

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



## Ecole Nationale Polytechnique

Département d'Electronique

Thèse de Doctorat d'Etat en Electronique,

Option : Signal et Communications.

# Modèles d'Ordres Réduits pour la Synthèse du Signal Parole

par Lahcène MITICHE

(Ingénieur et Magister Polytechnicien)

Soutenue le : 17 Février 2005

Devant le jury :

- M. Haddadi (Pr.) (Président, ENP)
- D. Berkani (Pr.) (Rapporteur, ENP)
- F. Boudjema (Pr.) (Examineur, ENP)
- A. Guessoum (Pr.) (Examineur, Univ. de Blida)
- M. Guerti (M.C.) (Examinatrice, ENP)

Thèse de Doctorat d'Etat:

MODELES D'ORDRES REDUITS POUR LA  
SYNTHESE DU SIGNAL PAROLE

Lahcène MITICHE <sup>1</sup>

Laboratoire Signal et Communications,

Département d'Electronique,

Ecole Nationale Polytechnique.

<sup>1</sup>Ingénieur d'état (Electronique, ENP, Juin 1991), Magister (Electronique, ENP, Octobre 1994).



# Remerciements

J'adresse mes remerciements à Monsieur Daoud Berkani, professeur au département d'Electronique de l'Ecole Nationale Polytechnique, et responsable du laboratoire Signal et Communications, de m'avoir proposé et d'avoir dirigé ce travail.

Je remercie vivement Monsieur le professeur Mourad Haddadi de l'ENP d'avoir accepté de présider le jury, ainsi que Messieurs les professeurs Abderrezak Guessoum de l'Université de Blida et Fares Boudjema de l'ENP, Madame Docteur Mhania Guerti de l'ENP, d'avoir accepté d'examiner mon travail.

Je tiens à témoigner toute ma reconnaissance à ma femme Amel.

*A ma femme Amel Baha Houda, et à mes filles Nour et Rym.*

# Notation

$n$  : Ordre du modèle *AR* original.

$r$  : Ordre du modèle d'ordre réduit.

$m$  : Ordre du modèle *ARMA*.

$k$  : Variable temporelle discrete.

$z$  : Variable complexe fréquentielle.

$(A, B, C, D)$  : Matrices d'état du modèle numérique d'ordre  $n$ .

$(A_r, B_r, C_r, D_r)$  : Matrices d'état du modèle numérique d'ordre réduit  $r$ .

$x(k)$  : Vecteur d'état discret.

$u(k)$  : Signal d'entrée discret.

$y(k)$  : Signal de sortie du modèle d'ordre  $n$ .

$y_r(k)$  : Signal de sortie du modèle d'ordre réduit  $r$ .

*AR* : Modèle Auto regressif.

*ARMA* : Modèle Auto regressif à Moyenne Adaptée.

*LSF* : Line spectral frequencies.

*SNR* : Signal to noise ratio (Rapport signal sur bruit).

$\Sigma$  : Matrice des valeurs singulières.

*SVD* : Singular values decomposition (Décomposition en valeurs singulières).

$\sigma_i$  : ième valeur singulière.

$\lambda(\cdot)$  : Valeur propre de  $(\cdot)$ .

$I_p$  :  $p \times p$  Matrice identité.

$F_s$  : Fréquence d'échantillonnage.

# Table des matières

<b>1</b>	<b>PRODUCTION ET TRAITEMENT DU SIGNAL DE LA PAROLE</b>	<b>15</b>
1.1	Introduction . . . . .	15
1.1.1	L'appareil phonatoire . . . . .	16
1.1.2	Fonctionnement général . . . . .	17
1.2	Modélisation du système de production de la parole . . . . .	19
1.3	Conclusion . . . . .	23
<b>2</b>	<b>MODELISATION DU SIGNAL DE PAROLE</b>	<b>25</b>
2.1	Introduction :	
	Modélisation d'un signal aléatoire stationnaire . . . . .	25
2.1.1	Modélisation d'un signal . . . . .	25
2.1.2	Prédiction linéaire . . . . .	26
2.1.3	Variance de l'erreur de prédiction . . . . .	26
2.1.4	Estimation des coefficients de prédiction . . . . .	27
2.1.5	Interprétation de la prédiction linéaire . . . . .	29



2.1.6	Filtre inverse . . . . .	29
2.1.7	Prédiction avant et prédiction arrière . . . . .	30
2.2	Modélisation AR du signal de parole . . . . .	31
2.2.1	Introduction . . . . .	31
2.2.2	Energie résiduelle de prédiction . . . . .	32
2.2.3	Systèmes d'équations . . . . .	32
2.2.4	Méthode de l'autocorrélation . . . . .	34
2.2.5	Méthode de la covariance . . . . .	35
2.2.6	Algorithme de <i>Burg</i> . . . . .	36
2.3	Les paramètres <i>LSF</i> (Line Spectral Frequencies) . . . . .	37
2.4	Conclusion . . . . .	38
<b>3</b>	<b>REDUCTION D'ORDRE DE MODELES PAR PROJECTIONS IN-</b>	
	<b>TERNES</b>	<b>39</b>
3.1	Introduction . . . . .	39
3.2	Méthode des Réalisations Equilibrées <i>MRE</i>	
	(Réalisations minimales) . . . . .	40
3.2.1	Introduction . . . . .	40
3.2.2	Hypothèses . . . . .	41
3.2.3	Algorithme de la réduction d'ordre par la <i>MRE</i> . . . . .	41
3.2.4	Conclusion . . . . .	43
3.3	Méthode de <i>Schur</i> (réalisations non minimales) . . . . .	44
3.3.1	Procédure de la simplification par <i>Schur</i> . . . . .	44

3.3.2	Algorithme de <i>Schur</i> modifié . . . . .	47
3.4	Conclusion . . . . .	50
<b>4</b>	<b>SIMULATIONS</b>	<b>53</b>
4.1	Introduction . . . . .	53
4.2	Exemple 1 . . . . .	54
4.3	Exemple 2 . . . . .	66
4.4	Exemple 3 . . . . .	75
4.5	Interprétations et conclusion . . . . .	84
	<b>Bibliographie</b>	<b>87</b>
	<b>A FACTORISATION DE CHOLESKY</b>	<b>93</b>
	<b>B VALEURS SINGULIERES ET ALGORITHME SVD</b>	<b>95</b>
	<b>C ALGORITHME D'EQUILIBRE (MODELES NUMERIQUES)</b>	<b>99</b>



# INTRODUCTION GENERALE :

## Position du problème

La parole est un phénomène essentiellement continu, les silences correspondent en général à des pauses de respiration dont l'occurrence est aléatoire. L'onde de la parole n'est ni Gaussienne, ni érgodique, ni stationnaire ; on peut faire cependant d'excellentes descriptions à court terme du comportement de l'appareil de production de la parole. Une caractéristique fondamentale du signal vocal est qu'il est très redondant. Cette redondance se traduit aussi par une surestimation de l'ordre du modèle classique *AR* de représentation du signal parole, ceci est justifié par le fait que le signal parole peut être synthétisé par la convolution du signal d'excitation avec la réponse impulsionnelle du filtre *AR*.

L'apport de notre travail est la modélisation mathématique du système de production de la parole par des modèles d'ordre réduit, ce qui revient à calculer, à partir d'un modèle classique *AR* pur d'ordre  $n$ , un autre modèle d'ordre réduit  $r$ , approximant avec une

faible erreur celui d'ordre complet. Pour cela, on exploitera le fait que le signal de parole est redondant, donc le modèle de représentation présente  $(n - r)$  modes ne contribuant que faiblement dans le comportement global du modèle.

L'idée est donc de faire ressortir ces modes (états) de faible contribution, en passant par la représentation d'état, afin de les éliminer par la suite par troncature. Pour faire ressortir ces modes non dominants, il nous faut choisir une représentation d'état spéciale, c'est à dire trouver une base de coordonnées adéquate où le vecteur d'état soit projeté. On opérera donc une réduction d'ordre par projections internes.

Dans notre recherche, un intérêt majeur est porté sur le comportement du critère de distorsion *SNR* (Signal to Noise Ratio) ou *RSB* (Rapport Signal sur Bruit) du modèle d'ordre réduit  $r$  comparé à celui du modèle *ARMA* de même ordre  $m = r$ , ainsi qu'au comportement des paramètres *LSF* et leur distribution géométrique dans le plan complexe. Nous avons élaboré une Toolbox (Speech Reduction) à base du *MATLAB* 6.5/Linux au Laboratoire de Physique Théorique et Modélisation (LPTM) de l'Université de Cergy Pontoise (France), pouvant à partir d'un fichier de parole (forme d'ondes) sous format \*.wav (speech.wav) calculer, visualiser ou comparer tous les paramètres du modèle classique *AR* de production de parole à ceux du modèle d'approximation d'ordre réduit, permettant ainsi une étude comparative des modèles d'ordre réduit  $r$  avec ceux *ARMA* synthétisés directement de même ordre  $m = r$ , et voir par la suite la qualité des signaux issus de ces deux modèles, les écouter, ou les comparer. Comme segments de parole, nous avons utilisé une base de données composée de syllabes de la langue française et des phrases en langue arabe et anglaise, sur lesquelles on opère dans un premier

temps une synthèse  $AR$  classique pour chaque segment, et par la suite une réduction d'ordre sur chacun des modèles, nous avons exploité aussi les avantages qu'apportent les modèles issus de la réduction d'ordre par rapport à ceux classiques de même ordre.

Il est possible, dans le futur proche, de développer la Toolbox sur un interface graphique, qui pourra être facilement exploitable que ce soit dans le domaine pédagogique (travaux pratiques par exemples), ou bien dans d'autres applications de recherche dans le domaine du "speech modeling".

Nos travaux de recherche ont été communiqués dans des conférences internationales [2,3] et ont été publiés dans des journaux internationaux [4,5,6]. La thèse est ordonnée comme suit :

Dans le chapitre 1, on donnera quelques notions sur la production et le traitement du signal de parole, dans le chapitre suivant, on traitera la modélisation  $AR$  du signal de la parole en donnant une description mathématique du système de production de la parole issu du modèle  $AR$  ainsi que les propriétés de ce dernier. Dans le chapitre 3, on développera divers algorithmes de réduction d'ordre par les projections internes, opérant tous dans l'espace d'état, permettant ainsi de calculer avec précision le modèle d'approximation d'ordre réduit à partir du modèle d'ordre complet. Des simulations sur des fichiers de parole phonétiquement équilibrés, ainsi que sur d'autres segments de parole, avec leurs interprétations sont données dans le chapitre 4.

Pour terminer notre travail, une conclusion générale, mettant en relief l'aboutissement de notre travail est présentée dans la fin de la thèse. En annexes, nous donnons la factorisation de *Cholesky* (Annexe A), l'algorithme *SVD* (Singular Values Decomposition)

en Annexe B, et l'algorithme d'équilibre des systèmes numériques en Annexe C.

# Chapitre 1

# PRODUCTION ET TRAITEMENT DU SIGNAL DE LA PAROLE

## 1.1 Introduction

Nous présentons dans ce chapitre, dans un premier volet, une description générale de l'appareil vocal humain, avec une modélisation du système de production de parole. Dans le deuxième volet, nous présenterons diverses méthodes de traitement du signal de parole, ainsi que leur domaines d'application.



### 1.1.1 L'appareil phonatoire

Schématiquement, l'appareil phonatoire humain [1] se compose de trois parties (Fig. 1.1)

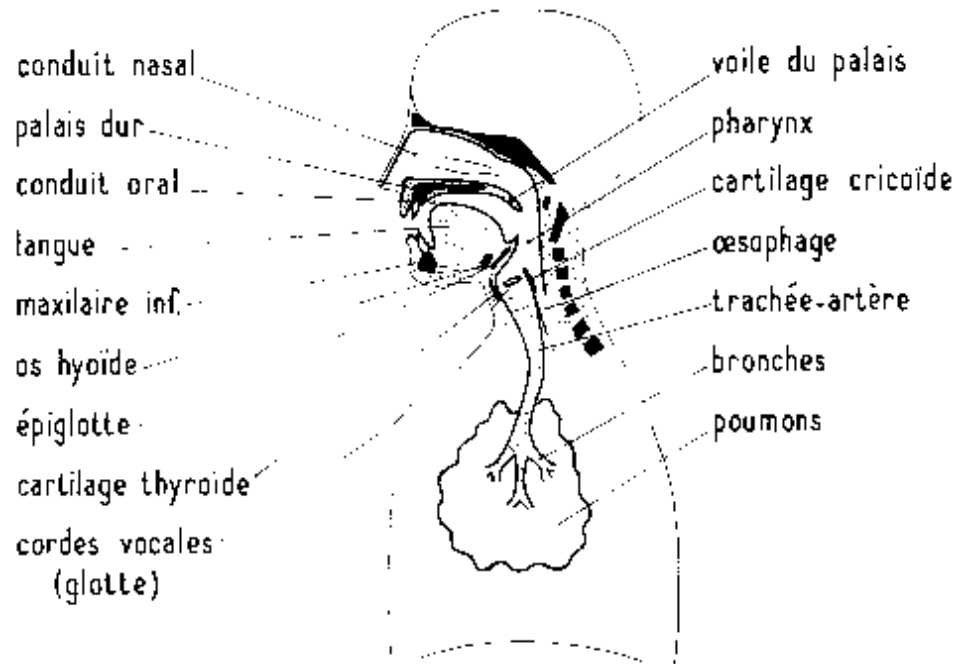


FIG. 1.1 – L'appareil phonatoire humain

#### Les poumons et le conduit trachéo-bronchique

Sous l'action des différents muscles abdominaux et notamment du diaphragme, la pression engendrée sous les poumons donne naissance à un flux d'air qui, via la trachée artère, se propage dans le larynx puis dans les cavités nasales et buccale.

#### Le larynx

Le larynx est le lieu de production de l'onde glottique, c'est-à-dire de la partie périodique

des sons voisés (présence du fondamental), cette onde est produite par les vibrations des cordes vocales (muscles élastiques du larynx) au passage du flux pulmonaire.

### **Le conduit vocal**

Le conduit vocal (pharynx, bouche, et cavité nasale) est la partie de l'appareil phonatoire qui imprime aux sons émis les caractéristiques spécifiques permettant de distinguer les divers phonèmes, ceci selon les deux fonctions suivantes :

- en tant que résonateur de l'onde glottique pour la production de voyelles,
- en tant que générateur de bruit pour la production des consonnes.

### **1.1.2 Fonctionnement général**

Les sons de la parole résultent de la rencontre du flux d'air pulmonaire avec un obstacle au niveau du conduit vocal.

En pratique, on distingue deux sources d'excitation de celui-ci.

#### **Excitation pseudo-périodique**

Dans ce cas, les cordes vocales oscillent suivant un phénomène de relaxation, des impulsions périodiques de pression sont ainsi appliquées au conduit vocal. Le conduit vocal peut être considéré comme une succession de tubes ou de cavités acoustiques de sections différentes.

Les sons voisés résultent donc de l'excitation du conduit vocal par des impulsions périodiques de pression liée aux oscillations des cordes vocales. Un son voisé est un signal quasi-périodique (Fig. 1.2-B), on y observe les raies correspondant aux harmoniques du fondamental  $F_0$  (structure du pitch). L'enveloppe de ces raies présente des maximums

appelés formants, qui correspondent aux fréquences propres du conduit vocal. Les trois premiers formants sont essentiels pour caractériser le spectre vocal, les formants d'ordre supérieur ont une influence plus limitée.

### **Excitation par un bruit blanc**

Lorsque la section devient trop faible en un point du conduit vocal, l'écoulement aérien crée des turbulences qui donnent naissance à une source de bruit, les sons ainsi émis sont de type non voisé.

Un son non voisé ne présente pas de structure périodique (Fig. 1.2-A), et peut être considéré comme un bruit blanc filtré par les transmittances de la partie du conduit vocal située entre la constriction et les lèvres, son spectre ne donne pas de structure du pitch.

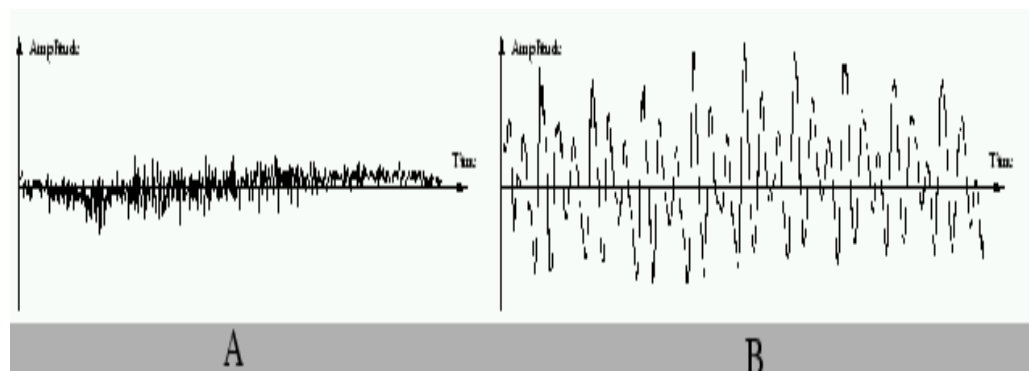


FIG. 1.2 – Signal voisé (B) et signal non voisé (A)

## 1.2 Modélisation du système de production de la parole

L'analyse d'un système complexe, requiert l'établissement préalable d'un modèle mathématique approprié permettant une mise en équation aussi aisée que possible. Dans le cas du système de production de parole, cette modélisation est basée sur l'étude de la propagation de l'onde acoustique, le long de l'axe du conduit vocal. De nombreuses recherches effectuées dans ce domaine ont montré que l'appareil phonatoire humain pouvait être modélisé, dans une première approximation, par un système linéaire, dont la fonction de transfert est composée de trois étages successifs, ces derniers correspondent à la modélisation de la source d'excitation du conduit vocal et de la cavité nasale et de la radiation des lèvres.

Nous allons décrire un tel schéma de modélisation et les résultats obtenus serviront par la suite à définir un modèle complet du système de production de la parole nécessaire pour une analyse efficace de ce dernier. La source d'excitation et le système de production de la parole peuvent être modélisés séparément à cause du faible couplage entre la glotte et le conduit vocal.

### **Modélisation de la source d'excitation**

Les différentes sources d'excitation du conduit vocal peuvent être modélisées par des générateurs d'impulsions pour les sons voisés et un bruit blanc pour les non-voisés. Pour les sons voisés, la source d'excitation d'un signal quasi-périodique est entièrement caractérisée par sa fréquence fondamentale  $F_0$ .

Soit  $E(z)$  et  $U_g(z)$  représentant respectivement le train d'impulsions périodique et la

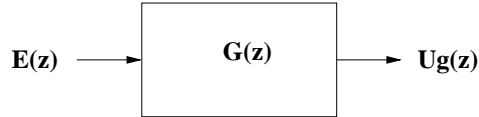


FIG. 1.3 – Modélisation de la source d'excitation pour les sons voisés

réponse à l'excitation périodique, et  $z$  étant la variable complexe fréquentielle. Le train d'impulsions périodique est modélisé par la réponse d'un filtre passe-bas d'ordre deux, à pôles réels, et dont la fréquence de coupure est de l'ordre de  $100Hz$ . Ce filtre est de la forme

$$G(z) = \frac{A}{(1 - \alpha z^{-1})(1 - \beta z^{-1})}, \quad (1.1)$$

où  $A$  est une constante,  $\alpha$  et  $\beta$  sont des pôles situés à l'intérieur du cercle unité.

Pour les sons non voisés, la source d'excitation du conduit vocal est dite 'bruité', et peut être modélisée par un générateur d'impulsions aléatoires produisant un bruit blanc.

### Modélisation du conduit vocal

Le conduit vocal est assimilé à une succession de tubes acoustiques élémentaires de diverses sections. L'étude de la propagation d'une onde acoustique plane le long d'un tel ensemble permet de dégager un modèle linéaire du conduit vocal. La fonction de transfert obtenue est celle d'un filtre tous pôles (Fig. 1.4), ce dernier permet de modéliser le conduit vocal d'une manière satisfaisante, et est donné par la fonction de transfert suivante

$$V(z) = \frac{U_v(z)}{U_g(z)} = \frac{B}{1 + \sum_{i=1}^{r-1} \nu_i z^{-i}}, \quad (1.2)$$

où  $B$  est une constante,  $\nu_i$  sont des coefficients réels,  $r_1$  est le nombre de pôles réels situés à l'intérieur du cercle unité,  $U_v$  est la réponse à l'excitation  $U_g$ .

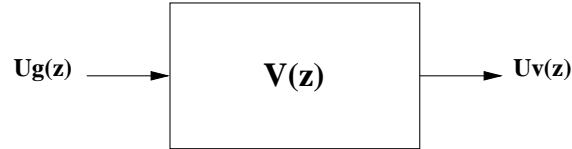


FIG. 1.4 – Modélisation du conduit vocal

### Modélisation de la radiation

Le son est finalement émis à travers l'ouverture des lèvres, celle-ci représente une charge acoustique, le rayonnement des lèvres peut être modélisé par la transmittance

$$R(z) = C(1 - z^{-1}), \quad (1.3)$$

où  $C$  est une constante.

