



**HAL**  
open science

# Méthodes d'analyse/synthèse et représentations optimales des sons musicaux basées sur la réduction de données spectrales

Thierry Rochebois

► **To cite this version:**

Thierry Rochebois. Méthodes d'analyse/synthèse et représentations optimales des sons musicaux basées sur la réduction de données spectrales. Traitement du signal et de l'image [eess.SP]. Université Paris XI, 1997. Français. NNT : . tel-01076386

**HAL Id: tel-01076386**

**<https://theses.hal.science/tel-01076386>**

Submitted on 21 Oct 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Orsay  
N d'ordre :

THESE

présentée pour obtenir le grade de  
Docteur de l'Université Paris XI

Spécialité : Electronique

par **Thierry Rochebois**

Sujet de la thèse :

**Méthodes d'analyse/synthèse et représentations  
optimales des sons musicaux basées sur la réduction  
de données spectrales**

soutenue le **19 décembre 1997**

devant le jury composé de :

M. Daniel Arfib	Rapporteur
M. Antoine Chaigne	Rapporteur
M. Gérard Charbonneau	Directeur de thèse
M. Christophe d'Alessandro	
M. Christian Deforeit	
M. Alain Mérigot	

*à mon épouse. ma famille et mes amis...*



# Remerciements

Ce travail n'aurait pas abouti sans la participation directe, indirecte, volontaire ou involontaire de nombreuses personnes. Je tiens réellement à les en remercier ici.

Je remercie, tout d'abord, Gérard Charbonneau de m'avoir proposé un sujet de thèse qui n'a eu de cesse d'évoluer. Son grand mérite est de m'avoir suivi et supporté durant mes cogitations.

Je remercie, par ordre alphabétique, tous les membres du jury : Daniel Arfib pour son enthousiasme, Antoine Chaigne pour sa rigueur et l'intérêt qu'il a su porter à mes travaux, Christophe d'Alessandro qui m'a donné le point de vue du traicteur de signal et du musicien. Christian Deforet dont les circuits intégrés hantent mon ordianteur personnel, et Alain Mérigot qui a accueilli notre activité au sein d'AXIS sans *a priori*.

Je remercie Christophe Hourdin qui par son acharnement a mis sur pieds une activité naissante. Je remercie Edith Léautey pour son travail, sa sincérité et son sens positif des réalités.

Je remercie mon ami François Le Coat pour les discussions acharnées où se mêlaient sons et paires stéréographiques. Je remercie également tous les "thésards" de l'IEF qui ont subi mes discussions acoustico-électroniques.

Je remercie tous les membres du département AXIS et de l'ex-département ex-TSI qui ont su supporter une activité aussi marginale que la nôtre.

Je remercie tous ceux qui animent le 220 tant en ce qui concerne la recherche que l'enseignement. Je remercie tout particulièrement Annick Giordano pour son efficacité et son aptitude rare à "arranger les choses". Je remercie Claudine Falcetta, Jorge Batista, Cédric Clerc, Elisabeth Bouyssy, Michel Fan et Samir Bouaziz pour leur agréable compagnie. Je remercie également Huguette Maréchal pour son don unique à multiplier les mémoires de thèses.

Je remercie mes étudiants de DEA, de DESS, de FIUPSO, de Licence EEA et de DUT de Chimie pour leurs questions et l'intérêt qu'ils ont daigné porter à mes cours.

Enfin, je tiens à remercier tous ceux qui m'entourent et me supportent : mon épouse, ma famille et mes amis.

*Il semble que la perfection soit atteinte  
non quand il n'y a plus rien à ajouter  
mais quand il n'y a plus rien à retrancher.  
Au terme de son évolution la machine se dissimule.*

Antoine de Saint-Exupéry

*Terre des hommes*

# Table des matières

<b>1</b>	<b>Etat de l'art</b>	<b>17</b>
1.1	Historique des techniques de synthèse . . . . .	18
1.1.1	L'ère électromécanique . . . . .	18
1.1.2	L'ère radioélectrique . . . . .	21
1.1.3	L'ère analogique . . . . .	23
1.1.4	L'ère numérique . . . . .	26
1.1.5	Conclusion . . . . .	29
1.2	Inventaire des méthodes de synthèse . . . . .	29
1.2.1	Synthèse additive . . . . .	29
1.2.2	Synthèse par tables d'ondes . . . . .	34
1.2.3	Synthèse par tables d'ondes multiples . . . . .	35
1.2.4	Synthèse soustractive . . . . .	36
1.2.5	Synthèse par modulation de fréquence . . . . .	38
1.2.6	Synthèse par formes d'ondes formantiques . . . . .	38
1.2.7	Synthèse granulaire . . . . .	40
1.2.8	Synthèse par distorsion ( <i>wave-shaping</i> ) . . . . .	41
1.2.9	Synthèse par formule de sommation discrète . . . . .	42
1.2.10	Synthèse par modélisation physique . . . . .	43
1.2.11	Implantations et processeurs . . . . .	43
1.2.12	Conclusions . . . . .	46
1.3	Quelques aspects de l'audition humaine . . . . .	46
1.3.1	Quelques aspects physiologiques de l'audition . . . . .	46
1.3.2	Quelques aspects psychoacoustiques de l'audition . . . . .	49
1.3.3	Conclusions . . . . .	56
1.4	Inventaire des méthodes d'analyse . . . . .	56
1.4.1	Filtrage hétérodyne . . . . .	56
1.4.2	Poursuite de pics de transformée de Fourier . . . . .	58
1.4.3	Modèle déterministe plus stochastique . . . . .	59
1.4.4	Méthode de Prony . . . . .	59
1.4.5	Bancs de filtres à facteur de qualité constant . . . . .	60
1.4.6	Transformée en ondelettes . . . . .	61
1.4.7	Résumé et conclusions . . . . .	61

<b>2</b>	<b>Modélisation par ACP de spectres réels</b>	<b>63</b>
2.1	Analyse spectrale de la composante harmonique . . . . .	63
2.1.1	Interpolation de transformée de Fourier discrète à court terme . . . . .	64
2.1.2	Effets de la troncature de la fenêtre . . . . .	66
2.1.3	Première sélection de la composante déterministe . . . . .	67
2.1.4	Poursuite des pics et resynthèse . . . . .	69
2.1.5	Sélection de la composante harmonique . . . . .	71
2.1.6	Conclusions . . . . .	72
2.2	Réduction des données spectrales par ACP . . . . .	72
2.2.1	Introduction aux méthodes d'analyse des données . . . . .	72
2.2.2	Analyse en composantes principales des données spectrales . . . . .	74
2.2.3	Interprétation géométrique . . . . .	75
2.2.4	Analyse de sons synthétiques . . . . .	77
2.2.5	Conclusions . . . . .	84
2.3	Application à l'étude des sons musicaux . . . . .	85
2.3.1	Analyse des sons de Grey . . . . .	85
2.3.2	Analyse et étude de quatre sons musicaux . . . . .	88
2.3.3	Représentation stéréoscopique des trajectoires . . . . .	95
2.3.4	Etude d'un <i>corpus</i> de 42 sons d'instruments acoustiques . . . . .	98
2.4	Application à la synthèse et à l'édition . . . . .	104
2.4.1	Procédé de reconstruction . . . . .	104
2.4.2	Dilatation temporelle et transposition . . . . .	104
2.4.3	Synthèse croisée . . . . .	110
2.4.4	Vers un modèle expressif . . . . .	111
2.4.5	Résumé . . . . .	112
2.5	Conclusions . . . . .	112
<b>3</b>	<b>Extensions du modèle</b>	<b>113</b>
3.1	ACP de spectres harmoniques complexes . . . . .	113
3.1.1	Phases et amplitudes complexes . . . . .	114
3.1.2	Analyse en composantes principales de $\mathcal{A}$ . . . . .	114
3.1.3	Sous-espace et trajectoire complexes . . . . .	115
3.1.4	Reconstruction par combinaison de tables d'ondes "complexes" . . . . .	115
3.1.5	Analyse de sons synthétiques . . . . .	117
3.1.6	Analyse de quatre sons musicaux . . . . .	121
3.1.7	Conclusions . . . . .	125
3.2	Etude de la composante bruitée . . . . .	125
3.2.1	Extraction de la composante bruitée . . . . .	126
3.2.2	Analyse de la composante bruitée . . . . .	127
3.2.3	Synthèse de la composante bruitée . . . . .	129
3.2.4	Conclusions . . . . .	132
3.3	Synthèse par combinaison de grains généralisée . . . . .	133
3.3.1	Dualité grain/table d'onde . . . . .	134
3.3.2	Transposition et enveloppe spectrale . . . . .	134
3.3.3	Combinaison de grains/tables d'onde . . . . .	135



3.3.4	Extension, conclusions et perspectives . . . . .	136
3.4	Conclusions et perspectives . . . . .	137
<b>A</b>	<b>Programmes</b>	<b>143</b>
A.1	Programmes d'analyse spectrale . . . . .	143
A.1.1	Analyse spectrale par FFT interpolée: <code>fftana</code> . . . . .	143
A.1.2	Visualisation: <code>visuana</code> . . . . .	143
A.1.3	Poursuite des pics: <code>trk</code> . . . . .	144
A.1.4	Détection du fondamental: <code>fnd</code> . . . . .	144
A.1.5	Détection du fondamental tenant compte de la phase: <code>fdc</code> . . . . .	144
A.1.6	Détection des harmoniques: <code>hrm</code> . . . . .	145
A.2	Programmes d'analyse en composantes principales . . . . .	145
A.2.1	Accumulation de la matrice de covariance: <code>cov</code> . . . . .	145
A.2.2	Accumulation de la matrice de covariance complexe: <code>covc</code> . . . . .	146
A.2.3	Analyse en composantes principales: <code>acp</code> . . . . .	146
A.2.4	Analyse en composantes principales complexe: <code>acpc</code> . . . . .	147
A.2.5	Projection dans le sous-espace $\mathcal{P}$ : <code>prj</code> . . . . .	147
A.2.6	Projection dans le sous-espace complexe $\mathcal{P}$ : <code>prjc</code> . . . . .	147
A.3	Programmes de synthèse . . . . .	148
A.3.1	Synthèse additive: <code>synd</code> . . . . .	148
A.3.2	Synthèse additive: <code>syn</code> . . . . .	148
A.3.3	Reconstruction à partir des données réduites: <code>rec</code> . . . . .	149
A.3.4	Reconstruction à partir des données réduites complexes: <code>recc</code> . . . . .	149
<b>B</b>	<b>Fiches de données réduites de 42 sons</b>	<b>151</b>
	<b>Index</b>	<b>212</b>

# Table des figures

1.1	Clavecin électrique de Jean-Baptiste de La Borde (1759)	18
1.2	Télégraphe harmonique de Gray (1874)	19
1.3	Flammes manométriques de Tyndall (1875)	20
1.4	Le Telharmonium de Thaddeus Cahill (1906)	20
1.5	Orgue radioélectrique Givelet Coupleux (1933)	22
1.6	Synthétiseur modulaire Moog (1964)	24
1.7	Synthétiseur analogique polyphonique Oberheim (1977)	25
1.8	Synthétiseur analogique polyphonique (1978)	26
1.9	Synthétiseur/échantillonneur Fairlight (1979)	27
1.10	Synthèse additive par banc d'oscillateurs	30
1.11	Oscillateur numérique à lecture de table	31
1.12	Synthèse additive par transformée de Fourier inverse	32
1.13	Transposition et repliement de spectre	34
1.14	Formes de synthèse par tables d'ondes multiples	35
1.15	Synthèse soustractive	37
1.16	Synthèse par modulation de fréquence	37
1.17	Synthèse par formes d'ondes formantiques	39
1.18	Différents types de grains	40
1.19	Waveshaping par un polynôme de Tchébychev	42
1.20	Architecture du DSP 56001 Motorola	44
1.21	Cochlée déroulée (d'après E.Leipp [Lei84])	47
1.22	Organe de Corti (d'après R.Pujol [Puj90])	48
1.23	Processus temporel actif de l'audition	48
1.24	Courbes isosoniques pour l'écoute binaurale en champ libre, d'après la norme AFNOR NFS30-003	49
1.25	Reconstruction géométrique à partir d'un tableau de distances	52
1.26	Espace de Grey	52
1.27	Plan XY de l'espace de Grey	54
1.28	Plan ZY de l'espace de Grey	55
1.29	filtre hétérodyne simple	56
1.30	filtre hétérodyne adaptatif	56
1.31	Poursuite de pics de transformées de Fourier	59
1.32	Transformée de Fourier et banc de filtres à Q constant	60
2.1	Interpolation de la transformée de Fourier	65
2.2	Transformées de Fourier de gaussiennes tronquées	67

2.3	Amplitude du premier lobe secondaire en fonction de $\alpha$ . . . . .	68
2.4	Seuillage des maxima locaux du spectre . . . . .	69
2.5	Interpolation cubique de la phase (d'après McAulay et Quatieri [MQ86]) . . . . .	70
2.6	Sélection de la composante harmonique . . . . .	71
2.7	Projections d'un objet à trois dimensions dans le plan . . . . .	74
2.8	Construction de la trajectoire harmonique . . . . .	75
2.9	Axes et trajectoire . . . . .	76
2.10	Données spectrales réduites du son <code>test1</code> non renormalisé . . . . .	78
2.11	Données spectrales réduites du son <code>test1</code> non renormalisé à enveloppe constante . . . . .	79
2.12	Données spectrales réduites du son <code>test1</code> renormalisé . . . . .	80
2.13	Données spectrales réduites du son FM <code>test2</code> . . . . .	81
2.14	Données spectrales réduites du son FM <code>test3</code> . . . . .	82
2.15	Données spectrales réduites du son FM <code>test4</code> . . . . .	83
2.16	Procédé d'analyse et de réduction de données . . . . .	84
2.17	Plan XY de l'espace de Grey et enveloppe $V_1(t)$ . . . . .	85
2.18	Plan XY de l'espace de Grey et spectre $P_1$ . . . . .	86
2.19	Plan ZY de l'espace de Grey et spectre $P_2$ . . . . .	87
2.20	Plan ZY de l'espace de Grey et enveloppe $V_2(t)$ . . . . .	87
2.21	Représentation spectrale en chutes d'eau . . . . .	89
2.22	Données spectrales réduites d'un son de piano . . . . .	91
2.23	Données spectrales réduites d'un son de hautbois . . . . .	92
2.24	Données spectrales réduites d'un son de flûte . . . . .	93
2.25	Données spectrales réduites d'un son de violoncelle . . . . .	94
2.26	Paire stéréographique de la trajectoire de la figure 2.9 . . . . .	95
2.27	Visualisation de la paire stéréographique . . . . .	96
2.28	Représentation stéréographique de la trajectoire du piano . . . . .	97
2.29	Représentation stéréographique de la trajectoire du hautbois . . . . .	98
2.30	Représentation stéréographique de la trajectoire de la flûte . . . . .	99
2.31	Représentation stéréographique de la trajectoire du violoncelle . . . . .	101
2.32	Spectres $P_1^{42}$ à $P_5^{42}$ . . . . .	101
2.33	Modèle simpliste d'instrument à cordes frottées . . . . .	102
2.34	Courbe paramétrique: $V_1(t), V_2(t)$ . . . . .	105
2.35	Courbe paramétrique: $V_1(t), V_3(t)$ . . . . .	106
2.36	Courbe paramétrique: $f_1(t), V_1(t)$ . . . . .	107
2.37	Courbe paramétrique: $f_1(t), V_2(t)$ . . . . .	108
2.38	Reconstruction par combinaison de tables d'ondes . . . . .	109
2.39	Interversion de trajectoires . . . . .	110
2.40	Synthèse de sons expressifs . . . . .	111
3.1	Spectre harmonique complexe . . . . .	115
3.2	Synthèse par combinaison de spectres harmoniques complexes . . . . .	116
3.3	Synthèse par combinaison de tables d'ondes "complexes" . . . . .	116
3.4	Son de profil spectral constant (phases relatives nulles) . . . . .	117
3.5	Son de profil spectral constant (phases relatives non nulles) . . . . .	117

3.6	Inharmonicités représentées sous forme de rotation de phases . . . . .	118
3.7	Battements représentés sous forme de modulation de phases . . . . .	118
3.8	Données réduites complexes d'un son FM (I variant de 4 à 0) . . . . .	119
3.9	Données réduites complexes d'un son FM (I variant de 8 à 0) . . . . .	120
3.10	Données réduites complexes d'un son FM présentant des battements . . . . .	120
3.11	Données réduites complexes d'un son de hautbois . . . . .	122
3.12	Données réduites complexes d'un son de piano . . . . .	123
3.13	Données réduites complexes d'un son de flûte . . . . .	124
3.14	Données réduites complexes d'un son de violoncelle . . . . .	124
3.15	Extraction de la composante bruitée à trois stades de l'analyse . . . . .	126
3.16	Analyse de la composante bruitée par transformées de Fourier à court terme	128
3.17	Synthèse de la composante bruitée par combinaison de tables . . . . .	129
3.18	Synthèse de la composante bruitée par combinaison de tables et modulation d'amplitude . . . . .	130
3.19	Synthèse de la composante bruitée par combinaison de grains à temps aléatoire . . . . .	131
3.20	Interprétation granulaire de la transposition d'une table d'onde . . . . .	133
3.21	Transposition granulaire d'une "table d'onde" . . . . .	134
3.22	Généralisation de la synthèse par grains à temps aléatoire . . . . .	136
B.1	Clarinette en si $\flat$ . . . . .	153
B.2	Clarinette en mi $\flat$ . . . . .	154
B.3	Hautbois . . . . .	155
B.4	Hautbois baroque . . . . .	156
B.5	Orgue baroque registre plenum . . . . .	157
B.6	Orgue baroque registre prinzipal . . . . .	158
B.7	Orgue baroque registre flûtes . . . . .	159
B.8	Flûte à bec basse . . . . .	160
B.9	Flûte traversière sans vibrato . . . . .	161
B.10	Flûte traversière avec vibrato . . . . .	162
B.11	Flûte alto sans vibrato . . . . .	163
B.12	Flûte alto avec vibrato . . . . .	164
B.13	Flûte jouée "flutter" . . . . .	165
B.14	Basson . . . . .	166
B.15	Cor Anglais . . . . .	167
B.16	Cor . . . . .	168
B.17	Cor avec sourdine . . . . .	169
B.18	Cornet . . . . .	170
B.19	Trombone . . . . .	171
B.20	Trombone avec sourdine . . . . .	172
B.21	Trompette baroque . . . . .	173
B.22	Trompette en Do . . . . .	174
B.23	Tuba . . . . .	175
B.24	Celesta . . . . .	176
B.25	Marimba . . . . .	177

B.26 Piano joué mezzo forte . . . . .	178
B.27 Piano joué fortissimo . . . . .	179
B.28 Guitare classique . . . . .	180
B.29 Guitare classique jouée sul tasto . . . . .	181
B.30 Clavecin . . . . .	182
B.31 Harpe . . . . .	183
B.32 Alto . . . . .	184
B.33 Violon . . . . .	185
B.34 Violoncelle joué martele . . . . .	186
B.35 Alto joué martele . . . . .	187
B.36 Violon joué martele . . . . .	188
B.37 Violoncelle joué avec vibrato . . . . .	189
B.38 Alto joué avec vibrato . . . . .	190
B.39 Violon joué avec vibrato . . . . .	191
B.40 Violoncelle joué pizzicato . . . . .	192
B.41 Alto joué pizzicato . . . . .	193
B.42 Violon joué pizzicato . . . . .	194

# Introduction

Ce mémoire de thèse présente les travaux que nous avons mené au sein du département AXIS<sup>1</sup> de l'Institut d'Electronique Fondamentale (URA 22) au cours des trois dernières années. Notre activité de Traitement du Signal Acoustique est encore récente et peut sembler originale dans un département voué à l'étude des architectures des circuits et des systèmes. En effet, de nombreuses activités proches de la nôtre sont généralement rattachées à des laboratoires d'acoustique disposant de dispositifs d'enregistrement et de mesures acoustiques de pointe. Malgré l'absence d'installations de ce type, l'originalité de notre association à AXIS est, de notre point de vue, un point fort : à long terme, nous pourrions bénéficier du savoir faire du département AXIS en matière de conception de circuits et systèmes pour réaliser un synthétiseur temps réel. L'activité TSA du département AXIS ne doit pas être considérée comme "concurrente" mais comme complémentaire des activités du même type.

L'analyse et la synthèse des sons musicaux, thème central de ce mémoire, suppose de multiples compétences dans des domaines aussi variés que : le traitement du signal, l'algorithmie, l'informatique, l'électronique numérique, la psychoacoustique, ou encore, l'acoustique instrumentale.

Considérant la spécificité de nos travaux, nous avons pris le parti de rédiger un mémoire de thèse accessible à toute personne, spécialiste ou non, désirant s'informer sur notre activité.

Le chapitre 1, "Etat de l'art", présente les multiples facettes de la synthèse et de l'analyse des sons musicaux. On y trouve, en guise d'entrée en matière, un historique retraçant cent vingt années d'ingéniosité et de créativité. La vision "verticale" de cet historique est complétée par un bilan des types de synthèse envisageables à l'heure actuelle. Ce bilan est suivi par un exposé de quelques aspects de l'audition humaine, destinataire ultime de tout synthétiseur de sons musicaux. Un court inventaire des méthodes d'analyse spectrale termine le chapitre.

Le thème principal de ce mémoire est un procédé — ou plus exactement une famille de procédés — permettant de simplifier la représentation des sons musicaux.

L'essentiel des idées sous-jacentes à ces procédés nous ont été inspirées par la confrontation des travaux de Christophe Hourdin — qui nous a précédé — et des travaux de Andrew Horner, James Beauchamp et Lippold Haken. Le premier a utilisé *l'analyse factorielle des données* afin d'obtenir une représentation des sons musicaux sous forme de "trajectoires". Son mémoire de thèse [Hou95] et ses publications [HCM97a, HCM97b] font état d'interprétations judicieuses et d'applications à l'analyse-synthèse des sons musicaux.

---

1. Architecture et Conception de Circuits Intégrés et de Systèmes

Les seconds ont évalués plusieurs techniques permettant de réduire le coût de calcul d'algorithmes de synthèse en effectuant des réductions de données [HBH93]. Cependant, leurs conclusions concernant une méthode proche de l'analyse factorielle — *l'analyse en composantes principales (ACP)* — semblaient s'opposer aux conclusions de Christophe Hourdin. A la lumière des travaux de Christophe Hourdin — et avec un certain esprit de contradiction — nous avons repris le procédé de réduction par analyse en composantes principales là où Andrew Horner, James Beauchamp et Lippold Haken l'ont laissé.

Les enjeux des procédés de réduction de données spectrales sont importants car ils permettent non seulement de simplifier l'utilisation d'un synthétiseur mais aussi de réduire considérablement le coût de calcul des algorithmes de synthèse. Ils font l'objet de nombreuses publications depuis cinq ans [SM95, HBH93, BH95, HB95, Hor95, CH96, Hor97, YH97].

L'exposé de notre étude des applications de l'analyse en composantes principales à la réduction des données spectrales s'articule autour des chapitres 2 et 3.

Le chapitre 2 décrit l'essentiel de nos travaux. On y trouve donc la description détaillée d'un procédé d'analyse-synthèse et de réduction de données. Cette description est suivie par une discussion des possibilités offertes par ce procédé pour l'étude et la synthèse des sons musicaux.

Le chapitre 3 propose plusieurs solutions permettant de compléter le procédé décrit au chapitre 2. Il tente de répondre à plusieurs questions induites par les hypothèses simplificatrices énoncées au début du chapitre 2 : comment représenter les sons dissonants, comment considérer la composante "bruitée" de certains sons, et, comment synthétiser plus efficacement certaines classes de sons ?

Ce mémoire n'a pas pour objectif de proposer une solution universelle au problème de la représentation des sons musicaux. Son but est de suggérer et d'évaluer de nouvelles voies d'investigation qui seront étudiées plus profondément dans l'avenir.

# Chapitre 1

## Etat de l'art

*Yes, it's a proprietary scheme.  
"proprietary" is a polite word for  
"We are not going to tell you what it is"!*

Robert A. Moog [Loh85]

Cet état de l'art a pour objectif de situer nos travaux de recherche dans leur contexte. Il comporte quatre sections : une introduction historique et un inventaire des méthodes actuelles de synthèse, un court exposé des fonctions auditives, et un inventaire des méthodes d'analyse. Il a pour finalité de faciliter la lecture de la suite du mémoire. Pour ce faire, de multiples détails techniques ont été volontairement omis. Les lecteurs intéressés par certains sujets pourront se référer à la bibliographie.

La section 1.1 présente l'historique des techniques électriques et électroniques appliquées à la synthèse des sons musicaux. Cette section est le fruit d'une recherche bibliographique "historique" menée en parallèle à une recherche bibliographique plus "actuelle". L'essentiel des documents utilisés pour cette section provient des fonds anciens de bibliothèques techniques (ESE et CNET) et de brevets. Cette introduction historique peut certes sembler déconcertante mais elle permet d'avoir une vision d'ensemble d'une activité qui paraît à certains originale, voire même marginale.

La section 1.2 complète notre vision du domaine en présentant de multiples procédés de synthèse des sons musicaux envisageables à l'heure actuelle. Certaines de ces méthodes ont été déjà exploitées, d'autres — faisant l'objet de recherches — nous semblent prometteuses. Cette section nous permettra d'une part de situer nos travaux par rapport aux méthodes existantes. D'autre part, certaines idées, inspirées par certaines méthodes, nous ont servi à étendre notre modèle.

La section 1.3 introduit très brièvement quelques aspects physiologiques et psychoacoustiques de l'audition humaine. Cette section permet d'appréhender quelques phénomènes concernant la perception des sons musicaux et introduit la notion de "timbre".



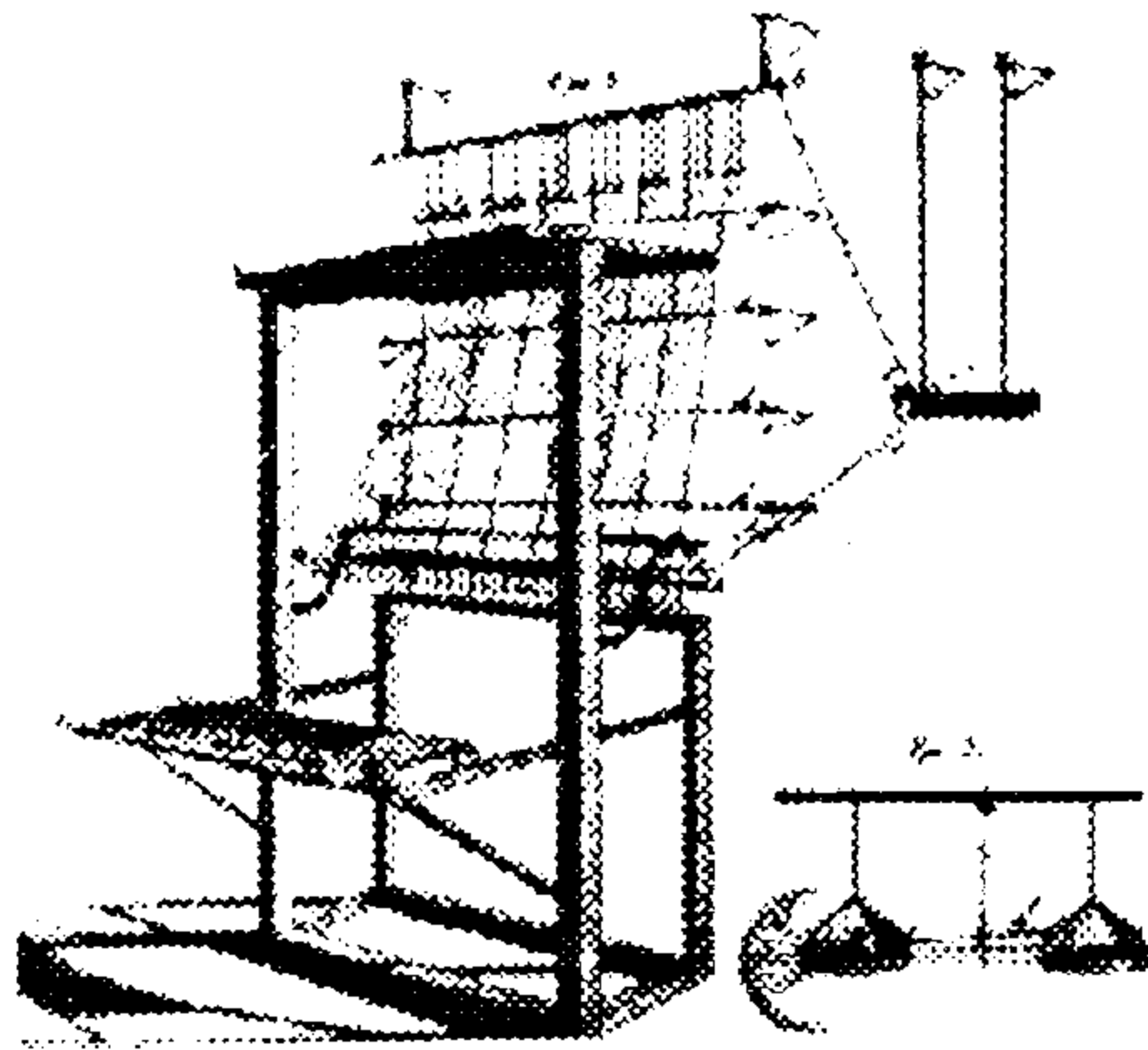


FIG. 1.1 – *Clavecin électrique de Jean-Baptiste de La Borde (1759)*

La section 1.4 aborde divers procédés d'analyse spectrale utilisables dans le cadre de l'étude des sons musicaux. La nature et la précision de ces analyses déterminent très fortement la validité des traitements effectués sur les données spectrales *a posteriori*. Cette section, bien que non exhaustive, permet de situer notre procédé d'analyse (§2.1, page 63) et d'envisager des méthodes alternatives pour l'étude de la composante "bruitée" des sons musicaux.

## 1.1 Historique des techniques de synthèse

En un peu plus d'un siècle, l'électricité, l'électronique puis l'informatique ont bouleversé notre quotidien. Ces technologies ont métamorphosé l'ensemble des activités humaines y compris l'art et en particulier la musique. Nous retraçons ici l'histoire des synthétiseurs musicaux de 1874 à nos jours. Nous abordons successivement les techniques électromécaniques, radioélectriques, analogiques et numériques.

L'histoire des synthétiseurs de musique est l'archétype de l'interaction entre une science, l'électronique, et un art, la musique. Cette interaction a été, et demeure, féconde. Elle revêt parfois quelques aspects surprenants.

### 1.1.1 L'ère électromécanique

Les tous premiers instruments à avoir utilisé l'électricité sont le "Denis d'or" du prêtre bohémien Prokop Diviš (1730) et le "Clavecin électrique" du prêtre jésuite Jean-Baptiste de La Borde (1759). Le premier était un instrument à corde dans lequel l'électricité jouait un rôle secondaire : il permettait à son inventeur d'électriser son malheureux élève en cas de fausse note... Le second, était en fait un carillon électrostatique capable de générer des sons entretenus (figure 1.1). Plus tard, à partir de 1860, de nombreuses orgues furent électrifiées. Tous ces instruments n'étaient pas à proprement parler des instruments électriques mais plutôt des instruments acoustiques électrifiés.

L'histoire du premier instrument électrique commence vers 1870. A cette époque, l'application principale de l'électricité est le télégraphe. Deux grands ingénieurs américains, Alexander Graham Bell et Elisha Gray se font une concurrence acharnée. Le premier est

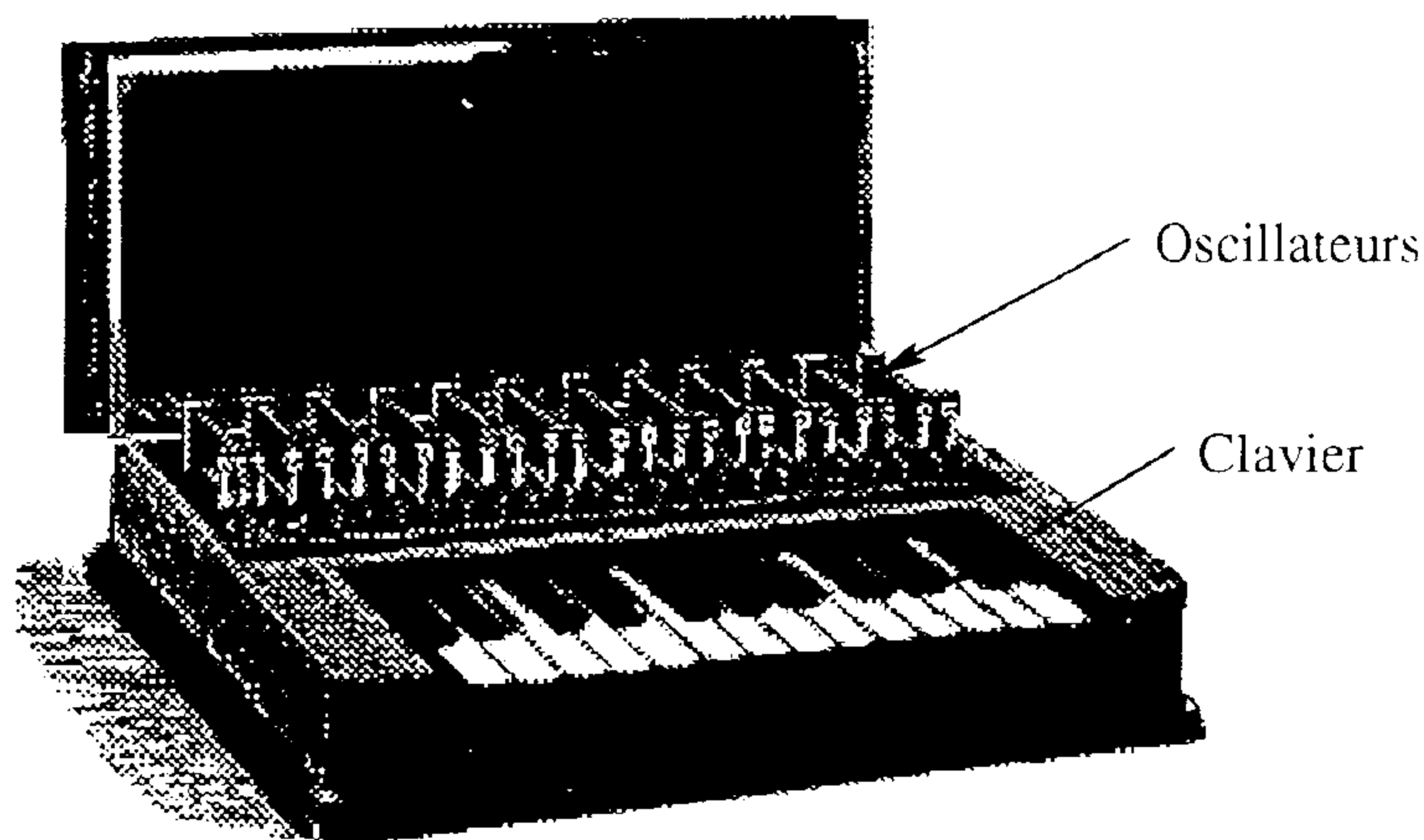


FIG. 1.2 – *Télégraphe harmonique de Gray (1874)*

resté célèbre grâce à l'invention du téléphone. Le second a été injustement oublié : il a eu le malheur de déposer une demande de brevet presque similaire à celle de Bell deux heures après celui-ci. Nous allons dans les paragraphes suivants nous pencher sur les travaux électroacoustiques de celui-ci.

A partir de 1870, Elisha Gray travaille au multiplexage des signaux télégraphiques. Les enjeux de ces travaux sont colossaux à l'époque : une telle technologie permettrait de multiplier le volume des transmissions télégraphiques sans augmenter le nombre de lignes.

Comme la plupart des ingénieurs électriciens de cette époque, Elisha Gray est essentiellement autodidacte, il a acquis la plus grande part de ses connaissances de façon empirique. Gardant cet état d'esprit, il incite son jeune neveu à manipuler électroaimants, piles et relais [Gra78b]. En 1874, son neveu lui montre une expérience amusante qu'il a réalisé autour d'un relais branché en oscillateur astable, d'une pile, ... et d'une baignoire. Ceci constitue le premier système de génération électro-acoustique comprenant oscillateur, amplificateur et transducteur électroacoustique. La pile et le relais constituent un oscillateur électrique. L'oscillation électrique est transmise aux muscles de la main par une électrode. La polarité négative de la pile étant reliée à la terre, le circuit est fermé lorsqu'on frotte la main sur le revêtement de la baignoire qui est implicitement reliée à la terre. L'oscillation électrique est transformée en oscillations mécaniques et acoustiques par stimulation électrique des muscles de la main !

Pour Elisha Gray, cette expérience ouvre la voie de la téléphonie, c'est-à-dire de la transmission de sons à distance. En 1874, il réalise le premier "télégraphe harmonique". Il s'agit de huit relais oscillateurs accordés selon les notes de la gamme. Cet instrument (figure 1.2) permet de transmettre de la musique sur les lignes télégraphiques dès 1874, deux ans avant l'invention du téléphone [Gra78b, Gra77, Gra76, Gra78a, Gra80].

De nombreuses expériences visant à analyser les sons furent menées dans la même période. Ces expériences ont établi les premières définitions "physiques" du timbre musical. Les expériences les plus connues sont celles de John Tyndall et de Hermann von Helmholtz. Les expériences de John Tyndall — dites des "flammes manométriques" — ont permis d'observer les formes d'ondes périodiques des sons musicaux à la manière des oscilloscopes contemporains (figure 1.3). Les travaux de Hermann von Helmholtz —

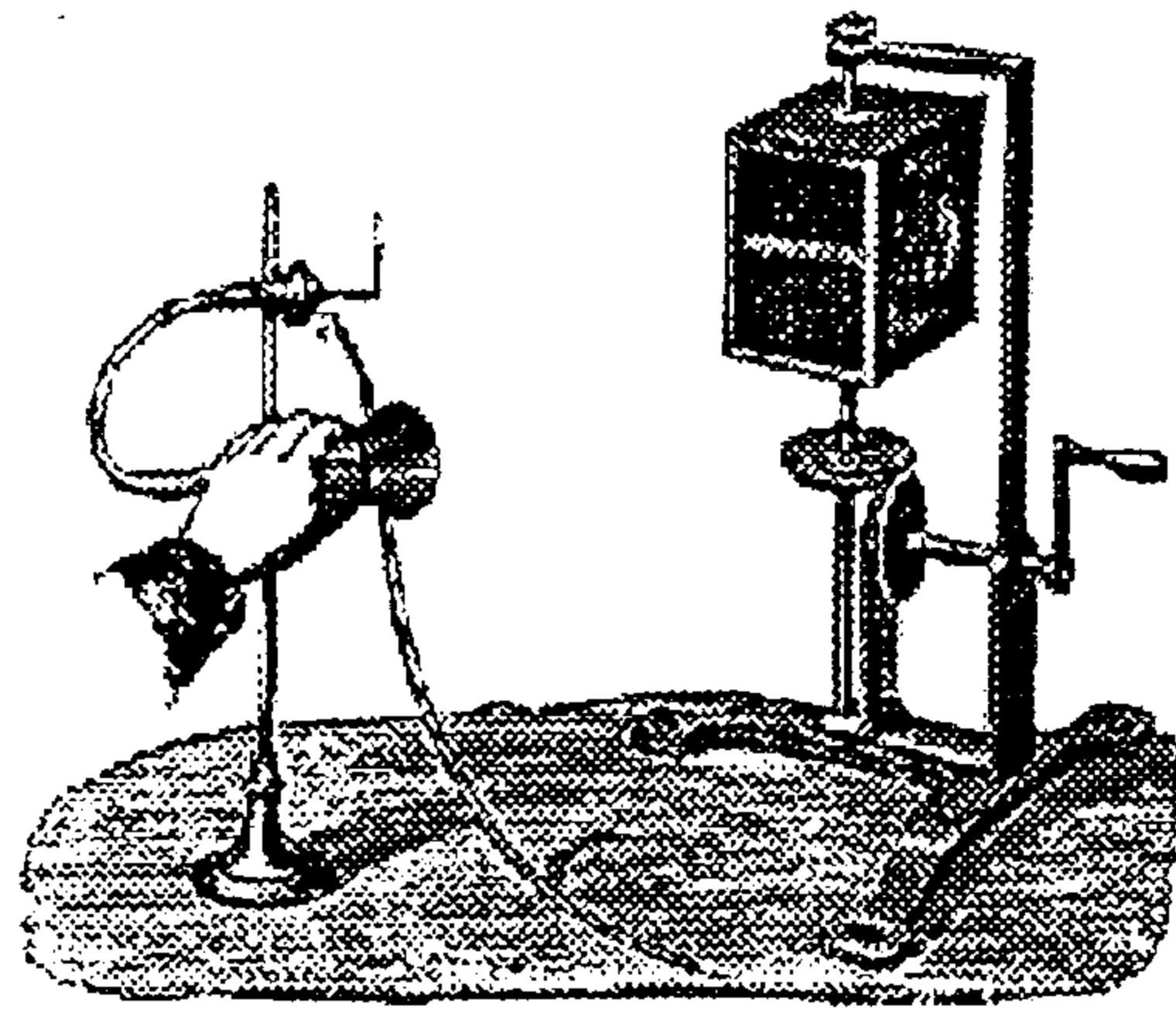
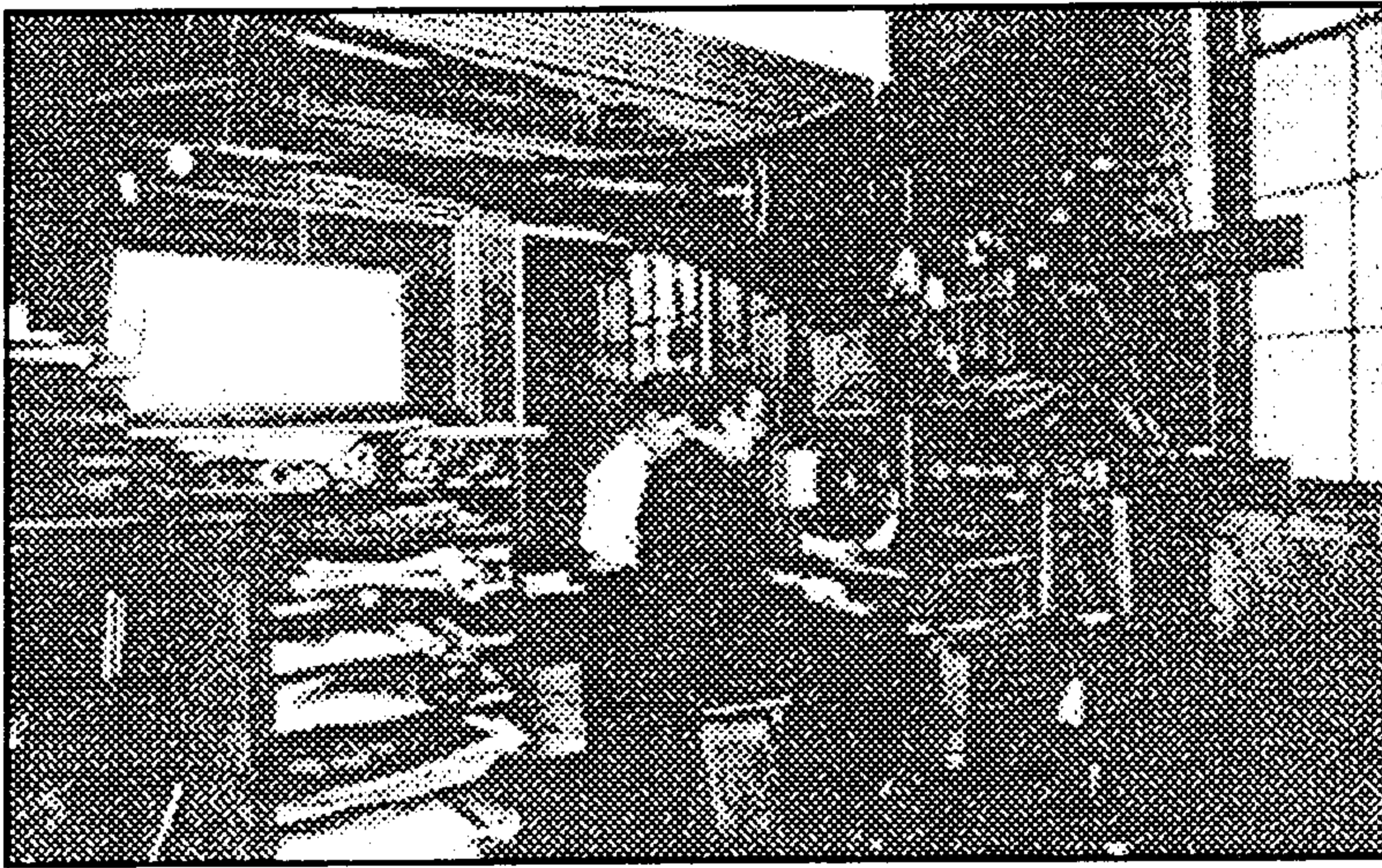


FIG. 1.3 – *Flammes manométriques de Tyndall (1875)*



a) Console du Telharmonium (1906)



b) un rotor du Telharmonium

FIG. 1.4 – *Le Telharmonium de Thaddeus Cahill (1906)*

menés à l'aide de résonateurs acoustiques — ont permis de mettre en évidence le rôle de la composition spectrale des sons. Ces deux physiciens ont mis en avant l'importance des phases "stables" des sons musicaux durant lesquelles le son est quasi-périodique. Bien que ne disposant pas de dispositif permettant l'étude, ils étaient conscients de l'importance des phases transitoires dans la perception du timbre [Roa95, page 500].

Les travaux de Gray et les travaux de Hermann von Helmholtz sur les sons musicaux seront la source d'inspiration de l'ingénieur américain Thaddeus Cahill [Wei95, Cah97, Cah17, Cah14, Cah19]. Au début de ce siècle, le téléphone ne sert pas uniquement à transmettre la parole, c'est aussi un moyen de distribution de musique. Le "telharmonium" de Thaddeus Cahill est considéré comme le premier synthétiseur digne de ce nom. Ce gigantesque instrument de deux cents tonnes est une véritable usine électrique. Il est constitué de multiples rotors délivrant chacun plusieurs signaux sinusoïdaux de fréquence fixe (figure 1.4b). Les sons musicaux sont obtenus par le mélange (synthèse additive) de ces sources d'oscillations. Le "telharmonic hall" ouvre ses portes le 26 septembre 1906.

Très rapidement, des problèmes financiers et techniques apparaissent (perturbation par induction des transmissions téléphoniques régulières). La dernière transmission de musique “telharmonique” a lieu en 1911. L’instrument sombre alors dans l’oubli. Hélas, aucun enregistrement de cet instrument ne subsiste.

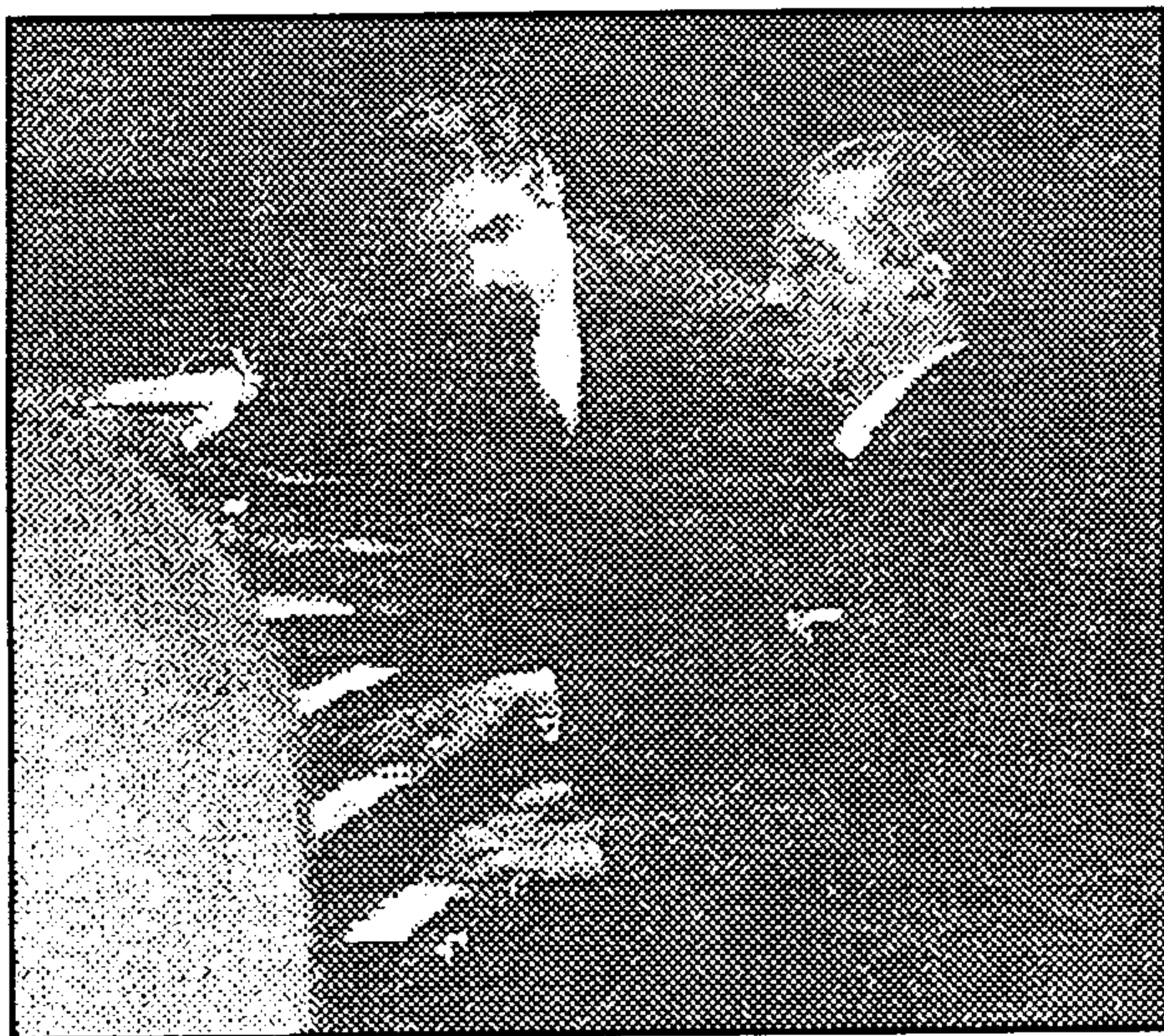
### 1.1.2 L’ère radioélectrique

La naissance de l’électronique est marquée par l’invention de la lampe triode (alors baptisée “audion”) en 1906 par Lee de Forest [For06, For20]. Cette invention ne sera réellement exploitée qu’à compter de 1915. Elle est alors produite en masse en France pour les transmissions hertziennes militaires.

C’est au cours de la première guerre mondiale que le jeune violoncelliste Maurice Martenot, affecté aux transmissions, s’initie aux lampes triodes. Au début de la guerre, les transmissions hertziennes sont réalisées par des “postes à étincelles” : les oscillations radioélectriques sont obtenues grâce à un arc électrique. Le son obtenu est une sorte de “Krrr Krrr” bien peu musical mais suffisant pour les transmissions morse. A partir de 1917, les postes à étincelles sont remplacés par des postes à lampes. Ces postes génèrent un son sinusoïdal presque pur. L’âme musicale de Maurice Martenot le pousse à expérimenter : pour le plaisir de ses compagnons d’infortune, il s’amuse à jouer “au clair de la Lune” en tournant le bouton de syntonisation de son poste [Lau90].

Au sortir de la guerre, Maurice Martenot consacre ses loisirs à la mise au point de ce qui deviendra “les ondes musicales Martenot”. Parallèlement à ces expérimentations, un jeune ingénieur russe, Lev Termen, conçoit un système d’alarme électronique. Cet appareil est constitué pour l’essentiel d’un oscillateur radiofréquence dont la fréquence peut être perturbée par la présence d’un corps dans le voisinage d’une antenne. Ces perturbations de fréquence sont ramenées dans le domaine audible par battement hétérodyne. L’appareil ne restera pas longtemps un système d’alarme, en mouvant ses mains autour de l’antenne Lev Termen parvient à jouer de la musique. Vers 1922, il fait une démonstration de cette invention, “l’etherophone”, devant Lénine. C’est le début d’une longue série de démonstrations. En 1927, il entame une tournée internationale qui le mène en Europe et aux Etats-Unis.

Maurice Martenot est un musicien rigoureux, bien que les résultats de ses expérimentations soient probants, il repousse jour après jour la date de présentation de son invention. Cette première version des “ondes” ressemble fort à l’etherophone, la principale différence étant la présence d’un fil entre la main de l’exécutant et l’instrument. 1927. Lev Termen fait deux démonstrations à Paris, la presse est enthousiaste, la revue “La Science et La Vie” titre “Vers une électrification de la musique” [Bol28]. Maurice Martenot a été pris de court. Sans doute aiguillonné par le succès de l’ingénieur russe [Lau90], il se lance dans une série de représentations et redouble d’efforts. Les “ondes Martenot” seront rapidement améliorées et équipées d’un clavier [Mar30, Kes34]. Le succès ne se fera pas longtemps attendre. C’est encore aujourd’hui l’instrument électronique considéré comme le plus abouti. Elles ont été fabriquées pendant plus de cinquante ans. L’etherophone, quant à lui, a été rebaptisé “thereminvox”, breveté et commercialisé aux Etats-Unis à partir de 1928 par la firme RCA [Thè28]. Le son de “voix surnaturelle” de cet instrument en a fait l’instrument de prédilection des cinéastes de science fiction américains dès 1933



a) A.Givelet et E.Coupleux à la console de leur orgue



b) Vue partielle des deux cents lampes

FIG. 1.5 – *Orgue radioélectrique Givelet Coupleux (1933)*

(*King Kong* de Max Steiner).

De nombreux autres instruments monophoniques et duophoniques sont développés durant les années trente tels que :

- de multiples "copies" de l'étherophone dont la *croix sonore* de Nikolay Obukhov (1926), *l'électronde* de Martin Taubmann (1929) et *l'ondium péchadre* de Péchadre (1930) ;
- le *dynaphone* de René Bertrand (1927), dont la fréquence était commandée par un rhéostat ;
- la série des *sphärophons* de Jörg Mager (1921 1928), dont la hauteur était commandé par une manette et un cadran gradué ;
- le *trautonium* de Friedrich Trautwein (1930) — premier instrument à intégrer une unité de réverbération à bande (ou plus exactement à fil) magnétique — commercialisé par la firme Telefunken [Tra38], dont une version améliorée, le *mixtur-trautonium* a servi à la réalisation des bruitages du film "les oiseaux" d'Alfred Hitchcock.

Martenot et Termen ont ouvert la voie de la musique électronique. Leurs instruments étaient monophoniques : ils ne pouvaient jouer qu'une note à la fois. La réalisation d'un orgue électronique polyphonique était un défi technologique presque insurmontable à l'époque en raison du manque de fiabilité des lampes triodes.

Le premier orgue électronique est le produit de la collaboration de l'ingénieur radio-électricien Armand Givelet et du facteur d'orgues Eloy Coupleux [Rou28, Wei31, Kes34, Giv34] (figure 1.5). Quatre orgues "radioélectriques" seront construites entre 1930 et 1935. Ces orgues, constituées d'environ deux cents lampes, fonctionnent sur les principes de la

synthèse additive et de la synthèse par filtrage formantique. Dans un premier temps, plusieurs oscillateurs sont combinés pour produire un signal riche en harmoniques. Le timbre de l'instrument est défini par cette combinaison d'harmoniques et par un banc de filtres qui affecte le spectre des signaux. Le choix du mélange des harmoniques et des filtres définit les jeux de l'orgue. Le plus important des orgues "radioélectriques" a été réalisé en 1933 pour le "Poste Parisien", il comportait plus de 200 lampes et 70 jeux, l'absence de microphone pour la prise de son permettait une transmission radiophonique directe de fidélité inégalable.

En 1934, Laurens Hammond conçoit une sorte de telharmonium miniaturisé : l'orgue Hammond. Dans cet orgue, les oscillations électriques sont obtenues au moyen de roues dentées en rotation, dites *roues phoniques* [Kes37, Ham34, Ham38]. Elles sont situées dans l'alignement de capteurs électromagnétiques similaires aux "pick-ups" des guitares électriques. Le faible signal généré est amplifié par un circuit à lampes. Ouvrons ici une parenthèse en signalant qu'Armand Givelet a déposé un brevet décrivant ce dispositif dès 1929 [Giv29]. L'orgue Hammond est d'abord destiné à remplacer les orgues à tuyaux. Malgré un coût intéressant et un encombrement réduit (200 kg), il essuie un échec cuisant : son timbre est trop différent de celui des orgues à tuyaux et de nombreux facteurs d'orgues mènent une campagne très active contre lui. L'orgue Hammond trouvera sa place dans le monde de la musique en devenant l'orgue de jazz par excellence.

En 1936, Laurens Hammond, sans doute inspiré par l'orgue "radioélectrique", conçoit le "Novachord". D'une taille similaire à l'orgue Hammond classique, cet orgue ne comporte pas moins de 170 lampes. Vers 1939, un ingénieur, John Eckert en fait l'acquisition. Il est étonné de la relative fiabilité de cet instrument, malgré le nombre important de lampes, il ne tombe presque pas en panne. Cette fiabilité est due à quelques astuces employées par Hammond (et précédemment par Givelet). En 1943, John Eckert travaille à la réalisation de calculateurs électromécaniques dédiés au calcul de tables balistiques. La fiabilité des lampes triodes ne permettait pas d'envisager sérieusement la réalisation d'un calculateur électronique et le choix d'un système électromécanique tombait sous le sens. Impressionné par la fiabilité du Novachord, John Eckert décide malgré tout de réaliser un calculateur électronique : l'ENIAC ne comportant pas moins de 18800 lampes, considéré comme le premier ordinateur.

### 1.1.3 L'ère analogique

Bien qu'inventé en 1948, le transistor mettra plus de quinze ans pour supplanter la lampe triode. Ce nouveau composant est capable de la remplacer dans de nombreuses applications. Le transistor permet un gain de coût et de place considérable.

Le premier synthétiseur analogique moderne est l'œuvre de Robert Arthur Moog. Durant les années 50, Robert A. Moog commercialise des thereminvox transistorisés en kits. Cette source de revenus lui permet de financer ses études. C'est dans ce contexte qu'il fait la connaissance d'un compositeur : Herbert Deutsch. En 1964, ils se réunissent pendant une semaine et travaillent à la conception d'un instrument de musique électronique capable de synthétiser une grande variété de sons. Le fruit de ces cogitations est le premier synthétiseur modulaire analogique (figure 1.6). Le terme "synthesizer" est pour la première fois employé pour décrire un générateur électronique de sons. Un tel synthétiseur repose

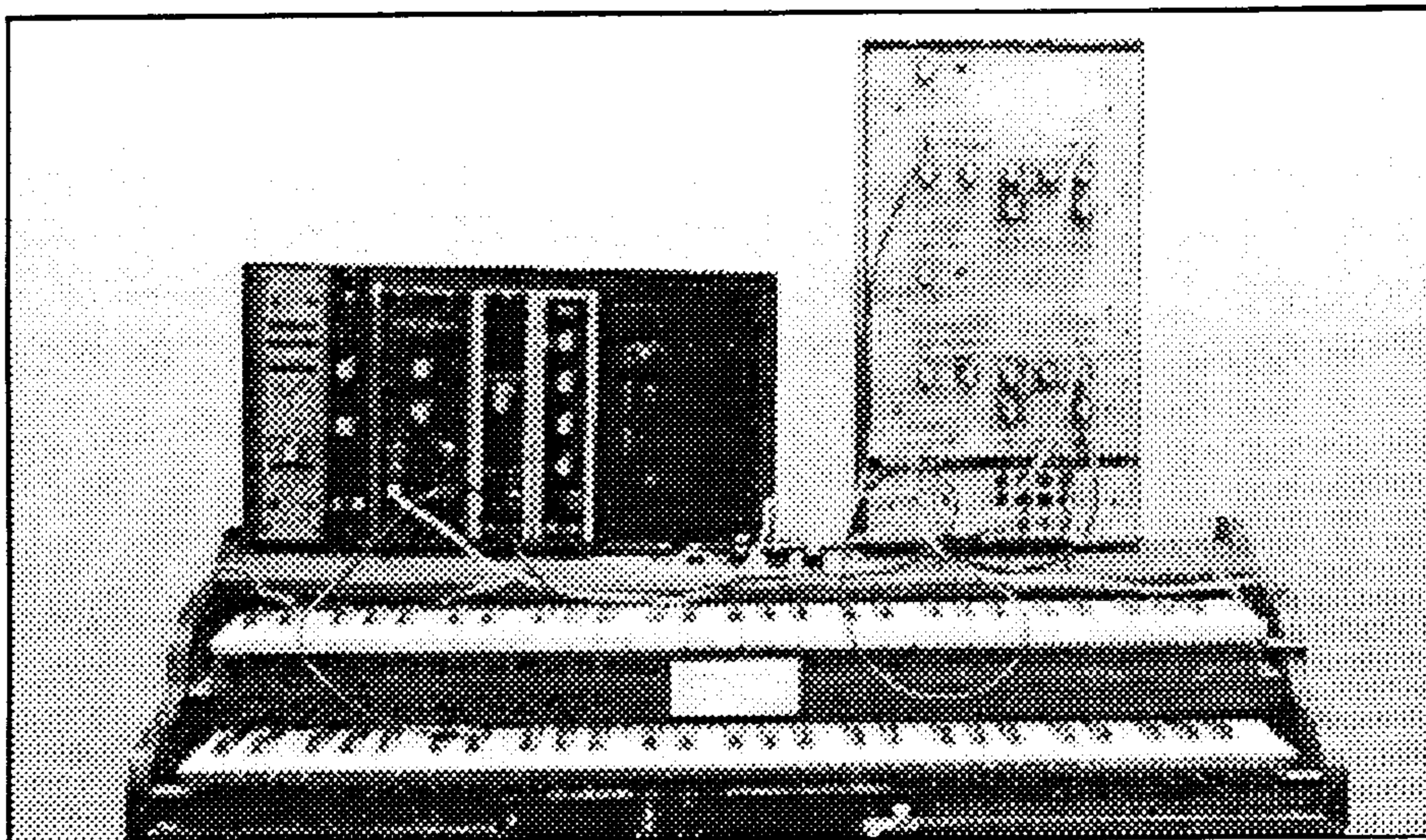


FIG. 1.6 – *Synthétiseur modulaire Moog (1964)*

sur le concept de module. Un module est un circuit électronique disposant en général d'entrées, de sorties et de boutons de réglage [Moo65]. Chaque module est capable de générer ou de modifier un signal. La chaîne de traitement du signal est définie par l'interconnexion des modules. Cette interconnexion se fait au moyen de câbles.

Les types de modules les plus importants sont :

- l'oscillateur (VCO) : c'est un générateur de fonctions dont la fréquence est commandée par une tension. Cette tension provient en général d'un clavier et détermine la note jouée ;
- l'amplificateur commandé en tension (VCA) : combiné au générateur d'enveloppe, il permet de faire varier l'amplitude d'un signal.
- le générateur d'enveloppe (EG ou ADSR) : il génère une tension de commande variable dans le temps. Cette variation est définie au moyen de boutons de réglage. Ce module est généralement utilisé pour piloter le VCA ou le VCF ;
- le filtre (VCF) : il permet d'affecter le spectre du signal. Sa fréquence de coupure est commandée par une tension. Elle peut, par exemple, provenir d'un générateur d'enveloppe. on obtient ainsi un son dynamique plus vivant qu'un son statique. Notons que ces filtres, conçus sur des critères esthétiques, étaient souvent bien différents du filtre "idéal" au sens physique et comportaient quelques non linéarités ;
- l'oscillateur basse fréquence (LFO) : tout comme le générateur d'enveloppe, ce module génère une fonction de commande destinée à affecter les paramètres d'un module. Appliqué à un oscillateur, il permet de créer un effet de vibrato ;
- Le générateur de bruit : ce module produit un signal bruité (en général blanc ou rose). Après filtrage, ce signal permet de réaliser de nombreux effets :

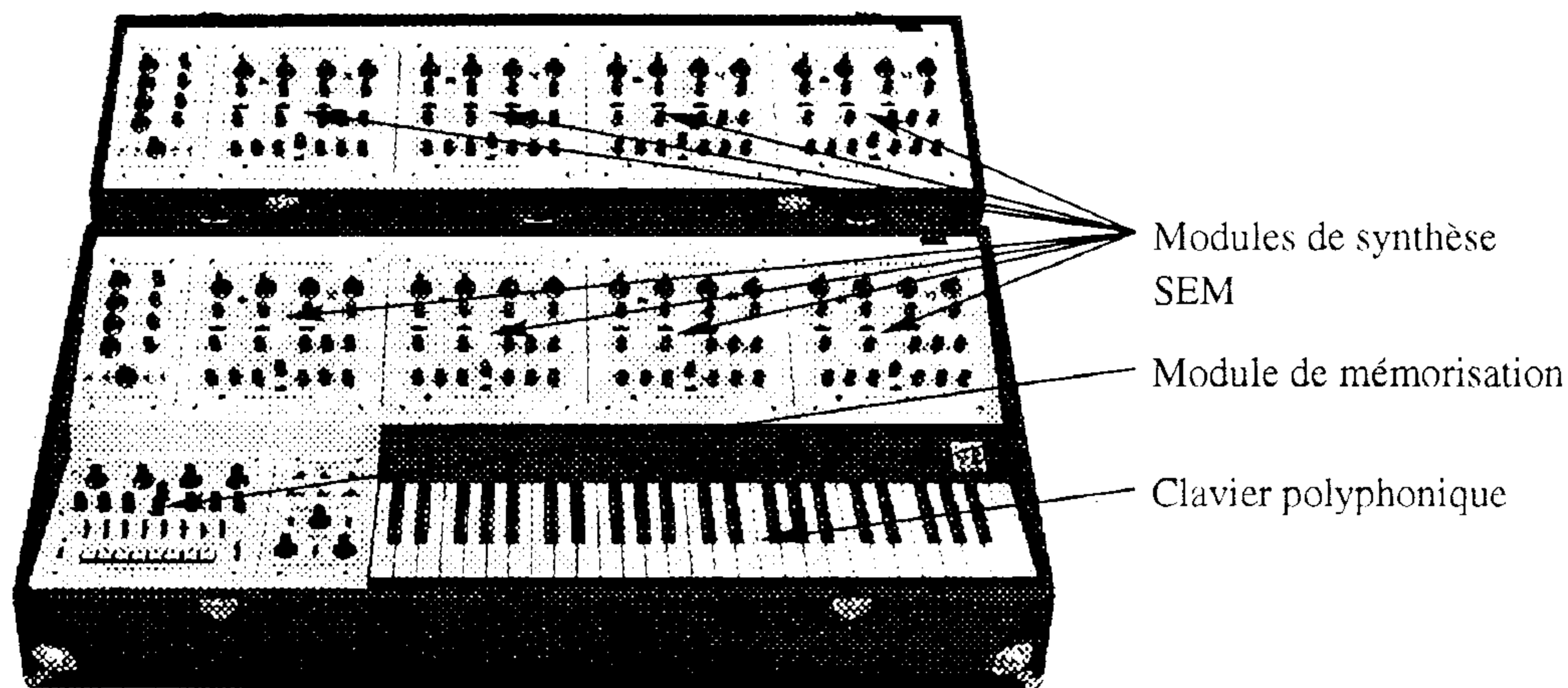


FIG. 1.7 – Synthétiseur analogique polyphonique Oberheim (1977)

- le modulateur en anneau (RM) : le signal de sortie de ce module est égal au produit des signaux d'entrée. L'effet de ce module est une translation dans le domaine spectral. Il est utile pour générer des sons dissonants tels que des sons de cloche.

