by the author are summarized here: starting from ideas proposed by Rao, Mendel and then Liporace, it is assumed that the coefficients of the model are expressed as linear combinations of known time functions. A class of estimators is described in this text, comprising autoregressive models, moving average models, and also lattice filters, parametrized through reflection coefficients or through Log Area Ratios. Speech synthesis, which is one of the applications for which these models are efficient, concludes the paper.*

## KEY WORDS

ARMA models, nonstationary signals, lattice filters, Log Area Ratios, speech coding

## TABLE DES MATIÈRES

### Introduction

#### 1. Description des modèles

- 1.1. Choix de la base
- 1.2. Méthodes d'estimation
- 1.3. Sélection de l'ordre
- 1.4. Extension au cas ARMA

#### 2. Modèles autorégressifs

- 2.1. Mise en équations
- 2.2. Estimation

#### 3. Modèles MA

- 3.1. Résolution par moindres carrés
- 3.2. Factorisation MA non stationnaire
- 3.3. Gain non stationnaire

#### 4. Modèles en treillis

- 4.1. Filtre en treillis, coefficient de réflexion
- 4.2. Fonctions d'aire logarithmiques

#### 5. Applications

### Introduction

Les méthodes paramétriques de traitement du signal n'ont plus à faire leurs preuves, tant que les signaux à traiter peuvent être considérés comme stationnaires, qu'ils le soient exactement ou que ce soit une approximation valable : les modèles autorégressifs et autorégressifs à moyenne ajustée ont de nombreuses applications dans ce contexte. Par contre lorsque la stationnarité est manifestement absente, la situation est moins simple. Diverses méthodes coexistent, les plus répandues sont les méthodes adaptatives où le modèle est réajusté à l'arrivée de chaque nouvel échantillon. Si leur efficacité est certaine en ce qui concerne les aspects purement liés à l'estimation, ils ont cependant l'inconvénient de ne pas se plier au principe de parcimonie, qui veut qu'on représente un signal par le nombre minimal de paramètres. En effet, ces méthodes adaptatives remplacent  $N$  échantillons par  $Np$  coefficients si  $p$  est l'ordre de chaque modèle instantané. Pour éviter cette accroissement, il est nécessaire de paramétrer non seulement chaque modèle, mais aussi leur évolution. L'approche qui fait l'objet de ce texte utilise une classe de modèles introduits dans les travaux de Mendel [36], Rao [43], Liporace [34], que nous conviendrons de dénommer dans ce qui suit « modèles évolutifs », pour les distinguer des autres modèles non stationnaires, modèles adaptatifs (méthodes récursives sur le temps avec oubli exponentiel ou mémoire finie) ou modèles à coefficients aléatoires.

Les cinq parties de ce texte recouvrent les contenus suivants : une description générale des modèles évolutifs, l'estimation des modèles autorégressifs, l'estimation des modèles MA, d'autres modèles dérivés du modèle ARMA, et enfin, une application aux signaux de parole.

### 1. Description des modèles

Le signal  $y_t$  mesuré sur l'intervalle  $[0, T]$  sera représenté soit par un modèle autorégressif :

$$(1) \quad y_t + a_1(t-1)y_{t-1} + \dots + a_p(t-p)y_{t-p} = \varepsilon_t$$

soit par un modèle ARMA :

$$(2) \quad y_t + a_1(t-1)y_{t-1} + \dots + a_p(t-p)y_{t-p} = b_0(t)\varepsilon_t + b_1(t-1)\varepsilon_{t-1} + \dots + b_q(t-q)\varepsilon_{t-q}$$

où  $\varepsilon_t$  est un bruit blanc gaussien, centré, de variance  $\sigma^2$  dans le modèle AR et 1 dans le modèle ARMA.

Le modèle autorégressif (1) et le modèle ARMA (2), seront dits évolutifs lorsque leurs coefficients sont combinaisons linéaires des fonctions  $f_j(t)$  :

$$(3) \quad a_i(t) = \sum_{j=0}^m a_{ij} f_j(t).$$

#### 1.1. CHOIX DE LA BASE

La première question qui se pose est celle du choix des fonctions  $f_i(t)$  pour une application donnée. Liporace [34], ne précise pas ce point, mais contraint les fonctions  $f_j(t)$  à être orthogonales. Ceci n'est pas nécessaire pour l'identification du modèle, mais améliore l'estimation par un meilleur conditionnement numérique des systèmes à résoudre. La base de fonctions la plus simple est celle des puissances du temps, issue du développement limité des  $a_i(t)$  :

$$(4) \quad f_m(t) = \frac{t^m}{m!}$$

Cette base est utilisée par plusieurs auteurs (tableau), de même que son équivalent après orthogonalisation, où les fonctions  $f_i(t)$  sont des polynômes de Legendre. Hall *et al.* [25, 26], ont introduit un second jeu de fonctions, issues de la décomposition en série de Fourier des coefficients  $a_i(t)$  :

$$(5) \quad \begin{cases} f_0(t) = 1 \\ f_{2m-1}(t) = \cos(mt), & m > 0 \\ f_{2m}(t) = \sin(mt), & m > 0. \end{cases}$$

Une troisième base, introduite par [17], est celle des fonctions sphéroïdales aplaties.

#### 1.2. MÉTHODES D'ESTIMATION

Cette diversité dans le choix des bases possibles se retrouve bien entendu aussi dans les méthodes d'esti-

# RECHERCHES

TABLEAU

Comparaison des méthodes évolutives publiées, vis-à-vis du modèle choisi (A=autorégressif, M=mixte autorégressif à moyenne ajustée, R=régression linéaire), de l'estimateur (G=gradient, C=moindres carrés, V=maximum de vraisemblance, A=approximations successives) et de la base (T=puissances du temps  $t$  ou développement limité, L=polynômes de Legendre, F=série de Fourier, S=fonctions sphéroïdales aplaties).

Auteurs	Année	Modèle			Estimateur				Base				
		A	M	R	G	C	V	A	T	L	F	S	
Mendel . . . . .	69, 73			*	*								
Rao . . . . .	70	*				*	*		*				
Liporace . . . . .	75	*				*				*			
Hall <i>et al.</i> . . . . .	77, 83	*				*			*		*		
Hinich, Roll . . . . .	77			*				*		*			
Kozin . . . . .	77	*					*			*			
Abdrabbo . . . . .	79	*					*		*				
Nakajima <i>et al.</i> . . . . .	80, 82	*				*			*				
Grenier . . . . .	81-85	*	*			*	*			*	*	*	*
Clergeot . . . . .	82, 84	*				*	*		*				
Gersch <i>et al.</i> . . . . .	82	*				*				*			
Ginn . . . . .	82	*											
Pearson . . . . .	82		*			*			*				
Martinelli . . . . .	83	*						*			*		
Pitarque . . . . .	84	*	*				*		*				
Sharman <i>et al.</i> . . . . .	84	*							*				
Xianya <i>et al.</i> . . . . .	84	*				*			*				

mation (tableau). La méthode la plus simple est la méthode du gradient ou approximation stochastique, utilisée par Mendel [36, 37]. Plus complexe mais plus efficace, est la méthode de moindres carrés utilisée par Rao [43], et plusieurs autres auteurs. La méthode du maximum de vraisemblance a été mise en œuvre par Rao [43], Kozin [31], Abdrabbo [1], Clergeot [8]. On trouve aussi chez Hinich et Roll [28], l'emploi d'une méthode robuste et non paramétrique d'estimation des  $a_i(t)$  suivie par leur approximation par des polynômes de Legendre. Certains auteurs utilisent une estimation en plusieurs étapes du modèle : à chaque étape est estimé une partie du modèle, en principe les composantes des  $a_i(t)$  sur une des fonctions de la base, cette partie est estimée sur le résidu obtenu par filtrage du signal avec les parties déjà identifiées du modèle. Nakajima et Suzuki [38-40], le font de façon optimale, et Martinelli [35], d'une manière sous-optimale. Pearson [41], est le seul auteur à avoir appliqué ces idées pour des signaux à temps continu. Gersch et Kitagawa [12], ont étendu le modèle AR évolutif au cas vectoriel. Ginn [13], identifie un modèle autorégressif complexe sur une sinusoïde modulée, sa procédure d'estimation est non linéaire. Xianya et Evans [49, 50], font de ces modèles une estimation adaptative sous une fenêtre glissante. Pitarque [42], étudie l'identification de modèles évolutifs sous forme d'équations d'état, en utilisant un filtre de Kalman étendu.

### 1. 3. SÉLECTION DE L'ORDRE

La mise en pratique de ces modèles évolutifs suppose que des procédures permettent une prise de décision purement mécanique entre deux bases de fonctions possibles, ou entre plusieurs ordres, plusieurs dimensions de la base de fonctions. Ce point a été très peu

étudié. Seuls Kozin, Nakajima [32], et Clergeot [9], ont considéré le problème de l'estimation de l'ordre du modèle, et ceci lorsque les paramètres sont estimés au sens du maximum de vraisemblance. Ils montrent que la consistance et la normalité asymptotique des paramètres assurent la validité du critère d'Akaike. Il semble alors raisonnable d'utiliser les divers critères d'ordre existants : le critère AIC d'Akaike [2], celui de Schwarz [47], et de Rissanen [45], celui de Hannan et Quinn [27], pour choisir la meilleure structure « ordre + dimension de la base » d'un modèle dans une base donnée, sur un ensemble d'échantillons  $[y_0 \dots y_T]$  d'un signal non stationnaire. Le choix entre deux ou plusieurs bases possibles se fait alors simplement en retenant pour chaque base la structure minimisant le critère d'ordre, puis la base associée au plus petit des critères ainsi obtenus. Un exemple de détermination de l'ordre d'un modèle AR [6] est présenté en section 5.

### 1. 4. EXTENSION AU CAS ARMA

Il est possible d'étendre à d'autres modèles l'idée de cette méthodologie qui est de postuler l'existence et la connaissance d'une famille de fonctions  $f_i(t)$  exprimant les évolutions des paramètres du modèle. L'extension la plus directe, mais pas nécessairement la plus facile à obtenir, du modèle autorégressif est donnée par le modèle mixte autorégressif à moyenne ajustée ou ARMA. L'identification de ces modèles peut se faire en deux temps. Dans le premier, c'est la partie autorégressive du modèle qui est estimée, par une équation du type Yule-Walker, rendue non symétrique [17, 19]. De même que, pour le modèle stationnaire, la partie autorégressive peut être estimée par une variante correspondant à un système d'équations surdéterminé, dans la méthode dite de « covariance

sur la corrélation » ou la méthode pondérée due à Cadzow, de même dans le cas non stationnaire peut s'écrire un système d'équations sur-déterminé [17], mais il semble alors que tout algorithme rapide soit à exclure.