### Modèle complet du système de production de la parole

Il est défini par la juxtaposition des trois modèles (Fig. 1.5), la fonction de transfert qui lui est associée est obtenue en mettant en cascade les fonctions de transfert de la glotte, du conduit vocal et de rayonnement aux lèvres (Fig. 1.5.a).

$$H(z) = G(z)V(z)R(z) = \frac{ABC(1 - z^{-1})}{(1 - \alpha z^{-1})(1 - \beta z^{-1})(1 + \sum_{i=1}^{r_1} \nu_i z^{-i})}. \quad (1.4)$$

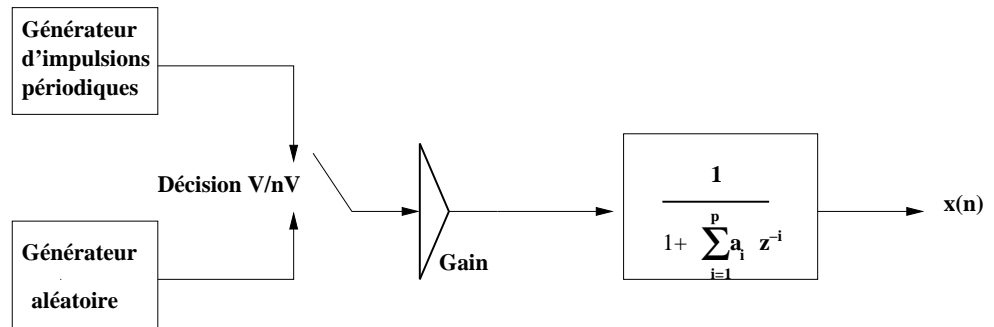


FIG. 1.5 – Modèle simplifié de production de la parole

Avec  $E_x(z)$  et  $S(z)$  sont les transformées en  $Z$  du signal d'excitation et du signal parole synthétisé. Si l'on considère que l'un des pôles de  $G(z)$  très voisin de l'unité, il vient

$$H(z) = \frac{\rho}{(1 - \alpha z^{-1})(1 + \sum_{i=1}^{r-1} \nu_i z^{-i})}, \quad (1.5)$$

en posant

$$A(z) = (1 - \alpha z^{-1})(1 + \sum_{i=1}^{r-1} \nu_i z^{-i}) = 1 + \sum_{i=1}^p a_i z^{-i}, \quad (1.6)$$

avec  $p = r1 + 1$ , et  $\rho = ABC$ , on obtient finalement un modèle *AR*, ou tout pôles, d'ordre  $p$ , de la forme

$$H(z) = \frac{\rho}{A(z)} = \frac{\rho}{1 + \sum_{i=1}^p a_i z^{-i}}. \quad (1.7)$$

### 1.3 Conclusion

Finalement, le modèle linéaire 'tous pôles' de représentation qu'on vient de décrire reste d'un intérêt majeur car ses paramètres peuvent être estimés d'une façon relativement simple et efficace grâce à la méthode d'analyse par prédiction linéaire notée *LPC* (Linear Predictive Coding).





## Chapitre 2

# MODELISATION DU SIGNAL DE PAROLE

### 2.1 Introduction :

#### Modélisation d'un signal aléatoire stationnaire

##### 2.1.1 Modélisation d'un signal

La modélisation d'un signal  $g(k)$  consiste à lui associer un filtre linéaire qui, soumis à une excitation particulière reproduit ce signal le plus fidèlement possible [1]. L'objectif essentiel de la modélisation d'un signal est de permettre la description de son spectre par un nombre fini de paramètres pertinents. La modélisation classique d'un signal est basée sur un filtre de fonction de transfert  $T(z) = B(z)/A(z)$  de degré  $(q)/(p)$  excité par

un bruit blanc  $b(k)$  de moyenne  $\nu_u$  nulle et de variance  $\sigma_u^2$  unitaire.

Dans le cas où ( $p > 0$  et  $q > 0$ ), le modèle est dit *ARMA* (Auto Regressive Moving Average). Lorsque  $B(z) = 1$  ( $p > 0, q = 0$ ), il s'agit d'un modèle *AR* (Auto Regressive) :  $T(z)$  est une fonction tous pôles. Lorsque  $A(z) = 1$  ( $p = 0, q > 0$ ), la fonction de transfert  $T(z)$  est celle d'un filtre *RIF*, et le modèle est dit *MA* (Moving Average).

### 2.1.2 Prédiction linéaire

La prédiction linéaire est basée sur l'hypothèse que chaque échantillon du signal original  $g(k)$  peut être estimé par une combinaison linéaire des  $p$  échantillons qui le précèdent

$$g(k) = - \sum_{i=1}^p a_i g(k-i) + e(k), \quad (2.1)$$

où les  $a_i$  sont les coefficients de prédiction d'ordre  $p$  et le signal

$$e(k) = \sum_{i=0}^p a_i g(k-i); a_0 = 1, \quad (2.2)$$

est l'erreur de prédiction (résidu) d'ordre  $p$ .

### 2.1.3 Variance de l'erreur de prédiction

Nous supposons que le signal  $g(k)$  est aléatoire et stationnaire ; les coefficients  $a_i$  sont donc indépendants du temps. L'estimation des coefficients de prédiction est basée sur la

minimisation de la variance de l'erreur de prédiction

$$\begin{aligned}
\sigma_e^2 &= E[e(k)^2] = E[\sum_{i=0}^p a_i g(k-i) \sum_{j=0}^p a_j g(k-j)] \\
&= E[\sum_{i=0}^p \sum_{j=0}^p a_i a_j g(k-i)g(k-j)] \\
&= E[\sum_{i=0}^p \sum_{j=0}^p a_i a_j \phi(i-j)],
\end{aligned} \tag{2.3}$$

où  $\phi(l)$  représente la fonction d'autocorrélation du signal  $g$ , et est définie par

$$\begin{aligned}
\phi(l) &= E[g(k)g(k+l)] = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{k=N+l}^{N-l} g(k)g(k+l), \\
\phi(0) &= \sigma_g^2.
\end{aligned} \tag{2.4}$$

Le vecteur des coefficients de prédiction d'ordre  $p$  sera noté

$$\mathbf{a} = \begin{bmatrix} 1 & a_1 & a_2 & \cdots & a_p \end{bmatrix}^T = \begin{bmatrix} 1 & \underline{\mathbf{a}}^T \end{bmatrix}^T, \tag{2.5}$$

et la matrice d'autocorrélation du signal  $g(k)$

$$\phi_g = \begin{bmatrix} \phi_g(0) & \phi_g(1) & \cdots & \phi_g(p-1) & \phi_g(p) \\ \phi_g(1) & \phi_g(0) & \cdots & \cdots & \phi_g(p-1) \\ \vdots & & & & \\ \vdots & & & & \\ \phi_g(p) & \cdots & \cdots & \phi_g(1) & \phi_g(0) \end{bmatrix}. \tag{2.6}$$

La variance de l'erreur de prédiction s'écrit alors

$$\sigma_e^2 = [1, \underline{\mathbf{a}}^T] \phi_g [1, \underline{\mathbf{a}}^T]^T. \tag{2.7}$$

Notons que la matrice d'autocorrélation a une structure de Toeplitz symétrique.

#### 2.1.4 Estimation des coefficients de prédiction

En dérivant l'expression (2.7) par rapport aux  $a_i$ , ( $i = \overline{1, p}$ ), on obtient un système d'équations linéaires en  $a_i$ , et la structure particulière de la matrice d'autocorrélation

conduit à un algorithme de résolution efficace et rapide basé sur une méthode récurrente sur l'ordre de prédiction : un vecteur  $\underline{a}$  d'ordre  $m$  va être calculé à partir de celui d'ordre  $m - 1$ , et ce pour  $m = \overline{1, p}$ . Dans ce qui va suivre, la matrice  $\phi_g$  sera notée  $\phi^{(p)}$  et le vecteur  $[1, \underline{a}]^T$  sera noté  $[1, \underline{a}_p]^T$  et nous posons

$$\phi^{(p)} = [\phi_g(1), \phi_g(2), \dots, \phi_g(p)]^T, \quad (2.8)$$

$$\phi^{(p)} = \begin{bmatrix} \sigma_g^2 & \phi^{(p)T} \\ \phi^{(p)} & \Phi^{(p-1)} \end{bmatrix}. \quad (2.9)$$

La forme quadratique (2.7) s'écrira

$$\sigma_e^2 = \sigma_g^2 + 2\phi^{(p)T} \underline{a}_p + \underline{a}_p^T \phi^{(p-1)} \underline{a}_p. \quad (2.10)$$

Nous avons

$$\frac{\partial \sigma_e^2}{\partial \underline{a}_p} = 2\phi^{(p)} + 2\phi^{(p-1)} \underline{a}_p = 0, \quad (2.11)$$

et le système à résoudre est

$$\phi^{(p-1)} \underline{a}_p = -\phi^{(p)}, \quad (2.12)$$

ou sous la forme développée

$$\phi_g(l) = \sum_{i=1}^p a_p(i) \phi_g(l-i), \quad (2.13)$$

avec  $\phi(0) = \sigma_g^2$ . La valeur minimale de la variance de l'erreur de prédiction est donc

$$\sigma_{e, \min}^2 = \sigma_g^2 + \phi_g^{(p)T} \underline{a}_p = \sum_{i=0}^p a_p(i) \phi_g(i) = \alpha_p. \quad (2.14)$$

D'autre part, les équations (2.12) et (2.14) donnent

$$\begin{bmatrix} \sigma_g^2 & \phi^{(p)T} \\ \phi^{(p)} & \phi^{(p-1)} \end{bmatrix} \begin{bmatrix} 1 \\ \underline{a}_p \end{bmatrix} = \begin{bmatrix} \alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (2.15)$$

### 2.1.5 Interprétation de la prédiction linéaire

L'équation (2.1) peut être définie comme un filtre excité par un signal  $e(k)$  et produisant un signal  $g(k)$ . Soient  $G(z)$  et  $E(z)$  les transformées en  $Z$  respectives de  $g(k)$  et  $e(k)$ , donc

$$G(z) = \frac{E(z)}{1 + \sum_{i=1}^p a_p(i)z^{-i}} = T(z)E(z). \quad (2.16)$$

La fonction de transfert  $T(z)$  est celle d'un filtre RII (tous pôles) ou modèle *AR*. Le signal élaboré par le modèle *AR* est en fait une approximation du signal original  $g(k)$ .

On introduit un facteur  $\delta$  dans la fonction de transfert du modèle, ce facteur est appelé gain du modèle, et est égal à  $\sqrt{\alpha_p}$ .

### 2.1.6 Filtre inverse

D'après l'équation (2.16), on a aussi

$$E(z) = A(z)G(z). \quad (2.17)$$

$A(z)$  est un filtre *RIF* qui, excité par  $g(k)$ , reproduit le signal  $e(k)$ , ce filtre est dit filtre inverse.

### 2.1.7 Prédiction avant et prédiction arrière

La prédiction avant prédit la valeur future  $g_F(k)$  à partir des  $p$  échantillons précédents  $g(k-i)$ ,  $i = \overline{1, p}$ , et la prédiction avant d'ordre  $m$  est

$$g_{Fm}(k) = - \sum_{i=1}^m a_m(i)g(k-i), \quad (2.18)$$

avec une erreur de prédiction avant d'ordre  $m$

$$f_m(k) = g(k) - g_{Fm}(k) = \sum_{i=0}^m a_m(i)g(k-i). \quad (2.19)$$

La prédiction arrière prédit (vérifie) une valeur passée  $g_B(k-p+1)$  à partir des mêmes échantillons, et la prédiction arrière d'ordre  $m$  est

$$g_{Bm}(k-m-1) = - \sum_{i=1}^m b_m(i)g(k-i), \quad (2.20)$$

avec une erreur de prédiction arrière d'ordre  $m$

$$s_m(k) = g(k-m+1) - g_{Bm}(k-m+1) = \sum_{i=1}^{m+1} b_m(i)g(k-i). \quad (2.21)$$

Les vecteurs de coefficients de prédiction avant et arrière d'ordre  $m$  sont respectivement avec une erreur de prédiction avant d'ordre  $m$

$$a_m = [1, a_m(1), a_m(2), \dots, a_m(m)] = [1, \underline{a}_m^T]^T, \quad (2.22)$$

$$b_m = [a_m(1), b_m(2), \dots, b_m(m), 1] = [\underline{b}_m^T]^T, \quad (2.23)$$

et les polynômes prédicteurs de l'erreur correspondants sont

$$A_m(z) = 1 + \sum_{i=1}^m a_m(i)z^{-i}, \quad (2.24)$$

$$B_m(z) = \sum_{i=1}^m b_m(i)z^{-i} + z^{-(m+1)}. \quad (2.25)$$

En appliquant la transformée en  $Z$  aux erreurs de prédiction, nous aurons

$$F_m(z) = A_m(z)G(z), \quad (2.26)$$

$$S_m(z) = B_m(z)G(z). \quad (2.27)$$

## 2.2 Modélisation AR du signal de parole

### 2.2.1 Introduction

La détermination d'un modèle *AR* pour le signal de parole est basée sur les résultats obtenus pour les signaux stationnaires. Il est clair que la modélisation de ce signal se présente dans un contexte particulier :

- Le signal de parole n'étant pas stationnaire, on va devoir travailler sur des tranches de temps de durées limitées, avec recouvrement : une procédure classique consiste à calculer un modèle pour des tranches successives de  $30ms$  décalées de  $10ms$ .
- Un signal voisé est engendré par un filtre *AR* excité par un train d'impulsions, et un signal non voisé est engendré par le même filtre *AR* excité par un bruit blanc (voir fig. 1.5).

Le critère pour l'estimation du modèle est la minimisation de la variance de l'erreur de prédiction [11,12] ; en fait, nous allons minimiser la somme des carrés des échantillons de l'erreur sur un intervalle de temps déterminé.



## 2.2.2 Energie résiduelle de prédiction

L'énergie résiduelle de prédiction est définie par la somme

$$E_p = \sum_{k=n_a}^{n_b} e^2(k), \quad (2.28)$$

dans laquelle il faut préciser les bornes  $n_a$  et  $n_b$  ; selon le choix qui est fait, la minimisation conduit à des systèmes linéaires semblables, mais non identiques.

L'hypothèse de base est de disposer de  $N$  échantillons indexés de  $h$  à  $h + N - 1$ , où  $h$  est l'origine pouvant être zéro si aucune ambiguïté n'est à craindre.

Selon (2.2) on a

$$\begin{aligned} E_p &= \sum_{k=n_a}^{n_b} [\sum_{i=0}^p a_p(i)g(k-i)]^2 \\ &= \sum_{i,j=0}^p a_p(i)a_p(j) \sum_{n=n_a}^{n_b} g(k-i)g(k-j). \end{aligned} \quad (2.29)$$

On pose

$$\phi_g(i, j) = \sum_{k=n_a}^{n_b} g(k-i)g(k-j); \quad i, j = \overline{1, p}, \quad (2.30)$$

et on obtient

$$E_p = \sum_{i,j=0}^p a_p(i)a_p(j)\phi_g(i, j). \quad (2.31)$$

## 2.2.3 Systèmes d'équations

La minimisation de (2.31) conduit au système

$$\sum_{j=0}^p \phi_g(i, j)a_p(j) = \phi_g(i, 0); \quad i = \overline{1, p}, \quad (2.32)$$

qui sous la forme compacte s'écrit

$$\Phi_g(p)a_p(j) = -\phi_g, \quad (2.33)$$

avec

$$\underline{\Phi}_g(p) = \begin{bmatrix} \phi_g(1,1) & \phi_g(1,2) & \cdots & \phi_g(1,p) \\ \phi_g(2,1) & \phi_g(2,2) & & \phi_g(2,p) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \phi_g(p,1) & \phi_g(p,2) & & \phi_g(p,p) \end{bmatrix}. \quad (2.34)$$

L'énergie résiduelle minimisée vaut

$$\alpha_p = \sum_{i=0}^p a_p(i) \phi_g(0, i). \quad (2.35)$$

Si on réunit (2.33) et (2.35), on obtient le système d'ordre  $(p+1)$

$$\Phi_g^{(p)} a_p = [\alpha_p, 0, 0, \dots, 0], \quad (2.36)$$

dans lequel on a posé

$$\Phi_g^{(p)} = \begin{bmatrix} \phi_g(0,0) & \phi_g(0,p) \\ \phi_g(1,0) & \phi_g(1,p) \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \phi_g(p,0) & \phi_g(p,p) \end{bmatrix} = \begin{bmatrix} \phi_g(0,0) & \phi_g^T \\ \phi_g & \underline{\Phi}_g(p) \end{bmatrix}. \quad (2.37)$$

## 2.2.4 Méthode de l'autocorrélation

Dans cette méthode, on prend  $n_a = -\infty$  et  $n_b = +\infty$ , cependant, pour respecter l'hypothèse sur le nombre fini d'échantillons, on posera  $g(k) = 0$  pour  $k < 0$  et  $k \geq N - 1$ . L'énergie résiduelle de prédiction est donc calculée sur l'ensemble des échantillons engendrés par le filtre inverse, qui est un filtre non récursif d'ordre  $p$ , nous aurons

$$E_p = \sum_{k=0}^{N-1+p} e^2(k). \quad (2.38)$$

D'autre part, la fonction d'autocorrélation (2.30) devient

$$\begin{aligned} \phi_g(i, j) &= \sum_{k=0}^{\infty} g(k-i)g(k-j) = \sum_{k=-\infty}^{\infty} g(k)g(k+i-j) \\ &= \phi_g(|i-j|) = \phi_g(l); \quad i, j = \overline{1, p}. \end{aligned} \quad (2.39)$$

En tenant compte sur l'hypothèse sur  $g(k)$ , on aura

$$\phi_g(l) = \sum_{k=0}^{N-1-l} g(k)g(k+l); \quad l = \overline{0, p}. \quad (2.40)$$

Cette expression correspond à une estimation de la fonction d'autocorrélation court terme, et la matrice du système devient

$$\underline{\Phi}_g(p) = \begin{bmatrix} \phi_g(0) & \phi_g(1) & \cdots & \phi_g(p-1) \\ \phi_g(1) & \phi_g(0) & & \phi_g(p-2) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \phi_g(p-1) & \phi_g(p-2) & & \phi_g(0) \end{bmatrix}, \quad (2.41)$$

et d'autre part

$$\phi_g = [\phi_g(1), \phi_g(1), \cdots, \phi_g(p)]^T. \quad (2.42)$$

Le système complet à résoudre est

$$\Phi_g^{(p)} a_p = [\alpha_p, 0, \dots, 0]^T. \quad (2.43)$$

Ainsi, grâce au fait que  $\phi_g(i, j)$  ne dépend que de  $|i-j|$ , la matrice  $\phi_g^{(p)}$  est une matrice de Toeplitz symétrique. Cependant, on peut observer que l'erreur de prédiction  $e(k)$  présente des valeurs exagérées en début de l'intervalle ( $k = \overline{0, p-1}$ ) car la prédiction porte sur un nombre insuffisant d'échantillons. Le même phénomène se produit en fin d'intervalle ( $k = \overline{N, N+p}$ ), où l'on tente de prédire des valeurs nulles pour  $g(k)$ .

Pour minimiser cet effet, on pondère les échantillons du signal vocal par une fenêtre  $w(k)$  symétrique, et on pose

$$g(k) = s(k)w(k)$$

où  $s(k)$  est le signal vocal avant fenêtrage.

La fenêtre la plus courante est celle de *Hamming* définie par

$$w_{Hm}(k) = \begin{cases} 0.54 - 0.46\cos(2\pi k/(N-1)) & \text{si } 0 \leq k \leq N-1, \\ 0 & \text{ailleurs.} \end{cases} \quad (2.46)$$

### 2.2.5 Méthode de la covariance

On pose  $n_a = p$  et  $n_b = N - 1$ . L'énergie résiduelle est calculée sur l'ensemble des échantillons de l'erreur de prédiction

$$E_p = \sum_{n=p}^{N-1} e^2(k). \quad (2.47)$$

Les effets d'extrémités évoqués dans la section précédente sont ici inexistantes ; la méthode n'exige aucune pondération préalable des échantillons.

L'expression (2.30) devient

$$\phi_g(i, j) = \sum_{k=p}^{N-1} g(k-i)g(k-j); i, j = \overline{0, p}. \quad (2.48)$$

On observe que  $\phi_g(i, j) \neq \phi_g(|i-j|)$ .

Nous aurons d'ailleurs

$$\begin{aligned} \phi_g(i+1, j+1) &= \sum_{k=p}^{N-1} g(k-i-1)g(k-j-1) \\ &= \sum_{k=p-1}^{N-2} g(k-i)g(k-j), \end{aligned} \quad (2.49)$$

d'où la récurrence

$$\phi_g(i+1, j+1) = \phi_g(i, j) + g(p-1-i)g(p-1-j) - \dots - g(N-1-i)g(N-1-j). \quad (2.50)$$

La matrice du système (2.43) reste symétrique et définie positive, mais n'est plus une matrice de Toeplitz.

### 2.2.6 Algorithme de Burg

L'algorithme de *Burg* est utilisé pour la mise à jour des coefficients de corrélation partielles  $k_m$  pour chaque nouvelle tranche du signal de parole.

