

1996/1997

RAPPORT DE STAGE

Stage ingénieur de 3^{ème} année effectué à l'I.R.C.A.M.

du 11 septembre au 31 décembre 1996.

Estimation des paramètres de partiels par modélisation de la Transformée de Fourier à court-terme utilisant des fenêtres spectrales sans lobes

Thomas HÉLIE

Date de soutenance: jeudi 16 janvier 1997

Responsable de stage (IRCAM) : Philippe DEPALLE
Correspondant de stage (ENST) : Antoine CHAIGNE

ÉCOLE NATIONALE SUPÉRIEURE DES TÉLÉCOMMUNICATIONS

Remerciements

Outre les précieux enseignements que mon stage à l'IRCAM m'a apportés, il est à souligner combien il est agréable de travailler en ces lieux : la compétence des chercheurs, la liberté d'organisation offerte, la chaleur de l'accueil, et le climat sympathique présent au quotidien y sont tous réunis.

Et, je tiens particulièrement à remercier, Xavier Rodet et Philippe Depalle du laboratoire d'*Analyse-Synthèse* pour la qualité de leur encadrement et de leurs conseils, ainsi que l'ensemble des thésards pour leur bonne humeur permanente.

Enfin, un nouveau remerciement à Philippe Depalle pour sa patience.

Résumé

Ce rapport s'attache à présenter une méthode d'analyse additive qui permet d'améliorer la qualité des paramètres (fréquence , amplitude , phase). Cette méthode utilise un modèle paramétrique de la *Transformée de Fourier à Court Terme*, ce qui permet non seulement de prendre en considération l'influence de chaque partiel sur le spectre dans l'estimation, mais aussi de réduire sensiblement la taille de la fenêtre d'analyse.

La méthode est d'abord décrite puis étudiée. Dans un soucis d'efficacité, il est ensuite accordé une attention particulière aux fenêtres utilisées. Enfin, un algorithme d'analyse, autonome et fondé sur cette méthode analytique est construit.

Ces travaux succèdent à la première étude de la méthode, qu'a effectuée Laurent Tromp au sein de l'IRCAM en 1995 dans le cadre d'un DEA [Tro95].

Table des matières

Préambule - Remerciements	3
Résumé	5
Table des matières	9
Liste des tableaux	11
Table des figures	13
Introduction	15
I Estimation des paramètres additifs par modélisation analytique du spectre à court-terme	19
1 Représentation et modélisation d'un signal sonore	21
1.1 Modélisation dans le domaine temporel	21
1.2 Modélisation dans le domaine fréquentiel	22
2 Description de la méthode	25
2.1 Théorie et estimation	25
2.1.1 Estimation des amplitudes et des phases	26
2.1.2 Estimation des fréquences	28
2.2 Formulation du procédé d'estimation	30
3 Validation première de la méthode	32
II Recherche de fenêtres de pondération adaptées à la méthode	35
1 Influence de la fenêtre de pondération sur la méthode	37
1.1 Considérations générales	37

1.1.1	De l'estimation des amplitudes et des phases	38
1.1.2	De l'estimation des fréquences	39
1.2	Étude de la convergence en utilisant les fenêtres classiques . .	39
1.2.1	Cas d'une bonne convergence	40
1.2.2	Cas d'une mauvaise convergence	41
2	Construction de fenêtres adaptées	44
2.1	Première famille	45
2.1.1	Définition	45
2.1.2	Caractéristiques	45
2.2	Seconde famille	47
2.2.1	Définition	47
2.2.2	Caractéristiques	47
2.3	Vérification première de l'adéquation des familles au problème posé	49
3	Tests comparatifs de fenêtres sur diverses situations sonores	51
3.1	Tests 1 & 2 : variation des amplitudes	53
3.1.1	Test 1 : variation non linéaire des amplitudes	53
3.1.2	Test 2 : variation croissante d'une seule des amplitudes	53
3.2	Tests 3 & 4 : variation des fréquences	56
3.2.1	Test 3 : variation non linéaire des fréquences	56
3.2.2	Test 4 : évolution d'une seule des fréquences vers la valeur d'une autre	56
3.3	Résultats	59
III	Construction d'une chaîne algorithmique complète pour la mesure des paramètres additifs	61
1	Mise en oeuvre de la chaîne algorithmique	63
1.1	Initialisation des fréquences	63
1.1.1	Détermination des fréquences approchées présentes dans le fragment sonore	63
1.1.2	Détermination du nombre de points du signal à prélever localement	65
1.2	Problèmes liés à l'exhaustivité de la recherche	67
1.2.1	Fusion de partiels lors d'une convergence	67
1.2.2	Détection de partiels oubliés	68
1.3	Concaténation de l'ensemble des résultats obtenus pour un fragment sonore	69

2 Réduction de la complexité des opérations en vue d'une implantation sur machine	71
2.1 Restriction de la méthode à une zone fréquentielle	71
2.1.1 Théorie et mise en œuvre	71
2.1.2 Le problème du filtrage des fréquences	72
2.2 Découpage de l'ensemble des fréquences en sous-ensembles selon un critère	72
2.2.1 Création des groupes selon le critère de "l'influence maximale"	73
2.2.2 Création des groupes selon le critère de "rupture énergétique"	74
3 Vers une chaîne complète d'analyse	76
3.1 Chaîne algorithmique de mesure des paramètres additifs d'un fragment sonore	76
3.2 Structure complète de la chaîne algorithmique	76
3.3 Proposition d'une méthode d'affinage de la résolution temporelle des enveloppes d'amplitudes	78
Conclusion	81
IV Annexes	83
A Caractéristiques des fenêtres de la première famille	85
B Caractéristiques des fenêtres de la seconde famille, à la limite de l'existence de l'inflexion	86
C Caractéristiques des fenêtres de la seconde famille, approchant au mieux une inflexion fixée	92
Bibliographie	99

Liste des tableaux

3.1	Tableau des fenêtres utilisées pour les tests comparatifs. . . .	52
3.2	Classement des fenêtres pour chaque partiel - Test 1	54
3.3	Classement global des fenêtres - Test 1	55
3.4	Résultats moyens observés sur l'estimation des amplitudes pour deux partiels arbitrairement proches - Test 4	58
1.1	Tableau de caractérisation de l'influence d'un partiel non dé- tecté sur l'estimation d'un autre, en fonction du rapport d'am- plitudes et de l'écart fréquentiel	67
A.1	Tableau des paramètres de la première famille de fenêtre	85
B.1	Tableau des paramètres b limites - Seconde famille de fenêtres à la limite d'existence d'inflexions	87
B.2	Tableau des bandes à -3dB (en bins) - Seconde famille de fenêtres à la limite d'existence d'inflexions	88
B.3	Tableau des bandes à -6dB (en bins) - Seconde famille de fenêtres à la limite d'existence d'inflexions	89
B.4	Tableau des bandes équivalentes de bruit (en bins) - Seconde famille de fenêtres à la limite d'existence d'inflexions	90
B.5	Tableau des taux de durée où se concentre 99% de l'énergie (en %) - Seconde famille de fenêtres à la limite d'existence d'inflexions	91
C.1	Tableau des paramètres b limites - Seconde famille de fenêtres approchant au mieux une inflexion fixée	93
C.2	Tableau des bandes à -3dB (en bins) - Seconde famille de fenêtres approchant au mieux une inflexion fixée	94
C.3	Tableau des bandes à -6dB (en bins) - Seconde famille de fenêtres approchant au mieux une inflexion fixée	95
C.4	Tableau des bandes équivalentes de bruit (en bins) - Seconde famille de fenêtres approchant au mieux une inflexion fixée	96
C.5	Tableau des taux de durée où se concentre 99% de l'énergie (en %) - Seconde famille de fenêtres approchant au mieux une inflexion fixée	97

Table des figures

3.1	Spectre à analyser - Cas d'un son simple, fenêtre de Blackman	33
3.2	Convergence de trois partiels - Cas d'un son simple, fenêtre de Blackman	33
3.3	Convergence des amplitudes - Cas d'un son simple, fenêtre de Blackman	34
1.1	Spectre à analyser - Cas d'un son simple, fenêtre de Hamming	40
1.2	Convergence des fréquences et des amplitudes - Cas d'un son simple, fenêtre de Hamming	41
1.3	Spectre à analyser - Cas d'un son simple, fenêtre rectangulaire	42
1.4	Mauvaise évolution des fréquences - Cas d'un son simple, fenêtre rectangulaire	42
1.5	Mauvaise évolution des amplitudes - Cas d'un son simple, fenêtre rectangulaire	43
2.1	Première famille de fenêtres sans lobes	46
2.2	Seconde famille de fenêtres sans lobes	48
2.3	Spectre à analyser - Cas d'un son simple, fenêtre de la famille 1 pour $a=2$ et $b=0$	49
2.4	Convergence des fréquences et des amplitudes - Cas d'un son simple, fenêtre de la famille 1 pour $a=2$ et $b=0$	50
3.1	Forme de la chaîne algorithmique obtenue pour la mesure des paramètres additifs	77

Introduction

Fabrication d'un outil de synthèse

La synthèse additive

Parmi les techniques de synthèse sonore utilisées dans le domaine musical, la synthèse additive occupe une place particulière. Elle regroupe des qualités indéniables tant pour l'auditeur, le compositeur que le scientifique. En effet, le modèle additif s'attache à reconstituer un son par une superposition de "sons purs" (qu'on appelle *partiels*), ondes sinusoïdales dont la fréquence et l'amplitude peuvent varier au cours du temps. Et, cette représentation qui, a priori, semble assez simple, offre une panoplie de propriétés intéressantes :

- Elle est tout d'abord particulièrement adaptée à notre perception puisque ses paramètres nous sont psychologiquement et directement accessibles. L'amplitude correspond au simple concept de "plus ou moins fort" ; et la fréquence renvoie à celui de "plus ou moins aigu".
- Les possibilités et transformations sonores qu'elle offre sont grandes : transposer sans changer la durée, changer l'évolution temporelle en conservant la hauteur, "découper" un son en plusieurs, élaborer un "morphing" d'un son vers un autre, ou encore de modifier l'inharmonicité. Et ceci, avec une très grande souplesse et facilité dans le contrôle du matériau sonore, comme le logiciel *DIPHONE* développé par l'IR-CAM peut déjà le montrer.
- La synthèse est très facilement implantable sur machine puisqu'elle ne nécessite qu'une table trigonométrique, et les opérateurs $+$ et \times . De plus, des algorithmes rapides existent [DR92].
- Mais surtout, elle permet par l'*analyse* d'obtenir les paramètres de sons réels enregistrés et donc, de les travailler.

Il apparaît dès lors, toute l'utilité de cette synthèse, et l'importance que lui donne les techniques d'analyse *analyse* correspondantes, pour apporter à l'art musical de nouvelles sonorités maniables...

L'analyse : un problème temps-fréquence

L'analyse consiste à déterminer l'ensemble des paramètres nécessaires à la synthèse. En l'occurrence, il s'agit de détecter les composantes fréquentielles pures du son (i.e. les partiels) au cours du temps et d'isoler le bruit résiduel en caractérisant sa forme. Aussi, l'analyse idéale consisterait à obtenir tous les paramètres instantanés du son analysé, permettant au modèle de s'en approcher au mieux.

Cependant, dans la plupart des cas, les propriétés mêmes du son interdisent d'approcher la ponctualité temporelle des paramètres: en effet, le son issu d'un instrument est souvent généré par un train d'impulsions filtrées par un résonateur. Par conséquent, l'information "minimale" ne peut correspondre à un intervalle temporel inférieur à une période. De plus, l'analyse spectrale classique, actuellement utilisée, nécessite un grand nombre de points du signal, de l'ordre de 4 périodes de la fréquence minimale contenue dans le son. On obtient donc des paramètres "moyens" associés à de grandes durées, ce qui est néfaste à la bonne qualité du modèle du signal sonore.

D'autre part, l'estimation des amplitudes correspond à la mesure d'un point du spectre localement interpolé, mesure qui ne tient pas compte de l'influence des partiels voisins et qui est peu résistante au bruit.

Dans le but d'améliorer essentiellement ce dernier point, une méthode analytique a été envisagée. Et une première étude effectuée par Laurent Tromp dans le cadre d'un DEA [Tro95] a montré l'intérêt et les apports bénéfiques qu'elle pouvait rendre quant à l'estimation des paramètres.

Vers une amélioration des paramètres d'analyse

Proposition d'une méthode analytique

Afin d'améliorer la caractérisation du signal, une solution analytique consistant à minimiser l'énergie d'un résidu dans le domaine fréquentiel a été proposée. Le principe repose sur une modélisation du spectre (et non plus uniquement un modèle temporel) qui permet d'estimer les paramètres additifs à partir des informations prélevées dans l'ensemble du spectre plutôt que sur des points isolés. Ce procédé a deux avantages principaux :

- l'estimation des paramètres est plus fine en précision, dans le sens où les influences mutuelles des partiels sur leur représentation spectrale n'est plus un handicap. En effet, si l'on considère deux partiels et les "pics spectraux" qui les représentent, il est clair que la forme du premier pic provoque une déformation du deuxième et réciproquement. Aussi, "regarder" l'information localement, c'est mesurer un point apparemment le plus représentatif mais sujet à déformation (sans compter la présence de bruit à l'enregistrement). Au contraire, avec un

modèle spectral, c'est tout le spectre qui permet l'évaluation simultanée de chaque paramètre. Par conséquent, les influences mutuelles des partiels sont prises en compte, et la perturbation due à un bruit est amoindrie.

- il permet de réduire la taille de la fenêtre d'analyse puisqu'il n'est pas nécessaire d'avoir une aussi bonne séparation des pics ni une résolution fréquentielle aussi importante que le requiert la première méthode. L'origine de cette propriété est encore la même : on ne cherche pas à optimiser la précision de chaque point du spectre, dans le but d'estimer des points particuliers, mais uniquement à ce qu'elle soit suffisante pour que l'information portée par le spectre permette de retrouver une bonne estimation.

Et ces propriétés amènent donc une meilleure qualité des estimateurs aussi bien dans la précision que dans la "localisation temporelle".

Sujet - Évolution du problème

Suite à la mise en place de la méthode, il s'agissait d'en étudier le principe et d'en concevoir une implantation pour aboutir à une *amélioration de l'estimation des paramètres additifs par analyse spectrale sur des fenêtres temporelles très courtes*.

Aussi, au regard des travaux déjà effectués, la première étape du stage a consisté à mettre en oeuvre des outils –fonctions développées en *Matlab*– afin d'étudier le comportement de la méthode face à diverses situations sonores et de déterminer les améliorations possibles. La forme de la fenêtre temporelle de pondération s'est alors révélée être un point particulièrement sensible et important pour le bon fonctionnement de la méthode. Ce constat a donc induit un travail axé sur la recherche de fenêtres bien adaptées au problème. Après avoir construit des fenêtres appropriées, il a été possible de s'engager dans l'élaboration d'une chaîne algorithmique d'analyse du son, qui tenait compte des autres remarques sur le comportement.

Finalement, l'outil développé permet effectivement une estimation qui prend en compte l'influence mutuelle des partiels et qui donne une localisation temporelle des paramètres proche de la période. Aussi, parmi les améliorations à venir, on peut compter :

- une pré-analyse déterminant les zones où le son est suffisamment stable¹ et sur lesquelles on pourra appliquer une recherche de paramètres ; les

1. Des travaux sont déjà engagés dans ce sens par DIGITAL MUSIC RESEARCH GROUP [BM95].

informations sur l'évolution permettront alors de mieux guider l'analyse.

- une méthode permettant d'obtenir une meilleure résolution des enveloppes et notamment en ce qui concerne les amplitudes². En effet, on dispose de procédés permettant d'affiner les fréquences estimées et de les interpoler convenablement, en utilisant les phases. Cependant, les amplitudes ne disposent pas de tels procédés qui permettraient, par exemple, de rétablir de forts écarts de dynamique sur une faible durée, adoucis par la mauvaise localisation temporelle qui n'offre que des amplitudes moyennes.

2. il sera proposé à ce sujet une première ébauche de solution.

Première partie

Estimation des paramètres additifs par modélisation analytique du spectre à court-terme

Chapitre 1

Représentation et modélisation d'un signal sonore

1.1 Modélisation dans le domaine temporel

Soit $s(n)$ un son échantillonné à la fréquence f_e , n étant l'indice temporel. On représente alors $s(n)$ selon le modèle additif par la somme de tous les partiels $P_k(n)$ qui le composent et d'un bruit $b(n)$, que l'on caractérise par une enveloppe spectrale évoluant au cours du temps. Chaque partiel P_k est en fait une sinusoïde "lentement" modulée en amplitude et en fréquence. Aussi, on le caractérise par une fonction temporelle $a_k(n)$ associée aux amplitudes, et une autre associée aux phases instantannées $\Phi_k(n)$, que l'on peut également représenter par les fonctions fréquences $f_k(n)$ et une phase à l'origine ϕ_k [DR92].

Aussi, si le son a K partiels, on pourra l'écrire de la façon suivante :

$$\begin{aligned} s(n) &= \sum_{k=1}^K P_k(n) + b(n) \\ &= \sum_{k=1}^K a_k(n) \cos(\Phi_k(n)) + b(n) \end{aligned} \quad (1.1)$$

avec la relation phase-fréquence : $\Phi_k(n+1) = \Phi_k(n) + 2\pi \frac{f_k(n)}{f_e}$ et $\Phi_k(0) = \phi_k$

Il y a deux points important à souligner concernant ce modèle :

- Le terme de bruit $b(n)$, dans le cas de la synthèse additive, n'a pas de sens péjoratif : il n'est pas à supprimer. Il constitue la partie du son de nature aléatoire. Cela peut être le souffle pour un instrument à

vent, la frappe du marteau pour un piano, etc... Sa présence est donc nécessaire quant au réalisme du son.

- Un partiel voit, comme il est dit ci-dessus, ses caractéristiques varier au cours du temps. Ceci mène à la notion de “trajets de paramètres”, chacun étant indexé par l’indice k . La présente étude ne comprend pas cette recherche de trajet. Elle se limite à l’estimation des triplets (a_k, f_k, Φ_k) sans se soucier du partiel auquel appartient chaque triplet. La détermination de ces trajets a déjà été l’objet d’une étude (reposant sur des processus de Markov cachés, [DGR93]).

Ce modèle est bien sûr la forme idéale que l’on souhaite obtenir (K pouvant être infini). Dans la pratique, K est fini, et l’estimation des triplets n’est pas faite sur le signal à chaque échantillon. Étant donné que les paramètres additifs sont déduits du spectre d’une partie du signal, ils sont significatifs de “l’information moyenne” de cette même partie. De plus, toujours pour des raisons pratiques mais aussi parce que ces paramètres sont supposés évoluer relativement lentement dans le temps par rapport aux fréquences f_k , les évaluations sont faites à un pas, le *pas d’avancement*, bien supérieur au pas d’échantillonnage.

Dans le cas de l’analyse, la précision fréquentielle obtenue est liée au nombre d’échantillons pris dans le signal pour l’évaluation du spectre : plus on prend de points, meilleure est la résolution. Mais ceci est au détriment de la localisation temporelle des paramètres qui est d’autant meilleure que l’on prend peu de points. Ainsi, la mesure des paramètres en utilisant le spectre ne permet pas de connaître le signal avec une précision maximale en temporel et en fréquentiel simultanément. L’expression $\sigma_t^2 \sigma_f^2 \geq 4\pi$ trouvée par Gabor [Gab46], reliant les variances temporelles et fréquentielles, permet de rendre compte de ce constat. Et par conséquent, les évolutions temporelles obtenues auront une résolution plus grossière que l’échantillon original et un aspect sensiblement “lissé”.

Plus précisément, on peut montrer que le pas d’avancement est lié à la largeur de bande de la fenêtre d’analyse [RS68].

1.2 Modélisation dans le domaine fréquentiel

Soit $I = [n_0, n_0 + N - 1]_{\mathbb{Z}}$ l’intervalle temporel du signal à analyser, et $w_I(n)_{n \in \mathbb{Z}}$ la fenêtre de pondération de support I à appliquer au signal.

Afin d’éviter la lourdeur des expressions calculées ci-dessous, on remplace l’origine du signal de façon à ce que l’expression du spectre $S_I(f)$ de $s(n)$ ait la forme suivante :

$$S_I(f) = \sum_{n \in I} w_I(n) s(n) \exp(-2\pi i \frac{f}{f_e} n) \quad (1.2)$$

Cette simplification revient, pour les paramètres qui nous intéressent, à ne plus considérer les phases à l'origine ϕ_k mais les phases correspondant au centre de I .