On conçoit aisément que l'interconnexion d'une telle variété de modules offre une quasi-infinité de sons. Aujourd'hui encore, ces synthétiseurs sont appréciés pour ces possibilités de recherche. Hélas, les synthétiseurs modulaires analogiques ont plusieurs défauts majeurs : ils sont compliqués à utiliser (surtout pour des musiciens), ils sont monophoniques, et, ils sont encombrants. Ces défauts limitent leur utilisation au cadre des studios.

Pour satisfaire à la demande d'un synthétiseur de scène, léger et simple, Robert Moog conçoit le *minimoog* en 1971. Ce petit synthétiseur monophonique dispose d'une architecture "soustractive" fixe : les modules sont interconnectés selon le schéma "oscillateurs filtre amplificateur", deux générateurs d'enveloppes viennent piloter le filtre et l'amplificateur. En dix ans, plus de 10000 minimoogs seront vendus.

En 1974, Tom Oberheim présente le SEM, il s'agit d'un petit module sans clavier comportant à peu près l'équivalent d'un minimoog. Une telle miniaturisation a été rendue possible grâce à l'utilisation des circuits intégrés analogiques tels que les amplificateurs opérationnels. En 1975, il présente le premier synthétiseur analogique polyphonique : il s'agit de quatre SEM reliés à un clavier. Pour pouvoir jouer correctement de cet instrument, il fallait régler à l'identique les quatre modules, opération délicate et pénible. La polyphonie de ce type de synthétiseur sera poussée jusqu'à 8 (figure 1.7).

A partir de 1975, l'utilisation du microprocesseur se généralise. Par ailleurs, la demande en circuits intégrés analogiques dédiés à la synthèse est suffisamment importante pour en lancer la production. En 1977, l'équivalent d'un minimoog peut être réalisé avec deux circuits intégrés analogiques spécialisés. Le premier synthétiseur analogique polyphonique à mémoire est commercialisé à partir de 1978. Ce synthétiseur, le Prophet 5, est composé d'un microprocesseur qui commande un banc de synthétiseurs analogiques (figure 1.8). Les nombreux réglages définissant le son (paramètres d'enveloppe A, D, S et R, fréquence et déviation de fréquence des oscillateurs, fréquence de coupure et coefficient de qualité du filtre) peuvent être mémorisés et rappelés grâce à une mémoire numérique. Le



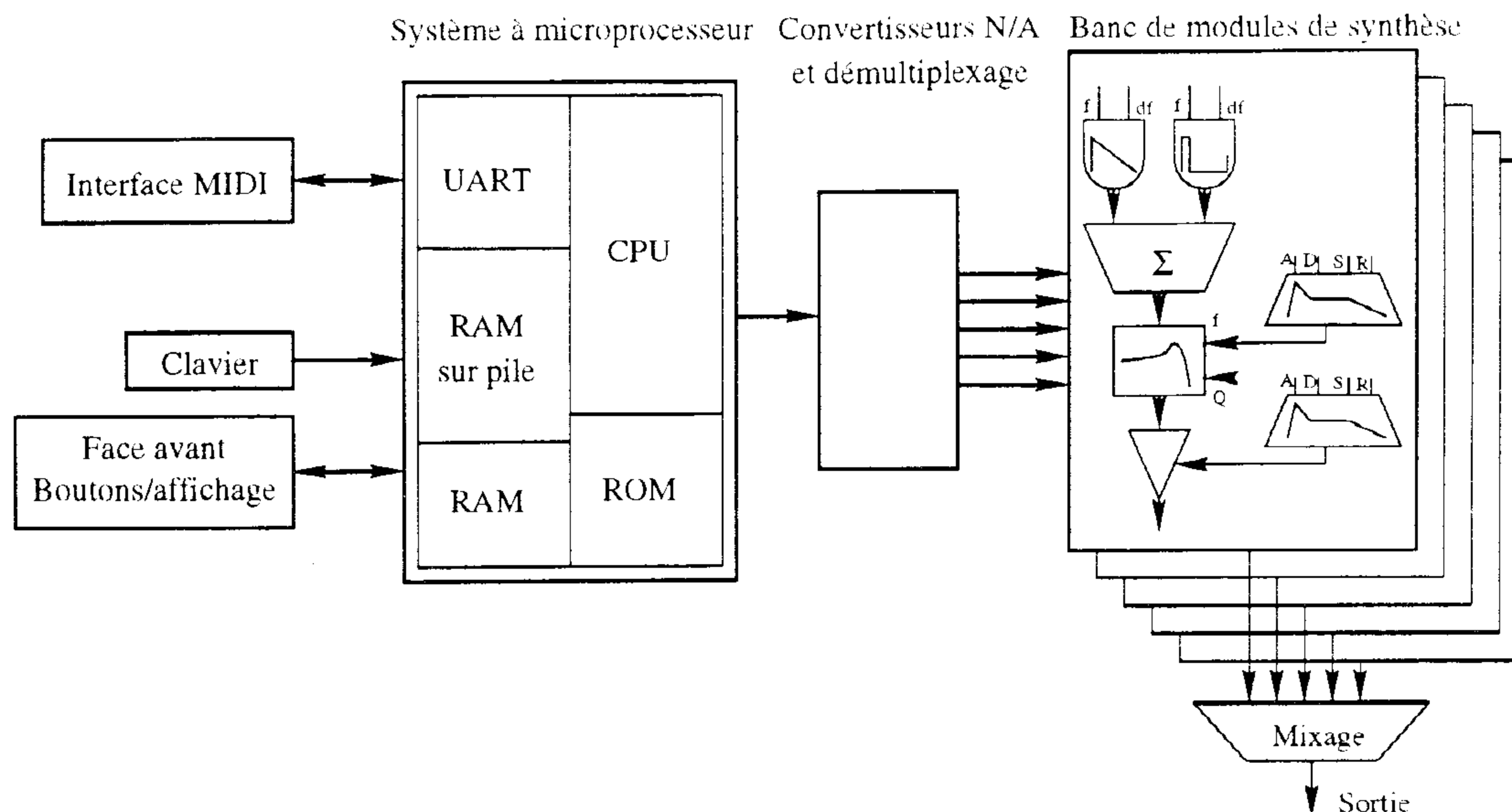


FIG. 1.8 – *Synthétiseur analogique polyphonique (1978)*

microprocesseur, dans ce contexte, permet de pallier aux inconvénients des synthétiseurs polyphoniques.

En une quinzaine d'année, le synthétiseur analogique a gagné en intégration, en polyphonie et en souplesse d'utilisation. Il a cependant perdu la richesse potentielle dont disposaient les premiers synthétiseurs modulaires.

#### 1.1.4 L'ère numérique

La première synthèse de son par ordinateur est l'œuvre de Max V. Mathews et de ses collègues des Bell Telephone Laboratories en 1957. L'ordinateur mis en œuvre pour les calculs était un IBM 704 à lampes. Le programme **MUSIC I**, écrit à cette occasion en assembleur par Max V. Mathews, était capable de produire une unique forme d'onde : le triangle. Néanmoins, cette première démonstration ouvrait la voie de la synthèse numérique des sons [Roa95, page 87].

Plusieurs programmes **MUSIC N** furent écrits par la suite par Max V. Mathews; chacun introduisait de nouveaux concepts et profitait des progrès de l'informatique. **MUSIC II** (1958) intégrait seize formes d'ondes. **MUSIC III** (1960) introduisait les "unités de génération", équivalents numériques des VCO de Robert A. Moog. **MUSIC IV** (1963) était une nouvelle version écrite en langage **BEFAP** (aujourd'hui tout à fait disparu). En 1969, **MUSIC V** était écrit en langage standard **FORTRAN IV** (aujourd'hui presque disparu). Il fut distribué vers une trentaine d'universités.

Durant les années soixante-dix, la synthèse numérique restera l'apanage des universités et des centres de calcul. En dehors des orgues numériques Allen (conçus par Ralph Deutsch dès 1971), il faudra attendre le début des années quatre-vingts pour voir les premières applications commerciales de la synthèse numérique.

Les premiers synthétiseurs échantillonneurs sont commercialisés à partir de 1979 par la société australienne Fairlight (figure 1.9). Ce type de synthétiseur repose sur

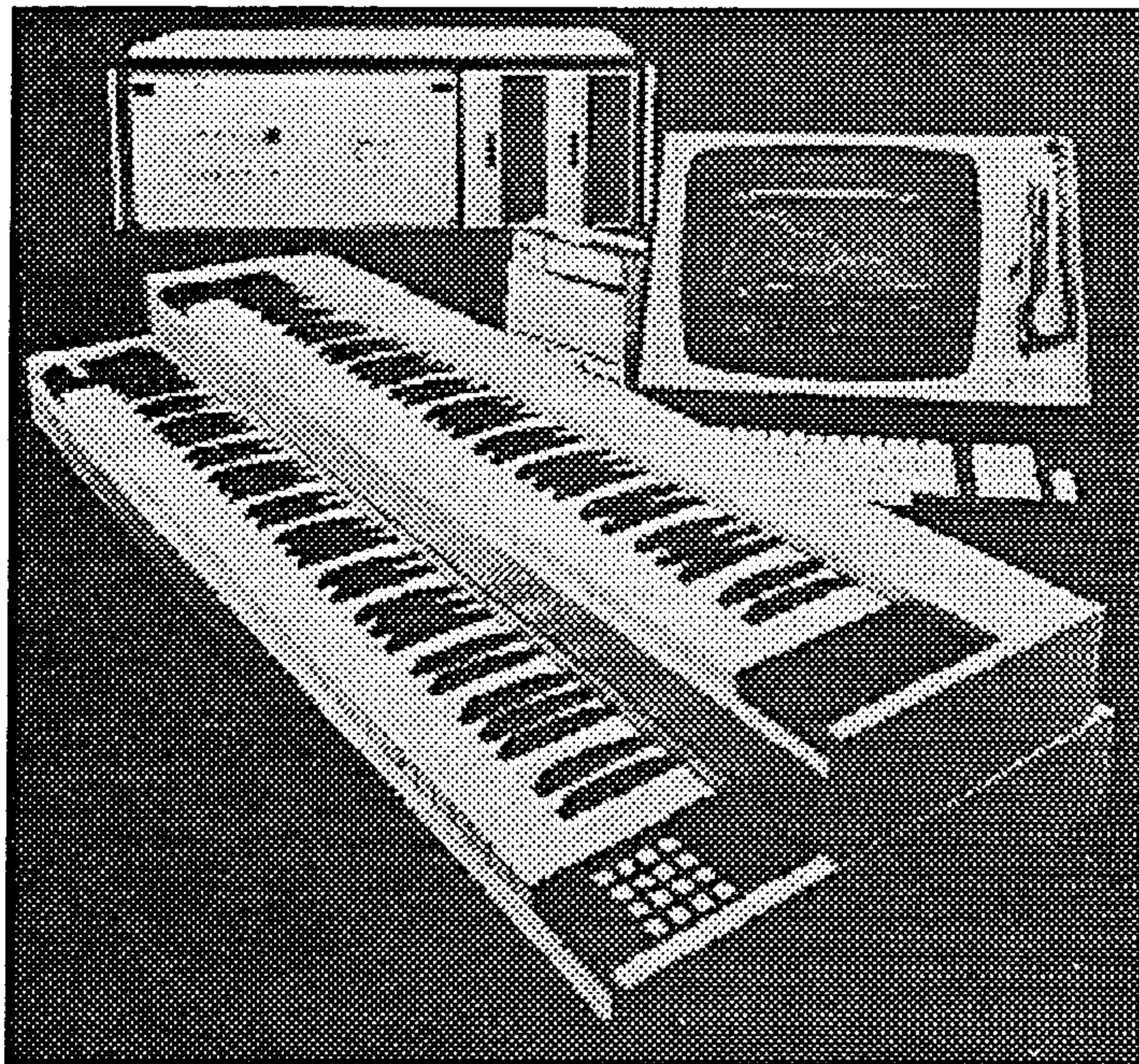


FIG. 1.9 – Synthétiseur/échantillonneur Fairlight (1979)

l'enregistrement de sons dans une mémoire numérique (échantillonnage). Les sons sont restitués par simple lecture de ces mémoires. En faisant varier la cadence de lecture, il est possible de transposer la fréquence des sons et ce faisant de jouer de la musique. L'enregistrement numérique de sons consomme beaucoup de mémoire. Une seconde de musique enregistrée en stéréo en 16 bits à 44,1 kHz (format CD) occupe 176400 octets. Constatant que la plupart des sons musicaux ont une forme d'onde presque constante durant leur maintien, les concepteurs d'échantillonneurs ont eu l'idée de "boucler" les sons. Une boucle élémentaire représente une période du son. On appelle une telle boucle "table d'onde" (§1.2.2, page 34). Il est possible de reconstituer certaines variations de spectre en faisant se succéder différentes tables d'ondes (*wave-sequencing*, §1.2.3, page 35).

Les années 80 voient la naissance de la micro-informatique. Cette décennie est marquée par l'essor des technologies numériques. Le premier synthétiseur numérique grand public apparaît en 1983. Il s'agit du DX7 de la firme Yamaha. Ce synthétiseur repose sur le principe de la modulation de fréquence (§1.2.5, page 38) qui a été avancé dès 1973 par John Chowning [Cho73, Cho77]. La modulation en fréquence d'un oscillateur par un autre oscillateur engendre un spectre de raies secondaires dont la richesse peut être réglée aisément. On comprend pourquoi cette forme de synthèse a séduit les ingénieurs de Yamaha : avec un minimum de calcul, la synthèse par modulation de fréquence permet de créer des sons dynamiques et complexes. Le synthétiseur DX7 est capable de générer 8 sons simultanément, chacun étant constitué de 6 oscillateurs agencés de 32 façons possibles.

La micro-informatique et la numérisation des synthétiseurs ouvre la voie à l'informatique musicale "grand public". Dès 1983, il est possible de piloter les générateurs de sons des synthétiseurs par micro-ordinateur au moyen de l'interface MIDI (Musical Instrument Digital Interface). Cette interface permet aux synthétiseurs et aux ordinateurs

de communiquer. Les informations échangées décrivent soit une action "virtuelle" sur le clavier (appuyer sur la touche 46 du son numéro 3 avec la vélocité 89), soit une action sur les paramètres du son synthétisé. En bref, l'interface MIDI ne transmet pas de son, elle transmet des instructions au générateur de sons du synthétiseur. La puissance, la simplicité et la souplesse de cette interface en ont fait un standard durable, toujours utilisé après quinze ans d'existence. L'apparition simultanée d'une forme de synthèse numérique relativement bon marché, de la micro-informatique et de l'interface MIDI a sonné le glas des synthétiseurs analogiques.

Contrairement à la synthèse soustractive, la synthèse par modulation de fréquence est très difficile à maîtriser. La plupart des musiciens n'exploitait pas les capacités de synthèse de leur synthétiseur et se contentait des sons "d'usine". Une alternative numérique à cette forme de synthèse est apparue à la moitié des années 80. Il s'agit en fait d'une forme de synthèse hybride combinant lecture de tables d'ondes et filtrage dynamique, c'est-à-dire dont les caractéristiques évoluent au cours du temps. L'architecture de ce type de synthétiseur reprend le schéma classique de la synthèse soustractive, la principale différence réside dans l'oscillateur qui est remplacé par un lecteur de tables d'ondes. Le lecteur de tables d'ondes permet d'obtenir des sons réalistes alors que le filtrage dynamique permet de leur donner une certaine vie. Ce type de synthèse s'est rapidement imposé en raison de sa simplicité de programmation et de son coût (les lecteurs de tables d'ondes nécessitent de la mémoire dont le coût n'a eu de cesse de baisser). La synthèse hybride, combinant lecture de tables d'ondes (§1.2.2, page 34) et filtrage dynamique (§1.2.4, page 36), est restée la plus courante, on la trouve dans la plupart des synthétiseurs et des cartes audionumériques, de moyen et de haut de gamme.

Le développement de la micro-informatique et des disques durs a permis de les utiliser comme support d'enregistrement audionumérique : le micro-ordinateur est devenu un véritable studio intégré. D'autre part, la production massive de lecteurs de disques audionumériques (Compact Discs) a permis de réduire considérablement le coût des composants audionumériques (en particulier les convertisseurs numérique-analogiques). Il a donc été possible de produire des cartes audionumériques à un coût réduit. De telles cartes permettent l'enregistrement (échantillonnage) et la lecture de sons sur le disque dur (direct to disc). Elles permettent aussi la synthèse de sons au moyen de synthétiseurs intégrés (les premières cartes utilisaient la modulation de fréquence, les cartes récentes utilisent la synthèse hybride). Notons que les circuits intégrés utilisés dans certaines cartes audionumériques le sont aussi dans certains synthétiseurs à vocation plus "professionnelle" (on peut citer le SAM9407 de la société DREAM/ATMEL-ES2 et le Emu8000 de la société E-mu/Creative technologies).

La synthèse hybride par lecture de tables d'ondes et filtrage occupe aujourd'hui une place prédominante sur le marché de la synthèse. Cette forme de synthèse est cependant fortement critiquée par les musiciens qui la trouvent peu expressive. En effet, un tel synthétiseur offre un comportement trop précis et trop répétitif. Une action engendre systématiquement le même son, ce qui n'est pas le cas avec un instrument acoustique car ces instruments sont très sensibles aux nuances de jeu. Une alternative à la synthèse hybride est la synthèse par modélisation physique. Ce type de synthèse a été développé par d'importants centres de recherche dont le CCRMA (Computer Center for Research in Music and Acoustics) de l'université de Stanford. Le principe de la modélisation physique

repose sur la simulation du fonctionnement d'un instrument acoustique. L'instrument est décrit en terme de source d'excitation, de guides d'ondes et de résonateurs. Les résultats obtenus par ce moyen sont parfois très réalistes et les sons sont très expressifs, c'est-à-dire sensibles aux nuances de jeu. Le principal écueil de la modélisation physique reste son coût : elle nécessite une quantité de calculs très importante. C'est pourquoi la plupart des synthétiseurs usant de cette forme de synthèse sont monophoniques ou se limitent à quelques aspects de la modélisation physique.

### 1.1.5 Conclusion

Le premier siècle d'évolution des techniques électroniques a engendré de nombreux instruments. Ces instruments, souvent cantonnés à une utilisation marginale, ont permis d'étendre la "palette" sonore disponible aux musiciens et d'affiner la connaissance du son.

Les vingt dernières années témoignent d'une évolution qui n'est pas sans rappeler celle de la micro-informatique. Le marché "grand public", de par son volume, a su bénéficier de moyens financiers de recherche et de développement considérables. Ces moyens ont permis aux sociétés dépendant de ce marché d'acquérir un niveau technique comparable à celui des sociétés dépendant du marché "professionnel".

L'intérêt que porte le "grand public" à ce domaine doit être considéré comme un élément positif. Dans cette perspective, la recherche dite universitaire doit contribuer non seulement au développement de nouvelles technologies mais aussi à leur vulgarisation auprès du public [Roc97].

## 1.2 Inventaire des méthodes de synthèse

La section 1.1 a présenté les diverses techniques de synthèse des sons musicaux sous un jour historique. Cette section a pour vocation de présenter divers types de synthèses sous un jour plus technique. Nous décrivons ici plusieurs méthodes de synthèse actuelles en explicitant leurs principes et leurs réalisations logicielles et/ou matérielles. Nous terminons par une brève description des divers supports matériels actuels.

### 1.2.1 Synthèse additive

#### Principe

La synthèse additive est la forme de synthèse la plus élémentaire. Elle procède par addition de composantes élémentaires, le plus souvent sinusoïdales. Cette formulation lui permet d'exploiter plus ou moins directement les données provenant d'analyses spectrales.

Chaque échantillon du signal synthétisé  $S(n)$  s'exprime sous la forme :

$$S(n) = \sum_{j=1}^J C_j(n) \quad (1.1)$$

avec

$$C_j(n) = a_j(n) \cos \left( 2\pi n \frac{f_j(n)}{F_e} + \alpha_j \right) \quad (1.2)$$

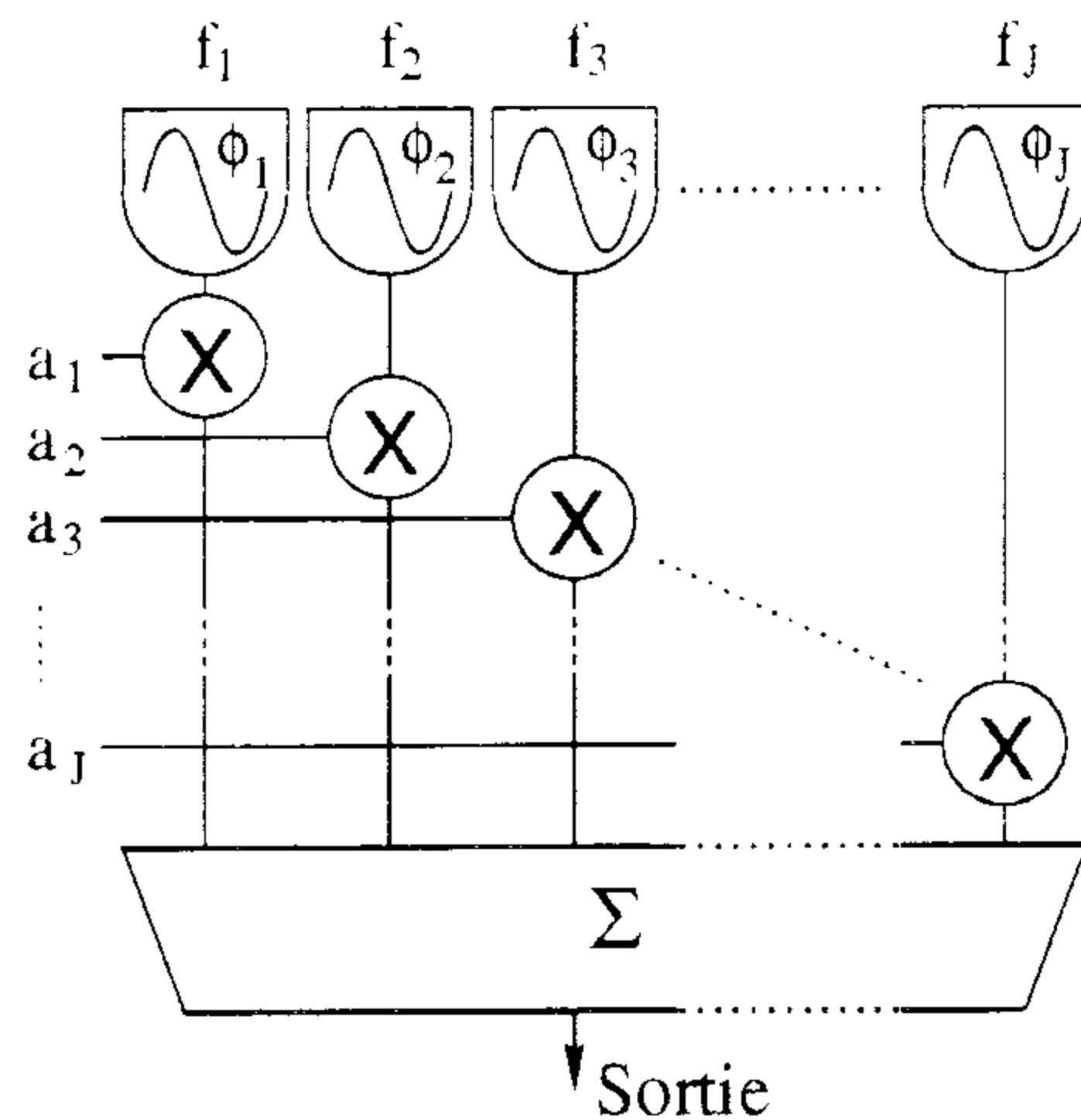


FIG. 1.10 – Synthèse additive par banc d'oscillateurs

- $J$  est le nombre de composantes sinusoïdales ;
- $F_e$  est la fréquence d'échantillonnage ;
- $f_j(n)$  est la fréquence de la composante  $j$  à l'instant  $n$  ;
- $a_j(n)$  est l'amplitude de la composante  $j$  à l'instant  $n$  ;
- $\alpha_j$  est la phase initiale de la composante  $j$ .

Les paramètres  $f_j(n)$ ,  $a_j(n)$  et  $\alpha_j(n)$  déterminent le signal généré. Dans le cadre des applications musicales, ces paramètres sont généralement "rafraîchis" à une cadence de l'ordre des 200 Hz.

### Oscillateurs sinusoïdaux à table d'onde

C'est la forme triviale de synthèse additive (figure 1.10). Chaque composante est générée par un oscillateur. La phase instantanée de chaque oscillateur est calculée par accumulation du terme de fréquence :

$$\phi_j(n) = \phi_j(n-1) + 2\pi \frac{f_j(n)}{F_e} \quad [2\pi] \quad (1.3)$$

Un oscillateur numérique à lecture de table peut être implémenté de deux façons (figure 1.11). Toutes les deux consistent à accumuler la fréquence dans un registre de phase au rythme d'une horloge  $H$  de fréquence fixe  $F_e$ . Ce registre peut, bien entendu, être étendu à une mémoire afin de synthétiser séquentiellement plusieurs composantes avec le même circuit [WEM79]. La phase est représentée sous la forme d'un réel en virgule fixe de partie entière  $\phi_{fort}$  et de partie décimale  $\phi_{faible}$ . L'excursion de la partie entière est équivalente à la longueur de la table en mémoire, c'est en général une puissance de deux.

La première méthode tronque purement et simplement la phase et ne considère que  $\phi_{fort}$  pour adresser une table de sinus en mémoire. Cette méthode provoque une erreur qui dépend de la taille de la table. Il est théoriquement possible de réduire cette erreur

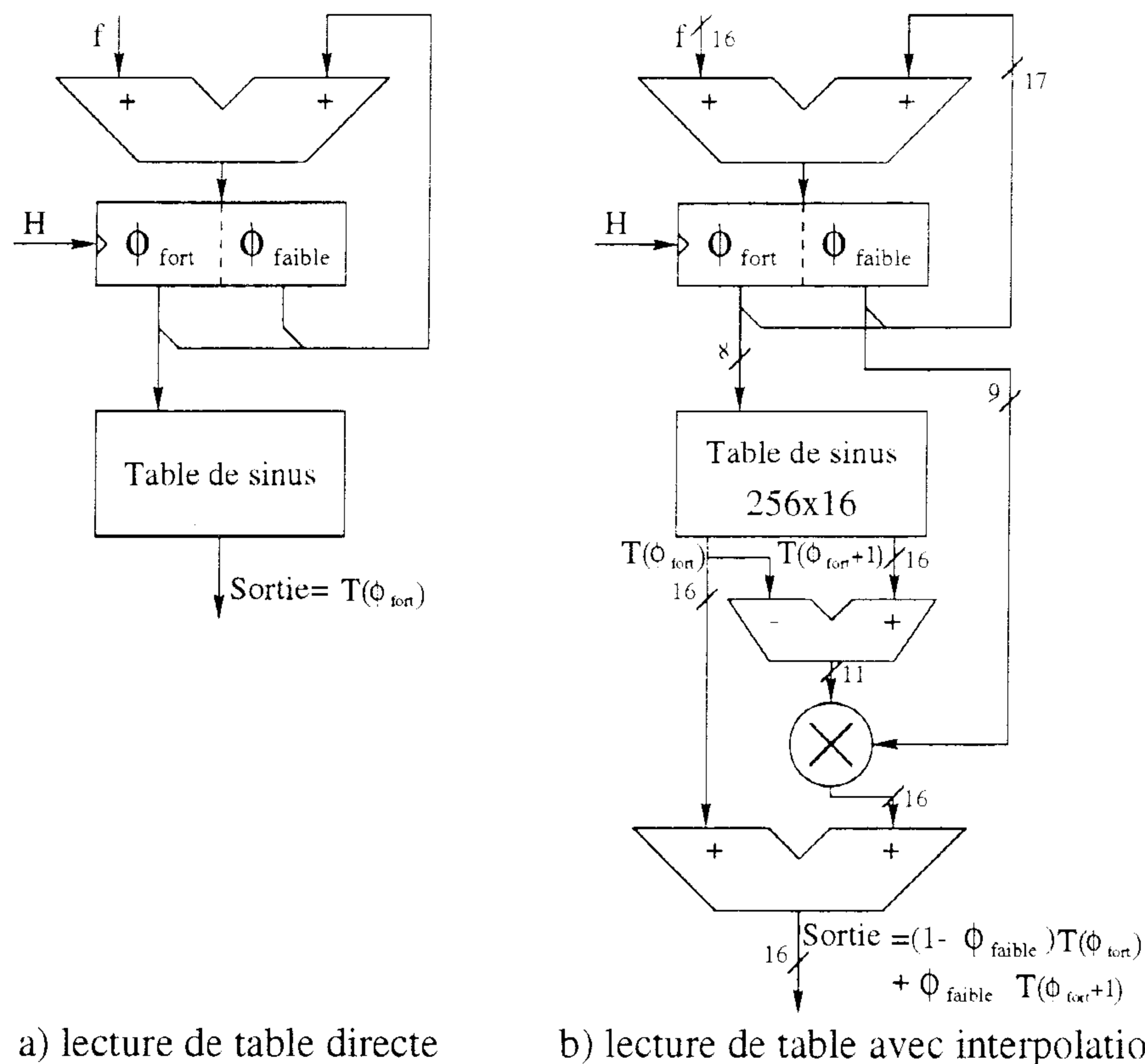


FIG. 1.11 – Oscillateur numérique à lecture de table

en utilisant une table d'onde de longueur conséquente. L'erreur est minimale, en regard de la résolution fréquentielle, lorsque la longueur de la table est équivalente au nombre de bits du registre de phase. Dans la pratique, la répartition spectrale et l'intensité du signal d'erreur dépendent fortement de la fréquence synthétisée (l'erreur est nulle si cette fréquence  $f$  est sous-multiple de la fréquence d'échantillonnage  $F_e$  [Chr90. Moo77a]).

La seconde méthode procède par interpolation linéaire de la table. Comme précédemment la partie entière  $\phi_{fort}$  sert de pointeur sur la table. Cependant, la partie décimale  $\phi_{faible}$  sert à pondérer les éléments de la table et permet ainsi de l'interpoler linéairement. L'interpolation permet d'utiliser des tables d'ondes assez courtes, son inconvénient est de nécessiter une multiplication et deux additions supplémentaires.

Le nombre de bits décrivant la fréquence, et par extension la phase, définit la résolution fréquentielle absolue du générateur. Dans la pratique, cette résolution doit être approximativement supérieure à la résolution fréquentielle absolue du système auditif humain. Celui-ci est capable de percevoir une variation de l'ordre des 0.3% de fréquence entre 150Hz et 5kHz [Gri94]. L'oscillateur numérique doit donc avoir une résolution fréquentielle d'approximativement  $150\text{Hz} \times 0.3\% \simeq 0.5\text{Hz}$ . Pour une fréquence d'échantillonnage de 48kHz, cette résolution est de l'ordre de  $\log_2(24000/0.5) = 15.55 \simeq 16$  bits pour la fréquence et 17 bits pour le registre de phase [HFM95]. Cette approximation justifie la taille des registres de phase de l'ordre des 16-20 bits courante dans les applications industrielles. La sortie de l'additionneur d'interpolation est codée sur 11 bits. Ces 11 bits sont suffisants pour coder la plus grande différence de deux échantillons successifs de la table

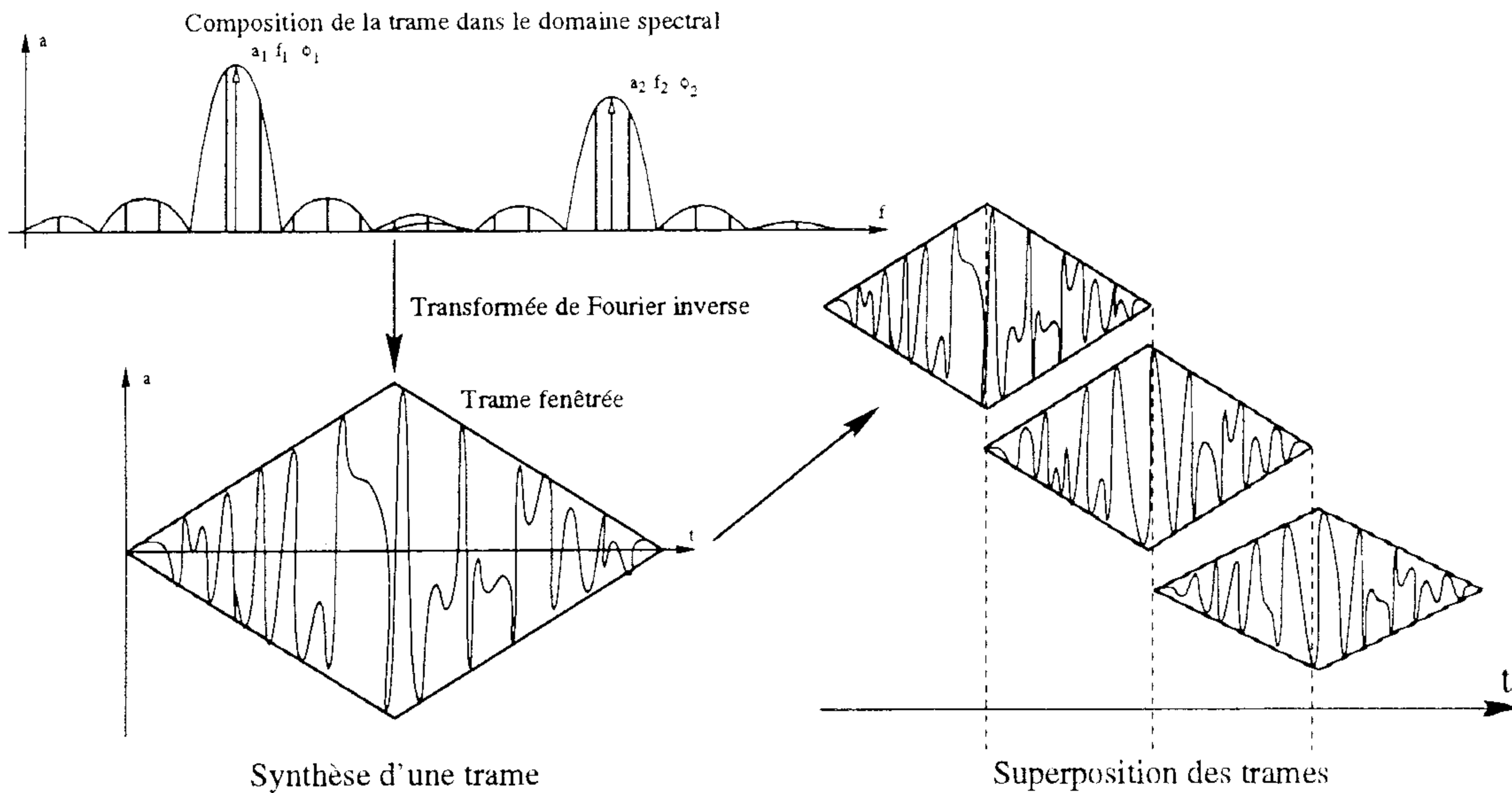


FIG. 1.12 – *Synthèse additive par transformée de Fourier inverse*

de sinus. Cette économie de nombre de bits est appréciable pour la réalisation du multiplicateur d'interpolation. D'autres optimisations, tenant compte des symétries de la table d'ondes peuvent être avantageusement exploitées [Sog97].

### Oscillateurs sinusoïdaux récursifs

Cette famille de méthodes permet de générer des signaux plus ou moins complexes. La méthode la plus connue permet de générer des signaux sinusoïdaux amortis avec un minimum de ressources de calcul et de mémoire. La série  $(X_n)_{n=1,\dots,N}$  définie par :

$$X_0 = A \sin(\phi) \quad (1.4)$$

$$X_1 = Aa \sin\left(\phi + \frac{\omega}{F_e}\right) \quad (1.5)$$

$$X_n = 2a \cos\left(\frac{\omega}{F_e}\right) X_{n-1} - a^2 X_{n-2} \quad (1.6)$$

ne nécessite pas de tables d'ondes mais seulement quatre emplacements mémoires, deux multiplications et une addition par échantillon. Elle définit une sinusoïde à amortissement exponentiel :

$$X_n = Aa^n \sin\left(\frac{\phi + n\omega}{F_e}\right) \quad (1.7)$$

Cette méthode a été largement utilisée dans le passé pour implanter un grand nombre d'oscillateurs sinusoïdaux [Bro82, page 55]. Son manque de souplesse — notamment l'impossibilité de faire varier les paramètres au cours du temps — et le coût décroissant des mémoires lui a fait perdre de l'intérêt. Notons que cette méthode peut être interprétée comme l'excitation d'un filtre du second ordre par une impulsion.

### Synthèse par transformée de Fourier inverse à court terme

Le coût de calcul de la synthèse additive par banc d'oscillateurs réserve cette forme de synthèse aux applications de laboratoire. L'utilisation d'algorithmes plus efficaces peut rendre la synthèse additive plus exploitable.

L'utilisation de la transformée de Fourier inverse à court terme pour la synthèse additive a été proposée et brevetée par Xavier Rodet et Philippe Depalle [DR94]. Cette méthode peut être interprétée comme la réciproque de l'analyse spectrale par interpolation de transformée de Fourier discrète abordée en 1.4.2.

Le son est décomposé en *trames temporelles* d'une durée de l'ordre des 10 ms. Pour chaque trame, on travaille dans le *domaine spectral*. Le domaine spectral est *discrétisé* en un nombre fini de canaux. Pour synthétiser une fréquence "continue" quelconque, il faut "exciter" plusieurs canaux. Cette "excitation" dépend directement de la *fenêtre* de la trame: afin de limiter le temps de calcul, on choisit une fenêtre qui minimise le nombre de canaux excités. Ainsi, chaque trame est composée dans le domaine spectral par accumulation "d'excitations" de canaux. La synthèse de la trame à proprement parler est effectuée par transformée de Fourier inverse rapide (figure 1.12). Pour pouvoir effectuer implicitement une interpolation linéaire des amplitudes des composantes entre deux trames successives, on doit utiliser une fenêtre triangulaire. Celle-ci est obtenue par "refenêtrage" de la trame. L'interpolation des fréquences et la cohérence des phases d'une trame à la suivante doit être assurée en ajustant convenablement la phase de "l'excitation" de chaque canal.

Comparée à la méthode précédente, cette méthode permet un gain de calcul de l'ordre de 13 (pour un nombre de sinusoides générées de l'ordre de 1000) [DR94]. Le travail dans le domaine spectral est une bonne solution *a priori* car il ne nécessite pas de bancs d'oscillateurs. Hélas, une partie du gain de calcul est perdue dans la gestion des "excitations" spectrales et des fenêtres.

### Conclusion

La synthèse additive, quel que soit sa réalisation, demeure difficilement exploitable dans les synthétiseurs commerciaux<sup>1</sup> pour deux raisons. La première est son coût de calcul qui reste encore important malgré le développement de procédés tels que l'utilisation de la transformée de Fourier inverse. La seconde réside dans le nombre de paramètres de synthèse mis en jeu: un son musical comporte en moyenne une trentaine de composantes dont il faut commander la fréquence et l'amplitude. Malgré ces défauts, la synthèse additive se révèle être un puissant outil de synthèse en temps différé: elle est souvent au cœur de procédés d'analyse synthèse qui permettent d'analyser et de modifier efficacement des sons numérisés. L'évolution des technologies et l'avènement de processeurs spécialisés rapides peut sérieusement modifier ce contexte et faire apparaître de nouveaux produits plus performants (tels que le "K5000" de la firme "Kawai").

Comme nous le verrons, la méthode présentée dans ce mémoire peut être interprétée comme une variante de la synthèse additive faisant abstraction de ces problèmes.

---

1. quelques tentatives telles que l'"Axcel" de la société "Technos" (1988) ou le "FDSS" de la société "Lyre" (1987) connurent des carrières commerciales moins qu'éphémères en raison de leurs coûts prohibitifs.



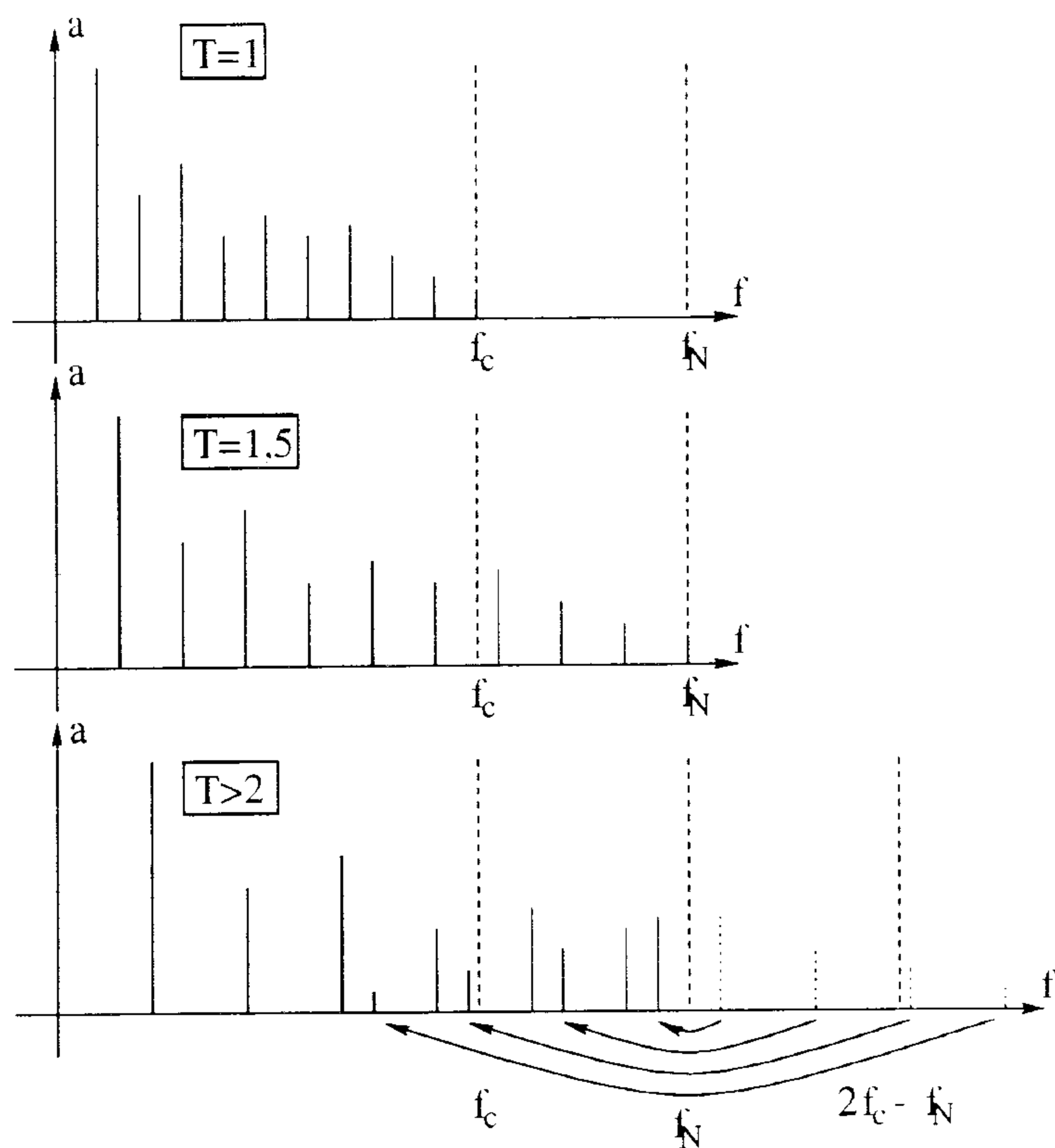


FIG. 1.13 – Transposition et repliement de spectre

### 1.2.2 Synthèse par tables d'ondes

Un son musical à hauteur définie peut être considéré, en première approximation, comme périodique. En synthèse par table d'onde simple, le son est généré par un oscillateur numérique qui diffère de ceux étudiés en 1.2.1 par le contenu de la table: la sinusoïde est remplacée par une fonction composite représentative du son synthétisé. Outre les problèmes inhérents aux calculs de phases et à l'interpolation évoqués en 1.2.1, la lecture de tables d'ondes non sinusoïdales pose un problème de repliement de spectre (*aliasing*).

Considérons, par exemple, une table d'onde comportant 32 harmoniques et une fréquence d'échantillonnage de 48kHz. En théorie, cette table ne pourra pas être transposée sans repliement de spectre au delà d'une fréquence fondamentale de  $24\text{kHz}/32 = 750\text{Hz}$ .

En pratique, on ne tient compte que du repliement de spectre audible. Considérons un convertisseur numérique analogique cadencé à une fréquence  $2f_N$  et suivi d'un filtre anti-repliement de fréquence de coupure  $f_c$ . Le repliement de fréquence ne sera physiquement audible qu'à partir d'une transposition d'un facteur  $2f_N/f_c - 1$ . Si  $f_N = 1.5f_c$ , la transposition maximale est donc d'une octave (figure 1.13).

L'interpolation linéaire peut aussi être avantageusement remplacée par une interpolation d'ordre plus élevé — équivalent à un filtre interpolateur ou décimateur — limitant distorsion et repliement de spectre [Ros92a].

La synthèse par tables d'ondes n'est pas très satisfaisante à elle seule. Elle sert souvent de source de signal en synthèse soustractive.

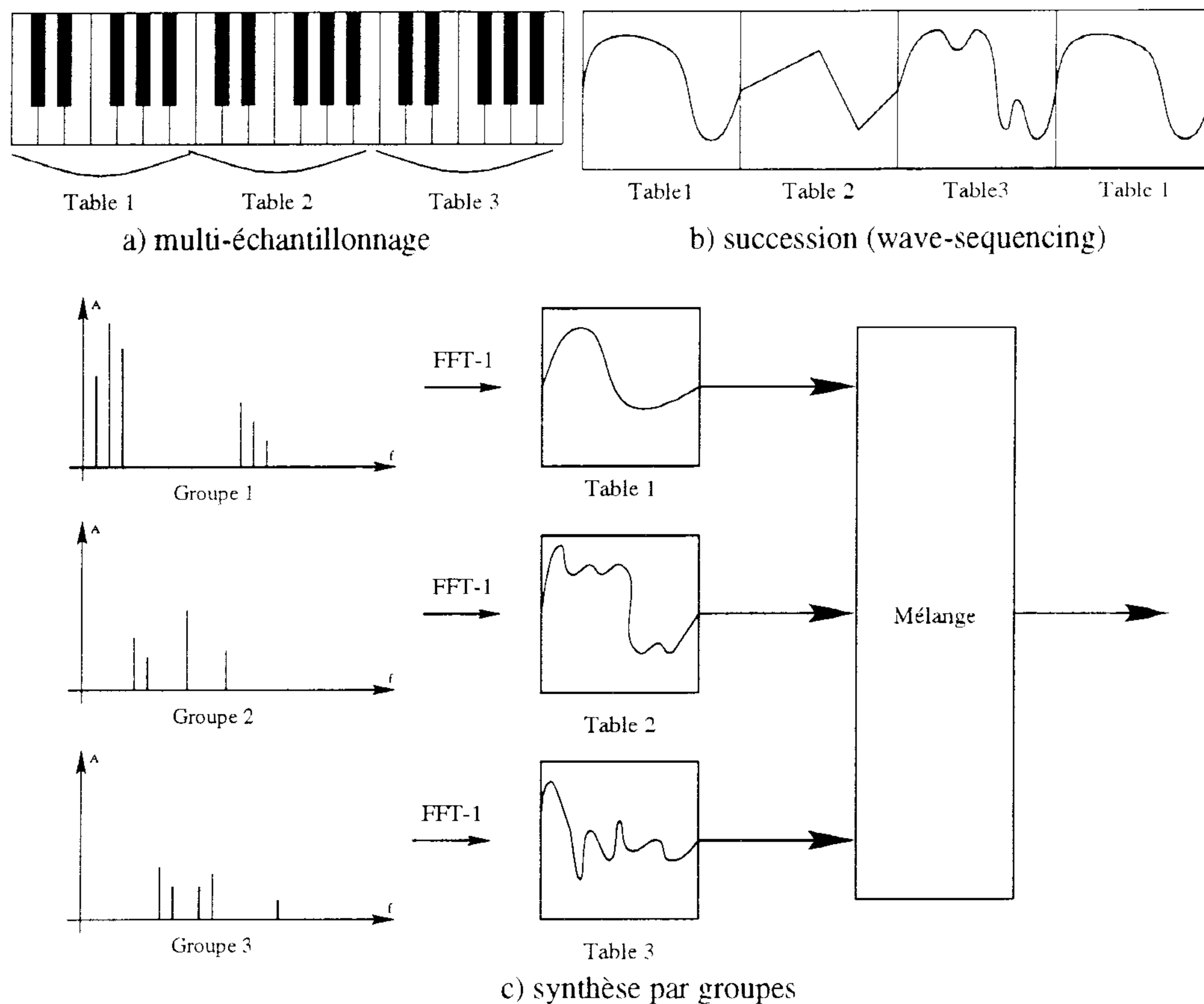


FIG. 1.14 – Formes de synthèse par tables d'ondes multiples

### 1.2.3 Synthèse par tables d'ondes multiples

La synthèse par tables d'ondes peut être sensiblement améliorée en combinant plusieurs tables d'ondes. Ces combinaisons peuvent prendre plusieurs formes : *multi-échantillonnage*, *succession*, *enchaînement*, *synthèse par groupes* et *combinaison dynamique*.

Le multi-échantillonnage (figure 1.14a) consiste à affecter plusieurs tables d'ondes sur l'étendue du clavier. Cette méthode permet de tenir compte des variations spectrales du son lors des transpositions. Elle limite ainsi les distorsions de timbre. Une variante du multi-échantillonnage consiste à utiliser différentes tables en fonction de la vitesse d'enfoncement de la touche (vitesse MIDI). Cette variante permet de restituer les variations de timbre liées à la dynamique du son. Ces deux types de multi-échantillonnage multiplient le nombre de tables d'ondes nécessaires. Afin de limiter la dimension des mémoires ROM destinées à leur stockage, on peut envisager de les synthétiser en léger temps différé par filtrage formantique — filtrage statique correspondant à la réponse de la caisse de résonance de l'instrument synthétisé — d'une table unique et de les charger dans une mémoire RAM [Shi95].

La synthèse par succession de tables (*wave-sequencing*, figure 1.14b) consiste à lire différentes tables séquentiellement afin d'obtenir un spectre évolutif. Cette forme de synthèse a notamment été utilisée par les firmes PPG et Korg.

La synthèse par enchaînement de tables consiste en un fondu enchaîné entre plusieurs

tables correspondant au spectre du son à des instants clés. Il s'agit d'une amélioration de la méthode précédente réduisant le bruit induit par le séquençement brutal des tables. Cette méthode est souvent combinée au multi-échantillonnage.

La synthèse par groupes (figure 1.14c) consiste à combiner différentes tables obtenues par regroupement d'harmoniques [Kle89]. Tous les harmoniques d'un groupe suivent donc les mêmes enveloppes d'amplitude et de fréquence. Les tables construites selon les groupes d'harmoniques sont intrinsèquement indépendantes (orthogonales). Bien que ce principe ait été commercialement exploité dans les années 80<sup>2</sup>, la principale difficulté consiste toujours à déterminer les regroupements nécessaires pour synthétiser un son donné. De nombreuses méthodes d'analyses telles que les algorithmes génétiques [CH96] sont évaluées à l'heure actuelle pour réaliser automatiquement ces regroupements.

La synthèse par combinaison dynamique de tables consiste à mélanger trois ou quatre tables. Ce mélange est commandé de façon dynamique selon l'action de l'interprète. Cette méthode peut être à la fois interprétée comme une généralisation de la synthèse par groupes et de la synthèse par enchaînement.

En raison de leurs faibles coûts de calcul, les trois dernières méthodes font l'objet d'études importantes. Ces études portent essentiellement sur la détermination de l'ensemble des tables d'ondes nécessaires à la synthèse d'un son particulier. De telles méthodes d'analyse ont été étudiées et évaluées notamment par Gregory Sandell, Andrew Horner, Lippold Haken et James Beauchamp [SM95, HBH93, BH95, HB95, Hor95, CH96, Hor97, YH97].

La méthode de synthèse que nous proposons dans ce mémoire (§2, page 63) peut être interprétée comme une forme de synthèse par combinaison dynamique. Cependant, notre étude se distingue des précédentes par :

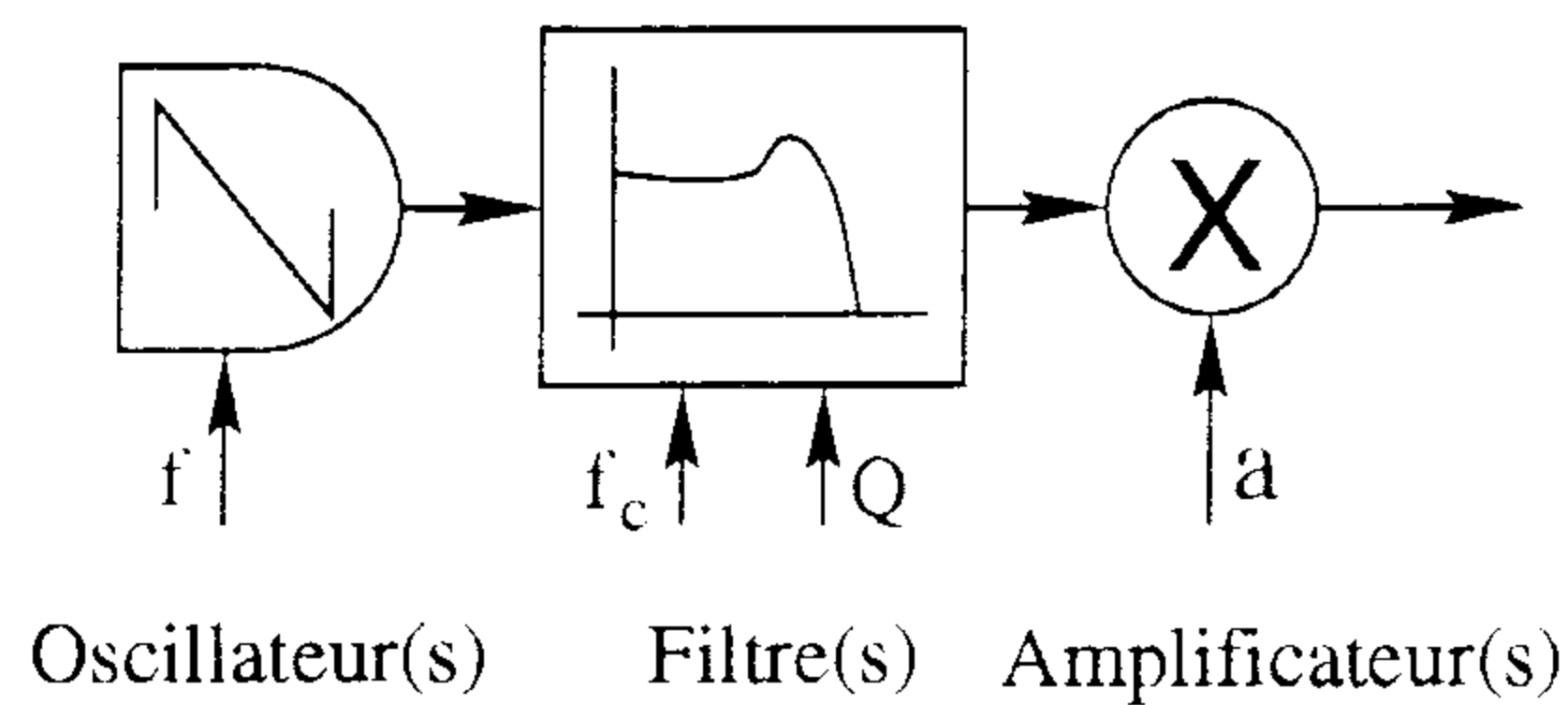
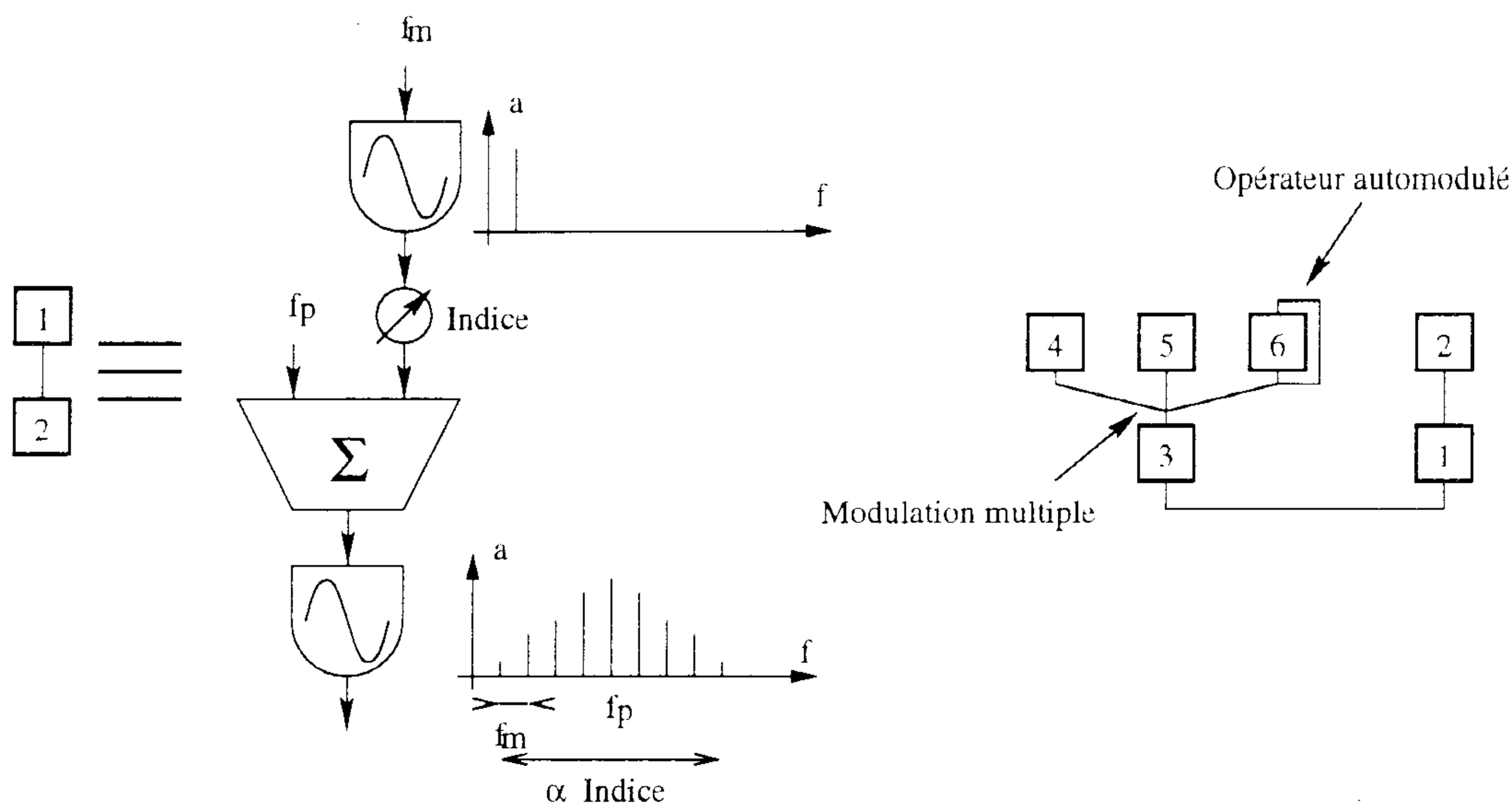
- l'utilisation de l'*analyse en composantes principales* (§2, page 63). Cette méthode a été écartée par certaines études précédentes au profit des algorithmes génétiques. Nous discutons plus avant de cette question dans les chapitres suivants ;
- l'attention portée à l'*interprétation des données réduites* (§2.3.1, page 85). Dans ce cadre, l'analyse en composantes principales devient un outil d'étude du timbre musical ;
- l'*extension et la généralisation* de la méthode aux sons *inharmoniques* et *bruités* (§3.1, page 113).

#### 1.2.4 Synthèse soustractive

La synthèse soustractive procède selon le modèle source/filtre. Un signal périodique, apériodique ou composite est filtré par un banc de filtres. Les caractéristiques de ces filtres, souvent variables dans le temps, définissent le timbre du son synthétisé. Théoriquement, cette méthode permet de synthétiser une grande variété de sons musicaux ou vocaux. Dans la pratique, c'est-à-dire dans le cadre des synthétiseurs commercialisés, la synthèse

---

2. le synthétiseur "K5" de la firme "Kawai" (1987) — improprement qualifié de synthétiseur additif — procédait par synthèse par quatre groupes.

FIG. 1.15 – *Synthèse soustractive*FIG. 1.16 – *Synthèse par modulation de fréquence*

soustractive est combinée à la synthèse par lecture de tables d'ondes pour lui donner plus de dynamique. Cette forme de synthèse et les architectures dédiées font toujours l'objet d'une activité de recherche industrielle importante [Ros92b, Sai93, MW93, Fuj96b, Fuj96a]. Dans ce cadre, elle est encore trop souvent bornée à un filtrage résonant passe-bas à quatre pôles.

L'utilisation de filtres multipôles, capables de simuler correctement les phénomènes de résonance reste rare en raison de leur coût de calculs. Une solution à ce problème consiste à réaliser ces calculs en léger différé pour produire des tables d'ondes utilisées en temps réel en synthèse par multi-échantillonnage [Shi95]. Une autre solution réside dans l'élaboration de circuits efficaces dédiés à cette application [Ros92b, IMS97]. Une autre difficulté de ce type de synthèse réside en l'absence d'algorithmes d'analyse efficaces permettant de séparer source et filtre. Par exemple, le codage par prédiction linéaire (LPC) se limite intrinsèquement à un filtrage tout-pôles. Des méthodes d'analyses plus adaptées aux sons musicaux conduisant à une synthèse par filtrage pôles/zéros font l'objet de recherches industrielles [MR93].

### 1.2.5 Synthèse par modulation de fréquence

La synthèse par modulation de fréquence permet de générer et de contrôler un signal riche en harmoniques avec un coût de calcul minimum. Le principe de la synthèse par modulation de fréquence à deux opérateurs (figure 1.16) a été proposée dès 1973 [Cho73] puis brevetée en 1977 [Cho77] par John Chowning.

Cette forme de synthèse consiste à moduler la fréquence d'un oscillateur porteur  $\omega_p$  par un oscillateur modulateur  $\omega_m$  selon la formule :

$$x(t) = \sin \left[ \left( \omega_p + I \omega_m \int_0^t \sin(\omega_m \tau) d\tau \right) t \right] \quad (1.8)$$

La richesse et les propriétés spectrales du son synthétisé dépendent des rapports de fréquences des oscillateurs et de l'indice de modulation. Elles sont données par l'expression :

$$x(t) = \sum_{n=0}^{+\infty} J_n(I) \sin[(\omega_p + n\omega_m)t] + \sum_{n=1}^{+\infty} (-1)^n J_n(I) \sin[(\omega_p - n\omega_m)t] \quad (1.9)$$

Deux extensions majeures ont été apportées à cette forme de synthèse avant sa commercialisation :

- l'utilisation d'un nombre plus important d'oscillateurs (quatre ou six) associés selon diverses architectures. Cette extension permet de multiplier les possibilités offertes par le procédé, elle en multiplie aussi la complexité ;
- l'utilisation d'oscillateurs s'automodulants (*feedback*) [HNY81, NNO81, Pee96]. Notons que cette notion a été présentée dès 1969 par Jean-Claude Risset dans "Introductory catalog of computer generated sounds" [Roa95, page 244].

La synthèse par modulation de fréquence est réputée être très compliquée auprès des musiciens en raison de son aspect mathématique peu intuitif (fonctions de Bessel, repliement de spectre en zéro, battements). Le succès des synthétiseurs de la firme Yamaha (série DX) dans les années quatre-vingts était essentiellement dû à leur fiabilité, aux nombres de sons préprogrammés disponibles plutôt qu'aux possibilités de synthèse. Ce manque d'ergonomie est la raison principale de la désaffection de cette forme de synthèse dès l'apparition de synthétiseurs numériques utilisant les deux précédentes.

La synthèse par modulation de fréquence peut encore être envisagée comme un élément de synthèse dans un générateur élaboré — combinant diverses formes de synthèse — tel qu'un synthétiseur formantique [LD95, MT97, Puc95]. Quelques variantes "assymétriques" de la modulation de fréquences peuvent aussi être envisagées [TG93, TGLT94].

### 1.2.6 Synthèse par formes d'ondes formantiques

La synthèse par formes d'ondes formantiques est conçue selon le modèle source/filtre de la synthèse soustractive (§1.2.4, page 36). Une *forme d'onde formantique* (FOF) est définie par la réponse impulsionnelle d'un filtre. La réponse d'un filtre du second ordre donne un exemple simple de FOF :

$$fof(t) = A e^{-t/T_{FOF}} \sin(2\pi t/P_{FOF} + \phi) \quad (1.10)$$

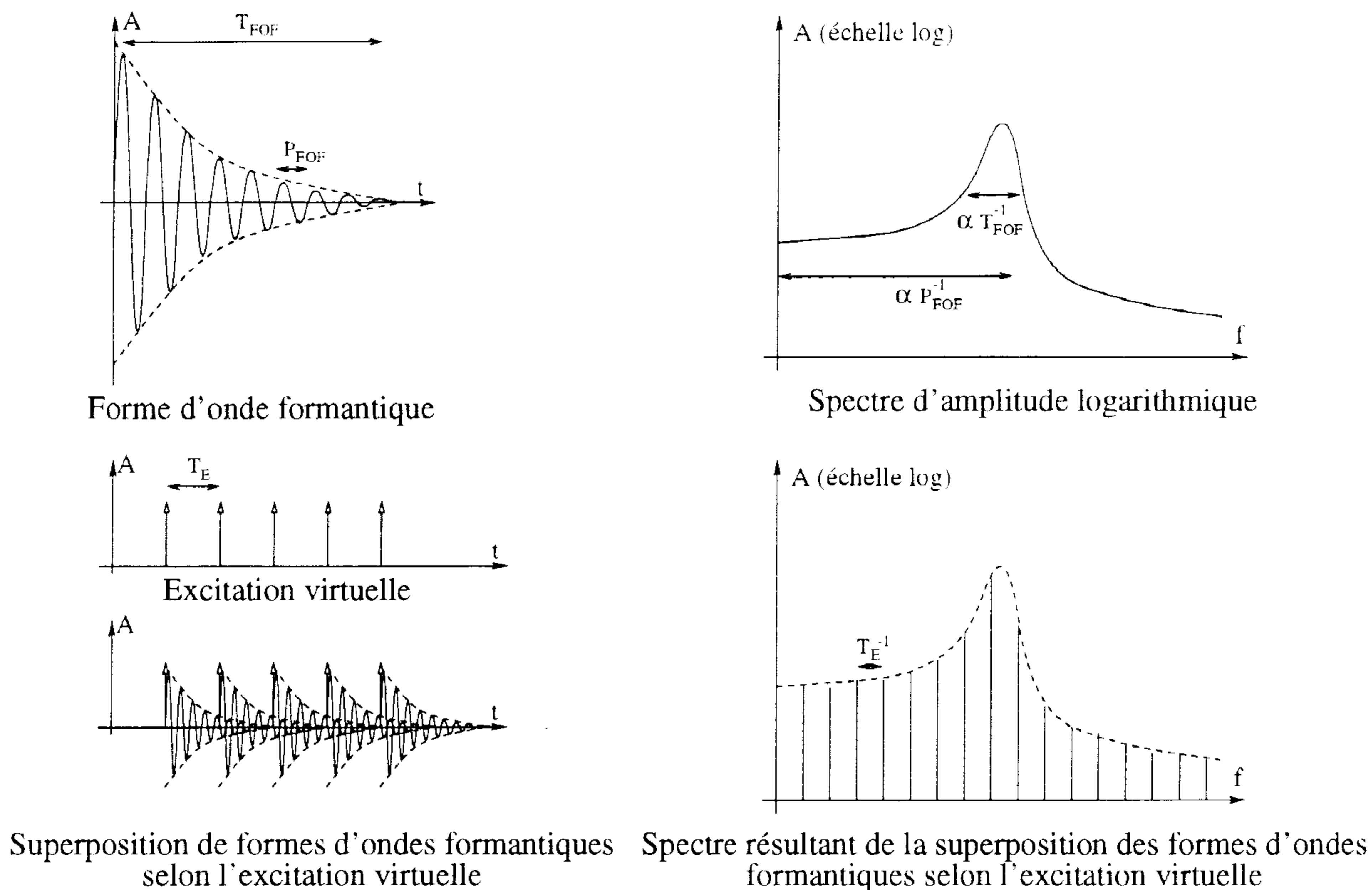


FIG. 1.17 – Synthèse par formes d'ondes formantiques

La synthèse par FOF simplifie le processus de filtrage en ne considérant comme source d'excitations que des signaux impulsionnels. Un signal équivalent au signal filtré est synthétisé par superposition des FOF au rythme du train d'impulsions (figure 1.17).