Le second temps, dans l'identification ARMA, sera celui de l'estimation de la partie MA, après qu'ait été éliminée, par filtrage inverse, la contribution de la partie AR du modèle. Deux solutions à ce problème peuvent être proposées. Dans l'une, le modèle MA est approximé par un modèle autorégressif d'ordre élevé, nettement supérieur à l'ordre du modèle MA. Le retour de ce modèle autorégressif au modèle MA se fait par l'estimation de l'entrée (filtrage inverse) et par régression du signal MA sur cette entrée [17, 19]. Dans l'autre résolution du modèle MA, on réalise une factorisation de la covariance du signal qui peut être vue comme une extension de la paramétrisation de Schur d'une covariance stationnaire, extension au cas non stationnaire, avec la contrainte que la solution soit exprimée sur la base des fonctions  $f_i(t)$ . Cette solution comporte une factorisation d'un signal vectoriel, suivie par une renormalisation pour obtenir le modèle scalaire non stationnaire [18].

Plutôt que cette approche en deux temps de la modélisation ARMA, on peut aussi concevoir une identification simultanée de la partie autorégressive et de la partie à moyenne ajustée. Une méthode de maximum de vraisemblance exacte a pu être proposée, assortie d'un calcul du gradient et du Hessien de la vraisemblance, permettant d'utiliser pour sa maximisation une méthode de Newton-Raphson [15, 16]. Le calcul de la vraisemblance étant cependant très coûteux, une méthode d'estimation simultanée des parties AR et MA était souhaitable. Une telle procédure a été proposée par [23] : le modèle ARMA est ajusté de manière à approcher au mieux la réponse impulsionnelle (à deux indices car non stationnaire) d'un modèle AR d'ordre élevé estimé sur un sur-ensemble de la base de fonctions. Cette méthode constitue une généralisation non stationnaire de celle proposée par Graupe *et al.* [14], pour le modèle ARMA stationnaire.

## 2. Modèles autorégressifs

Dans ce paragraphe est étudiée l'estimation du modèle autorégressif ou de la partie autorégressive du modèle mixte ARMA. Le point essentiel est l'introduction d'un vecteur dont les composantes sont les produits du signal par chacune des fonctions de base. Ce vecteur se retrouvera dans toutes les équations d'optimalité, et tous les estimateurs traités dans ce texte, qu'ils concernent le modèle autorégressif ou le modèle à moyenne ajustée MA.

### 2.1. MISE EN ÉQUATIONS

#### 2.1.1. Signal vectoriel associé au signal scalaire

Le modèle autorégressif évolutif est décrit par l'équation (1), où les coefficients  $a_i$  sont des fonctions

du temps obtenues comme combinaisons linéaires des fonctions  $f_i(t)$  connues. On remarquera le choix de l'origine des temps pour  $a_i$ , afin de faire coïncider l'indice temporel des deux termes du produit prédicteur-signal retardé, et multiplier  $y_{t-i}$  par  $a_i(t-i)$ . Quel est l'intérêt de cette écriture? C'est de faire apparaître dans chaque produit les variantes décalées d'un même signal vectoriel, et des vecteurs constants :

$$(6) \quad a_i(t-i)y_{t-i} = \left( \sum_{j=0}^m a_{ij} f_j(t-i) \right) y_{t-i} \\ = \sum_{j=0}^m a_{ij} (f_j(t-i) y_{t-i}) \\ = [a_{i0} \dots a_{im}] Y_{t-i}$$

Le vecteur  $Y_t$  est le vecteur des produits de  $y_t$  par chacune des fonctions de la base, ou en quelque sorte des projections de la fonction temporelle  $y_t$  sur la base :

$$(7) \quad Y_t = \begin{bmatrix} f_0(t) \\ \vdots \\ f_m(t) \end{bmatrix} y_t$$

Ceci permet de faire apparaître le modèle autorégressif non stationnaire comme une régression, cette fois stationnaire, du signal  $y_t$  sur le passé du vecteur  $Y_t$  :

$$(8) \quad y_t + [a_{10} \dots a_{1m} a_{20} \dots a_{pm}] \begin{bmatrix} Y_{t-1} \\ \vdots \\ Y_{t-p} \end{bmatrix} = \varepsilon_t$$

Toute la méthodologie des modèles évolutifs tient en ce tour de passe-passe : remplacer un signal scalaire non stationnaire par un signal vectoriel, donc plus complexe, mais pour lequel le modèle deviendra invariant dans le temps et par conséquent identifiable avec un effort modéré.

Un des relecteurs de cet article indique qu'il existerait une explication plus logique que l'argument « technique » employé ici. Il suffirait, dit-il, d'écrire l'équation d'état associée à un modèle ARMA stationnaire et de remplacer les coefficients dans l'équation d'état par des coefficients de la forme  $a_i(t)$  et  $b_i(t)$  tous synchrones : en écrivant alors la forme entrée-sortie équivalente, on obtient des termes de la forme  $a_i(t-i)y_{t-i}$ . Certes, la remarque s'avère exacte si l'on considère l'équation d'état sous forme canonique observable (9) : dans ce cas, si l'on indexe les coefficients stationnaires par le temps  $t$ , on obtient bien une équation telle que (2), donc avec des coefficients tous synchrones.

$$(9) \quad x_{t+1} = \begin{bmatrix} -a_1 & 1 & 0 & 0 \\ \vdots & 0 & 1 & 0 \\ & & & 1 \\ -a_p & 0 & & 0 \end{bmatrix} x_t + \begin{bmatrix} b_1 - b_0 a_1 \\ b_2 - b_0 a_2 \\ \vdots \\ b_p - b_0 a_p \end{bmatrix} e_t \\ y_t = [1 \ 0 \ \dots \ 0] x_t + b_0 e_t$$

Malheureusement, cet argument n'a pas le caractère de généralité qu'on attendrait de lui. Ainsi la forme canonique observable, si elle est écrite en ligne, au lieu d'être écrite en colonne comme cela vient d'être fait, réalise également un modèle ARMA stationnaire, mais nécessite, pour conduire à l'équation (2) dans le cas non stationnaire, le remplacement des  $a_i$  par  $a_i(t+p-i)$ . Quant aux formes canoniques commandables, leur situation est encore pire, car elles peuvent ne pas conduire à un modèle ARMA [20, p. 77-79], ou bien à un modèle ARMA d'ordre supérieur à la dimension de l'état. La plus grande prudence s'impose donc dans la manipulation des systèmes dépendant du temps, qu'ils soient décrits par une équation d'état, ou par une relation entrée-sortie. L'écriture de coefficients synchrones dans l'équation (6) reste ainsi ce qu'elle était, une convention d'écriture, arbitraire mais simplificatrice. Il est d'ailleurs possible, avec de nombreuses bases de fonctions de convertir les composantes  $a_{ij}$  d'une forme synchrone à une forme asynchrone, et vice-versa.

### 2.1.2. Équations d'optimalité du modèle AR

L'identification du modèle requiert le choix d'un critère d'optimalité. La triple signification de  $\varepsilon_t$  qui est à la fois l'entrée du modèle autorégressif, l'innovation du signal  $y_t$ , mais aussi une erreur de prédiction, rend le choix du critère immédiat :  $\varepsilon_t$  étant une erreur, il faut minimiser la variance de cette erreur. Soit  $\theta$  le vecteur des paramètres  $a_{ij}$  :

$$(10) \quad \theta = [a_{10} \dots a_{1m} a_{20} \dots a_{pm}]^T.$$

La minimisation de  $E(\varepsilon_t^2)$  par rapport à  $\theta$  conduit au système d'équations :

$$(11) \quad E \left( \begin{bmatrix} Y_{t-1} \\ \vdots \\ Y_{t-p} \end{bmatrix} [Y_{t-1}^T \dots Y_{t-p}^T] \right) \theta = -E \left( \begin{bmatrix} Y_{t-1} \\ \vdots \\ Y_{t-p} \end{bmatrix} y_t \right).$$

On reconnaît dans (11) la forme habituelle des équations de Yule-Walker pour la prédiction linéaire, mais les éléments du passé sont remplacés par les vecteurs  $Y_t$ . Par contre le vecteur inconnu ne dépend plus du temps. C'est cette remarque qui expliquera qu'on puisse utiliser des sommes temporelles pour faire l'estimation.