Considérons alors le spectre S_I d'un son $s(n)$ stationnaire sur I , de paramètres localement fixes A_k et F_k , d'expression pour $n \in I$,

$$s(n) = \sum_{k=1}^K A_k \cos\left(2\pi \frac{F_k}{f_e} n + \phi_k\right) + b(n) :$$

$$\begin{aligned} S_I(f) &= \sum_{n \in I} w_I(n) \left(\sum_{k=1}^K A_k \cos\left(2\pi \frac{F_k}{f_e} n + \phi_k\right) + b(n) \right) e^{-2\pi i \frac{f}{f_e} n} \\ &= \sum_{k=1}^K A_k \left(\sum_{n \in I} w_I(n) \cos\left(2\pi \frac{F_k}{f_e} n + \phi_k\right) e^{-2\pi i \frac{f}{f_e} n} \right) \\ &\quad + \underbrace{\sum_{n \in I} w_I(n) b(n) e^{-2\pi i \frac{f}{f_e} n}}_{B_I(f)} \\ &= \sum_{k=1}^K \frac{A_k}{2} \left(\sum_{n \in I} w_I(n) \left(e^{-2\pi i \frac{f-F_k}{f_e} n + i\phi_k} + e^{-2\pi i \frac{f+F_k}{f_e} n - i\phi_k} \right) \right) \\ &\quad + B_I(f) \\ &= \sum_{k=1}^K \frac{A_k}{2} \left(e^{i\phi_k} W(f - F_k) + e^{-i\phi_k} W(f + F_k) \right) + B_I(f) \end{aligned}$$

où $W(f) = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} w(n) e^{-2\pi i \frac{f}{f_e} n}$ représente la transformée de Fourier discrète

(TFD) de la fenêtre, expression qu'on appellera par la suite "fenêtre dans le domaine fréquentiel" ou encore de façon plus succincte "fenêtre fréquentielle".

Si l'on ne retient que la partie non bruitée du signal sonore, soit

$$\hat{s}(n) = \sum_{k=1}^K A_k \cos\left(2\pi \frac{F_k}{f_e} n + \phi_k\right),$$

on peut considérer $\hat{s}(n)$ comme un estimateur du signal $s(n)$. Sa TFD $\hat{S}_I(n)$ constitue alors un estimateur du spectre, estimateur dont l'expression ne dépend que des paramètres additifs :

$$\hat{S}_I(f) = \sum_{k=1}^K \frac{A_k}{2} \left(e^{i\phi_k} W(f - F_k) + e^{-i\phi_k} W(f + F_k) \right) \quad (1.3)$$

On pourra donc modéliser directement le spectre de la partie non bruitée du son, à la condition que les paramètres de fréquences et d'amplitudes soient

quasiment constants sur l'intervalle temporel correspondant. Ces propriétés seront généralement vérifiées sur des intervalles assez courts.

Chapitre 2

Description de la méthode

Avant d’expliciter la méthode, il convient de rappeler que dans la méthode usuelle permettant d’obtenir les paramètres additifs, on effectue une transformée de Fourier à court-terme, puis on détermine la valeur de ces paramètres à partir de la forme de pics spectraux observés. Comme il l’a déjà été signalé, cette méthode, géométrique et locale, ne donne qu’une information ponctuelle. Or, puisque l’on dispose dorénavant d’un modèle additif du spectre, il est intéressant de voir si l’on peut obtenir les paramètres en cherchant la meilleure adéquation entre le spectre observé et le modèle dont on dispose ; on considèrerait alors l’ensemble des points du spectre. (Du point de vue de “l’information”, cette proposition semble particulièrement intéressante.)

2.1 Théorie et estimation

Dans cette partie, on s’intéresse à déterminer à tour de rôle les amplitudes complexes et les fréquences les plus vraisemblables par rapport au spectre. On note alors, de façon générique le représentant de ces grandeurs \mathbf{p} (il symbolise donc soit les fréquences soit les amplitudes complexes). Et, on va chercher à maximiser le critère de vraisemblance de \mathbf{p} . Pour cela, on considère dans un premier temps que le bruit $b(n)$ est gaussien centré. On aboutit alors au critère suivant [Tro95] :

si $\hat{\mathbf{p}}$ représente les paramètres répondant au maximum de vraisemblance, alors

$$\frac{\partial \|S_I(f) - \hat{S}_I(f)\|^2}{\partial \mathbf{p}}(\hat{\mathbf{p}}) = \mathbf{0} \quad (2.1)$$

On peut généraliser ce résultat au-delà des hypothèses faites sur la caractérisation du bruit $b(n)$: le théorème de la “limite centrale” pouvant s’appliquer pour la caractérisation de $B_I(f)$, cette variable aléatoire a une loi presque gaussienne, et c’est justement le caractère gaussien de celle-ci

qui est nécessaire à l'obtention du critère. Ainsi, on pourra considérer ce critère énergétique qui exprime la minimisation de l'énergie résiduelle entre le spectre observé et modélisé, comme valide de façon générale. Il suffit que $b(n)$ soit centré...

Finalement, l'obtention de meilleurs paramètres, au sens où ils seront les plus vraisemblables, revient à minimiser la puissance résiduelle, autrement dit, la puissance du bruit. On va donc chercher \mathbf{p} tel que

$$\|B_I\|^2 = \|S_I - \hat{S}_I\|^2 \quad (2.2)$$

soit minimal.

2.1.1 Estimation des amplitudes et des phases

On cherche à formaliser l'estimation des amplitudes et des phases selon le critère (2.1). L'expression du modèle spectral (1.3) a une forme particulièrement intéressante puisqu'elle est linéaire en \mathbf{p} . En effet, on a d'après (1.3)¹:

$$\begin{aligned} \hat{S}(f) &= \sum_{k=1}^K \frac{A_k}{2} \left(e^{i\phi_k} W(f - F_k) + e^{-i\phi_k} W(f + F_k) \right) \\ &= \sum_{k=1}^K \frac{A_k}{2} \cos(\phi_k) (W(f - F_k) + W(f + F_k)) \\ &\quad + i \sum_{k=1}^K \frac{A_k}{2} \sin(\phi_k) (W(f - F_k) - W(f + F_k)) \\ &= \sum_{k=1}^{2K} p_k H_k(f) \end{aligned} \quad (2.3)$$

avec pour $k \in [1, K]_{\mathbb{N}}$:

$$\begin{cases} H_k(f) &= W(f - F_k) + W(f + F_k) \\ H_{K+k}(f) &= i(W(f - F_k) - W(f + F_k)) \end{cases}$$

et

$$\begin{cases} p_k &= \frac{A_k}{2} \cos \phi_k \\ p_{K+k} &= \frac{A_k}{2} \sin \phi_k \end{cases}$$

1. afin d'alléger l'écriture, on omettra le symbole I , témoin de la partie temporelle analysée

Dans la pratique, le spectre sera mesuré en chaque point f_m de $\mathbb{F}_M = \{\frac{m}{M}f_e / m \in [-\frac{M}{2}, \frac{M}{2} - 1]_{\mathbb{Z}}\}$. On déduit de (2.3) la relation matricielle suivante :

$$\begin{pmatrix} \hat{S}_{f_1} \\ \hat{S}_{f_2} \\ \vdots \\ \hat{S}_{f_M} \end{pmatrix} = \underbrace{\begin{pmatrix} H_1(f_1) & \cdots & H_K(f_1) \\ H_1(f_2) & \cdots & H_K(f_2) \\ \vdots & \ddots & \vdots \\ H_1(f_M) & \cdots & H_K(f_M) \end{pmatrix}}_{\text{partie réelle}} \underbrace{\begin{pmatrix} H_{K+1}(f_1) & \cdots & H_{2K}(f_1) \\ H_{K+1}(f_2) & \cdots & H_{2K}(f_2) \\ \vdots & \ddots & \vdots \\ H_{K+1}(f_M) & \cdots & H_{2K}(f_M) \end{pmatrix}}_{\text{partie imaginaire}} \begin{pmatrix} p_1 \\ \vdots \\ p_K \\ \text{---} \\ p_{K+1} \\ \vdots \\ p_{2K} \end{pmatrix}$$

En notant le vecteur colonne du modèle spectral échantillonné $\hat{\underline{S}}$, la matrice \mathcal{H} , et le vecteur colonne des paramètres \underline{p} , on a :

$$\hat{\underline{S}} = \mathcal{H} \cdot \underline{p} \quad (2.4)$$

Ainsi, on dispose d'un critère (2.1) pour un modèle linéaire linéaire en les paramètres cherchés. On peut alors optimiser le critère par une méthode des *moindres carrés*² :

- exprimons le critère (2.1) en utilisant l'expression (2.4). (2.1) s'écrit alors

$$\frac{\partial (\underline{S} - \mathcal{H} \cdot \underline{p})^H \cdot (\underline{S} - \mathcal{H} \cdot \underline{p})}{\partial \underline{p}}(\hat{\underline{p}}) = \underline{0}$$

où \cdot^H représente l'hermitien³ pour un vecteur ou une matrice.

- soit: $-2\mathcal{H}^H \cdot (\underline{S} - \mathcal{H} \cdot \hat{\underline{p}}) = 0$
- d'où, si l'on suppose que le rang et la dimension de $\mathcal{H}^H \cdot \mathcal{H}$ sont identiques (autrement dit, que $\mathcal{H}^H \cdot \mathcal{H}$ est inversible), on a :

$$\hat{\underline{p}} = \underbrace{(\mathcal{H}^H \cdot \mathcal{H})^{-1}}_{\text{pseudo-inverse de } \mathcal{H}} \cdot \mathcal{H}^H \cdot \underline{S} \quad (2.5)$$

Finalement, on obtient une équation (2.5) qui permet d'obtenir les amplitudes et les phases optimales à partir de fréquences et du spectre observé. Ceci permettra de trouver de façon systématique les meilleurs paramètres à fréquences données ; en effet, l'hypothèse qui conditionne la validité de cette équation par l'existence de la *pseudo-inverse de H* ne sera plus vérifiée lorsqu'un ensemble de colonnes de H seront liées, c'est-à-dire que deux fréquences seront trop proches, phénomène étudié par la suite et que l'on évitera.

2. voir [Tro95] & [LH74]

3. i.e. le conjugué de la transposée

2.1.2 Estimation des fréquences

Cette fois-ci, on s'intéresse à l'estimation des fréquences. Cependant, le problème est ici plus délicat : le modèle spectral (1.3) n'est pas linéaire en les fréquences, comme précédemment. On va alors chercher à obtenir un système linéaire permettant de résoudre ce problème, en effectuant un développement limité du critère.

linéarisation du problème

En fait, la non-linéarité du modèle en fonction de f vient de celle de la "fenêtre fréquentielle" W (Cf. 1.3). Aussi, dans un premier temps, on s'intéresse plus particulièrement au D.L. de W .

Considérons un partiel de fréquence vraie \mathcal{F}_k , représenté par la fréquence approchée F_k . En notant $\Delta_k = \mathcal{F}_k - F_k$ l'erreur fréquentielle, on peut réécrire $W(f - \mathcal{F}_k)$ de la façon suivante :

$$W(f - \mathcal{F}_k) = W(f - F_k - \Delta_k)$$

De plus, F_k étant la fréquence estimée, il est raisonnable de penser que Δ_k est "petit", et de développer l'expression précédente autour du point $f - F_k$:

$$W(f - \mathcal{F}_k) = W(f - F_k) - W'(f - F_k) \cdot \Delta_k + o(\Delta_k^2) \quad (2.6)$$

$$\text{où} \quad o(\Delta_k^2) = \frac{W''(f - F_k)}{2!} \cdot \Delta_k^2 + \dots$$

En négligeant la participation de $o(\Delta_k^2)$ devant celle des autres termes, et en considérant les Δ_k comme le paramètre \mathbf{p} à résoudre, l'expression (2.6) va permettre d'obtenir un système linéaire qui aboutit à la résolution des fréquences. On aura donc, si l'on résume :

- un système linéaire en Δ_k soluble par une méthode semblable à la précédente,
- une contrainte à respecter :

$$|o(\Delta_k^2)| \ll |W'(f - F_k) \cdot \Delta_k| \quad (2.7)$$

- une résolution des fréquences par : $F_{\text{solution}} = F_k + \Delta_k$

Cependant, on voit déjà, avant la mise en place du système, que cette résolution sera imparfaite sauf dans le cas où $o(\Delta_k^2)$ sera rigoureusement nul. En effet, les solutions optimales que donnera la résolution du système auront une qualité d'autant meilleure que la contrainte sera respectée, ce qui s'obtient pour :

- un type de fenêtre bien adaptée ;
- un Δ_k s'approchant de zéro, soit une erreur faible.

Obs: ceci indique que la résolution des fréquences ne pourra être envisagée que de façon itérative (on arrête lorsque Δ_k est quasi-nul).

Mise en forme du système et résolution

On considère désormais \mathbf{p} comme étant les Δ_k . On note A_k les amplitudes, ϕ_k les phases, $\hat{S}(f)$ le modèle du spectre pour les fréquences bien estimées \mathcal{F}_k , et $\tilde{S}(f)$ le modèle du spectre pour les fréquences approchées F_k . On écrit alors le modèle (1.3) en supposant la contrainte parfaitement respectée, au moyen de l'expression linéaire (2.6) :

$$\begin{aligned}\hat{S}(f) &= \sum_{k=1}^K \frac{A_k}{2} \left[e^{i\phi_k} (W(f - F_k) - W'(f - F_k) \cdot \Delta_k) \right. \\ &\quad \left. + e^{-i\phi_k} (W(f + F_k) + W'(f + F_k) \cdot \Delta_k) \right] \\ &= \underbrace{\sum_{k=1}^K \frac{A_k}{2} \left(e^{i\phi_k} W(f - F_k) + e^{-i\phi_k} W(f + F_k) \right)}_{\tilde{S}(f)} \\ &\quad + \sum_{k=1}^K \frac{A_k}{2} \left(-e^{i\phi_k} W'(f - F_k) + e^{-i\phi_k} W'(f + F_k) \right) \Delta_k \quad (2.8)\end{aligned}$$

Si on prélève à nouveau les valeurs de f dans \mathcal{F}_M , on aboutit d'après cette équation à :

$$\hat{\underline{S}} = \tilde{\underline{S}} + \Omega \cdot \underline{p} \quad (2.9)$$

$$\text{où } (\Omega)_{n,k} = \frac{A_k}{2} \left(-e^{i\phi_k} W'(f_n - F_k) + e^{-i\phi_k} W'(f_n + F_k) \right)$$

Ainsi, on est ramené à un système du même type que précédemment, et on peut de nouveau optimiser le critère (2.1) par la même méthode :

- exprimons le critère (2.1) en utilisant l'expression (2.9). (2.1) s'écrit alors

$$\frac{\partial (\underline{S} - \tilde{\underline{S}} - \Omega \cdot \underline{p})^H \cdot (\underline{S} - \tilde{\underline{S}} - \Omega \cdot \underline{p})}{\partial \underline{p}}(\underline{\hat{p}}) = \underline{0}$$

- soit : $-2\Omega^H \cdot (\underline{S} - \tilde{\underline{S}} - \Omega \cdot \underline{\hat{p}}) = 0$

- d'où, si $\Omega^H \cdot \Omega$ est inversible, on a :

$$\underline{\hat{p}} = (\Omega^H \cdot \Omega)^{-1} \cdot \Omega^H \cdot (\underline{S} - \tilde{\underline{S}}) \quad (2.10)$$

Finalement, on obtient les fréquences $\underline{\mathcal{F}}$ à partir des fréquences approchées \underline{F} de la façon suivante :

$$\underline{\mathcal{F}} = \underline{F} + (\Omega^H \cdot \Omega)^{-1} \cdot \Omega^H \cdot (\underline{S} - \tilde{\underline{S}}) \quad (2.11)$$

Et, si la contrainte (2.7) est respectée, on peut donc obtenir les bonnes fréquences à partir des amplitudes et des phases, du spectre observé, et du spectre modélisé dont les fréquences sont approchées.

Cependant, les amplitudes et les phases ne seront bien estimées que si l'on connaît les bonnes fréquences (Cf. 2.1.1). Aussi, si l'on dispose du spectre observé, du spectre modélisé, les autres paramètres ne seront qu'approchés. La prévision faite ci-dessus se confirme alors : l'estimation des paramètres additifs utilisant les deux méthodes analytiques qu'on vient d'expliciter ne pourra se faire que de façon itérative, chaque itération améliorant la précision des estimateurs.

2.2 Formulation du procédé d'estimation

On sait maintenant estimer les amplitudes et les phases les plus vraisemblables à partir de fréquences données [2.1.1], et réciproquement [2.1.2]. On construit alors un algorithme qui utilise ces deux méthodes et qui, à partir de fréquences approchées (que l'on suppose présentes dans le fragment sonore analysé), va :

- estimer les amplitudes et les phases associées [2.1.1],
- estimer les valeurs des fréquences pour les affiner [2.1.2],
- reprendre la première étape et recommencer, si les fréquences n'étaient pas stables, c'est-à-dire qu'au moins l'une d'elles a "trop changé".

Dans cet algorithme, toutes les qualités que l'on recherche (prise en compte des influences mutuelles des partiels, utilisation de l'information portée par tout le spectre, réduction possible de la taille de la fenêtre d'analyse, ...) sont présentes. Cependant, cette simple architecture ne peut faire face à de nombreux problèmes, et surtout, elle n'est pas autonome : tout d'abord, elle nécessite des fréquences d'initialisation et demande *a priori* de connaître l'existence de tous les partiels présents ; de plus, étant itérative, il faut étudier les cas éventuels d'oscillations, de divergence ou même de mauvaise convergence des paramètres ; enfin, toujours à cause de cet aspect itératif, il est possible que plusieurs représentants voient leur fréquence prendre une valeur identique (ou des valeurs très proches) à une certaine étape, ce qui interdit les méthodes précédemment décrites puisque la famille $(W(f - F_k))_{k \in [1, K]_{\mathbb{N}}}$ sera liée et que les pseudo-matrices de \mathcal{H} et Ω ne seront pas inversibles (ou de trop mauvais conditionnement pour une bonne estimation).

Par conséquent, cette méthode ne peut être que postérieure à une analyse classique, dans le but d'améliorer la qualité des estimateurs. Et, devant la diversité et la complexité des sons à analyser, l'étude de la convergence

est difficilement réalisable de façon théorique. Aussi, afin d'obtenir un algorithme "résistant", on utilise une approche heuristique utilisant des moyens informatiques pour développer la méthode⁴, la tester, déceler les points sensibles et remédier aux défauts observés.

4. la programmation a été effectuée en MATLAB afin de se concentrer sur l'amélioration de l'algorithme plutôt que sur la mise en œuvre du code

Chapitre 3

Validation première de la méthode

Avant de s'engager dans une étude générale qui comprend des cas délicats, il est important de s'intéresser à des cas simples : cela permet de valider la méthode et l'intérêt qu'on peut lui porter ; d'autre part, cela donne une bonne illustration de l'efficacité que l'on pourra espérer en général une fois les améliorations faites. Aussi, on ne tentera pas ici d'approcher les limites de fonctionnement, ni de "piéger" l'algorithme.

On s'intéresse d'abord à la qualité et à la vitesse de la convergence. On prend donc pour cette première étude un son stationnaire, avec peu de partiels, sans bruit. Le son analysé ci-dessous est défini par :

$$s(n) = \sum_{k=1}^3 \cos(2\pi \frac{f_k}{f_e} n) \quad (3.1)$$

où $f_e = 44100\text{Hz}$ et $(f_k) = (440, 1400, 4000)$

La fenêtre d'analyse utilisée est la fenêtre de Blackman standard [Har78]. On prend alors 100 points du son synthétisé ce qui correspond à une période de la plus basse fréquence. En effectuant une transformée de Fourier sur 1024 points, on obtient le spectre à analyser (fig. 3.1). Et, comme le montrent les marqueurs en pointillés qui indiquent la position des partiels, il est clair qu'une analyse classique serait mauvaise : les deux premiers pics sont noyés l'un dans l'autre et les maxima ne correspondent pas aux fréquences désirées.

Pour l'analyse on prend pour fréquences initiales, les fréquences du signal auxquelles on ajoute 100Hz (ce qui constitue respectivement une erreur fréquentielle de 22.7%, 7.14% et 2.5% sur la fréquence la plus basse. On stoppe la convergence lorsque que $\forall k, \frac{f_k^{(i)} - f_k^{(i-1)}}{f_k^{(i-1)}} < S$ avec $S = 0.0064\%$ ce qui est inférieur au 1/200^{ième} de comma.

La limite est atteinte à la 19^{ième} itération et les résultats obtenus sont très satisfaisants : $\max\{\frac{|f_k - f_k^{(19)}|}{f_k}\} < 1.1 \cdot 10^{-4}$ et $\max\{|1 - a_k^{(19)}|\} < 6.5 \cdot 10^{-5}$.

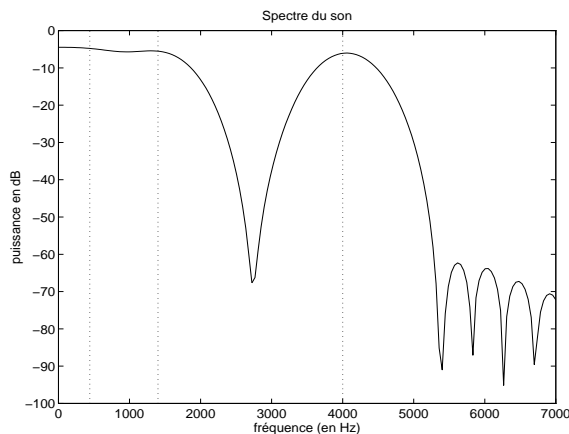


FIG. 3.1 – *Spectre à analyser - Cas d'un son simple, fenêtre de Blackman*

D'autre part, la convergence ne semble pas "difficile", à la vue de sa régularité et de sa monotonie (Cf. fig.3.2 & fig.3.3).