• *Acquisition* :

$$g(k) = s(k)$$

• *Initialisation* :

$$\text{pour } k = \overline{0, N-1}$$

$$f_0(k) = g(k)$$

$$s_0(k) = g(k-1)$$

• *Recursion* :

pour  $m = \overline{0, p}$

$$k_m = \frac{\sum_{k=0}^{N-2+m} f_{m-1}(k) s_{m-1}(k)}{\sum_{k=0}^{N-2+m} f_{m-1}^2(k) + \sum_{k=0}^{N-2+m} s_{m-1}^2(k)}$$

pour  $k = \overline{0, N-2+m}$

$$f_m(k) = f_{m-1}(k) + k_m s_{m-1}(k)$$

$$s_m(k) = s_{m-1}(k-1) + k_m f_{m-1}(k-1) .$$

## 2.3 Les paramètres *LSF* (Line Spectral Frequencies)

Dans l'analyse *LPC* du signal de la parole, on segmente le fichier de parole en plusieurs sections, durant lesquelles le signal est supposé stationnaire, et on modélise chaque segment par un modèle *AR*.

Soit  $A(z)$  le filtre inverse de *AR*, donné par

$$A(z) = 1 + \sum_{i=1}^n a_i z^{-i}, \quad (2.51)$$

où les  $a_i$ , ( $i = \overline{1, n}$ ) sont les coefficients de prédiction. Le polynôme (2.51) est utilisé pour construire deux polynômes  $P_1(z)$  et  $P_2(z)$ , en augmentant l'ordre  $n$  à  $n+1$ . Ceci est équivalent à mettre le  $(n+1)$ me coefficient de réflexion  $r_{m+1}$  à  $\pm 1$ . En fait, les valeurs de  $r_{m+1}$ , respectivement à  $+1$  et  $-1$ , correspondent à maintenir le modèle du tube acoustique entièrement fermé ou entièrement ouvert [13, 14, 15].

Les polynômes  $P_1(z)$  et  $P_2(z)$  sont donnés par

$$P_1(z) = A(z) + z^{-(n+1)} A(z^{-1}), \quad (2.52)$$

$$P_2(z) = A(z) - z^{-(n+1)}A(z^{-1}). \quad (2.53)$$

Les solutions des équations (2.52) et (2.53) donnent les *LSF* [14,15], déterminant alors les fréquences angulaires des deux polynômes.

#### **Propriétés des polynômes $P_1(z)$ et $P_2(z)$**

1. Les racines des polynômes  $P_1(z)$  et  $P_2(z)$  se trouvent ordonnés sur le cercle unité, et sont entrelacées.
2. Le filtre  $A(z)$  maintient sa propriété de phase minimale si la propriété (1) est vérifiée.

## **2.4 Conclusion**

Nous avons traité dans ce chapitre la modélisation *AR* du signal de parole. Le critère pour l'estimation du modèle est la minimisation de la variance de l'erreur de prédiction, qui conduit à un système d'équations. Pour résoudre ce système d'équations, nous avons proposé deux méthodes (méthode de l'autocorrélation et la méthode de covariance), et l'algorithme de *Burg* permet de calculer les coefficients du modèle *AR* sans passer par la fonction d'autocorrélation. Dans ce qui va suivre, le modèle *AR* d'ordre complet  $n$  va être approximé par un modèle d'ordre réduit  $r$ , ceci en utilisant des approches opérants dans l'espace d'état.

## Chapitre 3

# REDUCTION D'ORDRE DE MODELES PAR PROJECTIONS INTERNES

### 3.1 Introduction

Dans ce chapitre, nous traiterons le problème de réduction d'ordre de modèles discrets d'ordre assez élevé, par les projections internes. Selon que le modèle soit minimal ou non minimal, nous détaillerons respectivement la méthode des réalisations équilibrées [19,20,21,22], et la méthode de *Schur* [23,24] avec ses deux variantes.

Le principe de base de simplification de modèles d'ordre élevé par projections internes



est le suivant [20]

- Ayant un modèle discret d'ordre ( $n$ ), représenté dans l'espace d'état par

$$\begin{aligned}x(k+1) &= Ax(k) + Bu(k), \\y(k) &= Cx(k) + Du(k),\end{aligned}\tag{3.1}$$

où  $x(k) \in R^n$  est le vecteur d'état,  $u(k) \in R$  est le signal d'entrée,  $y(k) \in R$  est le signal de sortie, et  $A \in R^{n \times n}$ ,  $B \in R^{n \times 1}$ ,  $C \in R^{1 \times n}$ ,  $D \in R^{1 \times 1}$  sont les matrices d'état.

- Calculer un modèle d'ordre ( $r < n$ ), représenté dans l'espace d'état par

$$\begin{aligned}x_r(k+1) &= A_r x_r(k) + B_r u(k), \\y_r(k) &= C_r x_r(k) + D_r u(k),\end{aligned}\tag{3.2}$$

où  $x_r(k) \in R^r$  est le vecteur d'état réduit,  $u(k) \in R$  est le signal d'entrée,  $y_r(k) \in R$  est le signal de sortie du modèle réduit, et  $A \in R^{r \times r}$ ,  $B \in R^{r \times 1}$ ,  $C \in R^{1 \times r}$ ,  $D \in R^{1 \times 1}$  sont les matrices d'état du modèle d'ordre réduit.

## 3.2 Méthode des Réalisations Equilibrées *MRE* (Réalisations minimales)

### 3.2.1 Introduction

Introduite par **Moore** [20], cette approche est basée sur une transformation d'état particulière dite "transformation d'équilibre" [25] qui rend symétriques certaines propriétés du système, du point de vue entrées/sorties, dans un sens énergétique [22]. La *MRE* repose sur la diagonalisation des grammians de commandabilité  $W_c$  et d'observabilité

$W_o$ , (voir Annexe C), telles-que leurs transformées dans la base d'équilibre (Annexe C) soient égales à une matrice diagonale dite "matrice des valeurs singulières". Le modèle d'ordre réduit est obtenu en éliminant les valeurs singulières (Annexe B) les plus faibles, correspondant aux états qui contribuent faiblement dans la réponse impulsionnelle du système d'ordre complet.

### 3.2.2 Hypothèses

On suppose que le système est représenté par un modèle linéaire, invariant dans le temps, minimal, strictement propre, asymptotiquement stable et représenté dans l'espace d'état par le système d'équations (3.1).

### 3.2.3 Algorithme de la réduction d'ordre par la *MRE*

Une fois la réalisation équilibrée construite (Annexe C), c'est-à-dire; ayant obtenu les paramètres d'ordre  $n$  ( $A_e, B_e, C_e, \Sigma, n$ ), les éléments de la matrice  $\Sigma$  doivent être ordonnés en décroissance de façon à guider le choix de l'ordre.

La réalisation du modèle simplifié s'effectue par élimination des valeurs singulières les plus faibles [21]. Le modèle simplifié résultera de la troncature d'ordrer de la réalisation représentée dans la base d'équilibre. Le critère d'erreur utilisé est celui de la norme de *Hankel* [26].

**Procédure de la réduction d'ordre par la Méthode des Réalisations Equi-brées [19,21]**

**Entrées-** Ayant une réalisation d'ordre  $n$  ( $A_e, B_e, C_e, \Sigma, n$ ).

**Etape 1-** Choix de l'ordre  $r < n$  du modèle simplifié satisfaisant la norme de *Hankel*, définie par [26]

$$\sqrt{\sum_{i=1}^r \sigma_i^4} \gg \sqrt{\sum_{i=r+1}^n \sigma_i^4}, \quad (3.3)$$

ou bien

$$\xi_r = \frac{\sqrt{\sum_{i=r+1}^n \sigma_i^4}}{\sqrt{\sum_{i=1}^r \sigma_i^4}} \ll 1, \quad (3.4)$$

avec  $r \neq 0$  et  $r \neq n$ , et  $\sigma_i$  est la valeur singulière correspondant à l'état  $i$ . Pour plus d'efficacité, on définit une autre mesure d'erreur  $e$  telle-que

$$\zeta_i = |\xi_i - \xi_{i+1}| \leq e, \quad (3.5)$$

pour  $i = \overline{1, n-2}$  où  $e$  est une erreur fixée a priori.

**Etape 2-** Partition de la matrice  $\Sigma$  (dite matrice des valeurs singulières) comme suit

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}, \quad (3.6)$$

où  $\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_r)$  [partie à conserver],

$\Sigma_2 = \text{diag}(\sigma_{r+1}, \dots, \sigma_n)$  [partie à éliminer],

$r$  : ordre du modèle simplifié.

Les éléments de  $\Sigma$  doivent être ordonnés par ordre décroissant, c'est-à-dire;  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$ .

**Etape 3-** Introduction de la matrice  $\Sigma$  partitionnée dans les équations de *Lyapunov*, nous aurons [20]

$$\begin{bmatrix} A_{e11} & A_{e12} \\ A_{e21} & A_{e22} \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{e11}^T & A_{e12}^T \\ A_{e21}^T & A_{e22}^T \end{bmatrix} - \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} + \begin{bmatrix} B_{e1} \\ B_{e2} \end{bmatrix} \begin{bmatrix} B_{e1}^T & B_{e2}^T \end{bmatrix} = 0, \quad (3.7)$$

$$\begin{bmatrix} A_{e11}^T & A_{e12}^T \\ A_{e21}^T & A_{e22}^T \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{e11} & A_{e12} \\ A_{e21} & A_{e22} \end{bmatrix} - \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} + \begin{bmatrix} C_{e1}^T \\ C_{e2}^T \end{bmatrix} \begin{bmatrix} C_{e1} & C_{e2} \end{bmatrix} = 0. \quad (3.8)$$

**Etape 4-** Poser

$$(A_r, B_r, C_r, \Sigma_r) = (A_{e11}, B_{e1}, C_{e1}, \Sigma_1). \quad (3.9)$$

**Sorties-** La réalisation  $(A_r, B_r, C_r, r)$  représente le modèle simplifié d'ordre  $r$ .

**Remarques**

La réalisation simplifiée  $(A_r, B_r, C_r, r)$  correspond bien à la partie la plus commandable et la plus observable de la réalisation complète  $(A, B, C, n)$ .

Il est important de noter que la réduction d'ordre par l'approche *MRE* préserve les propriétés clés du modèle initial d'ordre complet, comme la minimalité et la stabilité.

### 3.2.4 Conclusion

Bien que l'approche *MRE* permet de donner un modèle simplifié possédant d'intéressantes propriétés, telles que la garantie de la stabilité des modèles d'ordres réduit, et de la minimalité, la détermination d'un tel modèle reste fortement limitée à cause des difficultés numériques, puisque les calculs demandés pour construire la matrice de transformation d'équilibre  $T$  sont compliqués et sensibles au bruit d'arrondi [30].

En particulier, l'équilibre peut devenir mal conditionné, lorsque les modes du système

sont presque non commandables ou/et mal observables [21].

### 3.3 Méthode de *Schur* (réalisations non minimales)

La méthode de *Schur* permet le calcul du modèle de **Moore** directement sans passer par l'étape d'équilibre du modèle d'ordre complet, ceci via des projections définies en termes de bases orthogonales pour les sous-espaces propres droite et gauche associés aux plus grandes valeurs propres du produit  $(W_c W_o)$ , des grammiens de commandabilité et d'observabilité.

Nous précisons que cette approche opère dans le domaine continu. Pour pouvoir simplifier un filtre numérique non minimal (où la *MRE* numérique est inutilisable), il faudra trouver l'équivalent continu du filtre numérique (en utilisant la transformation bilinéaire inverse), ceci pour pouvoir réduire son ordre par la méthode de *Schur* [31], et le filtre numérique à modèle simplifié sera obtenu en transformant (par transformation bilinéaire directe) le modèle simplifié continu. Dans ce qui va suivre nous présentons deux algorithmes de calcul des paramètres du modèle simplifié, dont la stabilité numérique est assurée grâce à l'utilisation des transformations orthogonales [32]. Les deux algorithmes que nous allons présenter permettent d'aboutir au même modèle simplifié.

#### 3.3.1 Procédure de la simplification par *Schur*

**Etape 1-** Calcul des grammiens de commandabilité  $W_c$  et d'observabilité  $W_o$ .

**Etape 2-** Calcul de la matrice réelle orthonormale  $V$ , telle que la matrice  $(V W_c W_o V^T)$  soit triangulaire supérieure ; c'est à dire, mettre sous la forme de *Schur*.

**Etape 3-** En utilisant les rotations orthogonales [32], calculer les transformations réelles orthogonales  $V_A$  et  $V_D$  qui ordonnent la forme de *Schur* respectivement en ordre ascendant et descendant, telles-que

$$V_A^T W_c W_o V_A = \begin{bmatrix} \lambda_{A_n} & * & * & * & * \\ 0 & \lambda_{A_{n-1}} & * & * & * \\ \vdots & 0 & * & * & * \\ \vdots & \cdots & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \lambda_{A_1} \end{bmatrix}, \quad (3.10)$$

$$V_D^T W_c W_o V_D = \begin{bmatrix} \lambda_{D_1} & * & * & * & * \\ 0 & \lambda_{D_2} & * & * & * \\ \vdots & 0 & * & * & * \\ \vdots & \cdots & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \lambda_{D_n} \end{bmatrix}, \quad (3.11)$$

avec

$$\lambda_{A_i} = \lambda_{D_i} = \sigma_i^2, (i = \overline{1, r}), \quad (3.12)$$

$$\lambda_{A_i} = \lambda_{D_i} = \sigma_i^2, (i = \overline{r+1, n}). \quad (3.13)$$

**Etape 4-** Partition de  $V_A$  et  $V_D$  telles que

$$V_A = \left[ \begin{array}{c|c} \overbrace{\quad}^{\text{n-r}} & \overbrace{\quad}^{\text{r}} \\ \hline \widehat{V}_{d,p} & \widehat{V}_{g,g} \end{array} \right], \quad (3.14)$$

$$V_D = \left[ \begin{array}{c|c} \overbrace{\quad}^{\text{r}} & \overbrace{\quad}^{\text{n-r}} \\ \hline \widehat{V}_{d,g} & \widehat{V}_{g,p} \end{array} \right]. \quad (3.15)$$

**Etape 5-** Former la projection [32]

$$E_g = V_{g,g}^T V_{d,g}, \quad (3.16)$$

et calculer sa décomposition en valeurs singulières *SVD* (Annexe B)

$$E_g = U_{E,g} \Sigma_{E,g} V_{E,g}. \quad (3.17)$$

**Etape 6-** Former les matrices

$$S_{g,g} = V_{g,g} U_{E,g} \Sigma_{E,g}^{-\frac{1}{2}}, \in R^{n \times r}, \quad (3.18)$$

$$S_{d,g} = V_{d,g} V_{E,g} \Sigma_{E,g}^{-\frac{1}{2}}, \in R^{n \times r}. \quad (3.19)$$

**Etape 7-** Construction de la réalisation d'état du modèle simplifié

$$\begin{bmatrix} A_r & B_r \\ C_r & D_r \end{bmatrix} = \begin{bmatrix} S_{g,g}^T A S_{d,g} & S_{g,g}^T B \\ C S_{d,g} & D \end{bmatrix}. \quad (3.20)$$

**Remarques**

- La réalisation d'état  $(A_r, B_r, C_r, D_r)$  présente des grammians de commandabilité et d'observabilité donnés respectivement par

$$W_{c,r} = S_{g,g}^T W_c S_{g,g} \in R^{r \times r}, \quad (3.21)$$

$$W_{o,r} = S_{d,g}^T W_o S_{d,g} \in R^{r \times r}. \quad (3.22)$$

- Les colonnes de  $V_{d,g}$  et  $V_{d,p}$  forment respectivement les bases orthonormales pour les sous-espaces propres droite de  $(W_c W_o)$  associés respectivement aux grandes valeurs

propres  $\left[ \sigma_1^2, \dots, \sigma_r^2 \right]$  et aux plus faibles valeurs propres  $\left[ \sigma_{r+1}^2, \dots, \sigma_n^2 \right]$ .

Les colonnes de  $V_{g,g}$  et  $V_{g,p}$  donnent une décomposition analogue des sous-espaces propres gauche.

– Notons que le modèle d'ordre réduit vérifie la borne de l'erreur de Hankel

$$\|H(z) - H_r(z)\|_{H_\infty} = \sup_{\omega \in R} \|H(j\omega) - H_r(j\omega)\|_2 \leq 2 \sum_{i=r+1}^n \sigma_i \quad (3.23)$$

où  $H(z)$  et  $H_r(z)$  sont les fonctions de transfert du modèle *AR* initial d'ordre  $n$  et celle du modèle d'ordre réduit  $r$ , et  $\|\cdot\|_2$  est la norme spectrale matricielle.

### 3.3.2 Algorithme de *Schur* modifié

Dans ce paragraphe, nous allons examiner une autre procédure que celle proposée par Safonov et Chiang [24]; au lieu de calculer les équations de *Lyapunov* pour  $W_o$  et  $W_c$ , nous calculerons uniquement une seule matrice dite "matrice intergrammien" notée  $W_{co}$ , définie par

$$W_{co} = \int_0^{T_f} \exp(At)BC \exp(A^T t) dt. \quad (3.24)$$

$W_{co}$  est solution de l'équation de *Lyapunov*

$$W_{co}A^T + AW_{co} + BC = 0, \quad (3.25)$$

et l'on montre que [30]

$$W_{co}^2 = W_c W_o. \quad (3.26)$$



### Remarque

$W_{co}$  enveloppe une information sur la commandabilité et l'observabilité du système et contrairement à  $W_o$  et  $W_c$ , il est invariant sous une transformation de similarité.

Une autre propriété qui fait la différence avec l'algorithme précédent est que la forme de Schur réelle est calculée une seule fois (au lieu de deux) et l'algorithme *SVD* (Singular Values Decomposition) (voir Annexe B) demandée dans la procédure proposée par Safonov et Chiang [30] est remplacée par la résolution de l'équation de *Sylvester* [34] de deux matrices triangulaires supérieures.

Ainsi, la matrice de transfert du modèle réduit  $G_r(s)$  est calculée avec un volume de calcul plus réduit.

### Procédure de *Schur* modifiée

**Entrées-** Ayant  $(A, B, C, n, r)$ .

**Etape 1-** Calcul de la matrice  $W_{co}$  par résolution de l'équation de *Lyapunov*.

**Etape 2-** Mettre  $W_{co}$  sous la forme de *Schur*

$$Q^T W_{co} Q = \tilde{W} = \begin{bmatrix} \tilde{W}_{11} & \tilde{W}_{12} \\ 0 & \tilde{W}_{22} \end{bmatrix}, \quad (3.27)$$

où les valeurs propres [33] de  $W_{co}$  apparaissent en décroissance tout au long de la diagonale de  $\tilde{W}$ .

Les matrices  $\tilde{W}_{11} \in R^{r \times r}$  et  $\tilde{W}_{22} \in R^{(n-r) \times (n-r)}$  sont triangulaires supérieures.

**Etape 3-** Calcul de la matrice  $X$  par résolution de l'équation de *Sylvester* [34], donnée par

$$\tilde{W}_{11} X - X \tilde{W}_{22} + \tilde{W}_{12} = 0. \quad (3.28)$$

**Etape 4-** Diagonalisation de  $\tilde{W}$  par blocs, par la matrice

$$\begin{bmatrix} I_r & X \\ 0 & I_{n-r} \end{bmatrix}, \quad (3.29)$$

telle-que la matrice de transformation totale  $V$  diagonalise  $W_{co}$  en blocs où  $W_{co}^g \in R^{r \times r}$

tel que  $\lambda(W_{co}^g) = \lambda_i$ , avec  $i = \overline{1, r}$  et

$W_{co}^p \in R^{(n-r) \times (n-r)}$  tel que  $\lambda(W_{co}^p) = \lambda_i$ , avec  $i = \overline{r+1, n}$ .

**Etape 5-** Partition de  $V$  et  $V^{-1}$ , telles-que

$$V = \begin{bmatrix} V_g & V_p \end{bmatrix}, \quad (3.30)$$

$$V^{-1} = W = \begin{bmatrix} W_g \\ W_p \end{bmatrix}, \quad (3.31)$$

avec

$$WV = \begin{bmatrix} I_r & 0 \\ 0 & I_{n-r} \end{bmatrix}. \quad (3.32)$$

**Etape 6-** Appliquer la transformation

$$x(t) = V\tilde{x}(t), \quad (3.33)$$

au système (3.1) et construire

$$\tilde{A} = V^{-1}AV = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}, \quad (3.34)$$

$$\tilde{B} = V^{-1}B = \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix}, \quad (3.35)$$

$$\tilde{C} = CV = \begin{bmatrix} \tilde{C}_1 & \tilde{C}_2 \end{bmatrix}. \quad (3.36)$$

**Sorties-** La réalisation simplifiée d'ordre  $r$  est donnée par  $(A_r, B_r, C_r, r) = (\tilde{A}_{11}, \tilde{B}_1, \tilde{C}_1, r)$ .

**Remarque**

Les colonnes de  $V_g$  (respectivement  $W_g$ ) engendrent les sous-espaces propres droite (respectivement gauche) associés aux plus grandes valeurs propres de  $W_{co}$ .