Dans le domaine temporel, le signal de sortie est égal au produit de convolution du train d'impulsions par la FOF. Dans le domaine spectral, on obtient le produit d'un spectre de raies harmoniques par la transformée de Fourier de la FOF, c'est-à-dire par la fonction de transfert du filtre. Pour simuler la mise en parallèle de  $N$  filtres, il suffit de combiner  $N$  FOF dans le domaine temporel.

Les caractéristiques de la fonction de transfert sont directement liés aux fréquences et aux taux d'amortissement des FOF. Il est ainsi possible de synthétiser un son dans lequel le filtrage joue un rôle prépondérant tel que la voix chantée.

La synthèse par FOF peut être étendue de diverses manières :

- l'utilisation de FOF plus élaborées permet de commander un plus grand nombre de caractéristiques du filtre modélisé ;
- l'utilisation de signaux d'excitation impulsionnels à temps aléatoire permet la synthèse de sons bruités tels que la voix chuchotée et certains transitoires [RdG92, Rd96, RdG93].

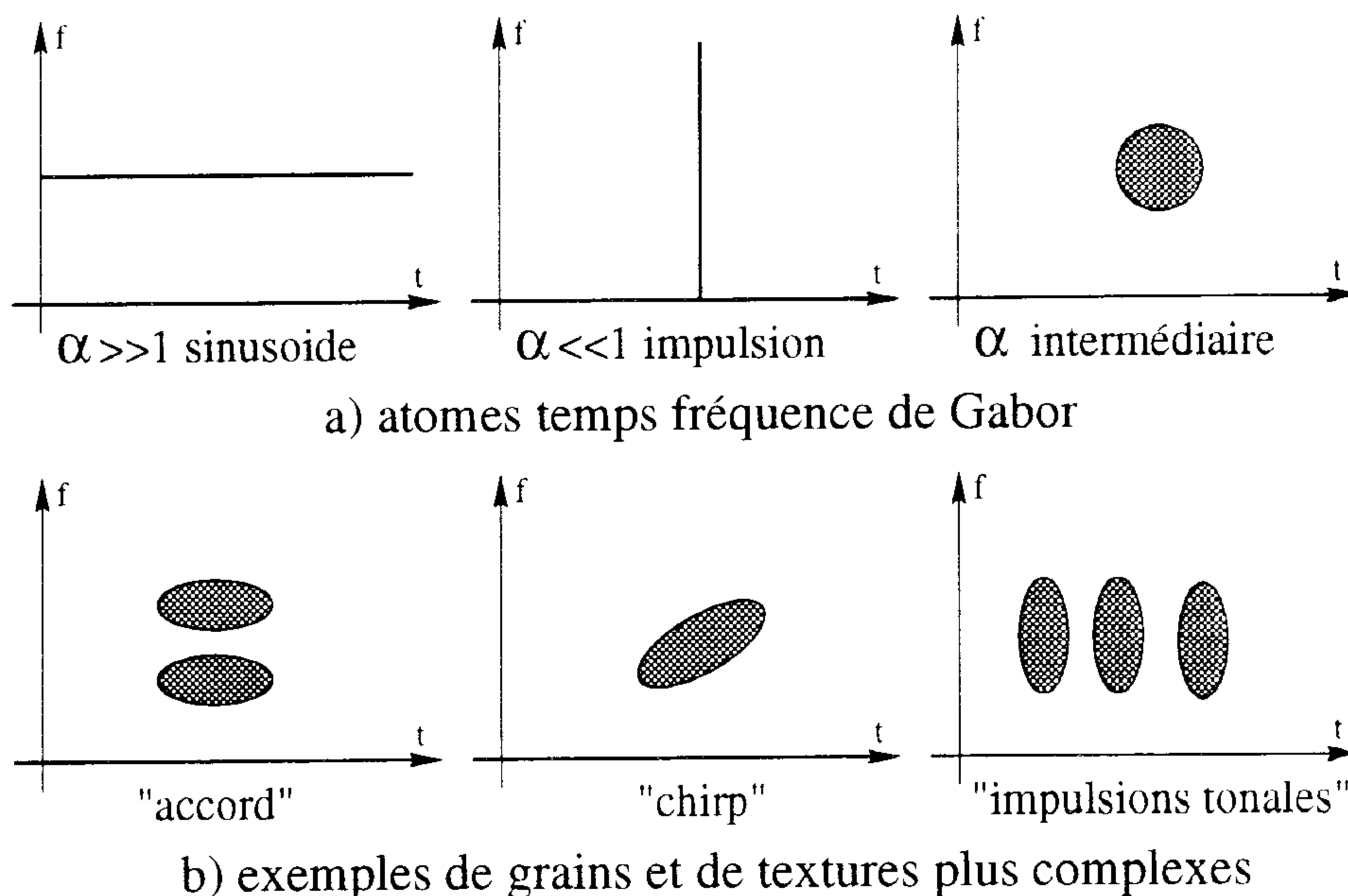


FIG. 1.18 - Différents types de grains

### 1.2.7 Synthèse granulaire

La synthèse granulaire est une forme de synthèse fortement apparentée à la synthèse par formes d'ondes formantiques. La synthèse granulaire considère qu'un signal acoustique est constitué d'une somme de "grains".

L'exemple de grain le plus simple est un atome temps/fréquence de Gabor [Roa88] (figure 1.18a):

$$G_g(t_0, \omega_0; t) = e^{-\alpha^2(t-t_0)^2} e^{i\omega_0 t + i\phi} \quad (1.11)$$

un tel atome est localisé dans le plan temps/fréquence en  $t_0, \omega_0$ , il "s'étale" sur ce plan suivant le principe de concentration spectro-temporelle:

$$\Delta t \Delta \omega = \pi \quad (1.12)$$

La constante  $\alpha$  règle le rapport entre  $\Delta t$  et  $\Delta \omega$ , il détermine la "forme" de l'étalement. Une valeur faible de  $\alpha$  engendre un signal sinusoïdal fenêtré par une gaussienne de durée importante bien localisé en fréquence et "étalé" sur l'axe temporel. Au contraire, une valeur importante de  $\alpha$  engendre un signal impulsionnel très localisé dans le temps.

D'autres formes de grains sont utilisables parmi lesquelles on peut compter (figure 1.18b):

- accord, assemblage vertical d'atomes de Gabor;
- série d'impulsions tonales, succession horizontale d'atomes de Gabor;
- "chirp", atome de Gabor modulé en fréquence [JP88] (§1.2.1, page 32).

La synthèse à proprement parler est effectuée en disposant les grains dans le plan temps fréquence [Tru88]. Cette disposition permet de constituer des "textures". Une texture

régulière engendre un son périodique à hauteur définie. Une texture irrégulière engendre un son bruité.

La synthèse de certains grains peut être réalisée par un algorithme récursif. Par exemple, la série complexe  $(X_n)_{n=1,\dots,N}$  définie par :

$$X_0 = \exp(C_0) \quad (1.13)$$

$$Y_1 = \exp(C_1 + C_2) \quad (1.14)$$

$$Y_{n+1} = Y_n \exp(2C_2) \quad (1.15)$$

$$X_{n+1} = Y_{n+1} X_n \quad (1.16)$$

engendre le signal sinusoidal gaussien complexe :

$$X_n = A \exp(C_0 + C_1 n + C_2 n^2) \quad (1.17)$$

La partie réelle de ce signal est un signal sinusoidal modulé en fréquence (“chirp”) à enveloppe gaussienne [JP88].

Pour conclure, la synthèse granulaire peut être interprétée comme une forme généralisée de synthèse par FOF. Pour ce faire, il suffit d’interpréter chaque grain comme la réponse impulsionnelle d’un filtre plus ou moins complexe. Dans cette perspective, la disposition d’un grain dans le temps correspond virtuellement à l’excitation par une impulsion du filtre associé.

### 1.2.8 Synthèse par distorsion (*wave-shaping*)

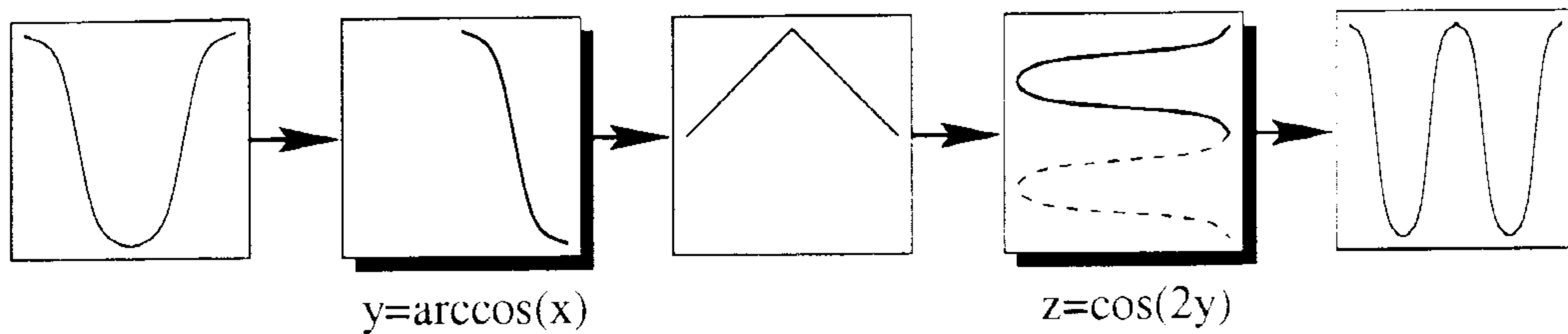
La synthèse par distorsion, souvent dénommée *wave-shaping*, consiste à distordre un signal  $s(t)$  selon une fonction de distorsion  $g$  donnée. Le résultat de cette distorsion  $g \circ s(t)$  est un signal dont le spectre est plus ou moins affecté. La nature de la modification spectrale dépend :

- de la fonction de distorsion  $g$  :
- de la forme d’onde du signal d’entrée  $s(t)$  :
- de l’amplitude du signal d’entrée :
- de la composante continue (*offset*) du signal d’entrée  $\bar{s}(t)$ .

L’avantage de la synthèse par distorsion est son faible coût de calcul. En effet, elle ne consiste qu’en la consultation d’une table (*look-up table*) — éventuellement agrémentée d’une interpolation — contenant la fonction de distorsion. A ce propos, on pourra noter que la synthèse par distorsion est équivalente à la lecture de tables d’ondes lorsque le signal d’entrée est une dents-de-scie.

L’inconvénient majeur de cette forme de synthèse est qu’il est en général difficile d’en prévoir l’effet sur une forme d’onde quelconque. La distorsion introduit de nouveaux harmoniques dont les fréquences et les amplitudes peuvent être importantes. Ce point ne doit pas être négligé en synthèse numérique car celle-ci exige de travailler avec des signaux



FIG. 1.19 – *Waveshaping par un polynôme de Tchébychev*

à bande limitée sous peine d'obtenir une distorsion, indésirable celle-là : le repliement de spectre.

La distorsion par polynômes, ou par combinaison de polynômes, de Tchébychev de signaux sinusoïdaux est un rare exemple pour lequel on maîtrise les effets spectraux de la distorsion. La série des polynômes de Tchébychev est définie par :

$$T_0(x) = 1 \quad (1.18)$$

$$T_1(x) = x \quad (1.19)$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad n \geq 1 \quad (1.20)$$

Le terme général de cette série peut être exprimé sous la forme :

$$T_n(x) = \cos(n \arccos x) \quad (1.21)$$

La distorsion par polynômes de Tchébychev est donc équivalente à l'application successive d'une fonction arccosinus — qui transforme le signal sinusoïdal en signal triangulaire — et d'une table d'onde en cosinus de fréquence multiple  $n$  (figure 1.19).

La synthèse par distorsion est très rarement utilisée seule. Elle peut être combinée à divers autres procédés pour obtenir des formes de synthèse élaborées telles que certains types de synthèse formantique [Puc95], de synthèse par modélisation physique ou encore de synthèse par formule de sommation discrète.

### 1.2.9 Synthèse par formule de sommation discrète

La synthèse par formule de sommation discrète (DSF : Discrete Summation Formula) a pour objectif de synthétiser des signaux numériques à bande limitée. Considérons l'expression :

$$\sum_{k=0}^{N-1} z^k = \frac{1 - z^N}{1 - z} \quad (1.22)$$

En développant cette expression selon l'identité  $2i \sin(x) = e^{ix} - e^{-ix}$  il vient :

$$\sum_{k=0}^{N-1} a^k \sin(kft) = \left[ \sin(ft) + a^N \sin(Nft) + a^{N+1} \sin((N-1)ft) \right] \times \frac{1}{1 - 2a \cos(ft) + a^2} \quad (1.23)$$

Il est ainsi possible de synthétiser un signal correspondant à la sommation discrète de  $N - 1$  sinusoïdes définie par l'expression de gauche avec l'expression de droite. Cette

expression consiste en 4 lectures de tables d'ondes, trois multiplications et un *wave-shaping* ( $w(z) = 1/1 - 2az + a^2$ ). Notons que le coût de calcul de cette méthode est indépendant du nombre d'harmoniques présents dans le signal.

Cette formulation rend possible la génération de signaux harmoniques de richesse variable — réglée par  $a$  — et comportant un nombre d'harmoniques commandé par  $N$ . Ce type de générateur peut donc fournir des signaux riches en harmoniques, exempts de repliement de spectre — il suffit pour cela de limiter convenablement  $N$  — fort utiles en synthèse soustractive [SS96b, SS96a].

### 1.2.10 Synthèse par modélisation physique

Cette forme de synthèse repose sur des modèles mathématiques d'instruments acoustiques. Il convient de distinguer deux catégories de modélisation physique.

La première consiste à élaborer des modèles de plus en plus fins d'instruments acoustiques dans le but d'étudier et de comprendre leur fonctionnement. Ces études permettent d'étendre nos connaissances en acoustique instrumentale et apportent de nouveaux outils aux luthiers et aux facteurs.

La seconde consiste à réaliser une synthèse convaincante de sons "acoustiques" en temps réel. Cette classe de synthèse par modélisation physique est un compromis entre la rigueur de la première et les possibilités offertes par les processeurs de traitement du signal disponibles. L'apparition de processeurs dédiés suffisamment puissants a permis le développement et la commercialisation de tels synthétiseurs à partir de 1993.

Les éléments de base d'un instrument modélisé sont typiquement :

- une source d'excitation non-linéaire ;
- un ou plusieurs guides d'ondes ;
- un ou plusieurs filtres ;
- des fonctions de couplage non-linéaires.

Synthèse soustractive et synthèse par modèle physique reposent toutes les deux sur le modèle excitation/résonance. La principale différence étant le couplage non-linéaire entre excitation et résonateur. Ces couplages — similaires aux distorsions vues en 1.2.8 — doivent être considérées avec le plus grand soin et ne pas impliquer de repliement de spectre.

La principale qualité de la modélisation physique est son aptitude à synthétiser des transitoires parfois très réalistes. Ses principaux écueils demeurent la complexité et le coût de calcul.

### 1.2.11 Implantations et processeurs

L'implantation pratique d'une méthode de synthèse et les problèmes pratiques qui lui sont liés déterminent en majeure partie son exploitation. Le succès commercial de la synthèse FM en est un excellent exemple : cette méthode tirait profit de l'électronique numérique des années 80. Les paragraphes qui suivent sont une brève présentation des supports d'implantation actuels.

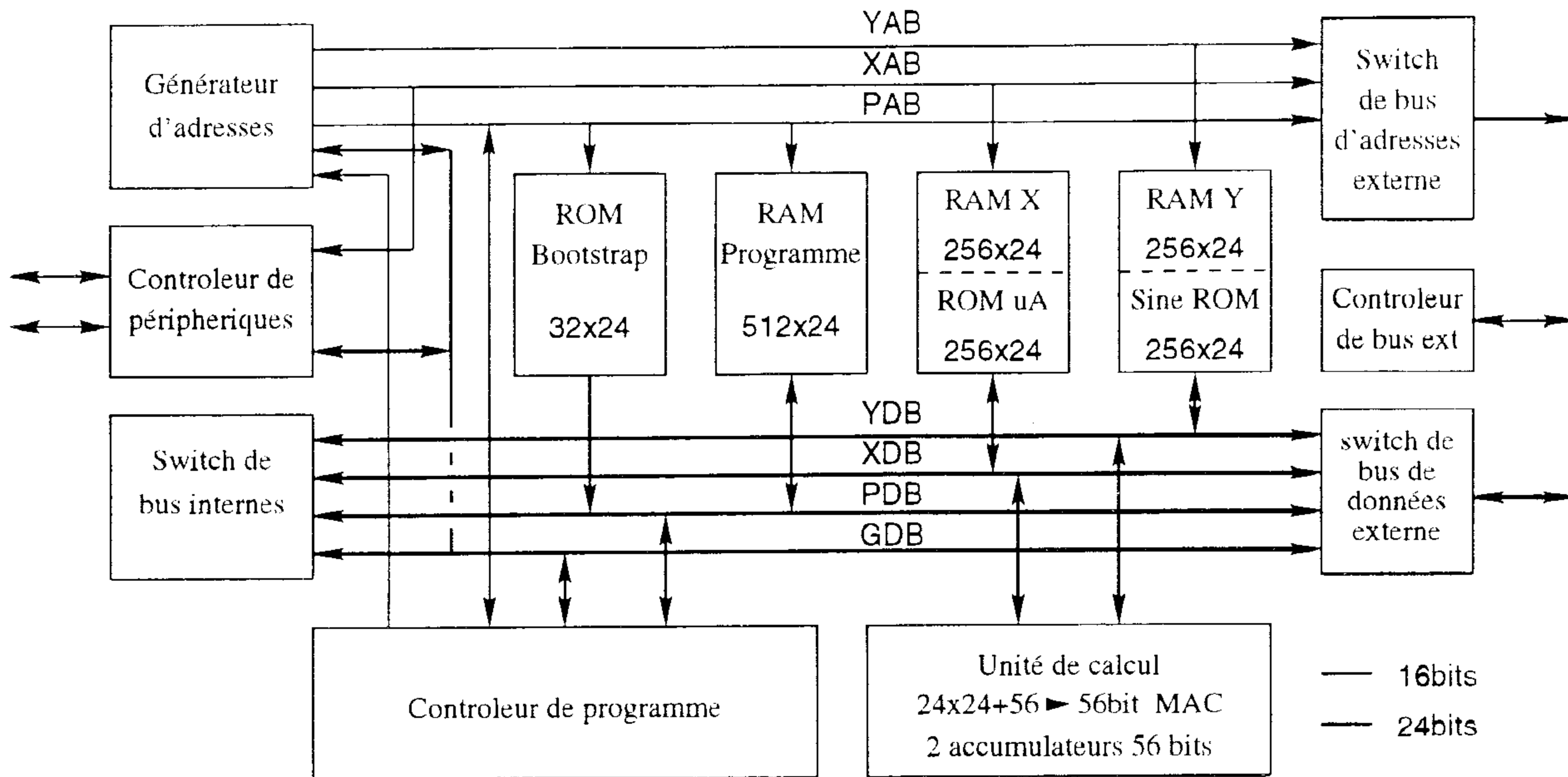


FIG. 1.20 - Architecture du DSP 56001 Motorola

## Processeurs et ordinateurs généralistes

Le jeu d'instructions et la conception physique des processeurs généralistes sont rarement optimisés pour les applications temps réel audio. Les ordinateurs généralistes, de par la disponibilité de langages de haut niveau, tels que le C, permettent la mise au point et le développement d'algorithmes d'analyse et de synthèse par calcul en temps différé. La réalisation sur ordinateur généraliste constitue souvent la première étape de mise au point d'un algorithme d'analyse ou de synthèse.

## Processeurs numériques de signaux (DSP)

Les processeurs numériques de signaux (DSP : Digital Signal Processor) sont une famille de processeurs dédiés au traitement du signal. Leurs architectures et leurs jeux d'instructions sont optimisés pour réaliser efficacement les opérations les plus courantes du traitement du signal : convolution, transformée de Fourier rapide et filtrage. Les instructions privilégiées dans ce type d'opérations sont la multiplication/accumulation et les déplacements de données. Les déplacements et les calculs sont généralement effectués en parallèle par une unité de gestion d'adresses et par une unité de calcul. Le 56001 de Motorola, largement utilisé dans le domaine audio, est conçu selon ce modèle (figure 1.20).

## Processeurs numériques de signaux spécialisés

Si les DSP standards sont efficaces pour réaliser certains traitements tels que réverbérations, échos et autres effets, ils montrent quelques faiblesses lors de la réalisation d'algorithmes de synthèse. La principale faiblesse tient à la dissociation de la gestion des adresses et des calculs. L'opération la plus élémentaire en synthèse est la lecture de tables d'ondes évoquée en 1.2.2. Cette opération nécessite le calcul de la phase par accumulation

du terme de fréquence et son utilisation, partielle ou complète, en pointeur d'adresse. On imagine les difficultés que représente une telle opération sur un DSP56001 dans lequel les registres d'adresse sont codés en entiers non signés sur 16bits et les registres d'accumulation en virgule fixe sur 56bits : la moitié des instructions se consacre à adapter les formats de données internes [Chr90].

L'avènement du multimédia et des applications audio grand public ont permis la production de DSP dédiés au traitement et à la synthèse de sons musicaux. De tels processeurs subviennent aux besoins de la lecture de tables d'ondes. Ils offrent de multiples spécificités telles que :

- un dimensionnement des unités de calcul adapté au traitement des signaux acoustiques ;
- la présence d'accumulateurs qui facilitent la génération de signaux de commande ;
- et la segmentation des programmes en "slots" — sous-programmes courts et disposants de registres propres — qui permet une gestion modulaire de la polyphonie [DH95].

### Circuits dédiés

La production industrielle de circuits dédiés, de par son coût de développement initial, suppose une production massive. Un tel volume de production suppose un marché important. Parmi les premiers circuits dédiés à la synthèse numérique des sons musicaux, on peut considérer les circuits dédiés à la synthèse FM de la série de synthétiseurs DX de la firme Yamaha. La généralisation de l'utilisation de la synthèse par lecture de tables d'ondes par les synthétiseurs et par les cartes audionumériques multimédia a permis le développement de circuits dédiés aux applications "grand public". L'évolution des techniques permet à ce jour la production de circuits qui intègrent toutes les fonctions d'un synthétiseur numérique : interface MIDI, mémoire et lecteur de tables d'ondes 64 voies, filtres et effet de réverbération [BBC<sup>+</sup>97].

La disponibilité d'outils logiciels (conception, simulation) et de circuits programmables, facilitent le développement et le prototypage de circuits dédiés. L'utilisation de circuits reconfigurables de façon dynamique (FPGA) est devenue envisageable pour les applications nécessitant des traitements spécifiques [VMS97].

### Conclusion

Les processeurs et les systèmes généralistes, de par leur souplesse, sont largement utilisés comme support de développement d'algorithmes. Les DSP généralistes permettent une première approche de mise au point temps-réel des algorithmes. Les DSP spécialisés permettraient, si des kits de développement complets étaient disponibles, de pousser plus avant cette approche temps-réel grâce à un jeu d'instructions adapté. Le développement de circuits dédiés reste encore très coûteux malgré la disponibilité de circuits programmables de type FPGA.

### 1.2.12 Conclusions

De nombreuses méthodes permettent de synthétiser des sons musicaux. Seule une petite proportion de ces méthodes est exploitée commercialement. En effet, les critères qui rendent une forme de synthèse exploitable ne se limite pas simplement à la qualité de la synthèse.

Le coût de calcul, et implicitement le coût économique, d'une méthode reste un facteur capital (le succès de la synthèse par modulation de fréquence en est un exemple). Un algorithme de synthèse, s'il a l'ambition d'être commercialement exploitable, devra être conçu en tenant compte de ces "basses" considérations.

Un second critère réside dans les possibilités offertes par le synthétiseur : celui-ci doit être capable de fournir une large "palette" de sons. Ce critère explique en partie la désaffection des synthétiseurs analogiques et des synthétiseurs à modulation de fréquence au profit des synthétiseurs à lecture de tables d'ondes proposant des centaines — voire des milliers — de timbres différents.

Les qualités d'un synthétiseur commercialement viable sont donc les suivantes : bonne qualité de synthèse, faible coût de calcul, et, large palette de timbres. Les méthodes que nous avons énumérées couvrent en partie ces qualités et sont souvent combinées pour obtenir des formes de synthèses plus efficaces.

## 1.3 Quelques aspects de l'audition humaine

L'analyse et la synthèse que nous allons effectuer ont pour destinataire le système auditif humain. Nous en présentons brièvement dans un premier temps le fonctionnement physiologique. Dans un second temps, nous nous intéressons à quelques phénomènes psychoacoustiques importants concernant les attributs de hauteur, d'intensité, et surtout, de timbre.

### 1.3.1 Quelques aspects physiologiques de l'audition

#### L'oreille externe et l'oreille moyenne

La fonction principale de l'oreille externe, composé du pavillon et du conduit auditif, est de canaliser le son. Un rôle plus secondaire de l'oreille externe est une accentuation des fréquences aux alentours des 3kHz.

La fonction de l'oreille moyenne est d'adapter l'impédance acoustique du conduit auditif (milieu aérien) à l'impédance de l'oreille interne (milieu aqueux). L'oreille moyenne est isolée du conduit auditif par une membrane souple : le tympan. L'équilibre des pressions entre oreille moyenne et externe est (plus ou moins) assurée par la trompe d'Eustache. La vibration acoustique est transmise par la chaîne des osselets (le marteau, l'enclume et l'étrier). La raideur de la chaîne, et ainsi l'adaptation d'impédance acoustique, sont commandés par un système musculaire. En fin de chaîne, l'étrier transmet l'onde sonore à l'oreille interne par la fenêtre ovale (figure 1.21).

L'oreille externe permet une canalisation/accentuation acoustique des sons tandis que l'oreille moyenne permet une adaptation du milieu aérien au milieu aqueux de l'oreille

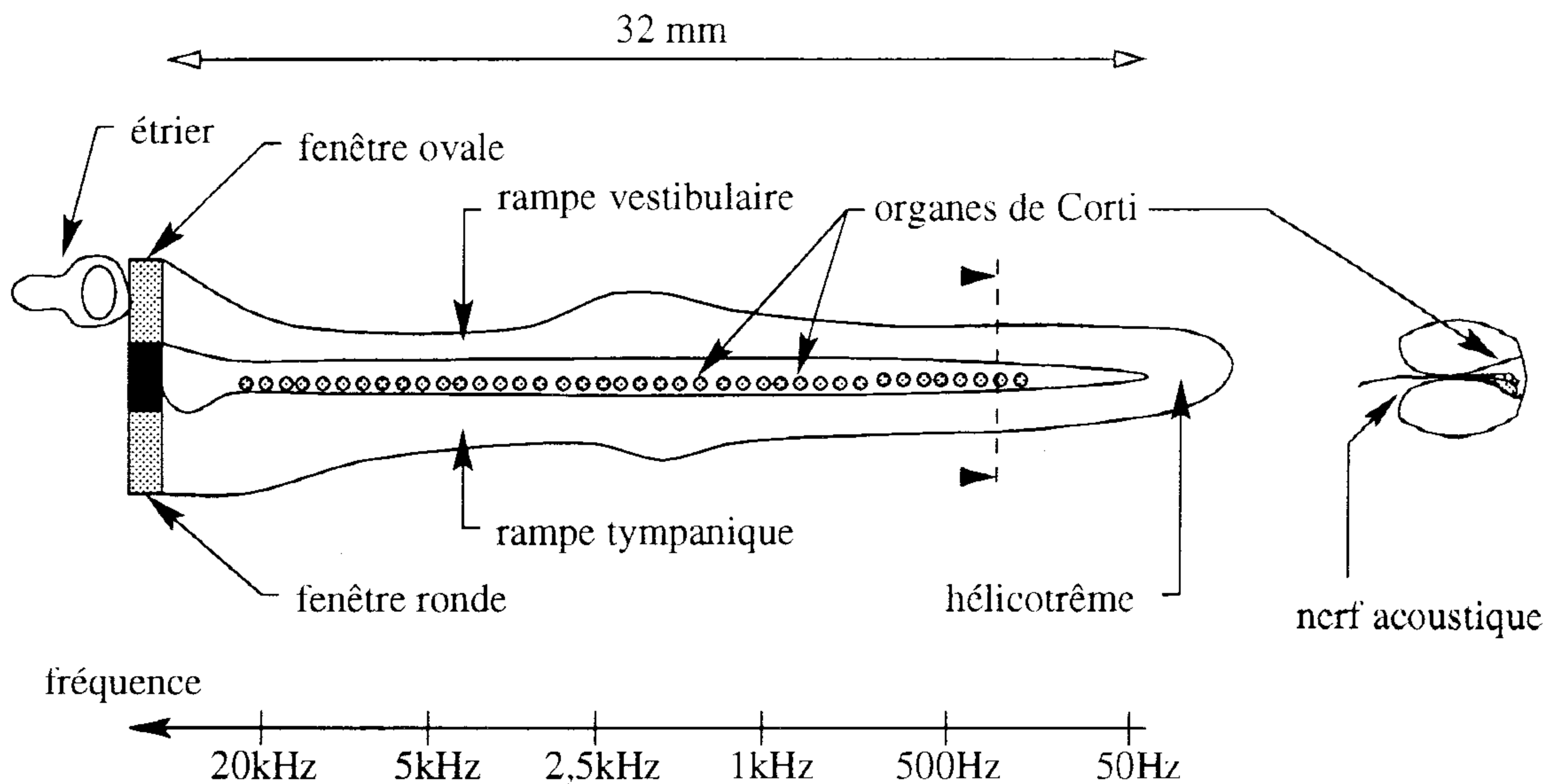


FIG. 1.21 – Cochlée déroulée (d'après E. Leipp [Lei84])

interne, siège de la perception.

### Fonctionnement mécanique de l'oreille interne

L'oreille interne est composée du limaçon, encore appelé cochlée. En première approximation, la cochlée joue le rôle d'un banc de résonateurs. Elle est séparée en deux rampes vestibulaire et tympanique (figure 1.21). Ces deux rampes contiennent un liquide aqueux : la périlymphe. La vibration acoustique se propage de l'étrier à la rampe vestibulaire au travers de la fenêtré ovale. Les rampes vestibulaires et tympaniques communiquent à l'hélicotreme, extrémité de la cochlée.

La propagation de l'onde acoustique provoque une déformation de la membrane basilaire et de la membrane de Reissner qui séparent les deux rampes. La localisation de la déformation dépend de la fréquence de l'onde acoustique (tonotopie).

### Physiologie de l'oreille interne

Les deux membranes — vestibulaire et tympanique — délimitent le canal cochléaire contenant de l'endolymphe, liquide épais, et les organes de Corti. La figure 1.22 montre une coupe transversale d'un organe de Corti [Puj90]. La transduction du signal acoustique en signal neuronal est effectuée par les cellules ciliées internes. L'émission d'une impulsion nerveuse est déclenchée par une déformation des cils de la cellule. L'information acoustique provenant des quelques 3.500 cellules ciliées internes est transmise au système nerveux central par le système afférent<sup>3</sup> radial du nerf cochléaire. L'activité de ces cellules est modérée par le système nerveux central au moyen du système efférent<sup>4</sup> latéral.

Le rôle des 12.000 cellules ciliées externes revêt un caractère indirect mais très important pour la perception des sons. En effet, ces cellules sont capables de provoquer

3. qui transmet l'influx nerveux de la périphérie vers les centres nerveux

4. qui transmet l'influx nerveux des centres vers la périphérie

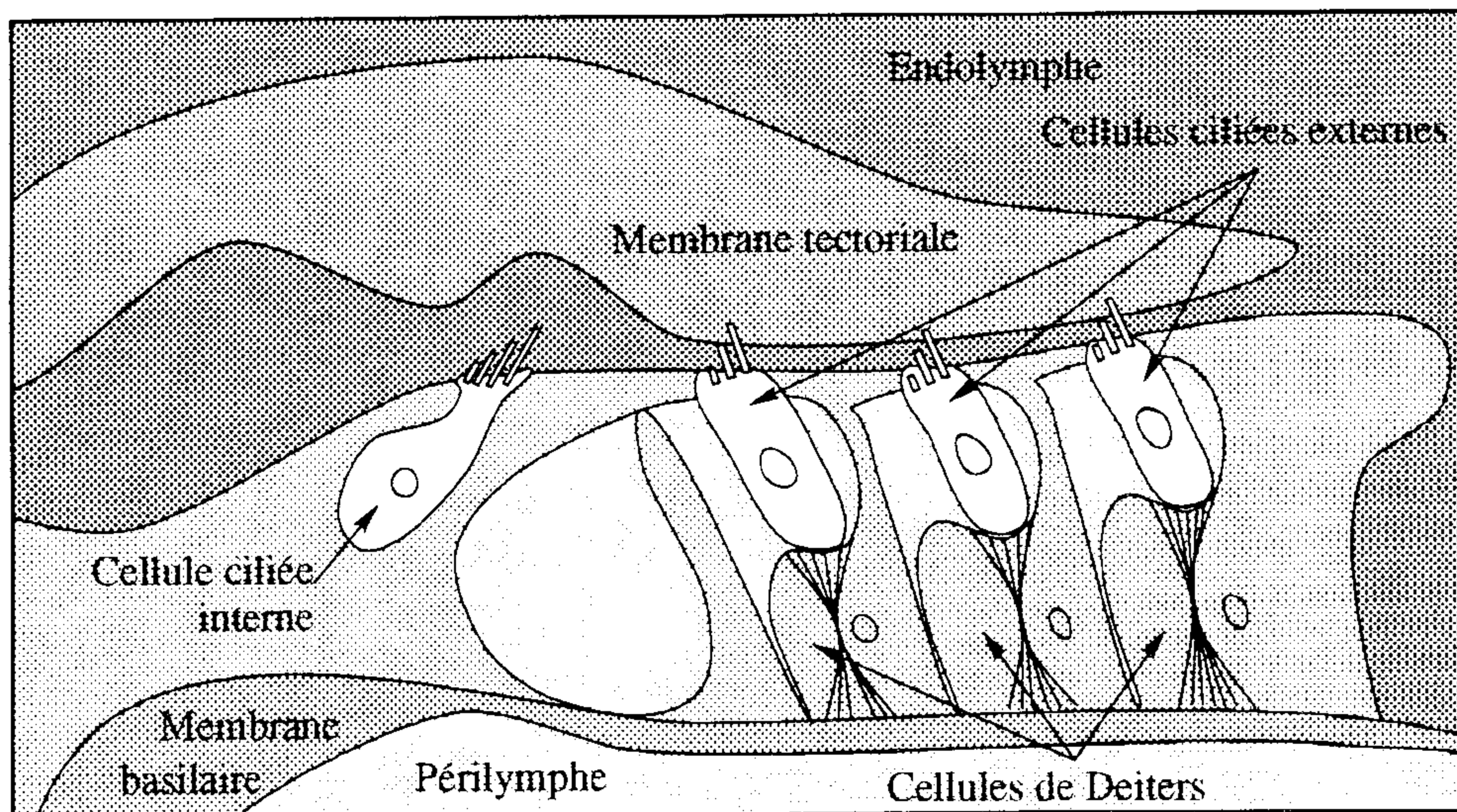


FIG. 1.22 – *Organe de Corti (d'après R.Pujol [Puj90])*

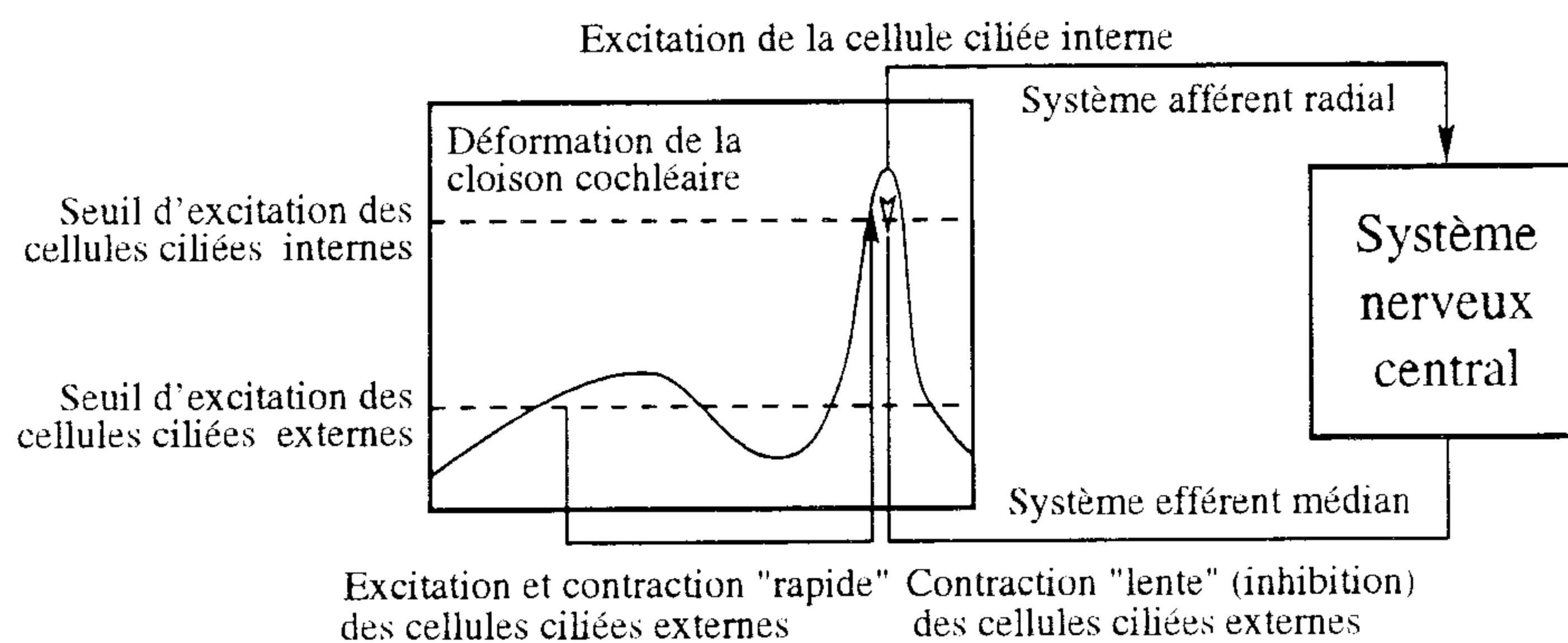


FIG. 1.23 – *Processus temporel actif de l'audition*

une contraction des cellules de Deiters sur lesquelles elles reposent. Ainsi, l'excitation des cellules ciliées externes provoque une contraction quasi-musculaire rapide qui accentue la déformation locale de la membrane basilaire. L'amplification active apportée par ces cellules peut atteindre les 50dB.

L'activité des cellules ciliées externes est modérée par le système efférent médian du nerf cochléaire qui provoque des contractions quasi-musculaires lentes (figure 1.23). L'activité amplificatrice des cellules ciliées externes, et donc les caractéristiques du "filtrage" effectué par l'oreille interne, est directement commandée par le système nerveux central.

L'oreille ne peut donc pas être assimilée à un simple banc de filtres statiques. C'est un organe actif dont la sensibilité est directement commandée par le système nerveux central. De plus, de nombreux "traitements" (corrélation, autocorrélation), plus ou moins bien connus, sont effectués par les centres nerveux sur les signaux acoustiques "bruts"

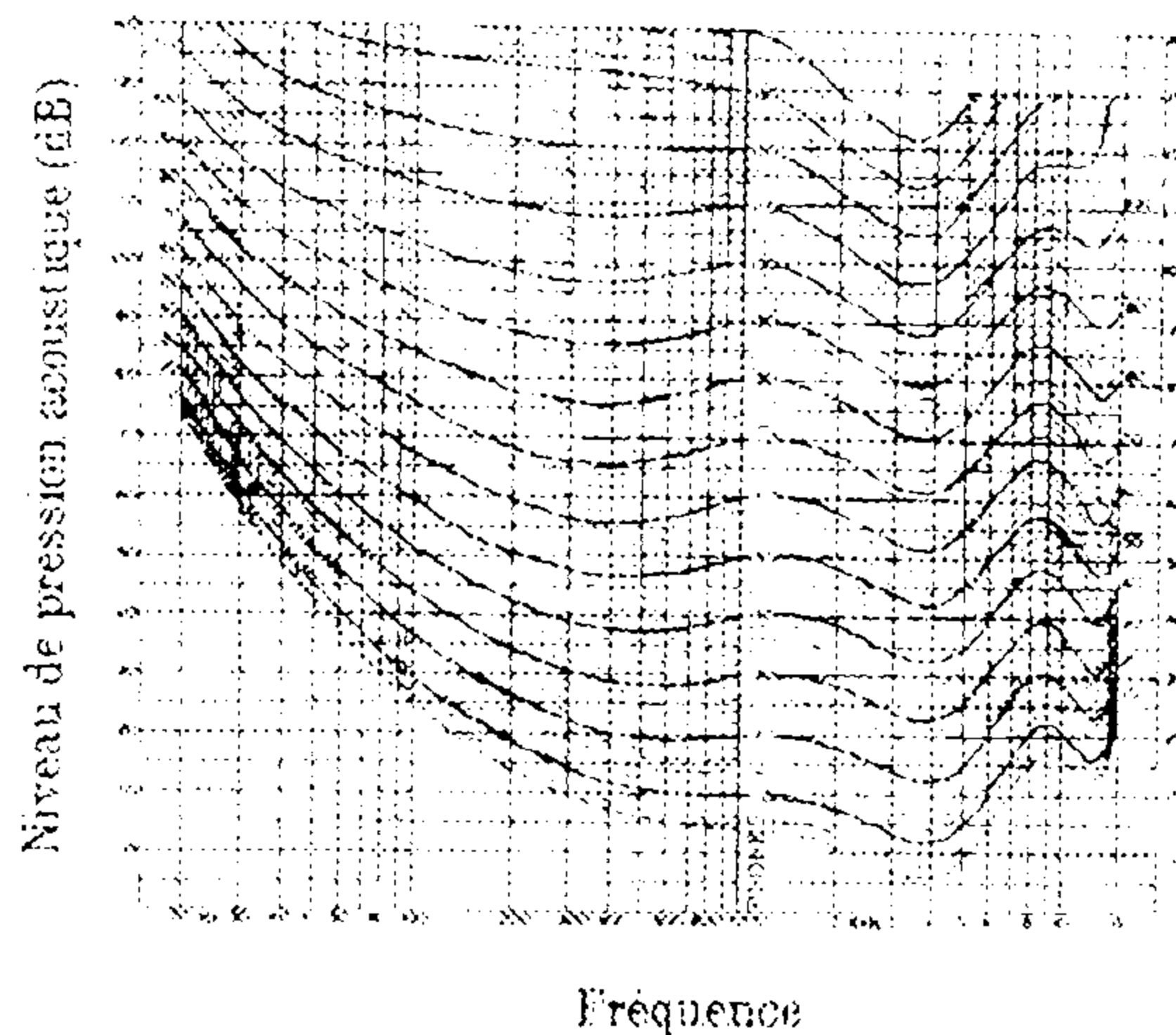


FIG. 1.24 – Courbes isosoniques pour l'écoute binaurale en champ libre, d'après la norme AFNOR NFS30-003

fournis par l'oreille interne.

L'audition de sons musicaux et par extension de musique est par nature un processus actif. Lors de l'audition d'une fugue de Bach, par exemple, l'auditeur est libre de "suivre" une ligne mélodique particulière ou de se laisser "porter" par l'harmonie globale de la pièce. La physiologie de l'oreille interne nous suggère que cette attention de l'auditeur va jusqu'à modifier le fonctionnement de l'oreille interne. La réaction de l'auditeur face à un son — ou à un ensemble de sons — affecte directement la physiologie de son oreille interne. Ce phénomène rend le processus d'audition aussi difficile à prévoir que les auditeurs eux-mêmes.

### 1.3.2 Quelques aspects psychoacoustiques de l'audition

La section précédente a mis en avant la complexité de l'audition sur le plan physiologique. Cette section introduit quelques résultats psychoacoustiques concernant trois attributs des sons musicaux : l'intensité, la hauteur, et le timbre.

#### Sensation d'intensité

La sensation d'intensité de sons purs est dénommée *sonie*. On dénomme *lignes d'isotonie* un ensemble de courbes établissant une relation entre intensité et niveau de pression acoustique pour l'ensemble des fréquences audibles [BCDS89].

Les conditions d'audition ont une influence importante sur les lignes d'isotonie. Ainsi, les lignes d'isotonies correspondant à une audition en champ libre (figure 1.24) présenteront quelques "irrégularités" qui n'apparaissent pas en audition sur écouteurs. Ces différences sont essentiellement de nature physique et correspondent au "filtrage" effectué par la tête et les pavillons des oreilles des auditeurs.

Outre sa complexité intrinsèque, ce filtrage dépend des positions relatives de l'auditeur et de la source. On désigne généralement ce filtrage par le sigle HRTF : Head Related



Transfer Function. Il est aujourd'hui établi que les HRTF sont un élément majeur de la perception tridimensionnelle [lsr95, WK89a, WK89b, ASS90, GM94, Mar95]. Une des applications les plus spectaculaires des HRTF est la simulation d'audition en champ libre sur écouteurs.

Le cas simple, voire simpliste, de l'étude de la sensation d'intensité pour les sons purs conduit à un grand nombre d'études psycho-physiologico-physiques. La sensation d'intensité est un attribut perceptif non seulement lié à la pression acoustique du son mais aussi à l'environnement de l'auditeur.

### Sensation de hauteur

La sensation de hauteur d'un son pur est en grande partie liée à sa fréquence. Cette sensation suit approximativement la règle de Fechner dans la zone la plus sensible de l'oreille (entre 1000 et 5000Hz). Cette relation simple entre phénomènes physiques et perception n'est cependant pas tout à fait valide en deçà de 1000Hz et au delà de 5000Hz. Cette relation non-linéaire définit l'échelle des *mels*:

$$f_{mels} = 2595 \log_{10} \left( 1 + \frac{f_{Hz}}{700} \right) \quad (1.24)$$

Cette échelle n'est certes valide que pour les sons purs mais elle permet d'appréhender certains phénomènes tels que la sensation proche de la dissonance perçue pour certains sons très riches en harmoniques (clavecin, instruments à cordes frottées). Un autre phénomène ayant trait à la sensation de hauteur de sons purs est une variation de cette sensation avec l'intensité du son.

### Perception du timbre

Le cas simple, voire simpliste, de la perception de sons purs montre de nombreux phénomènes complexes: les liens entre les paramètres physiques d'amplitude et de fréquence et les attributs perceptifs d'intensité et de hauteur ne sont ni directs ni linéaires.

L'étude des sons complexes introduit "un" troisième attribut perceptif: le *timbre*. Cet attribut est officiellement défini comme "l'attribut qui distingue deux sons de même intensité, durée, et hauteur"[Col60].

Grâce à une expérience psychoacoustique restée célèbre, John M. Grey a montré que cet attribut est en fait multidimensionnel [Gre75]. Nous allons décrire cette expérience et ses résultats dans les paragraphes suivants.

L'expérience psychoacoustique menée en 1975 par John M. Grey [GM77, Gre75] est une des premières à avoir mis en évidence l'aspect multidimensionnel de la perception du timbre musical. L'essentiel des conclusions de cette expérience a largement été confirmé par plusieurs expériences plus récentes. La méthodologie mise en jeu est toujours d'actualité, elle a été mise en œuvre pour de multiples études psychoacoustiques telles que l'étude de sons automobiles.

L'objectif de l'expérience de John M. Grey était d'établir une relation entre les attributs "timbraux" les plus significatifs et l'évolution du spectre d'un son musical. Pour

cette expérience, John M. Grey a utilisé 16 sons "perceptivement égalisés" en hauteur, intensité et durée. Ces 16 sons sont les suivants :

- BN : un basson ;
- C1 : une clarinette en  $Mi\flat$  ;
- C2 : une clarinette basse ;
- EH : un cor Anglais ;
- FH : un cor d'harmonie ;
- FL : une flûte ;
- O1 : un hautbois ;
- O2 : un second hautbois ;
- S1 : un violoncelle joué *sul ponticello* (sur le chevalet) ;
- S2 : un violoncelle joué normalement ;
- S3 : un violoncelle joué *sul tasto* (sur la touche) avec sourdine ;
- TM : un trombone avec sourdine ;
- TP : une trompette ;
- X1 : un saxophone *alto* joué *mf* ;
- X2 : un saxophone *alto* joué *p* ;
- X3 : un saxophone *soprano*.

Une vingtaine d'auditeurs — ayant des connaissances musicales — se sont prêtés à cette expérience, quinze d'entre eux ont procédé deux fois aux auditions. Chaque audition, d'une durée d'une heure, consistait en l'écoute de paires de sons choisis parmi les seize précités. Pour chaque paire, chaque auditeur devait évaluer la similarité, ou "distance perceptive", sur une échelle variant de 1 (très différents) à 30 (sons identiques).

Après audition, les distances perceptives collectées étaient regroupées dans un tableau de distances comportant 16 lignes et 16 colonnes. Ce type de tableau de distances n'est évidemment pas exploitable en l'état. *L'analyse factorielle des distances* permet d'obtenir une représentation "géométrique" des données du tableau : elle dispose un nuage de points — chaque point représentant un son — dans un espace à trois dimensions en respectant au mieux les distances. Les paragraphes suivants explicitent ce mode opératoire.

La figure 1.25 illustre comment il est possible de reconstituer la disposition des points à partir d'un tel tableau. Sachant que le point B est à une distance de 2 unités du point A, on choisit le point B sur le cercle de rayon 2 centré en A. Connaissant les distances entre le point C et les deux points précédents, on trace deux cercles centrés en A et B. Le

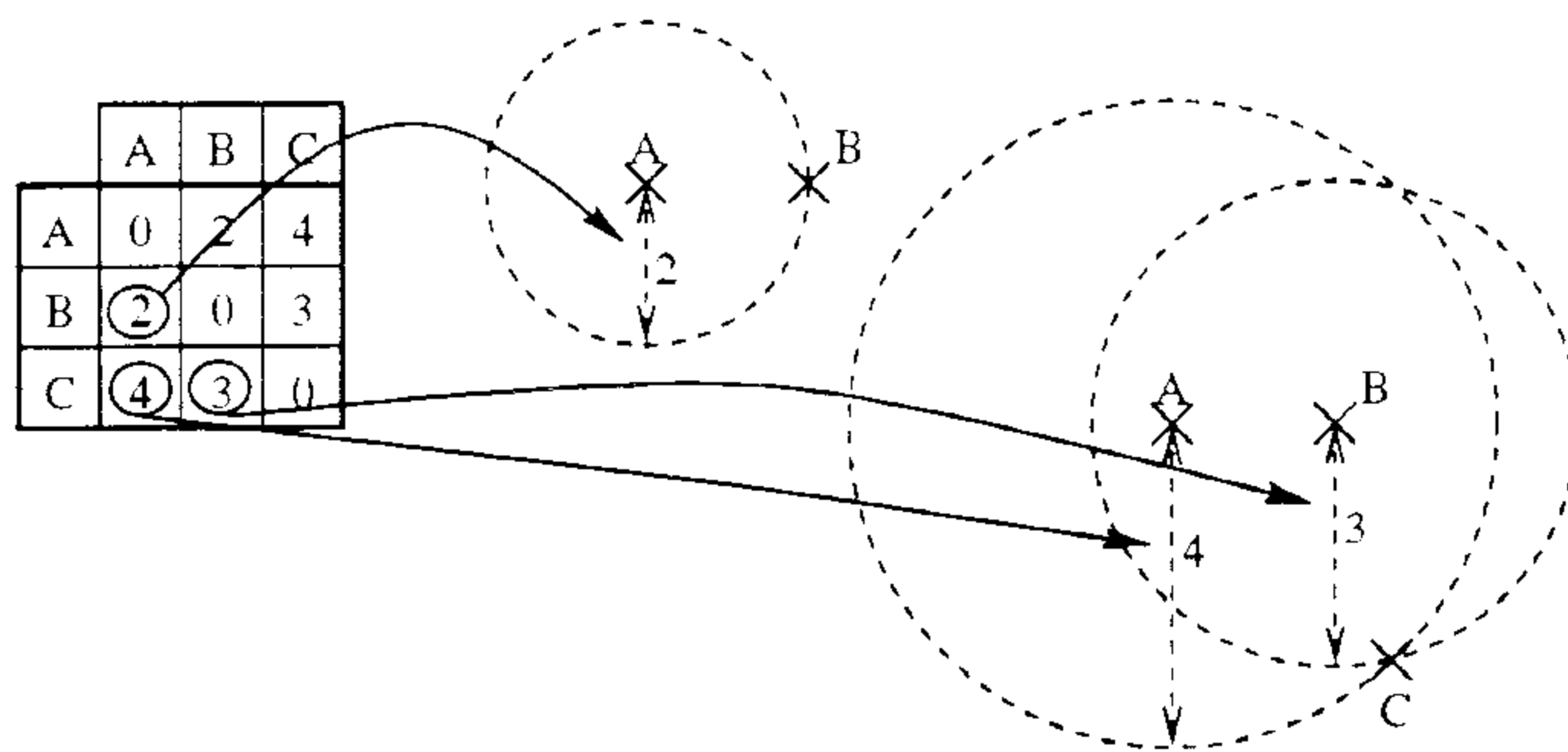


FIG. 1.25 – Reconstruction géométrique à partir d'un tableau de distances

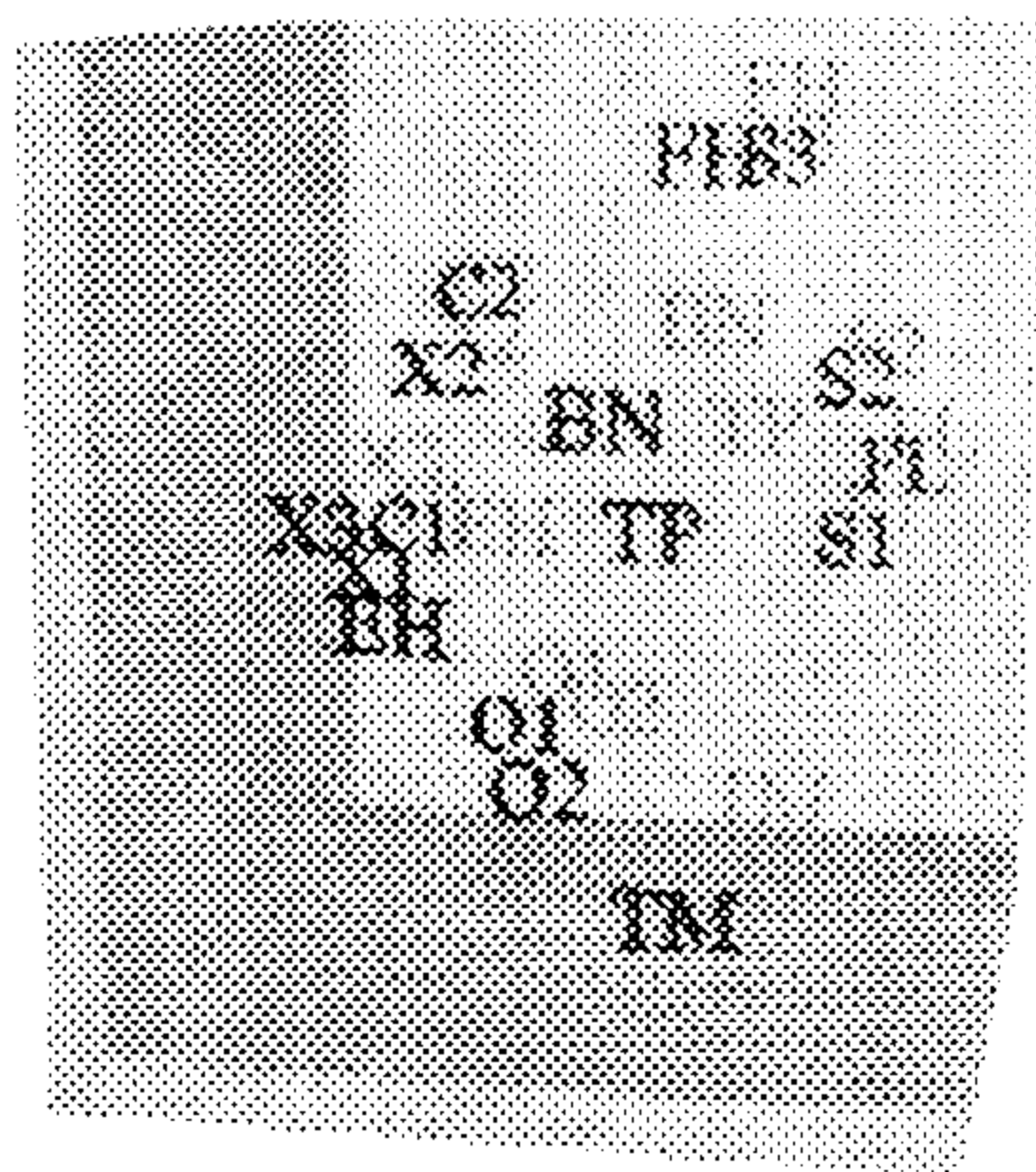


FIG. 1.26 – Espace de Grey

point C est choisi parmi les intersections de ces deux cercles. Notons qu'un choix différent pour les points A et B aurait conduit à une figure symétrique.

Plus généralement, à partir d'un tableau de distances de  $N$  points, il est possible de reconstruire leur disposition dans un espace à  $N - 1$  dimensions. Dans le cas des 16 sons de John M. Grey, on doit logiquement obtenir un espace à 15 dimensions.

Une réduction du nombre de dimensions est effectuée par analyse factorielle des distances. Cette analyse permet de trouver un sous-espace dans lequel les distances sont respectées au mieux. Les axes de ce sous-espace sont implicitement définis par le tableau de distances perceptives. Chaque axe a donc une signification perceptive intrinsèque.

La figure 1.26 représente l'espace perceptif de Grey en perspective. Cette représentation, aussi élégante soit elle, n'est pas facile à interpréter. Pour interpréter objectivement les axes et leur associer quelque attribut perceptif usuel (brillance, couleur, dynamique), il convient, en premier lieu, de les associer à des attributs physiques. Les figures 1.27 et 1.28 représentent les spectres dynamiques des sons disposés selon les plans XY et ZY de l'espace de Grey.

L'axe X, représenté horizontalement sur la figure 1.27, semble opposer les sons ayant un transitoire d'attaque rapide aux sons ayant un transitoire d'attaque plus "mou". Ainsi, ce premier axe perceptif semble être associé à la dynamique du son.

L'axe Y, représenté verticalement sur les figures 1.27 et 1.28, semble opposer les sons riches en harmoniques aux sons pauvres en harmoniques. Il semble donc que ce second

axe perceptif correspond à la richesse spectrale du son.

L'axe Z, représenté horizontalement sur la figure 1.28, est plus difficile à interpréter. John M. Grey lui a attribué des phénomènes de "synchronicité des harmoniques" lors des transitoires [Gre75, page 67].

L'étude de John M. Grey a permis d'associer trois attributs physiques à l'attribut perceptif de timbre :

- la dynamique des transitoires ;
- la richesse spectrale ;
- la synchronicité des harmoniques.

La méthode de réduction de données — analyse en composantes principales — que nous avons utilisé dans notre étude appartient à la même catégorie de méthodes que l'analyse factorielle des distances utilisée par John M. Grey. A ce sujet, la section 2.3.1, page 85, se propose d'établir une relation entre notre étude et celle de John M. Grey.

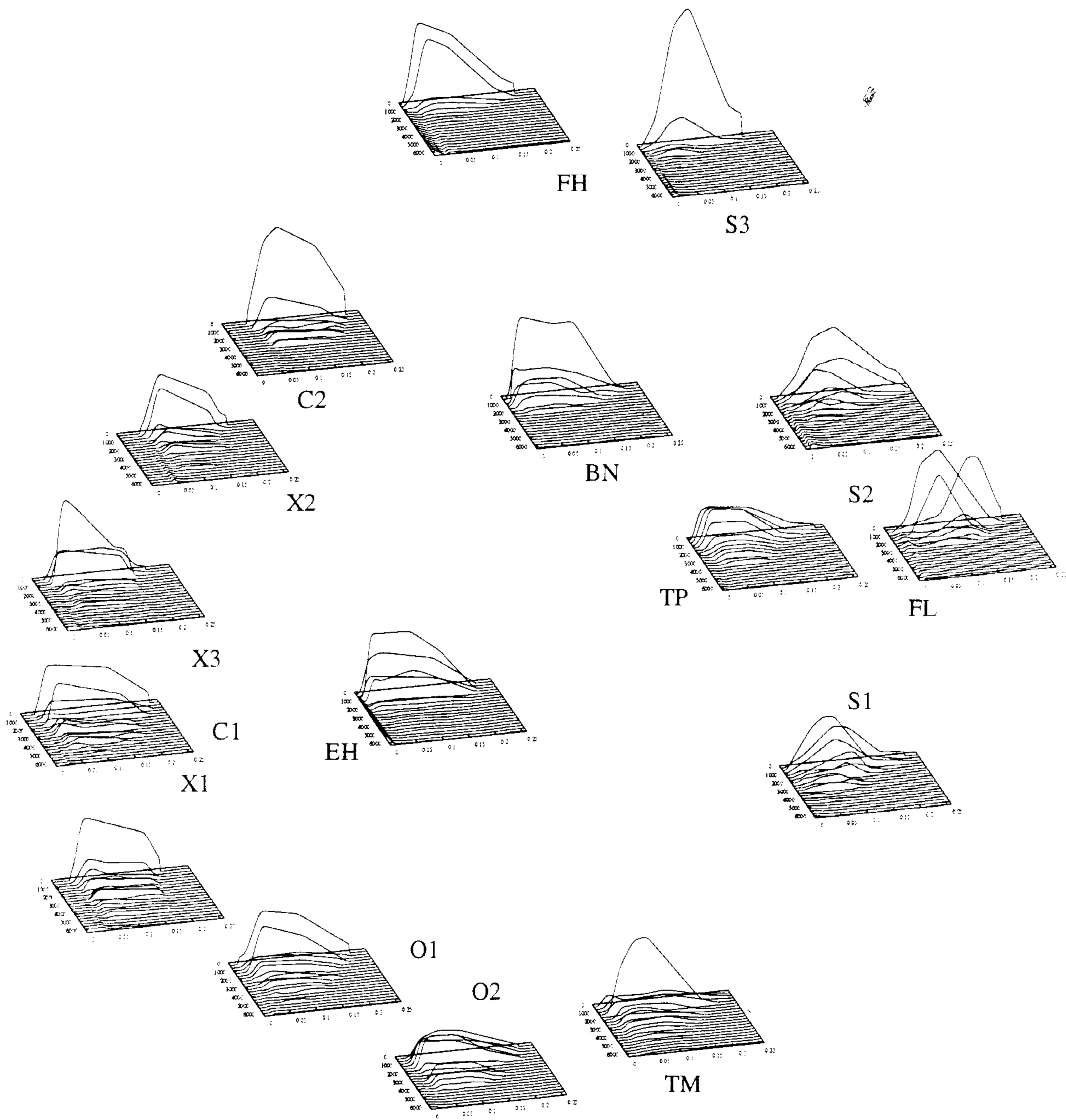


FIG. 1.27 - Plan XY de l'espace de Grey

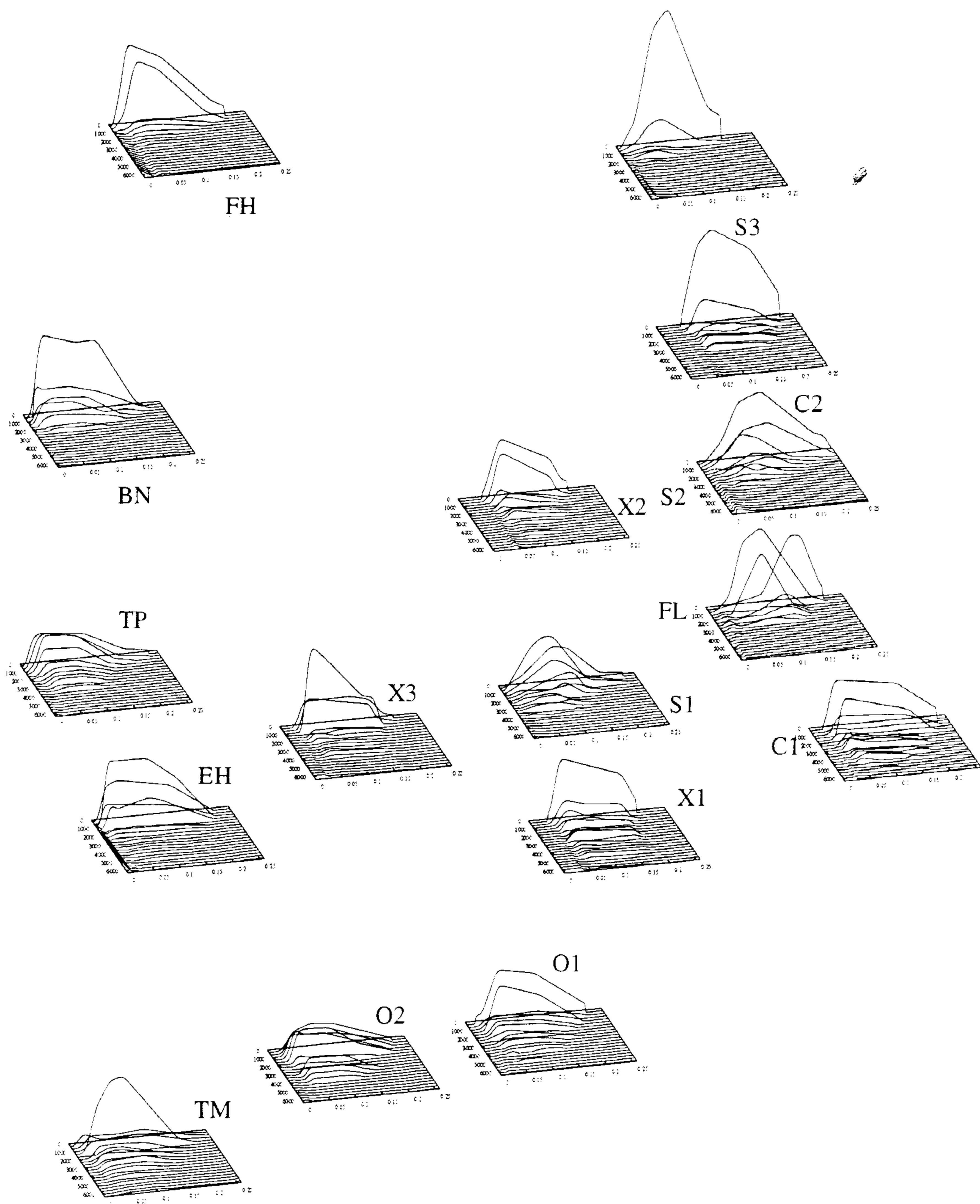
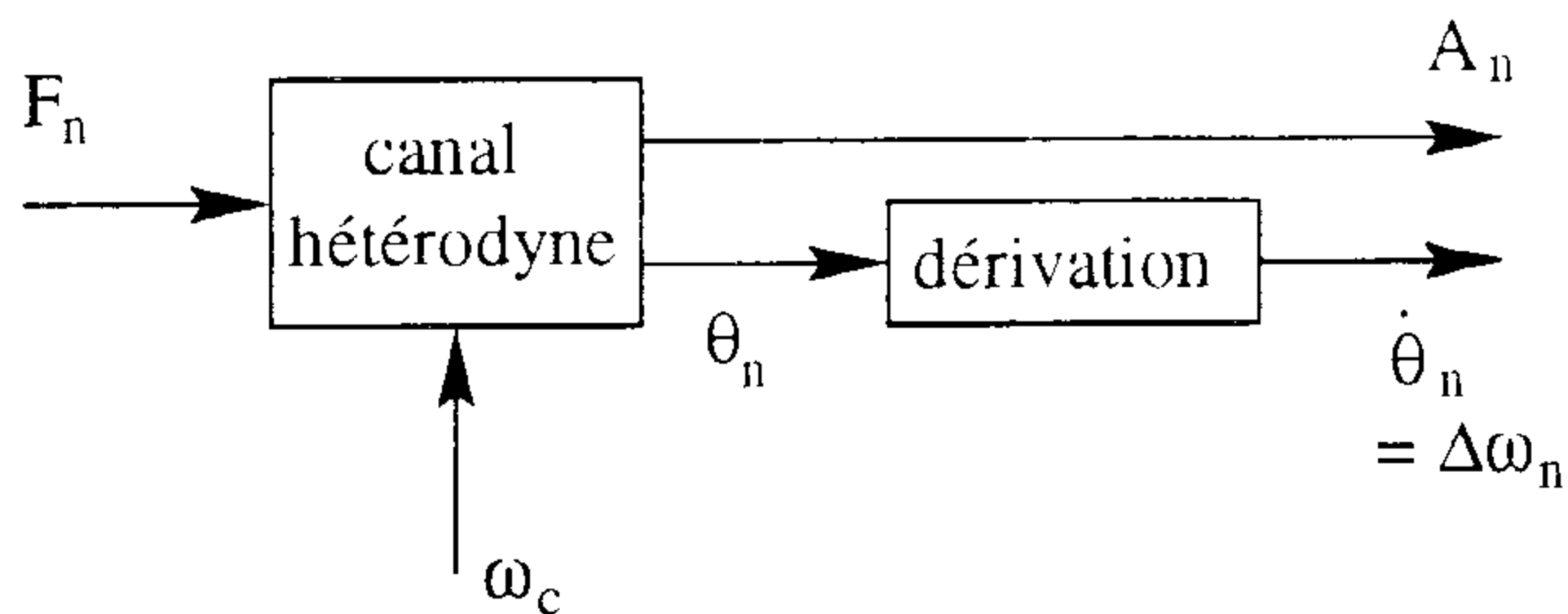
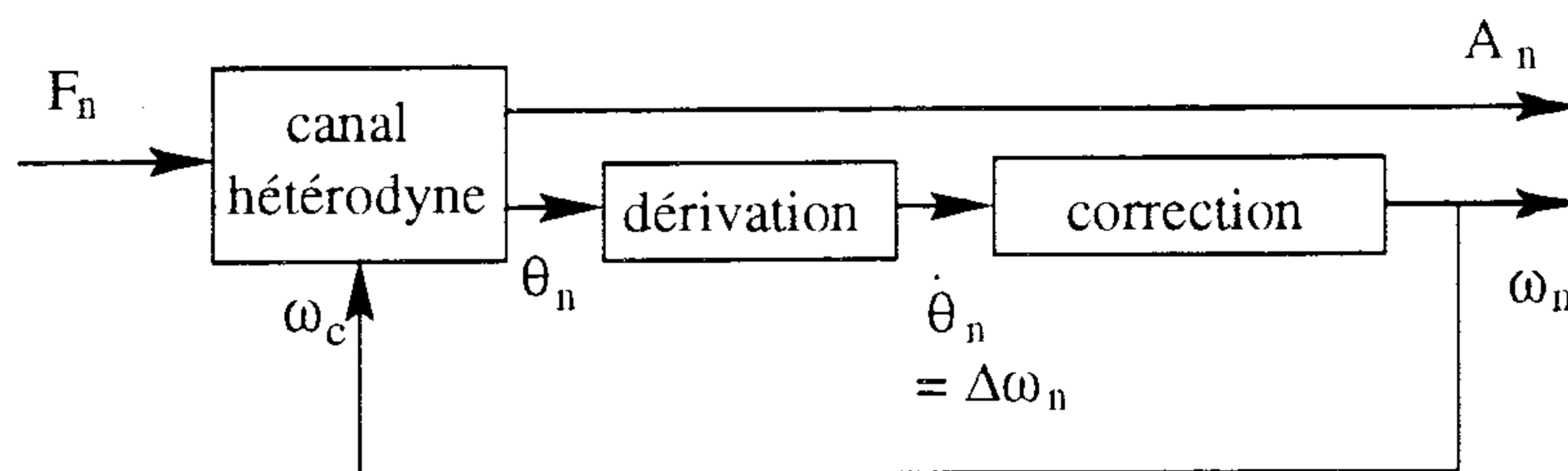


FIG. 1.28 - Plan ZY de l'espace de Grey

FIG. 1.29 – *filtre hétérodyne simple*FIG. 1.30 – *filtre hétérodyne adaptatif*

### 1.3.3 Conclusions

Le système auditif humain, destinataire final des sons synthétisés, se révèle être fort complexe. La physiologie de l'oreille interne présente des phénomènes actifs, commandés par les centres nerveux supérieurs, non négligeables.