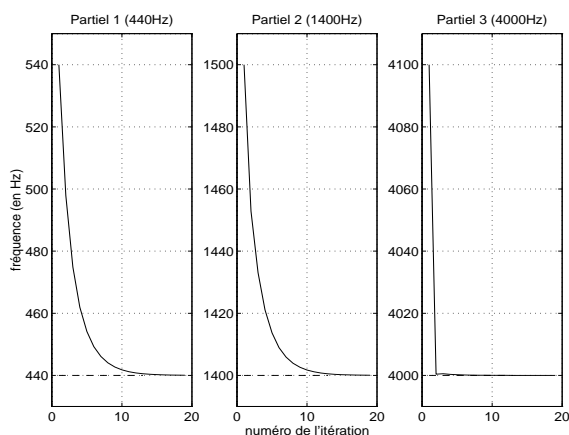


FIG. 3.2 – *Convergence de trois partiels - Cas d'un son simple, fenêtre de Blackman*

Finalement, cette première observation est encourageante et montre les possibilités que peut offrir la méthode développée : l'analyse a été effectuée sur une période de signal, et, les paramètres déterminés malgré une initialisation mauvaise et un spectre non exploitable de façon géométrique.

Cependant, si dorénavant il est possible d'obtenir l'information même lorsqu'elle n'est pas directement "visible", il n'en reste pas moins que plus elle est dissimulée et plus elle est difficile à trouver : c'est ce que semble exprimer la rapidité de la convergence du troisième partiel... Ainsi, un lobe principal plus fin (qui dépend du type de fenêtre ou du nombre de points du son pris) permettrait *a priori* d'obtenir des résultats plus rapidement et

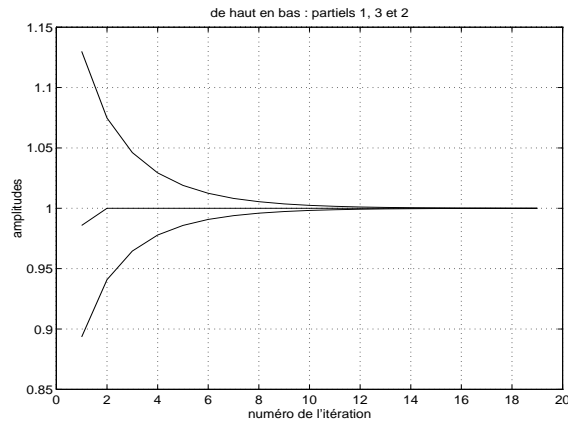


FIG. 3.3 – *Convergence des amplitudes - Cas d'un son simple, fenêtre de Blackman*

de façon plus fiable; et, la forme de la fenêtre participerait alors au bon fonctionnement de la méthode. Il semble donc qu'une étude sur ce point particulier soit utile.

Deuxième partie

Recherche de fenêtres de pondération adaptées à la méthode

Chapitre 1

Influence de la fenêtre de pondération sur la méthode

Dans la méthode utilisée, l'expression analytique de la fenêtre intervient dans tous les calculs.

Aussi, en ce qui concerne l'estimation des amplitudes et des phases, on a vu que pour des fréquences bien estimées, elle n'avait aucune influence et qu'elle n'était que "porteuse générique de l'information spectrale". Cependant, la méthode est itérative dès que l'on n'a pas accès aux bonnes fréquences dès l'initialisation ; on est alors hors de ce contexte. Qu'en est-il de la qualité de l'estimation des amplitudes et des phases : la forme de la fenêtre utilisée ne prend-elle pas part à la bonne convergence ?

Et, en ce qui concerne l'estimation des fréquences, aux conditions déjà exprimées (2.7) viennent s'ajouter les caractères approximatifs des amplitudes et des phases. Et les mêmes problèmes se posent...

Par conséquent, puisque l'expression de la fenêtre est un paramètre libre de la méthode, on s'intéresse à son étude afin de dégager des critères de bonne adaptation. Tout le travail qui s'en suivra sera alors de trouver une fenêtre combinant au mieux tous ces critères pour rendre la méthode plus robuste.

1.1 Considérations générales

Avant de se concentrer sur l'influence de la fenêtre dans chacune des deux estimations de paramètres, il est important de souligner deux points :

- tout d'abord, quelle que soit l'estimation considérée (*amplitudes & phases*, ou bien *fréquences*), le critère sur lequel elle repose est la minimisation de l'énergie entre le spectre observé et le modèle. Et c'est cela qui permet la convergence en lui donnant un sens : qualitativement, il s'agit de faire évoluer les paramètres à chaque itération dans

le but de combler au mieux les zones fortement énergétiques (les zones moins énergétiques ayant aussi leur importance). Il est alors possible que, surtout dans le cas où les fréquences initiales sont mal estimées, la présence de lobes secondaires trouble les évolutions des deux estimateurs puisque ceux-là constituent des zones localement énergétiques. Par cette remarque qui n'est encore qu'une hypothèse, il semble donc qu'il faille utiliser des fenêtres dont les lobes secondaires sont relativement faibles ;

- et, on sait que le *conditionnement* des pseudo-matrices sera d'autant plus élevé que plusieurs représentants seront semblables, autrement dit, que deux fréquences paraîtront proches. Afin d'augmenter le "pouvoir séparateur" des représentants et de repousser ainsi le problème, il faudra réduire tant que possible la largeur du lobe principal de la fenêtre. Ainsi, on pourra améliorer la résolution temporelle en conservant la même qualité des estimateurs. Il faut, à propos de ce critère, remarquer que les paramètres analysés sont significatifs de la zone temporelle où l'énergie du son pondéré est significative : si le fragment sonore pondéré par la fenêtre voit 99% de son énergie concentrée dans une durée 3 fois plus courte que la durée originale du fragment, il semble clair que les paramètres estimés seront significatifs du tiers de cette durée originale. Par conséquent, en prenant la *bande équivalente de bruit* (EQN) comme référence de la grandeur $\frac{\text{Durée du fragment}}{\text{Durée utile}}$ et la *bande passante* à -3dB (BP-3) comme référence de la largeur du lobe principal, la faiblesse de $\frac{\text{BP-3}}{\text{EQN}}$ ¹ rendra compte d'un bon "pouvoir séparateur" dans le cas auquel on s'intéresse.

1.1.1 De l'estimation des amplitudes et des phases

Dans le cas où les fréquences estimées sont assez proches des fréquences réelles, l'estimation des amplitudes complexes serait *a priori* d'autant meilleure que le lobe principal serait large : l'énergie du spectre couverte par le lobe principal du représentant serait effectivement d'autant plus forte et plus proche de l'énergie du bon représentant que le lobe principal de la fenêtre serait large.

Cependant, ceci va à l'encontre du "critère du pouvoir séparateur" et il faut alors faire un compromis. En fait, pour le moment, on peut situer la limite du "critère visant la bonne estimation des amplitudes" ainsi : le lobe principal doit couvrir la zone d'erreur fréquentielle initiale².

1. minorée par 0.89, valeur obtenue pour la fenêtre rectangulaire

2. on verra par la suite que ce critère pose rarement problème : le "critère du pouvoir séparateur" sera donc celui auquel on s'intéressera d'abord.

D'autre part, hormis la convergence, une bonne résistance au bruit sera obtenue si l'énergie de la fenêtre est assez répartie (selon les fréquences). Dans la pratique, il suffira que les zones assez "énergétiques" de la fenêtre dans le domaine fréquentiel contiennent suffisamment de points (sans nécessairement en maximiser le nombre).

Finalement, l'influence de la forme de la fenêtre dans l'estimation des amplitudes complexes n'est pas très forte en ce sens que celles-ci seront presque toujours réalisées sans que l'on s'en soucie. Et, par conséquent, les contraintes qui en découlent ne le seront également pas, dans la pratique.

1.1.2 De l'estimation des fréquences

L'estimation des fréquences est, en revanche, beaucoup plus dépendante de la forme de la fenêtre.

En effet, cette estimation repose sur un développement limité des représentants, en chaque point (2.11). Aussi, lorsqu'une fréquence initiale se trouve sur un lobe du spectre qui ne correspond pas à un partiel (mais qui est issu de lobes secondaires par ex.), le banc de dérivées prises en compte va provoquer une convergence sur ce "mauvais pic" : la dérivée positive puis négative va guider l'estimation fréquentielle sur ce lobe ; et l'estimation de l'amplitude complexe allant dans le même sens, il y aura stabilisation et donc convergence. Dans le cas où les valeurs seraient cependant très faibles, trop faibles pour bien stabiliser la convergence, cette mauvaise convergence ne sera que "locale" ; autrement dit, elle n'aura qu'un rôle de frein ou sera source d'instabilité en créant des oscillations.

Par conséquent, l'existence-même de lobes secondaires engendre sur tout l'axe fréquentiel des zones de mauvaise convergence ou de "frein", zones sur lesquelles la dérivée est négative, s'annule puis est positive.

De plus, dans la majorité des cas, l'oscillation de la dérivée accroît en quelque sorte la non-linéarité de celle-ci et éloigne donc de la contrainte sur les ordres supérieurs du développement limité (2.7).

Finalement, il semble que les lobes secondaires troublent la bonne convergence des estimateurs par leur simple existence, ce qui corrobore fortement l'hypothèse faite ci-dessus.

1.2 Étude de la convergence en utilisant les fenêtres classiques

Afin d'illustrer les remarques explicitées ci-dessus, on reprend le son inharmonique simple utilisé page 32. Cependant, on change le nombre de points de signal prélevé : on prend 400 points soit 4 périodes pour avoir des conditions similaires à celles d'une analyse "classique". D'autre part,

on prend pour fréquences initiales les fréquences de synthèse moins 100Hz, et pour le calcul du spectre, 2048 points. Ces conditions sont donc *a priori* meilleures que les précédentes du point de vue de l'information et de l'aspect géométrique du spectre...

On va alors, contre cet *a priori*, montrer que les remarques faites ci-dessus sont justes. Pour cela, on va analyser le son avec des fenêtres différentes.

1.2.1 Cas d'une bonne convergence

On utilise ici la fenêtre de Hamming. Les caractéristiques de cette fenêtre, auxquelles on s'intéresse sont les suivantes :

- le premier lobe secondaire a une puissance 43dB au-dessous du lobe principal, ce qui est moins bon que pour la fenêtre de Blackman précédemment utilisée (-58dB).
- sur l'axe des fréquences, le lobe principal sera environ 5.2 fois plus étroit que pour le test précédent³ (ceci n'étant qu'un ordre de grandeur).

On détermine le spectre (Cf. fig. 1.1) et on constate effectivement que les "pics" sont bien séparés et visibles.

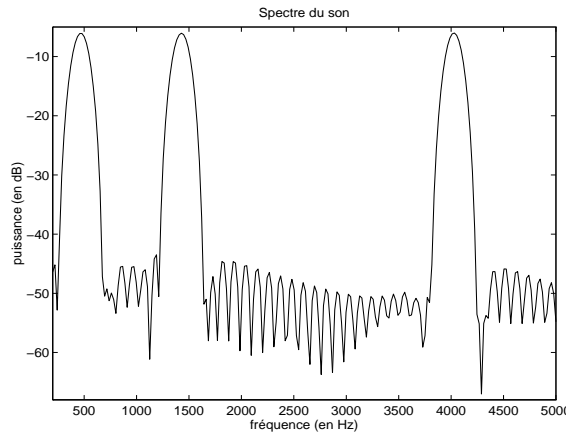


FIG. 1.1 – *Spectre à analyser - Cas d'un son simple, fenêtre de Hamming*

L'analyse est alors très satisfaisante: les paramètres sont bien trouvés comme précédemment et ceci en seulement 4 itérations (Cf. fig. 1.2).

Il semble alors, comme précédemment, que la finesse du lobe principal ajoute à l'efficacité de la méthode en réduisant sensiblement le nombre d'itérations utiles. Cependant, on remarque que les fréquences initiales font

³. $BP-3\{\text{Blackman}\}/\text{NbPoints1} \cdot \text{NbPoints2}/BP-3\{\text{Hamming}\} = 1.68/100 \cdot 400/1.3 = 5.1692$

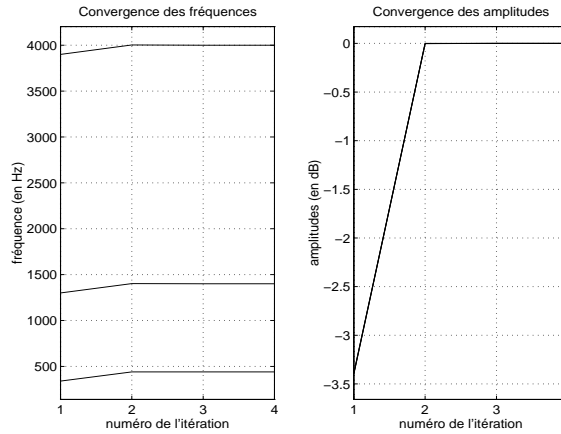


FIG. 1.2 – *Convergence des fréquences et des amplitudes - Cas d'un son simple, fenêtre de Hamming*

partie de “lobes signifiants” : le compromis “ finesse du lobe principal”, “appartenance des fréquences initiales aux lobes principaux” est en fait réalisé.

1.2.2 Cas d'une mauvaise convergence

On utilise ici la fenêtre rectangulaire. Les caractéristiques de cette fenêtre sont les suivantes :

- le premier lobe secondaire a une puissance 13dB au-dessous du lobe principal, ce qui est bien pire que pour la fenêtre de Blackman (-58dB).
- sur l'axe des fréquences, le lobe principal sera environ 7.6 fois plus étroit que pour le test précédent⁴.

On détermine le spectre (Cf. fig. 1.3) et on constate effectivement que les “pics” sont encore bien plus séparés et visibles.

Cependant, l'analyse est très mauvaise et il n'y a pas de convergence globale (Cf. fig. 1.4) :

- le 3^{ième} partiel tend vers la bonne valeur après quelques mouvements aux apparences chaotiques ;
- le 1^{er} converge de la même façon, mais tend vers une valeur apparemment quelconque ;
- tandis que le 2^{ième} ne converge pas, il évolue de façon chaotique. On peut cependant deviner des zones où, éventuellement, il y aurait “stabilité passagère”⁵.

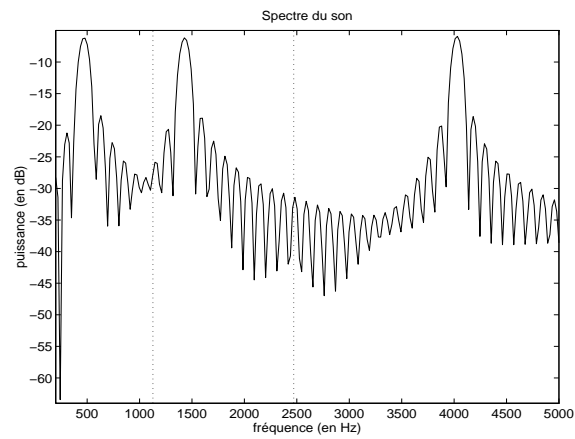


FIG. 1.3 – Spectre à analyser - Cas d'un son simple, fenêtre rectangulaire

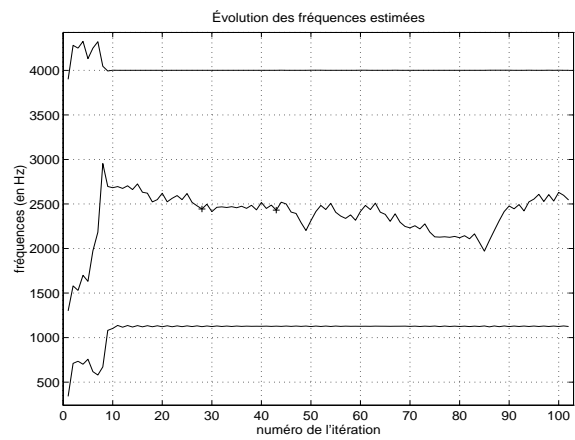


FIG. 1.4 – Mauvaise évolution des fréquences - Cas d'un son simple, fenêtre rectangulaire

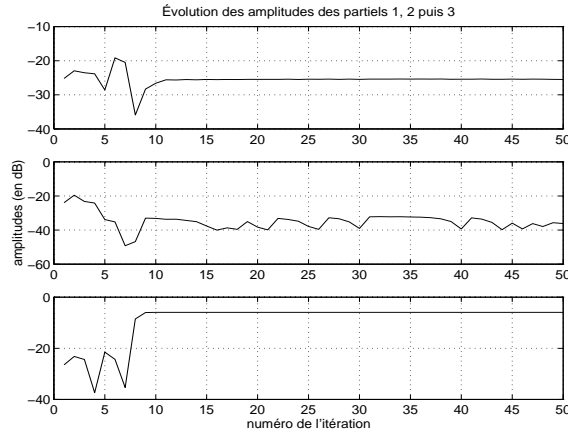


FIG. 1.5 – *Mauvaise évolution des amplitudes - Cas d'un son simple, fenêtre rectangulaire*

Or, on peut remarquer que les fréquences initiales sont hors des lobes principaux, ce qui a perturbé l'estimation des amplitudes dès le départ. D'autre part, les zones de “pseudo-stabilité” ou de mauvaise convergence sont associées aux lobes secondaires comme l'illustrent les marqueurs ajoutés à la figure représentant le spectre (fig. 1.3) : le marqueur de gauche correspond à la fréquence sur laquelle converge le 1^{er} partiel, et celui de droite correspond à la zone “pseudo-stable” délimitée sur la figure 1.5 par deux symboles +. On pourrait encore vérifier ce phénomène sur le 2^{ième} partiel, pour les itérations allant de 74 à 81.

Par conséquent, le fait que les fréquences initiales soient trop éloignées du lobe principal cause une mauvaise estimation des amplitudes. Ceci empêche (aussi par la proximité des lobes secondaires) une bonne correction des fréquences. On obtient ainsi un comportement d'apparence chaotique, pour lequel on observe des stabilisations passagères plus ou moins prononcées sur les lobes secondaires. Il faut donc absolument éviter ces situations, d'autant plus qu'il s'agit d'un son simple.

4. $BP-3\{\text{Blackman}\}/\text{NbPoints1} \cdot \text{NbPoints2}/BP-3\{\text{Rectangle}\} = 1,68/100 \cdot 400/0,89 = 7,551$

5. une de ces zones est délimitée par deux symboles + sur la figure 1.4

Chapitre 2

Construction de fenêtres adaptées

À la suite des observations précédentes, il semble donc que la meilleure façon de pallier les réactions chaotiques de la méthode soit d'utiliser des fenêtres sans lobes secondaires : c'est en fait la forme de la fenêtre dans le domaine fréquentiel qui importe ; affaiblir les lobes secondaires a tendance à élargir le lobe principal sans pour autant remédier définitivement au problème.

On compte parmi les fenêtres "standards" une seule famille de fenêtres sans lobes [Gab46], les fenêtres de Hanning-Poisson de formule :

$$w(n) = \frac{1}{2} \left[1 + \cos\left(\pi \frac{n}{N/2}\right) \right] \cdot e^{-\alpha \frac{|n|}{N/2}} \quad (2.1)$$

où $n \in \left[-\frac{N}{2}, \frac{N}{2} \right]_{\mathbb{Z}}$

Pour cette famille, l'absence des lobes secondaires est obtenue pour $\alpha \geq 2$. Et, bien qu'elles conviennent à la situation, ces fenêtres ne sont pas égales en qualité : la finesse du lobe principal, comme la rapide décroissance, sont optimales pour la même valeur $\alpha = 2$. Aussi, il existe une meilleure fenêtre de la famille dont les caractéristiques sont les suivantes :

- EQN = 2,02 ;
- BP-3 = 1,87 ;
- $\frac{\text{BP-3}}{\text{EQN}} = 0,926$.

Puisqu'on ne dispose finalement que d'une seule fenêtre sans lobes secondaires, une recherche d'autres familles répondant à ce critère semble particulièrement utile : on peut espérer obtenir d'abord de meilleurs paramètres, mais aussi une panoplie de caractéristiques différentes qui permettront un choix sur d'autres critères.

2.1 Première famille

La construction de cette famille n'a pas été effectuée de façon purement théorique du fait de la complexité du problème. Elle n'est cependant pas complètement hasardeuse : on a utilisé des propriétés bien connues sur la convolution, des remarques sur certaines formes de fenêtres, des méthodes numériques pour la détermination des paramètres et certes, une part d'intuition.

2.1.1 Définition

On sait que la transformée de Fourier d'une gaussienne est une gaussienne. Cette forme est à la fois positive et sans lobes secondaires. Mais cette fenêtre est de support infini et ne peut donc pas être directement utilisée (sur machine).

Considérons d'autre part la fenêtre triangulaire, qui est bien de support borné. Puisqu'elle est l'autoconvolution d'une fenêtre rectangulaire et que la *FFT* est un morphisme qui transforme l'opérateur “*” en l'opérateur “.”, elle a pour transformée une fenêtre positive (carré du noyau de Dirichlet)¹.

La multiplication des deux fenêtres précédentes est de support borné, et donnera en fréquentiel, la convolution de leur transformée respective, fenêtre nécessairement positive. Si la “variance” de la gaussienne est bien adaptée (suffisamment faible), le “lissage” qu'opère la convolution effacera les lobes secondaires. Cependant, la variance est le seul paramètre libre, et on obtiendra alors comme précédemment la meilleure fenêtre pour la valeur maximale des “variances solutions”.