### 2.1.3. Équations pour la partie AR du modèle ARMA

Pour passer d'un estimateur du modèle autorégressif à un estimateur de la partie autorégressive du modèle ARMA, il faut revenir sur la minimisation de la variance de  $\varepsilon_t$ . En écrivant l'équation (11) comme :

$$(12) \quad 0 = \text{Grad}_\theta E(\varepsilon_t^2) = 2 \cdot E(\varepsilon_t \text{Grad}_\theta \varepsilon_t)$$

avec :

$$(13) \quad \text{Grad}_\theta \varepsilon_t = \begin{bmatrix} Y_{t-1} \\ \vdots \\ Y_{t-p} \end{bmatrix}$$

on fait apparaître l'équation (11) comme une équation d'orthogonalité entre  $\varepsilon_t$  et les  $p$  derniers échantillons du passé de  $Y_t$ . Une telle relation résulte directement de l'équation du modèle AR, eq. (1), et de la blancheur de  $\varepsilon_t$ . Le modèle mixte ARMA sera donné par (2), où  $\varepsilon_t$  sera toujours un bruit blanc centré, mais ici sa variance sera 1. A cause de la moyenne mobile de  $\varepsilon_t \dots \varepsilon_{t-q}$ , on n'a plus l'orthogonalité de la partie autorégressive avec le signal  $y_{t-1} y_{t-2} \dots$ . La quantité  $w_t$  définie par (14) contenant les échantillons du passé de  $\varepsilon_t$  jusque  $\varepsilon_{t-q}$ , n'est orthogonale aux échantillons  $y_{t-i}$  (et donc aux échantillons  $Y_{t-i}$ ), que pour  $i > q$ .

$$(14) \quad w_t = y_t + a_1(t-1)y_{t-1} + \dots + a_p(t-p)y_{t-p}$$

Ainsi :

$$(15) \quad E \left( \left( y_t + \theta^T \begin{bmatrix} Y_{t-1} \\ \vdots \\ Y_{t-p} \end{bmatrix} \right) \cdot Y_{t-i}^T \right) = 0 \quad \text{pour } i > q.$$

La relation (11) se trouvera donc transformée en la relation (16) exprimant une condition nécessaire pour que  $\theta$  soit le vecteur des coefficients  $a_{ij}$  de la partie autorégressive.

$$(16) \quad E \left( \begin{bmatrix} Y_{t-q-1} \\ \vdots \\ Y_{t-q-p} \end{bmatrix} [Y_{t-1}^T \dots Y_{t-p}^T] \right) \theta = -E \left( \begin{bmatrix} Y_{t-q-1} \\ \vdots \\ Y_{t-q-p} \end{bmatrix} y_t \right)$$

Cette dernière relation est l'équivalent non stationnaire de l'équation utilisée par Hsia et Landgrebe [29], puis Gersch [11], pour estimer la partie autorégressive du modèle ARMA. Il faut cependant remarquer que les équations (11) donnant les conditions d'optimalité du modèle AR, ou (16) donnant celles de la partie AR du modèle ARMA ne constituent pas des conditions suffisantes pour que  $\theta$  les vérifiant soit le vecteur des paramètres du modèle. En effet, aucune des matrices figurant dans ces trois systèmes linéaires n'est de rang complet car le vecteur aléatoire  $Y_t$  est le produit d'un scalaire aléatoire  $y_t$  par le vecteur des fonctions  $f_0(t) \dots f_m(t)$  qui sont certaines. Les matrices multipliant  $\theta$  dans les trois équations sont de rang au plus égal à  $p$  dans (11) et (16) alors que leur dimension est  $p \cdot (m+1)$ . C'est le passage de l'espérance mathématique à la somme temporelle de l'estimateur ergodique qui rend ces matrices inversibles et de rang complet, les conditions deviennent des conditions suffisantes sur  $\theta$ .

## 2.2. ESTIMATION

L'emploi du terme « estimateur » à propos des équations (11) et (16) n'était pas justifié, car rien n'indiquait comment étaient estimées les espérances

mathématiques figurant dans ces équations. Il est clair, du fait de l'invariance temporelle de  $\theta$  que le remplacement dans les équations de l'espérance par une somme sur le temps, entraîne des équations qui expriment encore des conditions nécessaires sur  $\theta$ , et si la somme se fait sur un intervalle assez long, on obtient des conditions suffisantes pour que  $\theta$  soit solution au problème posé.

Pour l'estimation de la partie autorégressive du modèle ARMA, on retiendra l'équation nécessaire (16), et l'estimateur ergodique, ce qui donne l'équation (17) :

$$(17) \sum_{t \in \Gamma} \begin{bmatrix} T_{t-p-1} \\ \vdots \\ Y_{t-q-p} \end{bmatrix} [Y_{t-1}^T \dots Y_{t-p}^T] \cdot \theta = - \sum_{t \in \Gamma} \begin{bmatrix} Y_{t-q-1} \\ \vdots \\ Y_{t-q-p} \end{bmatrix} y_t$$

Ainsi les modèles AR et ARMA évolutifs permettent pour un signal non stationnaire de tourner la non-ergodicité inhérente à un tel signal, et de fabriquer un estimateur à partir de sommes temporelles. Il est visible que ceci est permis par la conjugaison de deux propriétés : l'existence d'une transformation du modèle dans laquelle les paramètres inconnus deviennent tous invariants vis-à-vis du temps, et la linéarité des équations d'optimalité, qui est liée au fait que ces équations expriment des conditions d'orthogonalité. L'estimation qui en résulte, biaisée par les variations de l'énergie résiduelle pourra être rendue sans-biais par une estimation *a posteriori* des variations incriminées, et leur incorporation dans les équations. Ainsi procède Clergeot [8], qui propose une minimisation de la vraisemblance d'un modèle autorégressif évolutif, dont le gain est lui aussi évolutif. Les équations d'optimalité conduisent cependant à un système non linéaire.

Ceci ne fixe pas totalement l'équation (17). Il est en effet possible de distinguer comme pour le modèle autorégressif stationnaire deux cas, connus sous les dénominations malencontreuses mais fermement établies de « méthode de corrélation » et « méthode de covariance ». Si on veut obtenir dans (17) une matrice de Toëplitz, on devra supposer les données  $y_0 \dots y_T$  prolongées par des zéros, utiliser une somme sur l'intervalle  $\Gamma = ]-\infty, +\infty[$  et ceci constituera la « méthode de corrélation ». Si on ne veut pas faire cette hypothèse, on limitera  $\Gamma$  à l'intervalle  $\Gamma = [p+q, T]$ , mais la matrice figurant dans (17) ne sera plus de Toëplitz. Ce sera la « méthode de covariance » pour laquelle existera cependant un algorithme rapide. Peut-on départager les deux méthodes? Hall *et al.* [26], à cause du biais de la « méthode de corrélation » penchent nettement pour l'autre méthode. Je ne partage pas cette opinion. Les simulations que j'ai réalisées [20], montrent un avantage certes pour la « méthode de covariance », mais le gain est si faible qu'il ne me paraît pas justifier le coût de

cette méthode : le code Fortran de la « covariance » est environ huit fois plus encombrant que celui de la « corrélation ».

### 3. Modèles MA

Ce paragraphe est consacré à l'estimation du modèle à moyenne ajustée MA (moving-average). Il pourra s'agir indifféremment d'un signal uniquement MA ou d'un signal ARMA dont la partie autorégressive a déjà été identifiée et éliminée par filtrage inverse du signal  $y_t$  d'origine. Deux méthodes ont été développées pour déterminer les paramètres du modèle MA. La première en réalise une approximation par un modèle autorégressif d'ordre élevé, suivie par l'approximation en retour du modèle AR par un modèle MA qui sera la solution cherchée. Cette méthode s'apparente à celle établie pour le modèle MA stationnaire par Durbin [10], Shanks [48]. La seconde méthode repose sur la décomposition de Schur de la transformée en  $z$  de la suite des matrices estimant la corrélation du vecteur  $Y_t$  projection du signal  $y_t$ . En d'autres termes, il s'agira d'utiliser une forme de modélisation MA du vecteur  $Y_t$ . Le passage de cette décomposition ou de ce modèle vectoriel au modèle MA scalaire non stationnaire se fait par une sélection sur les lignes des matrices du modèle vectoriel, en maximisant la vraisemblance de covariances déduites des observations relativement à celles déduites du modèle. Une troisième méthode concernera un cas particulier, celui du modèle MA d'ordre zéro ! Modèle qui est tout à fait trivial dans le cas stationnaire, où le coefficient  $b_0$  s'identifie au gain du modèle AR et n'est autre que l'écart-type de l'erreur de prédiction. Modèle cependant où la non-stationnarité du coefficient  $b_0(t)$  forcera inéluctablement l'estimateur à la non-linéarité. La solution tiendra en une transformation non-linéaire des données après laquelle un estimateur linéaire constituera un bon compromis.

#### 3.1. RÉOLUTION PAR MOINDRES CARRÉES

Dans cette méthode, on approxime le modèle MA par un modèle AR d'ordre élevé. On effectue successivement l'identification de ce modèle AR par un des estimateurs décrits au paragraphe précédent, puis un filtrage inverse du signal  $y_t$  pour reconstituer l'entrée  $\varepsilon_t$  du modèle MA, et enfin le calcul du modèle MA connaissant simultanément son entrée et sa sortie.

On introduit alors comme pour le modèle AR un vecteur produit du signal par chaque fonction de la base, mais ici le produit est sur l'entrée. Soit  $E_t$  le vecteur obtenu :

$$(18) \quad E_t = [f_0(t) \varepsilon_t \dots f_m(t) \varepsilon_t]^T$$

Avec ce nouveau vecteur le modèle MA se réécrit :

$$(19) \quad y_t = [E_t^T \dots E_{t-q}^T] \theta + v_t$$

où le vecteur  $\theta$  contient les paramètres inconnus  $b_{ij}$  :  $\theta = [b_{00} \dots b_{qm}]^T$  et  $v_t$  représente l'erreur due à l'approximation de l'entrée par le résidu du modèle AR.

Si on minimise cette erreur  $v_t$ , au sens d'un critère des moindres carrés, cela signifie qu'on réalise l'estimation de  $\theta$  comme une identification de système avec bruit sur la sortie. Les équations deviennent, en utilisant directement une somme temporelle pour l'espérance mathématique :

$$(20) \sum_{t \in \Gamma} \begin{bmatrix} E_t \\ E_{t-1} \\ \vdots \\ E_{t-q} \end{bmatrix} [E_t^T \dots E_{t-q}^T] \theta = \sum_{t \in \Gamma} \begin{bmatrix} E_t \\ E_{t-1} \\ \vdots \\ E_{t-q} \end{bmatrix} y_t.$$

La matrice multipliant  $\theta$  dans (20) sera, suivant le choix de  $\Gamma$ , une matrice de Toëplitz ou bien le produit de deux matrices de Toëplitz.

### 3. 2. FACTORISATION MA NON STATIONNAIRE

La famille de méthodes d'identification de modèles MA ou ARMA, qui vient d'être décrite reposait sur une approximation par un modèle autorégressif d'ordre élevé. Celle qui va maintenant être décrite est construite autour de l'idée de factorisation spectrale, autre méthode de modélisation MA dans le cas stationnaire. Le passage au cas non stationnaire se fait par l'interprétation de cette factorisation spectrale comme un recours implicite à un modèle autorégressif dont on fait tendre l'ordre vers l'infini. Ces méthodes de factorisation spectrale sont issues des travaux de Schur [46], Rissanen [44], Le Roux et Gueguen [33], Youla et Kazanjian [51].