En conclusion, l'algorithme de *Schur* donne toujours un modèle d'ordre réduit stable, même si le système tend vers la non observabilité et/ou la non commandabilité.

Reposant sur le même concept de base qu'est le calcul d'espaces propres droite et gauche, mais cette fois-ci associés aux plus grandes valeurs propres de l'intergrammien  $W_{co}$ , la technique de Schur modifiée permet de calculer la même fonction de transfert que celle de *Schur*, ceci avec un volume de calcul plus réduit, sans avoir à déterminer la transformation d'équilibre.

Les deux algorithmes de *Schur*, stables numériquement, sont surtout orientés vers le problème d'approximation de systèmes non minimaux.

### 3.4 Conclusion

Dans le cas où le modèle est complètement commandable et complètement observable (modèle minimal), la méthode des réalisations équilibrées peut être utilisée. Par contre, si le modèle est non minimal, c'est-à-dire; non observable et/ou non commandable, la méthode des réalisations équilibrées n'est plus utilisable, et l'approche de *Schur* vient pour y remédier. Nous en déduisons que la méthode de *Schur* n'est rien d'autre que la

généralisation de la méthode des réalisations équilibrées.



## Chapitre 4

# SIMULATIONS

### 4.1 Introduction

Dans ce chapitre, nous traitons la synthèse de système de production de parole avec des modèles simplifiés.

Dans chaque exemple, nous partons d'un signal de parole sous format \*.wav, utiliser pour obtenir un modèle *AR* (filtre tout pôles) d'ordre complet  $n$ , synthétisé par la méthode classique de *Burg* citée dans le chapitre 2. En appliquant à ce modèle la réduction d'ordre par projections internes développée dans le chapitre 3, nous obtenons ainsi un modèle d'ordre réduit  $r$ . Pour chaque exemple, nous étudierons le comportement temporel du signal original ainsi que celui issu du modèle réduit (formes d'onde), et le comportement fréquentiel des modèles de production de parole, ainsi que quelques caractéristiques clés de chaque modèle de représentation, voir rapport signal sur bruit *RSB*, lieu des pôles du

modèle d'ordre réduit de nature *ARMA* (critère de stabilité), lieu des zéros (minimalité de la phase), lieu des *LSF* (Line Spectrale Frequencies), ceci pour le modèle initial et son approximant d'ordre réduit. Notons que le modèle initial d'ordre complet *AR* est toujours stable, et que ses *LSF* se trouvent toujours sur le cercle unité. Pour chaque exemple, le signal d'excitation est obtenu par excitation du filtre *AR* inverse *RIF* par le signal de parole, ce signal d'excitation est appliqué au modèle d'ordre réduit pour avoir en sortie la forme d'onde du signal de parole issu du modèle simplifié.

Nous opérons pour chaque exemple à partir d'un signal de parole enregistré ou synthétisé sous format \*.wav.

Les réalisations d'état de chaque modèle *AR* sont obtenues en utilisant une transformation qui permet le passage de la fonction de transfert (numérateur et dénominateur) vers l'espace d'état (matrices  $A, B, C, D$ ). Il nous a semblé aussi important de comparer le modèle d'ordre réduit ( $r$ ) avec le modèle *ARMA* de même ordre, synthétisé directement en utilisant la fonction *ARMAX* du Matlab [34].

## 4.2 Exemple 1

Soit un signal de parole numérisé avec une fréquence d'échantillonnage  $F_s = 11025 Hz$ , correspondant à la Syllabe "ba" (Figure 4.1).

On construit un modèle *AR*, d'ordre 15, selon la méthode de *burg*, avec les coefficients suivants

$$a_i = [1, -1.6809, 1.2089, -8.7482e - 1, 9.7920e - 1, 8.9974e - 1, 8.5842e - 1, -9.9185e -$$

$1, 8.9750e - 1, 3.4562e - 1, 1.9411e - 1, -5.4363e - 1, 3.9457e - 1, -1.6627e - 1, 2.0950e - 1, -1.1255e - 1]$ .

Notons ici que l'ordre du modèle *AR* initial peut être fixé à des valeurs souhaitées, mais pour le cas du signal de parole, il est préférable de la prendre entre 10 et 15.

A partir de ce modèle d'ordre 15 et en utilisant la méthode de *Schur* (voir section 3.3), on synthétise un modèle d'ordre réduit d'ordre  $r = 5$ ; en général, l'ordre du modèle simplifié  $r$  est déterminé par le critère de la norme de *Hankel*, se basant essentiellement sur la répartition des valeurs singulières de *Hankel* (Fig. 4.2).

La fonction de transfert du modèle simplifié d'ordre  $r = 5$  est donnée par

$$H_r(z) = \frac{0.92 - 1.44z^{-1} + 0.92z^{-2} + 0.41z^{-3} - 0.814z^{-4} + 0.45z^{-5}}{1 - 3.42z^{-1} + 5.51z^{-2} - 5.06z^{-3} + 2.63z^{-4} - 0.6z^{-5}}$$

On remarque que le modèle d'ordre réduit 5 possède une structure *ARMA*, défini par 10 coefficients : 5 du numérateur et 5 du dénominateur, tandis que celui initial d'ordre complet de structure *AR* est défini par 15 coefficients. Les valeurs singulières de *Hankel* du modèle d'ordre complet sont données par

$\Sigma = [6.1201, 3.7887, 2.3960, 2.2460, 1.2042, 6.0826e - 1, 5.4957e - 1, 2.0269e - 1, 1.2123e - 1, 1.0325e - 1, 3.1003e - 2, 2.8773e - 2, 2.8117e - 2, 1.7032e - 3, 1.0003e - 3]$ .

Sur la même figure (Fig. 4.1), nous traçons le signal de parole initial  $y(k)$ , celui issu du modèle réduit d'ordre  $r = 5$   $y_r(k)$  ainsi que leur écart  $er(k)$ , ceci pour la syllabe 'ba'. Les Figures 4.3, 4.4, et 4.5 montrent respectivement le lieu des pôles et le lieu géométrique des *LSF*, et le lieu des zéros du modèle d'ordre réduit. Notons que nous voulons montrer par la fig. 4.4 que les *LSF* du modèle d'ordre réduit se trouvent bien sur le cercle unité par paires.



D'autre part, les signaux  $y(k)$  et  $y_r(k)$  se juxtaposent parfaitement, donc impossible de les tracer en même temps, c'est pour cela que leurs tracés sont séparés. Les figures 4.6 et 4.7 donnent les spectres d'amplitudes du modèle *AR* d'ordre 15, du modèle réduit d'ordre 5, du modèle *ARMA* directe issu de la fonction *ARMAX* du Matlab.

Le modèle d'ordre réduit 5 satisfait la borne de l'erreur de Hankel

$$\|H(z) - H_r(z)\|_{H_\infty} = \sup_{\omega \in R} \|H(j\omega) - H_r(j\omega)\|_2 = 1.1 \leq 2 \sum_{i=6}^{15} \sigma_i = 3.4$$

Pour une meilleure comparaison, nous permettant de ressortir quelques importants points sur la qualité de sons produits par nos modèles, on synthétise aussi des modèles d'ordre réduit  $r = 4$ , et  $r = 5$ , ainsi que des modèles *ARMA* de même ordre, pour quelques syllabes directement en utilisant la fonction *ARMAX* du *MATLAB* [36]. Cette comparaison est basée sur la mesure du rapport signal sur bruit *SNR*.

Les *SNR* associés au modèle d'ordre réduit et celui *ARMA* synthétisé directement sont définis respectivement par

$$SNR_r = \frac{\sigma_{y_r}^2}{\sigma_{e_1}^2}, \quad (4.1)$$

$$SNR_{ARMA} = \frac{\sigma_{y_{ARMA}}^2}{\sigma_{e_2}^2}, \quad (4.2)$$

où  $\sigma_x^2$  est la variance du signal  $x$ ,  $e_1$  et  $e_2$  sont les formes d'onde d'erreurs définies par

$$e_1(k) = y(k) - y_r(k), \quad (4.3)$$

$$e_2(k) = y(k) - y_{ARMA}(k), \quad (4.4)$$

où  $y(k)$ ,  $y_r(k)$  et  $y_{ARMA}(k)$  correspondent respectivement aux sorties du modèle d'ordre complet, du modèle d'ordre réduit, et celle du modèle *ARMA* directe.

Le tableau Tab. 4.1 donne la mesure des  $SNR_r$  et  $SNR_{ARMA}$  en unité  $dB$ , on remarque que le  $SNR_r$  du modèle d'ordre réduit est beaucoup plus important que celui  $SNR_{ARMA}$  du modèle  $ARMA$  de même ordre (voir Tab. 4.1).

La figure 4.7 montre que le  $SNR_r$  est toujours supérieur au  $SNR_{ARMA}$ , ceci pour n'importe quel ordre de réduction  $r$ , le tracé est donné pour la syllabe  $fa$ , et reste valable pour tous les autres segments de parole.

TAB. 4.1 – Comparaison du  $SNR_r$  et  $SNR_{ARMA}$

Syllabe	$SNR_{ARMA}(dB)$ , $SNR_r(dB)$ ,		$SNR_{ARMA}(dB)$ , $SNR_r(dB)$ ,	
	$m = 5$	$r = 5$	$m = 4$	$r = 4$
<i>a</i>	9.3	12.1	1.1	2.5
<i>ba</i>	10.5	19.5	7.4	8.4
<i>bo</i>	20.1	24.9	18.6	18.9
<i>sa</i>	12.9	19.8	7.3	14.7
<i>se</i>	21.9	24.0	11.3	12.4
<i>si</i>	21.1	25.0	1.5	2.4
<i>da</i>	15.7	18.5	8.3	8.7
<i>fa</i>	20.3	25.4	5.9	9.2
<i>ma</i>	14.1	15.9	10.2	12.0
<i>o</i>	8.2	25.9	6.4	7.1
<i>pa</i>	19.7	19.8	12.1	13.7
<i>po</i>	25.7	43.3	9.7	14.0
<i>i</i>	26.6	31.2	16.2	18.1

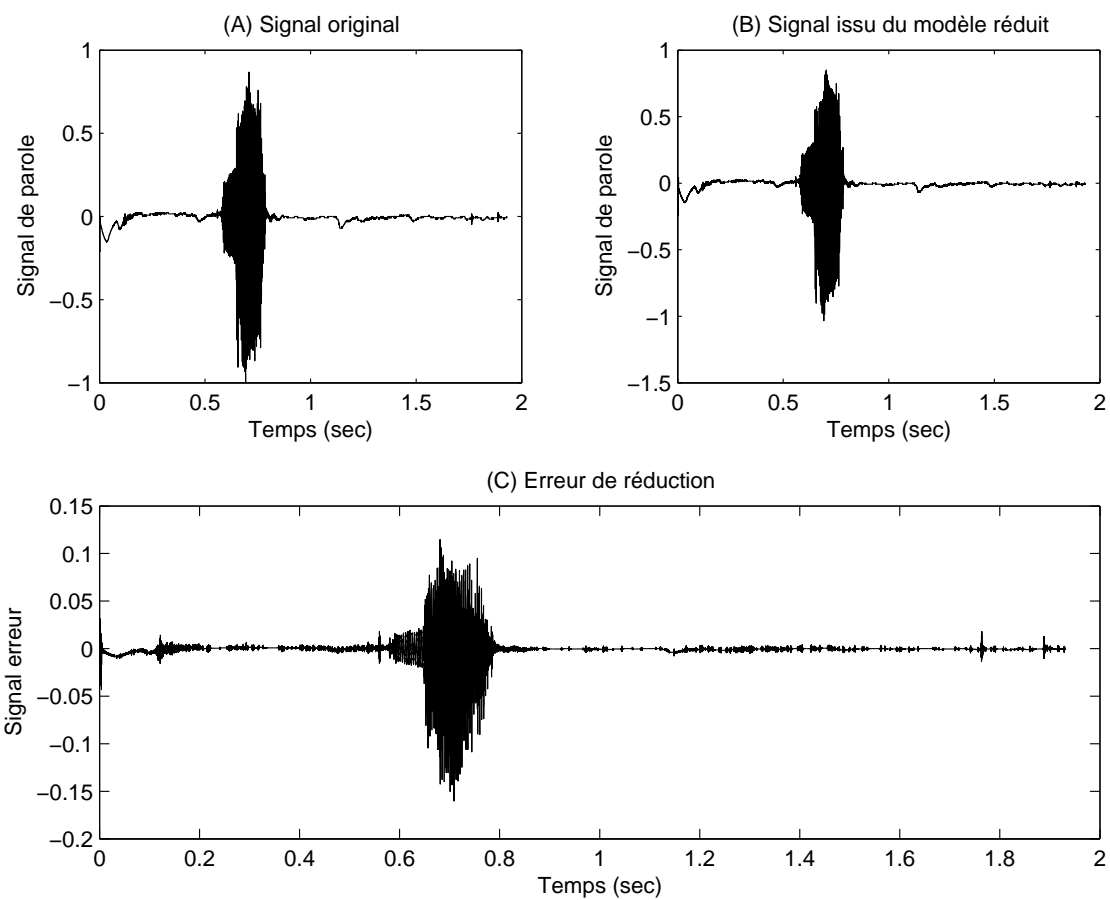


FIG. 4.1 – Formes d'onde de la phonème originale (A), d'ordre réduit (B) et de l'erreur (C)