Aussi, plutôt que de prendre une fenêtre triangulaire, on prend “une puissance” d'un triangle. Si l'on note cette puissance a , on peut supposer que pour a proche de 1, la distorsion étant faible, on pourra déterminer des variances validant le critère de non-existence de lobes secondaires. Les variances solutions les plus fortes donneront alors une meilleure fenêtre pour chaque a .

On définit alors une famille de fenêtres par la formule suivante :

$$w(n) = \left(1 - \frac{|n|}{N/2}\right)^a \cdot e^{-b \cdot \left(\frac{|n|}{N/2}\right)^2} \quad \text{où } n \in \left[-\frac{N}{2}, \frac{N}{2}\right]_{\mathbb{Z}} \quad (2.2)$$

2.1.2 Caractéristiques

On recherche d'abord les meilleurs paramètres. À a fixé, on détermine par dichotomie la meilleure valeur b correspondant à la limite de l'existence

1. on pourrait en fait considérer toute fenêtre temporelle de support borné, issue d'une autoconvolution. On prend ici le triangle car c'est l'autoconvolution de la fenêtre ayant le lobe principal le plus étroit, caractéristique que l'on recherche.

de lobes secondaires. Ainsi, pour chacun des a , on obtient une fenêtre à l'aspect lisse et dont le lobe est le plus fin possible.

On obtient alors une famille de fenêtres exploitables dont quelques représentants² sont sur la figure 2.1.

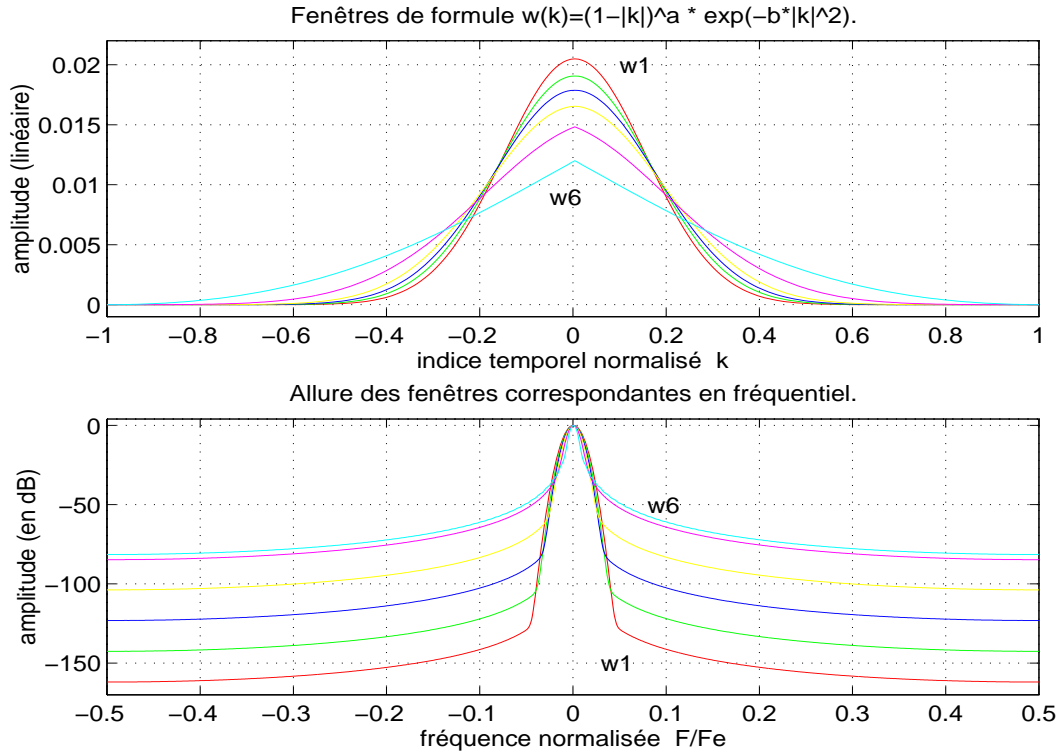


FIG. 2.1 – Première famille de fenêtres sans lobes

Et comme on peut le voir, contrairement aux fenêtres de Hanning-Poisson, le lobe le plus fin (w6) et le niveau le plus bas (w1) ne sont pas obtenus simultanément : on pourra donc sélectionner la fenêtre en fonction du critère qui sera le plus adapté. On peut remarquer également l'apparition d'une inflexion d'autant plus prononcée que le lobe est fin³.

On obtient d'autre part des caractéristiques intéressantes (cf. annexe A). Prenons par exemple la fenêtre w6 de paramètres $(a, b) = (1.8, 0.92)$:

- EQN = 1.88 ;
- BP-3 = 1.74 ;
- $\frac{\text{BP-3}}{\text{EQN}} = 0.925$.

2. w1 est obtenue pour le couple $(a, b) = (10^{-4}, 21, 6)$ et $(a, b) = (1.8, 0.92)$.

3. elle est totalement inexistante pour $a \leq 1$ et ne se prononce sensiblement que pour $a \geq 1.8$.

On peut également citer la fenêtre “limite” obtenue pour $(a, b) = (2, 0)$ (soit un “triangle au carré”) pour laquelle on a $BP-3 = 1.67$. On remarquera que cette valeur est même un peu meilleure que ce qu’on obtient pour la fenêtre de Blackman (1.68), souvent utilisée pour l’analyse.

Il semble donc, avant tout test, que la famille de fenêtres que l’on vient de déterminer pourra donner satisfaction.

2.2 Seconde famille

Pour cette famille, on s’intéresse à minimiser l’effet, connu sous le terme anglo-saxon “spectral leakage”, qui interdit aux périodes dont la durée du fragment sonore n’est pas multiple, de se projeter orthogonalement sur une unique sinusoïde. On pourrait penser que si l’on calcule la fenêtre temporelle de façon directe, puis la fenêtre fréquentielle par transformée de Fourier, plutôt qu’au moyen de l’expression analytique correspondante, ce phénomène intervenant sur les représentants même, il serait déjà pris en compte... Cependant, la phase intervient sensiblement sur la forme obtenue. Aussi, le meilleur moyen est bien de supprimer cet effet qui peut ajouter aux erreurs d’estimation.

2.2.1 Définition

Afin de supprimer l’effet de “spectral leakage”, on rend le signal pondéré C^∞ aux extrémités. On cherche donc à annuler le signal ainsi que toutes ses dérivées aux bornes du support. Finalement, on cherche une fenêtre qui conserve la propriété “lisse” dans le domaine fréquentiel, et qui s’annule avec toutes ses dérivées aux bornes de son support.

Aussi, plutôt que d’utiliser une gaussienne comme fenêtre de “lissage” sur la triangulaire, on prend une fonction en “cloche”. On définit ainsi une nouvelle famille à trois paramètres :

$$w(n) = \left(1 - \frac{|n|}{N/2}\right)^a \cdot e^{-\frac{b}{\left(1 - \frac{|n|}{N/2}\right)^c}} \quad \text{où} \quad \begin{cases} n \in \left[-\frac{N}{2}, \frac{N}{2}\right]_{\mathbb{Z}} \\ b > 0 \\ c > 0 \end{cases} \quad (2.3)$$

2.2.2 Caractéristiques

Cette fois, on cherche les paramètres limites b , pour a et c fixés. On effectue une recherche sur deux critères distincts qui donneront lieu à deux sous-familles : pour la première sous-famille, on se place à la limite de l’existence de points d’inflexion ; pour la seconde, on approche au mieux l’inflexion de la fenêtre de la première famille pour $a = 1, 8$ (fenêtre w6).

Dans chacune de ces recherches, on remarque la relation quasi-affine entre b et a à c fixé sur le domaine étudié (tableaux B.1 & C.1). Et, l’étude de la

bande à -3dB montre que, pour chaque c , il existe un unique couple (a, b) qui la minimise. On représente sur la figure 2.2 quelques fenêtres des deux sous-familles obtenues pour de tels couples (a, b) . On mesure notamment

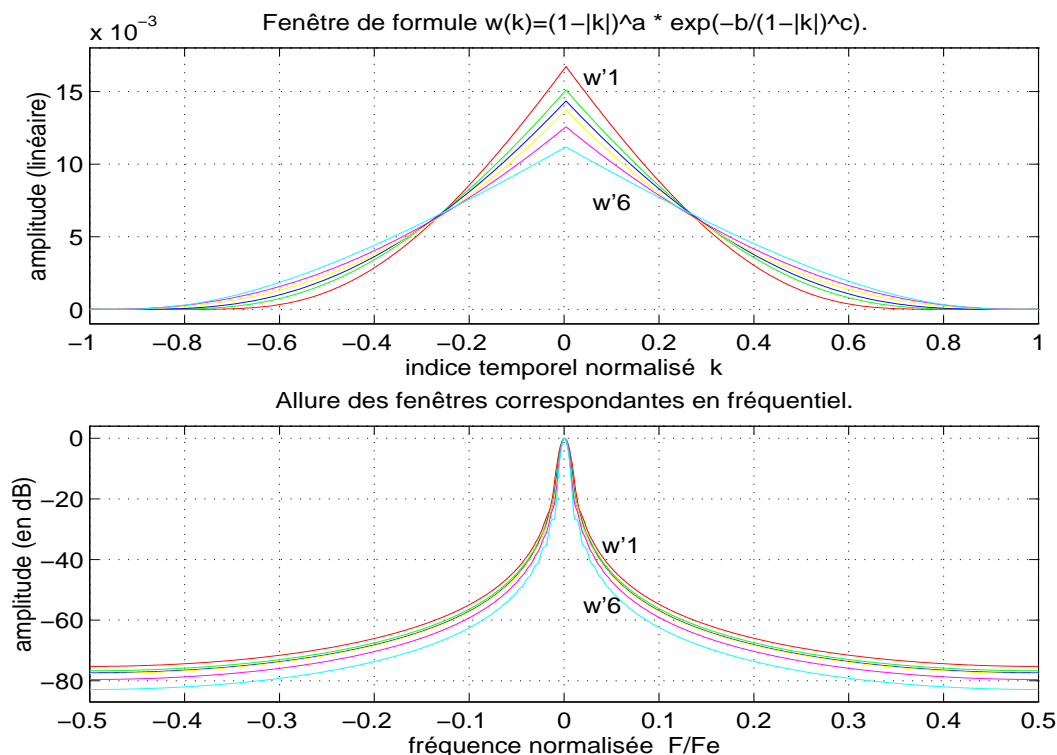


FIG. 2.2 – *Seconde famille de fenêtres sans lobes*

pour les fenêtres $w'1$ et $w'6$ les caractéristiques suivantes :

pour $w'1$, $(a, b, c) = (0.1, 2.543, 1)$, première sous-famille,

- EQN = 2.60 ;
- BP-3 = 2.41 ;
- $\frac{\text{BP-3}}{\text{EQN}} = 0.927$.

pour $w'6$, $(a, b, c) = (0.5, 2.287, 0.5)$, seconde sous-famille,

- EQN = 1.8 ;
- BP-3 = 1.69 ;
- $\frac{\text{BP-3}}{\text{EQN}} = 0.939$.

Cette famille semble elle-aussi répondre assez convenablement aux critères désirés. On dispose maintenant d'une panoplie intéressante qu'il s'agit alors de tester.

2.3 Vérification première de l'adéquation des familles au problème posé

Avant de s'engager dans un test comparatif, on vérifie que les fenêtres que l'on vient de déterminer évitent bien le comportement chaotique des estimateurs. On se place dans les mêmes conditions expérimentales que dans le chapitre 1, et en adoptant de plus, une situation défavorable : on prend la fenêtre présentant le plus d'inflexion, soit la fenêtre de la première famille pour $a = 2$ et $b = 0$; d'autre part, on se place en plus mauvaise posture à l'initialisation soit aux fréquences respectives $50Hz$, $2300Hz$ et $5000Hz$.

On représente le spectre à analyser sur la figure 2.3, spectre sur laquelle on peut effectivement remarquer la présence d'inflexions.

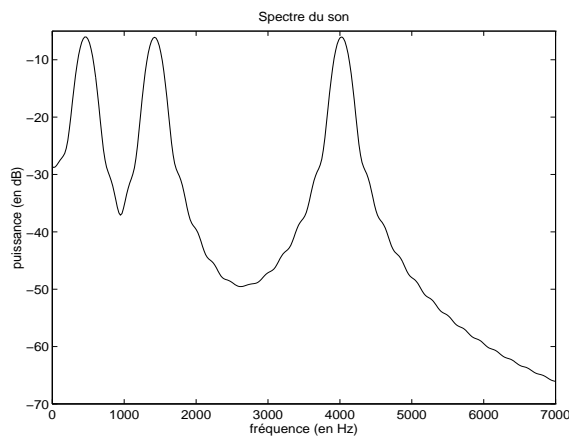


FIG. 2.3 – Spectre à analyser - Cas d'un son simple, fenêtre de la famille 1 pour $a=2$ et $b=0$

L'analyse effectuée est très concluante : comme on peut le voir sur la figure 2.4, la convergence est monotone pour chacun des estimateurs et on ne peut déceler aucune forme de comportement chaotique.

Aussi, ce test donnant une première confirmation de la propriété cherchée, on peut s'engager dans des bancs de tests comparatifs qui permettront de mieux caractériser l'adaptation des différentes fenêtres à la méthode.

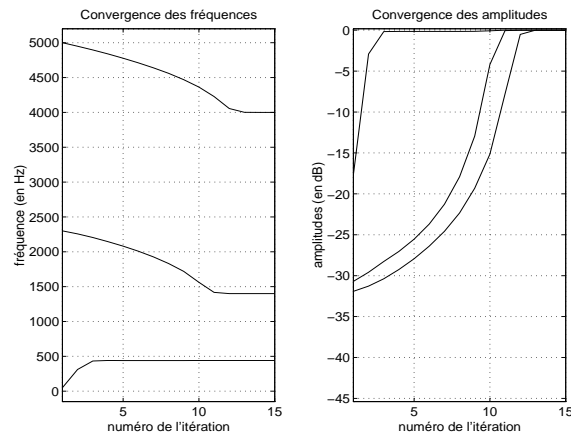


FIG. 2.4 – *Convergence des fréquences et des amplitudes - Cas d'un son simple, fenêtre de la famille 1 pour $a=2$ et $b=0$*

Chapitre 3

Tests comparatifs de fenêtres sur diverses situations sonores

Dans ce chapitre, on s'attache à mesurer les performances des fenêtres sur plusieurs sons parfaitement connus et dans des conditions équivalentes : l'initialisation est la même pour toutes les fenêtres, et on s'intéresse à avoir la même zone énergétique utile du signal sonore. Ce dernier point fait foi que les paramètres obtenus seront significatifs d'un même intervalle temporel. On obtient une bonne approximation de cette condition en prenant un intervalle de durée proportionnelle à la bande équivalente de bruit (EQN), centré en un point fixé.

Pour ces études, on effectue les tests sur 22 fenêtres présentées dans le tableau 3.1. D'autre part, les conditions communes à tous les tests qui suivent sont les suivantes :

- la fréquence d'échantillonnage du signal sonore est $F_E = 44100 Hz$;
- le sous-échantillonnage associé aux paramètres additifs de synthèse est $SE = 100$; aussi, l'évaluation des paramètres au pas $T_E = \frac{1}{F_E}$ est faite par interpolation linéaire ;
- le nombre de points du signal pris est donné par la formule $N_A = EQN \times N_{\text{période}}$, où $N_{\text{période}}$ représente le nombre de points du signal correspondant à une période ;
- le nombre de points du spectre est $N_S = 2^{10} = 1024$;
- enfin, le pas d'avancement¹ est de valeur SE .

1. indique (en nombre de points) la progression temporelle faite entre deux fragments successifs du signal sonore.

Numéro attribué	Type de la fenêtre	Paramètre a	Paramètre b	Paramètre c	EQN (en bins)
1	rectangle				1.00
2	Hanning d'ordre 2				1.50
3	Hanning-Poisson pour $\alpha = 2$				2.02
4	famille 1	0.0001	21.600		3.71
5	famille 1	0.0010	18.700		3.45
6	famille 1	0.0100	16.400		3.23
7	famille 1	0.1000	13.600		2.98
8	famille 1	1.0000	7.000		2.50
9	famille 1	1.2000	5.280		2.34
10	famille 1	1.4000	3.530		2.15
11	famille 1	1.6000	2.300		2.03
12	famille 1	1.8000	0.920		1.88
13	famille 2	2.5000	0.099	1.0	2.12
14	famille 2	2.4000	0.191	0.8	2.10
15	famille 2	2.3000	0.369	0.6	2.09
16	famille 2	2.0000	1.175	0.4	2.07
17	famille 2	1.2000	6.107	0.2	2.06
18	famille 2	0.1000	23.278	0.1	2.06
19	famille 2	1.4000	0.356	1.0	1.81
20	famille 2'	1.0000	0.973	0.7	1.80
21	famille 2'	0.8000	1.442	0.6	1.80
22	famille 2'	0.5000	2.287	0.5	1.80

TAB. 3.1 – Tableau des fenêtres utilisées pour les tests comparatifs.

On peut signaler également que pour se rapprocher des caractéristiques de signaux sonores généralement vérifiées, les partiels auront leur amplitude fonction décroissante de la fréquence dans la plupart des tests.

3.1 Tests 1 & 2 : variation des amplitudes

Pour ces deux tests, on prend trois partiels de fréquences fixes $\mathcal{F}_1 = 440Hz$, $\mathcal{F}_2 = 1400Hz$, $\mathcal{F}_3 = 4000Hz$. Et, les fréquences d'initialisation sont définies par $f_k = \mathcal{F}_k + 100$ (en Hz).

3.1.1 Test 1 : variation non linéaire des amplitudes

On définit pour $1 \leq k \leq 3$ et $n \in [0, 99]_{\mathbb{N}}$, les amplitudes (sous-échantillonnées) par $a_k(1 + n \cdot SE) = \frac{\mathcal{F}_1}{\mathcal{F}_k} \cdot 100^{\frac{n}{99}}$. On obtient ainsi des amplitudes croissant de plus en plus rapidement, ce qui permet de tester la réaction de la méthode pour des variations de dynamique faibles puis fortes.

On effectue alors la série de tests. Les résultats obtenus montrent que les erreurs fréquentielles $\left\| \frac{f_k(1+n \cdot SE) - \mathcal{F}_k}{\mathcal{F}_k} \right\|_{\infty}$ à k fixé, rendent bien compte de la qualité de l'ensemble des paramètres mesurés. C'est pourquoi, afin de ne pas surcharger les pages suivantes de nombreux tableaux et figures, on utilisera cette unique caractérisation.

À partir des mesures obtenues, on dresse un classement des fenêtres, d'abord pour chaque partiel (tabl. 3.2), puis globalement selon deux critères : la somme des erreurs sur les trois partiels, et le maximum des erreurs sur les trois partiels (tabl. 3.3).

Comme on peut le voir, on a des résultats tout à fait convenables pour l'ensemble des fenêtres, même si l'on remarque une nette dégradation en ce qui concerne la fenêtre rectangulaire. Plus particulièrement, la seconde famille semble la plus efficace, et encore plus précisément, sa deuxième sous-famille, pour laquelle le lobe principal est relativement étroit.

3.1.2 Test 2 : variation croissante d'une seule des amplitudes

Pour ce test, on s'intéresse à l'influence que peut avoir un partiel de plus en plus puissant, sur l'estimation des autres partiels. On définit alors les amplitudes ainsi, pour $n \in [0, 99]_{\mathbb{N}}$:

$$- \text{ pour } 2 \leq k \leq 3, a_k(1 + n \cdot SE) = \frac{\mathcal{F}_1}{\mathcal{F}_k};$$

$$- \text{ et, } a_1(1 + n \cdot SE) = 1 + 10 \cdot \frac{n}{99}.$$

Les résultats obtenus sont tout à fait similaires aux précédents à ces deux remarques près : les erreurs mesurées sont deux fois supérieures aux précédentes, mais celles obtenues pour la fenêtre rectangulaire sont bien pires (15 fois supérieures). Aussi, exception faite de cette dernière fenêtre, on peut à nouveau considérer ce test comme satisfaisant.