Qui plus est, des expériences psychoacoustiques simples, ne mettant en jeu que des sons purs, conduisent à des relations complexes entre grandeurs physiques et sensations.

L'attribut perceptif de timbre, qui est au cœur de nos travaux, est un attribut multidimensionnel dont seuls les deux premiers facteurs (composition spectrale et dynamique globale) sont clairement identifiés.

La perception auditive humaine demeure un domaine extrêmement complexe et difficile. C'est pourquoi, les procédés de réduction de données décrits aux chapitres 2 et 3 ne sont pas fondés sur des critères psychoacoustiques mais sur l'extraction de redondances présentes dans le signal.

## 1.4 Inventaire des méthodes d'analyse

Le développement du traitement du signal et de l'informatique permet d'utiliser une gamme croissante de méthodes d'analyses spectrales. Nous en décrivons succinctement quelques unes en explicitant leurs applications dans le cadre de l'étude des sons musicaux.

### 1.4.1 Filtrage hétérodyne

Le filtrage hétérodyne est une méthode d'analyse spectrale relativement ancienne. Il mérite toutefois notre attention en raison du nombre important d'études dans lesquelles il est entré en jeu.

Le filtrage hétérodyne permet de suivre l'évolution de composantes pseudo-harmoniques dans le temps. Chaque composante variant en fréquence et en amplitude.

Le signal — ici sous forme discrétisée — est donc considéré comme constitué d'une somme de partiels (composantes sinusoïdales) :

$$F_n = \sum_{p=1}^P A_{pn} \sin(\omega_{pn} nT + \theta_{pn}) \quad (1.25)$$

$F_n$  échantillon numéro  $n$

$T$  temps séparant deux échantillons

$P$  nombre de partiels

$A_{pn}$  amplitude du partiel  $p$

$\omega_{pn}$  pulsation du partiel  $p$

$\theta_{pn}$  phase du partiel  $p$

L'objectif de l'algorithme est de déterminer l'amplitude  $A_p$  et la phase  $\theta_p$  pour chaque partiel. Pour ce faire, on considère que le signal est quasi-harmonique et que la pulsation fondamentale est connue et ne varie pas beaucoup : la pulsation du partiel  $p$  est approximativement égale à  $\omega_c = p\omega$  où  $\omega$  est la pulsation moyenne du fondamental.

La solution du problème se résume dans ce cas aux équations :

$$a_{pn} = \sum_{i=n}^{n+N-1} F_i \sin(p\omega_i T) \quad (1.26)$$

$$b_{pn} = \sum_{i=n}^{n+N-1} F_i \cos(p\omega_i T) \quad (1.27)$$

$$A_{pn} = \sqrt{a_{pn}^2 + b_{pn}^2} \quad (1.28)$$

$$\theta_{pn} = \arctan\left(\frac{a_{pn}}{b_{pn}}\right) \quad (1.29)$$

$N$  représentant le nombre de points de la trame temporelle analysée. La variation temporelle des phases  $\theta_{pn}$  est ensuite interprétée comme "erreur" sur la pulsation  $\Delta\omega_{pn}$ . La pulsation "réelle" du partiel  $p$  s'écrit alors :

$$\omega_{pn} = p\omega + \Delta\omega_{pn} \quad (1.30)$$

Un filtre hétérodyne n'est jamais parfaitement centré sur le partiel. Cette "erreur" de centrage engendre une variation de phase d'une trame à la suivante. Cette variation de phase est prise en compte pour corriger "l'erreur" (figure 1.29). Un filtrage hétérodyne simple consiste en un banc de filtres hétérodynes à pulsations centrales  $\omega_c$  fixes.



On peut améliorer le procédé en le rendant adaptatif. Un filtrage hétérodyne adaptatif consiste en un banc de filtres hétérodynes à pulsations centrales  $\omega_c$  variables (figure 1.30). Les pulsations centrales sont "corrigées" dynamiquement en fonction des variations de phases.

Dans les deux cas, la phase étant définie modulo  $2\pi$ , une variation supérieure à  $\pi$  d'une trame à la suivante fausse l'analyse. Un taux de recouvrement plus important des trames devient nécessaire.

Cette méthode permet donc l'analyse de signaux harmoniques ou pseudo-harmoniques n'évoluant pas trop rapidement en fréquence et en amplitude. Pour la plupart des sons musicaux, la partie harmonique est prédominante. La simplicité du filtrage hétérodyne en a fait un outil "standard" d'analyse des sons musicaux [Gre75, GM77].

### 1.4.2 Poursuite de pics de transformée de Fourier

Cette méthode d'analyse spectrale consiste à découper le signal en sections temporelles dénommées *trames*. La présence et la localisation spectrale de composantes sinusoïdales sont évaluées pour chaque trame au moyen d'une transformée de Fourier discrète à court terme (figure 1.31). Pour affiner l'évaluation de la localisation des composantes, il est souvent pratiqué une interpolation quadratique [FH96]. Bien évidemment, la qualité de cette interpolation détermine en grande partie la qualité de l'analyse. Nous proposons en 2.1.1 une méthode d'interpolation précise basée sur les propriétés de la fenêtre gaussienne.

Après détection, chaque composante sinusoïdale d'une trame est soit [MQ86]:

- mise en relation avec un composante de la trame précédente, les critères de mise en correspondances portant généralement sur les différences de fréquence et d'amplitude;
- considérée comme "naissante" si son amplitude est faible et qu'elle ne possède pas d'homologue dans la trame précédente;
- considérée comme "mourante" si son amplitude est faible et qu'elle ne possède pas d'homologue dans la trame suivante.

Les résultats de l'analyse sont directement exploitables par synthèse additive (§1.2.1, page 29). La qualité de la resynthèse permet de contrôler la qualité de l'analyse effectuée et de procéder, le cas échéant, à une adaptation plus fine des paramètres d'analyse (largeur et taux de recouvrement des trames, types de fenêtre et d'interpolation).

Cette méthode, faisant un minimum d'hypothèses sur les variations de fréquence du signal analysé, est capable d'analyser des sons fortement inharmoniques. Cependant, la représentation de certains sons bruités reste difficilement exploitable: ils sont représentés par une multitude de sinusoïdes modulées en amplitude et en fréquence.

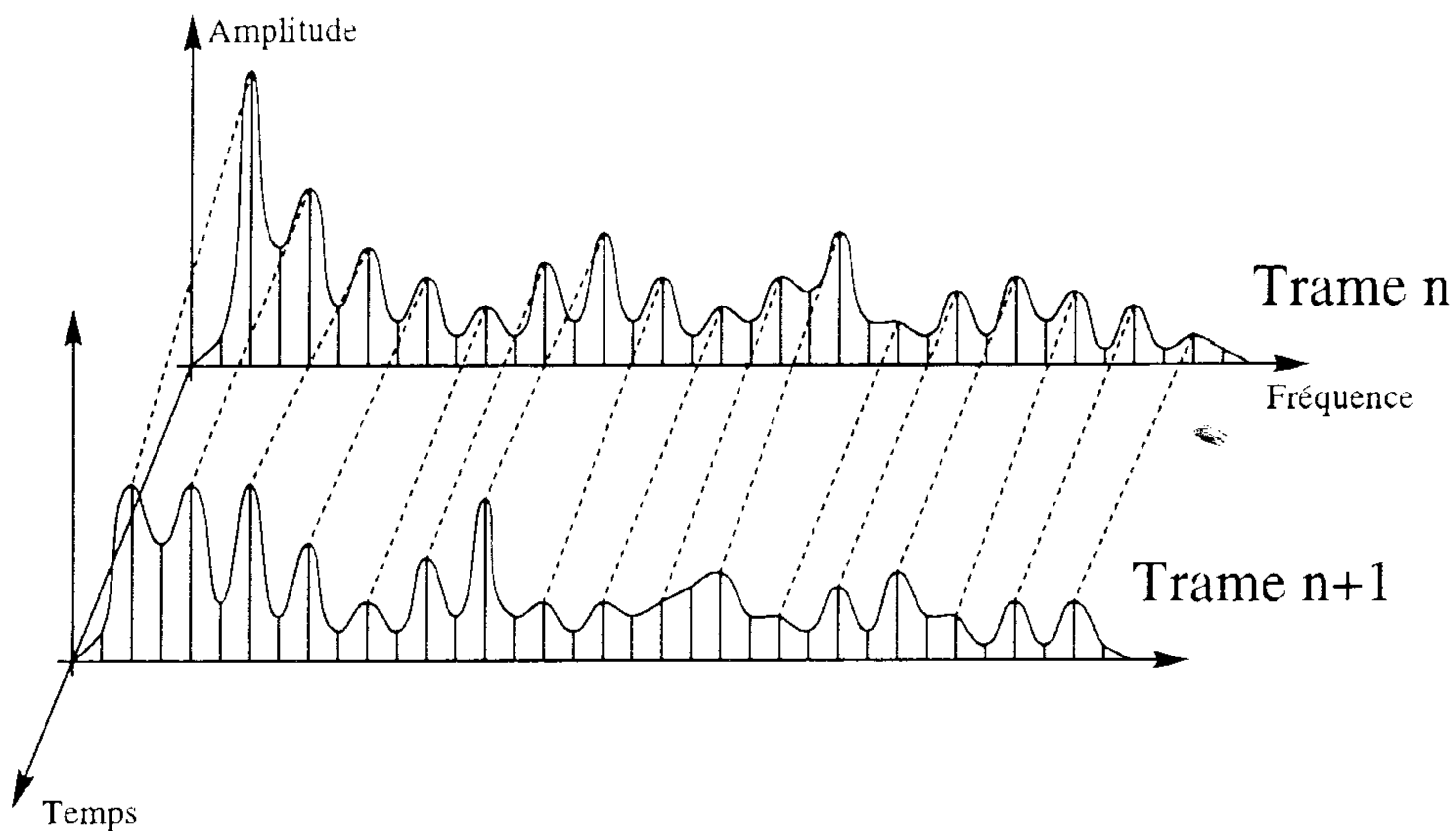


FIG. 1.31 – Poursuite de pics de transformées de Fourier

### 1.4.3 Modèle déterministe plus stochastique

La représentation de sons bruités par des composantes sinusoïdales est inadaptée et conduit à une synthèse additive trop coûteuse en temps de calcul. Une solution consiste à séparer partiels et composantes bruitées.

Comme dans le cas précédent, on commence par extraire les pics de la transformée de Fourier. On ne conserve que les pics ayant une amplitude importante. Ces pics seront par la suite considérés comme décrivant la partie “déterministe” du signal. L’étape suivante consiste à soustraire ces pics à la transformée de Fourier. Le spectre résiduel est alors considéré comme spectre de bruit. Il sera par la suite lissé. La synthèse de la partie bruitée sera effectuée par transformée de Fourier inverse de ce spectre lissé [SS90, dYD95].

### 1.4.4 Méthode de Prony

Ce type d’analyse considère un signal comme étant la somme de sinusoïdes à amortissement exponentiel de la forme :

$$F_n = \sum_{p=1}^P A_p a_p^n \sin \left( \frac{\phi_p + n\omega_p}{F_e} \right) \quad (1.31)$$

Ce type d’analyse est de toute évidence fort bien adapté aux sons percussifs tels que les sons de cordes frappées et grattées. De multiples algorithmes permettant d’extraire les paramètres  $A_p$ ,  $a_p$ ,  $\phi_p$ , et  $\omega_p$  sont possibles mais leur coût de calcul est longtemps resté prohibitif. Cette forme d’analyse est donc adaptée aux sons relativement brefs et percussifs ne dépassant pas les 1000 échantillons [Lar93].

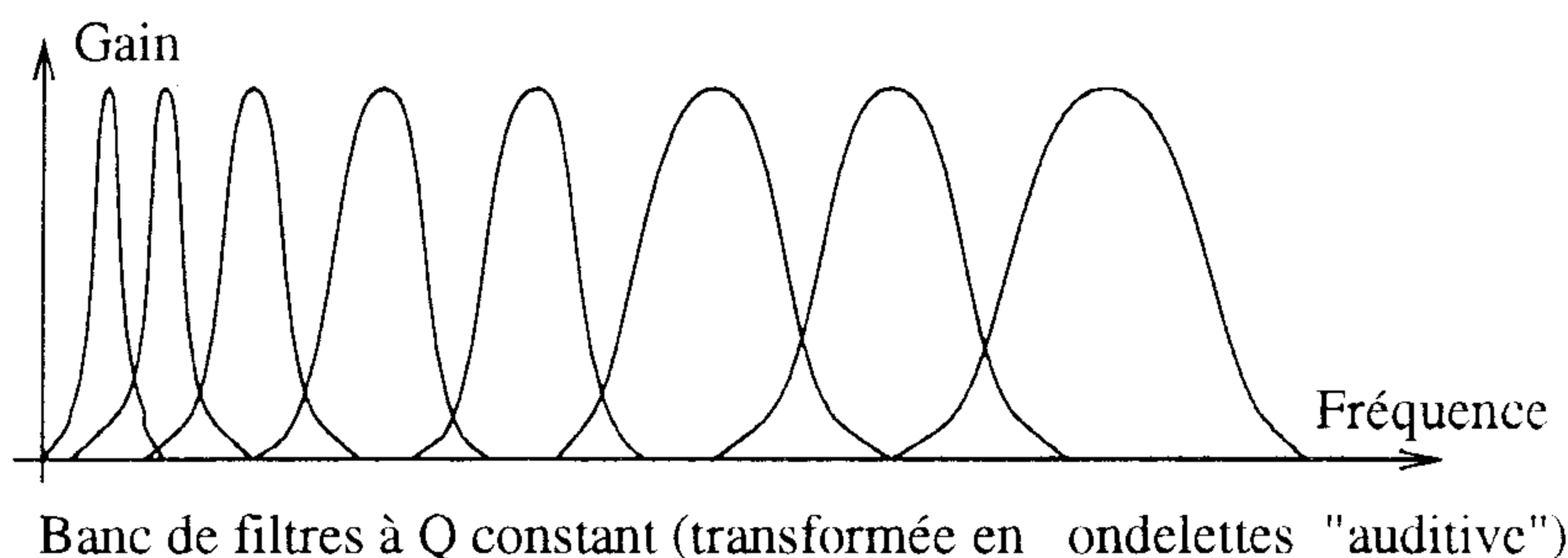
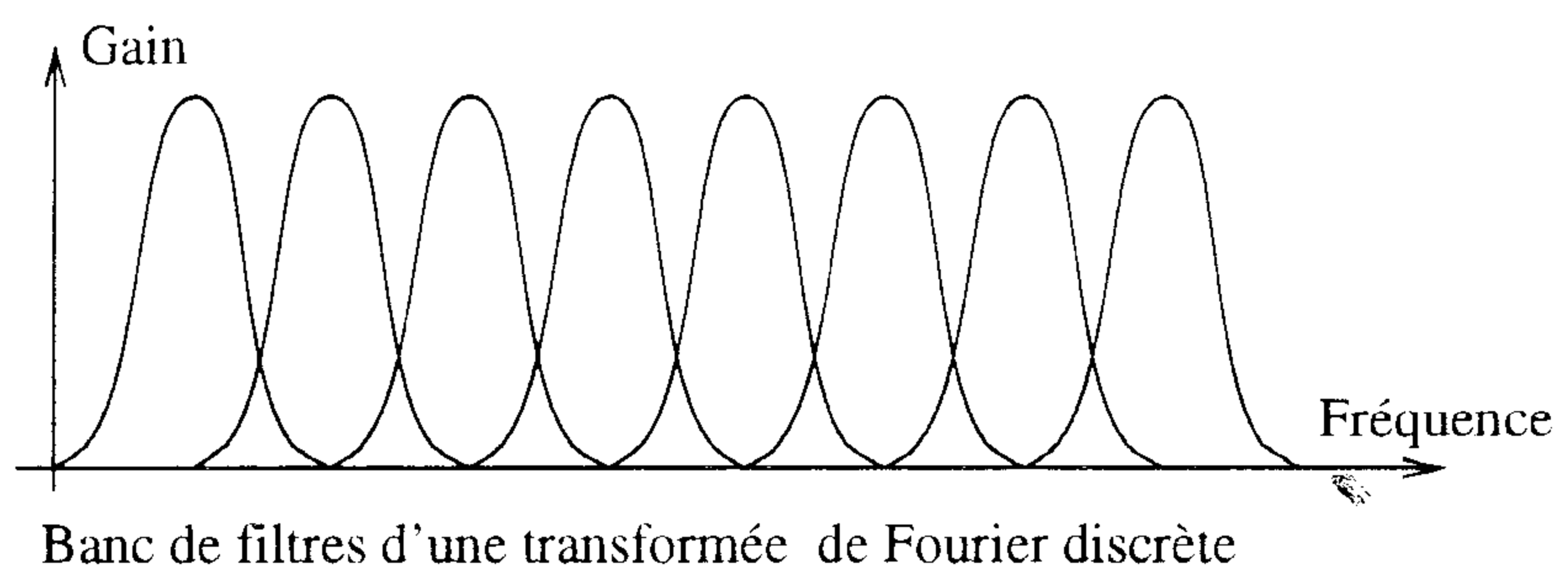


FIG. 1.32 – Transformée de Fourier et banc de filtres à Q constant

#### 1.4.5 Bancs de filtres à facteur de qualité constant

La transformée de Fourier suppose une résolution temps / fréquence linéaire : elle peut être interprétée comme un banc de filtres linéairement disposés sur l'axe des fréquences.

Cette représentation en échelle de fréquences linéaire n'est pas représentative de la résolution spectrale équivalente à l'audition humaine. L'analyse par banc de filtres à facteur de qualité  $Q$  constant utilise un banc de filtres tel que :

$$\frac{\Delta f}{f} = Q = \text{Constante} \quad (1.32)$$

$f$  représentant la fréquence du filtre. Une telle représentation, suivie d'un processus d'interpolation similaire à celui vu en 1.4.2, a été utilisée par Daniel Ellis pour modéliser la perception de sons multiples [Ell92]. Ce type d'analyse peut être étendu à un modèle plus fin de l'audition en :

- utilisant un banc de filtres dont les fonctions de transfert approchent celle de la cochlée, on parle alors de cochléogramme [YWS92, IK93a, IK93b] ;
- modélisant des traitements d'ordre supérieur tels que corrélation [Sla94, SLN95].

Certains de ces modèles permettent la resynthèse des sons. Il devient alors possible d'affiner leurs paramètres par des expériences psychoacoustiques. Ces expériences peuvent permettre de tenir compte des processus d'écoute attentive évoquées en 1.3. Ils deviennent alors de véritables modèles de l'audition humaine. Les enjeux de tels modèles sont considérables pour les télécommunications car ils permettraient de compresser le signal de parole en supprimant les informations non perçues [IK93b].

Dans le cadre des applications plus “musicales”, cette famille de méthodes permet de resynthétiser les transitoires de façon fidèle. Cette forme d'analyse reste cependant très coûteuse en calcul malgré quelques travaux intéressants menant à des algorithmes efficaces [TMM80].

### 1.4.6 Transformée en ondelettes

Les transformées en ondelettes généralisent — en quelques sortes — les notions introduites par l'analyse par banc de filtres à  $Q$  constant. Une telle transformée est définie par :

$$W_{\Psi}s(\alpha, \tau) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{+\infty} s(t) \Psi \left( \frac{\tau - t}{\alpha} \right) dt \quad (1.33)$$

$\Psi$  définit la “forme” de l'ondelette. Celle-ci doit respecter un certains nombres de critères tels que  $\int |\Psi(t)| dt < \infty$ ,  $\int |\Psi(t)|^2 dt < \infty$  et  $\int \frac{|\hat{\Psi}(\omega)|^2}{\omega} dt < \infty$ ,  $\hat{\Psi}$  représentant la transformée de Fourier de  $\Psi$ .

$\alpha$  définit le facteur d'échelle, il correspond à la localisation “spectrale” de l'ondelette.  $\tau$  définit le facteur de translation temporelle, il correspond donc à la localisation temporelle de l'ondelette.

L'expression de  $\Psi(t)$  détermine la nature de l'analyse pratiquée. Elle est en quelque sorte le pendant de la synthèse granulaire (§1.2.7, page 40).  $\Psi(t)$  peut par exemple prendre la forme d'un atome temps-fréquence de Gabor :

$$\Psi_G(t) = e^{-\frac{t^2}{2}} e^{i\omega t} \quad (1.34)$$

Sous cette forme, et sous de nombreuses formes simples, la transformée en ondelettes peut être interprétée comme une analyse par banc de filtres à  $Q$  constant. Dans d'autres cas,  $\Psi(t)$  peut prendre une forme plus élaborée, destinée, par exemple, à détecter une certaine classe de signaux. L'ondelette de détection d'octaves proposée par Richard Kronland-Martinet en est un bon exemple :

$$\Psi_O(t) = e^{-\frac{t^2}{2}} [e^{i\omega_0 t} + e^{2i\omega_0 t}] \quad (1.35)$$

Cette ondelette permet en effet de détecter efficacement des signaux musicaux en relation d'octave [KM88].

Les transformées en ondelettes à temps continu — ou considérant le signal à temps continu sous-jacent à un signal discrétisé — disposent que de très peu d'algorithmes rapides limités à certains types d'ondelettes [VLU95]. Les transformées en ondelettes à temps discret disposent parfois d'algorithmes très efficaces mais leur résolution fréquentielle est généralement de l'ordre de l'octave. Les transformées en ondelettes sont — dans leur forme la plus générale — encore réservées à des études en temps différé.

### 1.4.7 Résumé et conclusions

Cette section a introduit de multiples méthodes d'analyse permettant généralement d'obtenir une décomposition du signal sous formes de sommes de signaux sinusoidaux

élémentaires. Cette décomposition permet en général de resynthétiser le son et d'évaluer la qualité de l'analyse.

La méthode du filtrage hétérodyne est ancienne et ses limitations doivent être connues lorsqu'on se réfère à certains travaux. Elle est limitée aux sons dont les composantes n'évoluent pas trop rapidement en fréquence et en amplitude.

Les méthodes basées sur l'interpolation de transformées de Fourier à court terme et à la poursuite de composantes trame à trame sont aujourd'hui très courantes. Elles ont, en effet, les avantages d'être rapides et efficaces pour l'analyse des sons musicaux.

Les méthodes précédentes peuvent être avantageusement complétées par des méthodes d'analyse et de synthèse de la composante "bruitée".

Les méthodes de Prony et leurs dérivées considèrent le signal comme somme de sinusoides à amortissement exponentiel. Bien que fort coûteuses en calculs, ces méthodes sont très efficaces pour l'analyse de sons percussifs brefs.

Les méthodes basées sur des bancs de filtres à facteur de qualité constant rendent compte de façon plus correcte de la perception auditive. Hélas, leur coût de calcul reste encore prohibitif.

Les transformées en ondelettes permettent de caractériser plusieurs types de signaux selon la forme de l'ondelette mère. L'absence d'algorithmes généraux rapide en limite l'utilisation. Cependant, les avancées dans le domaine de la puissance des ordinateurs leur donne un intérêt grandissant.

Cette courte présentation nous donne un aperçu des formes d'analyse spectrale utilisables dans notre contexte et nous permettra de motiver nos choix.

## Chapitre 2

# Modélisation par ACP de spectres harmoniques réels

*In conclusion, though statistically optimal,  
PCA should by no means be regarded as  
the best matching method.*

A. Horner, J. Beauchamp et L. Haken

*Methods for Multiple Wavetable  
Synthesis of Musical Instruments [HBH93]*

Ce chapitre présente une méthode d'analyse des sons musicaux basée sur une analyse spectrale et une analyse en composantes principales (ACP). Cette méthode conduit à une représentation des sons musicaux sous forme de *données réduites* significatives du son analysé.

Les sections 2.1, page 63, et 2.2, page 72, décrivent les processus d'analyse spectrale et de réduction de données employés. La section 2.3.1 établit un lien entre l'espace perceptif de John M. Grey (§1.3.2, page 50) et les données réduites obtenues. La section 2.3, page 85, montre deux études du timbre musical menées grâce à notre méthode. La section 2.4, page 104, montre les applications de la méthode pour la synthèse des sons musicaux.

### 2.1 Analyse spectrale de la composante harmonique

L'objet principal de ces travaux est une optimisation, en terme de coût de calcul et "d'ergonomie", de la représentation des signaux musicaux. Pour obtenir une telle représentation, nous utiliserons l'analyse en composantes principales, d'abord sur un

spectre harmonique d'amplitudes réelles, puis, sur un spectre harmonique d'amplitudes complexes.

La qualité des résultats d'une telle analyse dépend en grande partie de la qualité des données qui lui sont fournies. C'est pourquoi, nous nous sommes attachés à développer une forme d'analyse spectrale précise et adaptée aux signaux pseudo-périodiques.

### 2.1.1 Interpolation de transformée de Fourier discrète à court terme

Les transformées en ondelettes (§1.4.6, page 61) sont à la base de nombreux modèles du système auditif humain [Ell92, IK93a, Sla94, SLN95]. Bien qu'ayant l'avantage d'être représentatives de la perception humaine, ces méthodes sont encore coûteuses en temps de calcul. Les algorithmes de transformées en ondelettes rapides restent cantonnés à une résolution spectrale égale à l'octave, ce qui est insuffisant dans notre contexte.

La transformée de Fourier discrète, quant à elle, a l'avantage de disposer d'algorithmes rapides et simples à programmer. Un autre atout de cette transformée tient à la nature même de la composante harmonique que nous étudions dans ce chapitre. Les harmoniques sont approximativement espacés par la fréquence fondamentale. Cet espacement est linéaire. L'échelle des fréquences de la transformée de Fourier est elle aussi linéaire. Il suffit donc d'utiliser une transformée de Fourier à court terme adaptée à cet espacement pour obtenir une représentation optimale de la composante harmonique en terme de résolution temps/fréquence.

L'objectif de notre analyse est d'extraire les caractéristiques des composantes sinusoïdales du signal continu sous-jacent au signal discrétisé. Chaque trame du signal discret analysé  $S(T)$  doit être exprimée sous la forme :

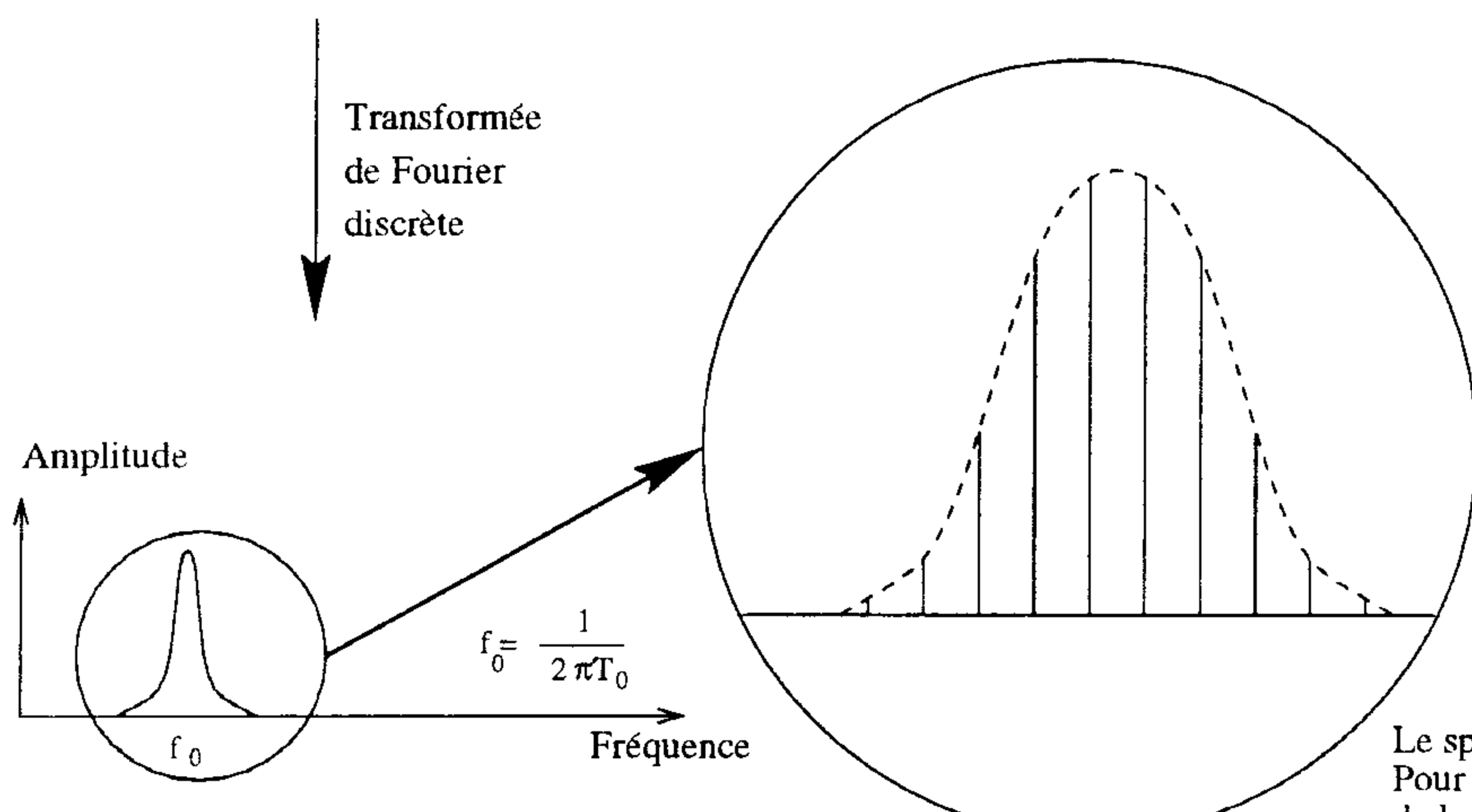
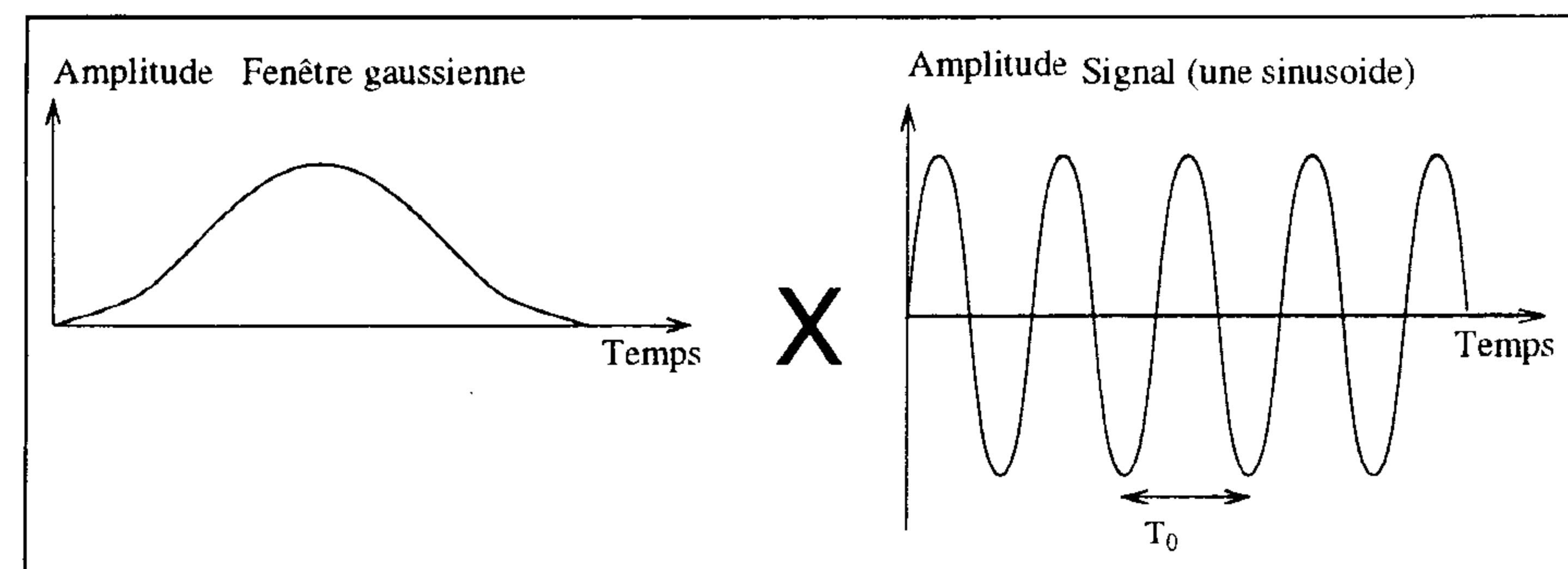
$$S(T, n) = \sum_{j=1}^{J(T)} C_j(T, n) \quad (2.1)$$

avec

$$C_j \left( T, n = \left[ -\frac{N-1}{2} .. \frac{N-1}{2} \right] \right) = a_j(T) \sin \left( \frac{2\pi n}{F_e} f_j(T) + \phi_j(T) \right) \quad (2.2)$$

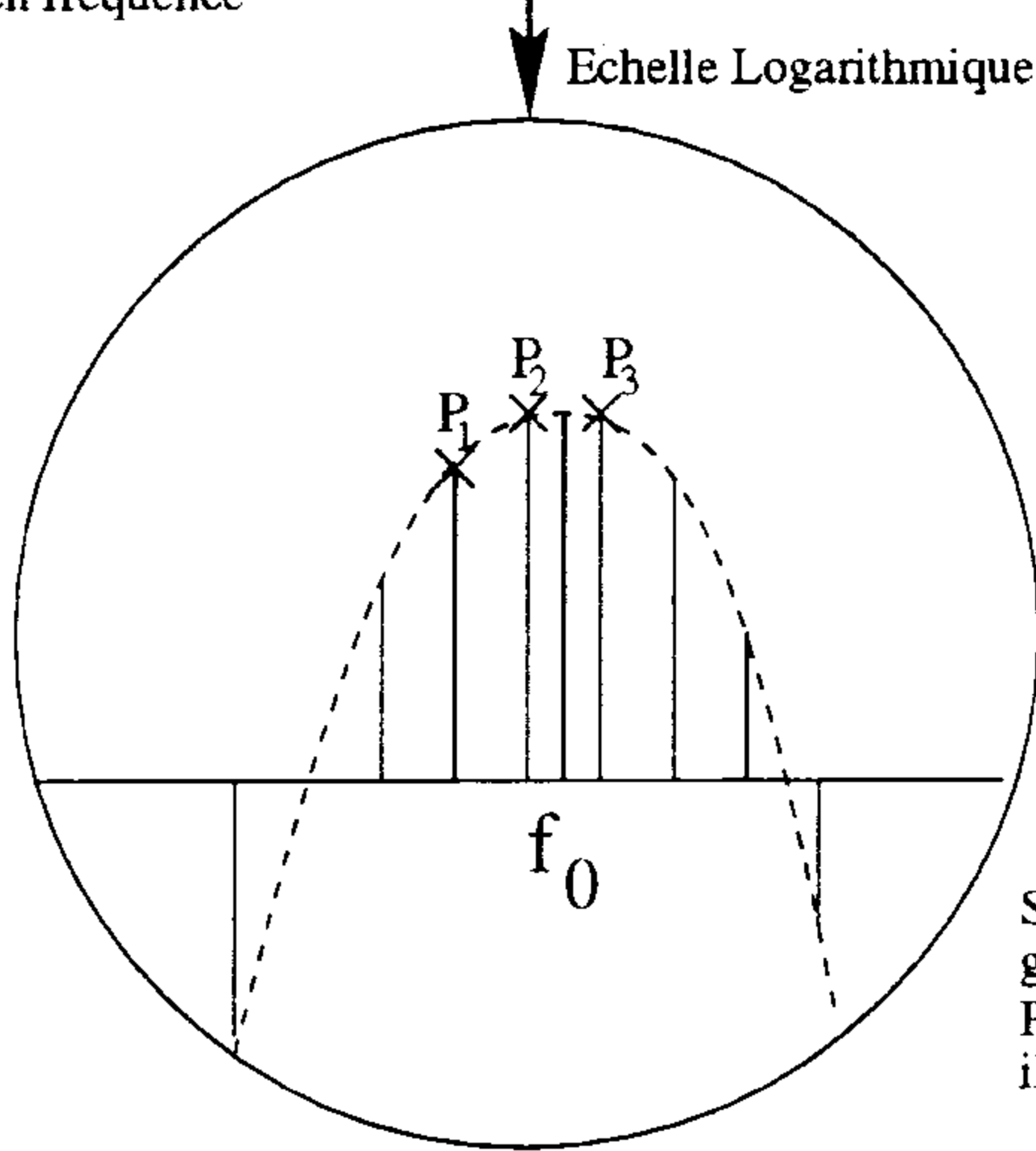
- $N$  est le nombre d'échantillons d'une trame,
- $F_e$  est la fréquence d'échantillonnage,
- $J(T)$  est le nombre de composantes sinusoïdales pour la trame  $T$ ,
- $f_j(T)$  est la fréquence de la composante  $C_j$  pour la trame  $T$ ,
- $a_j(T)$  est l'amplitude de la composante  $C_j$  pour la trame  $T$ ,
- $\phi_j(T)$  est la phase en milieu de trame  $T$  de la composante  $C_j$ .

Les paramètres  $f_j(T)$ ,  $a_j(T)$  et  $\phi_j(T)$  de chaque composante  $C_j(T)$  sont continus et variables d'une trame à la suivante. De plus, le nombre  $J(T)$  de composantes présentes dans le signal est lui aussi variable.



Multiplier le signal par une gaussienne dans le temps revient à le convoluer par une gaussienne en fréquence

Le spectre est discrétisé en fréquence. Pour obtenir l'amplitude et la fréquence de la composante sinusoïdale, il faut interpoler le spectre discret.



Sur une échelle logarithmique une gaussienne devient un arc de parabole. Pour interpoler parfaitement cet arc, il suffit d'en connaître 3 points.

FIG. 2.1 – Interpolation de la transformée de Fourier



L'algorithme utilisé pour cette étude est fondé sur l'interpolation d'une transformée de Fourier rapide à court terme (figure 2.1). En principe, pour chaque trame  $T$ , le signal est fenêtré par une fonction gaussienne. Une propriété remarquable de cette fonction est que sa transformée de Fourier est elle-même une fonction gaussienne. La transformée de Fourier  $\mathcal{F}(T)$  de la trame  $T$  étant discrète, chaque composante sinusoïdale  $C_j(T)$  sera représentée sous la forme d'une gaussienne "dicrotisée" et centrée en  $f_j(T)$ .

Chaque maximum local  $\mathcal{F}(T, m)$  tel que  $|\mathcal{F}(T, m-1)| < |\mathcal{F}(T, m)| > |\mathcal{F}(T, m+1)|$  sous-tend une composante sinusoïdale  $C_j(T)$ . Le nombre de maxima locaux définit  $J(T)$ .

Sur une échelle logarithmique des amplitudes, cette gaussienne devient un arc de parabole. Les trois points  $p_1 = \ln |\mathcal{F}(T, m-1)|$ ,  $p_2 = \ln |\mathcal{F}(T, m)|$ , et  $p_3 = \ln |\mathcal{F}(T, m+1)|$  suffisent pour l'interpoler exactement. La fréquence  $f_j(T)$  et l'amplitude  $a_j(T)$  de la composante  $C_j(T)$  sont données par l'interpolation quadratique :

$$f_j(T) = \frac{F_\varepsilon}{N}(m+x) \quad (2.3)$$

$$a_j(T) = e^{ax^2+bx+c} \quad (2.4)$$

avec

$$a = \frac{p_1 - 2p_2 + p_3}{2} \quad (2.5)$$

$$b = \frac{p_3 - p_1}{2} \quad (2.6)$$

$$c = p_2 \quad (2.7)$$

$$x = -\frac{b}{2a} \quad (2.8)$$

Nous justifions pleinement l'utilisation de l'interpolation quadratique par les propriétés du fenêtrage gaussien. L'interpolation quadratique a cependant parfois été utilisée pour d'autres études sans une telle justification [FH96].

La phase  $\phi_j(T)$ , quant à elle, est obtenue par simple interpolation linéaire.

### 2.1.2 Effets de la troncature de la fenêtre

Dans la pratique, la fenêtre gaussienne  $w$  est tronquée et discrétisée :

$$w(n) = e^{-\frac{1}{2}[\alpha \frac{n}{N/2}]^2} \quad -N/2 < n < N/2 \quad (2.9)$$

$$w(n) = 0 \quad -N/2 > n \text{ ou } n > N/2 \quad (2.10)$$

Le coefficient  $\alpha$  caractérise la largeur de la gaussienne par rapport à la fenêtre. L'énergie du signal étudié est donc concentrée dans cette région temporelle. On peut comparer l'effet du coefficient  $\alpha$  — qui détermine la largeur de la gaussienne — au *zero-padding* (bourrage de zéros) couramment utilisé en analyse spectrale avec d'autres enveloppes pour lisser et interpoler le spectrogramme. La dénomination *tail-padding* (bourrage de trainée) m'a été suggérée par Julius O. Smith, du Computer Center for Research in Music and Acoustics (CCRMA) de l'université de Stanford, lors de discussions sur ce sujet.

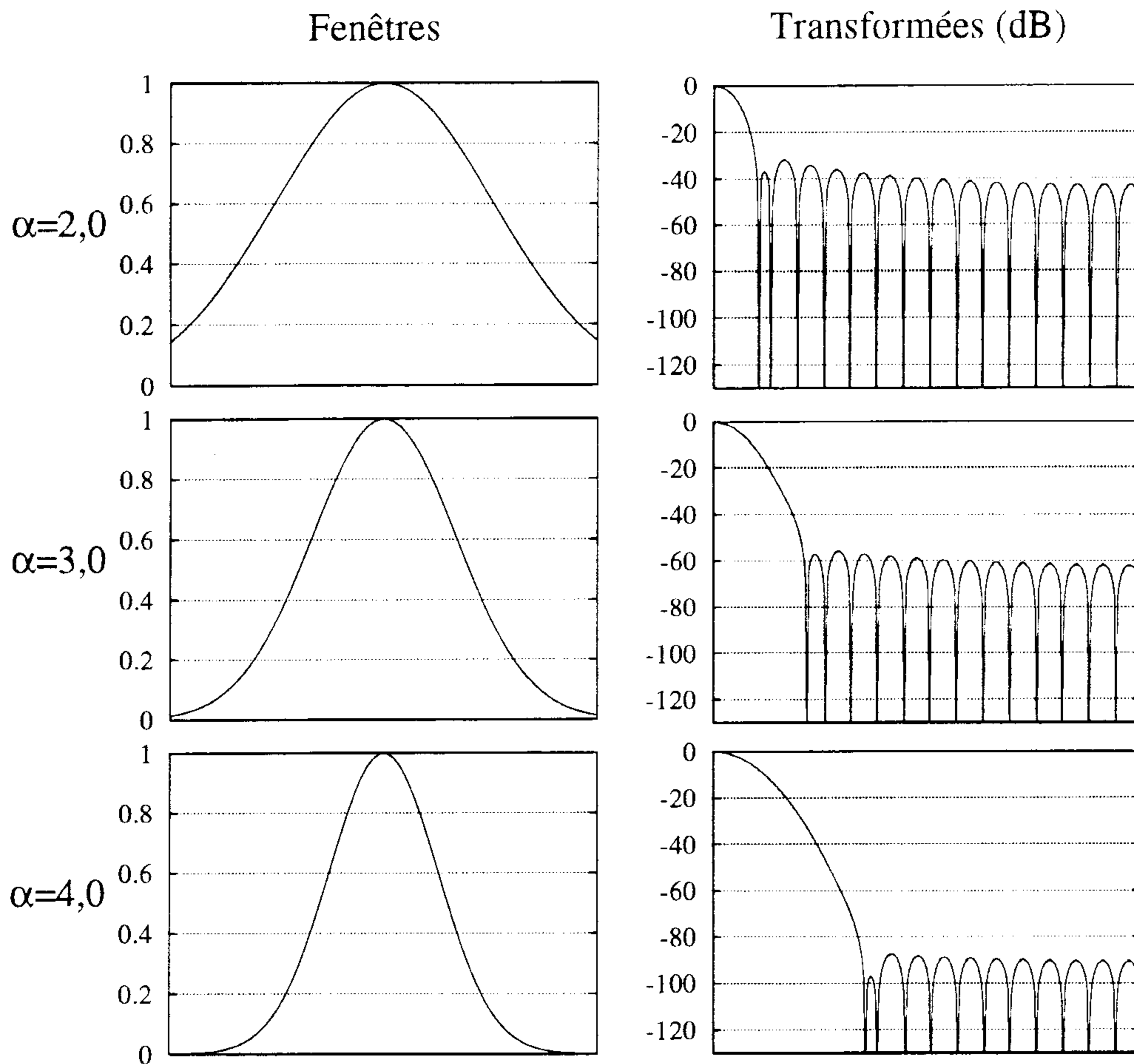


FIG. 2.2 – Transformées de Fourier de gaussiennes tronquées

Le *tail-padding* se distingue avantageusement du *zero-padding*: outre le lissage (qui permet d'améliorer l'interpolation quadratique), un *tail-padding* suffisamment important permet de minimiser l'amplitude des lobes secondaires. La figure 2.2 représente les spectres de gaussiennes tronquées pour un coefficient  $\alpha$  variant de 2 à 4. La figure 2.3 représente l'amplitude du lobe secondaire le plus important en fonction de  $\alpha$ .

Concluons en constatant qu'une valeur de  $\alpha = 4$  est suffisante pour étudier les sons échantillonnés sur 16 bits (qui procurent une dynamique maximale de 96dB). A l'inverse d'autres méthodes du même type [FH96], la précision de notre méthode d'analyse est justifiée. Elle peut s'affranchir de quelques défauts des méthodes dérivées de la transformée de Fourier discrète à court terme grâce à un procédé d'interpolation précis et au *tail-padding*.

### 2.1.3 Première sélection de la composante déterministe

Le résultat de cet algorithme est une liste des maxima locaux du spectre pour chaque trame. Chaque maximum local représente théoriquement une composante sinusoïdale

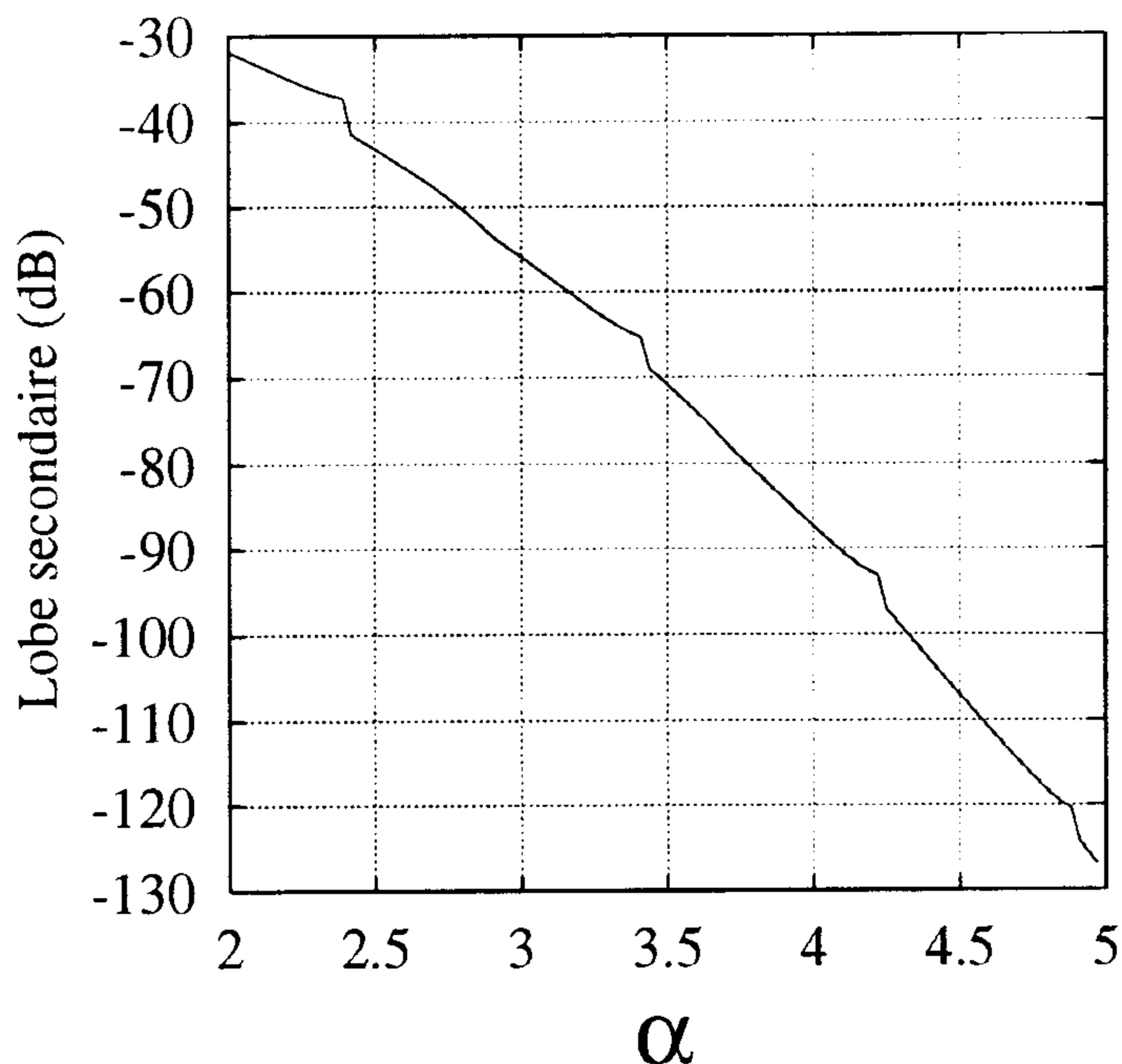


FIG. 2.3 – Amplitude du premier lobe secondaire en fonction de  $\alpha$

$C_j(T)$  du signal. Dans la pratique, tous les maxima locaux ne correspondent pas à des composantes harmoniques. Un signal “bruité” (souffle de flûtiste, bruits de marteaux... ) sera représenté sous la forme d’une multitude de composantes sinusoïdales de faibles amplitudes modulées en amplitude et en fréquence. Nous excluons, dans un premier temps, cette composante de notre étude; nous reviendrons sur cette question au chapitre 3.2, page 125.

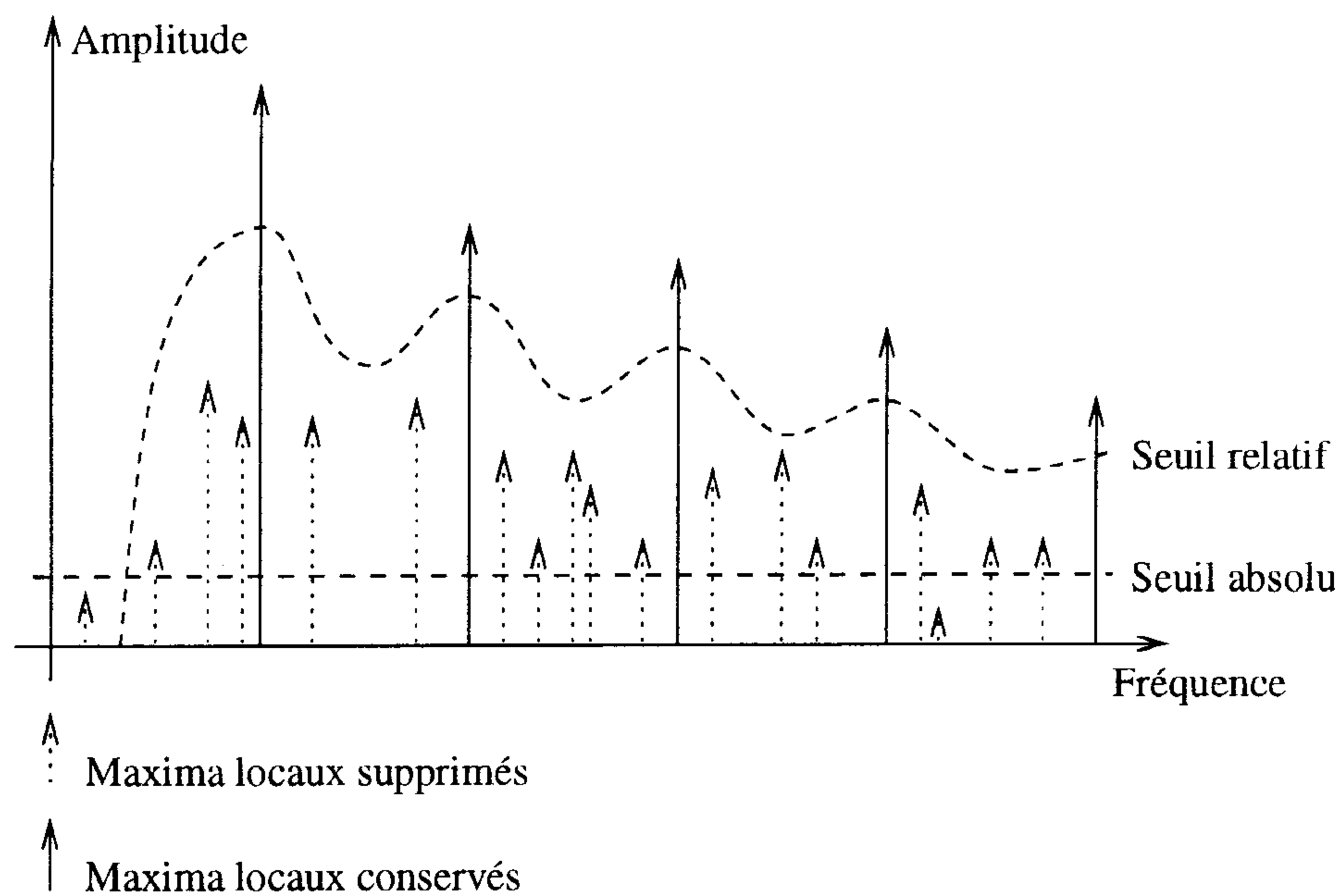
Afin d’éliminer le signal “bruité”, nous allons présélectionner les maxima locaux du spectre susceptibles d’être des composantes harmoniques par seuillages absolu et relatif (figure 2.4).

### Seuillage absolu

La présélection la plus simple consiste en un seuillage absolu : on ne conserve un maximum local que si son amplitude est supérieure à une amplitude minimale paramétrable. Cela permet d’exclure en grande partie le bruit de calcul (lobes secondaires de la fenêtre d’analyse et résolution de l’échantillonnage), et le bruit de fond.

### Seuillage relatif

L’amplitude des composantes aiguës est généralement beaucoup plus faible que celle des composantes graves. Ce phénomène n’est pas pris en considération par le seuillage absolu. Une solution consiste à effectuer un seuillage relatif : on compare chaque maximum local à une moyenne locale du spectre. Cette moyenne locale s’étend sur un nombre

FIG. 2.4 – *Seuillage des maxima locaux du spectre*

de canaux de transformée de Fourier paramétrable. Elle est pondérée par une courbe en cloche pour éviter les effets de bords. Un réglage judicieux consiste à prendre un nombre de canaux correspondant approximativement à la fréquence du fondamental. On parvient ainsi à supprimer la majeure partie des composantes bruitées se trouvant entre les composantes harmoniques.

#### 2.1.4 Poursuite des pics et resynthèse

La seule façon de vérifier l'efficacité de l'algorithme d'analyse spectrale décrit précédemment est d'effectuer la "resynthèse" du son. Cette resynthèse est effectuée selon la méthode suggérée par Robert J. McAulay et Thomas F. Quatieri [MQ86]. La première étape de cette démarche est la mise en relation des pics d'une trame avec les pics de la trame suivante. La seconde est la synthèse proprement dite.

##### Poursuite des pics

Nous avons déjà abordé ce sujet au §1.4.2. Une composante  $C_j(T)$  d'une trame est mise en relation avec une composante  $C_{j'}(T+1)$  de la trame suivante la plus proche en fréquence et en amplitude à condition qu'elle ne soit pas déjà mise en correspondance avec une autre.

Le principal problème est que le nombre de composantes  $J(T)$  peut varier d'une trame à l'autre. Il est donc nécessaire de gérer l'apparition et la disparition de composantes sinusoïdales. Les principes de cette gestion ont été décrits au §1.4.2, page 58.

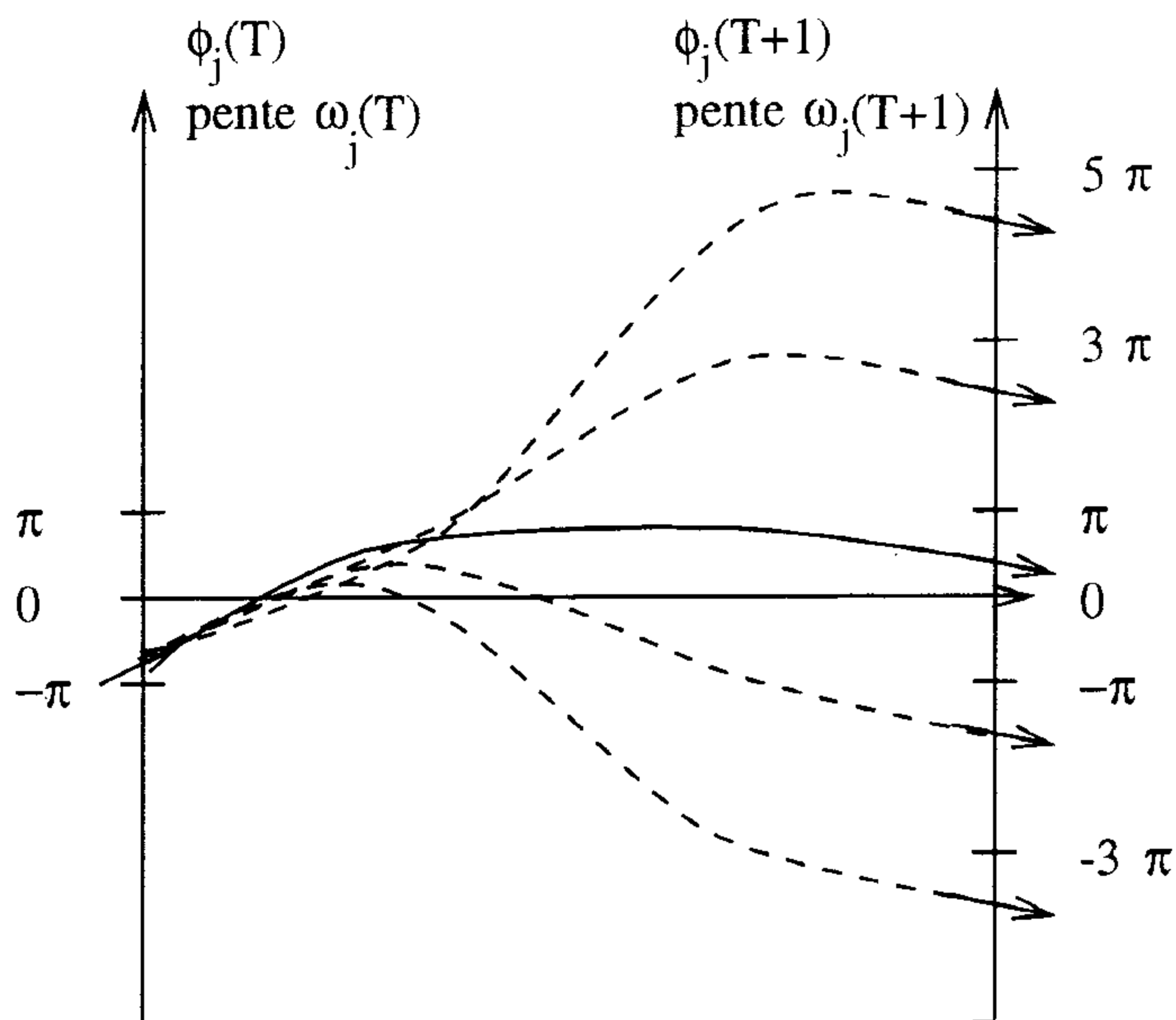


FIG. 2.5 – Interpolation cubique de la phase (d'après McAulay et Quatieri [MQ86])

## Resynthèse

La synthèse de la composante harmonique est indispensable pour vérifier le bon déroulement de l'analyse. Elle se montre aussi nécessaire à l'étude de la composante bruitée : nous allons par la suite soustraire la partie harmonique du son original afin d'en obtenir la composante bruitée.

Pour pouvoir soustraire le signal harmonique du signal d'origine, il faut un algorithme de synthèse respectant précisément les variations de fréquence, d'amplitude mais aussi de phase de chaque partiel.

La solution du problème proposée par Mac Aulay et Quatieri [MQ86] consiste à respecter la phase et la fréquence du signal en effectuant une interpolation cubique de la phase (figure 2.5). On connaît la phase  $\phi_j(T)$  et la pulsation  $\omega_j(T) = 2\pi f_j(T)$  d'une composante au milieu de la trame  $T$ . On connaît aussi la phase  $\phi_j(T+1)$  et la pulsation  $\omega_j(T+1)$  de la composante correspondante de la trame  $T+1$ . La pulsation  $\omega$  est la dérivée de la phase  $\phi$ . Pour tout intervalle séparant deux trames, on connaît la phase initiale, la pente initiale, la phase finale modulo  $2\pi$  et la pente finale de la phase. Puisqu'on connaît 4 paramètres de la courbe, il est possible de l'interpoler par un polynôme cubique. Le seul problème est que la phase finale est définie modulo  $2\pi$  (figure 2.5). La courbe, à choisir parmi une infinité de solutions, est la cubique la plus "lisse" [MQ86].

L'interpolation par polynôme cubique de la phase se révèle être très sensible à la précision de l'analyse spectrale. Son utilisation correcte suppose donc l'utilisation d'une analyse spectrale initiale très précise. Pour palier à cette sensibilité, on peut parfois avoir recours à une interpolation par polynôme quadratique moins sensible [DQ97] mais aussi intrinsèquement moins précise.