#### 3. 2. 1. Équivalence « AR long » et MA non stationnaire

Mettre cette démarche en œuvre dans la modélisation MA non stationnaire suppose donc de trouver un équivalent autorégressif au modèle MA dont les  $b_i(t)$  sont les coefficients. Le modèle autorégressif cherché s'écrira :

$$(21) \sum_{i=0}^{\infty} a_i(t-i) y_{t-i} = \varepsilon_t.$$

L'indépendance des  $\varepsilon_t$  entre eux permet en identifiant le modèle (21) au modèle MA de coefficients  $b_i(t)$ , d'écrire la relation :

$$(22) a_i(t) = - \sum_{j=1}^{\min(t, q)} a_{i-j}(t+j) \frac{b_j(t)}{b_0(t)}$$

avec  $a_0(t) = (b_0(t))^{-1}$ .

L'équivalence entre modèle MA d'ordre  $q$  et modèle AR d'ordre infini suppose ainsi deux conditions dans le cas non stationnaire qui sont déduites du calcul précédent et n'apparaissent pas dans le cas stationnaire. La première est que la normalisation  $a_0(t) = 1$

doit être abandonnée. La seconde est que la base de fonctions sur laquelle s'expriment les  $b_i(t)$  n'est pas la même que celle associée aux  $a_i(t)$ .

Le passage des  $b_i(t)$  aux  $a_i(t)$  se fait non seulement par l'accroissement de l'ordre  $q$  à l'infini, mais aussi par l'accroissement du degré  $m$  de la base à l'infini. Fort heureusement, ces deux effets n'ont pas la même amplitude, et l'accroissement du degré  $m$  se fait approximativement linéairement avec  $i$ , or les  $a_i(t)$  tendent vers 0 quand  $i$  croît, ce qui permet sans trop d'erreur de limiter ce degré à une valeur  $M$  finie, égale à deux ou trois fois  $m$ .

#### 3. 2. 2. Passage au modèle MA non stationnaire

Dans le cas stationnaire, la dualité entre l'algorithme de Levinson pour le calcul du modèle autorégressif, et l'algorithme de Schur pour le calcul du modèle MA ou la factorisation spectrale, s'explique en remarquant que les coefficients calculés à la  $k$ -ième itération de l'algorithme de Schur sont les intercorrélations entre l'entrée et la sortie du modèle autorégressif d'ordre  $k$ . Lorsque l'algorithme de Schur est appliqué à la corrélation d'un signal MA, on obtient, en faisant tendre  $k$  vers l'infini, la convergence de l'entrée du modèle autorégressif implicite d'ordre  $k$ , vers le bruit blanc excitateur du modèle MA, tandis que l'intercorrélations tend vers la réponse impulsionnelle du modèle MA, et donc vers les coefficients  $b_i$  eux-mêmes. C'est cette idée que l'on tente de réutiliser ici.

Observant alors que le modèle autorégressif non stationnaire sur une base de fonctions de dimension  $M$  fait appel au vecteur  $Y_t$  de dimension  $M$ , produit du signal  $y_t$  par la base, on prend comme point de départ ce vecteur qui est MA ( $q$ ) comme  $y_t$ , mais n'est pas stationnaire. La décomposition de Schur de sa corrélation conduira à des matrices  $E_i$  qui s'interprètent comme des matrices d'intercorrélations entre  $Y_t$  et l'entrée  $\xi_t$  d'un modèle AR ( $\infty$ ) vectoriel.

$$(23) E_i = E(\xi_t, Y_{t+i}^T)$$

De plus [18], il existe une relation entre l'entrée du modèle MA scalaire de  $y_t$  et celle du modèle MA vectoriel de  $Y_t$  : l'entrée  $\varepsilon_t$  du modèle MA scalaire est obtenue comme combinaison linéaire des composantes de  $\xi_t$  avec  $\varepsilon_t = L^T \xi_t$ . Utilisant alors l'expression de  $Y_t$  en fonction de  $y_t$ , et la nature MA de  $y_t$ , on en déduit une relation entre les matrices  $E_i$  et le modèle MA cherché :

$$(24) E_i^T L = \sum_t \begin{bmatrix} f_0(t+i) \\ \vdots \\ f_m(t+i) \end{bmatrix} [f_0(t) \dots f_m(t)] \begin{bmatrix} b_{i0} \\ \vdots \\ b_{im} \end{bmatrix}.$$

Ainsi peut-on calculer le modèle MA non stationnaire, dès que  $L$  est connu, à partir de la factorisation de  $Y_t$ , vecteur des projections de  $y_t$  sur la base  $[f_0(t) \dots f_m(t)]$ .

#### 3. 2. 3. Réduction du signal vectoriel au signal scalaire

Le seul point restant à préciser est le calcul du vecteur  $L$  permettant de passer du modèle vectoriel de  $Y_t$  au

modèle scalaire non stationnaire. Il faut pour cela réintroduire l'information : le vecteur  $Y_t$  est obtenu à partir du signal  $y_t$  à moyenne ajustée, d'ordre  $q$ , ou encore, le modèle AR implicite n'est pas quelconque, mais approxime l'inverse d'un modèle MA. Une façon de le faire est de contraindre les matrices de covariances du signal  $y_t$  ou  $Y_t$  à être égales à la même matrice calculée théoriquement à partir du modèle MA (qui dépend de  $L$ ). On ne pourra pas en réalité obtenir l'égalité entre la matrice mesurée et la matrice théorique, c'est-à-dire reconstruite à partir du modèle, du fait des approximations (il faut bien arrêter la récurrence sur  $p$  à un certain point, et de plus  $M$  étant limité supérieurement, l'évolution temporelle du modèle AR n'est pas exactement celle de l'inverse du modèle MA). Il suffira alors de maximiser la vraisemblance de l'hypothèse d'égalité des deux matrices. Anderson [3], donne l'expression de la vraisemblance  $V$  de l'hypothèse  $R=R^*$  où  $R$  et  $R^*$  sont deux matrices de covariance de dimension  $n$ , calculées sur  $N$  points :

$$(25) \quad \text{Log } V = N/2 \cdot (n + \text{Log Det}(R \cdot (R^*)^{-1}) - \text{Tr}(R \cdot (R^*)^{-1})).$$

La relation (25) peut être rendue plus symétrique sous la forme :

$$(26) \quad \text{Log } V = N/2 \times (n + \text{Log Det}((R^*)^{-1/2} \cdot R \cdot (R^*)^{-T/2}) - \text{Tr}((R^*)^{-1/2} \cdot R \cdot (R^*)^{-T/2})).$$

Dans le problème posé ici,  $R$  sera la matrice de covariance théorique, estimée à partir des  $b_j(t)$ , et  $R^*$  la même matrice, mais mesurée sur  $y_t$ . Plusieurs choix sont possibles pour cette matrice de covariance, par exemple :

$$R = E \left( \begin{bmatrix} Y_t \\ \vdots \\ Y_{t-q} \end{bmatrix} [Y_t^T \dots Y_{t-q}^T] \right).$$

Bien entendu, l'espérance est dans  $R^*$  remplacée par une somme temporelle, ce qui contraint, vu la non-stationnarité de  $y_t$ , à effectuer la même somme temporelle dans  $R$ . La matrice  $R$  est une fonction quadratique des  $b_{ij}$  et donc des  $\lambda_i$ , les composantes du vecteur  $L$ . On peut donc écrire la matrice  $(R^*)^{-1/2} R (R^*)^{-T/2}$  figurant dans le critère de vraisemblance comme :

$$(27) \quad C = (R^*)^{-1/2} R (R^*)^{-T/2} = \sum_{i,j=0}^M \lambda_i \lambda_j C_{ij}.$$

Les matrices  $C_{ij}$  dépendent uniquement des  $E_k$ ,  $k \in [0, q]$  et de  $R^*$  mesuré (voir [18] pour leur expression). Le problème est maintenant ramené à la maximisation de la quantité  $\text{Log Det}(C) - \text{Tr}(C)$  où  $C$  est une matrice carrée, fonction quadratique des  $\lambda_i$  :

$$(28) \quad V(\lambda_0 \dots \lambda_M) = \text{Log Det}(C) - \text{Tr}(C)$$

avec

$$C(\lambda_0 \dots \lambda_M) = \sum_{i,j=0}^M \lambda_i \lambda_j C_{ij} \quad (C_{ij} = C_{ij}^T = C_{ji}).$$

La présence du logarithme et du déterminant rendent cependant impossible cette maximisation de façon directe. Il faudra recourir à un algorithme du type Newton-Raphson pour la recherche du maximum de la fonction. Il est pour cela nécessaire de calculer son gradient et son Hessian. Le calcul montre en posant :

$$C_{ij}^* = C^{-1/2} C_{ij} C^{-T/2}$$

$$C_i = 2 \sum_{j=0}^M C_{ij}$$

$$C_i^* = 2 \sum_{j=0}^M C_{ij}^*$$

que le gradient et le Hessian sont :

$$(29) \quad \left\{ \begin{array}{l} \frac{\partial}{\partial \lambda_i} V(\lambda_0 \dots \lambda_M) = \text{Tr}(C_i^*) - \text{Tr}(C_i) \\ \frac{\partial^2}{\partial \lambda_i \partial \lambda_j} V(\lambda_0 \dots \lambda_M) \\ \quad = \text{Tr}(2C_{ij}^* - C_i^* C_j^*) - 2 \text{Tr}(C_{ij}). \end{array} \right.$$

On aura intérêt à démarrer cette partie itérative de l'algorithme en choisissant des valeurs initiales  $L_0$  voisines de l'optimum, ce qui pourra se faire en utilisant les valeurs de  $b_{ij}$  données par la méthode utilisant explicitement le modèle AR long et en calculant  $L$  au moyen de (24). Il est évident que cet algorithme consomme beaucoup plus de temps de calcul et de place mémoire que l'algorithme explicite, mais sa précision est nettement supérieure.

### 3.3. GAIN NON STATIONNAIRE

Il existe un cas très particulier du modèle MA, c'est le modèle MA(0). Dans le cas stationnaire, un tel modèle est réduit à un gain qui s'estime en général sans difficulté. Dans le cas non stationnaire, ce gain responsable des fluctuations de la variance  $\sigma^2(t)$  du résidu  $\varepsilon_t$  est une fonction du temps et il est possible de tenter de l'estimer comme le modèle MA habituel, avec l'hypothèse qu'il s'exprime comme combinaison linéaire des fonctions  $f_i(t)$  de la base. Aucune des méthodes MA décrites dans ce paragraphe ne permet cependant de calculer le gain  $b_0(t)$  tel que  $\sigma^2(t) = (b_0(t))^2$ . En effet les deux méthodes décrites font appel à un modèle AR long (explicite ou implicite) alors que le gain non stationnaire n'est pas équivalent à un modèle AR, ce qui contraint à développer une méthode spécifique pour ce cas.