Partiel 1		Partiel 2		Partiel 3	
Numéro fenêtre	Erreur fréquentielle (en %)	Numéro fenêtre	Erreur fréquentielle (en %)	Numéro fenêtre	Erreur fréquentielle (en %)
19	0.123242	20	$5.66459 \cdot 10^{-2}$	7	$3.1845 \cdot 10^{-5}$
20	0.124378	22	$5.66851 \cdot 10^{-2}$	6	$9.56395 \cdot 10^{-5}$
21	0.124861	21	$5.6709 \cdot 10^{-2}$	5	$1.75542 \cdot 10^{-4}$
22	0.125197	19	$5.75683 \cdot 10^{-2}$	4	$1.91426 \cdot 10^{-4}$
18	0.140322	17	$7.82225 \cdot 10^{-2}$	8	$3.29986 \cdot 10^{-4}$
17	0.140386	18	$7.83596 \cdot 10^{-2}$	9	$3.95541 \cdot 10^{-4}$
15	0.140779	15	$7.94041 \cdot 10^{-2}$	12	$5.34082 \cdot 10^{-4}$
16	0.140879	16	$7.95151 \cdot 10^{-2}$	10	$5.3472 \cdot 10^{-4}$
13	0.1415	14	$8.05574 \cdot 10^{-2}$	11	$5.50975 \cdot 10^{-4}$
14	0.142704	12	$8.08395 \cdot 10^{-2}$	22	$6.40585 \cdot 10^{-4}$
12	0.145579	13	$8.10095 \cdot 10^{-2}$	21	$6.502 \cdot 10^{-4}$
3	0.155668	2	$8.21878 \cdot 10^{-2}$	20	$6.58607 \cdot 10^{-4}$
11	0.169142	3	$8.90836 \cdot 10^{-2}$	19	$6.61415 \cdot 10^{-4}$
2	0.170531	11	$9.37058 \cdot 10^{-2}$	3	$6.81522 \cdot 10^{-4}$
10	0.183114	10	$9.78634 \cdot 10^{-2}$	17	$8.45216 \cdot 10^{-4}$
9	0.194305	9	0.10957	18	$8.46669 \cdot 10^{-4}$
6	0.201569	6	0.116193	16	$8.74108 \cdot 10^{-4}$
8	0.205468	8	0.116602	14	$9.17506 \cdot 10^{-4}$
5	0.215023	7	0.119503	15	$9.21054 \cdot 10^{-4}$
4	0.217162	5	0.119578	13	$9.5417 \cdot 10^{-4}$
7	0.219408	4	0.121461	2	$2.3423 \cdot 10^{-3}$
1	0.743721	1	0.931918	1	1.18771

TAB. 3.2 – Classement des fenêtres pour chaque partiel - Test 1

Numéro fenêtre	Somme des erreurs (%)	Numéro fenêtre	Maximum des erreurs (%)
19	0.181472	19	0.123242
20	0.181683	20	0.124378
21	0.18222	21	0.124861
22	0.182522	22	0.125197
17	0.219453	18	0.140322
18	0.219528	17	0.140386
15	0.221104	15	0.140779
16	0.221268	16	0.140879
13	0.223464	13	0.1415
14	0.224179	14	0.142704
12	0.226953	12	0.145579
3	0.245433	3	0.155668
2	0.255061	11	0.169142
11	0.263398	2	0.170531
10	0.281512	10	0.183114
9	0.304271	9	0.194305
6	0.317858	6	0.201569
8	0.3224	8	0.205468
5	0.334777	5	0.215023
4	0.338814	4	0.217162
7	0.338943	7	0.219408
1	2.86335	1	1.18771

TAB. 3.3 – Classement global des fenêtres - Test 1

3.2 Tests 3 & 4 : variation des fréquences

3.2.1 Test 3 : variation non linéaire des fréquences

On cherche ici à mesurer les erreurs d'estimation lorsque les partiels voient leur fréquence varier plus ou moins rapidement. De façon analogue au premier test, on définit les partiels de la façon suivante : soient les fréquences initiales F_k respectivement égales à $440Hz$, $1400Hz$ et $4000Hz$; on définit alors les amplitudes par les amplitudes fixes $A_k = \frac{F_1}{F_k}$, et les fréquences par $f_k(1 + n \cdot SE) = F_k \cdot 4^{\frac{n}{99}}$.

On obtient encore des résultats très semblables à ceux du premier test.

3.2.2 Test 4 : évolution d'une seule des fréquences vers la valeur d'une autre

Cette fois-ci, on veut étudier les phénomènes de convergence lorsqu'on rapproche une fréquence d'un partiel, de celle d'un autre. Le conditionnement des pseudo-matrices devenant arbitrairement élevé, on atteint alors les limites de fonctionnement notamment en ce qui concerne l'estimation des amplitudes et des phases (cf. p27).

On utilise pour cette étude les paramètres suivants :

- les amplitudes A_k sont constantes, égales à 1 ;
- les fréquences des partiels 1 & 3 sont constantes et respectivement égales à $440Hz$ et $2000Hz$; et, on définit celle du deuxième partiel de la façon suivante : $f_k(1 + n \cdot SE) = 440 \cdot 2^{\frac{99-n}{99}}$;
- les fréquences d'initialisations sont définies par les fréquences des partiels décalées de +5% de leur valeur.

Les résultats obtenus sont encore assez homogènes. On observe que les fréquences sont très bien estimées², mais il n'en est pas de même pour les amplitudes. Aussi, les résultats rassemblés dans le tableau 3.4, résultats obtenus pour la fenêtre 12 (que l'on peut considérer comme la fenêtre significative de la qualité moyenne de l'ensemble, cf. tabl. 3.3), montrent que les erreurs d'estimation des amplitudes croissent avec le conditionnement de la matrice $\mathcal{H}^H \cdot \mathcal{H}$, estimé pour la dernière itération.

On peut considérer que la méthode reste fiable pour $\frac{|f_2 - f_1|}{f_1} \geq S$ où S est de l'ordre de 0.1 à 0.3. Ceci semble donc indiquer que si l'on connaît les fréquences présentes dans le son à une faible erreur près, le fait de prendre un nombre de points d'analyse de la forme $N_A > \frac{F_P}{F} \cdot EQN \times S_{max}$ où F est le minimum de l'écart fréquentiel minimal entre les différents partiels et la

². on peut attribuer ces bons résultats au fait que les fréquences d'initialisation sont faiblement décalées par rapport aux valeurs réelles.

Indice temporel n	Proximité des fréquences 1 & 2 (en Hz)	Conditionnement de $\mathcal{H}^H \cdot \mathcal{H}$ (cf. p27)	Erreur d'amplitude (en %)	
			Partiel 1	Partiel 2
40	225.331	12.0713	0.545697	0.614459
41	220.688	12.9008	0.580737	0.627943
42	216.079	13.0906	0.650912	0.700613
43	211.501	14.0479	0.903791	0.972122
44	206.956	14.0891	1.25714	1.35778
45	202.442	15.3007	2.39777	2.52456
46	197.96	15.0966	2.48347	2.57388
47	193.509	16.6071	5.45376	5.48255
48	189.089	17.6017	0.139403	0.21747
49	184.699	18.2311	0.474408	0.480626
50	180.341	19.5643	0.63497	0.0391535
51	176.013	19.4356	1.37479	1.30871
52	171.715	21.2615	7.0602	7.20871
53	167.447	22.0111	0.921537	0.856586
54	163.209	23.3188	0.643598	0.472804
55	159	25.2828	1.60824	1.3233
56	154.821	24.3738	3.47521	3.5502
57	150.671	27.5471	0.72773	0.760522
58	146.55	30.0961	1.4722	0.668999
59	142.457	29.0307	3.22876	3.29999
60	138.393	32.6628	0.92249	0.888353
61	134.358	35.933	2.67392	2.14912
62	130.351	34.4298	3.96393	3.94489
63	126.371	38.7245	1.18153	0.207852
64	122.42	42.3766	9.87529	9.95715
65	118.496	46.7629	1.21909	0.801143
66	114.599	44.541	3.83349	3.83803
67	110.73	51.0915	1.3922	1.10992
68	106.887	56.6775	10.5643	10.4711
69	103.071	62.3624	1.36715	0.870673
70	99.2824	59.5132	4.28013	4.32002
...

Indice temporel n	Proximité des fréquences 1 & 2 (en Hz)	Conditionnement de $\mathcal{H}^H \cdot \mathcal{H}$ (cf. p27)	Erreur d'amplitude (en %)	
			Partiel 1	Partiel 2
...
71	95.5198	67.8403	1.97203	0.562935
72	91.7834	74.925	13.9339	13.9528
73	88.0732	88.6368	6.51777	4.28243
74	84.3888	86.767	2.17451	2.20403
75	80.7301	88.0986	5.6052	5.57531
76	77.097	103.137	2.69795	0.0839094
77	73.4892	118.153	14.9486	14.8533
78	69.9065	134.256	15.9684	15.7147
79	66.3489	162.762	6.64528	1.8766
80	62.8161	154.841	3.24559	3.33476
81	59.3079	164.202	6.42707	6.46901
82	55.8242	180.193	6.38906	5.29055
83	52.3648	232.882	2.58546	3.86788
84	48.9296	275.649	19.1184	18.71
85	45.5183	303.403	25.9915	26.0198
86	42.1308	355.292	23.6988	23.5679
87	38.767	448.258	19.5566	17.1127
88	35.4266	545.09	9.54039	6.95899
89	32.1096	632.432	0.731534	3.08903
90	28.8156	667.781	9.05866	9.02325
91	25.5447	818.673	11.4221	11.4332
92	22.2966	1002.85	14.3235	14.4117
93	19.0711	1258.1	15.4471	18.0488
94	15.8682	1826.3	4.32342	22.9375
95	12.6875	2686.04	10.2927	31.1587
96	9.52913	4452.3	28.5766	38.5949
97	6.39275	7198.14	45.2477	48.9321

TAB. 3.4 – Résultats moyens observés sur l'estimation des amplitudes pour deux partiels arbitrairement proches - Test 4

fréquence la plus basse, donnera de bonnes estimations sur les amplitudes. Dans la pratique, on prendra $S = 1$ ce qui cautionnera une bonne qualité.

Cependant, cette sécurité est prise surtout en considération de l'éventuelle faible qualité des estimations des fréquences d'initialisation. On peut alors penser pour les études qui poursuivront, que le seuil S pourra être utilisé jusqu'à sa limite pour les analyse faites aux "bords" du son : en effet, si l'on considère un instrument jouant deux notes succesives liées, l'analyse faite sur la transition ne correspondant pas à une zone stable sera très perturbée. En se référant aux études effectuées par DIGITAL MUSIC RESEARCH GROUP [BM95], on pourra mettre en œuvre une segmentation du signal et connaître ainsi les points de transition. Ainsi, on pourra utiliser un algorithme particulier aux extrémités du son, algorithme qui prendra pour fréquences d'initialisation les dernières fréquences estimées (pour N_A points de signal) ; puis en prenant $S = 0.3$, on pourra analyser le signal sur les $0.3 \cdot N_A$ points les plus à l'extrémité puisqu'on sera dans de bonnes conditions d'initialisation, et donc minimiser les "trous d'analyse" dus à la transition.

3.3 Résultats

Les tests effectués sur l'ensemble des fenêtres montrent que, finalement, les fenêtres construites sont bien adaptées à la méthode. Les fenêtres de la seconde famille, pour les paramètres autorisant une légère inflexion, semblent être les meilleures. Cependant, la première famille garde un grand intérêt, notamment pour les paramètres a élevés : particulièrement, la fenêtre obtenue pour le "carré d'un triangle", soit $a = 2$ et $b = 0$, correspond à la fenêtre sans lobes secondaires dont la bande passante à -3dB est la plus faible de toutes.

Finalement, on a déterminé des fenêtres qui empêchent tout comportement chaotique de la méthode, et qui de plus, semblent offrir les meilleurs résultats quant à la qualité des estimateurs. Il semble donc que la recherche fonctionnera bien dès lors que les initialisations seront suffisamment correctes. Ainsi, l'étape suivante consiste à déterminer une chaîne complète d'analyse, reposant sur cette méthode dont on peut espérer de bons résultats.

Troisième partie

Construction d'une chaîne algorithmique complète pour la mesure des paramètres additifs

Chapitre 1

Mise en oeuvre de la chaîne algorithmique

Dans ce chapitre, on construit des éléments qui vont permettre l'automatisation de la mesure des paramètres additifs sur chaque fenêtre de signal sonore à analyser, en utilisant la méthode précédemment décrite. Cependant, la structure de la chaîne complète d'analyse, chaîne qui contient notamment une fonction chargée de déterminer les partiels en formant des “trajets de paramètres additifs”¹, est décrite dans le dernier chapitre. D'autre part, la modélisation du bruit n'est pas abordée dans la mesure où, une fois les paramètres trouvés, la méthode employée est rigoureusement identique à la celle de “l'analyse classique”.

1.1 Initialisation des fréquences

Comme on l'a vu dans les chapitres précédents, la méthode qu'on veut utiliser nécessite des données d'initialisation. On s'intéresse donc ici à dénombrer les partiels existant dans la partie du signal à analyser, et à évaluer leur fréquence de façon approximative.

1.1.1 Détermination des fréquences approchées présentes dans le fragment sonore

On dispose, pour l'évaluation que l'on veut faire, d'un fragment sonore sur N_A points. Ce nombre N_A est également déterminé en fonction du contenu du signal puisqu'il est proportionnel au minimum de l'écart fréquentiel minimal et de la plus basse fréquence; cependant, on explicitera la manière dont on l'obtient dans le paragraphe suivant. Il est de plus proportionnel à la bande équivalente de bruit de la fenêtre d'analyse.

1. le principe de cette fonction n'est pas étudié ici; cependant, on pourra se référer à l'article [DGR93] pour plus de précisions.

L'évaluation des fréquences est évidemment faite sur le spectre en utilisant, en quelque sorte, une analyse classique simplifiée. Cependant, la faiblesse de la valeur N_A ne permet pas la séparation des "pics spectraux" puisque la méthode qu'on utilise ne le nécessite pas et que, d'autre part, on cherche à améliorer la localisation temporelle des paramètres additifs. Par conséquent, afin de rendre cette séparabilité, on utilise une fenêtre dont la bande passante à -3dB est beaucoup plus faible (environ 2 fois) que celle de la fenêtre d'analyse pour la nouvelle méthode; on prendra donc le plus souvent la fenêtre rectangulaire.

Principe

Une fois le spectre calculé au moyen de cette nouvelle fenêtre, on procède ainsi :

- on détermine d'abord tous les maxima locaux sur le module du spectre, en évaluant les fréquences associées² ;
- puis, on conserve les éléments de plus forte puissance dans une dynamique de $\mathcal{P}dB$.³
- enfin, on ne conserve des éléments restants, que ceux dont la forme est semblable à la forme dans le domaine fréquentiel, de la fenêtre employée. Pour cela, on utilise un critère de forme.

On obtient ainsi les éléments pertinents qui composent le fragment du signal sonore. Dès lors, on peut disposer des fréquences et des amplitudes approchées.

Le critère de forme

On considère en premier lieu, le nombre de points N_C sur lequel on va caractériser la forme⁴. Ce nombre est bien sûr proportionnel à $\frac{N_S}{N_A}$.

On définit le produit scalaire hermitien suivant :
pour $H_1(n)$ et $H_2(n)$ avec $-\frac{N_C-1}{2} \leq n \leq \frac{N_C-1}{2}$

$$\langle H_1, H_2 \rangle = \sum_{n=-\frac{N_C-1}{2}}^{\frac{N_C-1}{2}} H_1(n) \cdot \overline{H_2(n)} \quad (1.1)$$

Ce produit scalaire permet alors de définir le critère de forme par : si, pour $-\frac{N_C-1}{2} \leq n \leq \frac{N_C-1}{2}$, $W(n)$ est la partie de la fenêtre fréquentielle, référence

2. au bin près, mais sans effectuer d'interpolation.

3. ce paramètre \mathcal{P} est déterminé par l'utilisateur ; il a pour but d'autoriser un seuillage (pour 100dB par exemple) mais peut également être supprimé en le prenant infini.

4. ce nombre doit être impair pour pouvoir considérer un nombre de points identique autour du maximum de la fenêtre.

de la forme, et $S_b(n) = S(n + b)$ une partie du spectre où l'existence d'un "pic" déterminé au point b ⁵ doit être décidé, alors le critère de décision est

$$C = \sqrt{\frac{|\langle W, S_b \rangle|}{\|W\| \cdot \|S_b\|}} \quad (1.2)$$

où $\| \cdot \|$ est la norme associée au produit scalaire défini. Ce critère est toujours inférieur ou égal à 1, et caractérise une forme d'autant meilleure qu'il est proche de 1.

Ainsi, si $C \geq S_C$ où S_C est le seuil de décision (dans la pratique on prend $S_C \geq 0.7$), on considèrera le "pic" comme acceptable ; il sera non-acceptable sinon.

1.1.2 Détermination du nombre de points du signal à prélever localement

La bonne qualité de l'analyse requiert que l'information contenue dans le signal soit suffisante. Autrement dit, s'il n'est pas utile que les "pics spectraux" soient séparés, il n'en reste pas moins que le nombre N_A soit proportionnel à la fréquence présente la plus basse, ou même à l'écart fréquentiel minimal sur l'ensemble des partiels existants si celui-ci est plus petit (ce qui est souvent le cas à cause de l'inharmonicité). Aussi, il est important d'estimer la valeur fréquentielle à partir de laquelle on pourra calculer le nombre de points minimal mais nécessaire à l'analyse, N_A .

Détermination de l'ensemble partiels localement présents, sans aucune connaissance *a priori* du signal

N'ayant aucune connaissance *a priori* du signal à analyser, le seul moyen dont on dispose pour fixer la durée du fragment à pré-analyser sans masquer de fréquences importantes, ne peut se référer qu'au domaine de la psycho-acoustique. On définit alors la limite ainsi : puisqu'un "auditeur moyen" n'entend une fréquence en tant que hauteur qu'au-dessus de $20Hz$, on peut considérer que toute fréquence ou écart fréquentiel supérieurs doivent être détectés ; mais dans le cas contraire, toute fréquence inférieure peut être associée à une modulation des paramètres, qui sera modélisé dans l'évolution temporelle de ces-derniers.

Finalement, on utilise le détecteur défini précédemment pour un nombre de points associés à la fréquence limite de $27.5Hz$ ⁶ correspondant au la_0 . Puis, à partir des résultats obtenus, on déterminera N_A .

5. mesuré en bins

6. ce paramètre sera en fait laissé libre à l'utilisateur.

Évaluation de la fréquence limite en fonction de l'ensemble des fréquences détectées

La fréquence limite dépend de deux valeurs : la fréquence minimale, et l'écart fréquentiel minimal. Et, il s'avère qu'on peut, dans certains cas, augmenter la deuxième valeur sans nuire au bon fonctionnement de la méthode. Ceci est donc particulièrement intéressant dans le cas où ce sera cette seconde valeur qui sera la plus petite des deux (et donc retenue) : la localité temporelle des paramètres sera plus fine.

Le principe utilisé est le suivant : on détermine des sous-ensembles de fréquences en considération de leur amplitude, de façon telle que tous les éléments non représentés dans un sous-ensemble influenceront très peu la qualité des paramètres de l'élément de ce sous-ensemble de plus forte amplitude. On constitue ainsi un premier sous-ensemble qui contient l'élément de plus forte amplitude, puis un second qui ne contient pas l'élément précédent mais la "seconde amplitude la plus forte", (etc ...) jusqu'à ce que tous les éléments soient représentés au moins une fois. On calcule alors l'écart fréquentiel minimal pour chaque sous-ensemble, et on en prend le minimum. La valeur obtenue est alors supérieure ou égale à l'écart fréquentiel minimal déterminé sur l'ensemble des éléments. De plus, elle garantit que les paramètres analysés ne seront presque pas affectés. Il convient de remarquer que certains des éléments ne seront pas détectés à l'initialisation ; cependant, ils apparaîtront dans la différence entre le spectre observé et le spectre modélisé (i.e. le résidu), puisque les premiers paramètres ne seront pas dégradés par leur absence.

Il convient maintenant d'explicitier le critère qui permet de créer ce sous-ensemble : on considère un son constitué de deux partiels de fréquences fixes F_1 et F_2 , et on fixe l'amplitude du premier partiel à 1. On cherche alors l'amplitude limite $a(|F_2 - F_1|)$ du second partiel "potentiel", telle que l'erreur d'estimation faite sur F_1 lorsque le second partiel n'est pas représenté, n'excède pas un seuil que l'on note E_1 (valeur mesurée en comma).

Ainsi, si dans un fragment de signal l'élément de plus forte amplitude A a pour fréquence F , on obtiendra le premier sous-ensemble répondant à l'erreur maximale E_1 par $\{(F_i, A_i) / \frac{A_i}{A} \in [a(|F_i - F|), 1]\}$, ensemble qui contient (F, A) bien entendu.

On donne dans le tableau 1.1 l'ordre de grandeur de

$$a_{min} = \min_{F_2 \in \mathbb{F}_M} \{a(|F_2 - F_1|)\}$$

On peut ajouter que a_{min} représentant une valeur minimale, on obtiendra dans la plupart des cas des résultats meilleurs que ceux annoncés dans le tableau : un seuil compris entre 0.02 et 0.1 sera généralement suffisamment efficace.

E_1 (en comma)	a_{min}	a_{min} (en dB)
$\frac{1}{2}$	$1.07 \cdot 10^{-2}$	-39.38
1	$2.05 \cdot 10^{-2}$	-33.76
$\frac{3}{2}$	$4 \cdot 10^{-2}$	-27.95
$\frac{5}{2}$	$8.3 \cdot 10^{-2}$	-21.62

TAB. 1.1 – *Tableau de caractérisation de l'influence d'un partiel non détecté sur l'estimation d'un autre, en fonction du rapport d'amplitudes et de l'écart fréquentiel*

1.2 Problèmes liés à l'exhaustivité de la recherche

À l'issue de l'initialisation, l'ensemble des partiels sélectionnés peut contenir des partiels “parasites”. À l'inverse, des partiels existant peuvent ne pas avoir été détectés. On se propose donc, dans cette partie, de pallier ces deux problèmes.

1.2.1 Fusion de partiels lors d'une convergence

Lorsqu'il y a des partiels “parasites”, on observe dans la majeure partie des cas une “fusion de partiels” lors de la convergence. Ainsi, on va assister à un fort rapprochement de fréquences. Les représentants de ces partiels devenant très semblables, le conditionnement des matrices d'estimation (notamment pour les amplitudes et les phases) devient arbitrairement grand, et les amplitudes souvent très fortes (avec une “quasi-opposition” de phase).