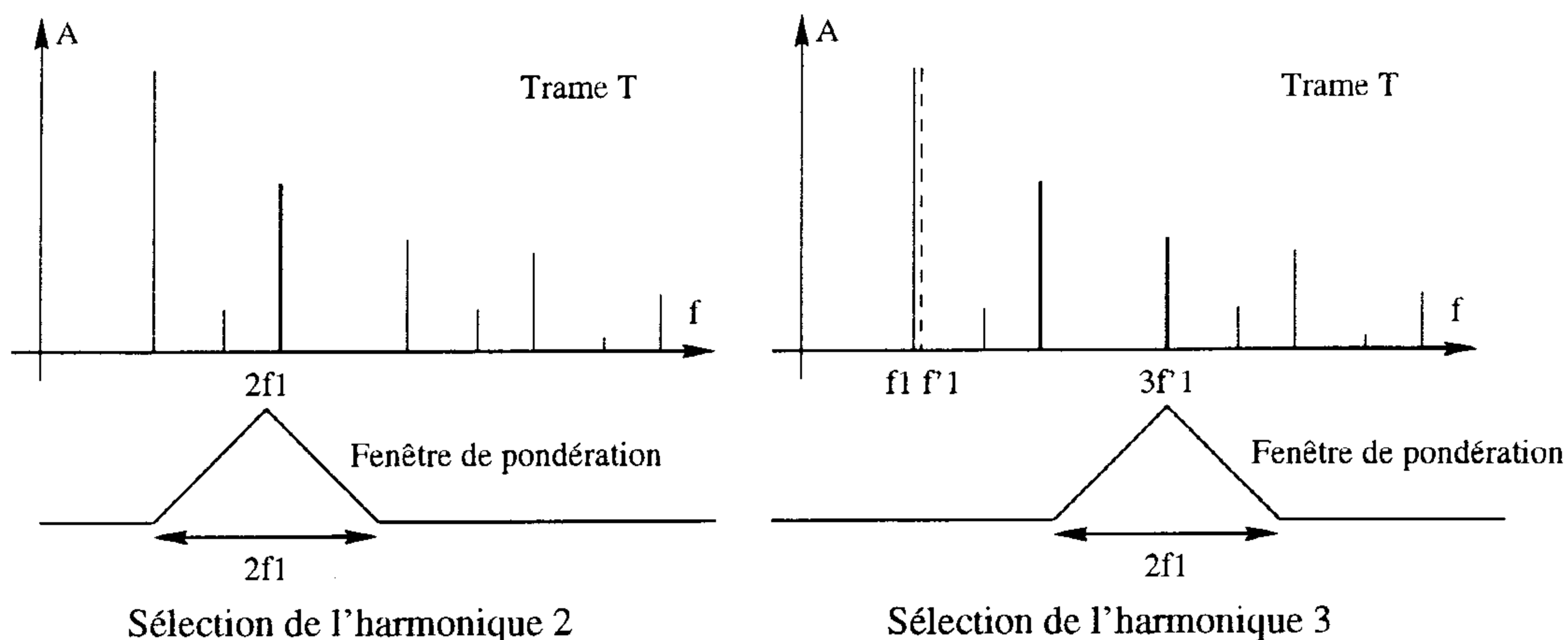


FIG. 2.6 – Sélection de la composante harmonique

## Résultats

La qualité de notre analyse spectrale nous a permis d'utiliser l'interpolation par polynôme cubique de la phase suggérée par Robert J. McAulay et Thomas F. Quatieri. La qualité de l'analyse peut directement être validée par la qualité de la resynthèse. Cette qualité peut être estimée empiriquement par une écoute des sons originaux et synthétisés. Dans la pratique, cette première estimation permet d'adapter les paramètres d'analyse spectrale (nombre d'échantillons de la trame, taux de recouvrement, coefficient  $\alpha$ , seuils absolu et relatif).

Une seconde estimation de la qualité de l'analyse peut être effectuée par soustraction directe, dans le domaine temporel, du signal original et du signal synthétique. Cette opération est rendue possible du fait que l'analyse et la resynthèse tiennent compte de façon précise des phases des signaux. Lorsque les paramètres d'analyse sont correctement définis, le son résiduel doit être composé de "bruit".

### 2.1.5 Sélection de la composante harmonique

#### Sélection du fondamental

La recherche du fondamental parmi les composantes  $C_{j=1..J(T)}(T)$  d'une trame est limité à un intervalle de fréquence prédéfini. Les critères de sélection sont l'amplitude et la proximité en fréquence de chaque composante détectée avec son homologue de la trame précédente. Cet algorithme, un peu simpliste, s'est révélé suffisant pour l'étude de la majorité des sons musicaux accordés. Il pourra être avantageusement amélioré par ajout d'autres critères tels que la présence de composantes harmoniques, l'absence de composantes sous-harmoniques et autres variations de phase [MB94, BP93].

#### Sélection des harmoniques

La sélection des  $H$  harmoniques  $C_{h=1..H}^H(T)$ , avec  $h = 1$  correspondant à la composante fondamentale, consiste à sélectionner les composantes de fréquences multiples ou quasi-multiples de la fréquence fondamentale  $f_1^H(T)$ .

Cette sélection s'opère pour chaque harmonique  $C_{h=2..H}^{\mathcal{H}}(T)$  selon un critère d'amplitude et de "multiplicité" de la fréquence par rapport à la fréquence fondamentale  $f_1^{\mathcal{H}}(T)$ . Le critère est basé sur une fenêtre de pondération de l'amplitude des composantes centrée à la fréquence  $h \times \bar{f}_1(T)$ ,  $\bar{f}_1(T)$  étant une fréquence fondamentale "moyenne" établie en tenant compte des déviations de fréquences des harmoniques précédemment détectés. Ce réajustement de la fréquence fondamentale "moyenne" permet de sélectionner des composantes quasi-harmoniques (figure 2.6).

Après cette étape de sélection, nous disposons de  $H$  composantes harmoniques  $C_h^{\mathcal{H}}(T)$ . Dans le cas général,  $H$  est de l'ordre de 40.

### 2.1.6 Conclusions

La méthode d'analyse spectrale que nous proposons a le double avantage d'être rapide (FFT) et précise (interpolation et *tail-padding*). La précision et la rapidité de cette méthode sont indispensables à l'étude d'un important *corpus* de sons musicaux. Le travail a été réalisé en langage C standard sous la forme d'un ensemble de programmes (annexe A). Cette réalisation "modulaire" permet d'ajouter aisément de nouveaux programmes en fonction des besoins (étude de la phase notamment) et de les porter sur diverses plateformes. Au terme de son évolution, elle pourra être intégrée à un logiciel graphique d'analyse et de synthèse des sons.

## 2.2 Réduction des données spectrales par ACP

L'analyse spectrale décrite précédemment fournit typiquement des informations de fréquence, d'amplitude et de phase d'une quarantaine de composantes sinusoïdales variant dans le temps. Cette masse d'information est d'une part difficile à interpréter et d'autre part conduit à une forme de synthèse additive coûteuse en calculs (§1.2.1, page 29).

Nous décrivons ici une méthode de réduction de données, fondée sur l'analyse en composantes principales (ACP), qui permet à la fois de simplifier l'interprétation des données et le processus de synthèse associé.

La section 2.2.1 introduit les principes généraux de l'analyse des données. La section 2.2.2, page 74, présente la réduction de données spectrales par analyse en composantes principales (ACP). La section 2.2.3, page 75, en présente une interprétation géométrique. Enfin, la section 2.2.4, page 77, illustre nos propos avec l'analyse de sons synthétiques dont l'on maîtrise les paramètres.

### 2.2.1 Introduction aux méthodes d'analyse des données

Les fondements mathématiques de l'analyse des données remontent au début du siècle. Ils sont l'œuvre de K. Pearson, initiateur de l'analyse en composantes principales, et de H. Hotelling, initiateur de l'analyse canonique. Ces méthodes étaient et sont encore largement utilisées en psychométrie, l'expérience psychoacoustique de John M. Grey (§1.3.2, page 50) en est un bon exemple. Les méthodes d'analyse des données reposent sur le fait que l'utilisation d'un volume important d'information rend indispensable leur traitement par des méthodes rigoureuses pour en dégager les paramètres importants.

La statistique descriptive classique ne permet que l'établissement de pourcentages et de diagrammes mesurant l'intensité d'une caractéristique en fonction d'un paramètre. Ces méthodes sont insuffisantes lorsqu'il s'agit, comme c'est le cas dans notre étude, d'extraire des informations d'un très grand volume de données d'où l'on veut, en extrayant les redondances, déduire des relations entre diverses caractéristiques.

L'objectif des méthodes d'analyse des données est de mettre en évidence les phénomènes importants en faisant un minimum d'hypothèses *a priori* sur les importances relatives des informations.

En analyse des données, on distingue généralement deux ensembles : *les individus et les caractères relatifs* à ces individus. L'ensemble des individus observés peut provenir d'un échantillonnage dans une population. Sur ces individus, on relève un large nombre de caractères qualificatifs ou quantitatifs significatifs de l'étude. Pour notre étude, les individus considérés sont les trames d'un son analysé et les caractères quantitatifs sont les amplitudes  $a_h^H(T)$  des  $H$  premiers harmoniques.

les méthodes d'analyse des données peuvent être séparées en deux catégories :

- les méthodes d'*analyse factorielle* qui supposent qu'il existe un petit nombre de caractères numériques indépendants sous-jacents aux caractères observés. Ces "*facteurs communs*" rendent compte des dépendances entre les quantités observées.
- les méthodes de *classification* qui visent à regrouper les individus en groupes homogènes.

L'*analyse en composantes principales* ou ACP, que nous utilisons dans notre étude, appartient à la première catégorie. Nous avons retenu cette méthode en raison de sa simplicité et de la nature des données analysées. Une extension de cette méthode, usant des nombres complexes est, en effet, illustrée en 3.1. Il est souvent admis que les différentes méthodes d'analyse factorielle fournissent des résultats voisins et que la qualité des données (précision, normalisation) est déterminante.

Lorsque le nombre de caractères est  $p$ , un individu  $i$  est un point dans un espace à  $p$  dimensions, de coordonnées  $x_i^1, x_i^2, \dots, x_i^p$ . Cet espace est appelé l'espace des individus. Si  $p$  est supérieur à trois, il est fort difficile, pour un être humain, de l'étudier.

En bref, ce type de problème est similaire au problème du dessinateur : il ne dispose que des deux dimensions du plan de travail pour dessiner un objet à trois dimensions. La solution à ce problème consiste à choisir le meilleur point de vue pour représenter au mieux un objet à trois dimensions sur un plan à deux dimensions (figure 2.7). "représenter au mieux" est un terme ambigu qui recouvre le critère de définition de l'espace de projection. Nous reviendrons plus loin sur ce critère.

L'analyse en composantes principales permet d'obtenir une représentation dans un espace de dimension usuelle en réduisant le nombre de caractères descriptifs. Ces caractères ne constituent pas une simple sélection des caractères initiaux, ce sont de nouveaux caractères composites appelés "*composantes principales*". Mathématiquement, les composantes principales sont définies par combinaisons linéaires des caractères d'origine. L'étude de ces combinaisons linéaires est une clé importante pour l'interprétation des résultats d'une telle analyse.



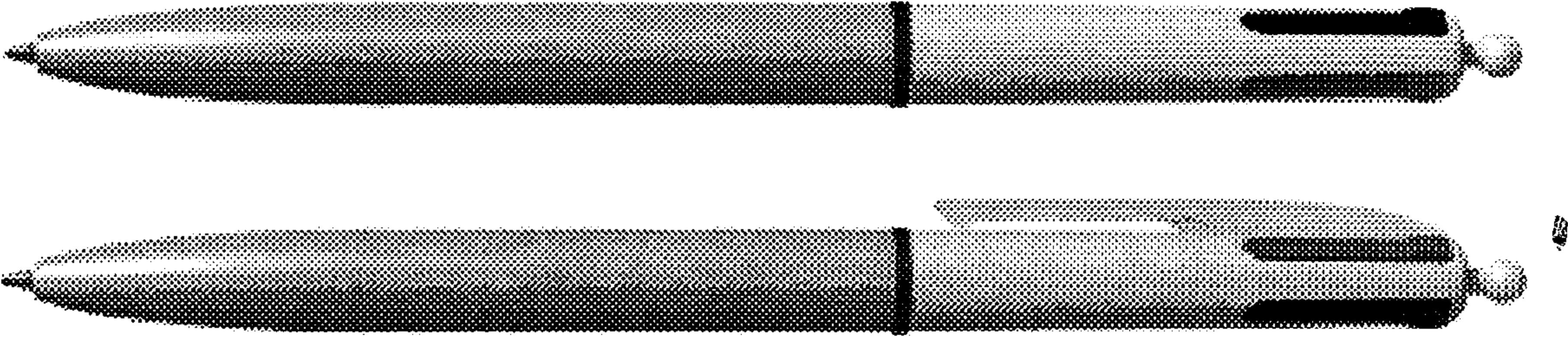


FIG. 2.7 – Projections d'un objet à trois dimensions dans le plan

Considérant un nuage de points-individus, défini par la matrice des individus :

$$\mathcal{X} = \begin{pmatrix} x_1^1 & x_1^2 & \cdots & x_1^P \\ x_2^1 & x_2^2 & \cdots & x_2^P \\ \vdots & \vdots & \ddots & \vdots \\ x_J^1 & x_J^2 & \cdots & x_J^P \end{pmatrix} \quad (2.11)$$

Il est possible de calculer les distances entre tous les points pris deux à deux. Le critère de définition des composantes principales est de conserver au mieux les distances initiales entre les points. Notons qu'une projection ne peut que réduire ces distances. Le critère peut donc être explicité comme : "rendre maximale la somme des carrés des distances", c'est-à-dire de choisir le sous-espace, pour lequel les points projetés sont le plus dispersés.

Le procédé explicite d'obtention des composantes principales est, comme nous l'avons précisé, assez simple : elles sont les vecteurs propres de la matrice de covariance des caractères. L'ACP se résume donc à l'établissement d'une matrice de covariance, à sa diagonalisation et à la projection du nuage de points individus sur les premiers vecteurs propres (classés selon les valeurs propres). Cette formulation simplifiera l'interprétation des résultats et l'extension de l'ACP aux nombres complexes.

### 2.2.2 Analyse en composantes principales des données spectrales

Pour représenter fidèlement un son, le volume de données harmoniques brutes est considérable : le nombre  $H$  de composantes harmoniques  $C_{h=1..H}^{\mathcal{H}}(T)$  étant de l'ordre de trente à cinquante. Dans la pratique, il faut tenir compte de l'évolution temporelle des amplitudes  $a_h^{\mathcal{H}}(T)$  et de la fréquence du fondamental  $f_1^{\mathcal{H}}(T)$  pour restituer les variations spectrales d'un son musical harmonique en synthèse additive.

La caractéristique essentielle de la représentation harmonique des sons musicaux est sa généralité : tous les sons pseudo-périodiques peuvent être représentés. En contrepartie, cette représentation n'est pas optimale — en terme de nombre de composantes — pour un son donné.

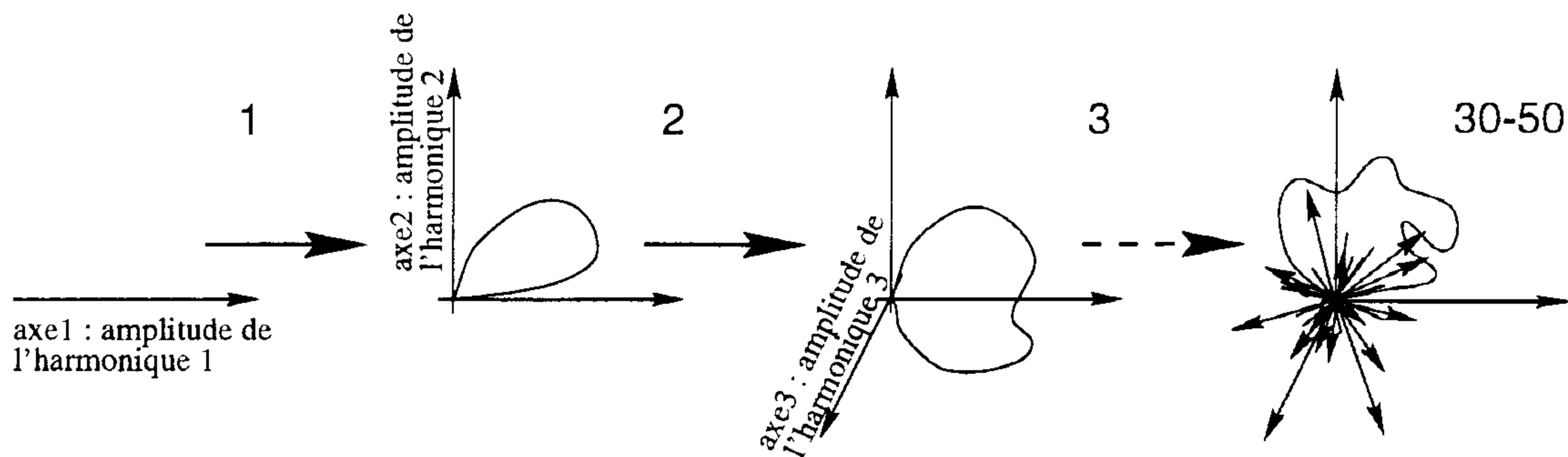


FIG. 2.8 – Construction de la trajectoire harmonique

L'analyse en composantes principales permet de trouver une représentation optimale du son considéré en terme de nombre de composantes le décrivant. Cette analyse consiste à extraire les redondances présentes dans les données spectrales. Notre analyse porte ici sur les amplitudes  $a_h^{\mathcal{H}}(T)$  des harmoniques. Le résultat de cette analyse est une nouvelle représentation du son. Comme nous le verrons, cette représentation peut être à la fois un outil d'analyse du timbre mais aussi servir de base à une forme de synthèse par combinaison de tables d'ondes.

La section 1.2.3, page 35, a déjà fait état des travaux de Gregory Sandell, Andrew Horner, Lippold Haken et James Beauchamp en partie comparables aux nôtres [SM95, HBH93, BH95, HB95, Hor95, CH96, Hor97, YH97]. N'oublions pas que toutes ces études avaient pour seule ambition de réduire le coût de calcul et de mémoire de l'opération de synthèse par tables d'ondes multiples.

Contrairement aux précédentes, l'étude que nous présentons ici se focalise avant tout sur l'interprétation de la réduction de données en termes psychophysiques. L'étude de A. Horner, J. Beauchamp et L. Haken allait jusqu'à affirmer que les résultats de l'analyse ont une nature "artificielle" [HBH93]<sup>1</sup>. Cette interprétation nous a permis d'améliorer le processus d'analyse en divers points, d'en comprendre les limitations, et, de proposer quelques solutions permettant de s'en affranchir.

### 2.2.3 Interprétation géométrique

Comme nous l'avons vu en 2.2.1, l'analyse en composantes principales peut être interprétée de façon géométrique. Pour représenter l'évolution des harmoniques sous la forme d'une courbe paramétrée — la trajectoire harmonique  $\mathcal{A}$  — nous pourrions associer au premier axe l'amplitude  $a_1^{\mathcal{H}}(T)$  du premier harmonique, au second axe l'amplitude  $a_2^{\mathcal{H}}(T)$  du second harmonique... et ce, jusqu'à  $H$ , nombre d'harmoniques (figure 2.8). Cette représentation, bien que mathématiquement correcte, n'est pas exploitable : le nombre de dimensions  $H$  de l'espace dans lequel  $\mathcal{A}$  est représenté est de l'ordre de trente à cinquante.

La trajectoire harmonique  $\mathcal{A}$  est une courbe à  $H$  dimensions, l'analyse en composantes principales va "choisir le meilleur point de vue" pour la représenter. En effet, elle permet d'obtenir une base de vecteurs  $P_1, \dots, P_H$ , optimale au sens de la consommation de variance.

1. [HBH93, page 342]: "On the other hand, these spectra may have a rather artificial relation to the original tone, since generally no basis spectrum will resemble any of the actual analysis spectra."

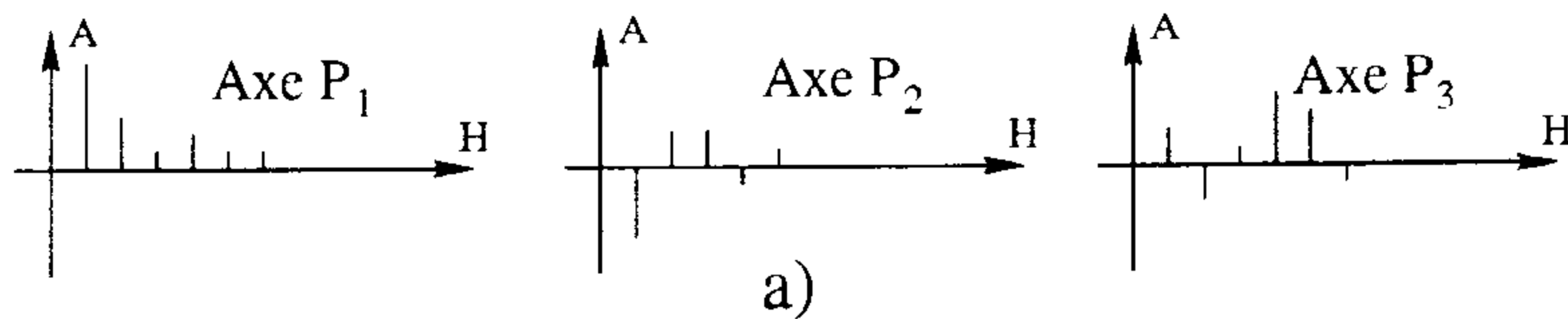
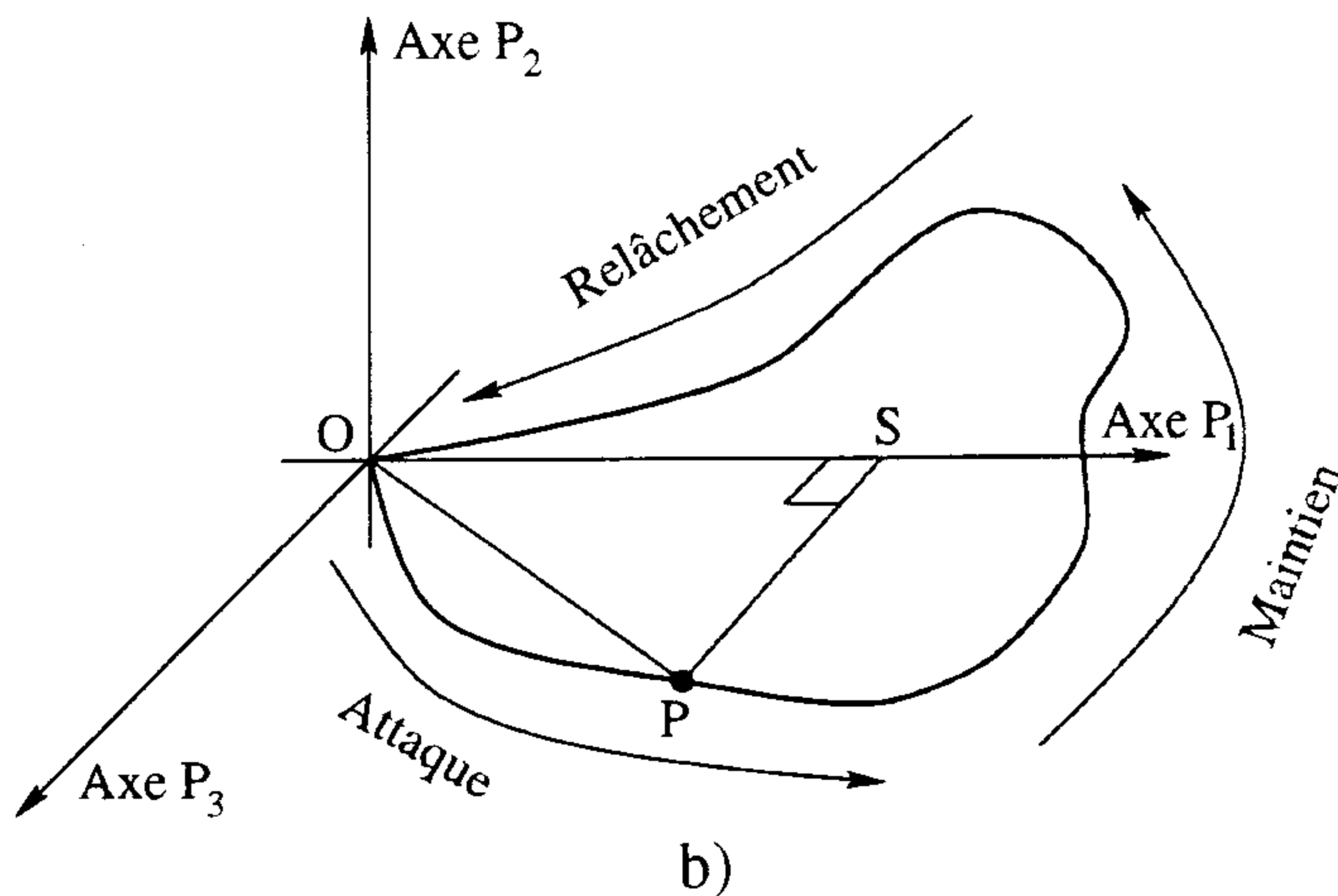
Définition des axes du sous espace principal  $\mathcal{P}$ Trajectoire du son dans le sous espace principal  $\mathcal{P}$ 

FIG. 2.9 – Axes et trajectoire

Ces vecteurs sont ordonnés selon leur “importance” ou plus exactement par la quantité de variance qu’ils “consomment”. Il est possible de se limiter aux premiers vecteurs pour obtenir une représentation significative de la plupart des sons musicaux. Cette simplification pouvant être validée en examinant l’amplitude du signal d’erreur résiduel. La trajectoire harmonique  $\mathcal{A}$  est ainsi projetée dans un *sous-espace harmonique principal*  $\mathcal{P}$  de dimension plus réduite. Cette projection est définie par :

$$\mathcal{A} \rightarrow \mathcal{V} : V_p(T) = \sum_{h=1}^H a_h^{\mathcal{H}}(T) \cdot P_p^h \quad (2.12)$$


Chaque axe du sous-espace harmonique principal  $\mathcal{P}$  peut être exprimé sous la forme d’une combinaison des axes de l’espace d’origine. Etant donné que chaque axe de l’espace d’origine représentait un harmonique du son, chaque axe du sous-espace harmonique principal  $\mathcal{P}$  représente une combinaison d’harmoniques, c’est-à-dire un *spectre harmonique principal*.

Dans la perspective d’analyse des sons, nous nous limiterons aux trois premiers axes  $P_1$ ,  $P_2$  et  $P_3$ . Ces trois axes qui définissent  $\mathcal{P}$  sont caractéristiques du son et peuvent être interprétés comme suit :

- $P_1$  correspond au profil spectral d’énergie moyenne du son représenté,
- $P_2$  et  $P_3$  correspondent aux déviations spectrales du son.

La figure 2.9a représente un exemple de profil typique des spectres principaux :

- $P_1$  est associé au spectre moyen, dans la majorité des cas, l’harmonique 1 est prédominant.

- $P_2$  oppose l’harmonique 1 aux autres harmoniques : son amplitude prend une valeur négative. La signification physique de cette amplitude négative est que l’harmonique 1 de  $P_2$  est en opposition de phase avec celui de  $P_1$  : l’addition des spectres harmoniques principaux de ces deux axes fera ressortir les harmoniques de rangs élevés au détriment de l’harmonique 1.
- Le rôle de  $P_3$  est en général plus complexe et parfois négligeable.
- Les axes d’ordre supérieur sont souvent négligeables. Leur contribution peut être quantifiée sous la forme d’une enveloppe d’erreur résiduelle. 

La figure 2.9b représente une trajectoire  $\mathcal{V}$  typique. Elle représente les trois étapes de l’évolution dynamique du son : attaque, maintien et relâchement. Chaque point de la trajectoire  $\mathcal{V}$  représente un instant du son. La trajectoire débute et se termine en  $O$ . Le point  $O$  représente le silence et le point  $P$  le son à un instant donné. La distance  $OP$  correspond à l’amplitude du son à cet instant. La distance  $SP$  entre le premier axe et ce point représente la déviation du spectre par rapport au spectre moyen : elle quantifie la variation instantanée de timbre du son.

En résumé, le sous espace principal  $\mathcal{P}$  définit le comportement spectral du son alors que la trajectoire  $\mathcal{V}$  en définit la dynamique. Ces données harmoniques réduites permettent d’analyser les sons musicaux, ils permettent aussi de les synthétiser de façon fidèle.

#### 2.2.4 Analyse de sons synthétiques

L’analyse de sons synthétiques — dont l’on connaît parfaitement la composition *a priori* — permet de vérifier, par la pratique, les assertions faites précédemment. Elle permet aussi d’affiner ces assertions et d’améliorer certains traits de l’analyse en composantes principales.

##### Son à deux composantes sous-jacentes

Considérons en premier lieu un son `test1` relativement simple composé de 3 composantes sinusoïdales définies par :

$$a_1(t) = A(t) \times (1 - t) \quad (2.13)$$

$$a_2(t) = A(t) \times 1 \quad (2.14)$$

$$a_3(t) = A(t) \times 1 \quad (2.15)$$

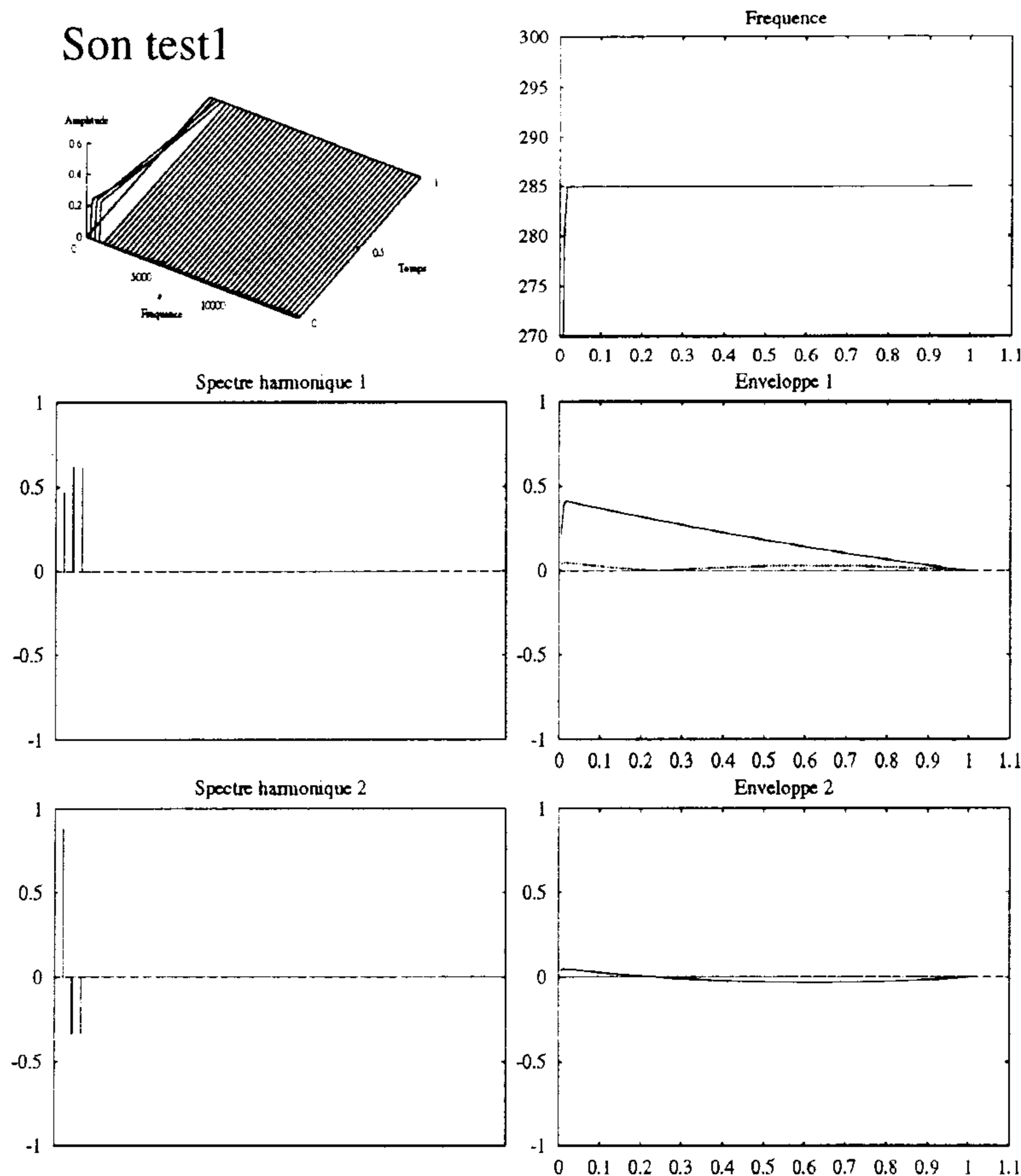
$$A(t) = 1 - t \quad (2.16)$$

$$f_1(t) = 285Hz \quad (2.17)$$

$$t \in [0, 1] \quad (2.18)$$

$a_2(t)$  et  $a_3(t)$  ont la même définition. Ainsi, le son `test1`, est constitué de deux composantes sous-jacentes. Le rôle de l’ACP va donc être de déduire ces composantes à partir des composantes sinusoïdales initiales.

La figure 2.10 représente les données spectrales réduites du son `test1` ainsi qu’une représentation en “chute d’eau” et de la fréquence fondamentale. Les deux premiers spectres

FIG. 2.10 – Données spectrales réduites du son `test1` non renormalisé

principaux  $P_1$  et  $P_2$  sont représentés sur la gauche. Leurs enveloppes temporelles respectives  $V_1(t)$  et  $V_2(t)$  sont représentées sur la droite, l'erreur résiduelle — correspondant à l'amplitude du signal d'erreur commis si l'on ne tenait compte que de la première composante — est représentée en pointillés.

On constate que le spectre harmonique  $P_1$  comporte toutes les composantes harmoniques, il correspond — à un coefficient de normalisation près — au spectre harmonique d'énergie moyenne défini par :

$$a_1^M = \sqrt{\int_0^1 [A(t) \times (1-t)]^2 dt} = \frac{1}{\sqrt{5}} = 0,45 \quad (2.19)$$

$$a_2^M = a_3^M = \sqrt{\int_0^1 A^2(t) dt} = \frac{1}{\sqrt{2}} = 0,58 \quad (2.20)$$

L'enveloppe  $V_1(t)$  est quasi-linéaire, elle correspond en grande partie à  $A(t)$ . L'erreur résiduelle est d'amplitude variable au cours du temps mais ses maxima sont d'amplitudes comparables quelle que soit l'amplitude globale du son.

Le spectre harmonique  $P_2$  oppose l'harmonique 1 et les harmoniques 2 et 3. Son enveloppe  $V_2$  concorde avec l'erreur résiduelle précédente. L'erreur résiduelle est nulle : les composantes  $P_1$  et  $P_2$  suffisent à décrire le son original.

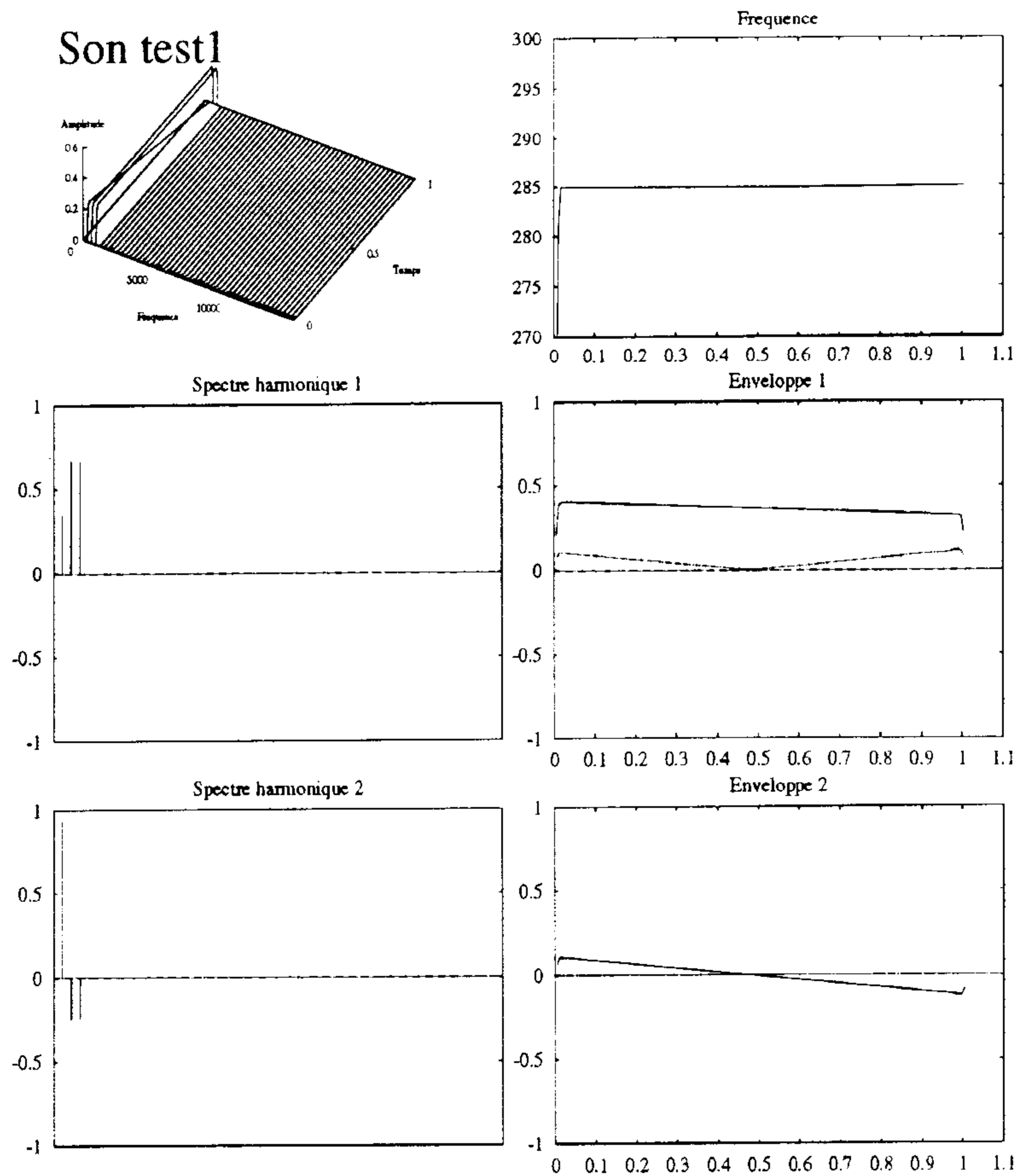


FIG. 2.11 – Données spectrales réduites du son test1 non renormalisé à enveloppe constante

### Remarques sur la normalisation des données

Les composantes principales sont censées “consommer” le maximum de variance — et implicitement d’erreur — et être ordonnées par ordre décroissant d’importance. La figure 2.11 représente l’analyse d’un son similaire à test1 dont l’enveloppe  $A(t)$  est constante.

$P_1$  est défini par :

$$a_1^M = \sqrt{\int_0^1 (1-t)^2 dt} = \frac{1}{\sqrt{3}} = 0,58 \quad (2.21)$$

$$a_2^M = a_3^M = \sqrt{\int_0^1 1 dt} = 1 \quad (2.22)$$

L’amplitude relative du premier harmonique dans  $P_1$  est plus faible que dans le cas précédent. L’enveloppe globale  $A(t)$  a affecté la définition du sous-espace  $\mathcal{P}$ . Les premières tentatives d’analyse/synthèse de sons réels par ce procédé ont conduit à des phénomènes de distorsion du timbre des phases transitoires d’attaque et de décroissance lorsque trois ou quatre axes étaient utilisés. Ces effets indésirables disparaissaient lorsque cinq ou six axes

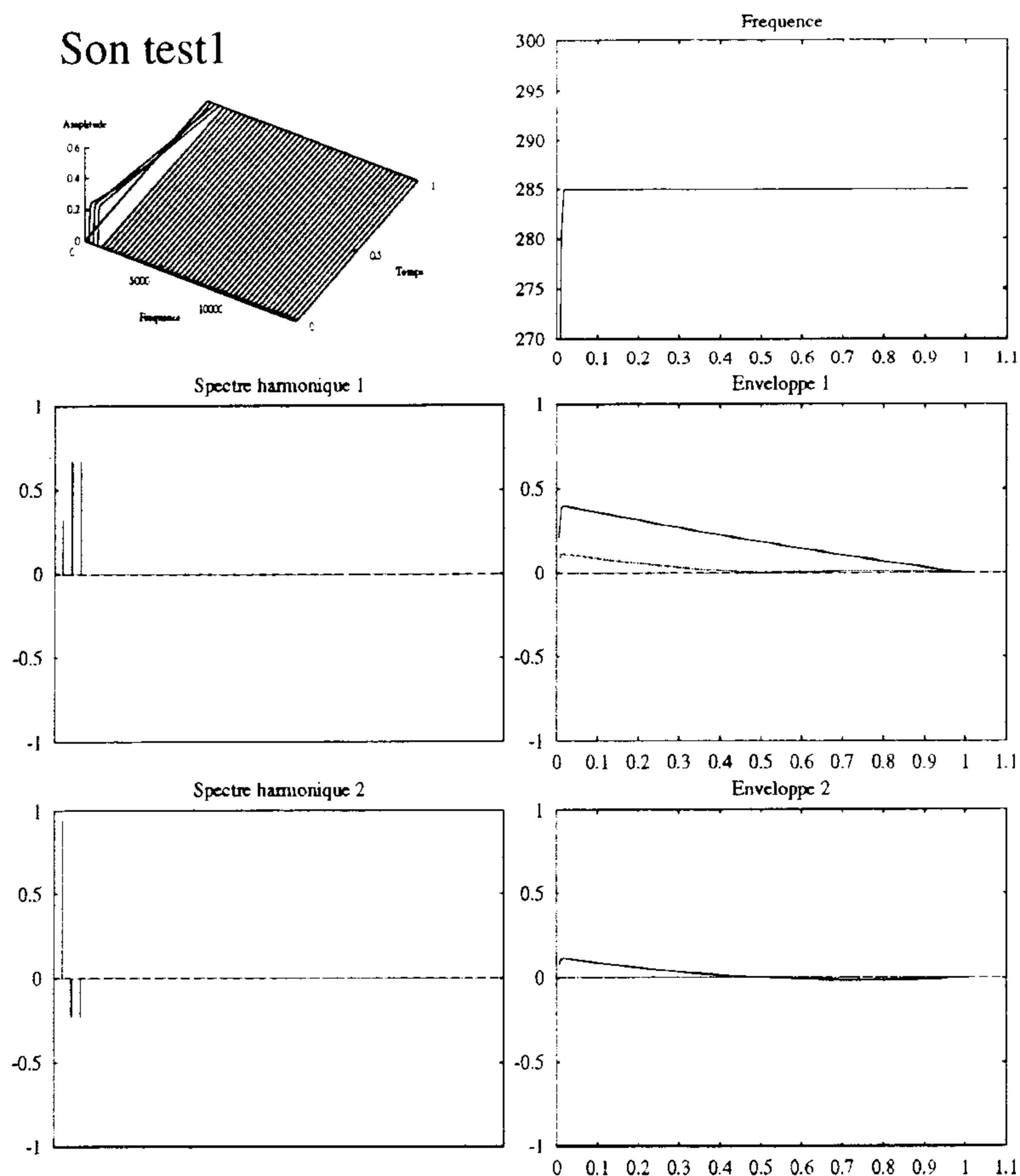


FIG. 2.12 – Données spectrales réduites du son `test1` renormalisé

étaient utilisés. De tels phénomènes ont été notamment décrits par A. Horner, J. Beauchamp et L. Haken [HBH93]<sup>2</sup>. Ils leur ont fait rejeter l'ACP au profit d'autres méthodes telles que les algorithmes génétiques. L'interprétation décrite dans les paragraphes suivants nous a permis de remédier à ces phénomènes.

Les variations spectrales les plus importantes se déroulent, pour la plupart des instruments, pendant les transitoires. Les procédés d'analyse spectrale et de synthèse ayant été validés par analyse/synthèse, seule la réduction de donnée peut être incriminée. Si les transitoires sont mal reconstruites, c'est parce qu'elles sont mal représentées : l'ACP ne leur accorde pas suffisamment "d'importance". En des termes plus précis, les individus sont intrinsèquement pondérés par leur amplitude globale  $A(t)$  : les spectres des phases transitoires de faible amplitude sont donc sous-pondérés. Nous avons adopté pour solution de normaliser les profils harmoniques selon la norme euclidienne. Géométriquement parlant,  $\mathcal{A}$  est projetée sur l'hypersphère de rayon unité de telle sorte que seules les *déviations angulaires*, caractéristiques des *variations de profil spectral*, soient prises en compte par l'analyse en composantes principales.

2. [HBH93, page 348] : "Also note that the error is consistently low in the sustain portion of the tone, but is much higher during the attack and decay."

La figure 2.12 représente les données réduites du son *test1* renormalisé. Les spectres harmoniques sont définis indépendamment de l'enveloppe  $A(t)$ . L'erreur résiduelle décroît avec l'enveloppe globale.

Dans la pratique, cette normalisation a conduit à une nette amélioration de la "qualité perceptive" des sons synthétisés en fonction du nombre d'axes considérés. Après normalisation, trois ou quatre axes suffisent pour reconstituer une bonne partie des sons musicaux "usuels". Ces résultats concordent, en terme de nombre de composantes nécessaires à une représentation correcte, avec les résultats obtenus par A. Horner, J. Beauchamp et L. Haken en usant des algorithmes génétiques [HBH93]. Bien que nous n'ayons pas eu l'occasion de le vérifier, il doit être possible d'étendre cette idée en tenant compte de facteurs psychoacoustiques. L'ACP "porterait son attention" sur les instants critiques du son grâce à une forme de "pondération psychoacoustique".

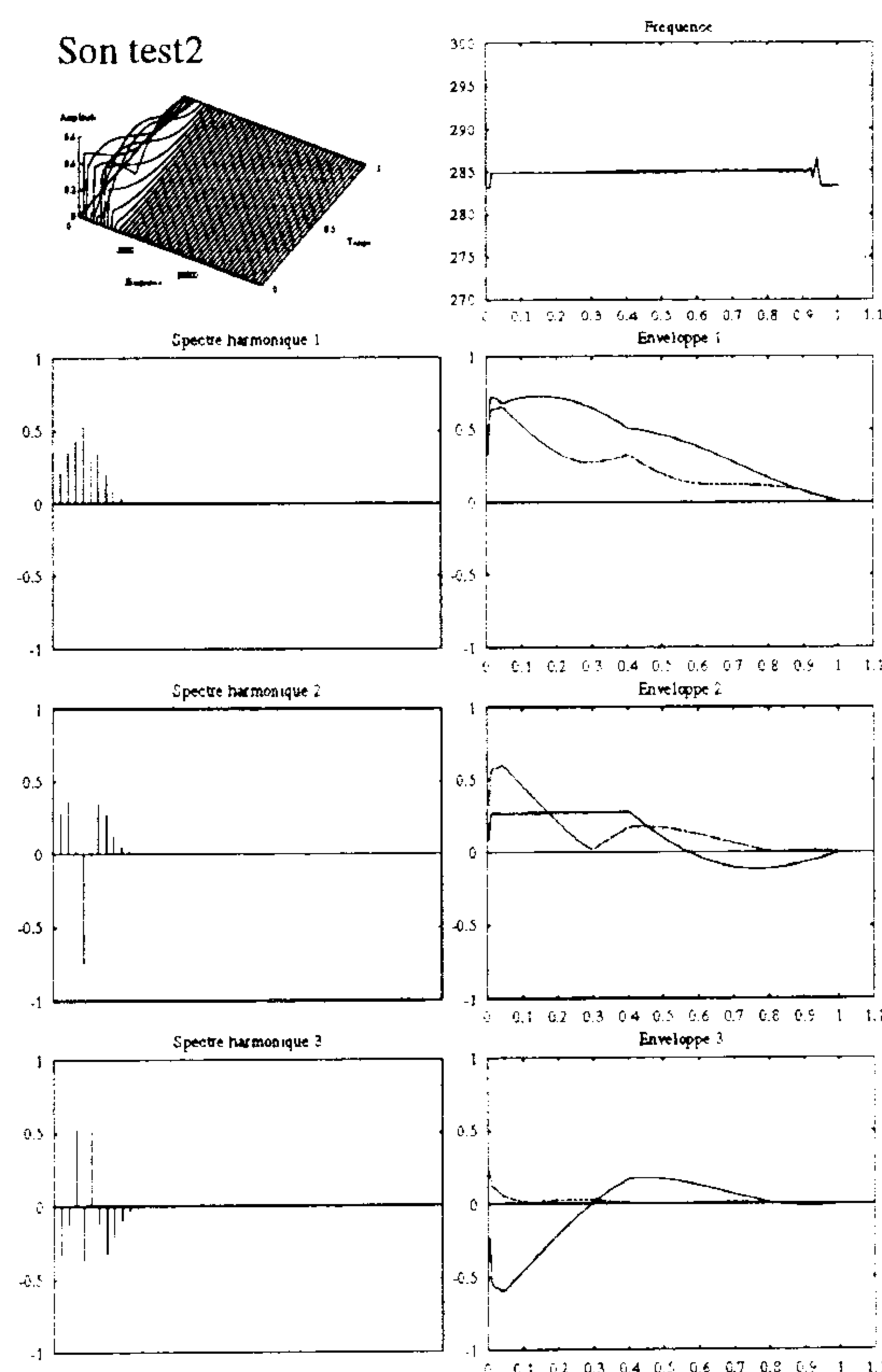


FIG. 2.13 – Données spectrales réduites du son FM *test2*

### Sons de modulation de fréquence à deux opérateurs

La figure 2.13 représente les données spectrales réduites d'un son synthétisé par modulation de fréquence à deux opérateurs (§1.2.5, page 38). La fréquence  $f_p$  de la porteuse est de 1140Hz, la fréquence de modulation  $f_m$  est de 285Hz, l'indice de



modulation  $I$  varie linéairement de 4 à 0. Nous allons maintenant observer comment les données réduites rendent compte des caractéristiques de ce son.

Une des caractéristiques de ce son est une quasi-symétrie spectrale autour de  $f_p$  : les harmoniques disposés de part et d'autre de  $f_c$  ont pour amplitudes les fonctions de Bessel  $J_1(I)$ ,  $J_2(I)$ ,... Les spectres harmoniques  $P_1$ ,  $P_2$  et  $P_3$  rendent compte de cette symétrie.

Une autre caractéristique de ce son est que "l'étalement" spectral global varie avec  $I$ . Les données spectrales réduites font clairement ressortir cette caractéristique :  $P_1$  et  $P_2$  ont des raies latérales sensiblement identiques qui s'additionnent puis se soustraient lorsque  $I$  diminue.

La dernière caractéristique de ce son est le comportement ondulatoire des fonctions de Bessel en fonction de  $I$ . Les ondulations sont plus importantes pour les raies centrales. Cette caractéristique se traduit par des oppositions de phase des raies centrales dans les spectres harmoniques  $P_1$ ,  $P_2$  et  $P_3$  et des ondulations dans leurs enveloppes.

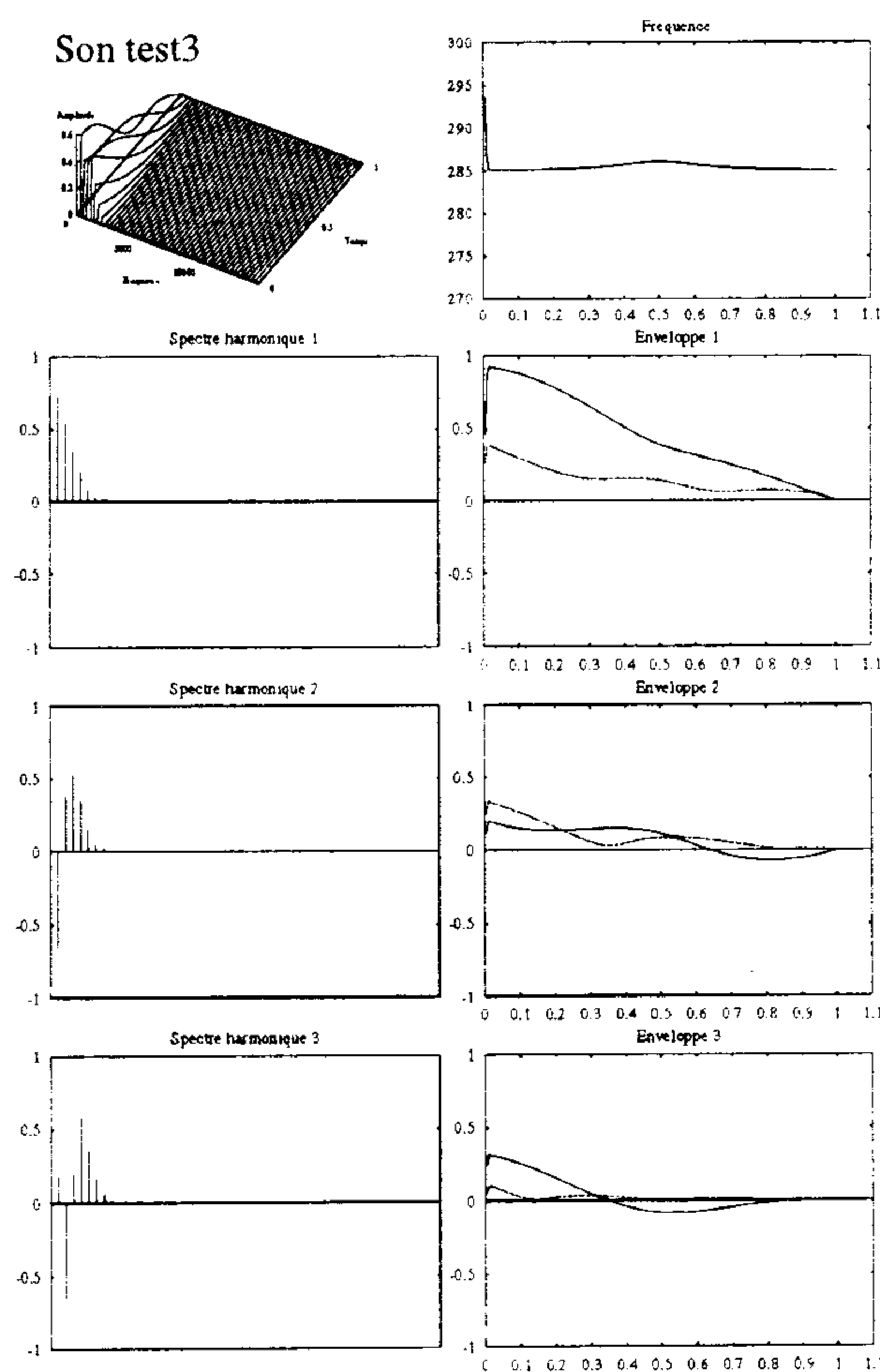


FIG. 2.14 – Données spectrales réduites du son FM test3

Le son `test3`, dont les données spectrales réduites sont représentées sur la figure 2.14 est similaire au son `test2` à l'exception de la fréquence de la porteuse  $f_p$  qui vaut ici 285Hz. La conséquence de cette modification est la disparition apparente de la symétrie spectrale du son : les raies spectrales de fréquences négatives se replient en 0 et se combinent aux raies spectrales de fréquences positives. Il en résulte un son en apparence plus complexe que le son `test2`. L'analyse en composantes principales appliquée au son `test3` donne des

résultats très similaires à ceux du son `test2`: l'enveloppe temporelle d'erreur résiduelle devient négligeable au delà de trois axes. Nous en concluons que la superposition des raies de fréquences positives et négatives n'a pas ajouté de degrés de liberté au son synthétisé, elle n'a fait que modifier la base de représentation du son.

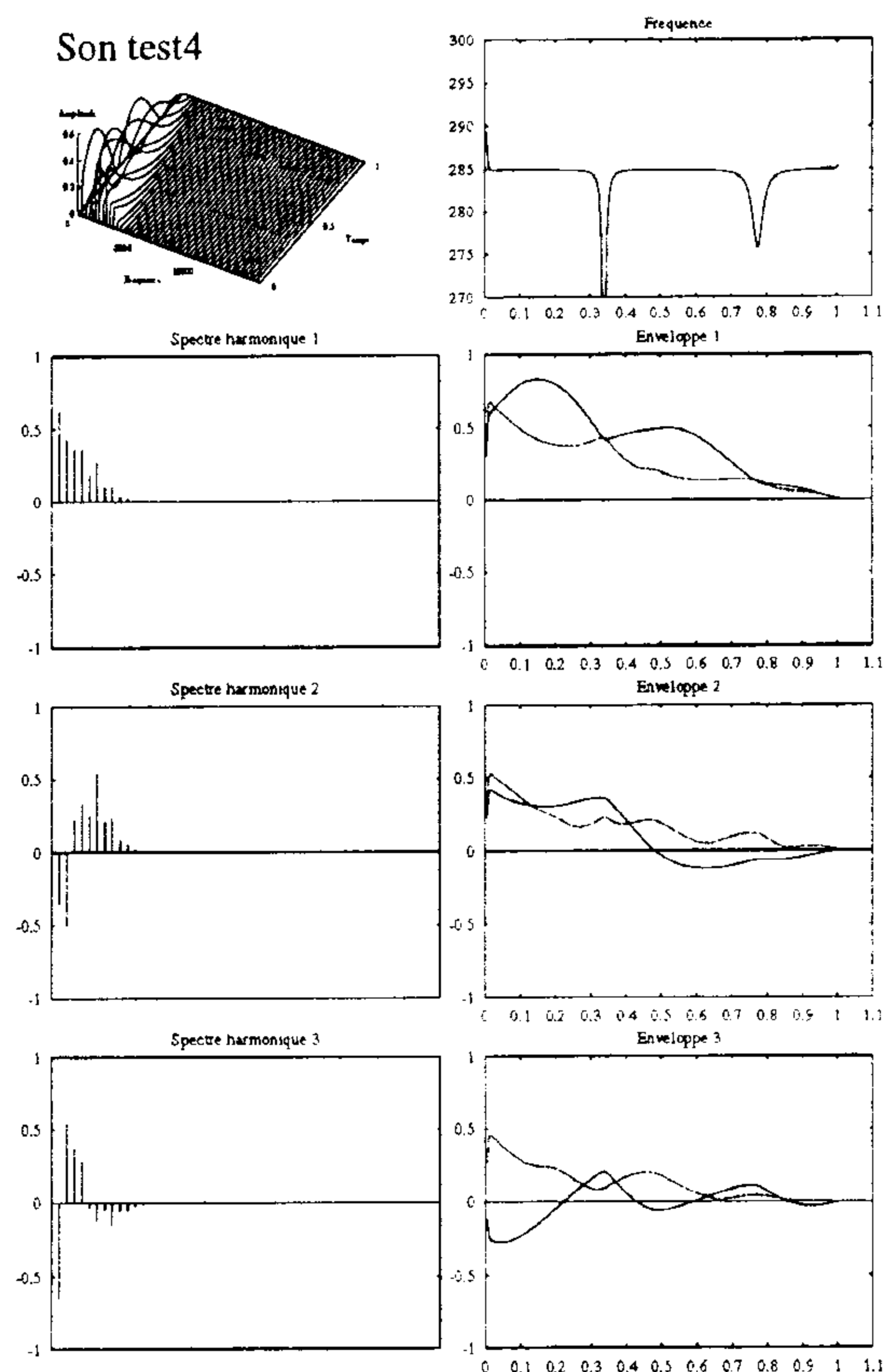


FIG. 2.15 – Données spectrales réduites du son FM `test4`

Le son `test4`, dont les données spectrales réduites sont représentées sur la figure 2.15 ajoute au son `test3` une variation d'indice de modulation de 8 à 0. On remarque d'emblée que l'erreur résiduelle reste considérable après le troisième axe. Nous avons constaté que cette erreur ne devenait négligeable — consommation de variance de l'ordre des 99% — qu'à compter de six axes. Il semble donc que l'étendue de variation de l'indice de modulation étende le nombre de degrés de liberté du son.

Pour un rapport de fréquences  $f_p/f_m$  donné, la variation de l'indice de modulation  $I$  définit à lui seul la "trajectoire" harmonique ou plus exactement l'ensemble des profils spectraux disponibles.

Notre méthode d'analyse permet donc de caractériser, au travers des sous-espaces correspondants, ces profils spectraux. Rappelons qu'une des difficultés majeures de la synthèse par modulation de fréquence à deux opérateurs consiste à déterminer le rapport  $f_p/f_m$  et l'indice  $I$  permettant la synthèse d'un profil spectral désiré (§1.2.5, page 38). La détermination automatique de ces paramètres permettrait peut être de remettre cette forme de synthèse au goût du jour. Elle pourrait, par exemple, être facilitée par

l'établissement d'un "catalogue" de sous-espaces caractéristiques des rapports  $f_p/f_m$ . La détermination des paramètres pourrait alors être effectuée par sélection de la meilleure projection du profil spectral désiré dans un sous-espace du catalogue. Ce sous-espace déterminerait donc le rapport  $f_p/f_m$ . La position du point représentatif du profil dans celui-ci permettrait de déterminer l'indice  $I$  (on choisirait le point de la "trajectoire" défini par  $I$  le plus proche du point représentatif).

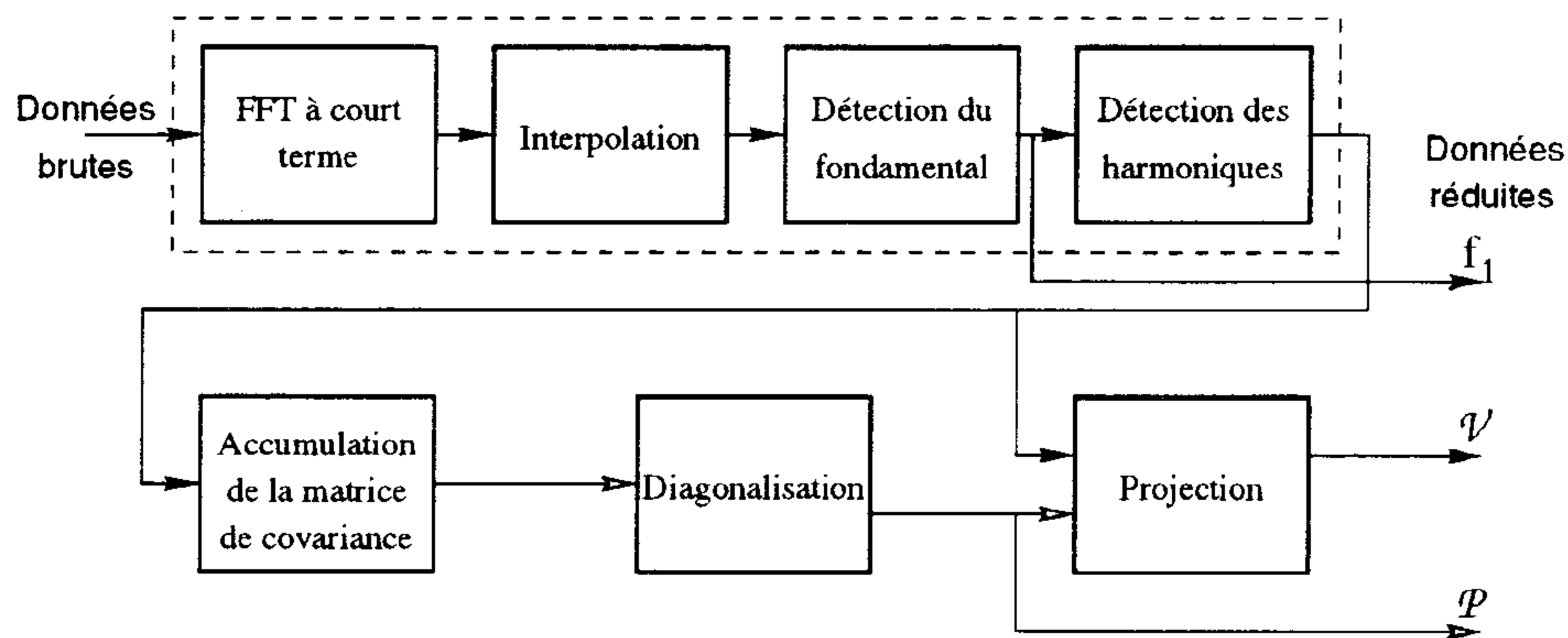


FIG. 2.16 – Procédé d'analyse et de réduction de données

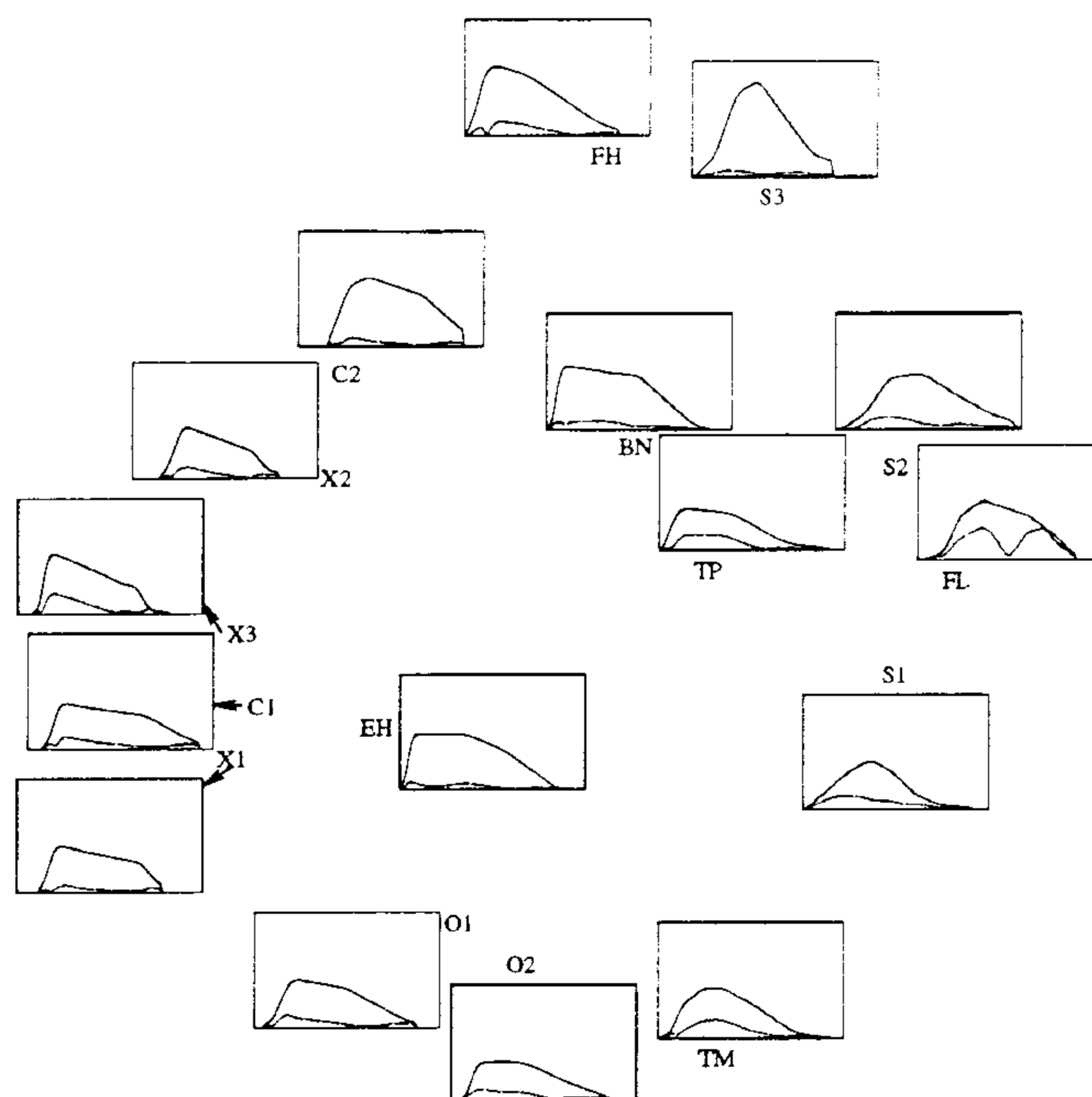
### 2.2.5 Conclusions

La figure 2.16 résume notre procédé d'analyse et de réduction de données. Cette procédure nous permet d'obtenir une représentation des sons harmoniques sous forme de données réduites. Ces données se résument à :

- la fréquence fondamentale du signal  $f_1(t)$ ,
- la trajectoire  $\mathcal{V}$  définie par les enveloppes dynamiques  $V_1(t)$ ,  $V_2(t)$  et suivantes,
- le sous-espace  $\mathcal{P}$  défini par les spectres  $P_1$ ,  $P_2$  et suivants.