Le problème étant non linéaire, il faut se ramener par une transformation des données à un problème linéaire. Un moyen consiste à remplacer  $\varepsilon_t$  par  $|\varepsilon_t|$ .

## RECHERCHES

La variable  $|\varepsilon_t|$  reste blanche, mais sa moyenne devient non nulle et vaut :

$$(30) \quad E(|\varepsilon_t|) = \sigma_t (2/\pi)^{1/2} = b_0(t) (2/\pi)^{1/2}.$$

Il est alors possible de calculer  $b_0(t)$  approximant au sens des moindres carrés cette moyenne, le critère sera :

$$(31) \quad J = \sum_t (|\varepsilon_t| - (2/\pi)^{1/2} b_0(t))^2.$$

Ce qui conduit, en introduisant les  $f_i(t)$ , au système :

$$(32) \quad \left( \sum_t \begin{bmatrix} f_0(t) \\ \vdots \\ f_m(t) \end{bmatrix} [f_0(t) \dots f_m(t)] \right) \begin{bmatrix} b_{00} \\ \vdots \\ b_{0m} \end{bmatrix} = (\pi/2)^{1/2} \sum_t \begin{bmatrix} f_0(t) \\ \vdots \\ f_m(t) \end{bmatrix} |\varepsilon_t|$$

Bien que le gain  $b_0(t)$  ainsi calculé soit la plupart du temps satisfaisant, on peut regretter qu'il lui soit possible de s'annuler, et même d'être négatif, par exemple si  $f_i(t) = t^i$ ,  $b_0(t)$  devient un polynôme dont la valeur peut sur certains intervalles de temps être en dessous de 0. Un moyen d'y remédier consiste à poser que le gain  $b_0(t)$  est en réalité une exponentielle :

$$(33) \quad b_0(t) = e^{g(t)} \quad \text{avec} \quad g(t) = \sum_{i=0}^m g_i f_i(t).$$

Il s'agit ici aussi de rendre le problème linéaire, ce qui peut se faire en retenant comme signal  $\text{Log}|\varepsilon_t|$ . L'espérance mathématique devient :

$$(34) \quad E(\text{Log}|\varepsilon_t|) = \text{Log}(b_0(t)) - \frac{C + \text{Log} 2}{2}$$

où  $C$  est la constante d'Euler,  $C = 0,577 215 665$ . Le critère d'estimation sera :

$$(35) \quad J = \sum_t \left( \text{Log}|\varepsilon_t| + \frac{C + \text{Log} 2}{2} - g(t) \right)^2.$$

Ce qui donne en utilisant l'expression de  $g(t)$  sur la base de fonctions  $f_i(t)$  :

$$(36) \quad \sum_t \begin{bmatrix} f_0(t) \\ \vdots \\ f_m(t) \end{bmatrix} [f_0(t) \dots f_m(t)] \begin{bmatrix} g_0 \\ \vdots \\ g_m \end{bmatrix} = \sum_t \begin{bmatrix} f_0(t) \\ \vdots \\ f_m(t) \end{bmatrix} \left( \text{Log}|\varepsilon_t| + \frac{C + \text{Log} 2}{2} \right).$$

Il est clair que ces deux estimateurs (31) et (36), bien que pris au sens des moindres carrés, ne sont pas optimaux car ni  $|\varepsilon_t|$ , ni  $\text{Log}|\varepsilon_t|$  ne sont gaussiens,

contrairement à  $\varepsilon_t$ . Cependant, dans la pratique, la dynamique de  $b_0(t)$  ou  $g(t)$ , qui est le rapport entre les amplitudes extrêmes prises par  $b_0(t)$  ou  $g(t)$ , est réduite, restant de l'ordre de 1 et ces estimateurs sont amplement suffisants.

### 4. Modèles en treillis

Les méthodes décrites dans les paragraphes précédents, concernant les parties, autorégressive puis à moyenne ajustée, du modèle ARMA peuvent être considérées comme une formulation standard de la modélisation non stationnaire. Tout en restant dans le cadre des modèles dont les paramètres s'expriment sur une base de fonctions, il est possible de réaliser des variations autour de la solution standard. Deux d'entre elles consistent la matière de ce paragraphe.

#### 4.1. FILTRE EN TREILLIS, COEFFICIENTS DE RÉFLEXION

La première méthode considérée dans ce paragraphe a son origine dans le modèle autorégressif classique avec son entrée en bruit blanc. La variante porte sur la forme de la réalisation du modèle, entraînant un remplacement du filtre transverse de coefficients  $a_i$ , par un filtre en treillis dont les coefficients sont les  $k_i$ , corrélations partielles. Le filtre en treillis, introduit à partir du modèle AR stationnaire par Itakura et Saito [30], se présente sous deux formes, d'une part un filtre « blanchissant » qui calcule le résidu  $\varepsilon_t$  du modèle AR à partir du signal  $y_t$ , et le filtre inverse, effectuant la synthèse de  $y_t$ . C'est le premier type qui interviendra ici, le seul pour lequel se pose la question de l'identification dans le cadre autorégressif.

Le filtre en treillis est très directement lié à l'algorithme de Levinson. Aux modèles AR( $p$ ) direct et rétrograde (forward, backward) sont associés respectivement les deux résidus  $\varepsilon_p^+(t)$  et  $\varepsilon_p^-(t)$ . Le filtre en treillis est une transcription du filtre autorégressif transverse en une structure de filtre constituée d'une cascade de  $p$  cellules, ayant en entrée les résidus d'ordre  $i-1$ , et en sortie, ceux d'ordre  $i$  (fig. 1).

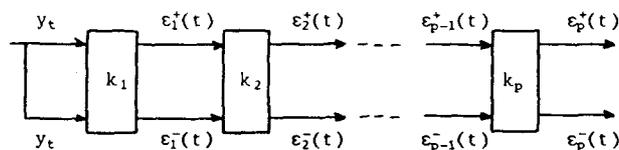


Fig. 1. — Structure du filtre en treillis.

Dans un contexte non stationnaire, les  $a_i$  et les  $k_i$  deviennent des fonctions du temps. La relation entrée-sortie pour chaque cellule est :

$$(37) \quad \begin{bmatrix} \varepsilon_i^+(t) \\ \varepsilon_i^-(t) \end{bmatrix} = \begin{bmatrix} 1 & k_i^+(t-1) \\ k_i^-(t) & 1 \end{bmatrix} \times \begin{bmatrix} \varepsilon_{i-1}^+(t) \\ \varepsilon_{i-1}^-(t) \end{bmatrix}.$$

Pour interpréter cette équation sous forme d'une cellule du filtre en treillis, il devient nécessaire d'introduire par rapport au cas stationnaire un second retard (fig. 2).

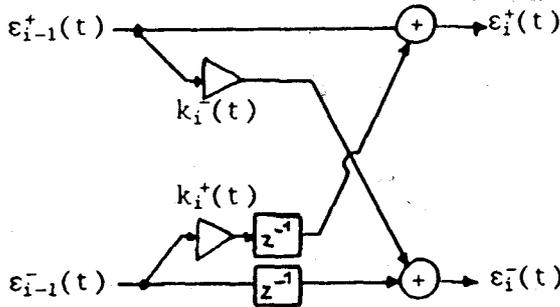


Fig. 2. — Structure d'une cellule à 2 retards.

L'identification du modèle (37) sera bien entendu calquée sur les méthodes existant dans le cas stationnaire, en postulant que chaque coefficient de réflexion s'exprime sur la base de fonctions :

$$(38) \quad \left\{ \begin{array}{l} k_i^+(t) = \sum_{j=0}^m k_{ij}^+ f_j(t) \\ \text{et} \\ k_i^-(t) = \sum_{j=0}^m k_{ij}^- f_j(t) \end{array} \right.$$

Comme dans la méthode de Burg [4], on cherchera à optimiser les cellules l'une après l'autre en initialisant avec  $\varepsilon_0^+(t) = \varepsilon_0^-(t) = y_t$ . Le calcul de la  $i$ -ième cellule se fait en minimisant l'énergie des sorties qui sont des erreurs de prédiction. Si les données  $y_t$  sont mesurées sur l'intervalle  $[0, T]$ , le critère  $J$  est :

$$(39) \quad J = \sum_{t=i}^T ((\varepsilon_i^+(t))^2 + \varepsilon_i^-(t))^2$$

Introduisant les vecteurs  $K_i^+$  et  $K_i^-$  des paramètres  $k_{ij}^+$  et  $k_{ij}^-$ , et les projections  $H_i^+$  et  $H_i^-$  des résidus  $\varepsilon_i^+(t)$  et  $\varepsilon_i^-(t)$ , sur la base des fonctions  $f_i(t)$ , il est possible de réécrire  $J$  :

$$(40) \quad J = \sum_{t=i}^T (\varepsilon_{i-1}^+(t) + (K_i^+)^T H_{i-1}^-(t-1))^2 + \sum_{t=i}^T (\varepsilon_{i-1}^-(t-1) + (K_i^-)^T H_{i-1}^+(t))^2$$

avec

$$K_i^+ = \begin{bmatrix} k_{i0}^+ \\ \vdots \\ k_{im}^+ \end{bmatrix}, \quad K_i^- = \begin{bmatrix} k_{i0}^- \\ \vdots \\ k_{im}^- \end{bmatrix},$$

$$H_i^+(t) = \begin{bmatrix} f_0(t) \varepsilon_i^+(t) \\ \vdots \\ f_m(t) \varepsilon_i^+(t) \end{bmatrix}, \quad H_i^-(t) = \begin{bmatrix} f_0(t) \varepsilon_i^-(t) \\ \vdots \\ f_m(t) \varepsilon_i^-(t) \end{bmatrix}.$$

La minimisation de  $J$  relativement aux paramètres  $k_{ij}^+$  et  $k_{ij}^-$  aboutit à deux systèmes d'équations distincts :

$$(41) \quad \Phi_i^+ \cdot K_i^+ = -\Psi_i^+ \quad \text{et} \quad \Phi_i^- \cdot K_i^- = -\Psi_i^-$$

avec

$$\Phi_i^+ = \sum_{t=i}^T H_{i-1}^-(t-1) (H_{i-1}^-(t-1))^T$$

$$\Phi_i^- = \sum_{t=i}^T H_{i-1}^+(t) (H_{i-1}^+(t))^T$$

$$\Psi_i^+ = \sum_{t=i}^T H_{i-1}^-(t-1) \varepsilon_{i-1}^+(t)$$

$$\Psi_i^- = \sum_{t=i}^T H_{i-1}^+(t) \varepsilon_{i-1}^-(t-1).$$

Il est clair que les vecteurs  $K_i^+$  et  $K_i^-$  généralisent les coefficients de corrélation partielle, car les équations (41) les présentent comme des « rapports » entre les intercorrélations  $\Psi$  et les autocorrélations  $\Phi$ .