Cependant, on ne peut pas éliminer directement les partiels qui semblent fusionner pour la raison suivante : il arrive régulièrement qu'une fréquence estimée à l'initialisation soit mauvaise (généralement lorsqu'elle est proche d'une autre, et d'amplitude relativement faible), et qu'on assiste alors à ce phénomène dans les premières itérations.

La solution envisagée a donc été de supprimer les fréquences “gênantes” uniquement lorsque la convergence se stabilise, c'est-à-dire que l'évolution des paramètres devient relativement faible. Dans la pratique, on fixe le “seuil de stabilité” à dix fois le seuil de sortie. Et, on effectue ce “filtrage” des fréquences en tenant compte de leur proximité. Finalement, si deux fréquences sont détectées trop proches dans une situation stable, on supprimera l'une des deux. Cependant, il peut y avoir des cas plus complexes pour lesquels plus de deux fréquences sont détectées “gênantes”. Aussi, parmi cet ensemble de fréquences, on doit en supprimer certaines, et ceci de la manière la plus pertinente possible. On détermine pour cela un procédé qui supprime les

fréquences une par une de la façon suivante :

- le choix de la fréquence à supprimer parmi les “fréquences incriminées” se fait d’abord en considération de “l’ordre de gêne” ; si une fréquence f_2 en gêne deux autres f_1 et f_3 , et que f_1 et f_3 ne gênent que f_2 , on définit les “ordres de gêne” $o_2 = 2$, $o_1 = o_3 = 1$. On sélectionne alors les fréquences dont l’ordre est maximal.
- puis, pour chaque fréquence obtenue, on calcule les écarts fréquentiels entre celle-ci et toutes les autres fréquences présentes. Chaque ensemble des écarts, attribué à une fréquence, est ordonné.
- on décide alors de la fréquence à supprimer en prenant celle qui est associée au plus petit “premier écart” ; si ce n’est pas suffisant, on considère le deuxième ; etc...
- si, à la fin, il reste encore plusieurs fréquences candidates, on décide arbitrairement de prendre la plus élevée (on suppose l’énergie des partiels décroissante avec la fréquence).

Ainsi, on supprime une par une les fréquences gênantes (selon cette priorité), jusqu’à ce qu’il n’en reste plus.

1.2.2 Détection de partiels oubliés

Il est encore plus fréquent que certains partiels ne soient pas détectés à l’initialisation : ceux-ci peuvent être masqués par des partiels de plus forte amplitude, volontairement (Cf. l’évaluation de N_A p 65) ou involontairement.

Pour pallier ces oublis, on effectue à nouveau une analyse sur le résidu obtenu : initialisation des fréquences, puis estimation par la méthode itérative sur le spectre résiduel. On peut itérer ce procédé jusqu’au résidu d’ordre R , si l’on y trouve des “partiels oubliés”. Dans la pratique, un ordre maximal R_{max} est défini. On peut ainsi limiter le nombre de recherches qui deviennent malgré tout de moins en moins fiables puisqu’un résidu cumule l’ensemble des éventuelles erreurs faites aux ordres plus faibles. Cependant, l’utilisateur peut le fixer à son gré.

On détermine ainsi plusieurs ensembles de paramètres additifs pour un même fragment de signal, chacun étant associé à l’ordre du résidu dans lequel l’estimation a été faite.

On appellera un ensemble dont les résultats ont été obtenus pour le résidu d’ordre r , ensemble des paramètres d’ordre r .

1.3 Concaténation de l'ensemble des résultats obtenus pour un fragment sonore

On se préoccupe ici, de rassembler les résultats obtenus pour les ensembles de paramètres des différents ordres, sur un même fragment de signal. Cette opération tient compte du fait que plus l'ordre de l'ensemble est élevé, plus les erreurs d'estimation voire d'existence peuvent se manifester. Aussi, on adopte une méthode qui vise à minimiser ces deux types d'erreurs.

Pour réduire les erreurs quant à l'existence de partiels appartenant aux ensembles d'ordres élevés, on procède de la façon suivante :

On concatène d'abord les deux ensembles d'ordre le plus élevé, puis l'ensemble obtenu avec l'ensemble d'ordre juste supérieur, et ainsi de suite. On concatène donc des ensembles deux à deux en partant de l'ordre le plus élevé...

1. On considère deux ensembles de paramètres, d'ordres successifs, \mathcal{E}_r et \mathcal{E}_{r+1} . On va alors concaténer les paramètres fiables à l'ordre r dans un ensemble qu'on note \mathcal{E}'_r .
2. On considère $\mathbf{f0} = (f_1^r, \dots, f_n^r, f_1^{r+1}, \dots, f_m^{r+1})$ le p-uplet dont les n premières composantes sont les fréquences de \mathcal{E}_r et les m dernières composantes sont celles de \mathcal{E}_{r+1} . De même, on considère $\mathbf{a0} = (a_1^r, \dots, a_n^r, a_1^{r+1}, \dots, a_m^{r+1})$ le p-uplet des amplitudes associées.
3. On évalue alors $\mathbf{a}'0$ à partir de $\mathbf{f0}$ et du résidu d'ordre r , en utilisant la méthode (non itérative) d'estimation des amplitudes et des phases décrite page 26.
4. On cacule l'erreur relative $\mathbf{e0}$ sur les amplitudes suivant la formule $\mathbf{e0}_i = \frac{|\mathbf{a0}_i - \mathbf{a}'0_i|}{\mathbf{a0}_i}$. On fixe un seuil de tolérance d'erreur S_e . Dans la pratique, ce seuil est encore une fois fixé par l'utilisateur; un bon ordre de grandeur est de 10%. Peuvent alors se présenter trois cas :

on élimine un partiel d'ordre $r+1$ qui dépasse le seuil de tolérance.
 $\exists i \geq n+1 / \mathbf{e}_i \geq S_e$; dans ce cas, on détermine l'erreur maximale pour $i \geq n+1$, et l'indice I à laquelle elle correspond. On définit alors $\mathbf{f1}$ et $\mathbf{a1}$ à partir de $\mathbf{f0}$ et de $\mathbf{a0}$ en supprimant le terme d'indice I (attaché à l'ensemble d'ordre $r+1$). On reprend à l'étape 3.

on élimine un partiel d'ordre $r+1$ provoquant un dépassement du seuil de tolérance pour un partiel d'ordre r .
 $\exists i \leq n / \mathbf{e}_i \geq S_e$, et $\forall i \geq n+1, \mathbf{e}_i < S_e$; dans ce cas, on détermine l'indice I pour laquelle l'erreur est la plus forte, puis l'indice $J \geq n+1$ du partiel dont l'influence sur le partiel I est la plus grande⁷. On

7. la définition de ce critère est donné page 73.

définit alors, de la même manière que dans le cas précédent, \mathbf{f}_1 et \mathbf{a}_1 en supprimant les termes d'indice J , puis on reprend à l'étape 3.

on concatène sans éliminer

$\forall i, \mathbf{e}_i < S_e$; dans ce cas, on construit l'ensemble \mathcal{E}'_r comme l'ensemble des paramètres restants. On reprend alors à l'étape 1 en considérant les ensembles \mathcal{E}_{r-1} et \mathcal{E}'_r .

(On peut remarquer que $\mathcal{E}_r \subseteq \mathcal{E}'_r$.)

On construit \mathcal{E}'_0 , ensemble des paramètres fiables du son, de proche en proche en utilisant ce procédé: si l'ordre le plus élevé des ensembles de paramètres est R , on estime \mathcal{E}'_{R-1} à partir de \mathcal{E}_R et \mathcal{E}_{R-1} , puis \mathcal{E}'_{R-2} à partir de \mathcal{E}'_{R-1} et \mathcal{E}_{R-2} , \dots , puis \mathcal{E}'_0 à partir de \mathcal{E}'_1 et \mathcal{E}_0 .

Finalement, les paramètres retenus comme représentant des partiels bien existants dans le fragment du signal sont donnés par \mathcal{E}'_0 .

Pour réduire les erreurs quant à la qualité de l'estimation, on prend les fréquences \mathbf{f} retenues par \mathcal{E}'_0 , et on réestime les amplitudes \mathbf{a} et les phases ϕ correspondantes par la méthode d'estimation associée. On n'utilise pas la méthode itérative qui réestimerait par la même les fréquences pour éviter les instabilités que subiraient les paramètres d'ordre supérieur ou égal à 1, à cause de la faiblesse relative de leur amplitude.

Chapitre 2

Réduction de la complexité des opérations en vue d'une implantation sur machine

Dans ce chapitre, on s'attache à modifier certains aspects des éléments de la chaîne, ou même certains aspects de la chaîne algorithmique, dans le but d'améliorer son rendement, sans détériorer la qualité des estimations de façon notable.

2.1 Restriction de la méthode à une zone fréquentielle

La méthode sur laquelle repose l'estimation des paramètres tire une des qualités du fait que l'information est prélevée sur l'ensemble du spectre. Cependant, l'énergie qui porte cette information n'est pas uniformément répartie: les "fenêtres fréquentielles" qu'on utilise, quelles qu'elles soient, voient leur énergie concentrée autour de la fréquence qu'elles représentent. De plus, les méthodes d'estimation utilisées sont assez coûteuses en calcul, et ceci d'autant plus que le nombre de points pris pour le calcul du spectre est élevé. Aussi, une première manière d'alléger la méthode est de ne considérer qu'une zone du spectre, où se concentre l'information cherchée, pour l'estimation des paramètres.

2.1.1 Théorie et mise en œuvre

Si l'on suppose que les fréquences d'initialisation sont bornées par une valeur minimale F_{min} , et maximale F_{max} . Si, d'autre part, on estime que l'erreur fréquentielle est inférieure à $\epsilon\%$ de F_{min} et que la majeure partie de l'énergie d'un partiel est concentrée de part et d'autre de sa fréquence de $u\%$ de F_{min} . On peut alors définir l'intervalle de validité des fréquences

par $\mathcal{I}^V = \left[F_{min} - \frac{\epsilon \cdot F_{min}}{100}, F_{max} + \frac{\epsilon \cdot F_{min}}{100} \right]$ et, de plus, la zone spectrale qu'on risque de nécessiter est

$$\mathcal{Z}^{(utile)} = \left[F_{min} - \frac{(e+u) \cdot F_{min}}{100}, F_{max} + \frac{(e+u) \cdot F_{min}}{100} \right]$$

En notant $\mathbb{F}_M^{(utile)} = \mathbb{F}_M \cap \mathcal{Z}^{(utile)} = \{f_p; f_{p+1}; \dots; f_q\}$, une bonne approximation des estimateurs \hat{p} déterminés par l'équation (2.5) page 27 sera obtenue en prenant :

$$\begin{aligned} - \underline{S} &= \begin{pmatrix} \hat{S}_{f_p} \\ \hat{S}_{f_{p+1}} \\ \vdots \\ \hat{S}_{f_q} \end{pmatrix} \\ - \mathcal{H} &= \begin{pmatrix} H_1(f_p) & \cdots & H_K(f_p) & \left| & H_{K+1}(f_p) & \cdots & H_{2K}(f_p) \\ H_1(f_{p+1}) & \cdots & H_K(f_{p+1}) & \left| & H_{K+1}(f_{p+1}) & \cdots & H_{2K}(f_{p+1}) \\ \vdots & \ddots & \vdots & \left| & \vdots & \ddots & \vdots \\ H_1(f_q) & \cdots & H_K(f_q) & \left| & H_{K+1}(f_q) & \cdots & H_{2K}(f_q) \end{pmatrix} \\ - \text{et, } (\Omega)_{n,k} &= \frac{A_k}{2} (-e^{i\phi_k} W'(f_{p+n-1} - F_k) + e^{-i\phi_k} W'(f_{p+n-1} + F_k)) \text{ où} \\ & 1 \leq n \leq q - p + 1. \end{aligned}$$

On réduit ainsi sensiblement le nombre de calculs à effectuer pour les estimations à effectuer ; et, on vérifie effectivement que les résultats sont inchangés.

2.1.2 Le problème du filtrage des fréquences

Cette amélioration ne peut cependant garantir un bon fonctionnement dans tous les cas : en supposant qu'on ait affaire à un partiel "parasite", puisque le spectre d'étude est tronqué, ce partiel peut sortir de l'intervalle de validité \mathcal{I}^V , seule zone dans laquelle les bonnes estimations sont garanties.

Par conséquent, au filtre des fréquences qui empêche la fusion de partiels (Cf. [1.2.1]), et qui ne s'active qu'une fois la stabilisation effectuée, on rajoute un filtre non conditionnel qui supprime toute fréquence sortant de l'intervalle \mathcal{I}^V .

2.2 Découpage de l'ensemble des fréquences en sous-ensembles selon un critère

La complexité des calculs et le temps qu'ils prennent n'étant pas fonction linéaire du nombre de partiels à déterminer K , le fait de découper l'ensemble des partiels en plusieurs sous-ensembles paraît être un moyen intéressant

d'accélération en soit mais aussi parce que cela accentue l'utilisation de l'optimisation précédente. Tout le problème est de savoir comment associer les partiels à un groupe pour que les estimations gardent leur qualité.

On étudie alors un premier algorithme qui ne s'est pas révélé être à la hauteur de nos espérances.

2.2.1 Création des groupes selon le critère de "l'influence maximale"

Puisque la méthode d'estimation a, entre autre, pour intérêt de considérer les influences mutuelles de partiels, il semble particulièrement intéressant de constituer les groupes sur ce critère : chacun contiendra des partiels qui ont une influence importante les uns sur les autres. Ainsi, deux groupes différents seront relativement indépendants et donc peu perturbés par l'absence des nombreux autres partiels et, les partiels sensibles à la présence d'autres étant regroupés, la qualité des estimations s'en trouvera peu affectée.

Critère de l'influence

On considère deux partiels P_1 et P_2 de fréquences respectives f_1 et f_2 et d'amplitudes complexes respectives a_1 et a_2 .

On définit alors le critère d'influence $C_{1,2} = C_{2,1}$ par :

$$C_{1,2} = 1 - \left| \frac{a_{max}}{a_{min}} \cdot W(|f_2 - f_1|) \right| \quad (2.1)$$

où $\begin{cases} a_{max} = a_1 & \text{si } |a_1| \geq |a_2| \\ = a_2 & \text{sinon} \end{cases}$, et W , la fenêtre en fréquentiel.

Plus ce critère (majoré par 1) est proche de 1, plus la dépendance des partiels est faible.

mise en œuvre

Après avoir déterminé les valeurs d'initialisation, on calcule le critère d'influence pour chaque couple de partiels. On ordonne alors les couples en fonction de ce critère. Puis, on crée les groupes en réunissant les partiels associés par les couples dont le critère est le plus bas. Ainsi, si les deux premiers couples sont représentés par quatre partiels distincts, on aura deux groupes de deux partiels. Si au contraire il n'y a que trois partiels représentés, on aura un groupe de trois partiels. On détermine ainsi, étape par étape, les groupes et, dès que l'un d'eux voit son cardinal dépasser un nombre d'éléments limite (fixé par l'utilisateur), il est analysé. On continue jusqu'à ce qu'on obtienne un unique groupe.

On peut remarquer que ce procédé fait appel à l'algorithme d'analyse un grand nombre de fois. Cependant, les paramètres étant déjà finement

estimés après le premier appel, le nombre d'itérations est beaucoup moins important lorsque les groupes ont un grand nombre d'éléments...

Cependant, ce procédé a l'inconvénient de donner parfois des groupes dont de nombreux éléments ont de faibles amplitudes, et pour lesquels l'intervalle de validité correspondant contient des zones fortement énergétiques non représentées. Aussi, le critère visant à minimiser l'énergie résiduelle, on assiste alors à des "changements d'identités" de partiels. Et, même en déterminant ces partiels, et en remettant leurs valeurs originales dans le nouveau groupe, on finit par ré-augmenter le temps de calcul.

Finalement, ce procédé attirant d'un point de vue théorique, ne donne pas de résultats intéressants quant à la diminution du temps de calcul.

2.2.2 Création des groupes selon le critère de "rupture énergétique"

Pour éviter d'avoir des zones énergétiques non représentées comme dans le cas précédent, et considérant que le critère de la méthode d'estimation repose sur la minimisation de l'énergie résiduelle, il peut sembler intéressant de découper l'axe des fréquences en plusieurs zones, les bornes étant déterminées aux endroits où l'énergie est très faible, relativement à celle des partiels les plus proches de ces bornes.

Aussi, on considère les minima locaux énergétiques m_k du spectre. Puis pour chacun des minima qu'on localise par leur position fréquentielle f_k , on considère le maximum local M_k^- , le plus proche pour les fréquences inférieures à f_k , et M_k^+ , le plus proche pour les fréquences supérieures à f_k . On calcule alors les écarts (en décibels) $\delta_{k\ dB}^- = M_k^- - m_{k\ dB}$ et $\delta_{k\ dB}^+ = M_k^+ - m_{k\ dB}$. Si $\delta_{k\ dB}^- \leq S_{dB}$ et $\delta_{k\ dB}^+ \leq S_{dB}$ où S_{dB} est un seuil fixé par l'utilisateur (20dB ou 40dB par exemple), on retient alors f_k comme borne b_l .

Une fois l'ensemble des bornes obtenues¹ on découpe l'axe des fréquences en $L - 1$ intervalles \mathcal{I}_l , L représentant le nombre de bornes. On détermine alors $L - 1$ ensembles de paramètres \mathcal{P}_l , chacun contenant les informations des partiels dont la fréquence appartient à \mathcal{I}_l .

On réunit ensuite les ensembles jusqu'à ce que les nouveaux ensembles formés possèdent au moins N éléments (N étant encore fixé par l'utilisateur) ou qu'il ne reste qu'un ensemble, de la façon suivante :

- on prend l'ensemble \mathcal{P}_l ayant le moins d'éléments (on en choisi un s'il y en a plusieurs) ;
- s'il possède au moins N éléments, on analyse séparément les ensembles déterminés ; sinon, on prend parmi les ensembles \mathcal{P}_{l-1} et \mathcal{P}_{l+1} celui qui a le moins d'éléments (on choisi \mathcal{P}_{l-1} si les deux on le même cardinal) ;

1. la première étant généralement 0 et la plus grande $F_E(\frac{1}{2} - \frac{1}{N_S})$.

et, on réunit les deux ensembles ; on obtient alors $L - 2$ ensembles ; on recommence alors ce procédé jusqu'à la condition décrite ci-dessus.

Ce procédé est très efficace pour réduire le coût en calculs. Les résultats obtenus sont de qualité légèrement moins bonne mais globalement satisfaisante. Aussi, si l'utilisateur ne cherche pas une analyse rapide mais précise, il suffit de prendre N infini.

Chapitre 3

Vers une chaîne complète d'analyse

3.1 Chaîne algorithmique de mesure des paramètres additifs d'un fragment sonore

Cf. figure 3.1.

3.2 Structure complète de la chaîne algorithmique

On a, pour le moment, déterminé une chaîne qui estime les paramètres additifs d'un fragment de signal. La mesure de l'ensemble des paramètres est donc simplement obtenue en évoluant le long du signal sonore : on peut utiliser un pas d'avancement fixe, ou proportionnel à la plus basse fréquence ; on peut également penser, pour les études futures, à effectuer une pré-analyse du son afin d'éviter d'analyser des zones de transition (intérêt a déjà été évoqué p 17).

Cependant, l'analyse reste encore incomplète : il est nécessaire à partir de l'ensemble des paramètres obtenus de déterminer les partiels en recherchant les trajets temporels auxquels ces paramètres appartiennent. Ceci a été l'objet d'une étude concluante [DGR93].

Enfin, une fois les partiels déterminés, on peut s'attacher à améliorer la qualité des enveloppes temporelles associées aux partiels. En ce qui concerne les phases, on dispose déjà de méthodes efficaces consistant à effectuer un "dewrapping" en utilisant un critère de régularité¹, ce qui minimise les modulations de fréquence indésirables. En ce qui concerne les amplitudes, on propose ici, une première méthode améliorant la résolution temporelle (méthode que l'on considérera comme une piste de réflexion).

1. on l'obtient en minimisant l'expression $\int_{t_i}^{t_{i+1}} \left| \frac{d^2 \phi}{dt^2} \right|^2 dt$ [MQ86].