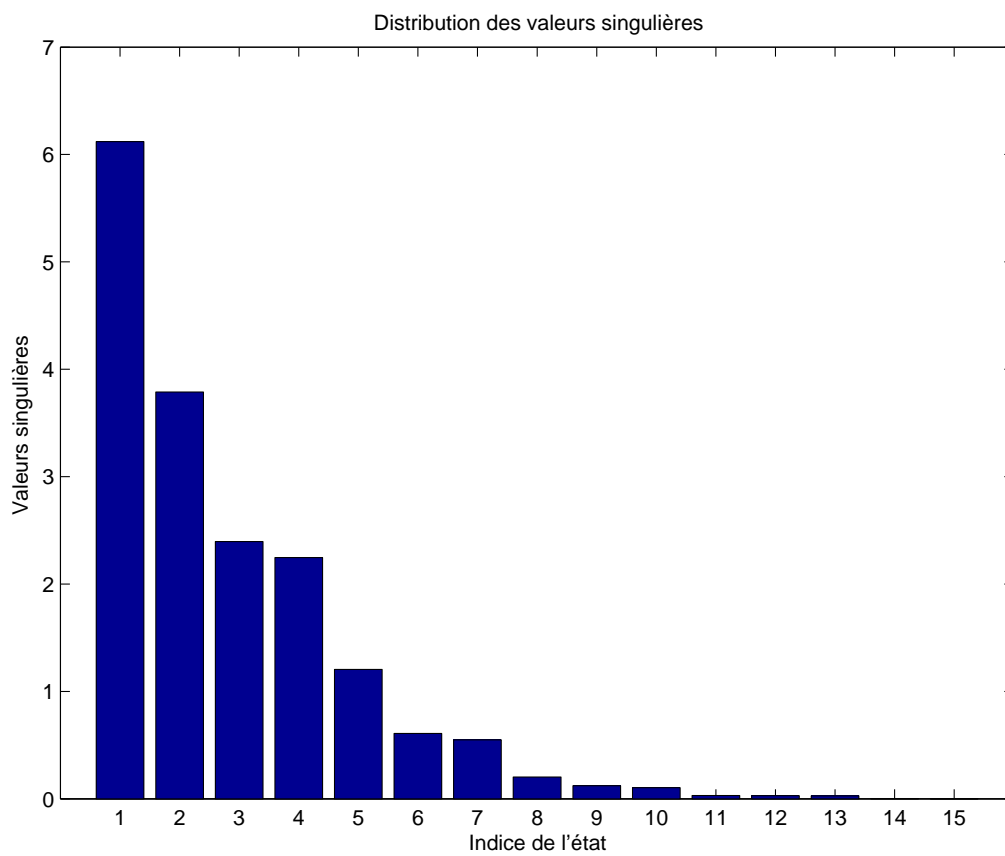


FIG. 4.2 – Répartition des valeurs singulières de Hankel

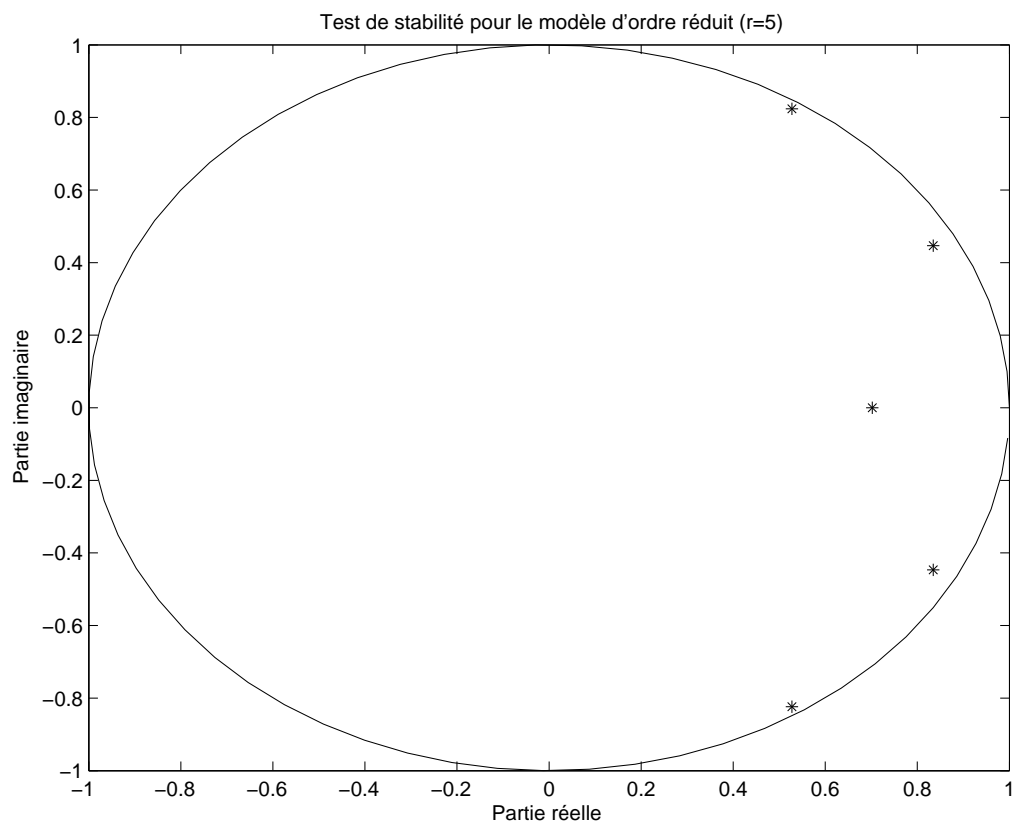


FIG. 4.3 – Lieu de stabilité du modèle d'ordre réduit

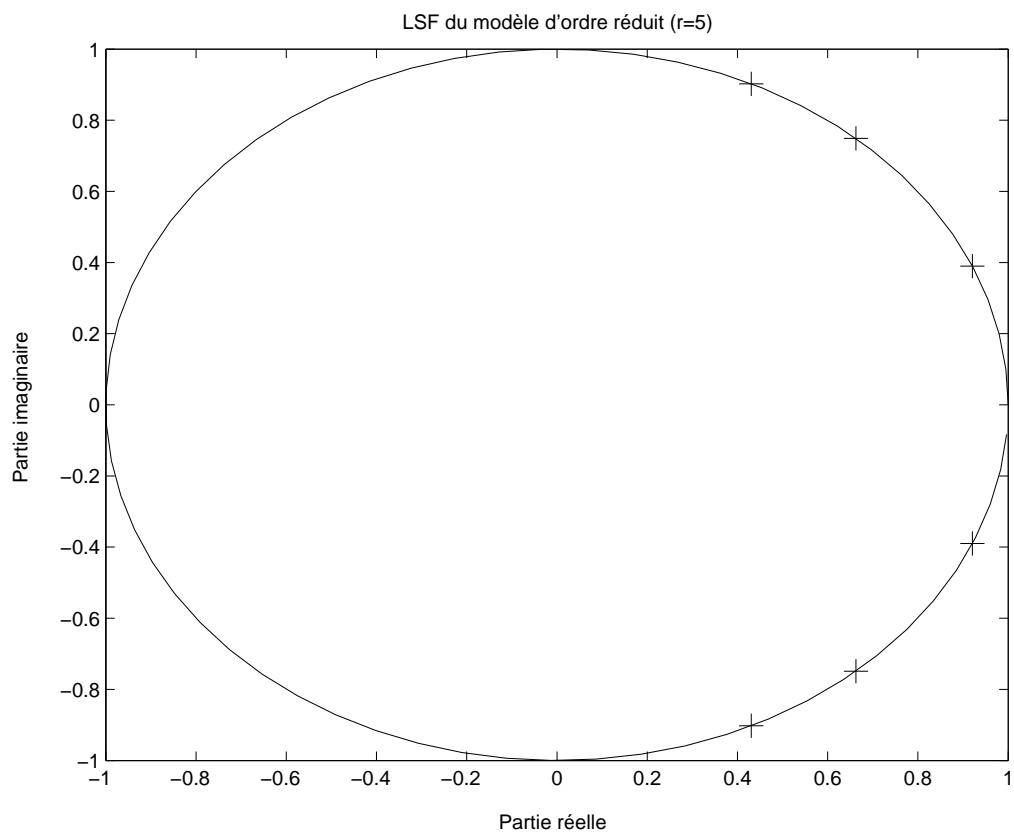


FIG. 4.4 – LSF du modèle d'ordre réduit

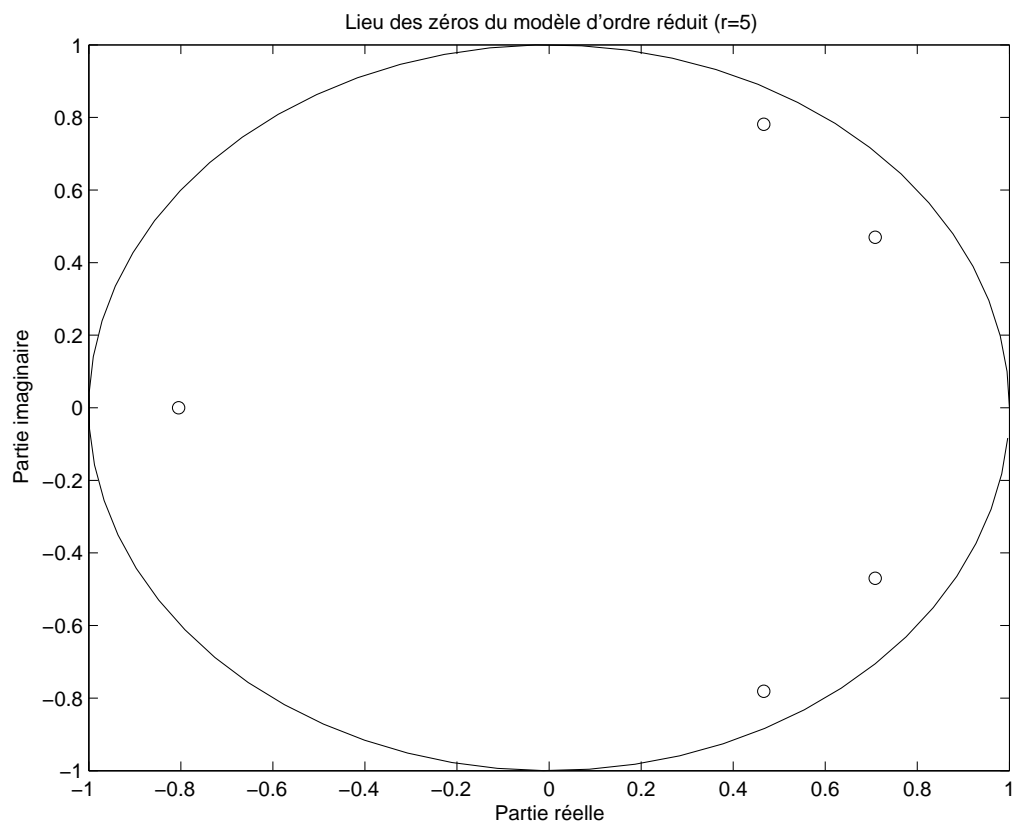


FIG. 4.5 – Lieu des zéros du modèle d'ordre réduit

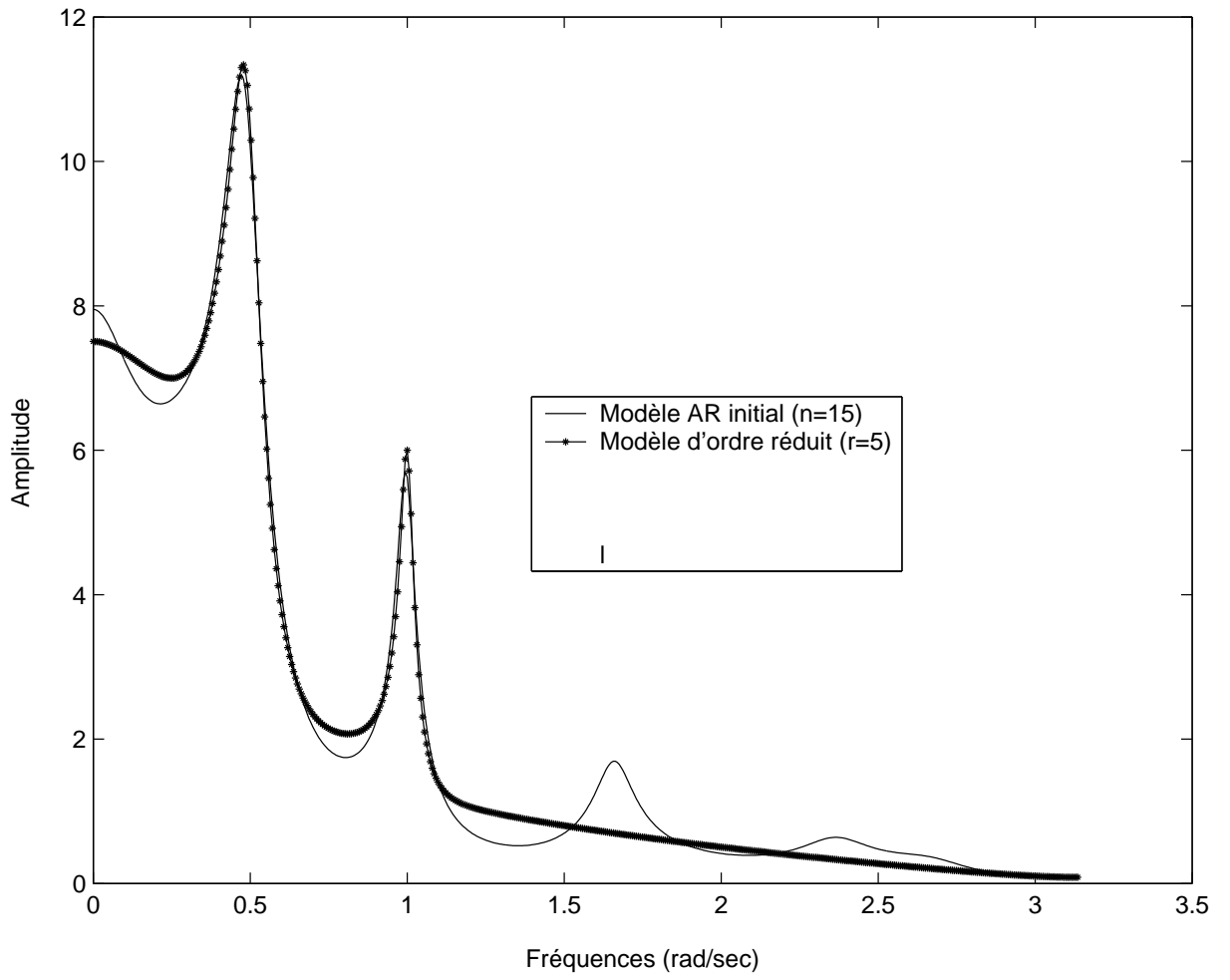


FIG. 4.6 – Spectres d'amplitude des modèles original et d'ordre réduit



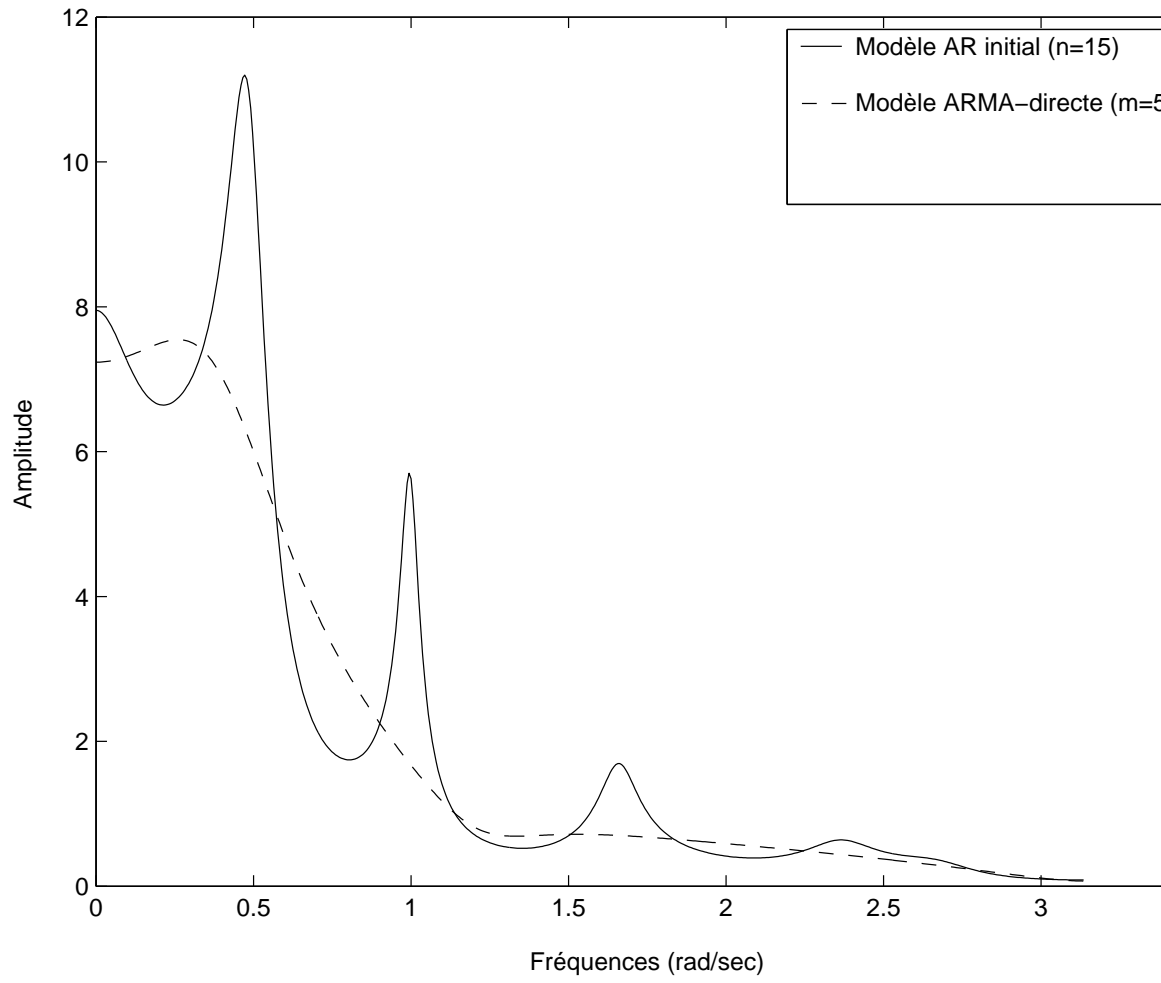


FIG. 4.7 – Spectres d'amplitude des modèles original et ARMA-directe

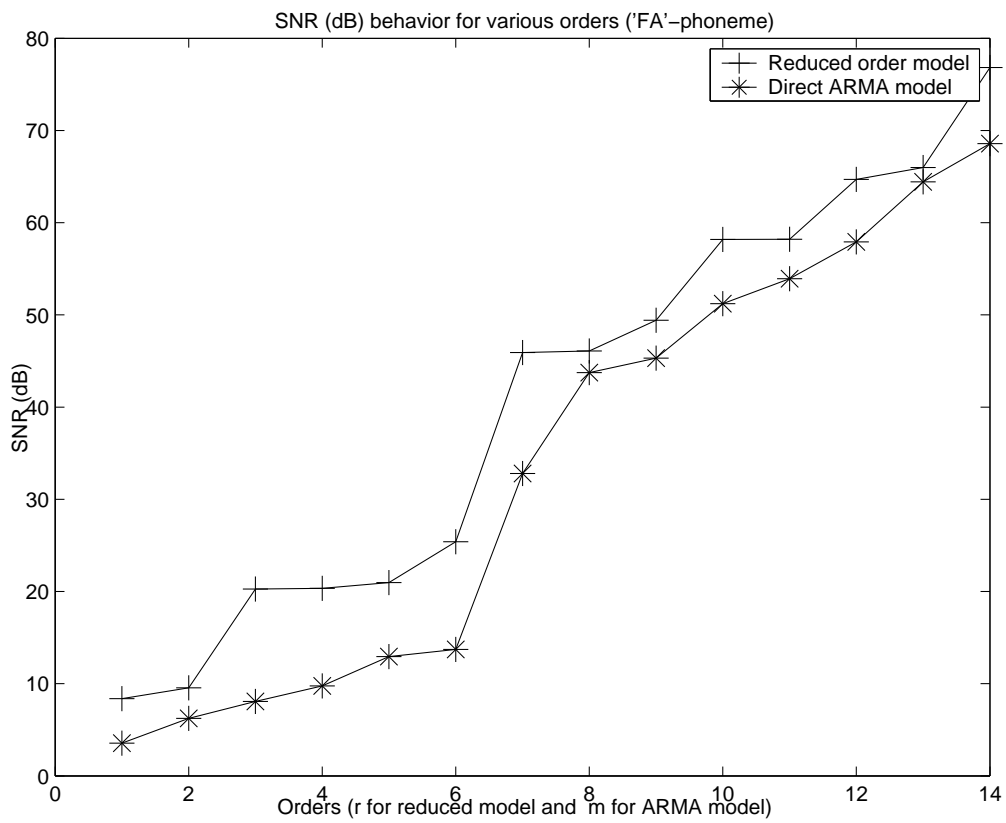


FIG. 4.8 – Comportement du SNR du modèle d'ordre réduit et celui direct *ARMA*

### 4.3 Exemple 2

Soit un signal de parole numérisé avec une fréquence d'échantillonnage  $F_s = 11025Hz$ , correspondant à la phrase en langue anglaise "*Here is an example of low order reduction of speech production system*". On construit un modèle *AR*, d'ordre 20, selon la méthode de *Burg*, dont les coefficients sont

$$a_i = [1, -6.6159e-1, -9.265e-2, -6.131e-1, 7.370e-1, -1.034e-1, -1.573e-1, 1.373e-1, -2.466e-2, 5.285e-2, -9.244e-2, 1.356e-1, 1.007e-1, -1.999e-1, 5.060e-2, 3.417e-2, 1.309e-1, -8.335e-2, -2.995e-2, -1.354e-2, 5.971e-2].$$

Les valeurs singulières de *Hankel* du modèle d'ordre complet (Fig. 5.14) sont données par  $\Sigma = [4.294, 3.324, 1.633, 1.494, 3.503e-1, 2.984e-1, 1.891e-1, 1.448e-1, 1.218e-1, 6.385e-2, 6.181e-2, 5.636e-2, 5.087e-2, 4.226e-2, 3.900e-2, 1.933e-2, 6.925e-3, 5.689e-3, 5.395e-4, 4.516e-4]$ .

A partir de ce modèle d'ordre 20 et en utilisant l'approche de *Schur* (voir section 3.3), on synthétise un modèle d'ordre réduit  $r = 5$ ; en général, l'ordre du modèle simplifié  $r$  est déterminé en fonction de la distribution des valeurs singulières de *Hankel*, et le modèle d'ordre réduit 5 satisfait la borne de l'erreur de *Hankel*

$$\|H(z) - H_r(z)\|_{H_\infty} = \sup_{\omega \in R} \|H(j\omega) - H_r(j\omega)\|_2 = 0.87 \leq 2 \sum_{i=6}^{20} \sigma_i = 2.20$$

La fonction de transfert du modèle simplifié d'ordre  $r = 5$  est donnée par

$$H_r(z) = \frac{0.96 - 0.62z^{-1} - 0.27z^{-2} + 0.23z^{-3} + 0.04z^{-4} - 0.243z^{-5}}{1 - 1.20z^{-1} - 0.007z^{-2} - 0.40z^{-3} + 1.05z^{-4} - 0.38z^{-5}}.$$

On remarque que le modèle d'ordre réduit 5 possède une structure *ARMA*, donc défini par 10 coefficients, 5 du numérateur et 5 du dénominateur, tandis que celui initial

d'ordre complet de structure  $AR$  est défini pas 20 coefficients. Le nombre de coefficients est très important dans la synthèse puis le codage d'un signal de parole, donc on cherche toujours à synthétiser un bon modèle de représentation du signal de parole avec un minimum de paramètres, correspondant ainsi à un nombre minimum de bits lors du codage. Un autre modèle  $ARMA$  synthétisé directement en utilisant la fonction *Matlab ARMAX*, de même ordre  $m = 5$  a pour fonction de transfert

$$H_{ARMA}(z) = \frac{0.99 + 0.07z^{-1} - 0.19z^{-2} - 0.20z^{-3} - 0.18z^{-4}}{1 - 0.58z^{-1} - 0.32z^{-2} - 0.70z^{-3} + 0.65z^{-4} + 0.11z^{-5}}$$

. La Fig. 4.9 donne, sur un même graphe, le signal parole original, celui issu du modèle d'ordre réduit ainsi que l'écart de réduction. La Fig. 4.10 montre les lieux des pôles des modèles initial et celui d'ordre réduit. La Fig. 4.11 montre bien que les  $LSF$  sont sur le cercle unité. La Fig. 4.12 donne les réponses fréquentielles du modèle initial d'ordre 20, et celui d'ordre réduit ( $r=5$ ), et la Fig. 4.13 donne le spectre du modèle initial et celui du modèle ARMA-directe synthétisé par la fonction *ARMAX*. La Fig. 4.15 montre bien que le modèle d'ordre réduit est à phase minimale.

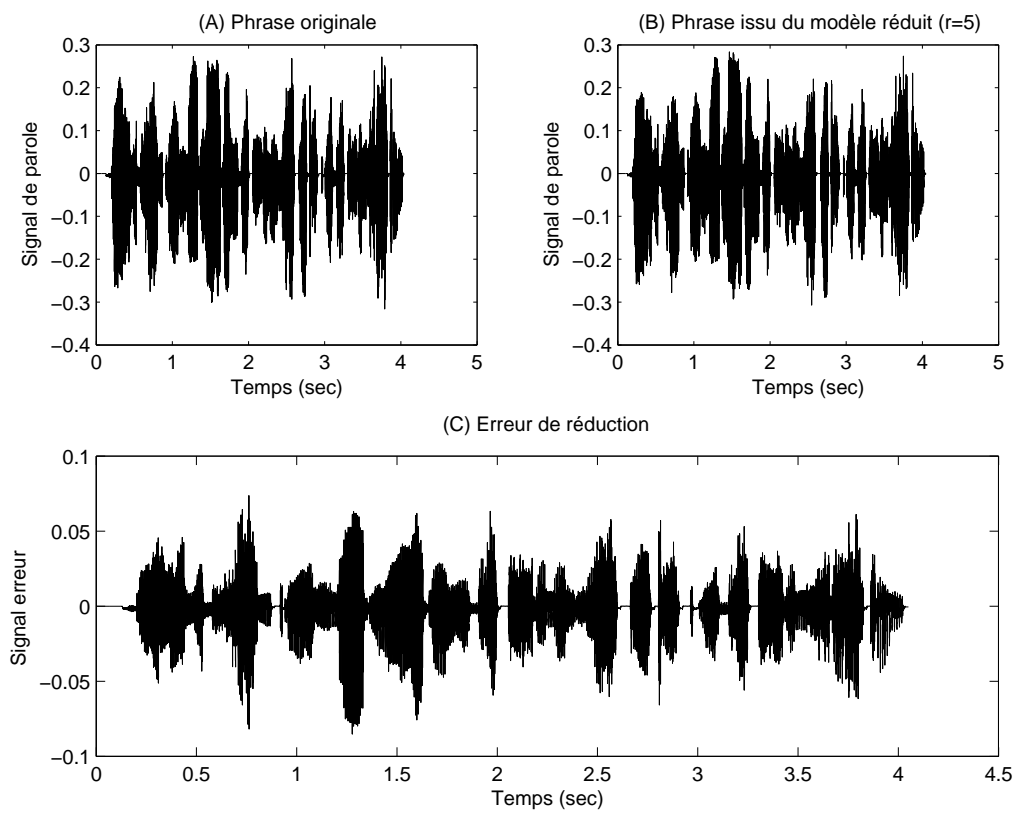


FIG. 4.9 – Signal original (A), d'ordre réduit (B) et erreur de réduction (C)