#### 4. 2. FONCTIONS D'ARE LOGARITHMIQUES

Ce troisième type de modèle, après le modèle autorégressif et le filtre en treillis, s'est avéré nécessaire sous la pression d'une application particulière qui est la synthèse de la parole, mais bien sûr, il n'est en rien spécifiquement lié au signal de parole. Ce qui était recherché, était un modèle qui reste toujours stable. Rien ne garantit en effet la stabilité du modèle autorégressif évolutif que l'on estime sur un signal. Il est même facile de construire des signaux tels que si on en sélectionne une portion, le modèle estimé sur celle-ci est instable. La stabilité se définit, rappelons le comme une sommabilité absolue de la réponse impulsionnelle  $h(t, s)$  du modèle, et ceci pour tout  $t$ , la sommation se faisant sur  $s$ . Dans la pratique, en analysant un signal qui s'annule aux deux extrémités de la fenêtre, on est assuré de trouver un modèle stable. C'est le cas par exemple, d'un segment de parole compris entre deux zones de silence (cf. section 5).

Pourtant, si le modèle tangent au modèle évolutif devient instable sur un petit intervalle de temps  $(t_1, t_2)$ , et même si la stabilité du modèle non stationnaire n'en est pas affectée, ceci a de graves inconvénients, ainsi en synthèse de parole, le signal synthétisé présentera sur l'intervalle  $(t_1, t_2)$  un comportement explosif, son amplitude divergera, ce qui créera une petite bouffée d'énergie à un niveau plusieurs fois supérieur à l'énergie moyenne avant et après cet incident. Il s'avère que l'oreille est très sensible à ces variations brèves d'amplitude qu'elle analyse comme des bruits de choc, ou des explosions. Il est donc nécessaire d'éliminer ces accidents.

Une méthode pour forcer la stabilité du modèle sans déformer son contenu spectral consiste à remplacer à tout instant où il est instable le modèle tangent par

un modèle équivalent où les pôles stables sont inchangés, mais où les pôles instables sont stabilisés, par une inversion qui les renvoie à l'intérieur du cercle unité. Une autre méthode, développée ici, consiste à remplacer les coefficients de réflexion du filtre en treillis, qui ne sont pas contraints, quand on les estime sur une base de fonctions, par l'estimateur (41), à rester dans l'intervalle  $[-1, +1]$ . Une application de cet intervalle sur l'intervalle  $]-\infty, +\infty[$  permettra de garantir la stabilité, lorsqu'on estimera les paramètres transformés : par construction, les coefficients de réflexion que l'on reconstruira ne pourront qu'être bornés en module par 1. La transformation la mieux adaptée semble être le recours aux fonctions d'aire logarithmiques (ou Log Area Ratios), définis par (42), et exprimés sur la même base de fonctions :

$$(42) \quad \left\{ \begin{array}{l} \gamma_i(t) = \text{Log} \left( \frac{1+k_i(t)}{1-k_i(t)} \right) \\ \text{et} \\ \gamma_i(t) = \sum_{j=0}^m \gamma_{ij} f_j(t). \end{array} \right.$$

On reconstitue les coefficients de réflexion à partir des fonctions d'aire logarithmiques par la relation inverse (43), exprimée au moyen d'une tangente hyperbolique :

$$(43) \quad k_i(t) = th \left( \frac{\gamma_i(t)}{2} \right)$$

Deux méthodes ont été considérées pour l'estimation de fonctions d'aire logarithmiques [5]. La première recourt à une linéarisation, et va être décrite ci-après. La seconde maximise la vraisemblance du modèle, mais présente l'inconvénient lorsque les  $k_i(t)$  sont trop proches de  $+1$  ou  $-1$ , de ne converger que très lentement, ce qui lui fait préférer la première procédure, moins précise, mais beaucoup plus rapide. A cause de la tangente hyperbolique figurant dans (43), il n'est pas possible de minimiser l'erreur quadratique moyenne en sortie d'une cellule du filtre, sans recourir à une méthode d'optimisation non linéaire. Pour éviter ceci, il s'agit de linéariser le problème. On cherchera à approximer la trajectoire des fonctions d'aire logarithmiques  $\gamma_i(t)$ , à partir de celles déduites, par la formule (42), d'une trajectoire de coefficients de réflexion. Ainsi s'obtient une procédure en deux étapes. La première étape réalise l'estimation d'un modèle en treillis, par l'estimateur (41). La seconde transforme ce modèle en fonctions d'aire logarithmiques, en ajustant au sens des moindres carrés les deux trajectoires. L'estimateur qui s'en déduit est donné par l'équation (44), où  $F(t) = [f_0(t) \dots f_m(t)]^T$  :

$$(44) \quad \sum_t F(t) F^T(t) \begin{bmatrix} \gamma_{i0}^+ & \gamma_{i0}^- \\ \vdots & \vdots \\ \gamma_{im}^+ & \gamma_{im}^- \end{bmatrix} = \sum_t F(t) \left[ \text{Log} \left( \frac{1+k_i^+(t)}{1-k_i^+(t)} \right) \text{Log} \left( \frac{1+k_i^-(t)}{1-k_i^-(t)} \right) \right].$$

Les sommations ci-dessus sont effectuées sur tous les instants où le membre de droite est défini, c'est-à-dire tous les instants où les coefficients de réflexion, direct et rétrograde, sont tous deux, en module, inférieurs à 1. Si  $T$  est le nombre d'échantillons, ce que la première étape (filtre en treillis) requiert est de l'ordre de  $m^2 T/2$  opérations, la seconde, de l'ordre de  $2mT$  opérations, car la matrice figurant au membre de gauche ne dépend que de la base et non du signal : elle peut être calculée exactement et avant toute modélisation. L'accroissement du coût de calcul lorsque l'on passe du filtre en treillis aux fonctions d'aire est largement contrebalancé par la certitude que le modèle obtenu sera stable tout au long de l'intervalle d'analyse. En synthèse de parole, par exemple, ceci éliminera la procédure de stabilisation du modèle, qui est très coûteuse.

## 5. Applications

Il eut été dommage de ne pas conclure cette description d'une méthodologie pour le traitement des signaux non stationnaires, sans décrire son application à un certain nombre de signaux issus de situations réelles et concrètes. Une application est suffisamment avancée pour que les expériences se soient stabilisées autour de résultats probants, qui vont être présentés dans ce qui suit : il s'agit du codage suivi de restitution de la parole par unités supra-phonémiques, dans un contexte de type messagerie vocale.

Puisque les modèles évolutifs sont à même de représenter une transition entre contenus spectraux différents, il a semblé naturel de tenter de réaliser une synthèse de parole où chaque modèle serait associé à la transition d'un phonème à un autre, ce qui correspond à une synthèse par diphones. On peut en détecter les zones où le signal a un contenu spectral stable au cours du temps, segmenter le signal en leur milieu. L'analyse se fera sur chaque segment commençant et finissant à un des points de segmentation. A la synthèse le problème sera de concaténer correctement les divers segments ou diphones synthétisés. Cette dernière opération est cependant délicate, car elle se fait dans les zones les plus énergétiques du signal où les défauts sont très audibles. Cette difficulté nous a conduits à nous tourner maintenant vers une synthèse utilisant des entités plus longues que les diphones, en interdisant la segmentation et donc la juxtaposition, au milieu des voyelles [6].

Il ne s'agit pas d'une véritable synthèse, mais dans un premier temps d'une analyse du signal suivie par une restitution de ce signal à partir des paramètres issus de l'analyse. En court-circuitant ainsi la longue phase d'acquisition du catalogue des modèles, on occulte un certain nombre de difficultés associées au choix du bon représentant de chaque élément du dictionnaire, à la concaténation de modèles issus de contextes très différents... Mais esquiver ces problèmes permet de se concentrer plus vite sur ceux liés à

la détermination des modèles, et qui sont les premiers à résoudre quand on souhaite, comme ici, mettre en œuvre une méthodologie nouvelle dont on ignore le comportement sur un signal de parole.

Chaque expérience comportera les mêmes phases :

(A) Segmentation, c'est-à-dire détection des limites entre syllabes. Dans un premier temps, on détermine à partir de l'histogramme des énergies de segments courts (32 échantillons), un seuil de détection silence/parole, qui permet donc de détecter les zones de parole. On ajoute à ces zones, celles contenant des fricatives ou des plosives non voisées, qui ont pu échapper à la détection précédente, en les repérant au moyen du taux de passage par zéro supérieur à un seuil fixe. Dans une troisième phase, cette segmentation est affinée, par un ensemble de règles permettant d'éliminer les trop courtes portions de silence (le plus souvent, entre la plosive et la voyelle qui suit). On obtient ainsi des zones de parole, séparées par des silences. La succession de segments qui en résulte est prête à être traitée, mais si des segments ont une longueur excédant 2 500 échantillons, on redécoupe ces unités en segments tous d'une durée inférieure à 2 500 échantillons, en détectant les instants où l'énergie est minimale, ceci pour atténuer les effets éventuels de la concaténation, opération pour laquelle on prévoit de plus un recouvrement de 64 échantillons entre deux segments non séparés par du silence.