L'analyse de sons synthétiques nous a permis d'une part d'appréhender les problèmes de normalisation des données et d'autre part, de constater que le nombre de dimensions de  $\mathcal{P}$  nécessaires à une représentation correcte dépend de la "complexité" du son analysé.

Cette pré-étude nous a donc donné la possibilité d'affiner notre méthode d'analyse et de nous préparer à l'interprétation des données réduites.

FIG. 2.17 - Plan XY de l'espace de Grey et enveloppe  $V_1(t)$ 

## 2.3 Application à l'étude des sons musicaux

Dans la section précédente nous avons donné quelques éléments permettant d'interpréter les données réduites. Nous allons pousser plus avant cette interprétation dans cette section. Pour ce faire, nous commencerons par établir un lien entre l'espace des timbres perceptifs de John M. Grey et les données réduites (§2.3.1, page 85). En section 2.3.2, page 88, nous examinerons dans le détail les données réduites de quatre sons musicaux. Enfin, en section 2.3.4, page 98, nous étendrons notre discussion au *corpus* de 42 sons musicaux analysés en annexe B.

### 2.3.1 Analyse des sons de Grey

La représentation des sons en "chutes d'eau" utilisée par John M. Grey permet de représenter leurs variations spectrales. En les disposant selon les axes de son espace perceptif, John M. Grey a observé certaines relations entre axes perceptifs et évolution spectrale (§1.3.2, page 50).

L'étude présentée ici substitue la représentation sous forme de données réduites à la représentation en "chutes d'eau"; cette opération nous permet de lier données réduites et axes de l'espace de John M. Grey.

La figure 2.17 dispose l'enveloppe  $V_1(t)$ , et l'erreur résiduelle correspondante, selon les axes X et Y de l'espace de John M. Grey. L'axe X, représenté horizontalement, semble opposer les sons pour lesquels  $V_1(t)$  varie rapidement (X3, X1, C1) aux sons pour lesquels  $V_1(t)$  varie lentement (S2, FL, S1). L'enveloppe  $V_1(t)$  correspond à l'amplitude du spectre

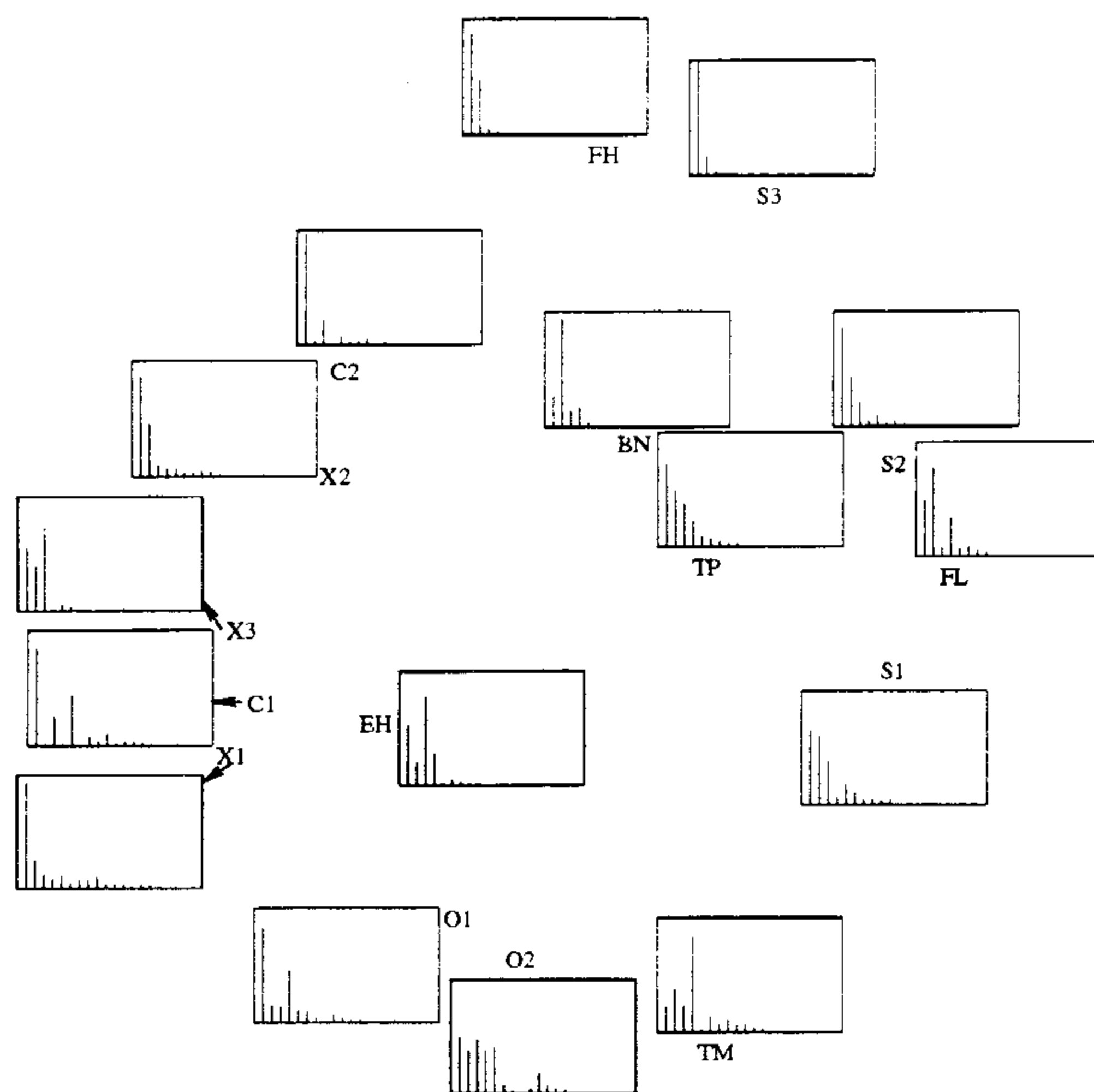


FIG. 2.18 – Plan  $XY$  de l'espace de Grey et spectre  $P_1$

harmonique principal  $P_1$ . Elle correspond donc grossièrement à l'enveloppe d'amplitude globale du son. Cette représentation indique une relation entre l'axe  $X$  de l'espace de John M. Grey et l'enveloppe  $V_1(t)$  du spectre harmonique  $P_1$ .

La figure 2.18 dispose le spectre harmonique  $P_1$  sur le plan  $XY$  de l'espace perceptif. L'axe  $Y$ , représenté verticalement tend à opposer :

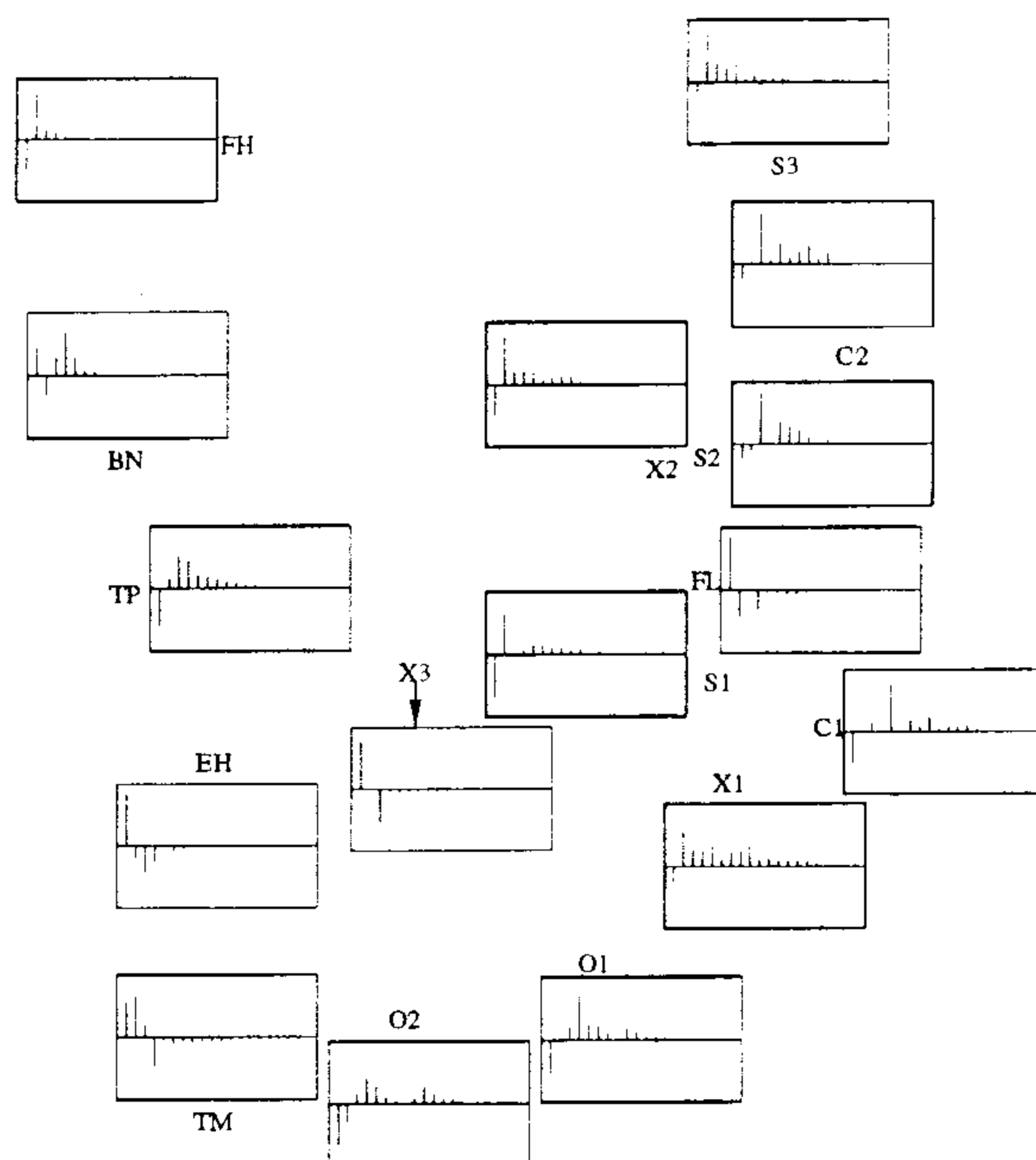
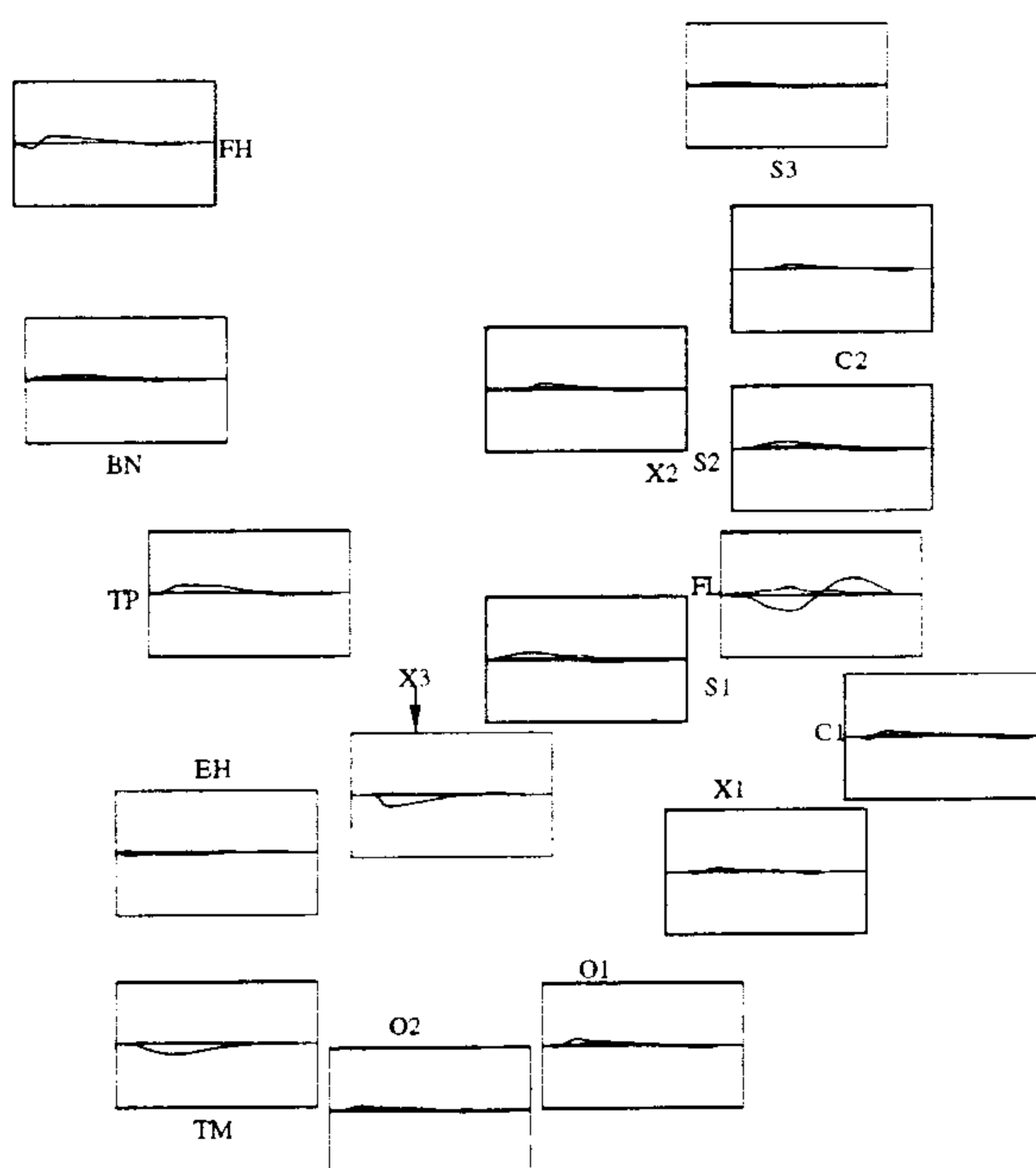
- les spectres  $P_1$  pauvres en harmoniques (FH, S3, C2),
- aux spectres  $P_1$  riches en harmoniques (O2, S1) et/ou pour lesquels le premier harmonique est faible (TM, EH).

Le spectre  $P_1$  caractérise le profil spectral moyen des sons analysés. Cette relation entre  $P_1$  et l'axe  $Y$  de l'espace de John M. Grey témoigne donc d'une relation entre cet axe et la richesse spectrale moyenne des sons.

La figure 2.19 dispose le spectre  $P_2$  selon les axes  $Z$  et  $Y$  de l'espace de John M. Grey. L'axe  $Z$ , représenté horizontalement, semble opposer les spectres  $P_2$  pauvres en harmoniques (FH, BN, EH), aux spectres  $P_2$  riches en harmoniques (C1, X1, S2). Certains sons (TP, FL, X3) semblent cependant se démarquer. La figure 2.20 représente  $V_2(t)$  selon la même disposition.  $P_2$  qualifie les déviations spectrales tandis que  $V_2(t)$  les quantifie. Il semble donc que l'axe  $Z$  soit lié à l'importance — étendue spectrale (figure 2.19) et/ou amplitude (figure 2.20) — des déviations spectrales. Cette constatation corrobore les conclusions de John M. Grey selon lesquelles ce troisième axe serait lié à la synchronicité des harmoniques lors des transitoires.

L'analyse des sons de John M. Grey nous a permis de constater des liens entre :

- l'axe  $X$  et l'enveloppe  $V_1(t)$ ,

FIG. 2.19 – Plan ZY de l'espace de Grey et spectre  $P_2$ FIG. 2.20 – Plan ZY de l'espace de Grey et enveloppe  $V_2(t)$

- l’axe Y et le spectre harmonique  $P_1$ ,
- l’axe Z et le spectre  $P_2$  et son enveloppe  $V_2(t)$ .

Les résultats de cette étude sont certes intéressants — puisqu’ils établissent des liens entre le timbre “perceptif” et les données réduites — mais nous pensons que l’étude mérite d’être étendue au moins sur deux plans.

D’une part, de nouvelles expériences psychoacoustiques, usant d’un plus large et plus riche *corpus* sonore permettraient d’affiner les résultats de John M. Grey en y ajoutant éventuellement des axes perceptifs. La recherche de relations similaires à celles décrites précédemment pourrait alors être entreprise pour les axes et enveloppes d’ordres supérieurs.

D’autre part, l’insertion de sons synthétiques obtenus par le processus de synthèse croisée décrit en 2.4.3, page 110, permettrait de renforcer le lien entre espace perceptif et données réduites. La validation d’une telle relation permettrait *in fine* d’utiliser l’espace perceptif comme base pour la synthèse des sons musicaux; la synthèse s’effectuant à partir des données réduites correspondant à un point sélectionné dans l’espace perceptif. Un tel espace deviendrait une sorte de “palette de timbres”.

### 2.3.2 Analyse et étude de quatre sons musicaux

Nous nous proposons ici d’étudier quatre sons, choisis pour la diversité de leurs caractéristiques. Cette étude nous permet d’illustrer l’utilisation des données réduites dans le cadre de l’analyse du timbre musical. Ces quatre sons, extraits du cédérom “Classical Sounds” de l’université McGill, sont :

- un son de piano (`PianomfD4`) joué mezzo forte
- un son de hautbois (`OboeD4`)
- un son de violoncelle (`CelloWvibC#4`) joué avec effet de *vibrato*
- un son de flûte (`FluteWvibD4`) joué avec effet de *trémolo*

On pourra avantageusement se référer à l’annexe B — qui présente 42 fiches de données réduites — pour étendre certaines observations à un plus large éventail de sons.

#### Représentation en “chutes d’eau”

La figure 2.21 est une représentation classique en “chute d’eau” des spectres dynamiques de ces quatre sons. L’échelle des amplitudes est linéaire. Ces figures permettent de remarquer certaines caractéristiques spectrales et dynamiques.

L’analyse du son de piano fait essentiellement ressortir une dynamique de son “percussif” : une attaque rapide, suivie d’une décroissance. Sur le plan spectral, on remarque que l’harmonique 4 a une amplitude assez faible (environ 45dB inférieure à l’amplitude du fondamental).

L’analyse du son de hautbois met en évidence une dynamique de son entretenu : une attaque lente suivie d’une période d’entretien durant laquelle le son est constant. Sur le

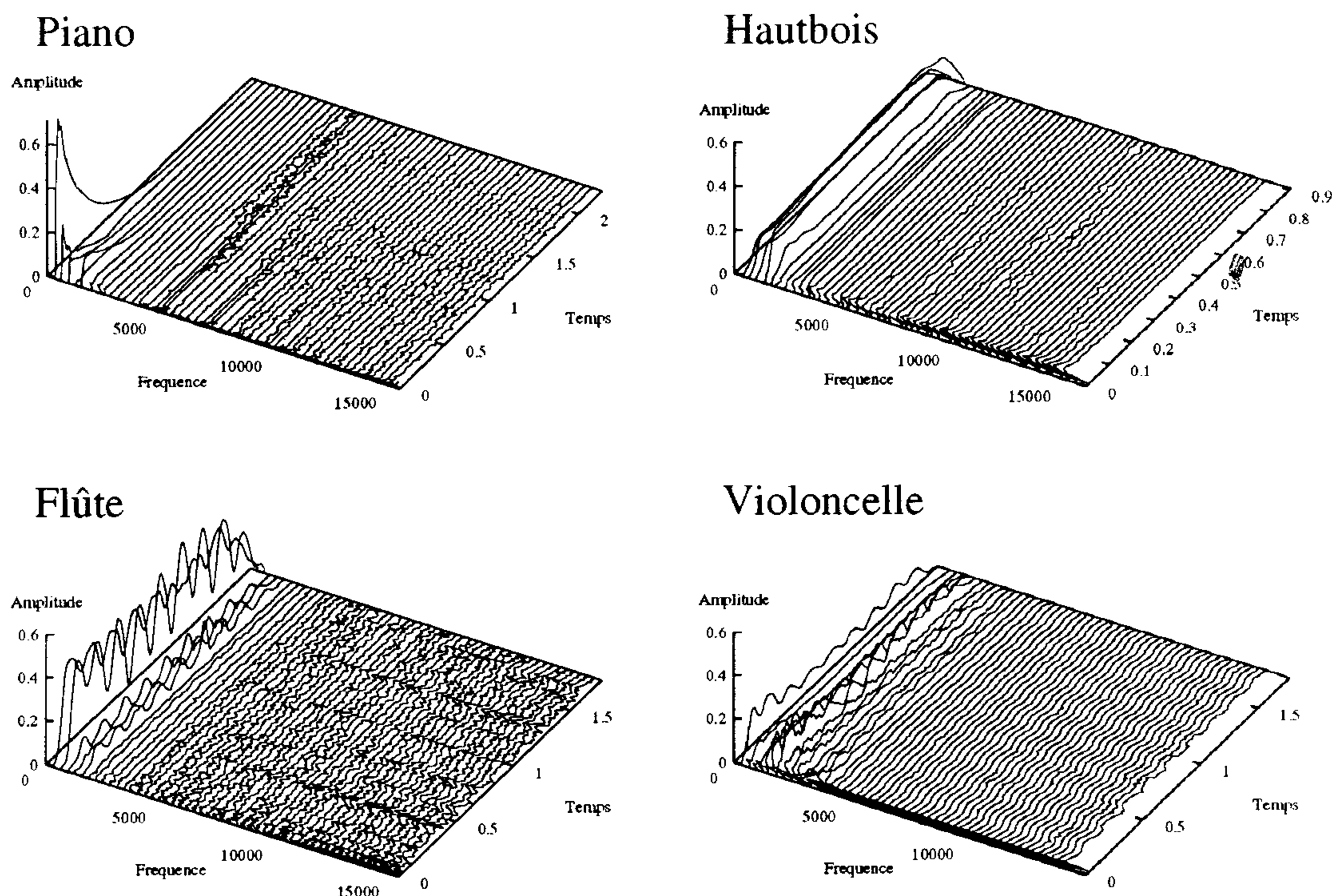


FIG. 2.21 – Représentation spectrale en chutes d'eau

plan spectral, on remarque que les 4 premiers harmoniques ont des amplitudes supérieures à l'amplitude du fondamental.

L'analyse du son de flûte révèle une dynamique de son entretenu joué avec un trémolo : une attaque lente suivie d'une période entretenu durant laquelle l'amplitude du son varie de façon périodique. Sur le plan spectral, on remarque que les 2 premiers harmoniques ont des amplitudes importantes et comparables.

L'analyse du son de violoncelle fait ressortir une dynamique de son entretenu joué avec un vibrato. Sur le plan spectral, on remarque que tous les harmoniques varient en fréquence et en amplitude en fonction du vibrato. Cependant, ces variations sont complexes : tous les harmoniques ne semblent pas varier de la même façon.

Bien quelle permette d'avoir une vue globale de chaque son, la représentation en "chute d'eau" ne permet pas d'analyser finement l'évolution dynamique du spectre. Même avec de l'expérience, il reste difficile de remarquer certains liens qui peuvent unir une cinquantaine d'harmoniques.

Les figures 2.22, 2.23, 2.24 et 2.25 représentent les données spectrales réduites des quatre sons ainsi qu'une représentation de la fréquence fondamentale et qu'une représentation en "chute d'eau". Les trois premiers spectres principaux  $P_1$ ,  $P_2$  et  $P_3$ , sont représentés sur la gauche tandis que leurs enveloppes  $V_1(t)$ ,  $V_2(t)$  et  $V_3(t)$ , sont représentées sur la droite. Les trois spectres correspondent aux axes du sous-espace principal  $\mathcal{P}$ ,



tandis que les trois enveloppes correspondent à la trajectoire  $\mathcal{V}$ . Les courbes en pointillé représentées sur les graphes d'enveloppes correspondent à l'erreur résiduelle. C'est l'erreur induite par la réduction de données lorsqu'elle est limitée aux axes précédents.

Nous allons commenter chacune de ces représentations.

### Piano (figure 2.22)

Le spectre harmonique  $P_1$  du son de piano correspond au spectre moyen<sup>3</sup>, il nous permet de remarquer la faible amplitude de l'harmonique 4. Son enveloppe  $V_1(t)$  caractérise la dynamique globale d'un son "percussif". Le spectre harmonique  $P_2$  qui correspond aux variations principales du spectre, fait ressortir une opposition entre le fondamental et le second harmonique. Son enveloppe  $V_2(t)$  montre que le second harmonique décroît plus rapidement que le fondamental. Ceci met en avant la principale variation spectrale de ce son. Notons que ce détail était assez difficile à remarquer avec la représentation en "chute d'eau". Le spectre harmonique  $P_3$  correspond aux variations spectrales d'ordre supérieur, son rôle est ici difficile à définir mais sa faible amplitude nous laisse croire qu'il est négligeable.

### Hautbois (figure 2.23)

Le spectre harmonique  $P_1$  du son de hautbois révèle l'importance relative des 4 premiers harmoniques. Son enveloppe  $V_1(t)$  est bien caractéristique d'un son entretenu. On remarque que les enveloppes  $V_1(t)$  et  $V_2(t)$  sont faibles, à l'exception de légères fluctuations durant l'attaque. Ce fait caractérise un son dont le contenu spectral est relativement constant. Par ce résultat trivial, l'ACP nous informe de la stabilité du profil spectral de ce son.

### Flûte (figure 2.24)

Le spectre harmonique  $P_1$  du son de flûte révèle l'importance relative des 2 premiers harmoniques. Son enveloppe  $V_1(t)$  est bien caractéristique d'un son joué trémolo. On remarque que le trémolo engendre une variation d'amplitude et une variation de fréquence. Le spectre harmonique  $P_2$  marque une opposition entre les deux premiers harmoniques. Son enveloppe varie  $V_2(t)$  aussi avec le trémolo. Cette variation spectrale peut correspondre aux modes de la flûte, l'oscillation "octaviant" plus ou moins selon l'amplitude de l'excitation. Le rôle du troisième spectre harmonique est moins explicite et semble négligeable (au regard de l'amplitude du signal d'erreur).

### Violoncelle (figure 2.25)

Le spectre harmonique  $P_1$  du son de violoncelle met en avant la richesse spectrale du son de violoncelle, on remarque une "résonance" autour des harmoniques 5 et 6. Son enveloppe  $V_1(t)$  n'est pas tout à fait constante mais ces fluctuations ne semblent pas périodiques. Le spectre harmonique  $P_2$  présente une forme assez complexe. Il oppose, entre autres, les harmoniques 5 et 6. Son enveloppe  $V_2$  est périodique et presque régulière.

---

3. Ou plus exactement au spectre correspondant au spectre d'énergie moyenne.

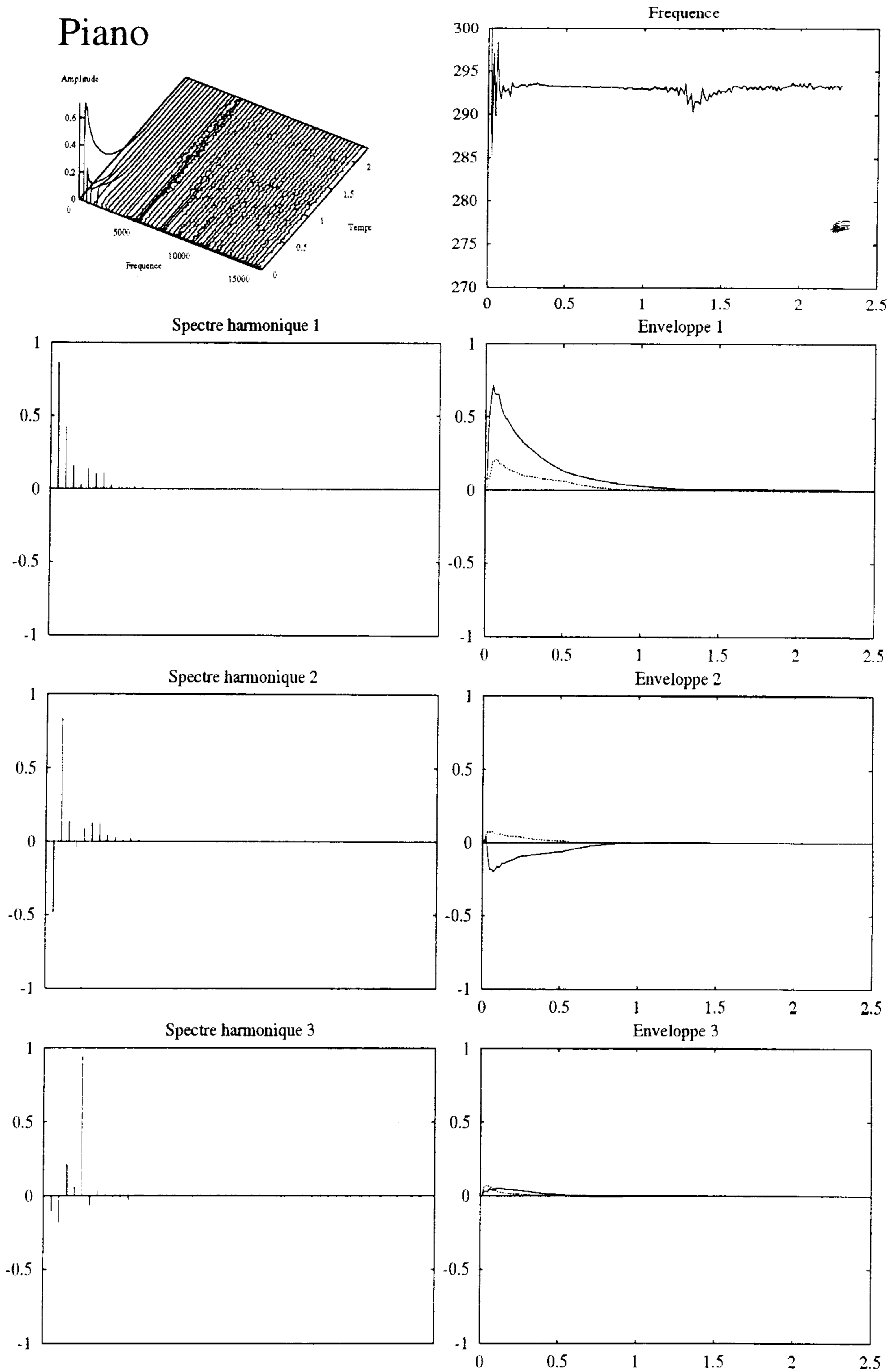


FIG. 2.22 – Données spectrales réduites d'un son de piano

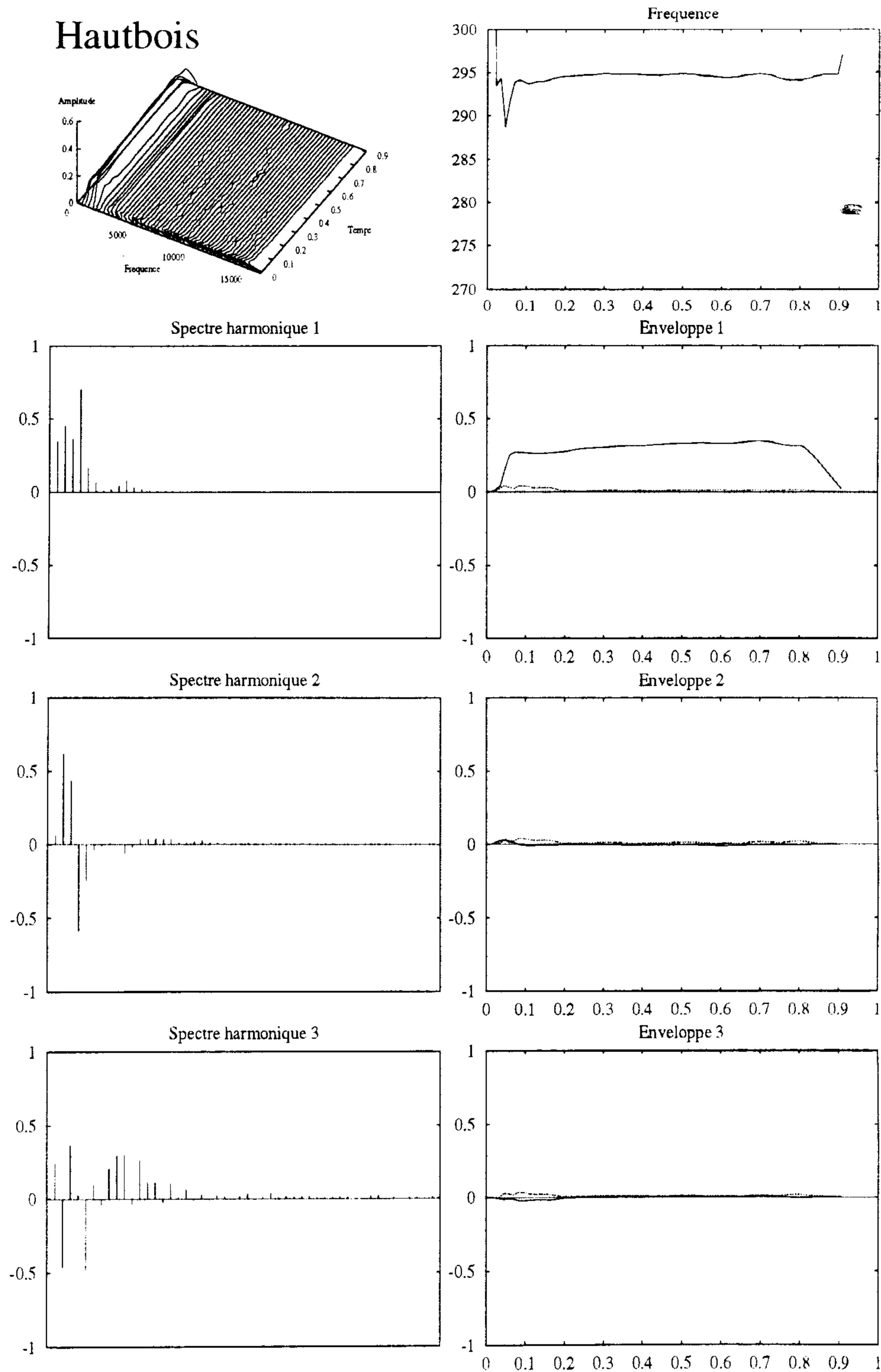


FIG. 2.23 – Données spectrales réduites d'un son de hautbois

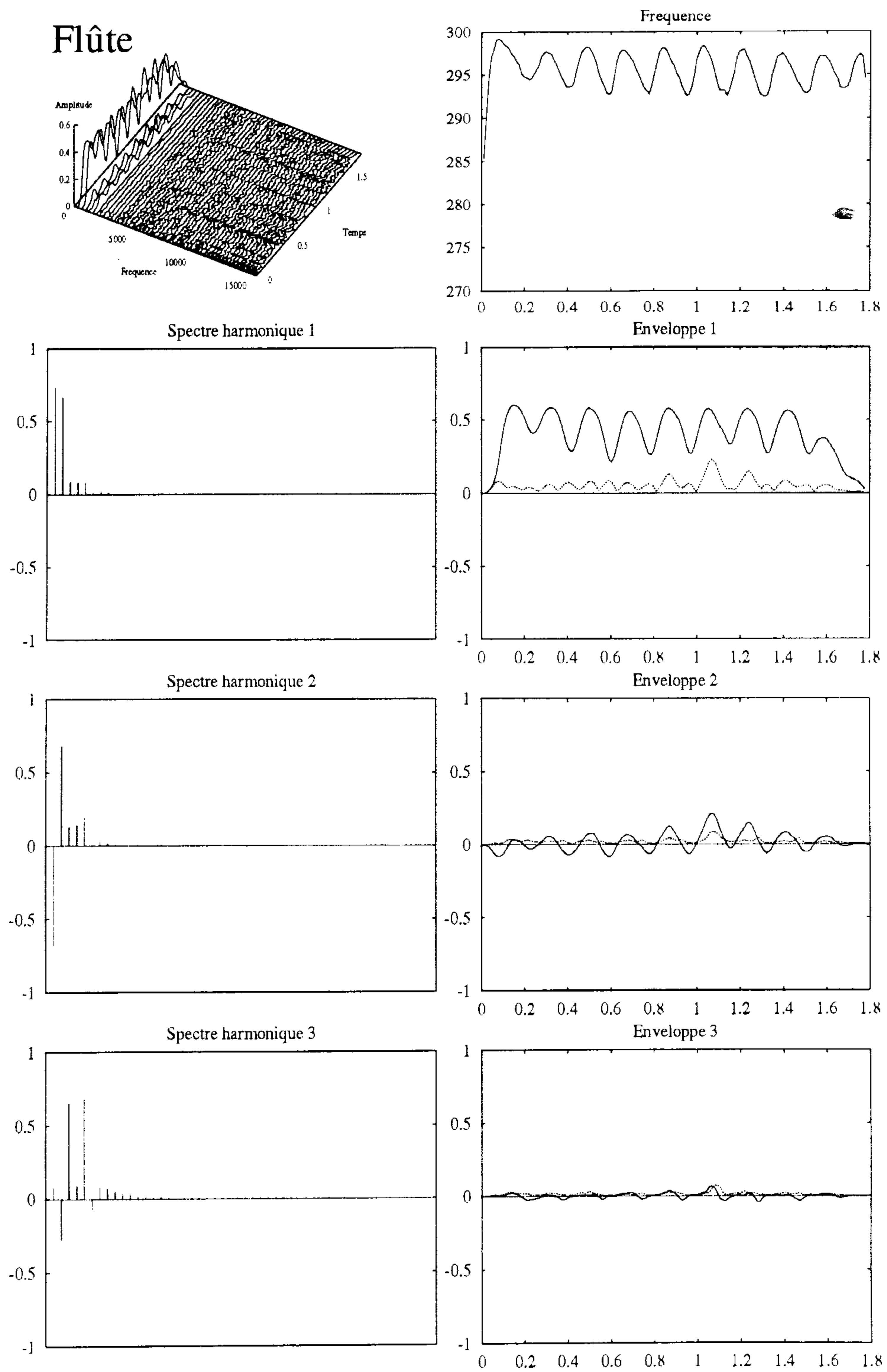


FIG. 2.24 – Données spectrales réduites d'un son de flûte

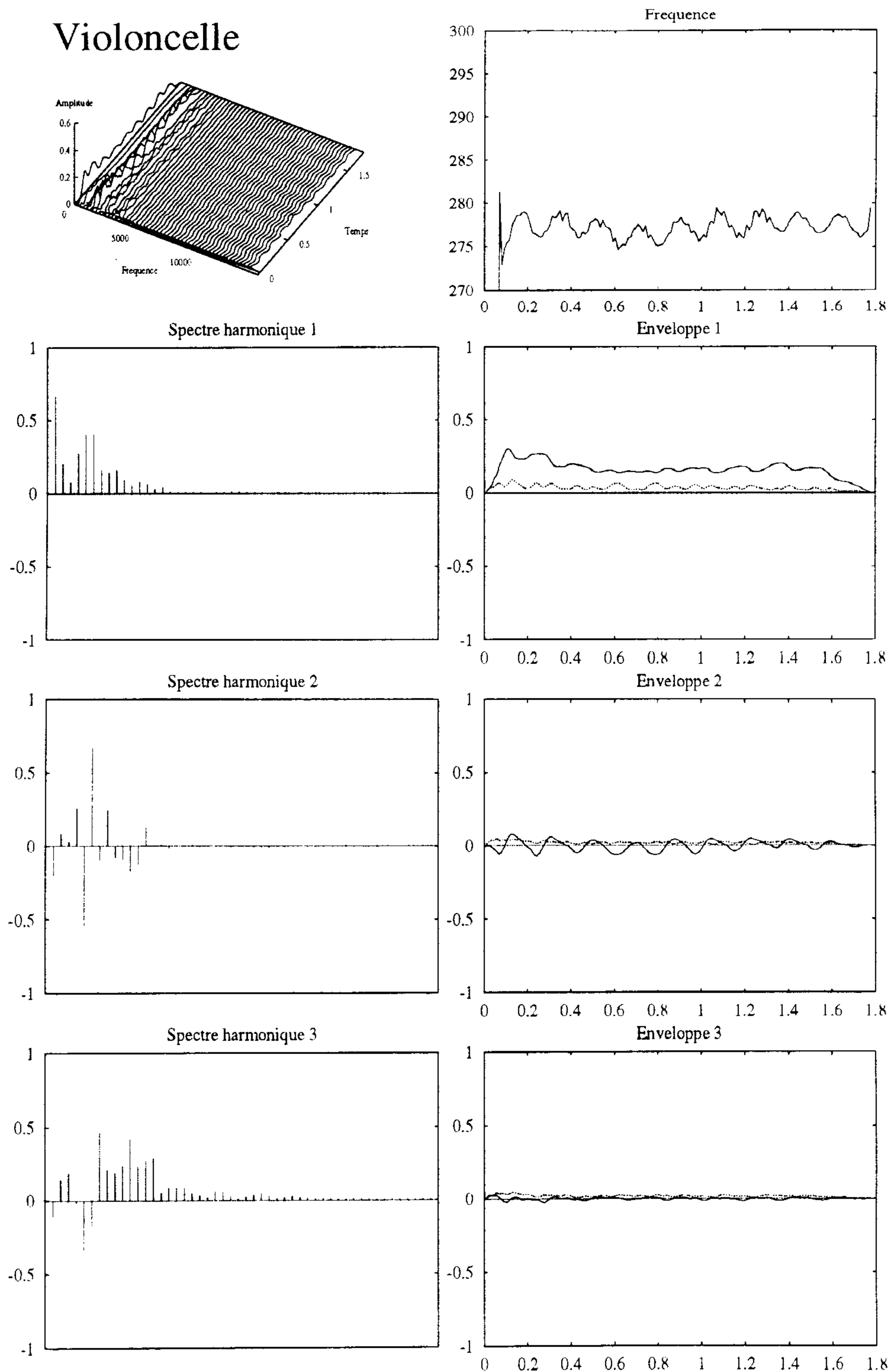


FIG. 2.25 – Données spectrales réduites d'un son de violoncelle

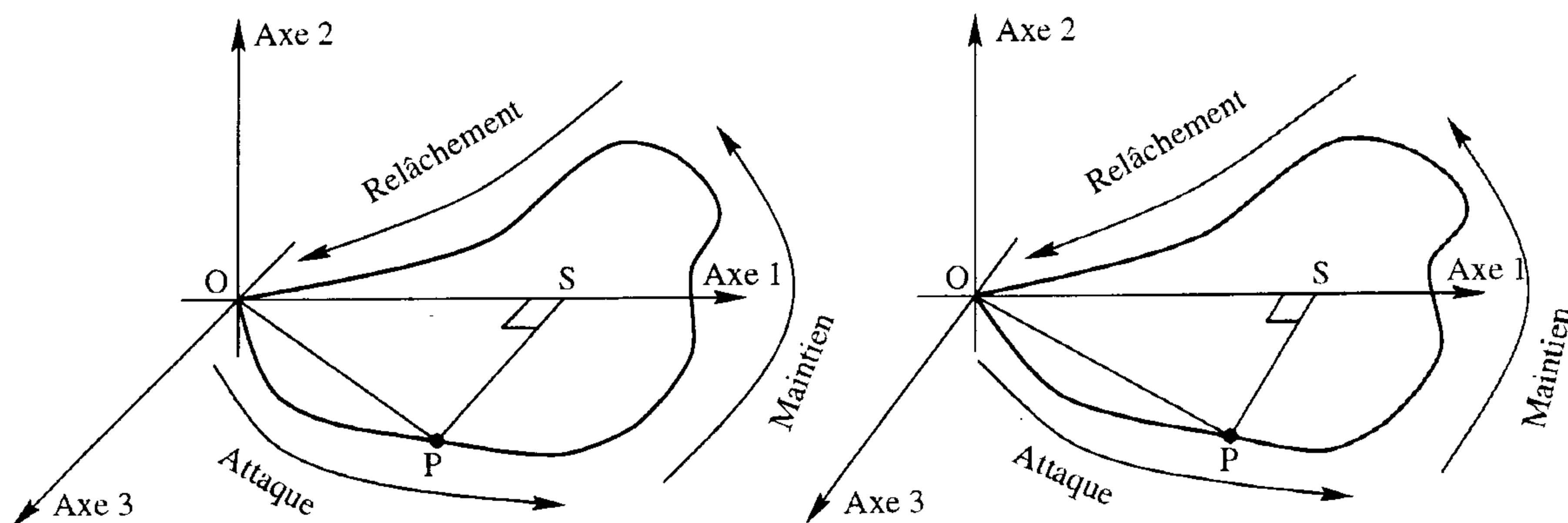


FIG. 2.26 – Paire stéréographique de la trajectoire de la figure 2.9

On peut déduire de ces observations que le vibrato engendre une modulation “spectrale” plutôt qu’une modulation d’amplitude globale. Ce phénomène peut en partie s’expliquer par les phénomènes de résonance des plaques de la caisse de l’instrument qui établissent des liens complexes entre fréquence fondamentale et amplitudes des harmoniques. Constatons aussi la faible diminution du signal d’erreur apportée par  $P_3$ .

## Conclusions

L’interprétation des spectres harmoniques  $P_1$  et de leurs enveloppes nous permet de faire les mêmes observations que la représentation en “chute d’eau”. L’interprétation des spectres harmoniques  $P_2$  et  $P_3$  permettent d’affiner ces observations en faisant ressortir des évolutions spectrales. Nous allons maintenant revenir à la représentation du son sous la forme de la trajectoire  $\mathcal{V}$ . Cette représentation alternative des enveloppes nous permettra d’observer les liens qui les unissent.

### 2.3.3 Représentation stéréoscopique des trajectoires

La représentation à deux dimensions d’une trajectoire tridimensionnelle est difficile. Dans la suite de cette section, nous utiliserons une représentation stéréoscopique des données. Ce type de représentation est, par exemple, largement employé en chimie organique pour représenter des structures moléculaires tridimensionnelles complexes.

Le principe de la stéréovision repose sur le fait que la perception tridimensionnelle est en partie liée à la disparité entre les images perçues par l’œil droit et par l’œil gauche. Une paire stéréographique est constituée de deux images, l’une destinée à l’œil gauche, l’autre destinée à l’œil droit.

Une paire stéréographique correspondant à la trajectoire de la figure 2.9b est donnée en figure 2.26. Pour percevoir l’image stéréoscopique correspondant à cette paire d’images, il faut placer la figure à une cinquantaine de centimètres des yeux et loucher légèrement de façon à voir trois images (figure 2.27). L’image centrale, qui correspond en fait à la superposition des deux images, est une image binoculaire; elle sera perçue en trois dimensions.

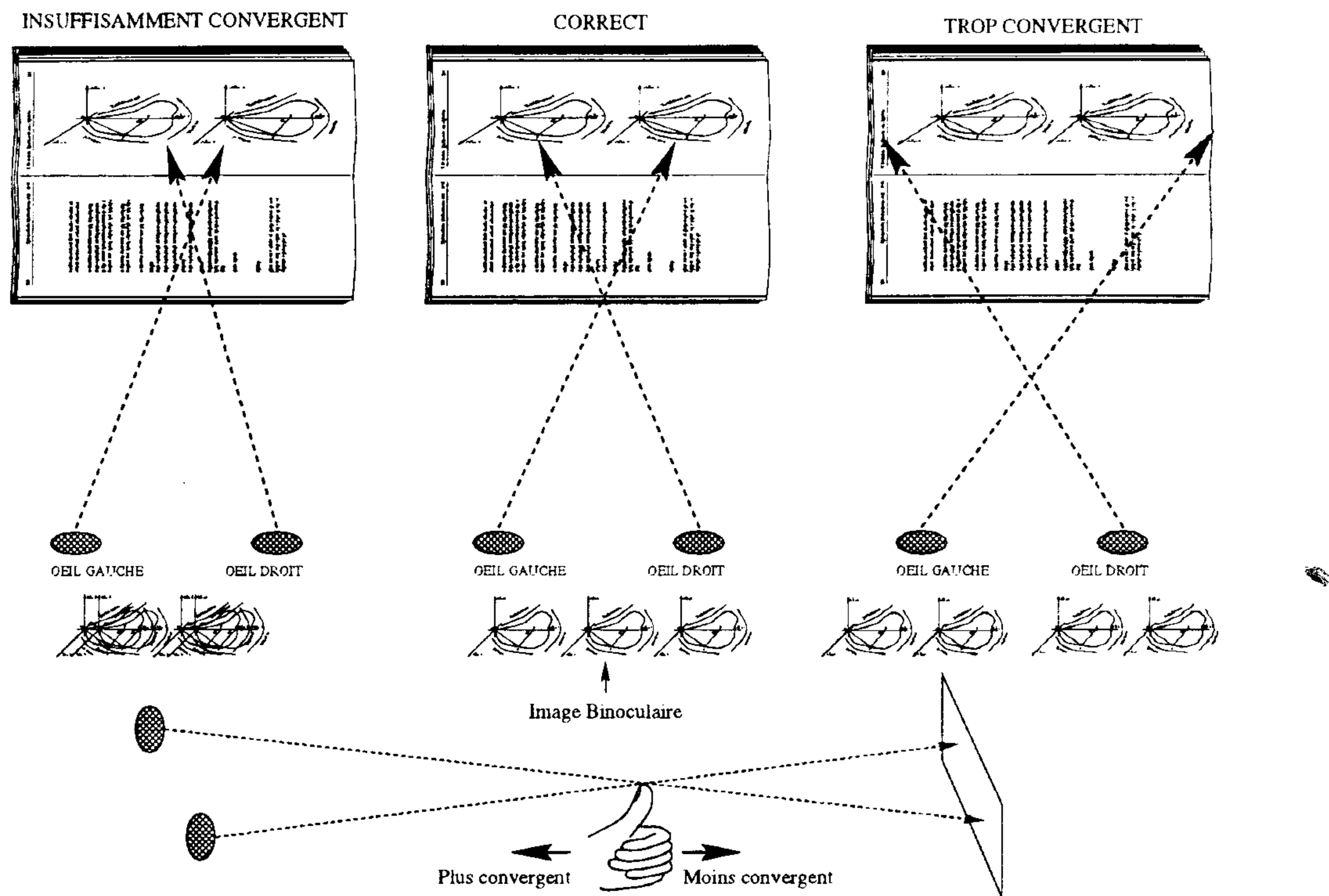


FIG. 2.27 – Visualisation de la paire stéréographique

Il existe plusieurs procédures pour obtenir la fusion des images. La plus simple consiste à poser un doigt en dessous et au milieu des deux images. Il faut ensuite fixer le regard sur l'extrémité du doigt, et lentement éloigner le doigt de la feuille. Les deux images se dédoublent alors et deviennent floues. A une certaine distance, on perçoit trois images floues. Dans ces conditions, il faut fixer son attention sur l'image centrale jusqu'à ce qu'elle devienne nette, ceci sans déconverger les yeux. Il suffit en général de quelques tentatives pour obtenir un résultat probant.

Les figures stéréographiques 2.28, 2.29, 2.30 et 2.31 représentent les trajectoires harmoniques  $\mathcal{V}$  de chaque son. Pour améliorer la visualisation des déviations spectrales, nous avons dilaté les échelles sur les axes  $P_2$  et  $P_3$ .

### Piano (figure 2.28)

L'axe  $P_2$  est dilaté par un facteur 1,5 et l'axe  $P_3$  par un facteur 6. On remarque que le spectre dévie assez peu : les points de la trajectoire sont presque alignés. Les variations les plus importantes se situent pendant l'attaque et vers la fin du relâchement.

### Hautbois (figure 2.29)

Les axes  $P_2$  et  $P_3$  sont fortement dilatés (facteur 4) car les fluctuations spectrales du hautbois sont très faibles. On constate que la plus importante variation a lieu pendant l'attaque. Pendant le maintien, le spectre reste constant, on observe une "grappe" de points localisés dans la même région de l'espace.

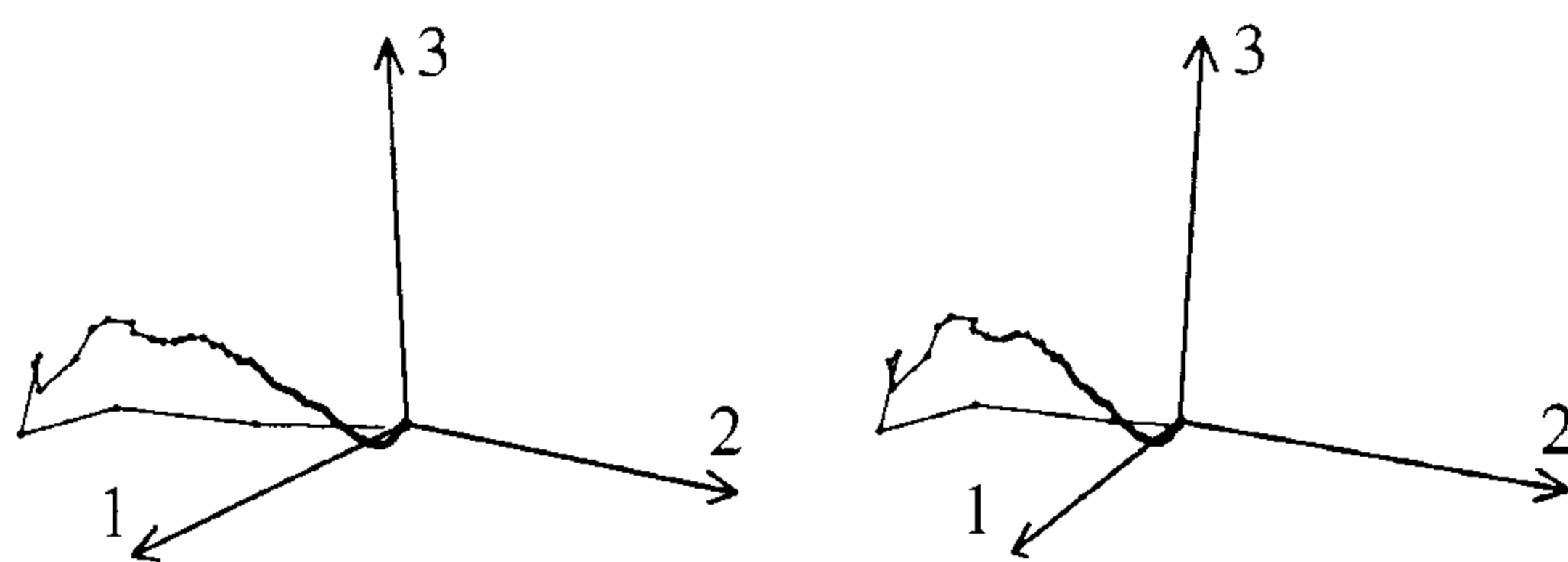


FIG. 2.28 – Représentation stéréographique de la trajectoire du piano

### Flûte (figure 2.30)

L'axe  $P_2$  est dilaté par un facteur 1,5 et l'axe  $P_3$  par un facteur 3. La trajectoire de ce son est bien plus complexe que celle des instruments précédents. Cette complexité est engendrée par le trémolo qui provoque des variations spectrales pendant le maintien. Le maintien n'engendre pas une "grappe" de points localisés mais plutôt une "pelote" plus ou moins compliquée. La forme de cette "pelote" correspond à la nature des variations spectrales. Ici, sa forme montre une dépendance entre la déviation spectrale et l'amplitude (c'est-à-dire la distance par rapport au point O). On remarque aussi un certain aspect "aléatoire".

### Violoncelle (figure 2.31)

L'axe  $P_2$  est dilaté par un facteur 1,5 et l'axe  $P_3$  par un facteur 3. La trajectoire de ce son présente des similitudes avec la précédente. La différence essentielle tient à la forme de la "pelote" : il semble qu'elle constitue une boucle ayant une forme assez bien définie. De plus, cette boucle est à une distance presque constante du point O. Bien que complexes, les déviations spectrales induites par le vibrato semblent donc avoir un caractère bien "déterministe" et indépendant de l'amplitude du son.

### Conclusion

La représentation des sons, sous forme de trajectoires, nous a permis d'observer les liens complexes qui unissent parfois les axes des sous-espaces lors de modulations d'amplitude ou de fréquence. La forme des trajectoires semble intimement liée à la dynamique spectrale des sons analysés.



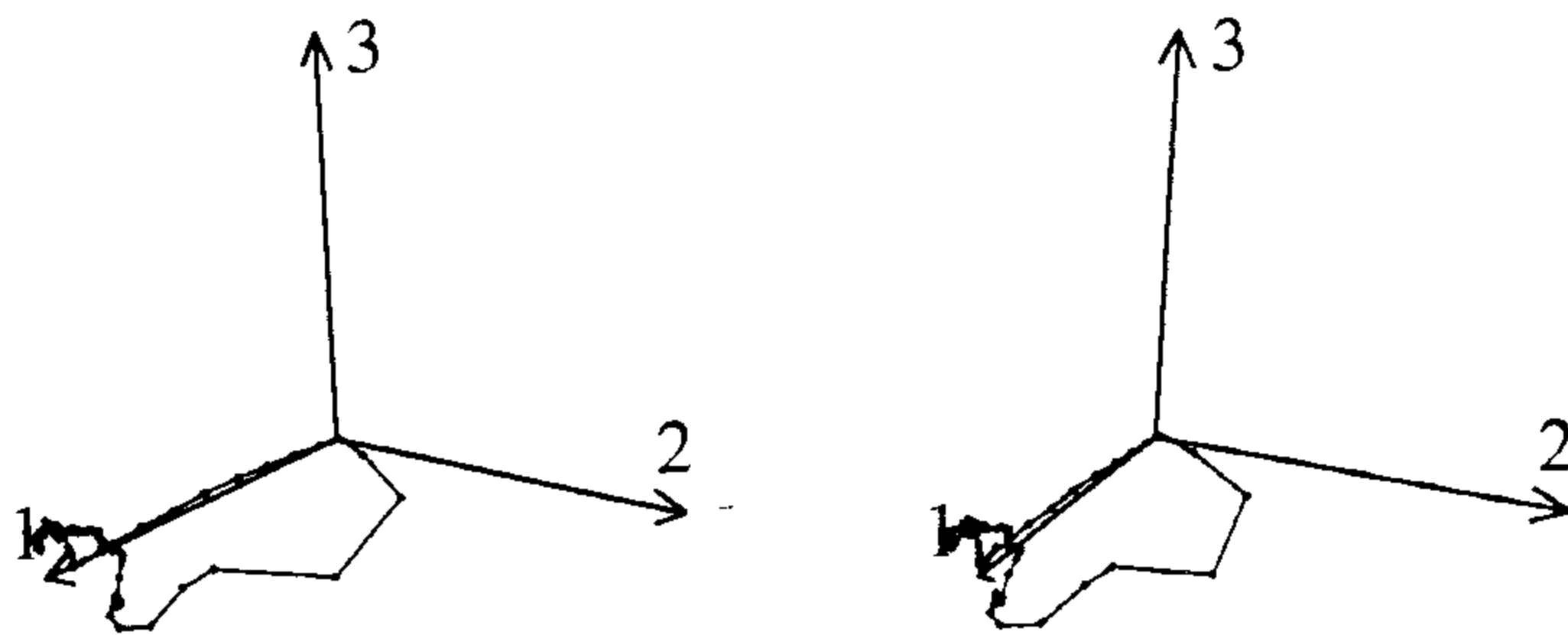


FIG. 2.29 – Représentation stéréographique de la trajectoire du hautbois

Cependant, notre démarche est restée volontairement descriptive. En effet, une démarche explicative nécessiterait une étude conjointe de l'acoustique des instruments ayant engendré les sons analysés.

### 2.3.4 Etude d'un corpus de 42 sons d'instruments acoustiques

La collection de fiches de données réduites de l'annexe B — et en particulier l'expression des spectres  $P_1$ ,  $P_2$ , et  $P_3$ , qui décrivent le sous-espace propre à chaque son — semble indiquer quelques similitudes entre certains sons. Nous allons tenter ici de caractériser plusieurs tendances et de regrouper certains sons.

#### Sous-espace commun au corpus des 42 sons

Afin de connaître la tendance générale de l'ensemble des 42 sons de l'annexe B, nous avons appliqué notre procédé d'analyse de la façon suivante :

- nous avons tout d'abord construit une matrice de covariance commune à tous les profils spectraux des 42 sons,
- dans un second temps, nous avons diagonalisé cette matrice de covariance.

Le résultat de cette analyse est une base de spectres  $P_1^{42}$ ,  $P_2^{42}$ , ...,  $P_H^{42}$  correspondant aux vecteurs propres de la matrice de covariance. La nature de ces spectres est similaire à celle des spectres des données réduites. Leur différence réside dans leur généralité. Les spectres  $P_1^{42}$ ,  $P_2^{42}$ , ...,  $P_H^{42}$  doivent qualifier les comportements spectraux de l'ensemble des 42 sons.

Les cinq premiers spectres sont représentés sur la figure 2.32. Le spectre  $P_1^{42}$  correspondant au spectre moyen des sons a une forme remarquable, il est, à peu de choses près,

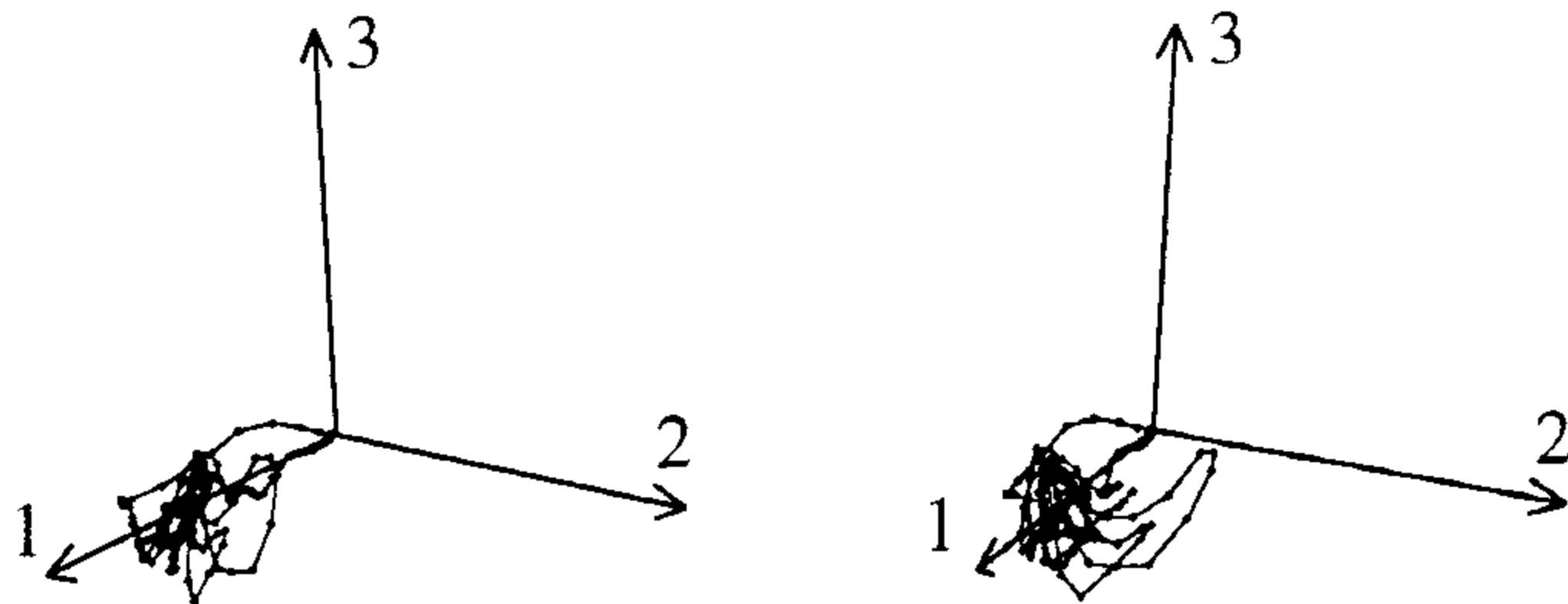


FIG. 2.30 – Représentation stéréographique de la trajectoire de la flûte

identique au spectre d'un signal en dents de scie. Le spectre  $P_2^{42}$  oppose l'harmonique 1 aux autres harmoniques. Les spectres  $P_n^{42}$  suivants semblent opposer les harmoniques de rangs  $n - 1$  aux harmoniques de rangs supérieurs. Il semble possible de revenir à la base des harmoniques en effectuant une soustraction du type :  $P_{n+1}^{42} - \sum_{k=1}^n P_k^{42} \approx H_n$ . Le passage des données "réduites" aux données harmoniques originales est donc assez trivial.

Notre méthode, appliquée à un large *corpus*, nous a fourni une base aussi générale que la base initiale. Ce fait démontre une chose importante concernant le *corpus* des sons analysés : il couvre une "grande partie" de l'espace des profils spectraux et interdit une réduction globale des données. En des termes plus clairs, le *corpus* de sons d'instruments acoustiques que nous avons choisi couvre une large diversité de profils spectraux.

### Classification des sons selon $P_2$

Rappelons que ce spectre caractérise la principale "direction" de variation spectrale et que les amplitudes des harmoniques le décrivant sont signées. L'observation des fiches de données réduites de l'annexe B, permet d'observer plusieurs tendances ayant trait au spectre  $P_2$ . Nous avons séparé les sons en trois classes selon différents profils de  $P_2$  :

- Classe 1 :  $P_2$  présente une opposition de signe entre l'harmonique 1 et les autres harmoniques.
- Classe 2 :  $P_2$  présente des groupes d'harmoniques contigus en opposition de signe.
- Classe 3 :  $P_2$  présente des harmoniques isolés en opposition de signe.