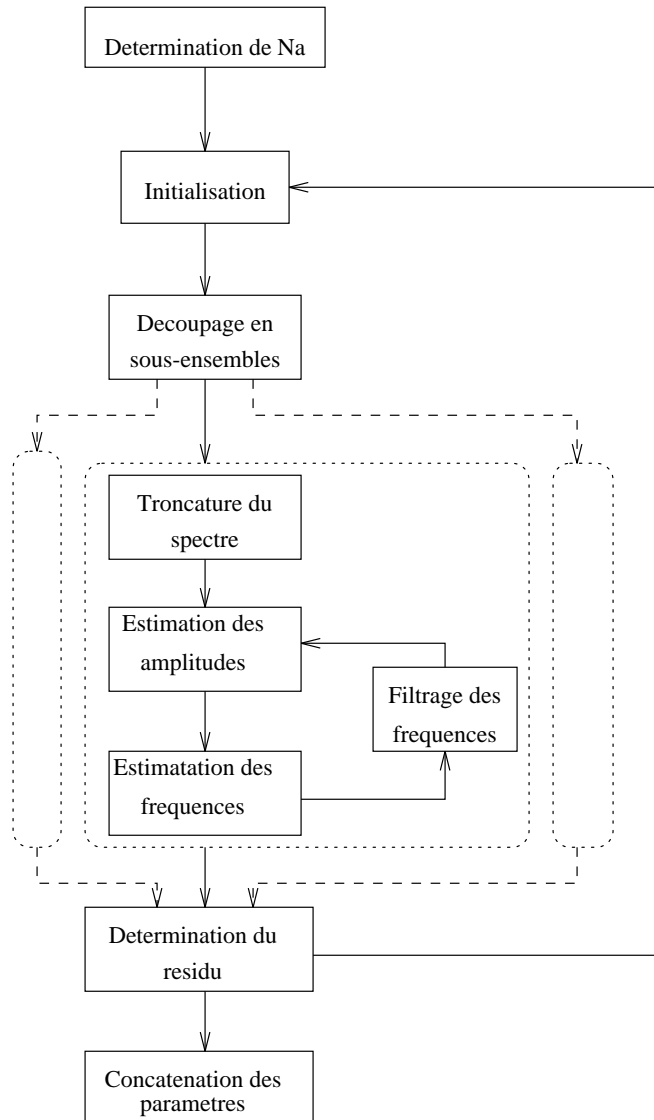


FIG. 3.1 – Forme de la chaîne algorithmique obtenue pour la mesure des paramètres additifs

3.3 Proposition d'une méthode d'affinage de la résolution temporelle des enveloppes d'amplitudes

On considère le signal localement stable en fréquence sur

$$I_m = [i_m, i_m + N - 1]:^2 \quad s(n) = \sum_{k=1}^K a_k(n) \cdot e^{i(2\pi \frac{F_k}{F_E} \cdot n + \Phi_k)}.$$

On estime alors la TFD du signal $s(n)$ pondéré par la fenêtre $w_{I_m}(n)$, en la fréquence F_p , et on cherche à y reconnaître l'amplitude complexe $A_p(m)$ déterminée par l'analyse sur l'intervalle I_m avec la même fenêtre. Afin de simplifier les expressions, on centre I_m en 0 quitte à ajuster les phases Φ_k en $\Phi'_k = \Phi_k + 2\pi \frac{F_k}{F_E} (i_m + \frac{N}{2})$:

$$\begin{aligned} (TFD_{F_p})_m &= \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{k \neq p} a_k(i_m + \frac{N}{2} + n) \cdot w(n) \cdot e^{i(2\pi \frac{F_k}{F_E} \cdot n + \Phi'_k)} \cdot e^{-i2\pi \frac{F_p}{F_E} \cdot n} \\ &+ \underbrace{\sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} a_p(i_m + \frac{N}{2} + n) \cdot w(n) \cdot e^{i(2\pi \frac{F_p}{F_E} \cdot n + \Phi'_p)} \cdot e^{-i2\pi \frac{F_p}{F_E} \cdot n}}_{A_p(m)} \end{aligned}$$

Le terme $A_p(m)$, indépendant des termes obtenus pour $k \neq p$, est bien celui évalué par la méthode puisque la présence de tous les partiels est prise en compte dans l'estimation. On a donc :

$$A_p(m) = e^{i \cdot \Phi'_p} \cdot \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} a_p(i_m + \frac{N}{2} + n) \cdot w(n)$$

On suppose d'autre part que le pas d'avancement est constant, et égal à $\frac{N}{C}$. Si on décompose la somme précédente en C termes de la façon suivante :

$$A_p(m) = \sum_{l=0}^{C-1} \left(\sum_{n=N \cdot (\frac{l}{C} - \frac{1}{2})}^{N \cdot (\frac{l+1}{C} - \frac{1}{2}) - 1} a_p(i_m + \frac{N}{2} + n) \cdot w(n) \right)$$

chaque intervalle $[N \cdot (\frac{l}{C} - \frac{1}{2}), N \cdot (\frac{l+1}{C} - \frac{1}{2}) - 1]$ étant très court, on peut considérer que a_p y est constante. On note alors $\mathcal{A}_p(m \cdot C + l)$ cette valeur moyenne.

2. le fait que le signal soit complexe n'est pas gênant dans la mesure où il s'agit simplement d'une représentation à partir des paramètres associés aux partiels (paramètres dont on disposera).

3.3. PROPOSITION D'UNE MÉTHODE D'AFFINAGE DE LA
RÉSOLUTION TEMPORELLE DES ENVELOPPES D'AMPLITUDES79

On peut alors écrire, en notant $p_l = \sum_{n=N.(\frac{l}{C}-\frac{1}{2})}^{N.(\frac{l+1}{C}-\frac{1}{2})-1} w(n)$:

$$\begin{aligned} A_p(m) &\approx \sum_{l=0}^{C-1} \mathcal{A}_p(m.C+l) \left(\sum_{n=N.(\frac{l}{C}-\frac{1}{2})}^{N.(\frac{l+1}{C}-\frac{1}{2})-1} w(n) \right) \\ &\approx \sum_{l=0}^{C-1} p_l . \mathcal{A}_p(m.C+l) \end{aligned} \quad (3.1)$$

Finalement, en exprimant cette relation pour chacun des $m \in [0, M-1]_{\mathbb{N}}$, et en connaissant (ou extrapolant) les valeurs des dernières $C-1$ amplitudes, on obtient la relation matricielle suivante :

$$\underline{A}_p = \mathcal{M} . \underline{\mathcal{A}}$$

$$\text{où } \underline{A}_p = \begin{pmatrix} A_p(0) \\ \vdots \\ A_p(M-1) \end{pmatrix}, \underline{\mathcal{A}} = \begin{pmatrix} \mathcal{A}_p(0) \\ \vdots \\ \mathcal{A}_p(M+C-1) \end{pmatrix},$$

et, \mathcal{M} est la matrice de dimension $M \times (M+C-1)$ définie par :

$$\mathcal{M} = \begin{pmatrix} p_0 & \cdots & p_{C-1} & 0 & 0 & 0 \\ 0 & p_0 & \cdots & p_{C-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \cdots & \ddots & 0 \\ 0 & 0 & 0 & p_0 & \cdots & p_{C-1} \end{pmatrix}$$

Le nombre d'inconnues ($M+C-1$) est supérieur de $C-1$ au nombre d'équations données par le système. Cependant, si l'on connaît $C-1$ valeurs de \mathcal{A}_p , où qu'elles soient, le système devient soluble. Ceci peut-être obtenu si, par exemple, une zone du son est stable sur au moins $C-1$ mesures consécutives, ou encore si le son est défini depuis son origine, ou jusqu'à sa fin (ce qui signifie que les $C-1$ premières (resp. dernières) coordonnées de $\underline{\mathcal{A}}$ sont nulles).

Finalement, si l'on considère l'ensemble I des indices des $C-1$ coordonnées connues, on définit $\underline{\mathcal{B}}$ le vecteur qui représente ces coordonnées, et $\underline{\mathcal{A}'}$ le vecteur qui représente les coordonnées d'indices appartenant au complémentaire de I . On construit encore la matrice \mathcal{M}_2 à partir des colonnes de \mathcal{M} d'indice appartenant à I , et \mathcal{M}_1 avec les autres colonnes, et tout ceci en respectant l'ordre des indices. On réécrit la relation précédente de la façon suivante :

$$\underline{A}_p = \mathcal{M}_1 . \underline{\mathcal{A}'} + \mathcal{M}_2 . \underline{\mathcal{B}}$$

\mathcal{M}_1 étant carrée et inversible, on peut déterminer les M inconnues par :

$$\underline{\mathcal{A}'} = (\mathcal{M}_1)^{-1} . (\underline{A}_p - \mathcal{M}_2 . \underline{\mathcal{B}})$$

Finalement, on obtient le vecteur cherché \underline{A} , vecteur symbolisant l'enveloppe temporelle de l'amplitude du partiel p avec une résolution temporelle C fois plus importante qu'à l'analyse. Ce procédé appliqué à chaque partiel, lorsque les conditions l'autorisent, permet donc *a priori* d'améliorer notablement les enveloppes d'amplitudes surtout en ce qui concerne les variations brutales telles que les attaques. Cependant, par manque de temps, il n'a pu être testé.

Conclusion

L'analyse additive est un outil qui est déjà utilisé dans le milieu musical depuis plusieurs années, et qui donne de bons résultats dans de nombreux cas. Cependant, pour consolider cet outil et le rendre efficace de façon aussi générale que possible, l'IRCAM continue d'effectuer des recherches. C'est justement dans ce but, que le stage, duquel sont issus les travaux présentés dans ce rapport, a été réalisé.

Les améliorations désirées consistaient essentiellement en : la réduction de la taille de la fenêtre d'analyse, qui était jusqu'alors de l'ordre de quatre périodes, à environ une période afin d'affiner la localisation temporelle des paramètres ; et, accroître la qualité des paramètres estimés, notamment en tenant compte de l'influence mutuelle des différents partiels. Ces désirs ont d'abord amené une première étude [Tro95] sur une modélisation paramétrique du spectre à court terme, permettant de développer une méthode minimisant l'erreur mesurée entre le spectre modélisé et le spectre observé.

Suite à ce travail, il a alors fallu étudier la méthode, puis considérer les différents aspects qui limitaient son bon fonctionnement afin de les pallier. Cette recherche a tout d'abord mené à la mise au point de nouvelles fenêtres interdisant tout comportement chaotique, et assurant la bonne convergence des estimateurs. La seconde étape a alors consisté à établir un premier algorithme d'analyse afin de tester l'efficacité et le comportement de la méthode pour diverses situations sonores. La mise à l'épreuve de cet algorithme de base a ensuite suscité de nombreuses améliorations : en premier lieu, les améliorations garantissant la validité et l'exhaustivité des partiels contenus dans le signal sonore –on peut, entre autre, citer les problèmes liés à la fusion de partiels– ; puis, les améliorations permettant la réduction du coût de calcul en vue d'une implantation sur machine.

Finalement, l'algorithme obtenu a répondu assez favorablement aux améliorations visées puisque la localisation des paramètres estimés a pu être affinée sensiblement et que la qualité d'estimation a très peu souffert de la réduction correspondante –réduction d'un facteur de l'ordre de trois à quatre–. De plus, on a pu vérifier l'efficacité de la prise en compte de l'influence mutuelle des partiels, notamment lorsque le découpage de l'ensemble des paramètres d'initialisation ne sépare pas de partiels s'influçant de fa-

çon non négligeable. Enfin, ce dernier point a permis d'envisager un procédé susceptible d'améliorer encore la résolution temporelle pour les amplitudes, ce qui serait utile pour faire face à des variations rapides telles que les "attaques".

Cependant, il reste de nombreux travaux à effectuer pour améliorer l'analyse. On peut compter au moins deux pistes de recherche : la première concerne l'établissement de méthodes d'affinage des enveloppes temporelles de paramètres, une fois les partiels déterminés ; la seconde concerne l'aide de l'analyse dans ces recherches, en utilisant une "synchronisation avec les transitoires". Cette dernière direction a d'autant plus d'intérêt qu'il s'agit d'un procédé qui se situe en amont de l'analyse, et qu'il est donc susceptible d'augmenter la quantité d'information obtenue dans les paramètres mesurés ; au contraire, un traitement de "post-analyse" ne peut découvrir que l'information déjà incluse dans les paramètres. D'ailleurs, on a pu remarquer que le traitement proposé dans le dernier chapitre est conditionné par un apport d'information.

Enfin, un autre type de travail à venir, est l'implantation des algorithmes d'analyse en langage C, afin d'obtenir des applications rapides et accessibles à l'utilisateur.

Quatrième partie

Annexes

Annexe A

Caractéristiques des fenêtres de la première famille

Paramètre <i>a</i>	Paramètre <i>b</i>	Bande à -3dB (en bins)	Bande à -6dB (en bins)	Bande éq. de bruit (en bins)	Taux de la durée où se concentre 99% de l'énergie (en %)
10^{-4}	21,6	3,48	4,92	3,71	27,8
10^{-3}	18,7	3,24	4,58	3,45	29,9
$5 \cdot 10^{-3}$	17	3,09	4,37	3,29	31,3
10^{-2}	16,4	3,03	4,29	3,23	31,9
$5 \cdot 10^{-2}$	14,4	2,85	4,04	3,04	33,9
0,1	13,6	2,79	3,95	2,98	34,7
0,5	9	2,41	3,41	2,58	2,5
1	7	2,33	3,3	2,5	41,8
1,1	6,15	2,25	3,2	2,42	43,3
1,2	5,28	2,16	3,08	2,34	45,1
1,3	4,55	2,1	2,99	2,27	46,6
1,4	3,53	1,99	2,82	2,15	49,4
1,5	3,09	1,96	2,78	2,12	50,2
1,6	2,3	1,88	2,66	2,03	52,7
1,7	1,84	1,84	2,62	2	53,8
1,8	0,92	1,74	2,45	1,88	57,5
1,9	$1,1 \cdot 10^{-2}$	1,62	2,29	1,74	62,1
2	0	1,67	2,35	1,8	60,3

TAB. A.1 – Tableau des paramètres de la première famille de fenêtre

Annexe B

Caractéristiques des fenêtres de la seconde famille, à la limite de l'existence de l'inflexion

Paramètres (a, b, c) .

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	23.278	11.26	7.616	5.781	4.651	3.964	3.452	3.048	2.775	2.543
0.2	22.352	10.759	7.289	5.519	4.464	3.801	3.299	2.933	2.667	2.437
0.3	21.425	10.251	6.959	5.249	4.275	3.633	3.135	2.816	2.554	2.322
0.4	20.498	9.782	6.625	4.968	4.083	3.459	3.002	2.695	2.435	2.231
0.5	19.568	9.333	6.287	4.731	3.887	3.277	2.868	2.569	2.305	2.138
0.6	18.634	8.881	5.944	4.496	3.685	3.087	2.732	2.437	2.205	2.041
0.7	17.721	8.426	5.594	4.257	3.479	2.934	2.59	2.292	2.101	1.937
0.8	16.813	7.969	5.24	4.015	3.263	2.778	2.442	2.169	1.993	1.824
0.9	15.903	7.508	4.918	3.769	3.034	2.617	2.284	2.053	1.877	1.719
1	14.995	7.042	4.59	3.518	2.831	2.45	2.118	1.931	1.751	1.626
1.1	14.098	6.572	4.251	3.261	2.644	2.276	1.987	1.804	1.63	1.529
1.2	13.199	6.107	3.936	2.996	2.453	2.09	1.851	1.664	1.527	1.423
1.3	12.31	5.66	3.652	2.742	2.256	1.891	1.709	1.515	1.419	1.301
1.4	11.425	5.224	3.363	2.503	2.052	1.738	1.558	1.4	1.3	1.204
1.5	10.559	4.798	3.069	2.251	1.838	1.578	1.388	1.279	1.161	1.108
1.6	9.697	4.367	2.769	2.01	1.642	1.412	1.23	1.148	1.051	1.001
1.7	8.838	3.935	2.472	1.807	1.452	1.233	1.095	0.994	0.943	0.866
1.8	7.987	3.524	2.185	1.598	1.241	1.073	0.95	0.855	0.821	0.773
1.9	7.152	3.114	1.919	1.38	1.081	0.917	0.799	0.733	0.668	0.67
2	6.329	2.713	1.652	1.175	0.921	0.747	0.675	0.597	0.562	0.525
2.1	5.504	2.321	1.395	0.977	0.755	0.622	0.522	0.494	0.444	0.436
2.2	4.677	1.948	1.149	0.792	0.601	0.49	0.424	0.366	0.354	0.326
2.3	3.845	1.584	0.913	0.618	0.463	0.369	0.313	0.281	0.248	0.248
2.4	3.01	1.242	0.694	0.458	0.333	0.263	0.216	0.191	0.174	0.165
2.5	2.17	0.904	0.493	0.313	0.221	0.168	0.137	0.117	0.106	0.099
2.6	1.324	0.558	0.309	0.19	0.126	0.091	0.071	0.059	0.052	0.048
2.7	0.474	0.203	0.116	0.074	0.052	0.036	0.026	0.019	0.016	0.013

TAB. B.1 – Tableau des paramètres b limites - Seconde famille de fenêtres à la limite d'existence d'inflexions

Bande passante à -3dB.

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	1.89	1.91	1.97	2.03	2.08	2.15	2.22	2.27	2.35	2.41
0.2	1.89	1.9	1.96	2.02	2.08	2.14	2.2	2.27	2.33	2.4
0.3	1.89	1.9	1.96	2.01	2.07	2.13	2.18	2.26	2.32	2.37
0.4	1.89	1.9	1.95	2	2.06	2.12	2.18	2.24	2.3	2.36
0.5	1.89	1.9	1.95	1.99	2.06	2.11	2.17	2.23	2.28	2.35
0.6	1.89	1.9	1.94	1.99	2.05	2.09	2.16	2.22	2.27	2.34
0.7	1.89	1.9	1.93	1.98	2.03	2.08	2.15	2.19	2.26	2.32
0.8	1.9	1.9	1.92	1.97	2.02	2.07	2.13	2.18	2.25	2.3
0.9	1.9	1.89	1.92	1.97	2.01	2.07	2.11	2.17	2.23	2.28
1	1.9	1.89	1.91	1.96	2	2.05	2.09	2.16	2.21	2.27
1.1	1.9	1.89	1.9	1.95	1.99	2.04	2.08	2.14	2.19	2.25
1.2	1.9	1.89	1.9	1.94	1.98	2.02	2.07	2.12	2.18	2.23
1.3	1.9	1.89	1.9	1.93	1.97	2	2.06	2.1	2.16	2.2
1.4	1.91	1.89	1.9	1.92	1.96	1.99	2.04	2.09	2.14	2.19
1.5	1.91	1.89	1.9	1.91	1.94	1.98	2.02	2.07	2.11	2.17
1.6	1.91	1.89	1.9	1.91	1.93	1.97	2	2.05	2.09	2.15
1.7	1.92	1.9	1.9	1.91	1.93	1.95	1.99	2.02	2.08	2.11
1.8	1.92	1.9	1.89	1.9	1.91	1.94	1.97	2	2.05	2.09
1.9	1.93	1.9	1.9	1.9	1.91	1.93	1.95	1.98	2.01	2.07
2	1.93	1.91	1.9	1.9	1.91	1.92	1.94	1.96	1.99	2.02
2.1	1.94	1.91	1.9	1.9	1.9	1.91	1.92	1.95	1.97	2
2.2	1.94	1.92	1.91	1.9	1.9	1.91	1.92	1.93	1.95	1.97
2.3	1.95	1.93	1.91	1.91	1.91	1.91	1.92	1.93	1.93	1.96
2.4	1.95	1.94	1.92	1.92	1.91	1.91	1.92	1.92	1.93	1.94
2.5	1.96	1.95	1.93	1.93	1.92	1.92	1.92	1.92	1.93	1.94
2.6	1.96	1.95	1.95	1.94	1.94	1.94	1.93	1.93	1.94	1.94
2.7	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.95	1.95	1.95

TAB. B.2 – *Tableau des bandes à -3dB (en bins) - Seconde famille de fenêtres à la limite d'existence d'inflexions*

Bande passante à -6dB.

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	2.68	2.7	2.79	2.88	2.95	3.05	3.14	3.22	3.32	3.41
0.2	2.68	2.7	2.79	2.86	2.94	3.04	3.12	3.21	3.3	3.39
0.3	2.68	2.69	2.78	2.85	2.93	3.02	3.09	3.19	3.28	3.36
0.4	2.69	2.69	2.77	2.83	2.92	3.01	3.08	3.18	3.26	3.35
0.5	2.69	2.69	2.76	2.82	2.91	2.99	3.07	3.16	3.23	3.33
0.6	2.69	2.69	2.75	2.81	2.9	2.96	3.06	3.14	3.22	3.31
0.7	2.69	2.69	2.74	2.81	2.88	2.95	3.04	3.11	3.2	3.29
0.8	2.69	2.69	2.73	2.8	2.87	2.94	3.02	3.09	3.18	3.26
0.9	2.7	2.69	2.72	2.79	2.84	2.93	3	3.08	3.16	3.23
1	2.7	2.69	2.71	2.78	2.83	2.91	2.97	3.06	3.13	3.21
1.1	2.7	2.68	2.7	2.76	2.82	2.89	2.95	3.04	3.1	3.19
1.2	2.71	2.68	2.7	2.75	2.81	2.87	2.94	3.01	3.08	3.16
1.3	2.71	2.68	2.7	2.73	2.79	2.84	2.92	2.97	3.06	3.12
1.4	2.71	2.69	2.7	2.73	2.78	2.83	2.9	2.96	3.03	3.1
1.5	2.72	2.69	2.7	2.71	2.76	2.81	2.86	2.94	2.99	3.08
1.6	2.73	2.69	2.69	2.7	2.74	2.79	2.84	2.91	2.97	3.04
1.7	2.73	2.7	2.69	2.7	2.73	2.77	2.82	2.87	2.94	2.99
1.8	2.74	2.7	2.69	2.7	2.71	2.75	2.8	2.84	2.91	2.97
1.9	2.75	2.71	2.7	2.7	2.71	2.74	2.77	2.81	2.85	2.93
2	2.75	2.71	2.7	2.7	2.71	2.72	2.76	2.78	2.83	2.86
2.1	2.76	2.72	2.71	2.7	2.71	2.72	2.73	2.77	2.79	2.84
2.2	2.77	2.73	2.71	2.71	2.71	2.71	2.73	2.74	2.78	2.8
2.3	2.78	2.74	2.73	2.72	2.71	2.71	2.72	2.74	2.75	2.78
2.4	2.78	2.76	2.74	2.73	2.72	2.72	2.72	2.73	2.75	2.76
2.5	2.79	2.77	2.76	2.75	2.74	2.74	2.74	2.74	2.75	2.76
2.6	2.8	2.79	2.78	2.77	2.76	2.76	2.76	2.76	2.76	2.76
2.7	2.8	2.8	2.8	2.8	2.79	2.79	2.79	2.79	2.79	2.79

TAB. B.3 – Tableau des bandes à -6dB (en bins) - Seconde famille de fenêtres à la limite d'existence d'inflexions

Bande équivalente de bruit.