(B) Analyse de chaque segment, de façon à obtenir un modèle autorégressif évolutif pour chaque segment. C'est à cette étape que se situe la variable essentielle dans cette expérience, à savoir le choix de la structure du modèle. Celle-ci se caractérise par :

(B1) Le choix de l'ordre de la partie autorégressive. On peut s'aider pour celui-ci de la connaissance acquise dans la modélisation stationnaire sous fenêtres brèves où l'ordre AR est compris entre 10 et 16. La figure 3 montre l'application sur un segment de parole du critère d'Akaike, pour quatre degrés  $m$  différents.

(B2) Le choix de la base de fonctions et de son degré  $m$ . Deux bases seront utilisées ici : Legendre et Fourier, mais ceci ne préjuge pas de l'expérience d'une base de fonctions idéalement adaptée à la parole. Le degré  $m$  est la grande inconnue, il devra être au moins 1, pour avoir avec les deux fonctions  $f_0(t)$  et  $f_1(t)$  une valeur moyenne et une transition. Sur l'exemple de la figure 3, la valeur optimale de  $m$  est 5, pour un ordre 10.

(B3) Le choix de l'algorithme et du type de modèle. Le modèle peut avoir trois formes, du moins pour sa partie autorégressive qui peut être transverse avec les prédicteurs  $a_i(t)$  ou en treillis avec soit les coefficients de réflexion  $k_i(t)$ , soit les fonctions d'aire  $\gamma_i(t)$ . Dans l'estimation du modèle autorégressif, on retiendra pour son compromis qualité de l'estimateur/coût de calcul, la méthode de « corrélation ». Les deux autres modèles sont estimés par les estimateurs (41) et (44).

(C) Calcul d'une excitation, qui peut être obtenue avec une détection du fondamental, soit par une mesure de sa hauteur  $F_0$ , soit par une localisation

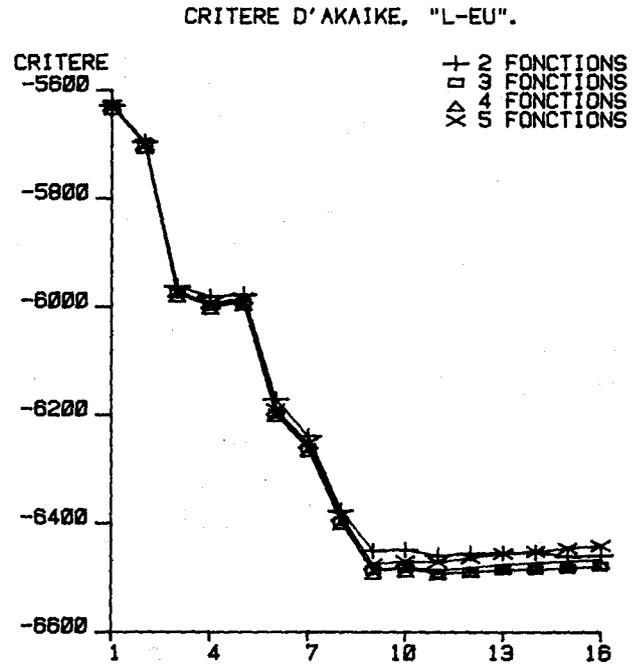


Fig. 3. — Application du critère d'Akaike à la détermination de l'ordre du modèle (fonctions d'aire logarithmiques), et du nombre de fonctions dans la base (polynômes de Legendre). Le segment de parole contient le diphone « LE ».

temporelle des impulsions. Une variante est l'excitation de type multi-impulsionnelle, qui s'utilise maintenant couramment avec les méthodes stationnaires, et pour lequel nous avons élaboré une extension non stationnaire [7].

(D) Synthèse (ou plutôt restitution) du signal par passage de l'excitation (impulsions/bruit blanc ou multi-impulsions) dans le système correspondant à chaque segment. Les segments ainsi reconstruits sont alors concaténés, avec soit insertion des silences intermédiaires, soit (lorsqu'il n'y avait pas de silence), passage progressif d'un signal au suivant, au moyen d'une pondération en cosinus (de  $t$ , le temps) sur la zone où les deux segments successifs se recouvrent. Ceci réduit fortement les parasites dus à la concaténation.

Quelles étapes de ce processus ont engendré des difficultés spécifiques aux modèles non stationnaires ? Sûrement par l'étape A, indépendante de la modélisation. L'étape B ne présente pas plus de difficultés. On prendra seulement la précaution de pré-accentuer le signal de parole, pour réduire sa dynamique spectrale et améliorer le conditionnement de la matrice de covariance du signal  $Y_t$ , des projections sur la base. La pré-accentuation retenue a été la même que dans la LPC classique : on remplace le signal  $y_t$  par sa différence première  $y_t^* = y_t - y_{t-1}$  (le signal est échantillonné à 8 khz, sur 12 e. b). Les ordres et le degré de la base ont été fixés pour chaque expérience, en les choisissant dans les limites données ci-dessus. Les meilleurs résultats avaient été obtenus pour la base de Legendre, un modèle d'ordre 12 sur les fonctions  $f_0(t) \dots f_4(t)$ . La phase C soulève dans le cas multi-

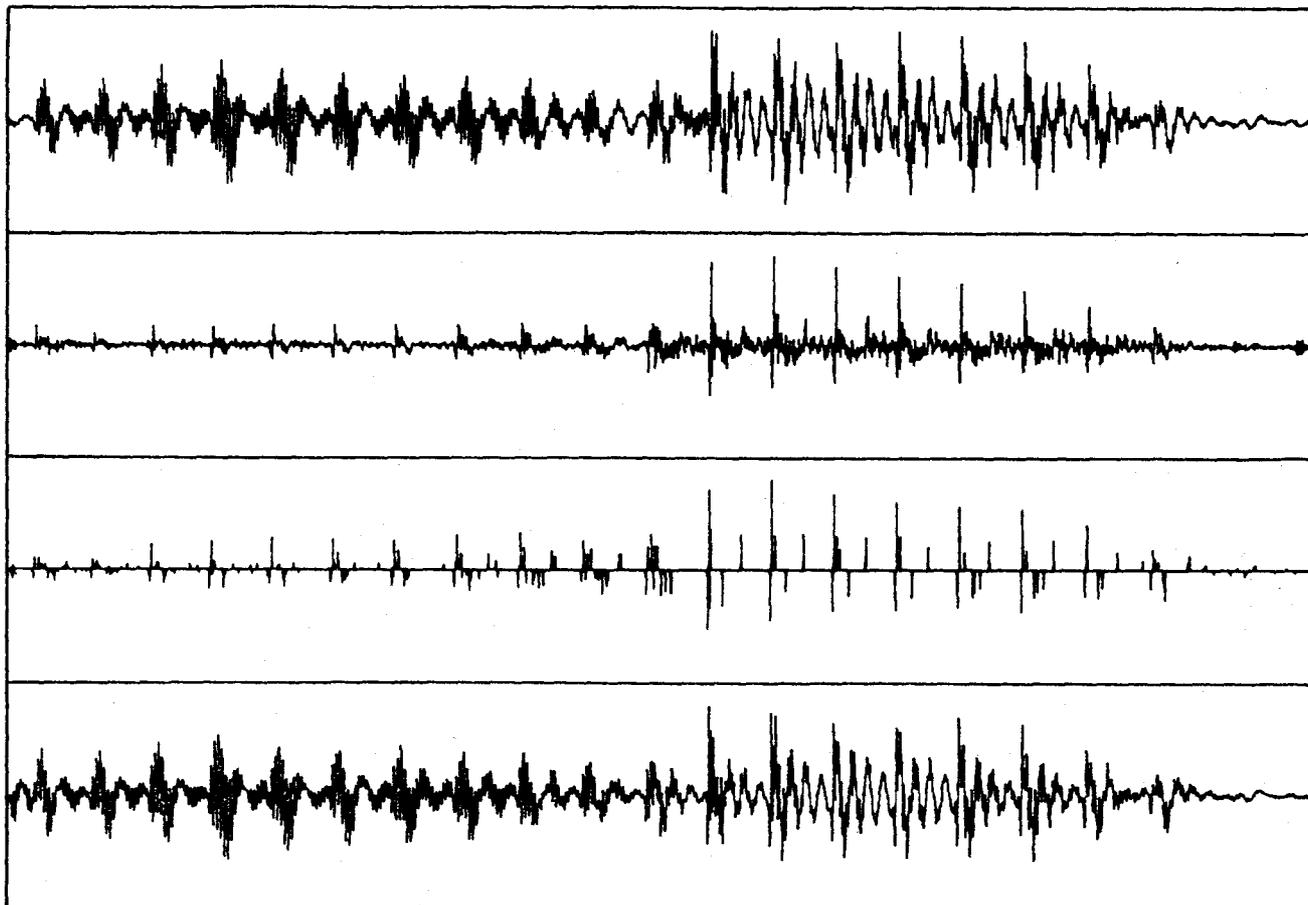


Fig. 4. — Analyse et restitution de parole. De haut en bas, le signal original « LE », le résidu ou erreur de prédiction, le codage multi-impulsionnel de l'entrée, et le signal synthétisé.

impulsionnel des problèmes de temps de calcul et de taille de matrices qui n'ont pu être résolus qu'en rendant la résolution séquentielle [7]. Mais les difficultés essentielles se sont situées dans la phase D, de synthèse, tant que nous travaillions avec les modèles AR ou en treillis, du fait de la stabilité non garantie par ces modèles. Le recours au modèle des fonctions d'aire logarithmiques s'est donc imposé pour cette raison.

Les résultats, à l'écoute de cette synthèse semblent satisfaisants. Les fenêtres ont une durée moyenne comprise entre 800 et 2000 échantillons, les ordres des modèles sont en moyenne de 12, avec 3 à 6 fonctions dans la base de Fourier. Ceci assure un codage du contenu spectral du signal sur environ 250 paramètres par seconde, ce qui est à mettre en parallèle avec les 500 paramètres par seconde que représente la LPC-10 (prédiction linéaire à l'ordre 10, un modèle toutes les 20 ms). Ceci fait espérer un gain probable d'un facteur 2 sur cette méthode, avec une qualité analogue sinon meilleure. La figure 4 montre un exemple de segment avec le résidu, l'excitation multi-impulsionnelle calculée par la méthode décrite en [7], et le signal reconstruit.