D'après cette classification, la classe 1 semble regrouper les sons de cordes frappées, de bois et de certains cuivres. Pour cet ensemble de sons,  $P_2$  oppose le premier harmonique aux autres. En d'autres termes, ces sons sont caractérisés par une variation de l'amplitude des

Son	Page
Clarinette en si $\flat$	153
Clarinette en mi $\flat$	154
Orgue baroque registre prinzipal	158
Orgue baroque registre flûtes	159
Flûte traversière avec vibrato	162
Flûte alto sans vibrato	163
Flûte jouée "flutter"	165
Basson	166
Trombone	171
Trombone avec sourdine	172
Tuba	175
Celesta	176
Marimba	177
Piano joué mezzo forte	178
Piano joué fortissimo	179
Guitare classique jouée sul tasto	181
Harpe	183
Violon	185
Violoncelle joué martele	186

TAB. 2.1 – *Sons de classe 1*

Son	Page
Hautbois	155
Hautbois baroque	156
Flûte à bec basse	160
Flûte alto avec vibrato	164
Cor Anglais	167
Cor	168
Cor avec sourdine	169
Cornet	170
Trompette baroque	173
Trompette en Do	174
Guitare classique	180
Alto joué martele	187

TAB. 2.2 – *Sons de classe 2*

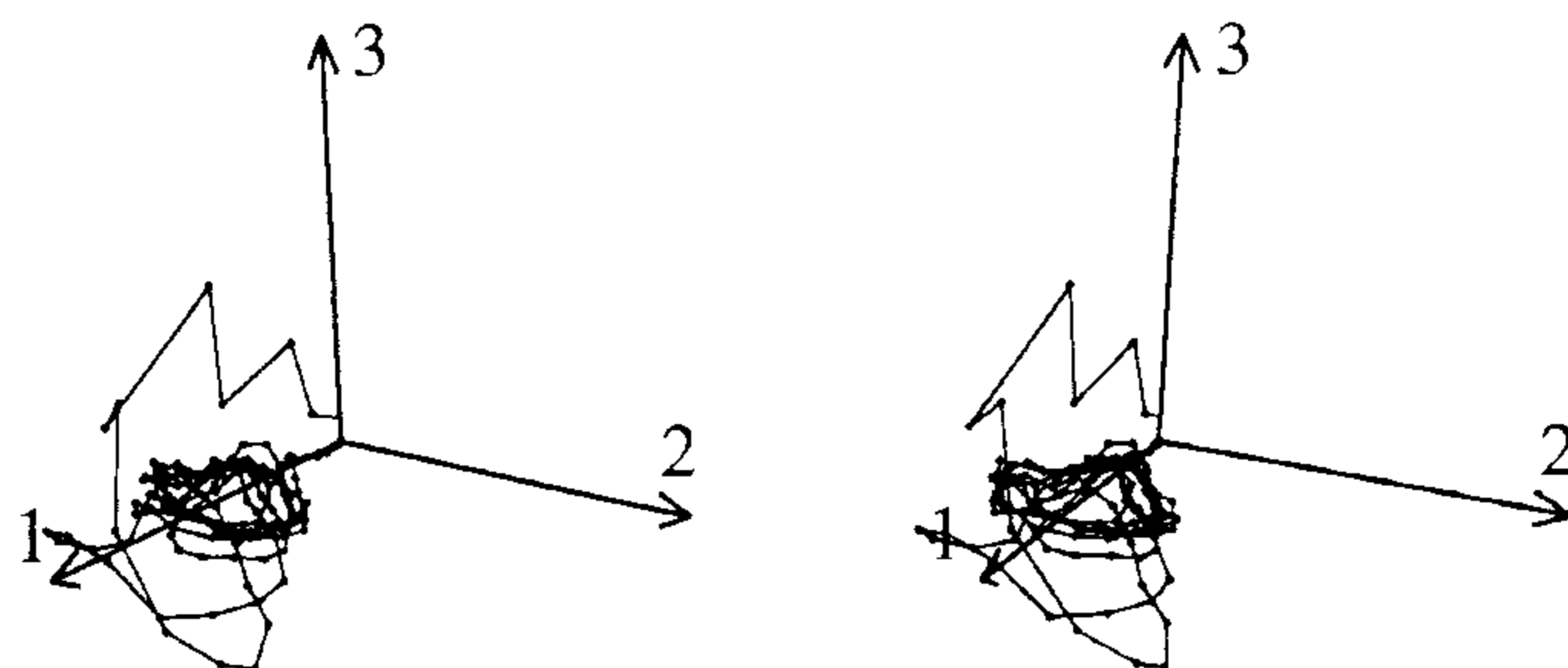
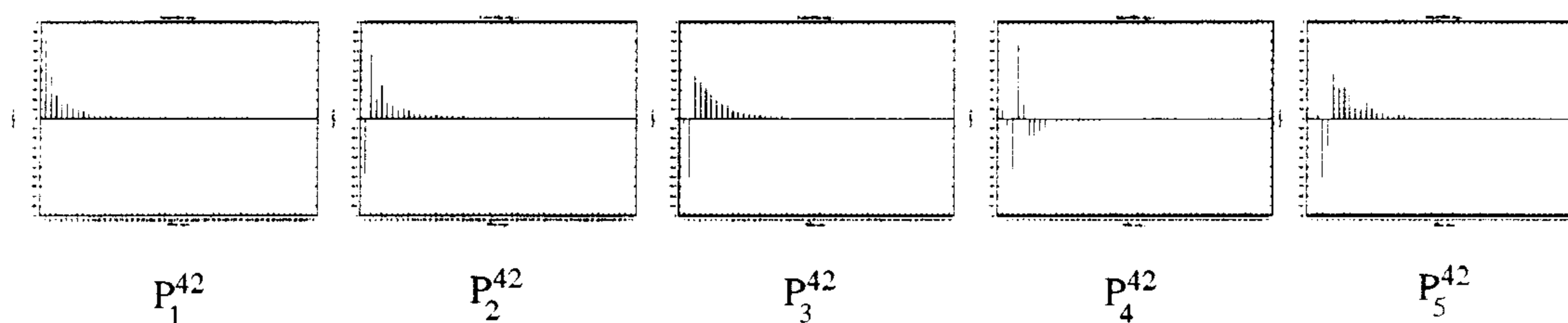


FIG. 2.31 – Représentation stéréographique de la trajectoire du violoncelle

FIG. 2.32 – Spectres  $P_1^{42}$  à  $P_5^{42}$ 

harmoniques au détriment de l'amplitude de l'harmonique 1 c'est-à-dire du fondamental. Ce fait peut être interprété physiquement de différentes façons :

- le mode fondamental est rayonné de façon efficace et son amplitude décroît plus rapidement que celles des autres modes (cordes),
- une part de l'énergie du mode fondamental est transmise aux autres modes (cordes),
- les variations d'intensité de la source d'excitation favorisent ou non le mode fondamental (vents).

La classe 2 semble regrouper les cuivres, certains bois et certaines cordes frottées ou pincées. Ici,  $P_2$  oppose des groupes d'harmoniques contigus. En d'autres termes, l'enrichissement d'une partie du spectre est accompagné d'un appauvrissement d'une autre partie. Ce phénomène est connu chez les cuivres depuis longtemps et a fait l'objet de multiples démonstrations par Max V. Mathews et Jean-Claude Risset [RM69]. Leur interprétation de ce phénomène est la suivante : l'instrument peut être, en première approximation, assimilé à un guide d'ondes excité par une source d'énergie non-linéaire ; cette non-linéarité entretient une oscillation dans le guide d'onde dont certains modes sont

Son	Page
Orgue baroque registre plenum	157
Flûte traversière sans vibrato	161
Clavecin	182
Alto	184
Violon joué martelé	188
Violoncelle joué avec vibrato	189
Alto joué avec vibrato	190
Violon joué avec vibrato	191
Violoncelle joué pizzicato	192

TAB. 2.3 – Sons de classe 3

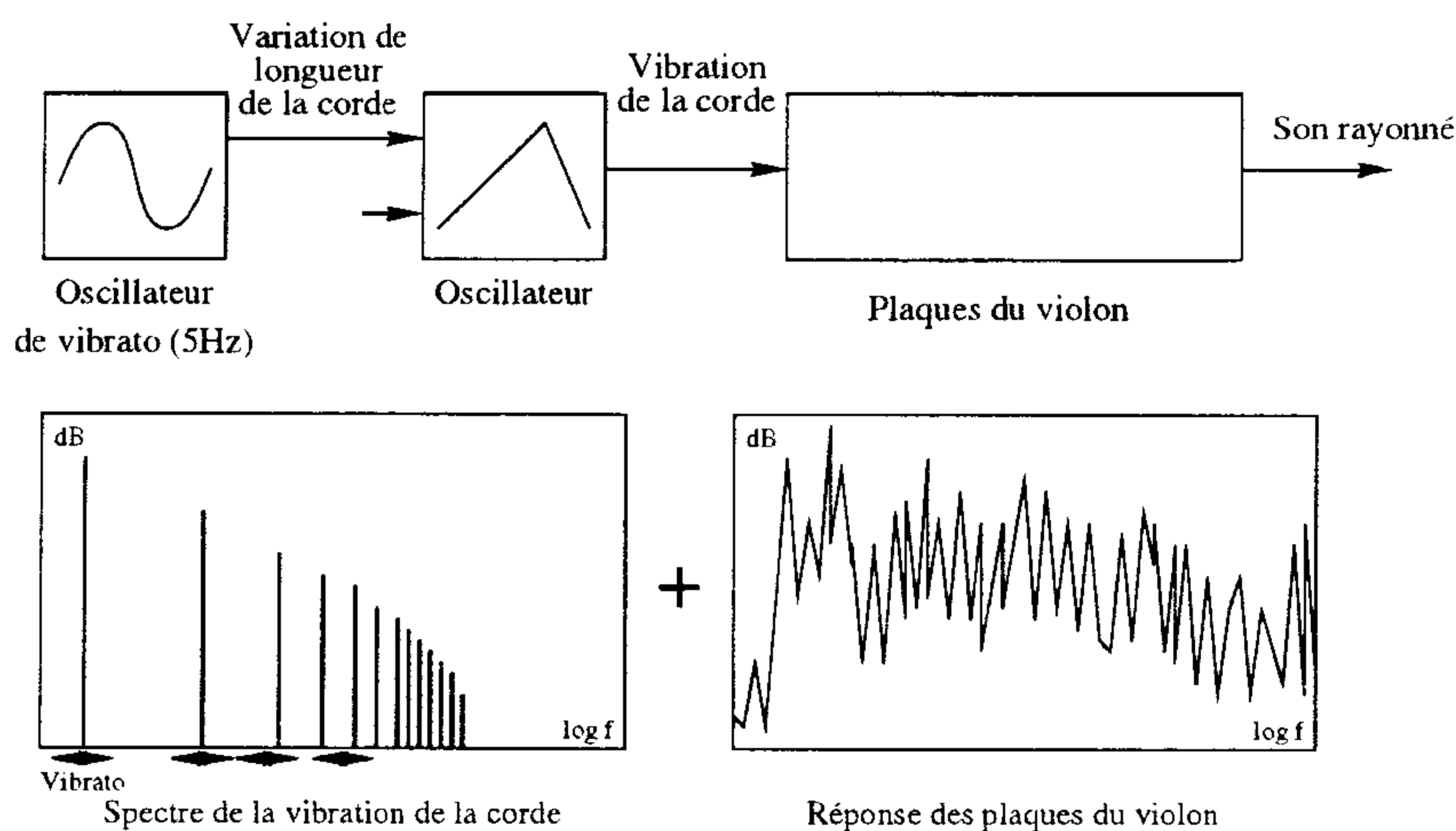


FIG. 2.33 – Modèle simpliste d'instrument à cordes frottées

rayonnés ou réfléchis par le pavillon ; la nature de la non-linéarité et de son couplage avec le guide d'onde varie en fonction de l'intensité de l'excitation ; cette variation conduit à une variation du contenu harmonique du son tendant à favoriser les modes aigus pour les sons de forte intensité.

La classe 3, quant à elle, tend à regrouper les sons d'instruments à cordes frottées et en particulier ceux joués avec des effets de *vibrato*. Notons aussi que les courbes d'amplitude du signal d'erreur — représentées en pointillés conjointement aux enveloppes — demeurent importantes même au delà de  $P_3$ . Il semble donc que notre méthode d'analyse ne soit pas appropriée à cette classe de signaux.

Recherchons maintenant les causes de ce problème dans les caractéristiques particulières de ces instruments. La corde frottée peut être grossièrement assimilée à un générateur de signal triangulaire dont la symétrie varie selon la position et l'adhérence de l'archet. Un tel signal est très riche en harmoniques. Autre élément, l'usage d'effets de *vibrato* et autres *glissando* induisant une variation continue de la fréquence fondamentale sont très usités. Dernier élément de notre modèle simpliste, le signal triangulaire est "filtré" et rayonné par les deux plaques de la caisse de l'instrument. La nature du filtrage opéré par

ces plaques aux formes irrégulières est fort compliquée en termes d'amplitudes mais aussi de phases. La figure 2.33 récapitule notre modèle : un signal riche en harmoniques, dont la fréquence varie, est filtré de façon très élaborée. Les variations d'amplitude des harmoniques correspondent à la réponse du filtre, ces variations sont imposées par la variation de la fréquence fondamentale. En conséquence, le spectre du signal de "sortie" varie de façon déterministe mais compliquée en fonction de la fréquence du son (notons que des observations de ce type ont pu être formulées en 2.3.3, page 95, grâce à la visualisation par paires stéréographiques). Les variations d'amplitudes des harmoniques sont liées entre elles, de façon non linéaire, par l'intermédiaire de la réponse du filtre et des variations de la fréquence fondamentale. L'ACP ne permet que d'extraire des redondances "linéaires", elle est donc inapte à extraire les relations non linéaires qui unissent les harmoniques de cette classe de sons.

### Etude des relations non-linéaires entre $V_1(t)$ , $V_2(t)$ , $V_3(t)$ et $f_1(t)$

Les figures 2.34 et 2.35 représentent respectivement les courbes paramétriques  $V_1(t)$ ,  $V_2(t)$  et  $V_1(t)$ ,  $V_3(t)$  des 42 sons analysés.  $V_1(t)$  étant essentiellement lié à l'amplitude du son, ces courbes paramétriques font ressortir des relations entre amplitude et déviations spectrales. Ces relations, lorsqu'elles existent, sont non-linéaires grâce aux propriétés intrinsèques de l'analyse en composantes principales.

Les instruments à cordes frappées et grattées (piano, guitare, clavecin, harpe) ainsi que certains instruments à vent (cors, trombone, trompette) présentent des relations assez similaires entre  $V_1(t)$  et  $V_2(t)$  :  $V_2(t)$  semble lié à  $V_1(t)$  par une relation non-linéaire relativement régulière en forme d'arc. Les relations entre  $V_1(t)$  et  $V_3(t)$  sont parfois plus complexes, surtout lorsque l'amplitude de  $V_3(t)$  est faible relativement à  $V_1(t)$  et  $V_2(t)$ . Pour ces instruments, il semble donc qu'une relation non-linéaire lie certains paramètres de déviation spectrale à l'amplitude du son.

Les instruments joués avec des effets de *vibrato* ou de trémolo font ressortir des relations plus complexes que les précédentes.

La flûte jouée avec un effet de *vibrato* présente une relation  $V_1(t)$ ,  $V_2(t)$  moyenne en forme d'arc. La "trajectoire"  $V_1(t)$ ,  $V_2(t)$  semble "tourner autour" de cette relation moyenne. Cette observation peut être interprétée comme un "accrochage" sur divers modes de résonance de la flûte présentant un phénomène d'hystérésis. La trajectoire  $V_1(t)$ ,  $V_3(t)$  semble avoir un caractère plus aléatoire, la faible amplitude de  $V_3(t)$  peut nous laisser penser que  $P_3$  correspondrait à une modulation d'amplitude par un signal bruité.

Les figures 2.36 et 2.37 viennent compléter les courbes précédentes en présentant respectivement les courbes paramétriques  $f_1(t)$ ,  $V_1(t)$  et  $f_1(t)$ ,  $V_2(t)$  des 42 sons analysés. Ces courbes établissent une relation quasi linéaire entre  $f_1(t)$ ,  $V_1(t)$  et  $f_1(t)$ ,  $V_2(t)$  pour les sons de flûtes jouées avec un effet de *vibrato*. Il semble donc que pour ces instruments, le nombre de relations entre les paramètres dynamiques soit important.

Observons maintenant les instruments à cordes frottées joués avec un effet de *vibrato*. Les "trajectoires"  $V_1(t)$ ,  $V_2(t)$  et  $V_1(t)$ ,  $V_3(t)$  ne révèlent pas de relations simples entre ces paramètres. Par contre, les courbes  $f_1(t)$ ,  $V_1(t)$  et  $f_1(t)$ ,  $V_2(t)$  semblent dégager certaines relations assez complexes. Il semble donc que la fréquence fondamentale  $f_1(t)$  joue un rôle

déterminant vis-à-vis de la composition harmonique de ces instruments.

## Conclusions

La réduction globale du *corpus* des 42 sons est rendue impossible à cause de la diversité des profils spectraux  $P_1$ . L'observation des profils de déviation spectrale  $P_2$  nous laisse supposer qu'il est possible de séparer les sons en trois classes. Enfin, l'observation des courbes paramétriques de  $V_1(t)$ ,  $V_2(t)$ ,  $V_3(t)$ , et,  $f_1(t)$  nous permettent d'envisager des relations non-linéaires — pouvant inclure des phénomènes d'hystérésis, entre ces paramètres dynamiques.

## 2.4 Application à la synthèse et à l'édition

### 2.4.1 Procédé de reconstruction

La reconstruction s'effectue en deux étapes (figure 2.38). L'étape d'initialisation consiste à transformer les spectres harmoniques  $P_1, P_2, \dots$  en tables d'ondes par transformée de Fourier inverse. Après initialisation, les tables d'ondes contiennent les *fonctions périodiques principales*  $T_1, T_2, \dots$ . L'étape de synthèse proprement dite consiste simplement à combiner ces tables d'ondes selon leurs enveloppes respectives  $V_1(t), V_2(t), \dots$ . La simplicité de cette étape permet la synthèse en léger temps différé, voire en temps réel [HBH93, HB95]. Cette méthode n'exige que trois ou quatre oscillateurs. Comparée à la synthèse additive qui en exige une quarantaine, nous obtenons un gain de calcul de l'ordre de dix.

### 2.4.2 Dilatation temporelle et transposition

Dilatation temporelle et transposition sont deux opérations indispensables au jeu musical : elles permettent de modifier la durée et la hauteur des sons. Voyons comment pratiquer ces deux opérations à partir des données réduites.

#### Dilatation temporelle

En première approximation, il suffit de modifier l'échelle temporelle de  $V_1(t)$ ,  $V_2(t), \dots$  et de  $f_1(t)$  pour réaliser cette opération. Dans la pratique, au moins trois cas doivent être considérés :

- Le cas où cette assertion est vérifiée ; il s'agit pour l'essentiel des instruments à vent pour lesquels les variations de dynamique sont pleinement contrôlées par le musicien ;
- le cas où la durée du transitoire d'attaque est indépendante de la durée de la note ; c'est par exemple le cas des instruments à cordes frappées et grattées. Il convient d'effectuer la dilatation temporelle sur la décroissance du son ;

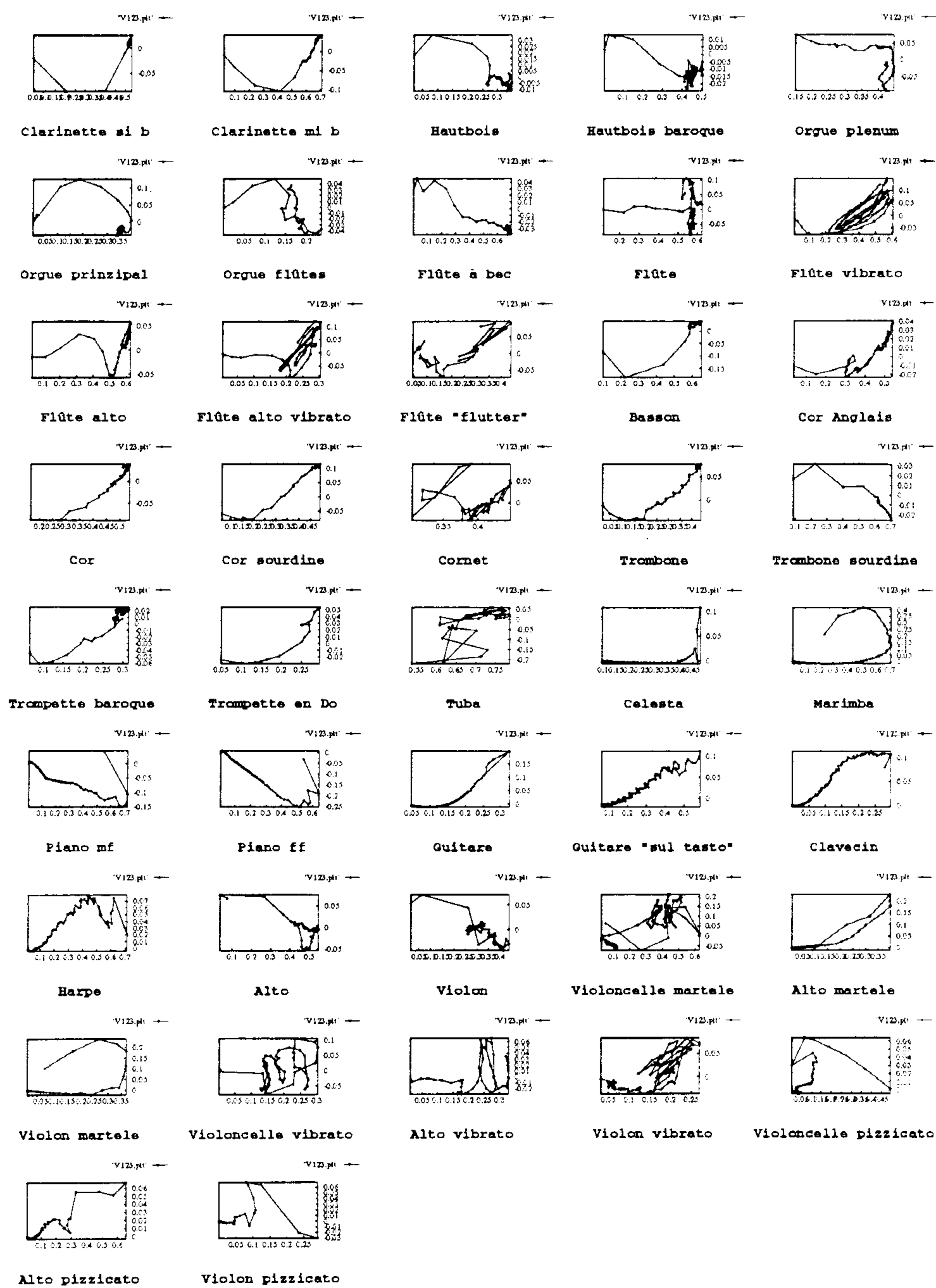
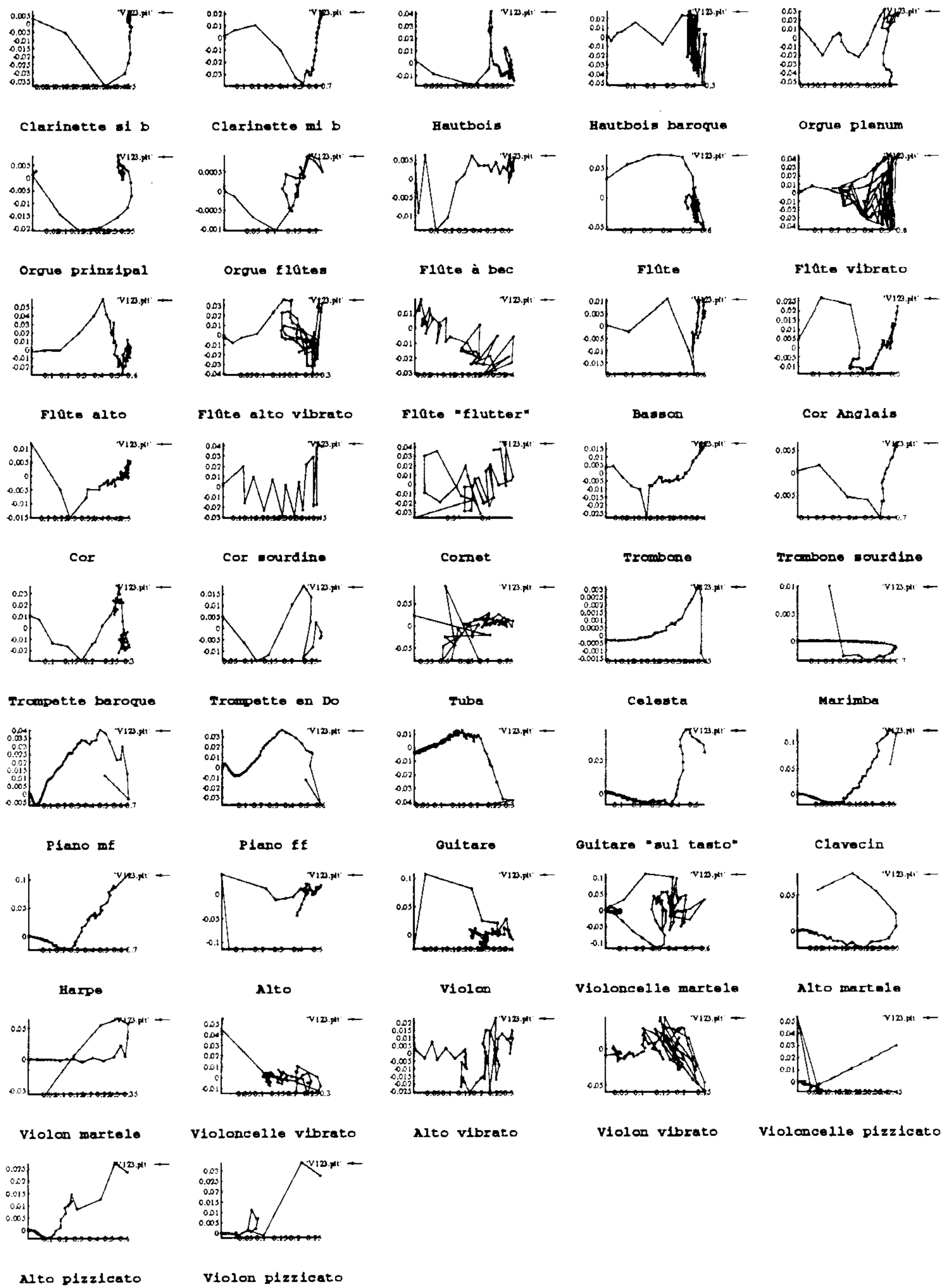


FIG. 2.34 - Courbe paramétrique:  $V_1(t)$ ,  $V_2(t)$



FIG. 2.35 - Courbe paramétrique:  $V_1(t)$ ,  $V_3(t)$

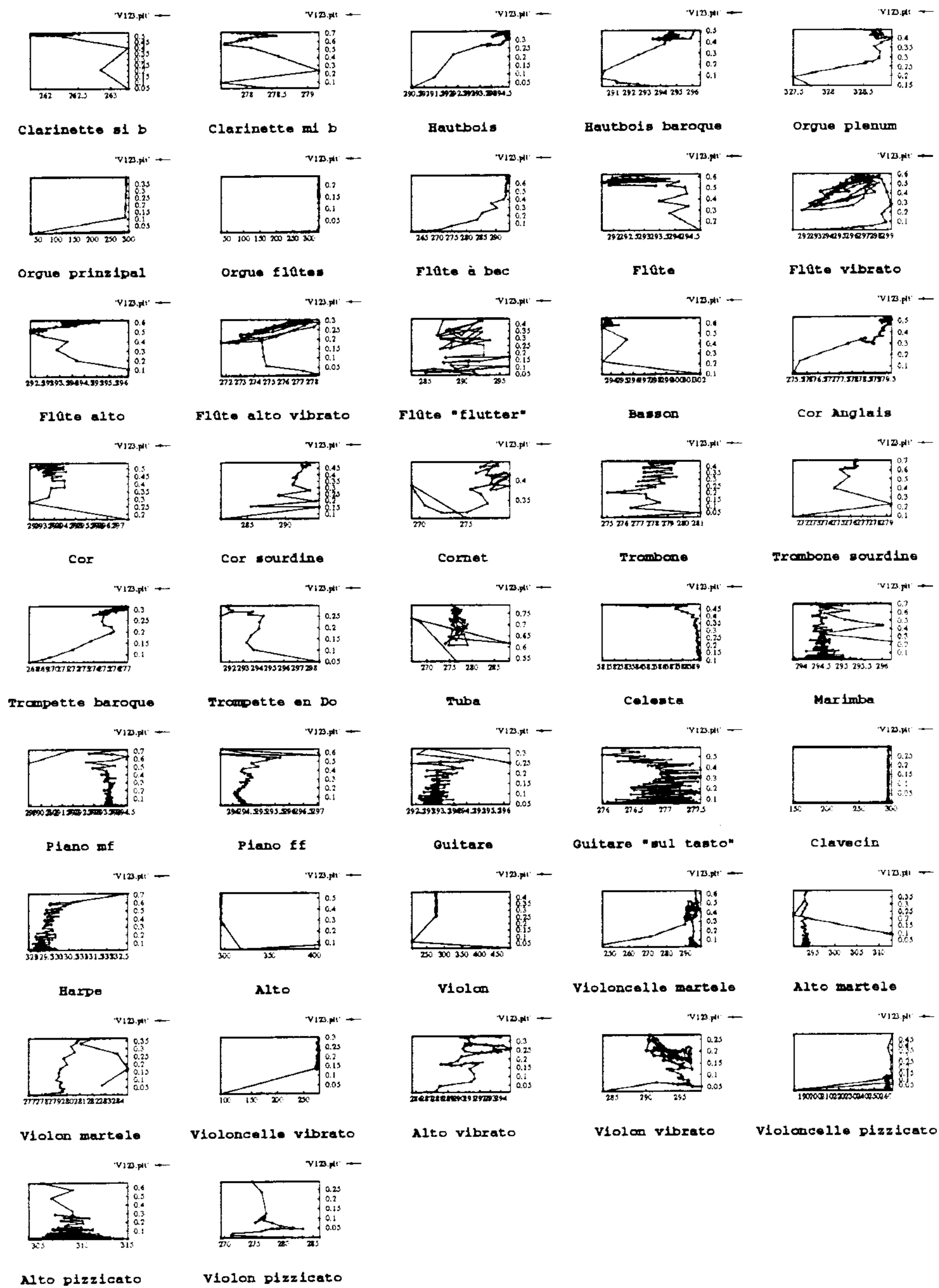


FIG. 2.36 - Courbe paramétrique :  $f_1(t)$ ,  $V_1(t)$

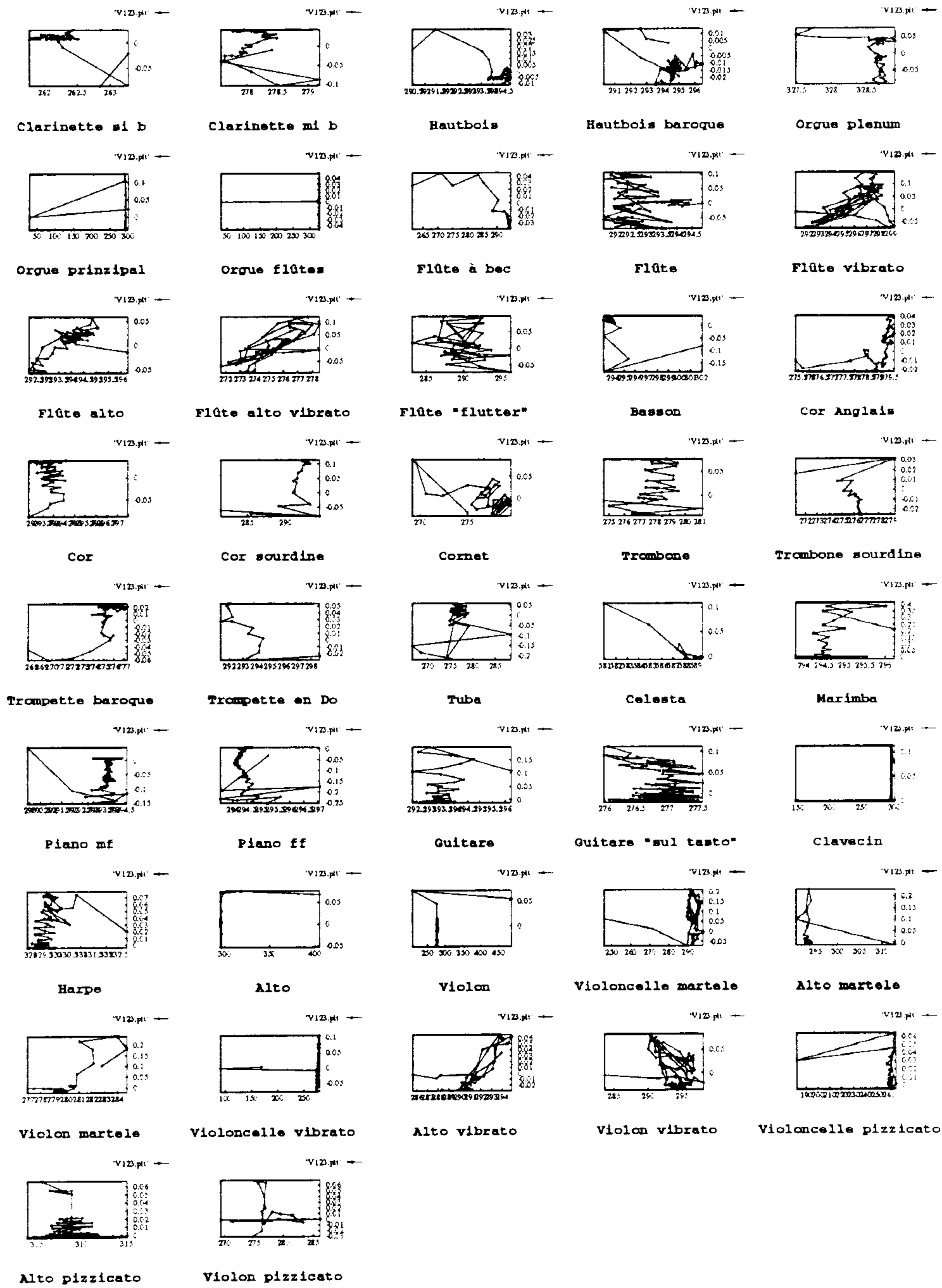


FIG. 2.37 - Courbe paramétrique :  $f_1(t)$ ,  $V_2(t)$

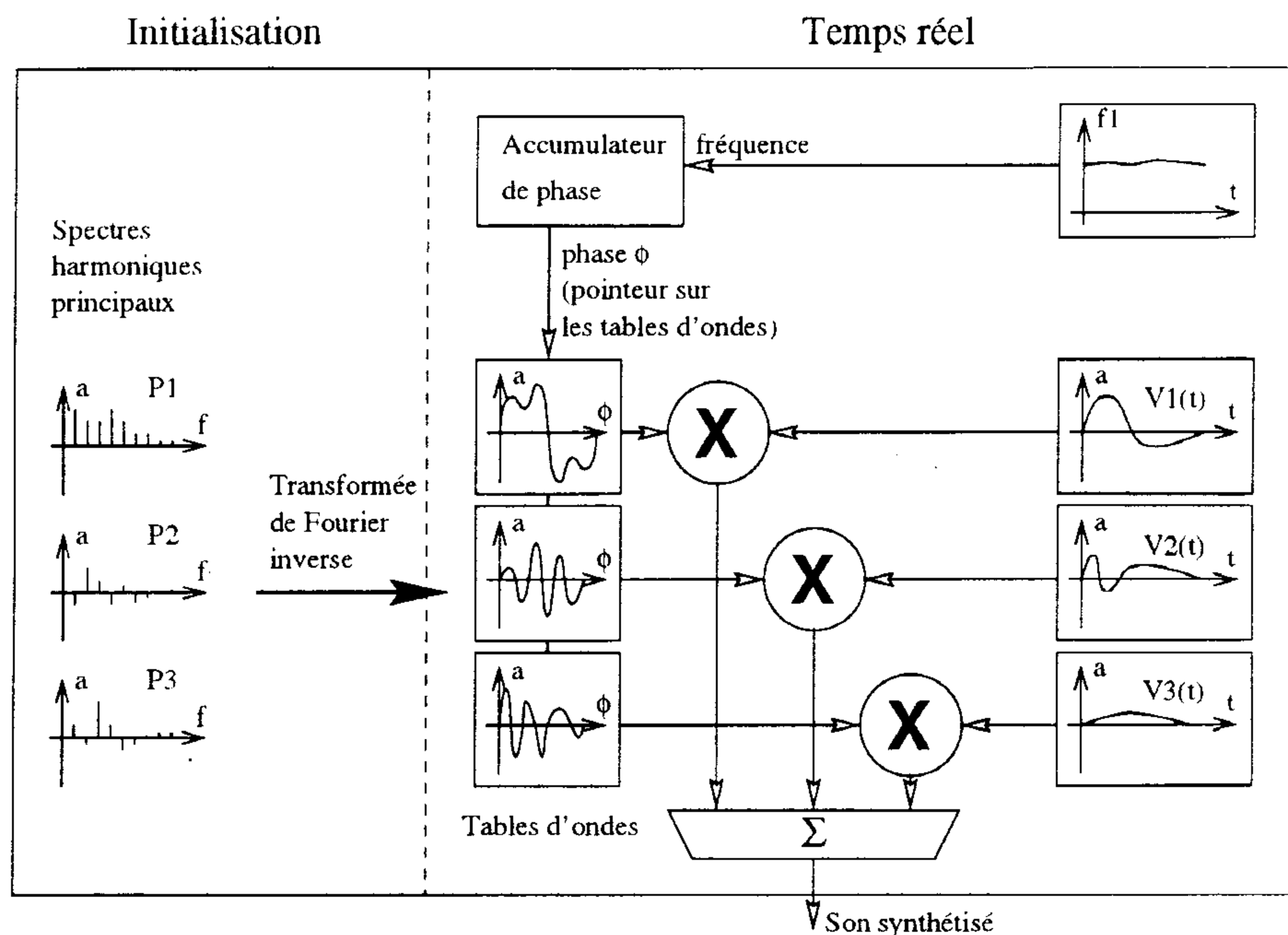


FIG. 2.38 – Reconstruction par combinaison de tables d'ondes

- Le cas des sons présentant des effets de *vibrato* et autre *tremolo*. Effectuer une dilatation ou une contraction temporelle provoquerait une variation artificielle des vitesses de ces effets. Une solution consiste à effectuer un bouclage des paramètres  $V_1(t)$ ,  $V_2(t)$ ,... et de  $f_1(t)$  sur une période d'effet.

## Transposition

*A priori*, l'opération de transposition est simple : il suffit d'effectuer une multiplication de la fréquence  $f_1(t)$  par  $2^{n/12}$ ,  $n$  étant le nombre de demi-tons chromatiques.

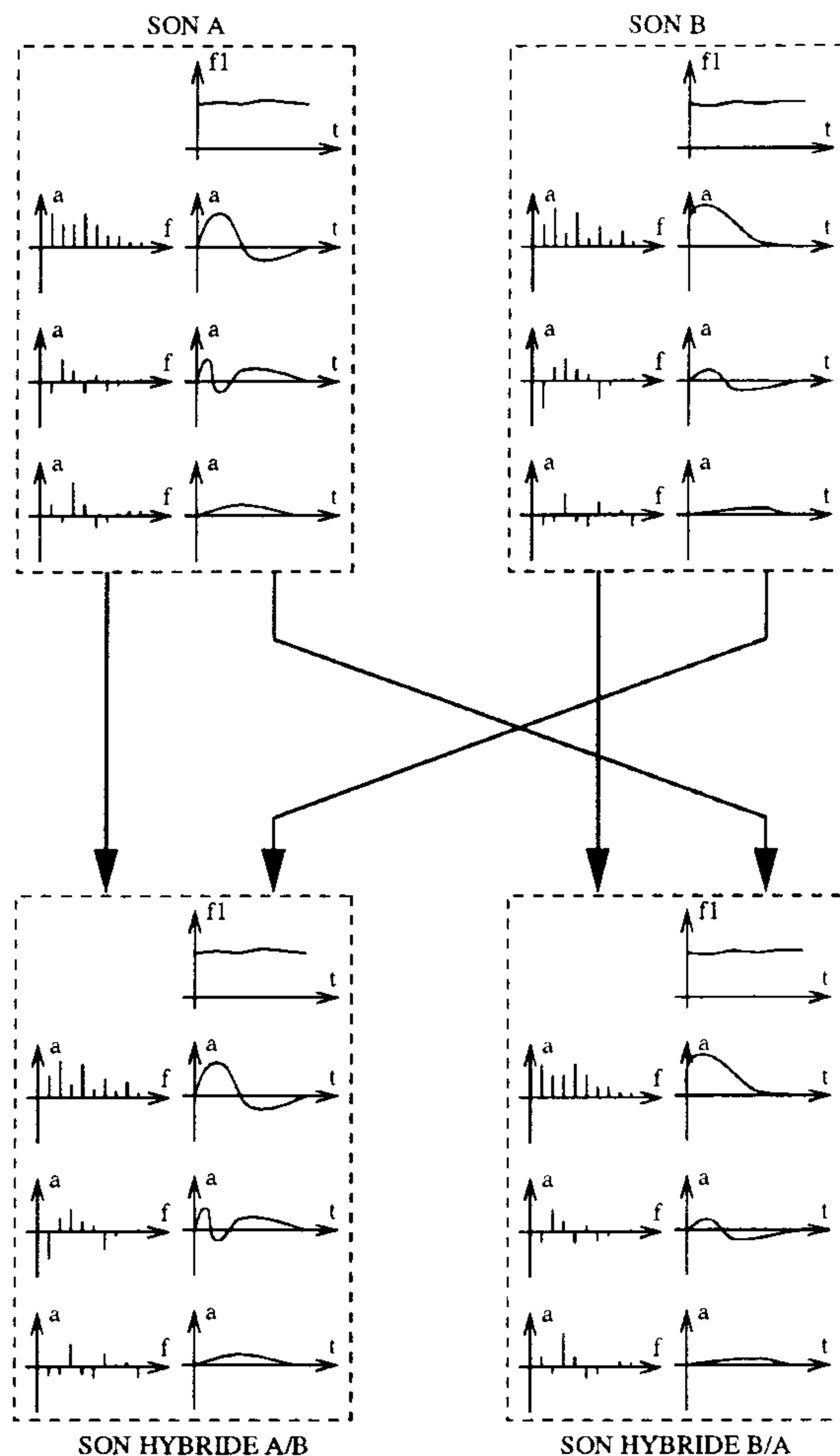
Dans la pratique, certains sons présentent de fortes "distorsions de timbre". Ces distorsions impliquent en général les sons ayant un "aspect formantique", c'est-à-dire issus d'instruments pour lesquels la caisse ou les plaques de résonance jouent un rôle important.

Une solution simple à ce problème est un multi-échantillonnage (§1.2.3, page 35) des tables  $T_1, T_2, \dots$ . En d'autres termes, le sous-espace  $\mathcal{P}$  n'est plus considéré comme constant mais variable selon la note jouée.

Une autre solution — basée sur certains principes de la synthèse par FOF — est proposée en 3.3, page 133.

## Conclusions

Les opérations de dilatation temporelle et de transposition sont relativement simples à effectuer à partir des données réduites. Cependant, ces opérations ne sont pas suffisantes pour créer des sons "expressifs" ou créer de nouveaux sons car on se limite aux données de dynamiques originales.

FIG. 2.39 – *Intervention de trajectoires*

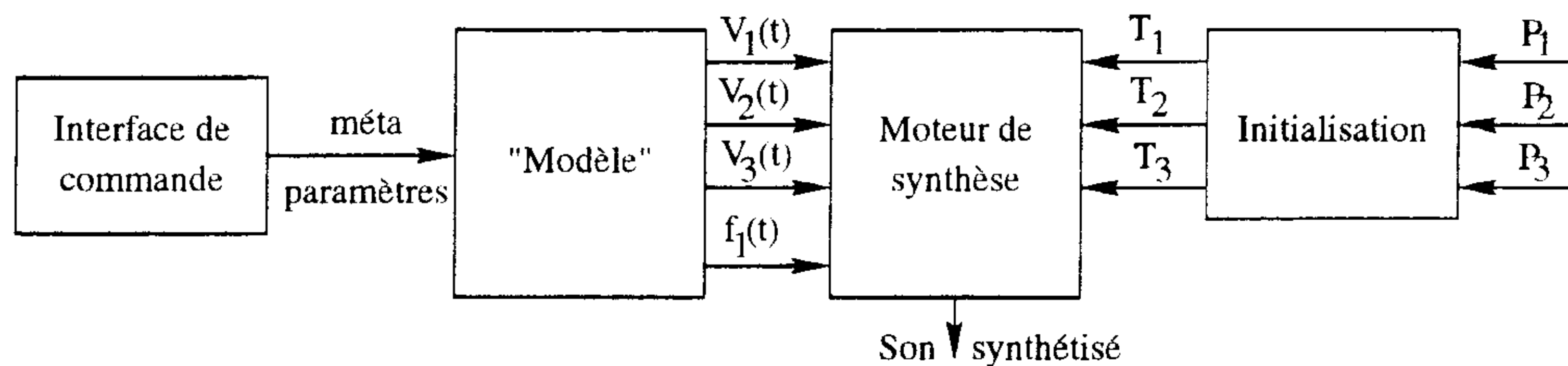
### 2.4.3 Synthèse croisée

La synthèse de sons hybrides consiste souvent à appliquer la réponse formantique d'un son à un autre son. Ce procédé fait généralement appel à un banc de filtres d'analyse et à un banc de filtres de synthèse commandé par le premier. Ce type de synthèse croisée est souvent appelé "vocodeur" car il est souvent utilisé pour faire "chanter" des instruments. Cette technique peut être de qualité satisfaisante lorsqu'elle est réalisée en temps différé sur ordinateur. L'implémentation "analogique" temps réel de cette méthode est souvent de piètre qualité et engendre des voix "robotiques".

Les quatre sons que nous avons choisis pour illustrer nos propos précédents nous font ressortir de multiples caractéristiques physiques du timbre des sons musicaux. Ces caractéristiques sont de deux natures : spectrales (spectres harmoniques) et dynamiques (trajectoire ou enveloppes).

Dans le cas général, la synthèse de sons hybrides consiste à :

- analyser deux sons ;
- intervertir une partie des données d'analyse ;
- resynthétiser deux sons à partir de ces données d'analyse "croisées".

FIG. 2.40 – *Synthèse de sons expressifs*

Ici, nous avons choisi d'invertir les trajectoires de deux sons (figure 2.39). Cette opération consiste simplement à “dessiner” la trajectoire  $\mathcal{V}$  d'un instrument dans le sous-espace harmonique  $\mathcal{P}$  d'un autre. Les conséquences de cette manipulation vont mettre en avant les rôles respectifs de  $\mathcal{V}$  et de  $\mathcal{P}$ .

Dans le cas simpliste de sons ne comportant qu'une enveloppe d'amplitude et un profil spectral constant, cette interversion consisterait à appliquer l'enveloppe d'un son à un autre son. Ainsi, notre manipulation peut être interprétée comme la généralisation de ce procédé simpliste de synthèse croisée.

Dans la pratique, les sons hybrides synthétisés à partir des quatre sons d'exemple ont des caractéristiques assez variées. Par exemple, le son résultant de la trajectoire de la flûte dans le sous-espace du hautbois est une sorte de “hautbois soufflé” assez naturel. A l'opposé, le son résultant de la trajectoire du piano dans le sous-espace du violoncelle fait penser à un son de clavecin plutôt qu'à un son de violoncelle joué *pizzicato*.

La synthèse croisée permet de valider en partie l'association caractéristiques spectrales / sous-espace harmonique et l'association caractéristiques dynamiques / trajectoire. Cette partie de nos travaux a fait l'objet de deux publications [RC97b, RC97a].

#### 2.4.4 Vers un modèle expressif

Un instrument acoustique — joué par un être humain — ne produit jamais deux fois le même son. Les variations de hauteur, d'intensité et de timbre de l'instrument sont influencées par le jeu du musicien. Pour être crédible — ou plutôt musicalement intéressant — la synthèse de sons musicaux doit pouvoir tenir compte de ces aspects.

Dans notre cas, notre objectif sera de nous affranchir des données dynamiques  $V_1(t)$ ,  $V_2(t)$ ,... et de trouver des “méta-paramètres” de jeu d'un niveau supérieur. L'étude menée en 2.3.4 apporte quelques éléments permettant d'atteindre en partie cet objectif: pour certains instruments, les enveloppes  $V_1(t)$ ,  $V_2(t)$ ,... et  $f_1(t)$  sont liées par des relations non-linéaires. Ainsi, il semble possible, à partir de ces relations, de réduire encore davantage le nombre de données dynamiques.

Etant donné les limitations expérimentales — normales dans le cadre d'une activité naissante — dans lesquelles nous avons menés nos travaux, nous nous sommes limités à quelques tentatives, assez empiriques, visant à générer les paramètres  $V_1(t)$ ,  $V_2(t)$ ,  $V_3(t)$ , et,  $f_1(t)$  à partir des données MIDI de note et de vélocité. Ces tentatives ont été menées en temps différé grâce au logiciel CSOUND [Ver92] et ont conduit à des résultats plus (flûtes et autres vents) ou moins (cordes frottées) probants.

### 2.4.5 Résumé

La synthèse à partir des données réduites conduit à une forme de synthèse par combinaison dynamique de tables d'ondes économique en calcul. Les opérations élémentaires de dilatation temporelle et de transposition sont relativement aisées et consistent à manipuler les données dynamiques  $V_1(t)$ ,  $V_2(t)$ ,... et  $f_1(t)$ . La synthèse de sons hybrides nous suggère que des manipulations plus avancées de ces paramètres peuvent conduire à des résultats dignes d'intérêt. Enfin, l'établissement d'un "modèle" de plus haut niveau, générant dynamiquement les paramètres  $V_1(t)$ ,  $V_2(t)$ ,... et  $f_1(t)$  à partir de "méta-paramètres" de jeu, permettrait d'obtenir une forme de synthèse expressive, c'est-à-dire tenant compte des nuances de jeu du musicien.

Bien que quelques tentatives aient été menées, il nous semble indispensable de compléter nos analyses afin d'affiner ces premiers modèles. A cette fin, il nous semble indispensable de disposer d'un plus grand nombre d'échantillons sonores — enregistrés dans des conditions parfaitement contrôlées — et d'élaborer un procédé automatique "d'extraction" de ces "méta-paramètres".

## 2.5 Conclusions

La représentation des sons musicaux sous forme de données réduites est un puissant outil d'analyse du timbre musical. Elle permet notamment de faire ressortir des liens qui unissent de nombreux paramètres physiques tant spectraux que dynamiques et d'en extraire des informations pertinentes.

En simplifiant la représentation des sons musicaux, la réduction de données permet de simplifier la resynthèse des sons analysés. Les sons ne sont plus décrits sous forme d'une somme de nombreuses sinusoïdes, mais comme une somme de quelques formes d'ondes. Il en résulte un gain de coût de calcul de l'ordre de dix (par comparaison à la synthèse additive).

Par opposition à la synthèse additive, cette méthode n'utilise qu'un nombre réduit de paramètres significatifs. En première approximation, le sous-espace harmonique définit la "couleur" et les variations de couleurs du son — il serait donc l'équivalent d'une "palette" — la trajectoire, qui en définit la dynamique serait l'équivalent du "tracé d'un dessin". Une étude psychoacoustique — du type de celle de John M. Grey — pourrait étendre cette idée en liant ces paramètres à des attributs perceptifs.

Nuançons quelque peu nos propos en rappelant que notre méthode se limite à la représentation et à la synthèse des sons musicaux pseudo-périodiques : elle ne peut ni tenir compte des déviations de fréquence de chaque harmonique ni considérer la composante "bruitée" des sons. Des extensions de la méthode sont exposées dans le chapitre suivant.

L'une d'elle consiste à généraliser l'analyse en composantes principales aux amplitudes complexes afin de tenir compte des phases. En considérant les phases et de leurs variations, cette extension rend compte des légères déviations de fréquence des harmoniques.

Une seconde extension consiste à extraire, analyser et synthétiser la composante bruitée des sons musicaux. Cette composante, négligée pendant assez longtemps en synthèse des sons musicaux, peut se révéler importante, notamment durant les phases transitoires.

## Chapitre 3

# Extensions du modèle

Nous avons proposé dans le chapitre précédent une méthode d’analyse/synthèse basée sur une représentation optimale — au sens de la consommation de variance — de sons harmoniques. Ce modèle est contestable à plus d’un titre :

- il ne rend pas compte de l’inharmonicité de certains sons ;
- il ne tient pas compte des composantes “bruitées”<sup>1</sup> présentes dans les transitoires ;
- il n’est pas toujours adapté à la transposition de certains sons à fort comportement formantique.

Ce chapitre a pour objectif de proposer plusieurs solutions permettant de généraliser, compléter, valider et adapter notre méthode à une plus large catégorie de sons musicaux.

La section 3.1 propose une généralisation du principe de réduction de données aux sons inharmónicos tenant compte de l’information de phase de chaque composante pseudo-harmonique. Par ailleurs, cette méthode permet de resynthétiser le son en respectant les phases initiales de chaque composante. Il est donc possible d’effectuer une soustraction directe dans le domaine temporel entre le son original et le son resynthétisé. La nature du son “résiduel” permet de valider ou non l’analyse.

Lorsque l’analyse et la synthèse sont correctement pratiqués, le son résiduel correspond à la composante bruitée du son analysé. La section 3.2 envisage plusieurs solutions permettant d’analyser et de synthétiser cette composante.

Enfin, la section 3.3 propose une méthode — inspirée de la synthèse granulaire et de la synthèse par FOF — permettant de transposer les sons à fort comportement formantique en limitant la distorsion de timbre.

### 3.1 ACP de spectres harmoniques complexes

L’étude présentée au chapitre 2 ne tenait compte que de sons pseudo-périodiques strictement harmoniques, c’est-à-dire comportant un ensemble d’harmoniques dont les amplitudes évoluent dans le temps. Il est largement reconnu que certains instruments

---

1. Bruits d’impacts et/ou sons à “spectres continus” posant des problèmes de résolution temps/fréquence.



accordés ne répondent que très partiellement à cette description. Le son du piano, par exemple, en raison de phénomène de dispersion du milieu de propagation de l'onde (les cordes), comporte des “harmoniques” dont les fréquences ne sont plus tout à fait multiples de la fréquence fondamentale. On dénomme habituellement de telles composantes “*partiels*”. Pour des raisons de cohérence qui apparaîtront plus loin, nous userons toujours du terme “harmonique”. En effet, la méthode que nous présentons ici est une généralisation de la méthode d'analyse par ACP permettant de tenir compte des déviations de fréquence des harmoniques.

### 3.1.1 Phases et amplitudes complexes

Une des possibilités permettant d'inclure les phénomènes d'inharmonicité serait d'inclure dans l'ACP les fréquences des harmoniques — ou leur rapport avec la fréquence fondamentale — aux côtés de leurs amplitudes. Les travaux de Christophe Hourdin [Hou95] ont montré qu'une telle démarche conduit à une méthode d'étude du timbre très efficace mais parfois difficile à maîtriser en raison de l'inhomogénéité — au sens des dimensions physiques — des données analysées. De plus, cette inhomogénéité interdit la transposition de la méthode à la synthèse par combinaison de tables d'ondes.

La solution que nous adoptons ici repose sur le principe utilisé par le filtrage hétérodyne: une déviation de fréquence peut être représentée comme une variation de phase d'une trame à la suivante. Nous considérerons donc que chaque “partiel” correspond à un harmonique déviant de sa fréquence “théorique” multiple de la fréquence fondamentale.

Le procédé d'analyse spectrale utilisé (§2.1, page 63) donne une mesure de l'amplitude, de la fréquence, mais aussi de la phase de chaque composante sinusoïdale en milieu de trame. Pour être utilisable *a posteriori* en synthèse par combinaison dynamique de tables d'ondes, la phase de chaque harmonique  $h$  doit être renormalisée par rapport à la phase du fondamental :

$$\tilde{\phi}_h^{\mathcal{H}}(T) = \phi_h^{\mathcal{H}}(T) - h\phi_1^{\mathcal{H}}(T) \quad [2\pi] \quad (3.1)$$

Cette renormalisation facilite la lecture de tables d'ondes car la phase du fondamental y fait office de pointeur de référence.

Pour chaque trame  $T$ , chaque harmonique  $C_h^{\mathcal{H}}(T)$  est maintenant représenté par un terme d'amplitude  $a_h^{\mathcal{H}}(T)$  et un terme de phase relative  $\tilde{\phi}_h^{\mathcal{H}}(T)$ . Ces informations peuvent être regroupées en une amplitude complexe :

$$\tilde{a}_h^{\mathcal{H}}(T) = a_h^{\mathcal{H}}(T)e^{2\pi i\tilde{\phi}_h^{\mathcal{H}}(T)} \quad (3.2)$$

L'ensemble des  $\tilde{a}_h^{\mathcal{H}}(T)$  — par analogie à la trajectoire harmonique  $\mathcal{A}$  — définit la *trajectoire harmonique complexe*  $\tilde{\mathcal{A}}$ .  $\tilde{\mathcal{A}}$  et les paramètres du fondamental  $f_1^{\mathcal{H}}(T)$ ,  $\phi_1^{\mathcal{H}}(T)$  permettent de reconstruire le signal en respectant strictement les phases trame à trame et donc les déviations de fréquence des harmoniques.

### 3.1.2 Analyse en composantes principales de $\tilde{\mathcal{A}}$

L'analyse en composantes principales est effectuée sur l'ensemble  $\tilde{\mathcal{A}}$  des  $\tilde{a}_h^{\mathcal{H}}(T)$  d'une façon très similaire à ce qui a été effectué sur l'ensemble  $\mathcal{A}$  des  $a_h^{\mathcal{H}}(T)$  au chapitre 2 :

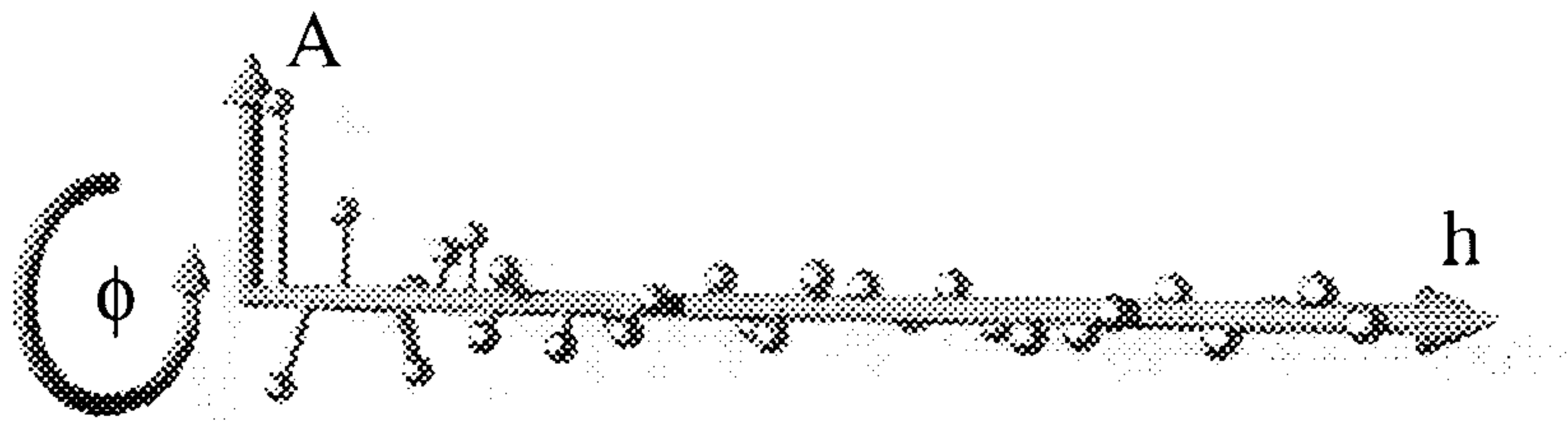


FIG. 3.1 – Spectre harmonique complexe

normalisation des modules, calcul de la matrice de covariance, diagonalisation et projection sur les axes principaux. Toutes ces opérations sont effectuées ici sur des nombres complexes. Les vecteurs propres complexes de la matrice de covariance (implicitement hermitienne), définissent les *composantes principales complexes*  $\tilde{P}_p$  du *sous-espace harmonique principal complexe*  $\tilde{\mathcal{P}}$ . La projection  $\tilde{V}$  de la trajectoire initiale  $\tilde{A}$  est obtenue par produit scalaire :

$$\tilde{A} \rightarrow \tilde{V} : \tilde{V}_p(T) = \sum_{h=1}^H \tilde{a}_h^H(T) \cdot \tilde{P}_p^{h*} \quad (3.3)$$

$p$  représentant l'ordre du spectre,  $h$  le rang de l'harmonique,  $H$  le nombre d'harmoniques, et,  $\tilde{P}_p^{h*}$  le complexe conjugué de  $\tilde{P}_p^h$ .

### 3.1.3 Sous-espace et trajectoire complexes

Chaque axe  $\tilde{P}_p$  du sous-espace harmonique complexe  $\tilde{\mathcal{P}}$  est un *spectre harmonique complexe*. la figure 3.1 représente un tel spectre : chaque harmonique est maintenant affecté d'un module et d'une phase.

Chaque coordonnée  $\tilde{V}_p(T)$  de la *trajectoire complexe*  $\tilde{V}$  représente une *enveloppe complexe*. La multiplication de  $\tilde{V}_p(T)$  par  $\tilde{P}_p$  consiste à multiplier  $\tilde{P}_p$  par le module de  $\tilde{V}_p(T)$  et à faire “tourner” les phases de  $\tilde{P}_p$  selon l'argument de  $\tilde{V}_p(T)$ .

Le processus de reconstruction du spectre du signal d'origine est représenté en figure 3.2. Il peut être interprété comme une généralisation du procédé décrit au chapitre précédent : les composantes des axes réels  $P_p$  pouvaient être négatives — c'est-à-dire en opposition de phase — et ainsi se soustraire. A. Horner, J. Beauchamp et L. Haken avaient noté des difficultés d'utilisation des phases sans s'attacher à obtenir une solution [HBH93]<sup>2</sup>. Dans la méthode proposée ici, l'information de phase est pleinement maîtrisée et exploitée : les composantes se combinent dans le plan complexe.

### 3.1.4 Reconstruction par combinaison de tables d'ondes “complexes”

La transposition de la réduction de données pratiquée sur  $\tilde{A}$  à la synthèse par combinaison dynamique de tables d'ondes est moins triviale qu'au chapitre précédent.

2. [HBH93, page 338]: “Note that the phases of the corresponding harmonics of multiple wavetables must be the same to avoid inadvertent phase cancellation.”

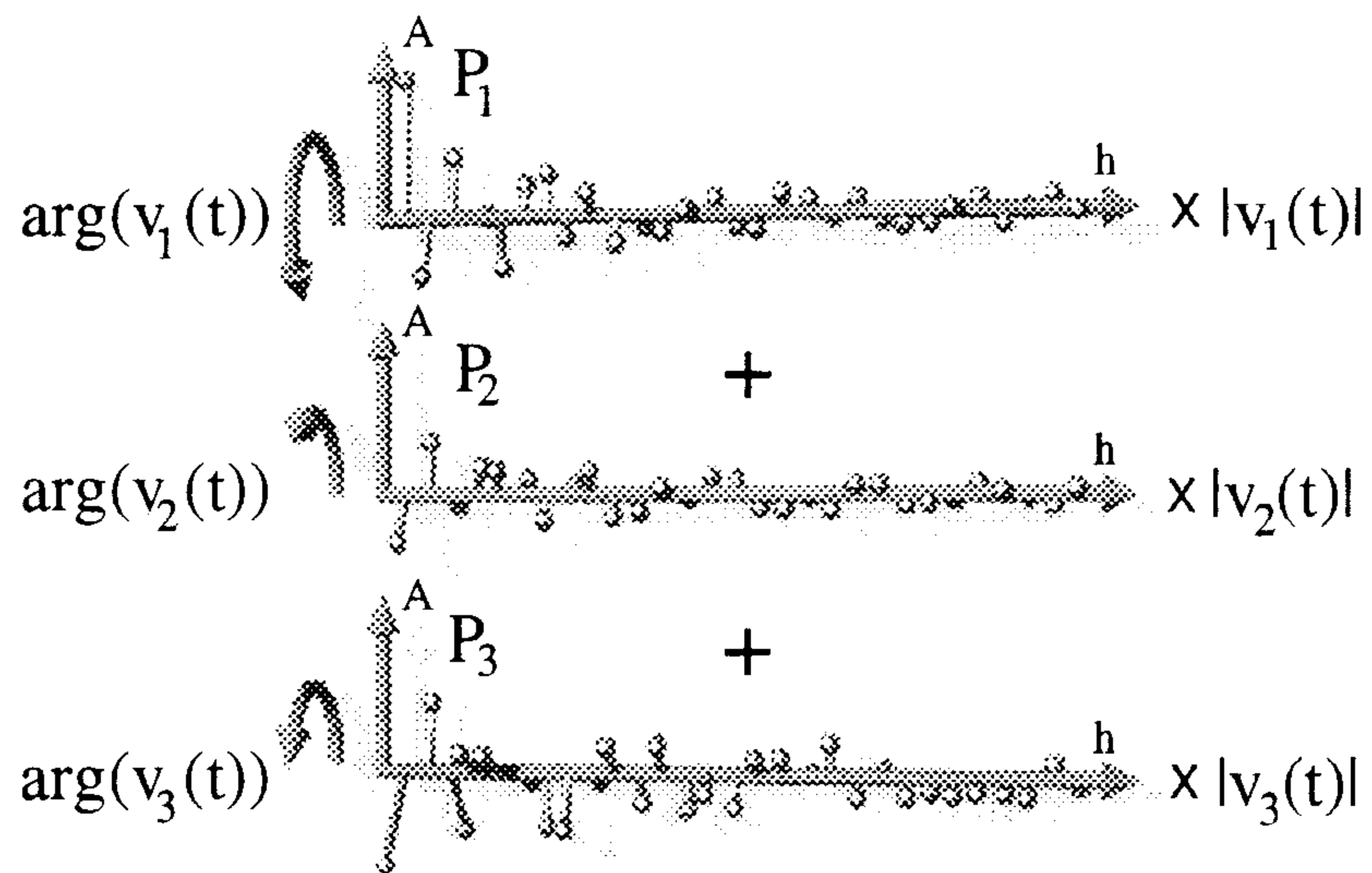


FIG. 3.2 – Synthèse par combinaison de spectres harmoniques complexes

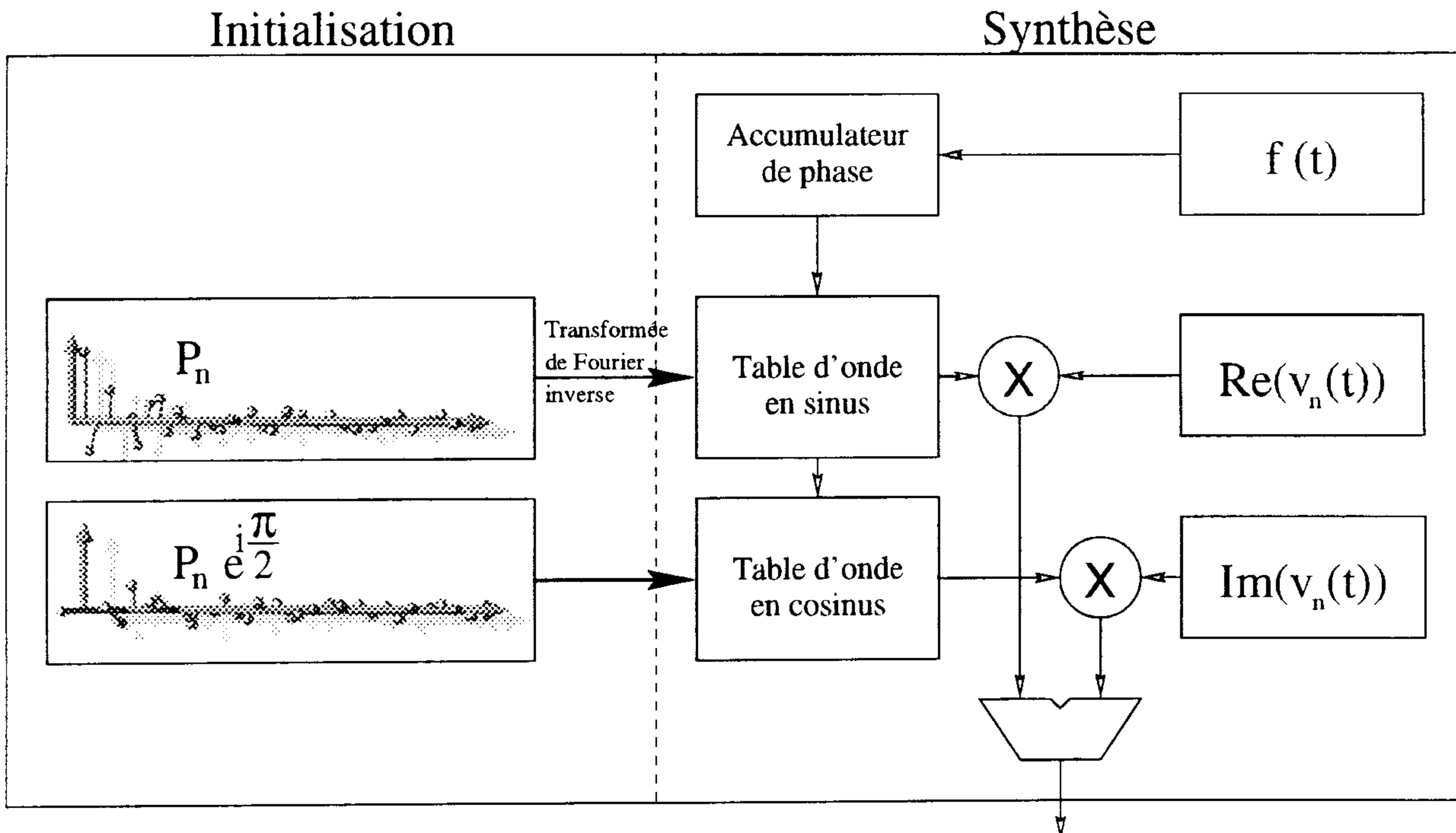


FIG. 3.3 – Synthèse par combinaison de tables d'ondes "complexes"

La synthèse ne consiste plus simplement à pondérer et à additionner des tables d'ondes, elle consiste aussi à faire varier leurs phases. L'opération à réaliser pour chaque spectre harmonique est illustrée en figure 3.3 : la rotation des spectres est obtenue par combinaison de tables d'ondes en quadrature de phase.