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	2.06	2.07	2.14	2.2	2.26	2.33	2.4	2.46	2.53	2.6
0.2	2.06	2.07	2.13	2.19	2.25	2.32	2.38	2.45	2.52	2.59
0.3	2.06	2.06	2.13	2.18	2.24	2.31	2.36	2.44	2.51	2.56
0.4	2.06	2.06	2.12	2.16	2.24	2.3	2.35	2.43	2.49	2.55
0.5	2.06	2.06	2.11	2.16	2.23	2.28	2.35	2.41	2.47	2.54
0.6	2.07	2.06	2.11	2.15	2.22	2.26	2.34	2.4	2.46	2.53
0.7	2.07	2.06	2.1	2.15	2.21	2.26	2.32	2.37	2.45	2.51
0.8	2.07	2.06	2.09	2.14	2.19	2.25	2.31	2.36	2.43	2.49
0.9	2.07	2.06	2.08	2.14	2.18	2.24	2.29	2.35	2.41	2.47
1	2.07	2.06	2.08	2.13	2.16	2.23	2.27	2.34	2.39	2.45
1.1	2.08	2.06	2.07	2.12	2.16	2.21	2.26	2.32	2.37	2.44
1.2	2.08	2.06	2.07	2.11	2.15	2.2	2.25	2.3	2.36	2.42
1.3	2.08	2.06	2.07	2.1	2.14	2.17	2.24	2.27	2.34	2.39
1.4	2.09	2.06	2.07	2.09	2.13	2.16	2.22	2.26	2.32	2.37
1.5	2.09	2.07	2.07	2.08	2.11	2.15	2.19	2.25	2.29	2.35
1.6	2.1	2.07	2.07	2.07	2.1	2.14	2.17	2.23	2.27	2.33
1.7	2.1	2.07	2.07	2.07	2.1	2.12	2.16	2.2	2.25	2.29
1.8	2.11	2.08	2.07	2.07	2.08	2.11	2.14	2.17	2.23	2.27
1.9	2.11	2.08	2.07	2.07	2.08	2.1	2.12	2.16	2.18	2.25
2	2.12	2.09	2.07	2.07	2.08	2.09	2.12	2.13	2.17	2.2
2.1	2.13	2.09	2.08	2.08	2.08	2.09	2.09	2.12	2.14	2.18
2.2	2.13	2.1	2.09	2.08	2.08	2.08	2.1	2.1	2.13	2.15
2.3	2.14	2.11	2.1	2.09	2.09	2.09	2.09	2.1	2.11	2.14
2.4	2.15	2.13	2.11	2.1	2.1	2.09	2.1	2.1	2.11	2.12
2.5	2.15	2.14	2.12	2.11	2.11	2.11	2.11	2.11	2.11	2.12
2.6	2.16	2.15	2.14	2.13	2.13	2.13	2.12	2.12	2.13	2.13
2.7	2.16	2.16	2.16	2.16	2.15	2.15	2.15	2.15	2.15	2.15

TAB. B.4 – *Tableau des bandes équivalentes de bruit (en bins) - Seconde famille de fenêtres à la limite d'existence d'inflexions*

Taux des durées où se concentre 99% de l'énergie (en temporel).

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	53.05	52.49	50.76	49.21	47.92	46.32	44.96	43.83	42.45	41.26
0.2	53.03	52.61	50.9	49.45	48.06	46.52	45.26	44.01	42.68	41.56
0.3	53	52.75	51.05	49.73	48.22	46.76	45.65	44.2	42.95	41.95
0.4	52.98	52.79	51.21	50.06	48.4	47.04	45.84	44.43	43.27	42.14
0.5	52.96	52.79	51.39	50.21	48.6	47.38	46.04	44.7	43.7	42.36
0.6	52.94	52.8	51.59	50.34	48.84	47.78	46.25	45.03	43.89	42.62
0.7	52.9	52.82	51.82	50.5	49.11	47.95	46.52	45.47	44.13	42.95
0.8	52.85	52.84	52.07	50.68	49.44	48.15	46.83	45.75	44.4	43.37
0.9	52.81	52.86	52.2	50.87	49.85	48.38	47.24	45.98	44.75	43.74
1	52.76	52.91	52.36	51.1	50.13	48.66	47.71	46.27	45.21	44
1.1	52.7	52.96	52.57	51.35	50.31	49	47.93	46.61	45.64	44.31
1.2	52.65	53	52.68	51.66	50.53	49.43	48.2	47.07	45.91	44.72
1.3	52.58	53	52.69	51.92	50.78	49.98	48.52	47.65	46.24	45.32
1.4	52.5	52.96	52.71	52.1	51.09	50.2	48.93	47.92	46.69	45.68
1.5	52.41	52.91	52.75	52.36	51.48	50.48	49.53	48.26	47.38	46.04
1.6	52.31	52.86	52.81	52.57	51.75	50.82	50.04	48.72	47.79	46.55
1.7	52.2	52.82	52.86	52.58	52	51.27	50.37	49.44	48.2	47.43
1.8	52.09	52.73	52.88	52.62	52.39	51.59	50.8	50.04	48.79	47.83
1.9	51.96	52.63	52.81	52.71	52.45	51.88	51.31	50.48	49.81	48.38
2	51.81	52.51	52.75	52.73	52.5	52.31	51.57	51.1	50.28	49.58
2.1	51.67	52.37	52.65	52.72	52.6	52.35	52.15	51.36	50.95	50.02
2.2	51.53	52.18	52.5	52.63	52.61	52.46	52.16	51.95	51.25	50.83
2.3	51.39	51.97	52.31	52.48	52.51	52.47	52.31	52.01	51.82	51.16
2.4	51.26	51.71	52.06	52.25	52.35	52.34	52.31	52.13	51.92	51.6
2.5	51.14	51.44	51.74	51.94	52.06	52.11	52.1	52.05	51.92	51.75
2.6	51.02	51.19	51.35	51.51	51.64	51.72	51.75	51.74	51.7	51.63
2.7	50.91	50.9	51.01	51.05	51.08	51.12	51.16	51.19	51.19	51.2

TAB. B.5 – Tableau des taux de durée où se concentre 99% de l'énergie (en %) - Seconde famille de fenêtres à la limite d'existence d'inflexions

Annexe C

Caractéristiques des fenêtres de la seconde famille, approchant au mieux une inflexion fixée

Paramètres (a, b, c) .

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	20.863	9.49	5.703	3.827	3.005	2.617	2.224	2.031	1.795	1.686
0.2	19.948	9.079	5.46	3.669	2.828	2.409	2.1	1.869	1.7	1.511
0.3	19.033	8.669	5.218	3.508	2.649	2.165	1.975	1.705	1.603	1.429
0.4	18.117	8.257	4.975	3.348	2.469	2.023	1.848	1.6	1.494	1.347
0.5	17.202	7.845	4.731	3.187	2.287	1.88	1.669	1.494	1.309	1.263
0.6	16.285	7.432	4.487	3.026	2.166	1.735	1.461	1.385	1.219	1.176
0.7	15.368	7.02	4.241	2.864	2.052	1.587	1.342	1.254	1.127	1.001
0.8	14.451	6.605	3.995	2.701	1.937	1.442	1.221	1.05	1.032	0.921
0.9	13.532	6.19	3.748	2.537	1.822	1.358	1.097	0.949	0.906	0.84
1	12.613	5.775	3.5	2.372	1.706	1.273	0.973	0.844	0.741	0.754
1.1	11.695	5.359	3.252	2.207	1.59	1.188	0.909	0.737	0.651	0.579
1.2	10.775	4.941	3.003	2.041	1.472	1.101	0.844	0.66	0.558	0.5
1.3	9.855	4.523	2.752	1.873	1.354	1.014	0.779	0.609	0.483	0.417
1.4	8.934	4.104	2.501	1.705	1.234	0.927	0.713	0.559	0.444	0.356
1.5	8.013	3.685	2.248	1.536	1.114	0.838	0.647	0.507	0.403	0.324
1.6	7.09	3.264	1.995	1.365	0.992	0.749	0.579	0.455	0.363	0.292
1.7	6.168	2.842	1.74	1.193	0.869	0.657	0.51	0.403	0.322	0.26
1.8	5.245	2.42	1.484	1.02	0.745	0.566	0.44	0.348	0.279	0.226
1.9	4.301	1.996	1.226	0.845	0.62	0.472	0.368	0.293	0.236	0.192
2	3.32	1.571	0.967	0.669	0.492	0.377	0.296	0.236	0.191	0.156
2.1	2.337	1.144	0.706	0.49	0.362	0.279	0.22	0.177	0.145	0.119
2.2	1.354	0.67	0.443	0.309	0.229	0.178	0.142	0.116	0.095	0.08
2.3	0.368	0.183	0.122	0.093	0.075	0.065	0.059	0.049	0.042	0.036

TAB. C.1 – Tableau des paramètres b limites - Seconde famille de fenêtres approchant au mieux une inflexion fixée

Bande passante à -3dB.

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	1.79	1.76	1.73	1.69	1.72	1.79	1.82	1.9	1.93	2.01
0.2	1.79	1.76	1.73	1.7	1.71	1.77	1.82	1.87	1.92	1.95
0.3	1.79	1.77	1.74	1.7	1.71	1.73	1.81	1.84	1.91	1.94
0.4	1.79	1.77	1.74	1.71	1.7	1.72	1.8	1.83	1.9	1.93
0.5	1.79	1.77	1.75	1.72	1.69	1.72	1.77	1.82	1.84	1.92
0.6	1.8	1.77	1.75	1.72	1.7	1.71	1.73	1.81	1.83	1.91
0.7	1.8	1.78	1.76	1.73	1.71	1.7	1.72	1.79	1.82	1.85
0.8	1.8	1.78	1.76	1.74	1.71	1.69	1.71	1.73	1.81	1.83
0.9	1.8	1.78	1.76	1.74	1.72	1.7	1.7	1.72	1.78	1.82
1	1.8	1.78	1.77	1.75	1.73	1.71	1.69	1.71	1.73	1.81
1.1	1.8	1.79	1.77	1.76	1.74	1.72	1.7	1.7	1.72	1.74
1.2	1.8	1.79	1.78	1.76	1.75	1.73	1.71	1.7	1.7	1.72
1.3	1.8	1.79	1.78	1.77	1.75	1.74	1.73	1.71	1.7	1.71
1.4	1.8	1.79	1.78	1.77	1.76	1.75	1.74	1.72	1.71	1.7
1.5	1.8	1.8	1.79	1.78	1.77	1.76	1.75	1.74	1.73	1.72
1.6	1.81	1.8	1.79	1.78	1.77	1.77	1.76	1.75	1.74	1.73
1.7	1.81	1.8	1.79	1.79	1.78	1.77	1.77	1.76	1.75	1.75
1.8	1.81	1.8	1.8	1.79	1.79	1.78	1.78	1.77	1.76	1.76
1.9	1.81	1.8	1.8	1.8	1.79	1.79	1.78	1.78	1.78	1.77
2	1.81	1.81	1.8	1.8	1.8	1.8	1.79	1.79	1.79	1.78
2.1	1.8	1.81	1.81	1.8	1.8	1.8	1.8	1.8	1.8	1.8
2.2	1.8	1.81	1.81	1.81	1.81	1.81	1.81	1.81	1.81	1.81
2.3	1.8	1.8	1.8	1.8	1.81	1.81	1.81	1.81	1.81	1.81

TAB. C.2 – Tableau des bandes à -3dB (en bins) - Seconde famille de fenêtres approchant au mieux une inflexion fixée

Bande passante à -6dB.

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	2.53	2.48	2.43	2.37	2.41	2.51	2.56	2.66	2.71	2.82
0.2	2.54	2.49	2.44	2.38	2.4	2.48	2.55	2.62	2.7	2.74
0.3	2.54	2.49	2.44	2.39	2.39	2.43	2.54	2.57	2.68	2.72
0.4	2.54	2.5	2.45	2.4	2.38	2.42	2.52	2.56	2.66	2.71
0.5	2.54	2.5	2.46	2.41	2.37	2.41	2.48	2.55	2.58	2.7
0.6	2.54	2.51	2.47	2.43	2.38	2.4	2.43	2.54	2.57	2.68
0.7	2.55	2.51	2.47	2.44	2.4	2.39	2.42	2.51	2.56	2.59
0.8	2.55	2.52	2.48	2.45	2.41	2.38	2.4	2.43	2.54	2.57
0.9	2.55	2.52	2.49	2.46	2.42	2.39	2.39	2.42	2.5	2.56
1	2.55	2.53	2.5	2.47	2.44	2.4	2.38	2.4	2.43	2.54
1.1	2.55	2.53	2.5	2.48	2.45	2.42	2.39	2.39	2.41	2.44
1.2	2.56	2.53	2.51	2.49	2.46	2.43	2.41	2.39	2.39	2.42
1.3	2.56	2.54	2.52	2.5	2.47	2.45	2.43	2.41	2.39	2.39
1.4	2.56	2.54	2.52	2.5	2.48	2.46	2.44	2.43	2.41	2.39
1.5	2.56	2.55	2.53	2.51	2.5	2.48	2.46	2.44	2.43	2.41
1.6	2.56	2.55	2.54	2.52	2.51	2.49	2.48	2.46	2.45	2.43
1.7	2.56	2.55	2.54	2.53	2.52	2.51	2.49	2.48	2.47	2.46
1.8	2.57	2.56	2.55	2.54	2.53	2.52	2.51	2.5	2.49	2.48
1.9	2.57	2.56	2.55	2.55	2.54	2.53	2.52	2.52	2.51	2.5
2	2.56	2.56	2.56	2.55	2.55	2.54	2.54	2.53	2.53	2.52
2.1	2.56	2.57	2.56	2.56	2.56	2.55	2.55	2.55	2.55	2.54
2.2	2.56	2.56	2.57	2.57	2.56	2.56	2.56	2.56	2.56	2.56
2.3	2.56	2.56	2.56	2.56	2.56	2.57	2.57	2.57	2.57	2.57

TAB. C.3 – Tableau des bandes à -6dB (en bins) - Seconde famille de fenêtres approchant au mieux une inflexion fixée

Bande équivalente de bruit.

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	1.94	1.89	1.85	1.8	1.83	1.9	1.94	2.02	2.05	2.13
0.2	1.94	1.9	1.85	1.81	1.82	1.88	1.93	1.98	2.04	2.07
0.3	1.94	1.9	1.86	1.82	1.81	1.84	1.92	1.95	2.03	2.06
0.4	1.95	1.91	1.87	1.83	1.81	1.83	1.91	1.94	2.02	2.05
0.5	1.95	1.91	1.87	1.84	1.8	1.83	1.88	1.93	1.96	2.04
0.6	1.95	1.92	1.88	1.84	1.81	1.82	1.84	1.92	1.95	2.03
0.7	1.95	1.92	1.89	1.85	1.82	1.81	1.83	1.9	1.94	1.96
0.8	1.95	1.92	1.89	1.86	1.83	1.8	1.82	1.84	1.93	1.95
0.9	1.95	1.93	1.9	1.87	1.84	1.81	1.81	1.83	1.9	1.94
1	1.96	1.93	1.91	1.88	1.85	1.83	1.8	1.82	1.84	1.92
1.1	1.96	1.94	1.91	1.89	1.86	1.84	1.82	1.81	1.83	1.85
1.2	1.96	1.94	1.92	1.9	1.87	1.85	1.83	1.81	1.82	1.83
1.3	1.96	1.94	1.92	1.91	1.89	1.87	1.85	1.83	1.81	1.82
1.4	1.96	1.95	1.93	1.91	1.9	1.88	1.86	1.85	1.83	1.81
1.5	1.96	1.95	1.94	1.92	1.91	1.89	1.88	1.86	1.85	1.83
1.6	1.97	1.95	1.94	1.93	1.92	1.9	1.89	1.88	1.87	1.85
1.7	1.97	1.96	1.95	1.94	1.93	1.91	1.9	1.89	1.88	1.87
1.8	1.97	1.96	1.95	1.94	1.94	1.93	1.92	1.91	1.9	1.89
1.9	1.97	1.96	1.96	1.95	1.94	1.94	1.93	1.92	1.92	1.91
2	1.97	1.97	1.96	1.96	1.95	1.95	1.94	1.94	1.93	1.93
2.1	1.97	1.97	1.97	1.96	1.96	1.96	1.96	1.95	1.95	1.95
2.2	1.97	1.97	1.97	1.97	1.97	1.97	1.97	1.96	1.96	1.96
2.3	1.96	1.96	1.97	1.97	1.97	1.97	1.97	1.97	1.97	1.97

TAB. C.4 – *Tableau des bandes équivalentes de bruit (en bins) - Seconde famille de fenêtres approchant au mieux une inflexion fixée*

Taux des durées où se concentre 99% de l'énergie (en temporel).

a \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	56	56.9	58	59.1	58	55.7	54.6	52.4	51.4	49.5
0.2	56	56.9	57.8	58.9	58.3	56.5	54.8	53.3	51.7	50.9
0.3	56	56.8	57.7	58.7	58.5	57.6	55.1	54.2	52	51.2
0.4	55.9	56.7	57.5	58.5	58.8	57.9	55.4	54.5	52.5	51.5
0.5	55.9	56.6	57.4	58.3	59.1	58.1	56.3	54.8	54	51.8
0.6	55.9	56.5	57.3	58.1	58.9	58.4	57.6	55.2	54.3	52.2
0.7	55.8	56.4	57.1	57.9	58.6	58.8	57.9	55.9	54.7	54
0.8	55.8	56.3	57	57.7	58.4	59.1	58.3	57.6	55.1	54.3
0.9	55.8	56.3	56.8	57.5	58.1	58.8	58.7	57.9	56	54.7
1	55.7	56.2	56.7	57.3	57.9	58.5	59.1	58.4	57.7	55.3
1.1	55.7	56.1	56.6	57.1	57.7	58.2	58.8	58.9	58.2	57.5
1.2	55.7	56	56.5	56.9	57.4	57.9	58.4	58.9	58.7	58
1.3	55.6	56	56.3	56.8	57.2	57.6	58.1	58.5	59	58.7
1.4	55.6	55.9	56.2	56.6	57	57.4	57.8	58.1	58.5	58.9
1.5	55.6	55.8	56.1	56.4	56.7	57.1	57.4	57.8	58.1	58.5
1.6	55.5	55.8	56	56.3	56.5	56.8	57.1	57.4	57.7	58
1.7	55.5	55.7	55.9	56.1	56.3	56.6	56.8	57	57.3	57.5
1.8	55.5	55.6	55.8	56	56.1	56.3	56.5	56.7	56.9	57.1
1.9	55.5	55.6	55.7	55.8	55.9	56.1	56.2	56.4	56.5	56.7
2	55.5	55.5	55.6	55.7	55.8	55.9	56	56.1	56.2	56.3
2.1	55.6	55.5	55.5	55.6	55.6	55.7	55.7	55.8	55.8	55.9
2.2	55.6	55.6	55.5	55.5	55.5	55.5	55.5	55.5	55.6	55.6
2.3	55.7	55.7	55.6	55.6	55.6	55.5	55.4	55.4	55.4	55.3

TAB. C.5 – Tableau des taux de durée où se concentre 99% de l'énergie (en %) - Seconde famille de fenêtres approchant au mieux une inflexion fixée

Bibliographie

- [BM95] A. Bateman and P. Matri, *Identification of nonstationary audio signals using the fft, with application to analysis-based synthesis of sound*, IEE-95 Colloquium on Audio Engineering, May 1995.
- [DGR93] P. Depalle, G. Garcia, and X. Rodet, *Tracking of partials using hidden markov models*, IEEE ICASSP-93, Minneapolis, Minesota, April 1993.
- [DR92] P. Depalle and X. Rodet, *A new additive synthesis method using inverse fourier transform*, Proc. Int. Computer Music Conf. (ICMC'92), October 1992, pp. 410–411.
- [Gab46] D. Gabor, *Theory of communication*, Journal of the IEEE No 93 Londres, 1946.
- [Har78] Frederic J. Harris, *On the use of windows for harmonic analysis with the discrete fourier transform*, IEEE-78, vol. 66, January 1978, pp. 51–83.
- [LH74] Ch. L. Lawson and R. J. Hanson, *Solving least squares problems*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.
- [MQ86] R. J. McAulay and T. F. Quatieri, *Speech analysis/synthesis based on sinusoidal representation*, IEEE-86, August 1986, pp. 744–754.
- [RS68] L. R. Rabiner and R. W. Schafer, *Digital processing of speech signals*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1968.
- [Tro95] Laurent Tromp, *Amélioration de l'extraction de partiels dans les signaux sonores*, Master's thesis, Université Paris-Sud Orsay, 1995.