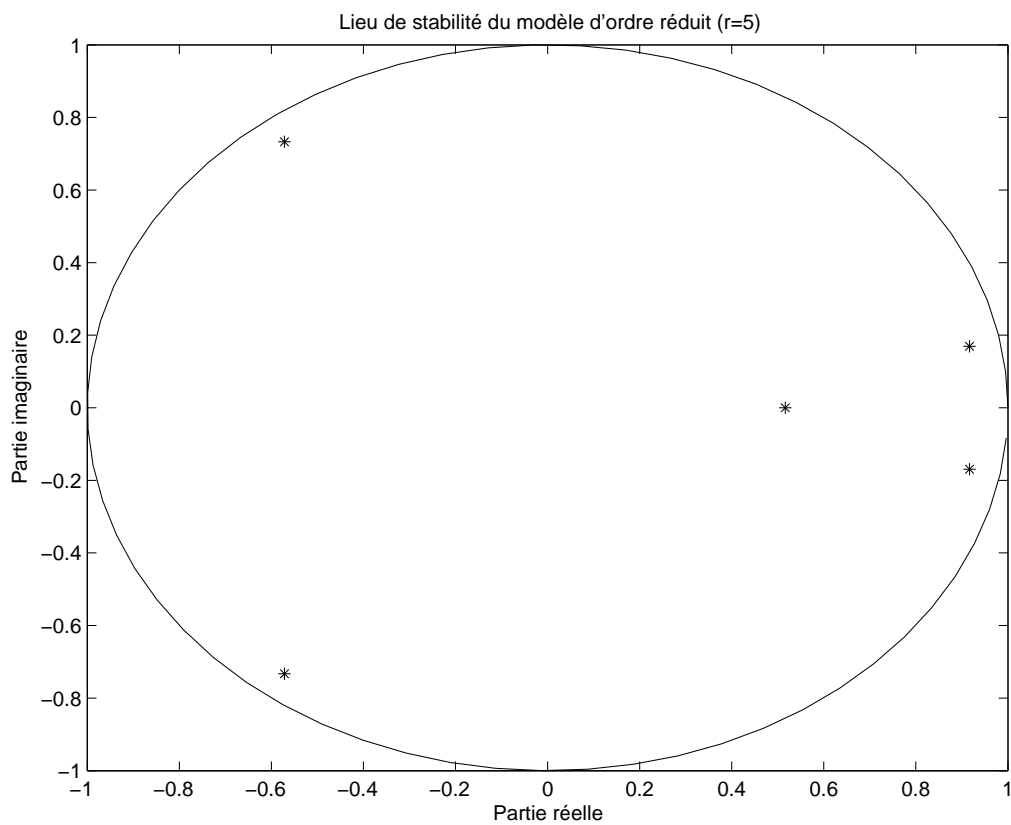


FIG. 4.10 – Lieu de stabilité du modèle d'ordre réduit

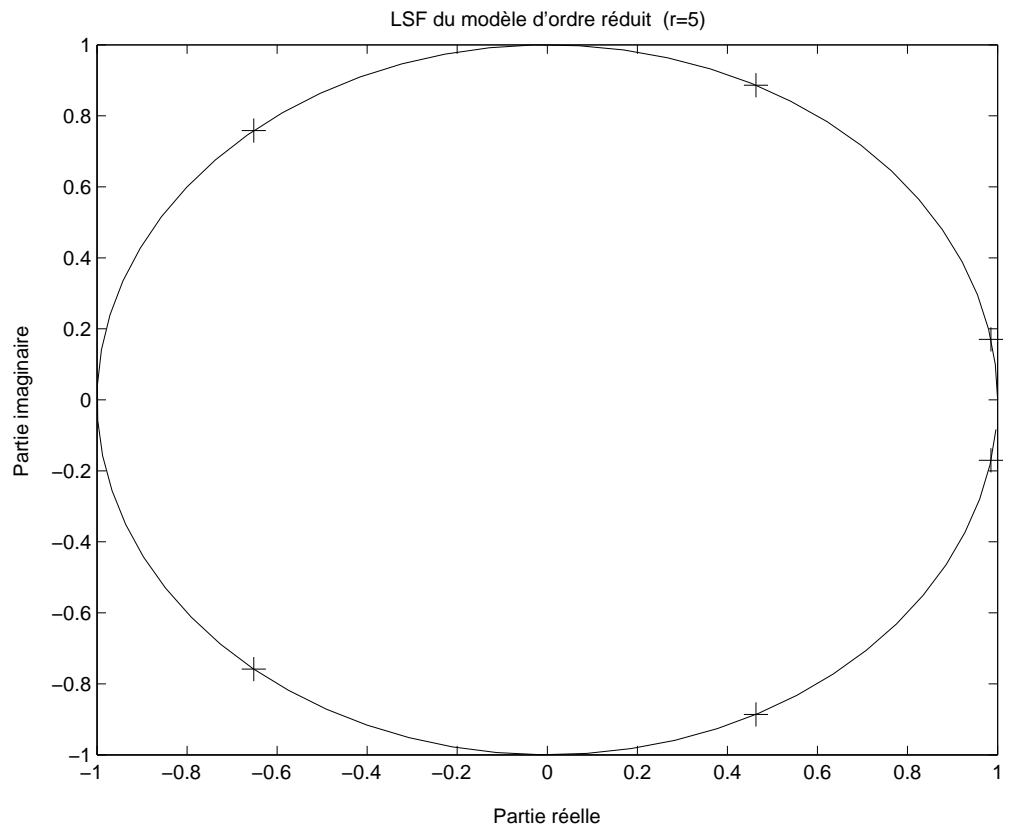


FIG. 4.11 – Lieu des LSF du modèle d'ordre réduit

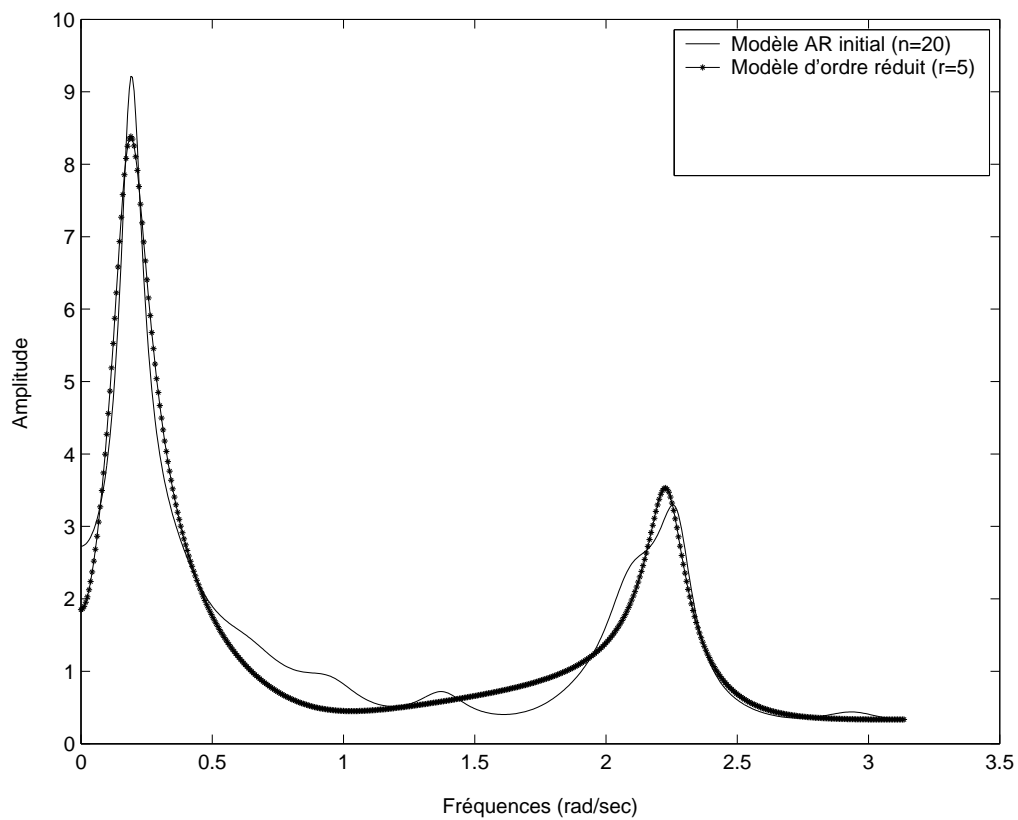


FIG. 4.12 – Spectres d'amplitude du système original, et celui d'ordre réduit



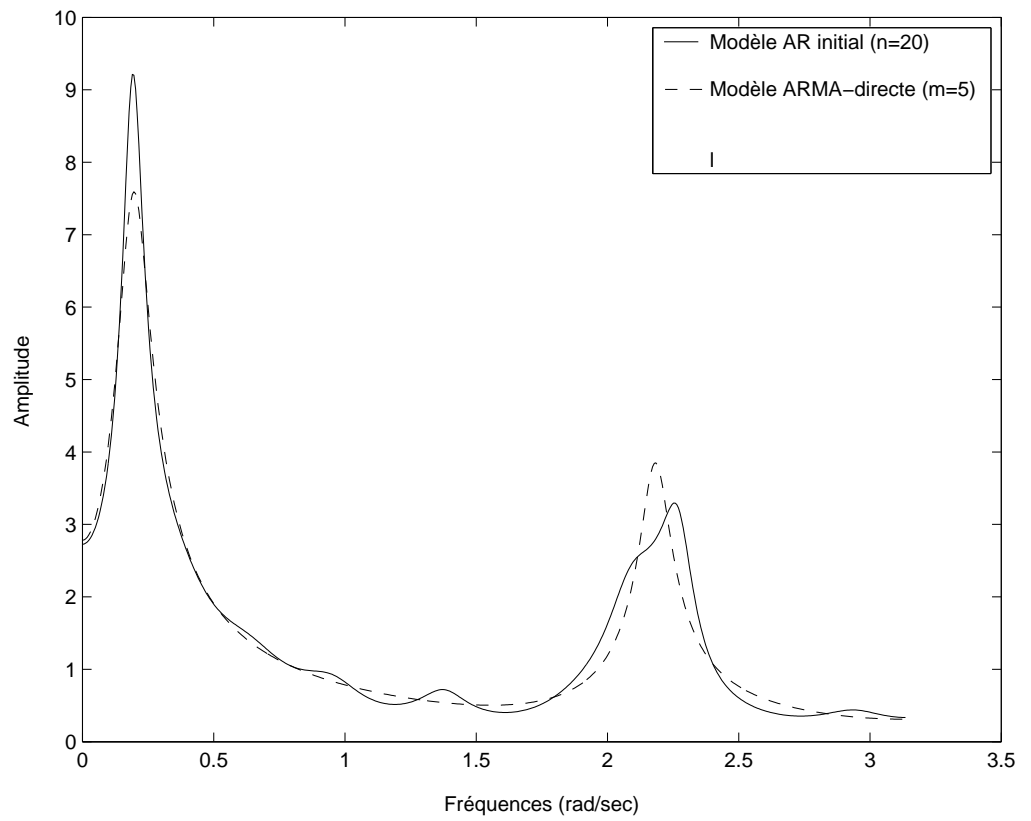


FIG. 4.13 – Spectres d'amplitude du système original, et celui ARMA-directe

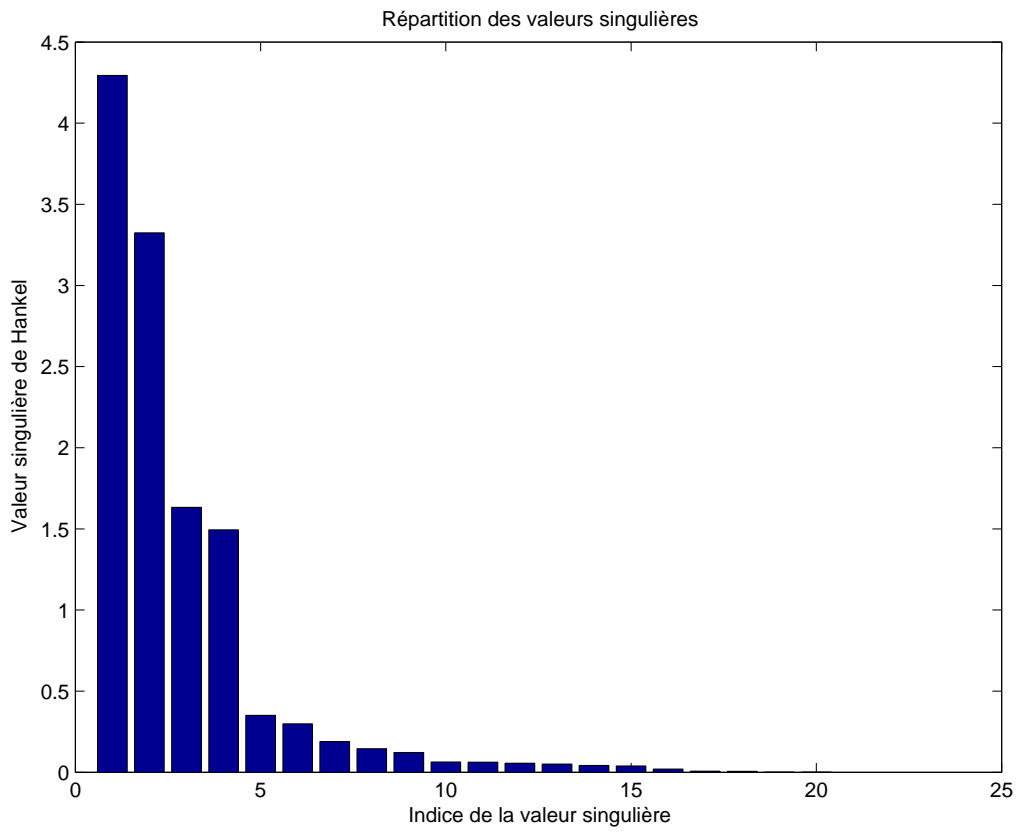


FIG. 4.14 – Répartition des valeurs singulières

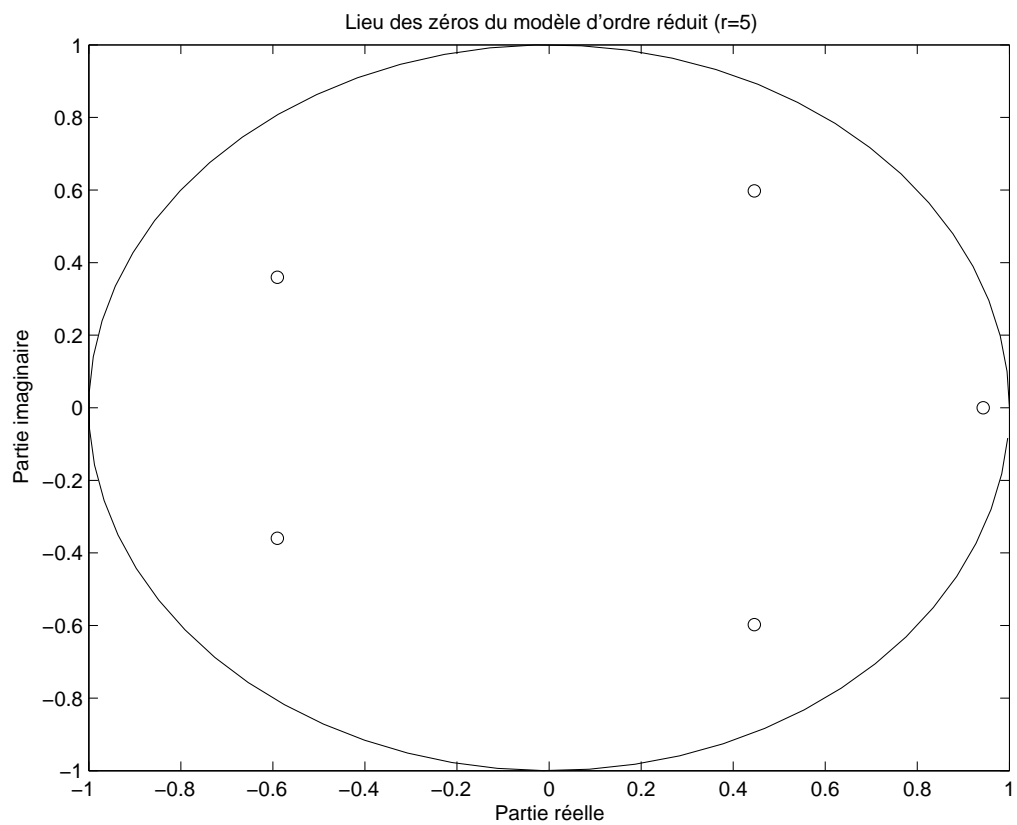


FIG. 4.15 – Lieu des zéros du modèle d'ordre réduit

## 4.4 Exemple 3

Soit un signal de parole numérisé avec une fréquence d'échantillonnage  $F_s = 11025Hz$ , correspondant à la phrase en langue arabe "*Essalam Alaikoum*". On construit un modèle  $AR$ , d'ordre 20, selon la méthode de *Burg*, dont les coefficients sont

$$ai=[1, -1.598, 1.064, -7.569e-1, 8.170e-1, -6.439e-1, 3.756e-1, -2.483e-1, 3.342e-1, -6.429e-1, 5.644e-1, -1.567e-1, 1.412e-1, -2.737e-1, 5.821e-2, 1.660e-2, 6.265e-3, -6.707e-2, 8.537e-2, -6.117e-2, -2.123e-3]$$

Les valeurs singulières de *Hankel* du modèle d'ordre complet (Fig. 4.21) sont données par  $\Sigma=[4.026e+1, 3.813, 2.499, 1.313, 1.296, 5.727e-1, 5.340e-1, 3.904e-1, 2.771e-1, 8.455e-2, 6.355e-2, 5.129e-2, 3.080e-2, 2.451e-2, 1.633e-2, 1.496e-3, 7.372e-4, 7.157e-4, 2.982e-4, 0]$ .

A partir de ce modèle d'ordre 20 et en utilisant l'approche de *Schur* (voir section 3.3), on synthétise un modèle d'ordre réduit  $r = 5$ ; en général, l'ordre du modèle simplifié  $r$  est déterminé en fonction de la distribution des valeurs singulières de *Hankel*. Le modèle d'ordre réduit 5 satisfait la borne de l'erreur de Hankel

$$\|H(z) - H_r(z)\|_{H_\infty} = \sup_{\omega \in \mathbb{R}} \|H(j\omega) - H_r(j\omega)\|_2 = 0.90 \leq 2 \sum_{i=6}^{20} \sigma_i = 4.09$$

La fonction de transfert du modèle simplifié d'ordre  $r = 5$  est donnée par

$$H_r(z) = \frac{0.88 - 2.05z^{-1} + 1.48z^{-2} + 0.31z^{-3} - z^{-4} + 0.41z^{-5}}{1 - 4.13z^{-1} + 7.32z^{-2} - 6.97z^{-3} + 3.56z^{-4} - 7.78z^{-5}}$$

On remarque que le modèle d'ordre réduit 5 possède une structure *ARMA*, donc défini par 10 coefficients, 5 du numérateur et 5 du dénominateur, tandis que celui initial d'ordre complet de structure *AR* est défini pas 20 coefficients.

Un autre modèle *ARMA* synthétisé directement en utilisant la fonction *Matlab ARMAX*, de même ordre  $m = 5$  a pour fonction de transfert

$$H_{ARMA}(z) = \frac{0.98 - 0.26z^{-1} - 0.08z^{-2} - 0.035z^{-3} - 0.36z^{-4}}{1 - 1.86z^{-1} + 1.41z^{-2} - 0.94z^{-3} + 0.60z^{-4} - 0.21z^{-5}}$$

La Fig. 4.16 donne, sur un même graphe, le signal de parole original, celui issu du modèle d'ordre réduit ainsi que l'écart de réduction. La Fig. 4.17 montre le lieu des pôles du modèle d'ordre réduit. La Fig. 4.18 montre bien que les *LSF* sont sur le cercle unité. La Fig. 4.19 donne les réponses fréquentielles du modèles initial d'ordre 20, et celui d'ordre réduit ( $r=5$ ), tandis que la Fig 4.20 donne les spectres initial et celui synthétisé directement par la fonction *ARMAX*. La Fig. 4.22 montre bien que le modèle simplifié présente une phase minimale.