Pour ne pas allonger inutilement cet article, seule l'application à la restitution de parole a été décrite

ici. Cependant diverses études ont montré que la méthodologie des modèles AR ou ARMA évolutifs pouvait aussi s'appliquer à l'analyse temps-fréquence de nombreux signaux. Des simulations dans des cas difficiles peuvent être trouvées dans [22], pour des signaux dont l'évolution n'est pas continue, ce qui contredit l'hypothèse d'existence d'une base de fonctions, aussi dans [21], où la loi de fréquence instantanée de signaux très fortement bruités a pu être reconstituée grâce à des modèles AR évolutifs, ainsi que dans [24], où est présentée une analyse de signaux multi-composantes, qui sont des émissions sonars produites par des chauves-souris.

#### BIBLIOGRAPHIE

- [1] N. A. ABDRAHBO, A new approach to the problem of estimating spectral parameters of non-stationary time-series models, *J. of Comput. And Appl. Mathematics*, 5, n° 2, 1979, p. 125-129.
- [2] H. AKAIKE, Information theory and an extension of the maximum likelihood principle, *2nd Int. Symp. on Inform. Th.*, Tsakhadsor, Arménie, 1971.

- [3] T. W. ANDERSON, *An introduction to multivariate statistical analysis*, J. Wiley, New York, 1958.
- [4] J. P. BURG, A new analysis technique for time-series data, *NATO Adv. Study Inst. on Sig. Proc. with emphasis on Underwater Acoustics*, 1968.
- [5] M. C. CHEVALIER et Y. GRENIER, Autoregressive models with time-dependent Log Area Ratios, *IEEE-ICASSP 85*, Tampa, Florida, 1985, p. 1049-1052.
- [6] M. C. CHEVALIER, G. CHOLLET et Y. GRENIER, Speech analysis and restitution using time-dependent autoregressive models, *IEEE-ICASSP 85*, Tampa, Florida, 1985, p. 501-504.
- [7] M. C. CHEVALIER-OMNES, Y. GRENIER et G. CHOLLET, Codage multi-impulsionnel pour la restitution de parole par modèles évolutifs, *Colloque GRETSI sur le Traitement du signal et ses Applications*, Nice, 1985, p. 887-892.
- [8] H. CLERGEOT, Estimation du spectre d'un signal aléatoire gaussien par le critère du maximum de vraisemblance ou du maximum de probabilité a posteriori, *Thèse de Doctorat d'État*, Université de Paris-Sud, 1982.
- [9] H. CLERGEOT, Filter order selection in adaptive maximum likelihood estimation, *IEEE Trans. on IT*, 30, n° 2, 1984, p. 199-210.
- [10] J. DURBIN, Efficient estimation of parameters in moving-average models, *Biometrika*, 49, 1959, p. 306-316.
- [11] W. GERSCH, Estimation of the autoregressive parameters of a mixed autoregressive moving-average time-series, *IEEE Trans. on AC*, 15, n° 5, 1970, p. 583-588.
- [12] W. GERSCH et KITAGAWA, A time-varying multivariate autoregressive modeling of econometric time series, 1982, Submitted to *JASA*.
- [13] J. T. GINN, Time-varying autoregressive signal models with an application to chirped signals, *ICASSP 82*, 1982, p. 335-338.
- [14] D. GRAUPE, D. J. KRAUSE et J. B. MOORE, Identification of autoregressive moving-average parameters of time-series, *IEEE Trans. on AC*, 20, n° 1, 1975, p. 104-107.
- [15] Y. GRENIER, Estimation de spectres rationnels non-stationnaires, *Colloque GRETSI sur le Traitement du signal et ses Applications*, Nice, 1981-a, p. 185-192.
- [16] Y. GRENIER, Rational non-stationary spectra and their estimation, *1st ASSP Workshop on Spectral Estimation*, Hamilton, Ontario, 1981-b, p. 6.8.1-6.8.8.
- [17] Y. GRENIER, Identification de modèles ARMA non stationnaires, *Séminaire « Algorithmes Rapides pour le Traitement des Systèmes Linéaires »*, RCP CNRS-567, Aussois, 1981-c, p. 10.1-10.20.
- [18] Y. GRENIER, Estimation of non-stationary moving-average models, *IEEE ICASSP-83*, 1983-a, p. 268-271.
- [19] Y. GRENIER, Time-dependent ARMA modeling of non-stationary signals, *IEEE Trans. on ASSP*, 31, n° 4, 1983-b, p. 899-911.
- [20] Y. GRENIER, Modélisation de signaux non stationnaires, *Thèse de Doctorat d'État*, Université de Paris-Sud, 1984-a.
- [21] Y. GRENIER, Time-frequency analysis using time-dependent ARMA models, *IEEE ICASSP-84*, 1984-b.
- [22] Y. GRENIER et D. ABOUTAJDINE, Comparaison des représentations temps-fréquence de signaux présentant des discontinuités spectrales, *Annales des Télécommunications*, 38, n° 11-12, 1984, p. 429-442.
- [23] Y. GRENIER, Estimation simultanée AR et MA d'un modèle non stationnaire, *Colloque GRETSI sur le Traitement du signal et ses Applications*, Nice, 1985-a, p. 41-45.
- [24] Y. GRENIER, Nonstationary signal modeling with application to Bat echolocation calls, *Acustica*, 61, n° 3, 1986.
- [25] M. HALL, A. V. OPPENHEIM et A. WILLSKY, 1977, Time-varying parametric modelling of speech, *IEEE Decision and Control Conf.*, New Orleans, 1977, p. 1085-1091.
- [26] M. HALL, A. V. OPPENHEIM et A. WILLSKY, Time-varying parametric modelling of speech, *Signal Processing*, 5, n° 3, 1983, p. 267-285.
- [27] E. J. HANNAN et B. G. QUINN, The determination of the order of an autoregression, *J. Royal Statist. Soc., Series B*, 41, n° 2, 1979, p. 190-195.
- [28] M. J. HINICH et R. ROLL, Measuring non-stationarity in the parameters of a linear model with applications to asset return, *Rapport AD. A 042137*, 1977.
- [29] T. C. HSIA et D. A. LANDGREBE, On a method for estimating power spectra, *IEEE Trans. on Instr. and Measurement*, 16, n° 3, 1967, p. 255-257.
- [30] F. ITAKURA et S. SAITO, Digital filtering techniques for speech analysis and synthesis, *7th Int. Congr. Acoustics*, Budapest, 1971, p. 261-264.
- [31] F. KOZIN, Estimation and modelling of non-stationary time-series, *Symp. on Appl. of Computer Methods in Eng.*, Los Angeles, 1, 1977, p. 603-612.
- [32] F. KOZIN et F. NAKAJIMA, The order determination problem for linear time-varying AR models, *IEEE Trans. on AC*, 25, n° 2, 1980, p. 250-257.
- [33] J. LE ROUX et C. GUEGUEN, A fixed point computation of partial correlation coefficients, *IEEE Trans. on ASSP*, 25, n° 3, 1977, p. - .
- [34] L. A. LIPORACE, Linear estimation of non-stationary signals, *J. Acoust. Soc. Amer.*, 58, n° 6, 1975, p. 1288-1295.
- [35] G. MARTINELLI, Non-stationary AR model identification by batch estimation, *IEEE ICASSP-83*, 1983, p. 265-267.
- [36] J. M. MENDEL, 1969, A priori and a posteriori identification of time-varying parameters, *2nd Hawaii Int. Conf. on Syst. Sciences*, 1969, p. 207-210.
- [37] J. M. MENDEL, Discrete techniques of parameter estimation: the equation error formulation, *Control Theory*, 1, chap. 6: estimation of time-varying parameters, Marcel Dekker Inc., 1973.
- [38] T. NAKAJIMA et T. SUZUKI, A non-steady state speech analysis method, *Progress Report on Speech Research*, PIPS-R-27, 1980-a, p. 14-17.
- [39] T. NAKAJIMA et T. SUZUKI, Non-steady state speech analysis by recursive method, *Progress Report on Speech Research*, PIPS-R-27, 1980-b, p. 34-38.
- [40] T. NAKAJIMA, T. SUZUKI et H. OHMURA, Non-steady state speech analysis method with dynamic feature enhancing effect, *IEEE ICASSP-82*, 1982, p. 1299-1302.
- [41] A. PEARSON, Identification of linear time-varying differential systems amid modal disturbances, *6th IFAC Symp. on Ident. and Syst. Param. Estim.*, 1982, p. 925-930.
- [42] T. PITARQUE, Étude d'algorithmes de filtrage et de lissage en traitement du signal: application à l'extraction de potentiels évoqués, *Thèse de 3<sup>e</sup> cycle*, Université de Nice, 1984.
- [43] T. S. RAO, The fitting of non-stationary time-series models with time-dependent parameters, *J. of the Royal Statist. Soc., Series B*, 32, n° 2, 1970, p. 312-322.
- [44] J. RISSANEN, Algorithms for triangular decomposition of block Hankel and Toeplitz matrices with application to factoring positive matrix polynomials, *Math. of Comput.*, 27, n° 121, 1973, p. 147-154.

## RECHERCHES

- [45] J. RISSANEN, Modeling by shortest data description, *Automatica*, 14, n° , 1978, p. 465-471.
- [46] J. SCHUR, 1917, Uber Potenzreihen, die im Innern des Einheitskreises beschränkt sind, *Z. Reine Angew. Math.*, 147, 1917, p. 205-232, et 148, 1918, p. 122-145.
- [47] G. SCHWARZ, Estimating the dimension of a model, *The Annals of Statistics*, 6, n° 2, 1978, p. 461-464.
- [48] J. L. SHANKS, Recursion filters for digital processing, *Geophysics*, 32, n° 1, 1967, p. 33-51.
- [49] X. XIANYA et R. J. EVANS, 1967, Discrete-time stochastic adaptive control for time-varying systems, *IEEE Trans. on AC*, 29, n° 7, 1984-a, p. 638-640.
- [50] X. XIANYA et R. J. EVANS, Discrete-time adaptive control for deterministic time-varying systems, *Automatica*, 20, n° 1, 1984-b, p. 5-17.
- [51] D. C. YOULA, N. N. KAZANJIAN, 1978, Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle, *IEEE Trans. on CAS*, 25, n° 2, 1978, p. 57-69.