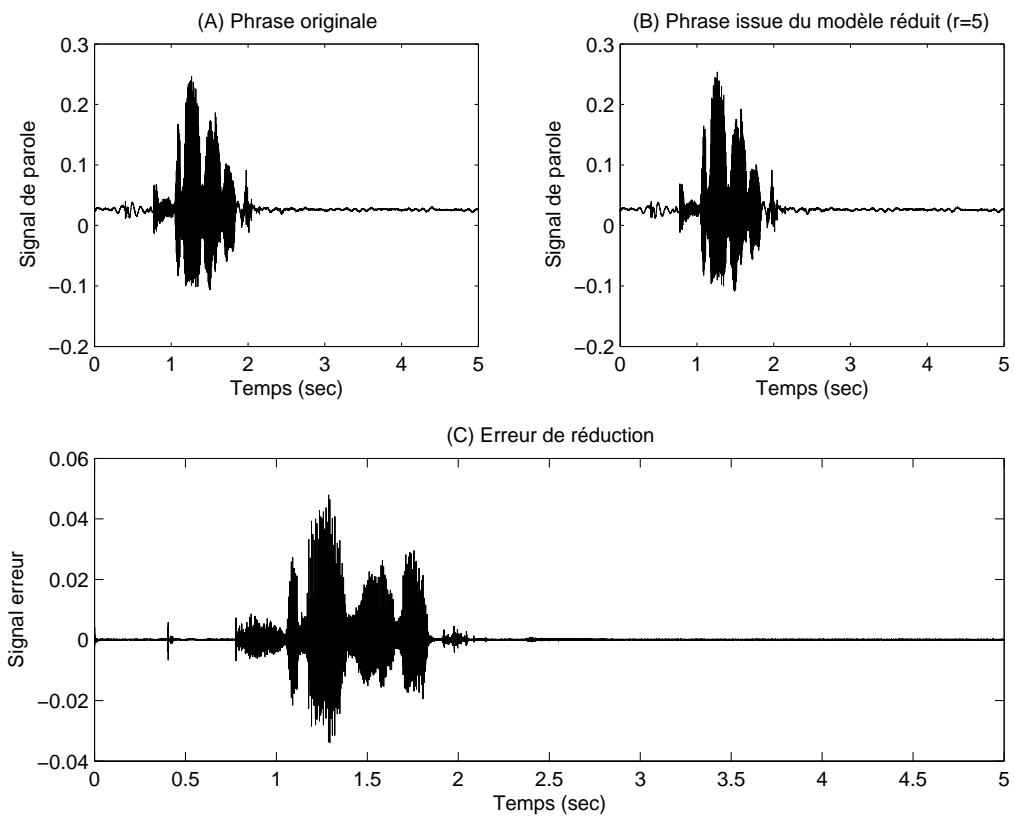


FIG. 4.16 – Signal original (A), d'ordre réduit (B) et erreur de réduction (C)

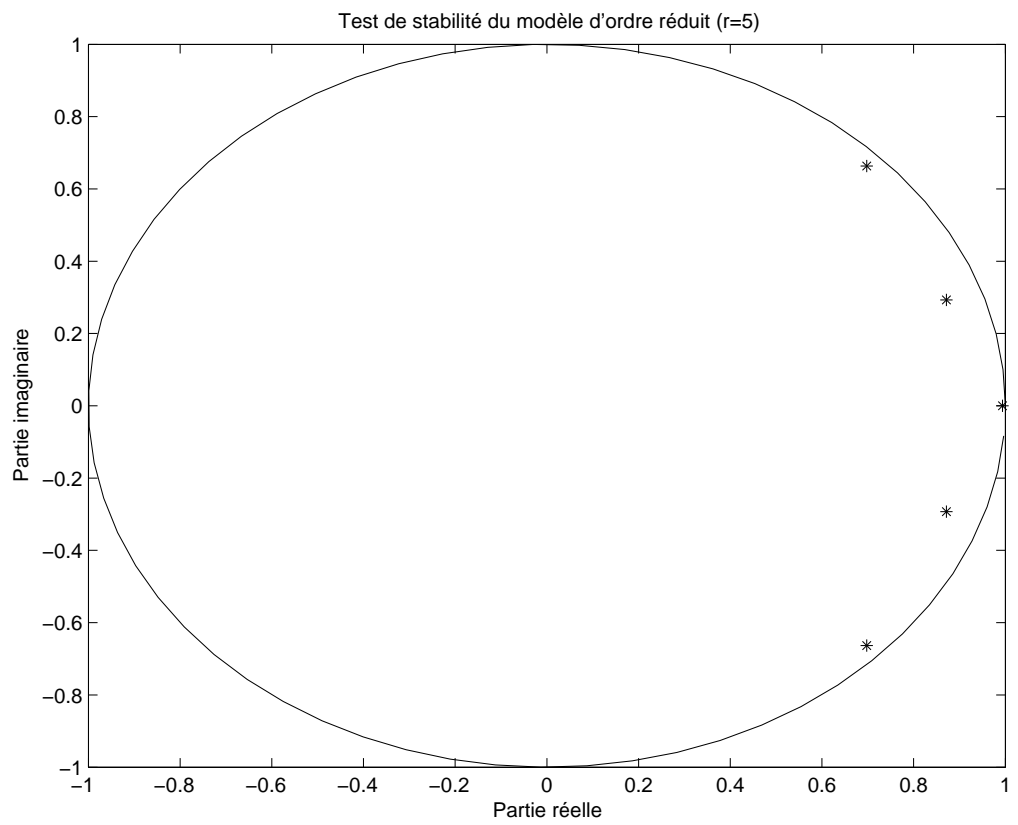


FIG. 4.17 – Lieu de stabilité du modèle d'ordre réduit

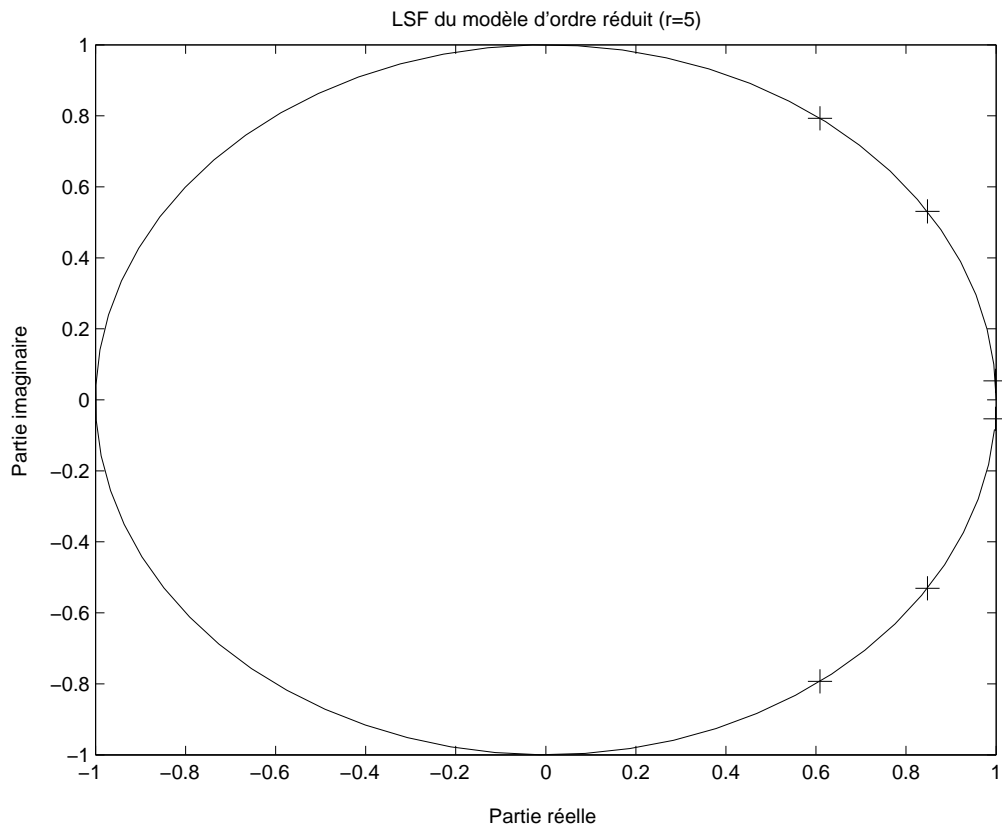


FIG. 4.18 – Lieu des LSF du modèle d'ordre réduit



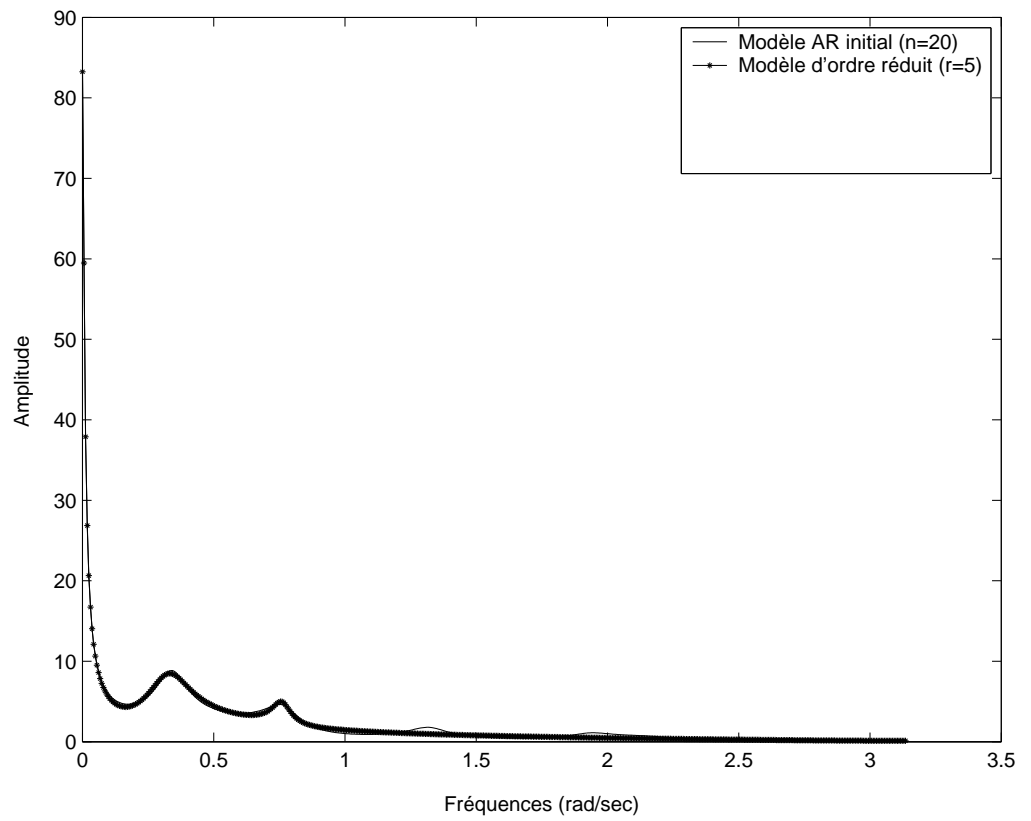


FIG. 4.19 – Spectres d'amplitude du système original, et celui d'ordre réduit

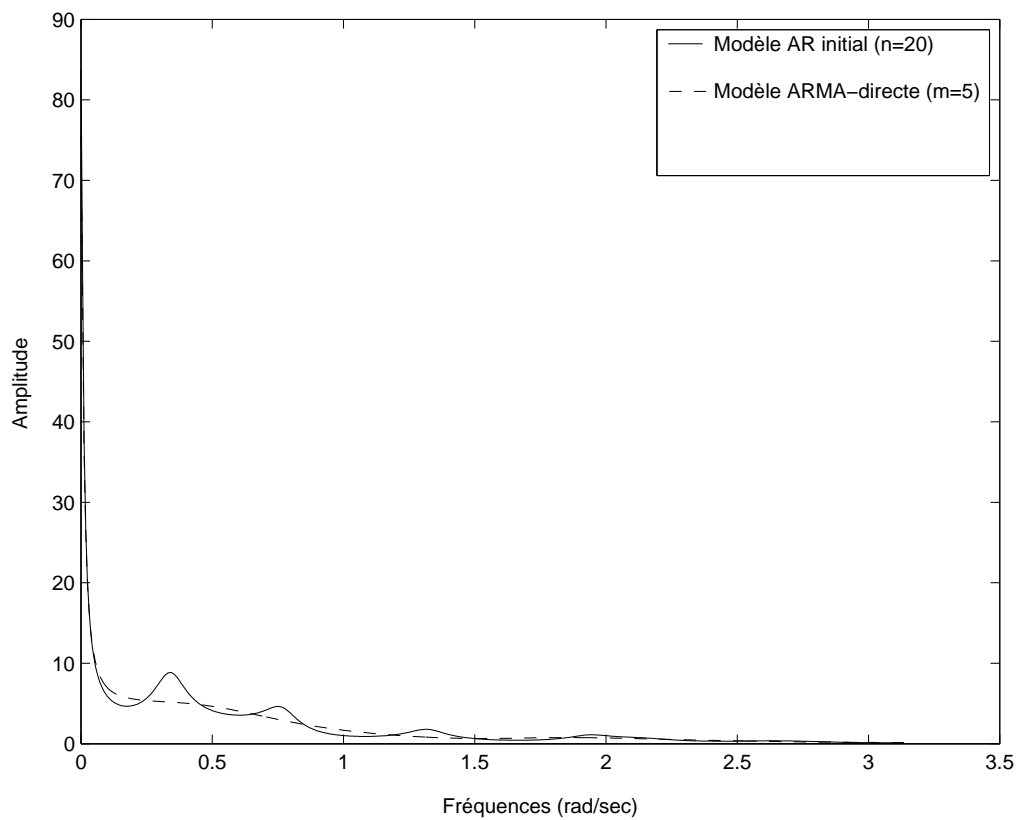


FIG. 4.20 – Spectres d'amplitude du système original, et celui ARMA-directe

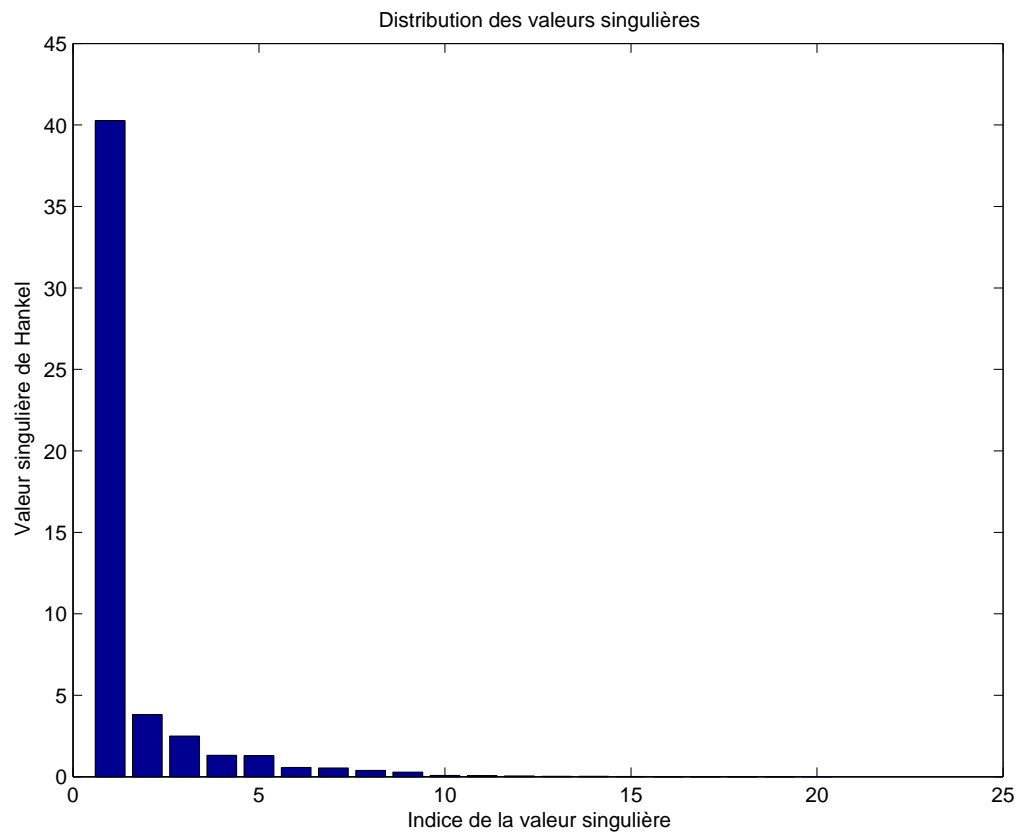


FIG. 4.21 – Répartition des valeurs singulières

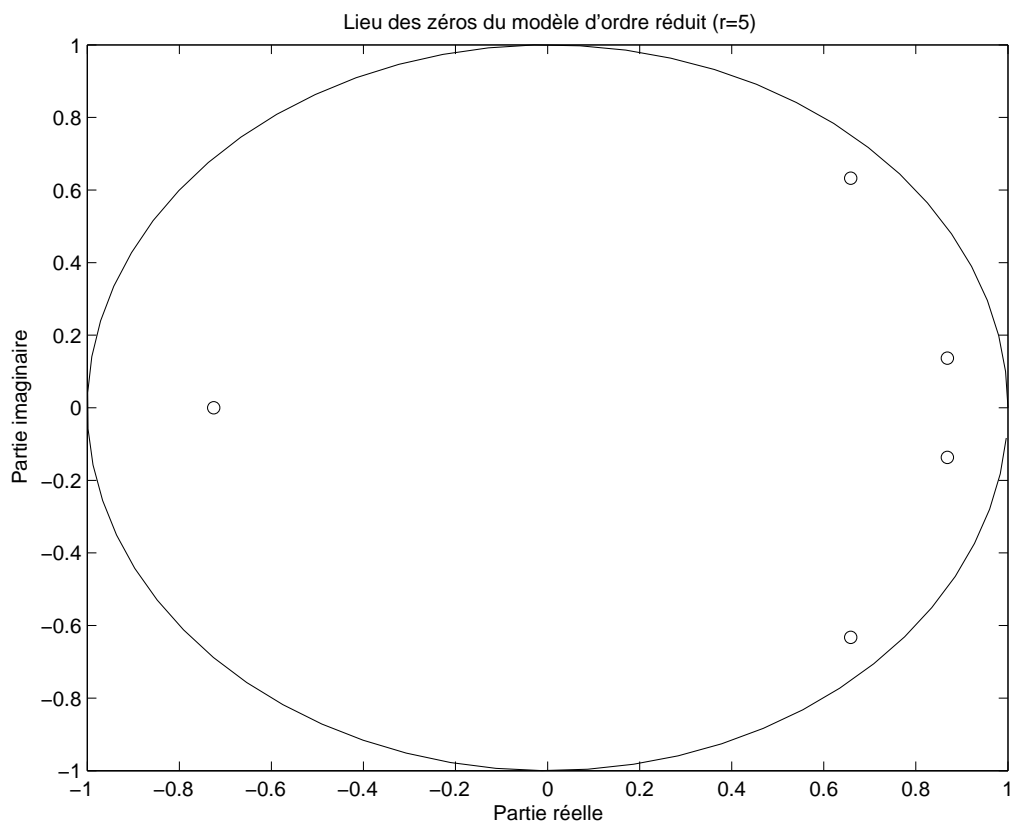


FIG. 4.22 – Lieu des zéros du modèle d'ordre réduit

## 4.5 Interprétations et conclusion

D'après les résultats de simulation, on tire les conclusions suivantes :

- Le  $SNR_r$  du modèle d'ordre réduit  $r$ , est supérieur comparé à celui du modèle  $ARMA$  généré directement par la fonction  $ARMAX\ SNR_{ARMA}$  du Matlab de même ordre  $m = r$ , ceci pour l'ensemble des syllabes et le segment de parole (voir Tab. 4.1, Fig. 4.7).
- Le modèle d'ordre réduit  $r$  génère un signal de parole dont la forme d'onde temporelle approche de très près celle originale issue du modèle  $AR$  d'ordre complet  $n$  (voir Fig. 4.1, Fig. 4.9 ,et Fig. 4.16), ceci est vérifié d'après la courbe d'écart d'erreur. Cette bonne approximation est due d'une part au fait que l'approche de réduction développée, basée sur le critère  $SVD$ , présente l'avantage d'éliminer les modes redondants, donc inutiles du système d'ordre complet, et d'autre part au fait que la modélisation classique  $LPC$  donne un modèle  $AR$  avec une surestimation de l'ordre initial.
- Les modèles d'ordre réduit sont toujours stables, à phase minimale et ayant des  $LSF$  situés sur le cercle unité.

# CONCLUSION GENERALE

Le présent travail traite une nouvelle approche de la synthèse du signal de parole par des modèles d'ordre réduit. L'idée est basée sur le fait que le signal de parole présente des redondances. Cette redondance se traduit par quelques états internes dans le modèle de représentation classique *AR* qu'on appelle aussi "modes" qui sont presque non observables à la sortie du modèle (notons que la sortie du modèle de représentation n'est rien d'autre que le signal de parole). Nous avons exploité la théorie de la réduction de modèles, en particulier celle qui opère dans l'espace d'état. Nous avons généralisé l'approche de réduction basée sur les projections internes, et les formes de *Schur*, aux systèmes numériques dans le but de construire des systèmes de production de parole. Elle reste aussi valable pour les modèles non minimaux (non observables et/ou non contrôlables). D'après les résultats de simulation, on tire les conclusions suivantes

- Le modèle d'ordre réduit  $r$  génère un signal de parole dont la forme d'onde temporelle approche de très près celle originale issue du modèle *AR* d'ordre complet  $n$ , ceci est vérifié par la courbe d'écart d'erreur. Cette bonne approximation est due d'une part au fait que l'approche de réduction développée, basée sur le critère

*SVD*, présente l'avantage d'éliminer les modes redondants, donc inutiles du système d'ordre complet, et d'autre part au fait que la modélisation classique *LPC* donne un modèle *AR* avec une surestimation de l'ordre initial.

- Les modèles d'ordre réduit correspondant à chaque syllabes (voir Tab. 4.1) donnent un meilleur rapport signal sur bruit *SNR* comparés à ceux *ARMA* synthétisés directement par la fonction *ARMAX* du *Matlab*, et de même ordre. Notons que le rapport *SNR* est une mesure directe de l'effet du bruit sur le signal utile lors d'une transmission, car il est défini comme étant le rapport de la puissance totale du signal utile (dans notre cas le signal de parole ciblé) sur la puissance totale du bruit.

# Bibliographie

- [1] R. Boite, H. Boulard, T. Dutoit, J. Hancq, et H. Leich, *Traitement de la parole*, Presses polytechniques et universitaires romandes, Lausanne, 2000.
- [2] L. Mitiche, B. Derras, and A. B. H. Mitiche-Adamou, '*Speech modeling via model reduction*', Proceeding of the IEEE-EURASIP Seventh International Symposium on Signal Processing and its Applications ISSPA 2003, Vol. I, pp. 381-384, Paris, France, 1-4 July 2003.
- [3] L. Mitiche, A.B.H. Adamou-Mitiche, '*A new approach for speech modeling based on model reduction*', Proceeding of the IEEE-EURASIP Firt International Symposium on Control, Communications, and Signal Processing, ISCCSP 2004, pp. 607-610, Hammamet, Tunisia, 21-24 March 2004.
- [4] L. Mitiche, B. Derras, and A.B.H. Adamou-Mitiche, '*Speech modeling via model reduction : A SNR behaviour*', Electronics Letters, IEE, Vol. 39, No. 17, pp. 1288-1290, 21st August 2003.
- [5] L. Mitiche, B. Derras, and A.B.H. Adamou-Mitiche, '*Efficient low-Order auto regressive moving average (ARMA) Models for Speech Signals*', ARLO-ASA, Acoustic Re-



- search Letters, Acoustical Society of America, Vol. 5, Issue 2, pp. 75-81, April 2004.
- [6] L. Mitiche, D. Berkani, " *Speech Modeling via Model Reduction*", International Journal of Speech Technology, Kluwer Academic Publishers, Issue 7(4) : pp. 311-318, October 2004.
- [7] D. G. Childers, *Modern spectrum analysis*, IEEE press, New York, 1978.
- [8] L. R. Rabiner and B. Gold, *Theory and application of digital signal processing*, Prentice-Hall, Inc. Englewood Cliffs, New Jersey, USA, 1975.
- [9] M. Nadjim, *Modélisation et identification en traitement du signal*, Edition Masson, Paris, 1988.
- [10] C. Ma, Y. Kamp, et L.F. Willems, *Robust signal selection for linear prediction analysis of voiced speech*, Speech Communication, n° 12, pp. 69-81, 1993.
- [11] J. Makoul, *Linear prediction : a tutorial review*, Proc. IEEE V. 63 n°4 (April 1975), pp. 561-580.
- [12] J.D. Markel, and A.H. Gray, *Linear prediction of speech*, Springer Verlag, New York, 1980.
- [13] G. S. Kang and L. J. Franssen, 'Application of Line Spectrum Pairs to low bit rate speech encoder', in Proc. Int. Conf. Acoust. Speech and Signal Processing, pp. 7.3.1-7.3.4, Tempa. FL., Apr. 1985.
- [14] F. Itakura and Sagamura, 'LSP Synthesizer, its principles and implementation', Tech Rep S79.46, Speech study group. Acoust. Soc., Japan, Nov. 1979.

- [15] H. Wakita, '*Linear predictive voice synthesizers : Line Spectrum Pairs (LSP) is the newest of the several technique*', Speech Tech., Vol. 1., pp. 17-22, 1981.
- [16] A.V. Oppenheim, et R.W. Schaffer, '*Digital Signal Processing* ', Prentice-Hall , Englewood Cliffs, NJ, 1975.
- [17] N. S. Jayant and P. Noll, '*Digital coding of waveforms principles and applications to speech and video*', Prentice-Hall, Inc. Englewood Cliffs, New Jersey, USA, 1984.
- [18] K. Feher (Editor), '*Advanced digital communications systems and signal processing techniques*', Prentice-Hall, Inc. Englewood Cliffs, New Jersey, USA, 1987.
- [19] L. Mitiche and A. B. H. Adamou, '*Model reduction via balanced realizations : Application to digital prototype filters*', International Conference on Signals and Systems ICSS'94, Vol. 1, pp. III.54-III.57, Algiers Sept. 24-26, 1994.
- [20] B. C. Moore, '*Principal component analysis in linear systems : Controllability, Observability and model reduction*', IEEE Trans. on Automatic Control, Vol. AC-26, pp. 17-32, Feb. 1981.
- [21] L. Pernebo and L. Silverman, '*Model reduction via balanced state space representation*', IEEE Trans. on Automatic Control, Vol. AC-27, No. 2, pp. 382-387, April 1982.
- [22] M. Bettayeb, '*New interpretation of balancing state space representation as an input-output energy minimization problem*', International Journal Systems SCI, pp. 325-331, 1991.

- [23] L. Mitiche, A. B. H. Adamou, '*A Schur method for Model reduction (nonminimal case)*', International Conference on Signals and Systems, Vol. 2, pp. VI.63-VI.66, Algiers Sept. 24-26, 1994.
- [24] M. G. Safonov and R. Y. Chiang, '*A Schur method for Balanced model reduction*', IEEE Transactions On Automatic Control, Vol. 34, No. 2, July 1989.
- [25] A. Laub, M. T. Heath, C. C. Paige and R. C. Ward, '*Computation of systems balancing transformations and other applications of simultaneous diagonalization algorithms*', IEEE Trans. on Automatic Control, Vol. AC-32, No. 2, pp. 115-122, Feb. 1987.
- [26] K. Glover, '*All optimal Hankel-norm approximations of linear multivariable systems and their  $L^\infty$  error bounds*', International Journal of Control, Vol. 39, No. 6, pp. 1115-1193, 1984.
- [27] J. E. Mason, '*Identification using low order systems*', Ph.D Dissertation, University of California, Berkely, 1988.
- [28] P. T. Kabamba, '*Balanced gains and their significance for  $L^2$  model reduction*', IEEE Trans. on Automatic Control, Vol. AC-30, pp. 690-693, July 1985.
- [29] P. T. Kabamba, '*Balanced forms : Canonicity and parametrization*', IEEE Trans. on Automatic Control, Vol. AC-30, No. 1, pp. 1106-1109, Nov. 1985.
- [30] C. T. Mullis and R. A. Roberts, '*Synthesis of minimum Roundoff noise fixed point digital filters*', IEEE Transactions on circuits and systems, Vol. Cas-23, No. 9, September, 1976.

- [31] Rabah W. Aldhaheri, '*Model order reduction via real Schur-form decomposition*', International Journal of Control, 1991, Vol. 53, No. 3, 709-716.
- [32] G. Strang, *Linear algebra and its applications*, Academic Press, Inc., USA, 1976.
- [33] T. Kailath, *Linear systems*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1980.
- [34] L. Ljung, *System identification theory for the user*, Prentice-Hall Ptr, Upper Saddle River, New Jersey, USA, 1987.
- [35] Nicolas Moreau, *Techniques de compression des signaux*, édition Masson, Paris, 1994
- [36] '*Matlab 6.5 Release 13, System Identification Toolbox, Ver. 5.02 Release 13*', The Mathworks, Inc.
- [37] L. Ljung, *System Identification - Theory for the User*, Prentice Hall, Upper Saddle River, N.J. 2nd edition, 1999.
- [38] A.B.H. Adamou-Mitiche, L. Mitiche, '*Comparative study of model reduction schemes-application to the digital filters synthesis*', Proceeding of the IEEE sixth International Symposium on signal processing and its applications ISSPA 2001, Vol. 2, pp. 675-678, 13-16 August 2001, Kuala Lumpur, Malaysia.



## Annexe A

# FACTORISATION DE CHOLESKY

Les opérations de factorisation et de triangularisation de matrices constituent les opérations de base de nouveaux algorithmes de contrôle ou de traitement de signal.

Ainsi la procédure de décomposition  $LU$  d'une matrice  $A(n, n)$ , avec

- $L$  : matrice triangulaire inférieure,
- $U$  : matrice triangulaire supérieure,

permet par exemple de ramener la résolution du système d'équations linéaires

$Ax = B$  où  $A = n \times n$  - matrice,  $B, x = n$  - vecteur à la résolution de deux systèmes particuliers

- Système triangulaire inférieur  $Lx = B$ .

– Système triangulaire supérieur  $Ux = Y$ .

Si  $A$  est une matrice symétrique définie positive, alors l'algorithme de *Cholesky* nous affirme que  $A$  peut se mettre sous la forme

$$A = SS^T, \quad (\text{A.1})$$

où  $S$  est une matrice triangulaire inférieure.

$$\begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \\ \vdots & \cdots & \ddots & \cdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} s_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \cdots \\ \vdots & \cdots & \ddots & 0 \\ s_{n1} & \cdots & \cdots & s_{nn} \end{bmatrix} \begin{bmatrix} s_{11} & \cdots & \cdots & s_{n1} \\ 0 & \ddots & \cdots & \cdots \\ \vdots & \cdots & \ddots & \cdots \\ 0 & \cdots & 0 & s_{nn} \end{bmatrix} \quad (\text{A.2})$$

Certains auteurs considèrent  $S$  comme étant la racine carrée de la matrice  $A$ , du fait qu'une telle décomposition (*Cholesky*) est unique. L'algorithme de factorisation de *Cholesky* permet de calculer les coefficients  $s_{i,j}$  de la matrice  $S$  par récurrence.

### Algorithme

– **Entrée**-  $A = n \times n$  – matrice

For  $i = 1$  to  $n$  Do

$$s_{i,j} = \sqrt{a_{i,j} - \sum_{k=1}^{i-1} s_{i,k}^2}$$

If  $s_{i,j} = 0$ , then Halt ( $A$  est non définie positive)

$$\text{while } j > i \text{ and } j \leq n \text{ Do } s_{i,j} = \frac{1}{s_{i,i}} [a_{i,j} - \sum_{k=1}^{i-1} s_{i,k} s_{j,k}]$$

End loop on  $i$ .

– **Sortie**-  $S = n \times n$  – matrice triangulaire inférieure.

## Annexe B

# VALEURS SINGULIERES ET ALGORITHME SVD

Les valeurs singulières d'une matrice  $A \in C^{m,n}$  de rang  $q$ , notées  $\sigma_i$ , sont les racines carrées non négatives des valeurs propres de  $A^T A$ , ordonnées telles que :

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0.$$

Si  $q < n$  nous aurons  $n - q$  valeurs singulières nulles, c'est à dire :

$$\sigma_{q+1} = \sigma_{q+2} = \dots = \sigma_n = 0.$$

Il existe deux matrices orthogonales  $U \in C^{m,m}$ ,  $V \in C^{n,n}$  et une matrice diagonale  $\Sigma \in C^{m,n}$  telles que [32]

$$A = U\Sigma V^T = U \begin{bmatrix} \Sigma_q & 0 \\ 0 & 0 \end{bmatrix} V^T, \quad (\text{B.1})$$



et les valeurs singulières  $\sigma_i$ ,  $i = \overline{1, q}$  sont les racines carrées des  $q$  valeurs propres positives (non nulles) de  $A^T A$ .

$U$  et  $V$  sont les matrices orthogonales ayant pour colonnes les vecteurs propres de  $A^T A$  et  $AA^T$  respectivement.

Cette décomposition est dite "Décomposition en Valeurs Singulières" *SVD* de la matrice  $A$ .

### Remarque

Si  $A = U\Sigma V^T$  est la décomposition en valeurs singulières de la matrice  $A$ , alors les décompositions en valeurs propres de  $A^T A$  et de  $AA^T$  sont données par [32]

$$A^T A = V(\Sigma^T \Sigma)V^T = V\Sigma^2 V^T, \quad (\text{B.2})$$

$$AA^T = U(\Sigma \Sigma^T)U^T = U\Sigma^2 U^T. \quad (\text{B.3})$$

Notons

$\bar{\sigma}(A) = \sigma_1$  [la plus grande valeur singulière de la matrice  $A$ ],

$\underline{\sigma}(A) = \sigma_n$  [la plus petite valeur singulière de la matrice  $A$ ].

Les valeurs singulières maximum et minimum d'une matrice  $A$ , notées respectivement  $\bar{\sigma}$  et  $\underline{\sigma}$ , sont équivalentes en terme de norme spectrale à

$$\bar{\sigma}(A) = \|A\|_2, \quad (\text{B.4})$$

$$\underline{\sigma}(A) = \begin{cases} \|A^{-1}\|_2^{-1} & \text{si, } \det(A) \neq 0, \\ 0 & \text{si, } \det(A) = 0 \end{cases} \quad (\text{B.5})$$

La valeur singulière minimale  $\underline{\sigma}(A)$  = donne une mesure de la tendance à la singularité de la matrice  $A$ .

### Propriétés des valeurs singulières [32]

1.

$$\underline{\sigma}(A) = \min \frac{\|Ax\|}{\|x\|}, x \in C^n, \quad (\text{B.6})$$

2.

$$\bar{\sigma}(A) = \max \frac{\|Ax\|}{\|x\|}, x \in C^n, \quad (\text{B.7})$$

3.

$$\underline{\sigma}(A) \leq |\lambda_i(A)| \leq \bar{\sigma}(A), \quad (\text{B.8})$$

où  $\lambda_i(\cdot)$  est la  $i^{\text{ème}}$  valeur propre de  $(\cdot)$ ,

4. Si  $A^{-1}$  existe,

$$\underline{\sigma}(A) = \frac{1}{\bar{\sigma}(A^{-1})}, \quad (\text{B.9})$$

5. Si  $A^{-1}$  existe,

$$\bar{\sigma}(A) = \frac{1}{\underline{\sigma}(A^{-1})}, \quad (\text{B.10})$$

6.

$$\bar{\sigma}(\alpha A) = |\alpha| \bar{\sigma}(A), \quad (\text{B.11})$$

7.

$$\bar{\sigma}(A + B) \leq \bar{\sigma}(A) + \bar{\sigma}(B), \quad (\text{B.12})$$

8.

$$\bar{\sigma}(AB) \leq \bar{\sigma}(A) \bar{\sigma}(B), \quad (\text{B.13})$$

9.

$$\max[\bar{\sigma}(A), \bar{\sigma}(B)] \leq \bar{\sigma}(AB) \leq \sqrt{2} \max[\bar{\sigma}(A), \bar{\sigma}(B)], \quad (\text{B.14})$$

10.

$$\max |a_{i,j}| \leq \bar{\sigma}(A) \leq n \max |a_{i,j}|, \text{ pour } i, j \quad (\text{B.15})$$

11.

$$\sum_{i=1}^n \sigma_i^2 = \text{tr}[A^T A]. \quad (\text{B.16})$$

## Annexe C

# ALGORITHME D'EQUILIBRE (MODELES NUMERIQUES)

### Equilibre de modèles

L'équilibre des modèles est l'étape fondamentale sur laquelle repose la *MRE*. Il consiste en la détermination d'une transformation non singulière  $T$  dont le calcul est essentiellement basé sur les grammiens de commandabilité et d'observabilité du modèle [21].

### Procédure de l'équilibre [41]

– **Entrées**- Ayant la réalisation d'ordre  $n$  ( $A, B, C, n$ ).

– **Etape 1**-

Calcul des grammiens  $W_c$  et  $W_o$ , solutions des équations de *Lyapunov* [41]

$$A^T W_o A - W_o + C^T C = 0, \quad (\text{C.1})$$

$$AW_c A^T - W_c + BB^T = 0. \quad (\text{C.2})$$

Ces deux grammien sont donnés par les quantités suivantes

$$W_c = \sum_{k=0}^{\infty} A^k BB^T (A^T)^k, \quad (\text{C.3})$$

$$W_o = \sum_{k=0}^{\infty} (A^T)^k C^T C A^k. \quad (\text{C.4})$$

**Remarque.**

Si le modèle est asymptotiquement stable,  $W_c$  et  $W_o$  convergent pour un nombre fini de termes.

– **Etape 2-**

Factorisation de *Cholesky* (Annexe A) de la paire  $(W_c, W_o)$  telle que

$$W_o = L_o L_o^T, \quad (\text{C.5})$$

$$W_c = L_c L_c^T, \quad (\text{C.6})$$

avec  $L_o$  et  $L_c$  des matrices triangulaires inférieures.

– **Etape 3-**

Décomposition en valeurs singulières *SVD* (Annexe B) de la quantité

$$M = L_o^T L_c, \quad (\text{C.7})$$

telle que

$$M = U \Sigma V^T, \quad (\text{C.8})$$

où  $U$  et  $V$  sont des  $n \times n$  – *matrices orthogonales*.

$\Sigma = n \times n$  – *matrice diagonale* (matrice des valeurs singulières du système) telle

que

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) \quad \text{avec} \quad \sigma_i > \sigma_j \quad \text{pour} \quad i > j.$$

– **Etape 4-**

Obtention de la transformation d'équilibre  $T$  et de son inverse  $T^{-1}$

$$T = L_c V \Sigma^{-1/2}, \quad (\text{C.9})$$

$$T^{-1} = \Sigma^{-1/2} U^T L_o^T. \quad (\text{C.10})$$

– **Etape 5-**

Construction de la réalisation équilibrée

$$A_e = T^{-1} A T = \Sigma^{-1/2} U^T L_o^T A L_c V \Sigma^{-1/2}, \quad (\text{C.11})$$

$$B_e = T^{-1} B = \Sigma^{-1/2} U^T L_o^T B, \quad (\text{C.12})$$

$$C_e = C T = C L_c V \Sigma^{-1/2}. \quad (\text{C.13})$$

– **Sorties-**

La réalisation équilibrée du modèle d'ordre complet  $(A_e, B_e, C_e, n)$ .

On montre bien que les grammien d'observabilité et de commandabilité s'expriment dans la base d'équilibre par

$$\begin{aligned} W_{oe} &= T^T W_o T \\ &= \Sigma^{-1/2} V^T L_c^T L_o L_o^T L_c V \Sigma^{-1/2} \\ &= \Sigma^{-1/2} V^T M^T M V \Sigma^{-1/2} \\ &= \Sigma, \end{aligned} \quad (\text{C.14})$$

et

$$\begin{aligned}W_{ce} &= T^{-1}W_cT^{-T} \\ &= \Sigma^{-1/2}U^T L_o^T L_c L_c^T L_o U \Sigma^{-1/2} \\ &= \Sigma^{-1/2}U^T M M^T U \Sigma^{-1/2} \\ &= \Sigma,\end{aligned}\tag{C.15}$$

d'où  $(W_{oe}, W_{ce}) = (\Sigma, \Sigma)$ .

De point de vue mathématique,  $\Sigma$  est aussi solution des équations de *Lyapunov*

$$A_e \Sigma A_e^T - \Sigma + B_e B_e^T = 0,\tag{C.16}$$

$$A_e^T \Sigma A_e - \Sigma + C_e^T C_e = 0.\tag{C.17}$$

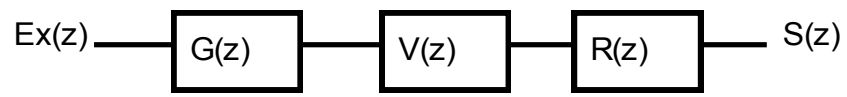


FIG. 1.5.a- Mise en cascade des fonctions de transfert



## ملخص

تتطرق هذه الأطروحة إلى تمثيل إشارة الكلام باستعمال نماذج جديدة مصغرة من نوع ARMA إنطلاقاً من النموذج الكلاسيكي AR وباستعمال خوارزمية جديدة و المرتكزة على الإسقاطات في فضاء الحالات وعلى صيغة شور نحصل على نموذج جديد مصغر ذو درجة أقل .  
تتمن خصوصية فكرتنا في إقصاء الحالات الأقل ملاحظة و الأقل مراقبة من النموذج الأصلي .  
من مميزات نموذجنا الجديد كونه دوماً مستقرًا ذو أدنى طور كما أنه ينتج إشارة الكلام بجودة عالية .

## Abstract

Using model reduction, a new approach for low-order speech modelling is presented. In this approach, the modeling process starts with a relatively high-order (full-order) autoregressive AR model obtained by some classical methods. The AR model is then reduced using the a state projection method, operating in the state space. The model reduction yields a reduced-order autoregressive moving-average ARMA model which interestingly preserves the key properties of the original full-order model such as causality, stability, minimality, and phase minimality. Signal-to-noise ratio SNR behavior is also investigated. To illustrate the performance and the effectiveness of the proposed approach, some computer simulations are conducted on some practical speech segments.

**Keywords:** Autoregressive AR model, Autoregressive moving average ARMA model, Model reduction, singular values, Speech modeling, State space.

## Résumé

Nous présentons dans ce travail une nouvelle approche de modélisation des systèmes de production de parole par des modèles d'ordres réduits. Dans cette approche, la modélisation commence par un modèle d'ordre assez élevé de type AR (Auto régressif) obtenu en utilisant les méthodes classiques. Ce modèle AR est ensuite réduit en utilisant les projections internes, qui opèrent dans l'espace d'état du modèle initial. La réduction d'ordre conduit à un modèle d'ordre réduit de type ARMA (Auto régressif à moyenne ajustée) qui préserve les propriétés clé du modèle initial d'ordre complet, voire la stabilité, la minimalité, et la phase minimale. La réduction d'ordre s'effectue dans le sens où les états internes du modèle initial les faiblement observables et faiblement contrôlables, correspondant aux valeurs singulières faibles, sont éliminées. Pour faire ressortir les valeurs singulières du modèle original d'ordre complet, l'algorithme SVD (Singular Values Decomposition) est utilisé. Nous nous sommes particulièrement intéressés aux comportements des paramètres SNR (Signal to Noise Ratio) issus de la réduction d'ordre. Afin d'illustrer et de mettre en valeur les performances et l'efficacité de notre approche, diverses simulations sur des segments de signaux de parole (phonèmes, diphtonges et phrases) ont été effectuées.

**Mots clés:** Traitement de la parole, Modèle Auto Régressif AR, Modèle Auto Régressif à Moyenne Ajustée ARMA, Réduction de modèles, Modélisation de la parole, Espace d'état.