Le nombre d'opérations à effectuer est ainsi multiplié par deux mais chaque composante harmonique se combine dans le plan complexe ce qui offre plus de "liberté" qu'une simple addition/soustraction.

### 3.1.5 Analyse de sons synthétiques

Nous reprenons ici la procédure expérimentale déjà utilisée en 2.2.4 : nous allons analyser et interpréter les données réduites provenant de sons synthétiques.

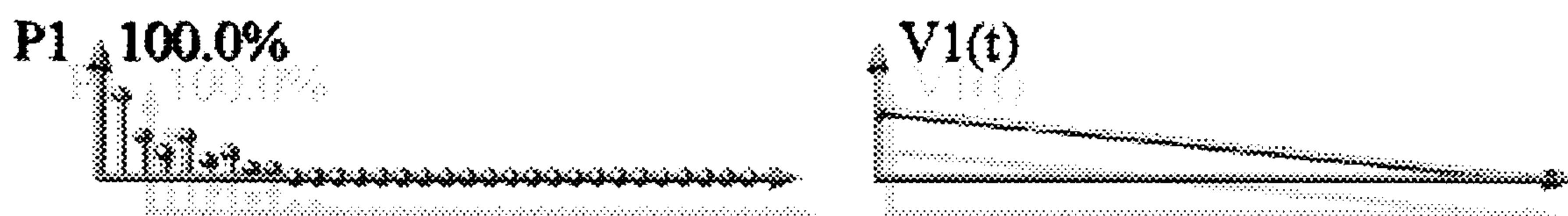


FIG. 3.4 – Son de profil spectral constant (phases relatives nulles)

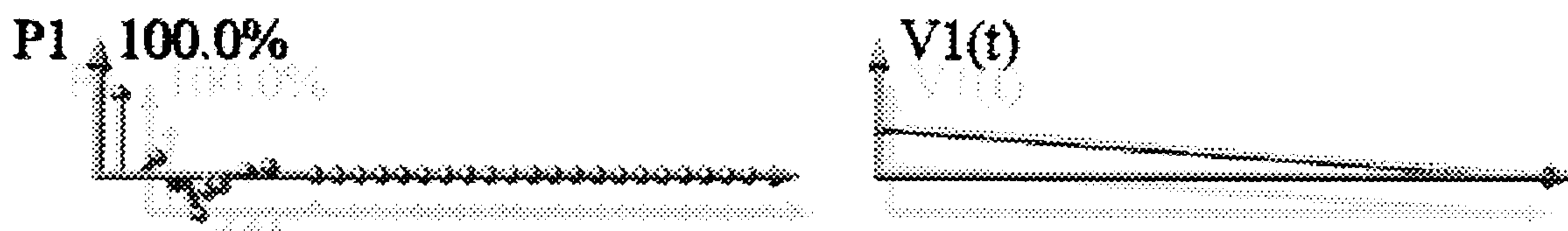


FIG. 3.5 – Son de profil spectral constant (phases relatives non nulles)

#### Observation de la phase

Considérons en premier lieu un son `test1` relativement simple. Il dispose d'un profil spectral constant dans le temps composé de 8 composantes sinusoïdales de phases relatives au fondamental nulles. La figure 3.4 représente les données réduites complexes de ce son qui se résument, dans ce cas simple, à  $\tilde{P}_1$  et  $\tilde{V}_1(t)$ . La grandeur représentée en pourcentage correspond à la "consommation" de variance de l'axe considéré. On constate que cette représentation est en tout point similaire à celle étudiée au chapitre précédent.

Considérons maintenant un son ayant le même spectre d'amplitude mais dont les harmoniques sont successivement déphasés de 45 degrés. La figure 3.5 correspond à un tel son : le déphasage relatif de chaque harmonique de  $\tilde{P}_1$  est représenté par l'angle qu'il décrit par rapport à l'axe du fondamental. L'enveloppe  $\tilde{V}_1(t)$ , quant à elle, a une phase constante et varie en amplitude.

La représentation que nous avons adoptée permet donc d'observer les relations de phase des sons à profil spectral constant ce que ne permettait pas la représentation présentée au chapitre 2.

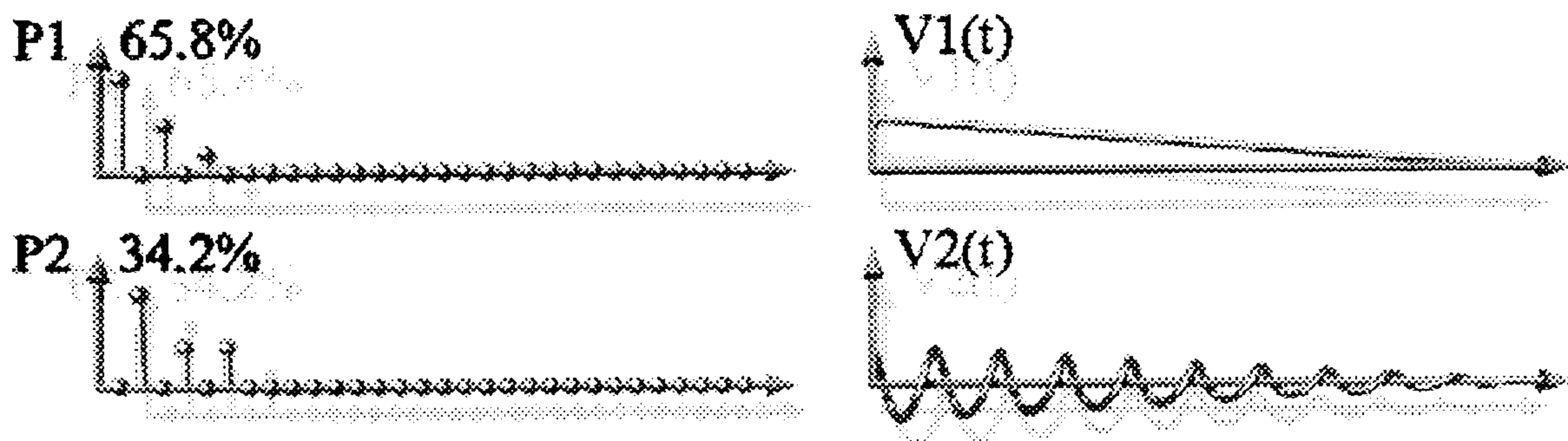


FIG. 3.6 – *Inharmonicités représentées sous forme de rotation de phases*

### Rotations de phase et inharmonicité

Considérons maintenant un son légèrement inharmonique composé de 6 partiels de fréquences : 295Hz, 600Hz, 885Hz, 1190Hz, 1475Hz, 1780Hz. La figure 3.6 nous permet de constater une séparation entre les harmoniques de rangs pairs et les harmoniques de rangs impairs. En effet, la déviation de fréquence (de 10Hz) des harmoniques de rangs pairs peut être représentée par une rotation de phase. Le spectre harmonique  $P_2$  regroupe ces harmoniques et leur associe une vitesse de rotation de phase correspondant à 10Hz (10 rotations sur une durée d'une seconde).

Un spectre harmonique comportant certaines raies déviant de façon égale de leur fréquence "théorique" se caractérise par des regroupements d'harmoniques et des enveloppes complexes en forme de tire-bouchons.

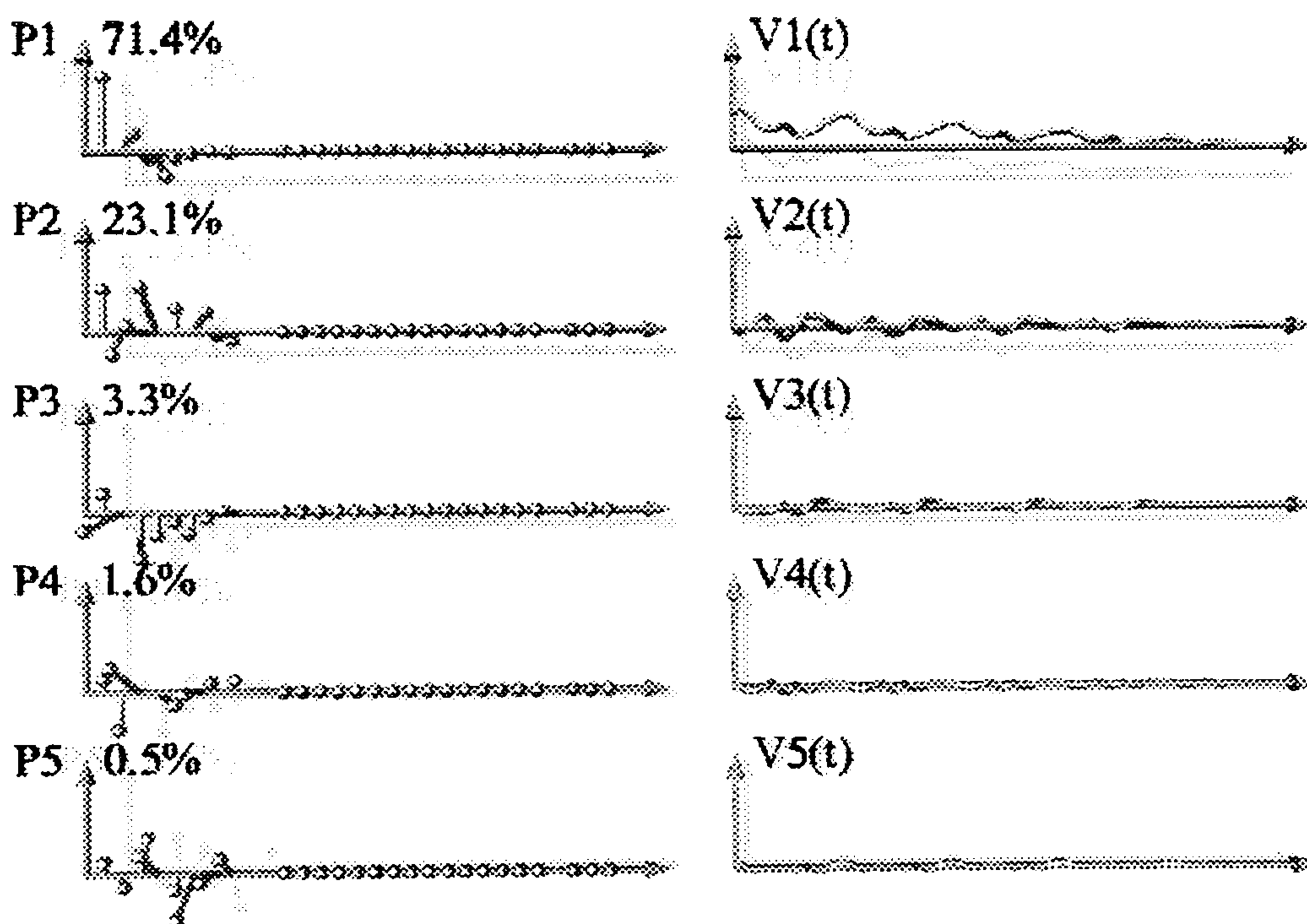


FIG. 3.7 – *Battements représentés sous forme de modulation de phases*

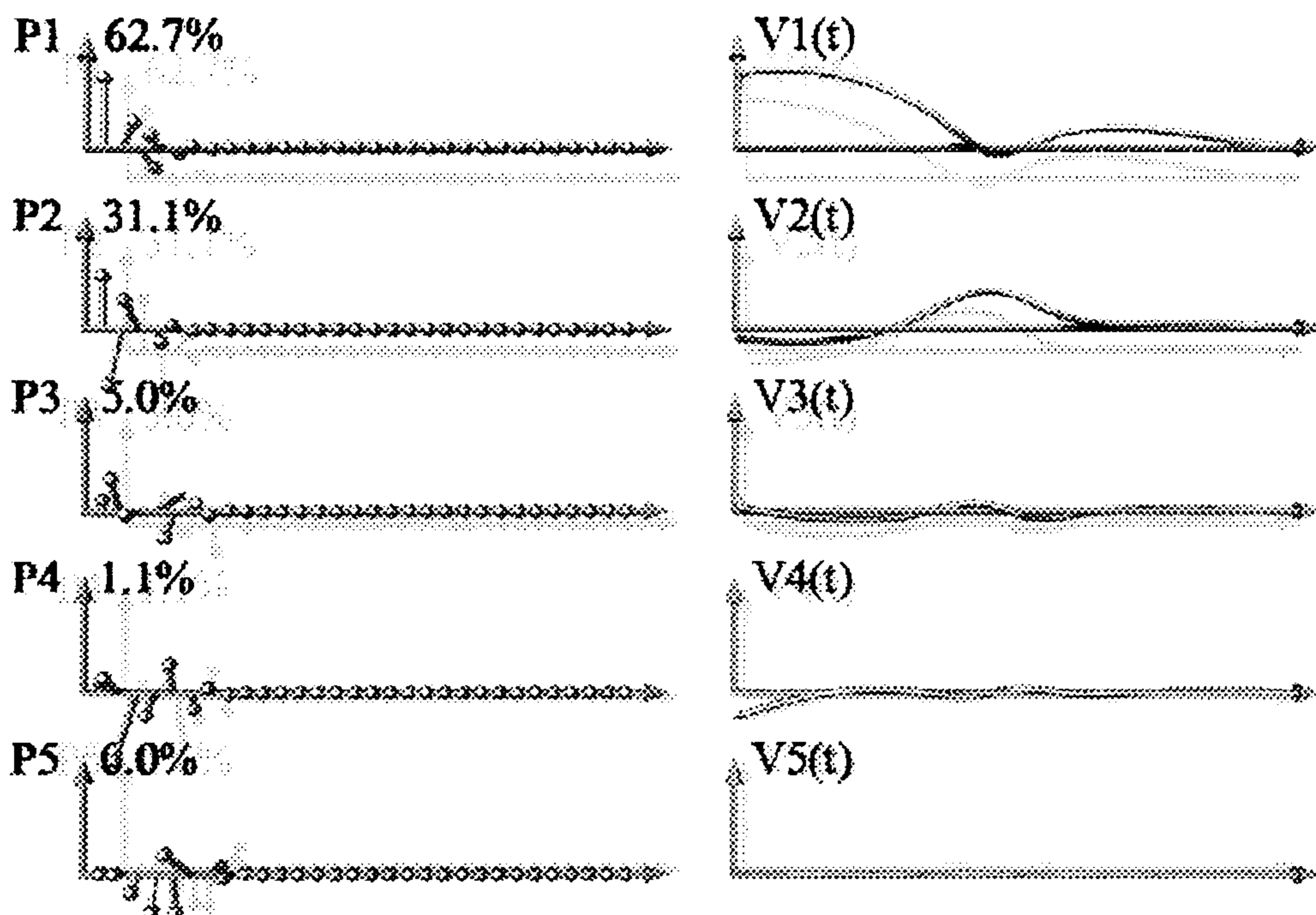


FIG. 3.8 – Données réduites complexes d'un son FM ( $I$  variant de 4 à 0)

### Rotations de phase et battements

Considérons maintenant un son constitué par superposition de deux signaux de fréquences différentes, l'un à 295Hz, l'autre à 300Hz. Cette superposition engendre des battements entre les harmoniques. La figure 3.7 représente les données réduites d'un tel son. Les enveloppes complexes prennent un aspect périodique — correspondant à la période des battements — et compliqué : chaque enveloppe est sujet à modulation de phase et d'amplitude.

Ainsi, la superposition de deux sons de fréquences proches se traduit par des enveloppes spectrales compliquées et nécessite l'utilisation de nombreux axes.

### Sons de modulation de fréquence à deux opérateurs

Nous reprenons ici les sons analysés en 2.2.4, page 81. La figure 3.8 représente les données réduites d'un son obtenu avec une fréquence porteuse  $f_p$  de 285Hz, une fréquence de modulation  $f_m$  de 285Hz et un indice de modulation  $I$  variant linéairement de 4 à 0.

Outre les ondulations déjà observées en 2.2.4 sur les enveloppes  $V_1(t)$ ,  $V_2(t)$ ,... nous observons des phénomènes de rotation de phase dépendant de l'indice de modulation. La forme des spectre  $\tilde{P}_1$ ,  $\tilde{P}_2$ ,... est très similaire à celle observée précédemment, les phases ont cependant une forme assez compliquée.

La figure 3.9 correspond à un son synthétisé dans les mêmes condition mais selon un indice de modulation variant de 8 à 0. On remarque une accentuation du phénomène de rotation de phase des enveloppes.

La figure 3.10 correspond à un son synthétisé avec une fréquence porteuse  $f_p$  de

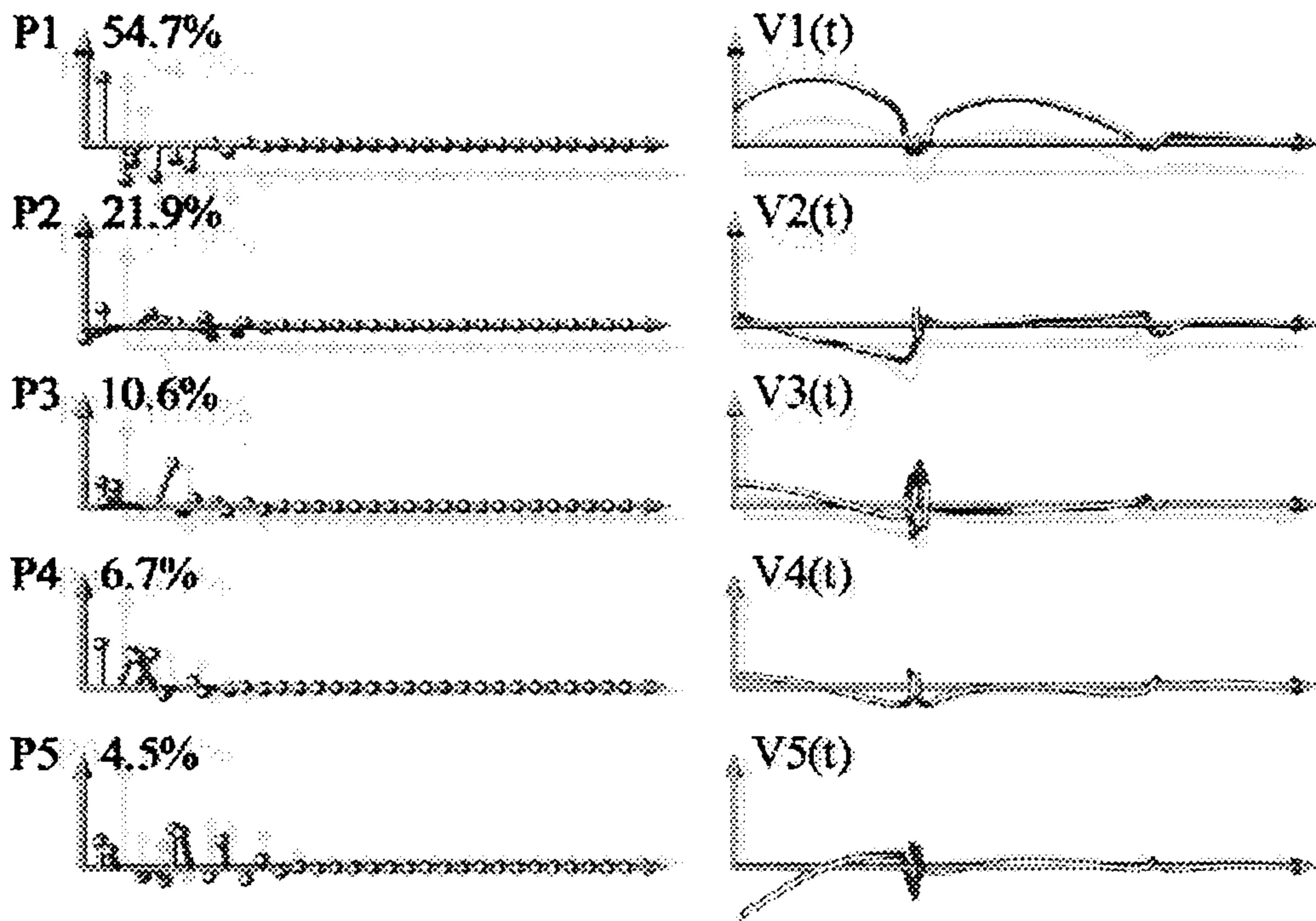


FIG. 3.9 – Données réduites complexes d'un son FM ( $I$  variant de 8 à 0)

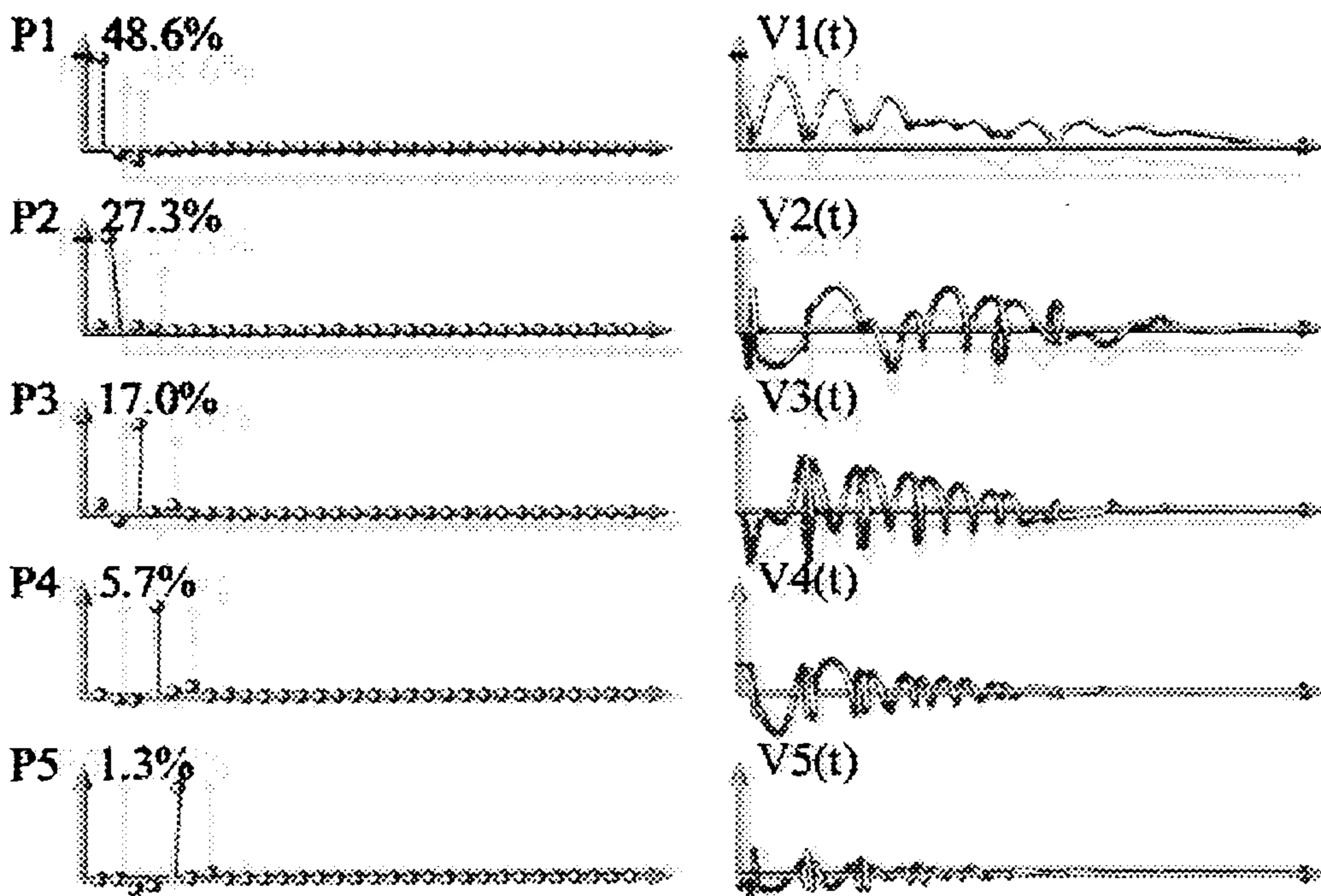


FIG. 3.10 – Données réduites complexes d'un son FM présentant des battements

285Hz, une fréquence de modulation  $f_m$  de 290Hz et un indice de modulation  $I$  variant linéairement de 4 à 0. La différence de 5Hz entre  $f_m$  et  $f_p$  engendre des phénomènes de battement bien connu des adeptes de la synthèse par modulation de fréquence<sup>3</sup>. La représentation sous forme de données réduites complexes est en quelque sorte une combinaison des résultats observés en figure 3.6 et en figure 3.10. La complexité des enveloppes provient de la superposition des phénomènes liés aux variations d'indice et aux battements.

L'analyse en composantes principales nous donne des résultats très similaires à ceux observés en 2.2.4, page 81 : le nombre de composantes nécessaires à une représentation correcte est du même ordre de grandeur et les ondulations des enveloppes ont une forme comparable. Cependant, cette représentation ajoute une visualisation des phases des enveloppes qui permet d'observer les phénomènes de battement de "raies repliées".

### Conclusions

L'analyse de sons synthétiques nous a permis d'observer plusieurs caractéristiques des données réduites complexes :

- ces données permettent d'observer le spectre harmonique moyen  $\tilde{P}_1$  ainsi que ses phases ;
- les phénomènes de inharmonicité (déviations de fréquence de groupes d'harmoniques) se traduisent par des enveloppes en tire-bouchon et un groupement des harmoniques concernés dans le spectre harmonique correspondant ;
- les phénomènes de battements — provenant de superposition de sons de fréquences légèrement différentes — provoquent un phénomène de même nature que les inharmonicités mais sensiblement plus compliqué ;
- les sons obtenus par modulation de fréquence engendrent des données réduites assez compliquées mais comparables à celles obtenues au chapitre 2.

Ces observations — somme toute simplistes — nous permettront de mieux interpréter les données réduites complexes de sons d'instruments acoustiques.

### 3.1.6 Analyse de quatre sons musicaux

Nous reprenons encore une fois le processus expérimental du chapitre précédent (§2.3.2, page 88) en pratiquant l'analyse de quatre sons musicaux. Ces quatre sons sont extraits du cédérom "Classical Sounds" de l'université McGill :

- un son de piano (PianomfD4) joué *mezzo forte* ;
- un son de hautbois (OboeD4) ;

---

3. les raies de fréquences "négatives" sont repliées et interfèrent avec les raies de fréquences positives. Ce procédé astucieux permet d'obtenir des effets de *tremolo* sans faire intervenir un oscillateur basse fréquence supplémentaire.



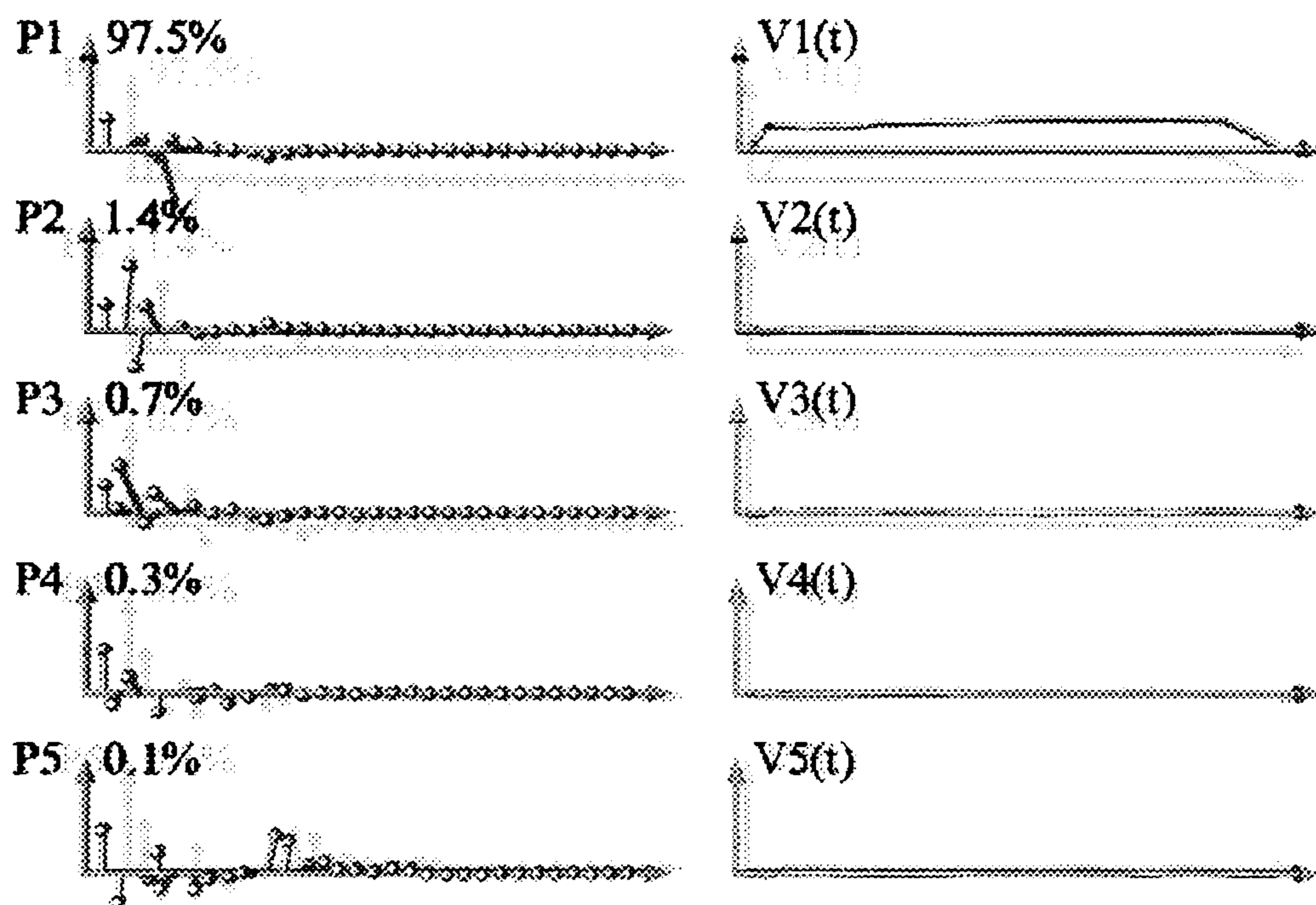


FIG. 3.11 – Données réduites complexes d'un son de hautbois

- un son de violoncelle (CelloWvibC#4) joué avec effet de *vibrato*;
- un son de flûte (FluteWvibD4) joué avec effet de *tremolo*.

### Son de hautbois

La figure 3.11 représente les cinq premières composantes d'un son de hautbois. 97,5% de variance sont consommés par le spectre  $\tilde{P}_1$ , les trois premiers spectres suffisent à consommer 99,6% de la variance. Le son de hautbois est donc fort bien représenté par cette méthode. Le spectre  $\tilde{P}_1$  dispose des mêmes amplitudes que  $P_1$  observé au chapitre 2. En outre,  $\tilde{P}_1$  apporte une information supplémentaire sur les phases relatives des harmoniques : ils ne sont pas exactement en phase avec le fondamental. Cette observation des phases relatives "statiques" a relativement peu d'intérêt dans le sens que nous n'avons pas d'informations précises sur les conditions d'enregistrement du son analysé. En effet, une application de la méthode pourrait être l'observation des phases relatives des harmoniques en fonction de la disposition du micro par rapport à l'instrument ; ce type d'observation permettrait de caractériser en partie les phénomènes de déphasage liés au rayonnement par l'instrument.

### Son de piano

Les résultats de l'analyse du son de piano sont représentés en figure 3.12. Tout comme les spectres  $P_1$  et  $P_2$  observés au chapitre 2,  $\tilde{P}_1$  et  $\tilde{P}_2$ , "gèrent" l'opposition entre les deux premiers spectres harmoniques. Leurs enveloppes  $\tilde{V}_1(t)$  et  $\tilde{V}_2(t)$  ne sont

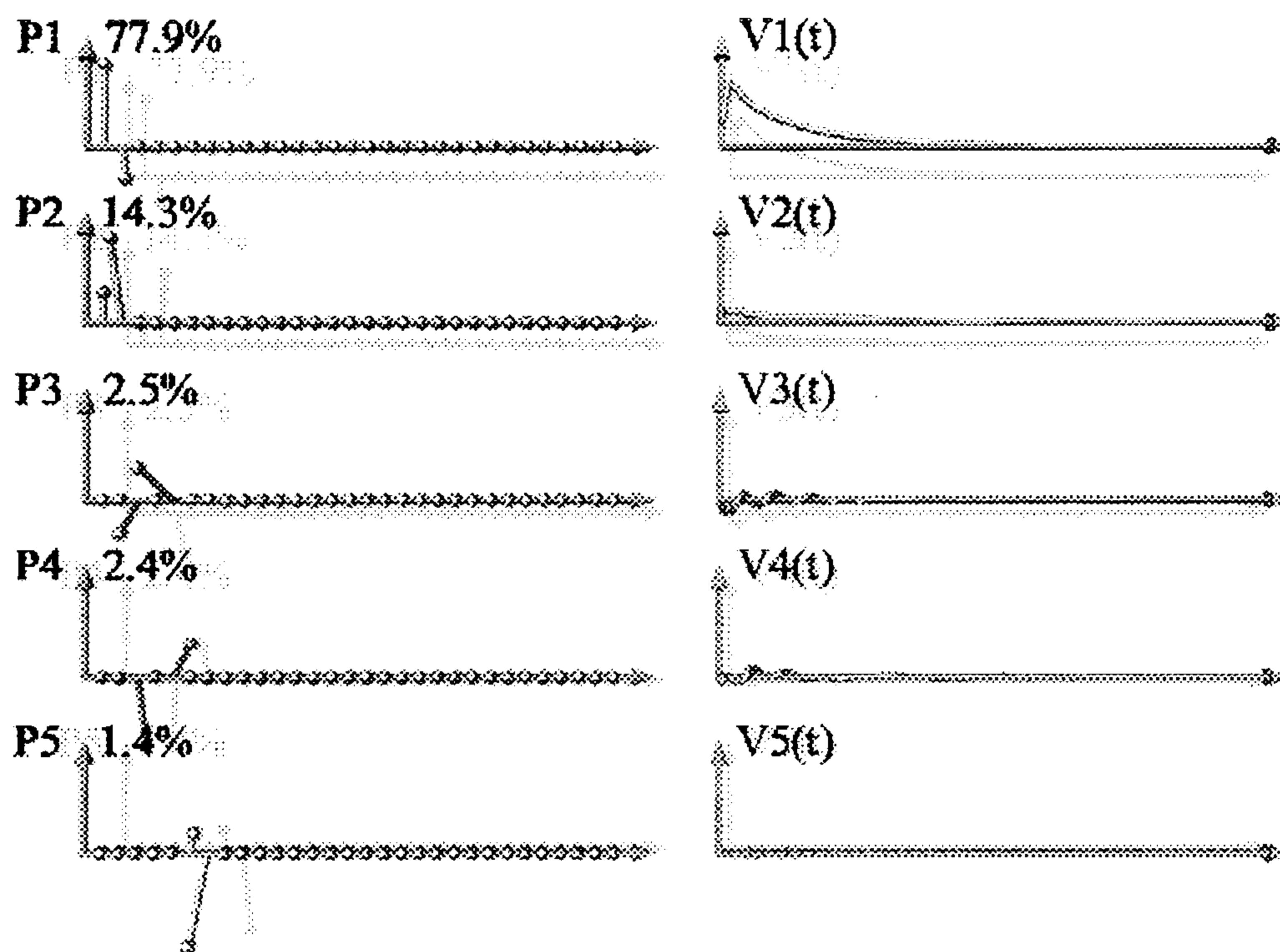


FIG. 3.12 – Données réduites complexes d'un son de piano

pas sans rappeler les enveloppes observées aux chapitre précédent (§2.3.2, page 90). Les axes d'ordres supérieurs révèlent des enveloppes  $\tilde{V}_3(t)$ ,  $\tilde{V}_4(t)$  et  $\tilde{V}_5(t)$  en "tire-bouchon". Cette forme de variation des phases est caractéristique des inharmonicités. On remarque qu'en ce qui concerne  $\tilde{V}_3(t)$  et  $\tilde{V}_4(t)$ , la vitesse de rotation du "tire-bouchon" diminue lorsque l'amplitude du son diminue. Cette observation peut s'expliquer par le fait que le phénomène de dispersion lié à la raideur des cordes de piano — responsables des phénomènes d'inharmonicité — varie avec l'amplitude de l'oscillation. Une étude plus poussée de ce type de phénomènes serait à envisager à condition de disposer d'enregistrements de piano joué à diverses notes et intensités<sup>4</sup>.

### Son de flûte

Les données réduites complexes du son de flûte (figure 3.13) sont très similaires aux données réduites réelles : elles opposent l'harmonique 1 à l'harmonique 2.  $\tilde{P}_1(t)$  et  $\tilde{P}_2(t)$  permettent d'observer un déphasage relatif de 90 degrés entre les deux premiers harmoniques. En outre,  $\tilde{V}_1(t)$  et  $\tilde{V}_2(t)$  — de formes très similaires à  $V_1(t)$  et  $V_2(t)$  — montrent de légères variations de phases relatives lors du trémolo. Ces variations semblent liées à l'amplitude du son. Une étude plus poussée de ce phénomène déterminerait dans quelle mesure ces variations de phases sont caractéristiques de l'instrument.

4. Nous pensons que l'utilisation d'un piano acoustique piloté par MIDI permettrait d'obtenir une telle série d'échantillons de manière relativement aisée.

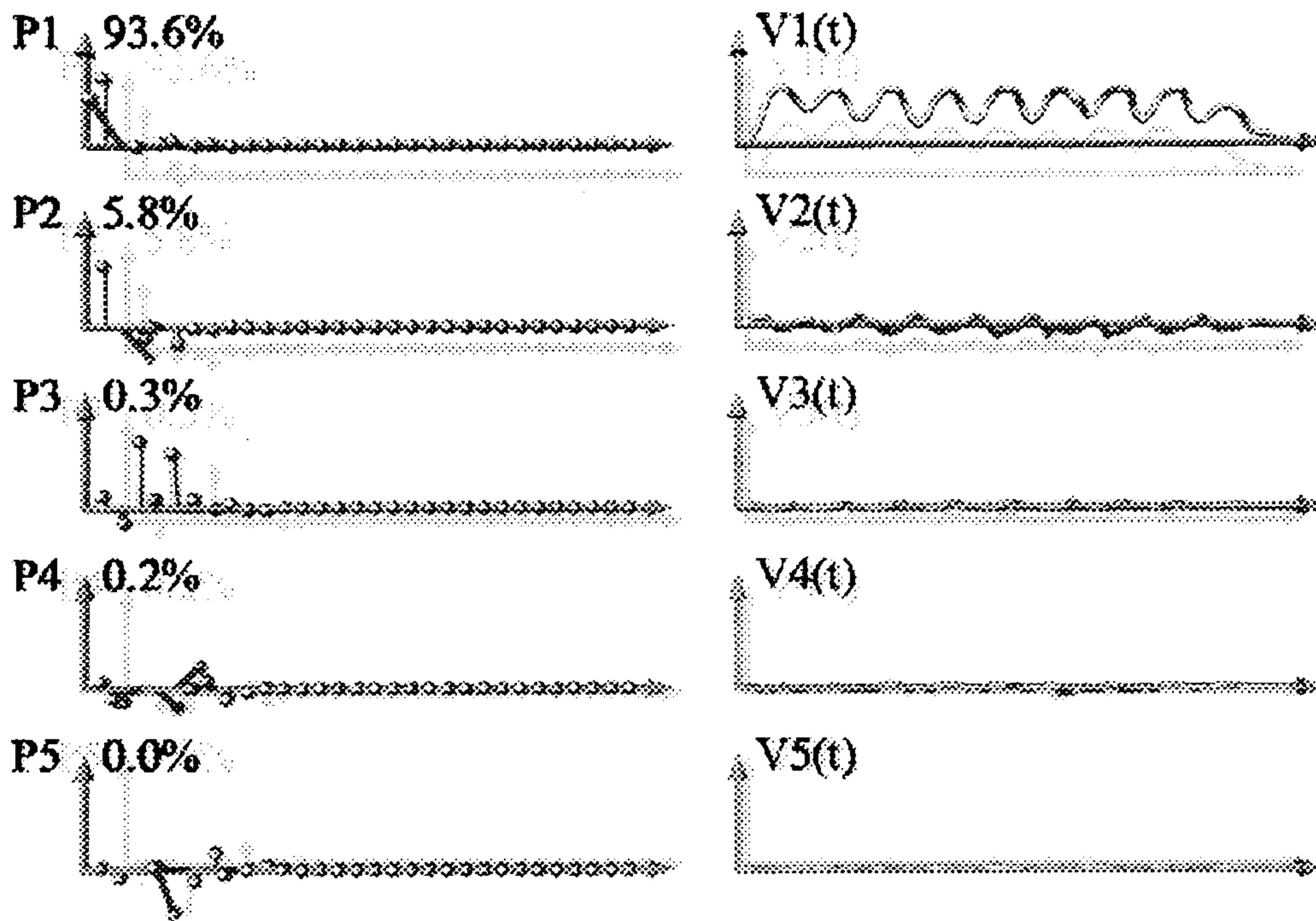


FIG. 3.13 – Données réduites complexes d'un son de flûte

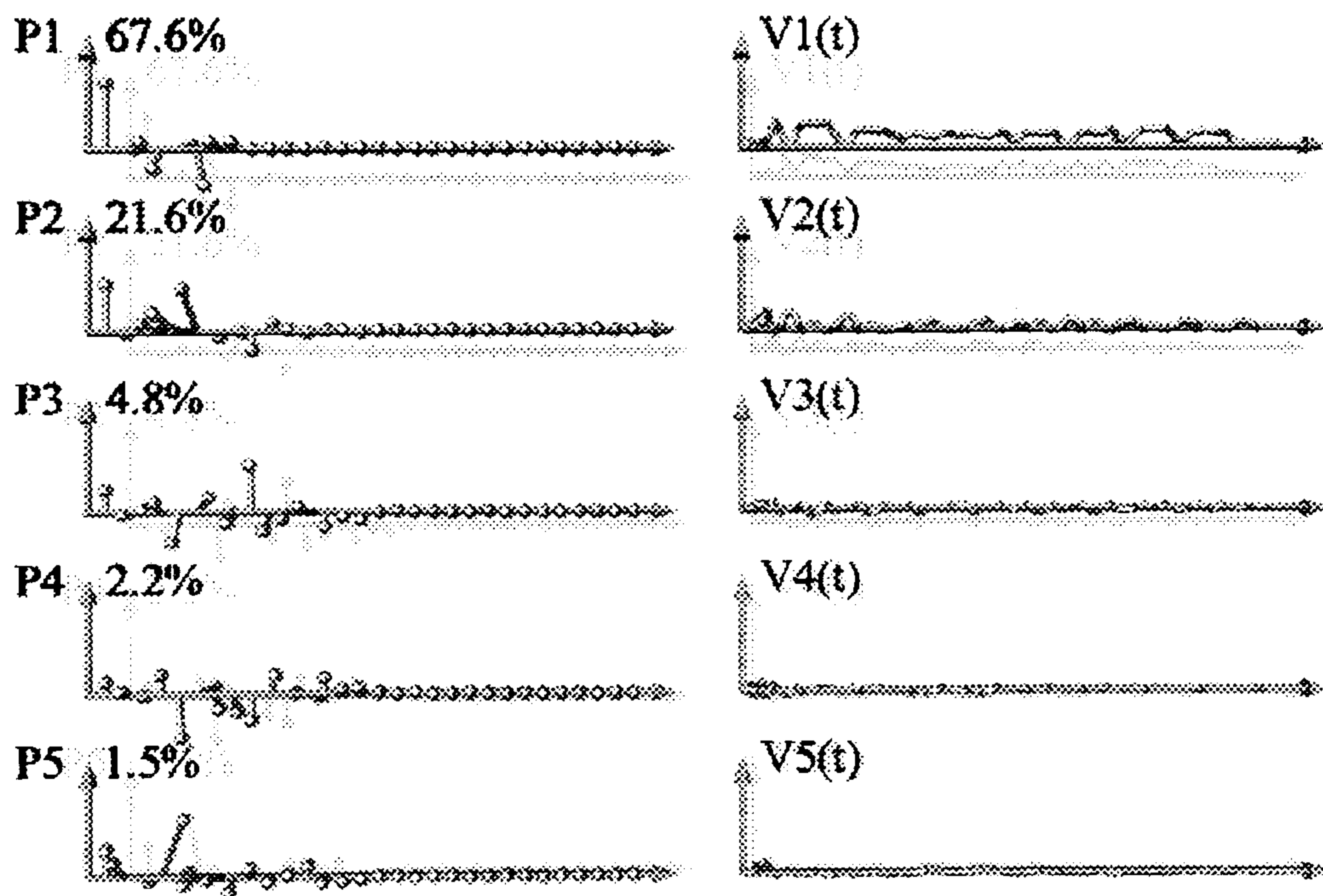


FIG. 3.14 – Données réduites complexes d'un son de violoncelle

### Son de violoncelle

Le son de violoncelle est relativement difficile à synthétiser par combinaison de tables d'ondes. La figure 3.14 nous révèle des données spectrales et dynamiques compliquées. Tout d'abord, les spectres harmoniques  $\tilde{P}_p$  ont des formes moins triviales que ceux observés précédemment, tant au point de vue des amplitudes que des phases. Deuxièmement, les enveloppes  $\tilde{V}_p(t)$  révèlent des phénomènes complexes de modulation d'amplitude et de phase. Pour terminer, la consommation de variance de chaque composante est moins importante que dans les cas précédents. Ainsi, outre des difficultés de réduction de données liées à la complexité des variations d'amplitudes, s'ajoutent les difficultés liées aux variations de phases.

Bien que la perception humaine des phases relatives soit souvent considérée comme faible — par rapport aux grandeurs physiques d'amplitude et de fréquence —, nous émettons l'hypothèse que ces variations contribuent au timbre particulier des instruments à cordes frottées. Nous pouvons en partie justifier cette assertion en considérant que le nombre important d'harmoniques composant ce type de sons multiplie le nombre de "phases relatives". Cette assertion semble se confirmer à l'écoute de sons synthétisés à partir des données réduites réelles et complexes<sup>5</sup>.

#### 3.1.7 Conclusions

L'analyse en composantes principales de spectres harmoniques complexes conduit à une représentation plus générale que celle présentée au chapitre 2. Outre les informations concernant le profil d'amplitude spectrale, elle ajoute des informations sur les phases relatives des harmoniques. Ces informations permettent non seulement d'observer des déphasages pouvant être caractéristiques des instruments joués mais aussi d'observer des rotations de phases caractéristiques d'inharmonicités, de battements ou de déphasages induits par une caisse de résonance.

Cette représentation peut donc être considérée comme un nouvel outil d'étude des phénomènes de déphasage intervenant dans certains instruments acoustiques. Elle pourrait se révéler utile en acoustique musicale (les phénomènes de déphasage sont en effet cruciaux en prise de son de proximité).

D'autre part, elle permet de resynthétiser certains sons inharmoniques, tels que les sons de piano, pour un coût de calcul approximativement double de la méthode présentée au chapitre 2. De façon pragmatique, on est en droit de se demander si ce surcoût est réellement indispensable. La réponse à ce type de question ne peut être apportée que par une série d'expériences psychoacoustiques visant à déterminer les différences de fidélité des deux méthodes.

## 3.2 Etude de la composante bruitée

La section précédente permet de faire état de certains sons inharmoniques en représentant les déviations de fréquences sous forme de rotations de phases. Ce modèle n'est

---

5. Une écoute au casque — parce qu'elle limite les déphasages liés au milieu de propagation du son — est plus démonstrative qu'une écoute sur enceintes.

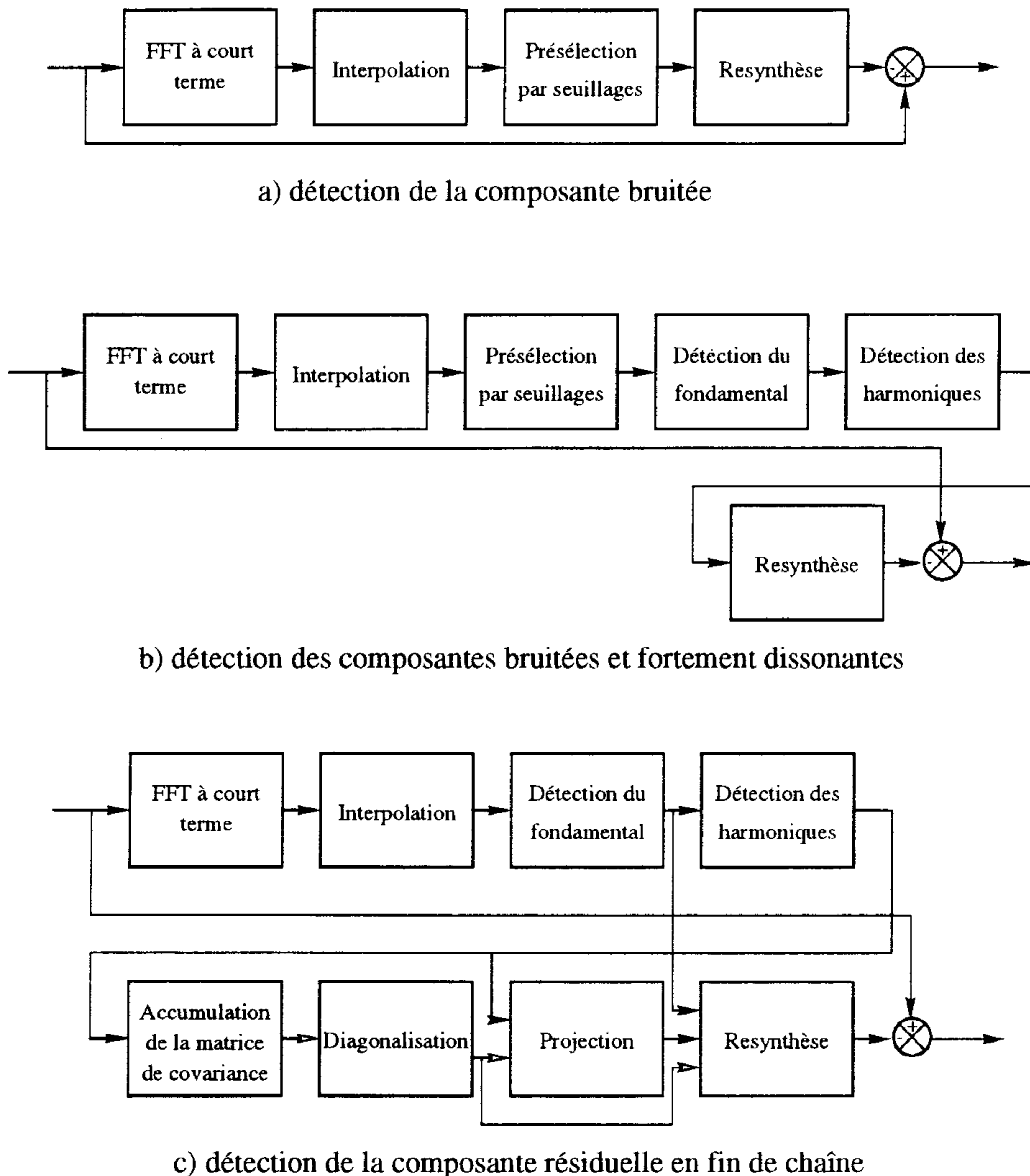


FIG. 3.15 – Extraction de la composante bruitée à trois stades de l'analyse

cependant pas encore complet car il ne tient pas compte de la composante “bruitée” du son.

La section 3.2.1 présente plusieurs méthodes d'extraction de cette composante fondées sur notre procédé d'analyse-synthèse. La section 3.2.2, page 127, présente une méthode d'analyse de la composante bruitée inspirée par le procédé employé pour la composante harmonique. Enfin, la section 3.2.3, page 129, propose plusieurs procédés visant à synthétiser cette composante.

### 3.2.1 Extraction de la composante bruitée

L'extraction de la composante bruitée peut être effectuée dans le domaine temporel par soustraction directe d'un son synthétique à la condition expresse que celui-ci respecte les phases du signal original. Comme nous l'avons vu précédemment, nos méthodes d'analyse-synthèse permettent de tenir compte de ces paramètres.

Grâce à nos outils, une telle soustraction peut avoir lieu à trois stades, chacune ayant

son utilité. La figure 3.15 les présente.

La soustraction peut être réalisée après la présélection par seuillages relatif et absolu, on obtient alors un bruit exempt de toute composante sinusoïdale inharmonique (figure 3.15a).

Elle peut être effectuée après la sélection des harmoniques, le son résiduel peut alors contenir quelques composantes fortement inharmoniques (figure 3.15b).

Enfin, elle peut être réalisée en bout de chaîne lorsqu'on utilise le procédé décrit dans la section précédente, cette manipulation a l'intérêt de démontrer — ou non — de façon directe l'efficacité du processus d'analyse-synthèse complet (figure 3.15c).

Les résultats des second et troisième processus doivent conduire à des résultats similaires lorsque la réduction de données est correcte, c'est-à-dire lorsqu'un nombre de composantes suffisant est utilisé. Dans la pratique, on observe de bons résultats pour de nombreux instruments à l'exception notable des instruments à cordes frottées joués avec des effets de *vibrato*.

Dans la pratique, nous utiliserons les deux premières méthodes, la première nous permettant d'obtenir un spectre de bruit "continu" synthétisable par les méthodes proposées en section 3.2.3, la seconde permettant de détecter la présence de composantes sinusoïdales fortement inharmoniques qui seraient mieux resynthétisée à l'aide d'oscillateurs supplémentaires<sup>6</sup>.

La synthèse de la composante bruitée a longtemps été considérée comme secondaire. L'écoute comparative des sons originaux, de leur composante harmonique et de leur composante bruitée, nous a fait constater que cette composante participait de diverses manières au timbre. La composante bruitée est notamment un élément de dynamique important pour les sons de cordes frappées ou grattées. Elle joue un rôle plus secondaire et plus variable — selon les modes de jeu — pour les autres instruments.

### 3.2.2 Analyse de la composante bruitée

L'analyse de la composante bruitée que nous décrivons ici reprend le schéma que nous avons suivi au chapitre 2 : analyse spectrale à court terme et réduction de données. La principale différence réside dans le caractère "continu"<sup>7</sup> du spectre de bruit.

#### Analyse spectrale de la composante bruitée

L'audition révèle deux catégories importantes de composantes bruitées :

- les composantes entretenues provenant d'une source d'excitation (souffle du flûtiste) ;
- les composantes impulsionnelles résultant d'impacts (bruit de marteau d'un piano).

Ces deux types de signaux ont en commun de nécessiter une faible résolution spectrale et une bonne résolution temporelle. Dans un souci de simplicité et de cohérence avec le reste

---

6. Ne disposant que d'échantillons provenant d'un cédérom de l'université McGill, il nous est difficile de savoir si de telles composantes — observées pour certains sons de piano — sont caractéristiques de l'instrument ou sont des artefacts d'enregistrement.

7. par opposition au spectre de "raies" des harmoniques.

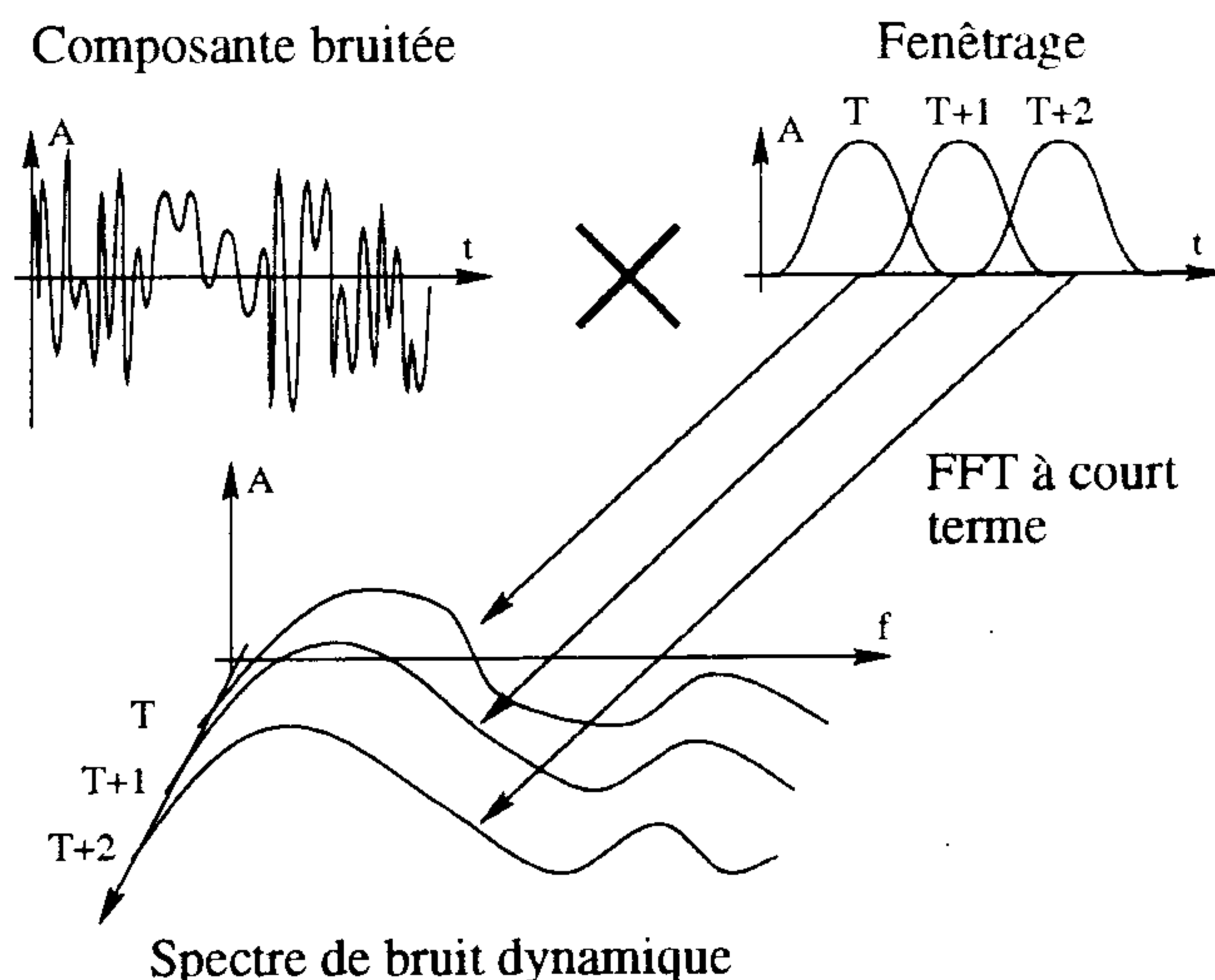


FIG. 3.16 – Analyse de la composante bruitée par transformées de Fourier à court terme

du mémoire, nous avons choisi la transformée de Fourier à court terme comme base de notre analyse. D'autres méthodes, telles que les transformées en ondelettes —, sont sans doute plus adaptées à ce type de signaux en raison de leur résolution temps/fréquence.

La figure 3.16 résume notre procédé d'analyse spectrale. Il est basé sur une série de transformées de Fourier à court terme d'une longueur de 128 points (soit environ trois millisecondes à 44100Hz), avec un taux de recouvrement d'un demi (soit environ 1,5 millisecondes).

D'autres traitements — tels qu'un lissage et un sous-échantillonnage du spectre — visant à réduire le nombre de données et à simplifier l'analyse en composantes principales sont envisageables.

### Analyse en composantes principales de la composante bruitée

L'analyse en composantes principales des 64 canaux de la transformée de Fourier est effectuée d'une façon similaire à celle illustrée au chapitre 2 : nous ne tenons compte que du spectre d'amplitude et nous ne tenons pas compte du spectre de phase.

Ouvrons une parenthèse en précisant que cette exclusion du spectre de phase est justifiée dans le cas des composantes bruitées "entretenues", elle est beaucoup plus discutable dans le cas des composantes bruitées impulsionnelles pour lesquelles la synchronicité des phases peut être importante<sup>8</sup>. Une extension du procédé décrit ici devra donc être envisagée dans l'avenir.

Le résultat de l'analyse en composantes principales est un sous-espace de spectres de bruit  $\mathcal{S}$  et une trajectoire de bruit  $\mathcal{B}$ . Les spectres principaux de bruit  $S_1, S_2, \dots$  définissent  $\mathcal{S}$ .  $S_1$  représente le spectre de bruit d'énergie moyenne, son enveloppe  $B_1(t)$ , représente donc approximativement l'enveloppe du spectre de bruit. Les spectres  $S_2$  et suivants représentent les déviations du spectre de bruit.

8. En effet, des déphasages aléatoires conduisent à un "étalement" du bruit d'impact similaire à l'effet produit par une courte réverbération.

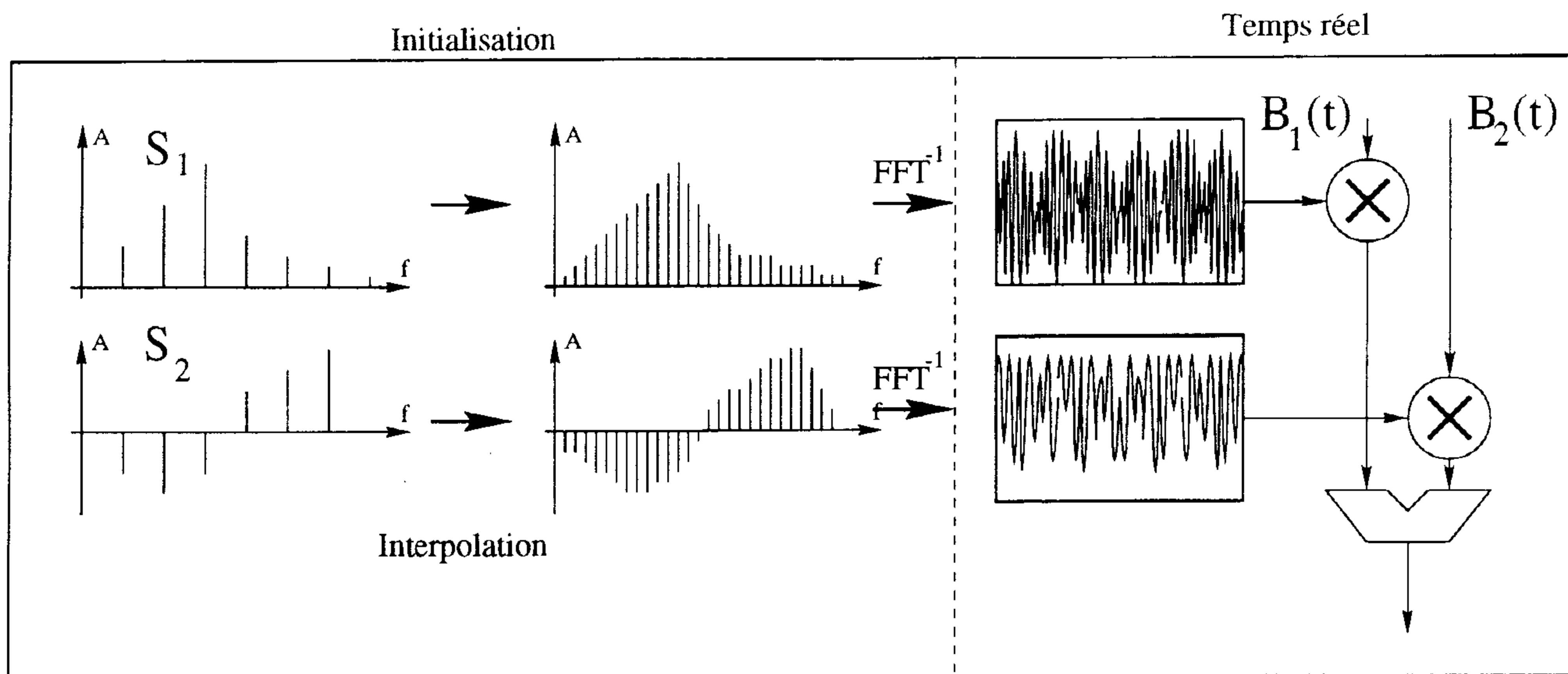


FIG. 3.17 – Synthèse de la composante bruitée par combinaison de tables

### 3.2.3 Synthèse de la composante bruitée

La synthèse de la composante bruitée à partir des données spectrales réduites doit remplir deux conditions : elle doit être économique en calculs, et, elle doit être de bonne qualité ou plus précisément éviter des artefacts tels que la périodicité et le “*phasing*”.

#### Synthèse de la composante bruitée par combinaison de tables

Cette méthode, décrite en figure 3.17, est similaire au procédé employé pour la synthèse de la composante harmonique. La principale différence étant que nous ne voulons pas synthétiser un spectre de raies mais un spectre continu.

Pour approcher ces conditions, il faut considérer un spectre discret disposant d'un nombre de canaux importants. Les amplitudes de ces canaux sont obtenues par interpolation des spectres de données réduites et leurs phases sont choisies aléatoirement (chaque spectre de bruit devra disposer du même jeu de phases aléatoires afin de permettre leur combinaison).

Le résultat de la transformée de Fourier inverse est une série de tables de bruit de durée importante. En effet, un artefact indésirable serait d'entendre une périodicité dans un signal bruité entretenu. Dans la pratique, une durée de table de l'ordre de la seconde est nécessaire. Cette durée correspond à une table de 44100 échantillons. En conclusion, cette méthode permet une première approche de la synthèse de la composante bruitée mais n'est pas satisfaisante en raison de la taille des tables nécessaires.

#### Synthèse de la composante bruitée par combinaison de tables et modulation d'amplitude

Le problème de la méthode précédente réside dans la dimension des tables nécessaires à la resynthèse. Une solution à ce problème pourrait être d'utiliser des tables d'ondes de courte durée modulées par une source de bruit à bande limitée.



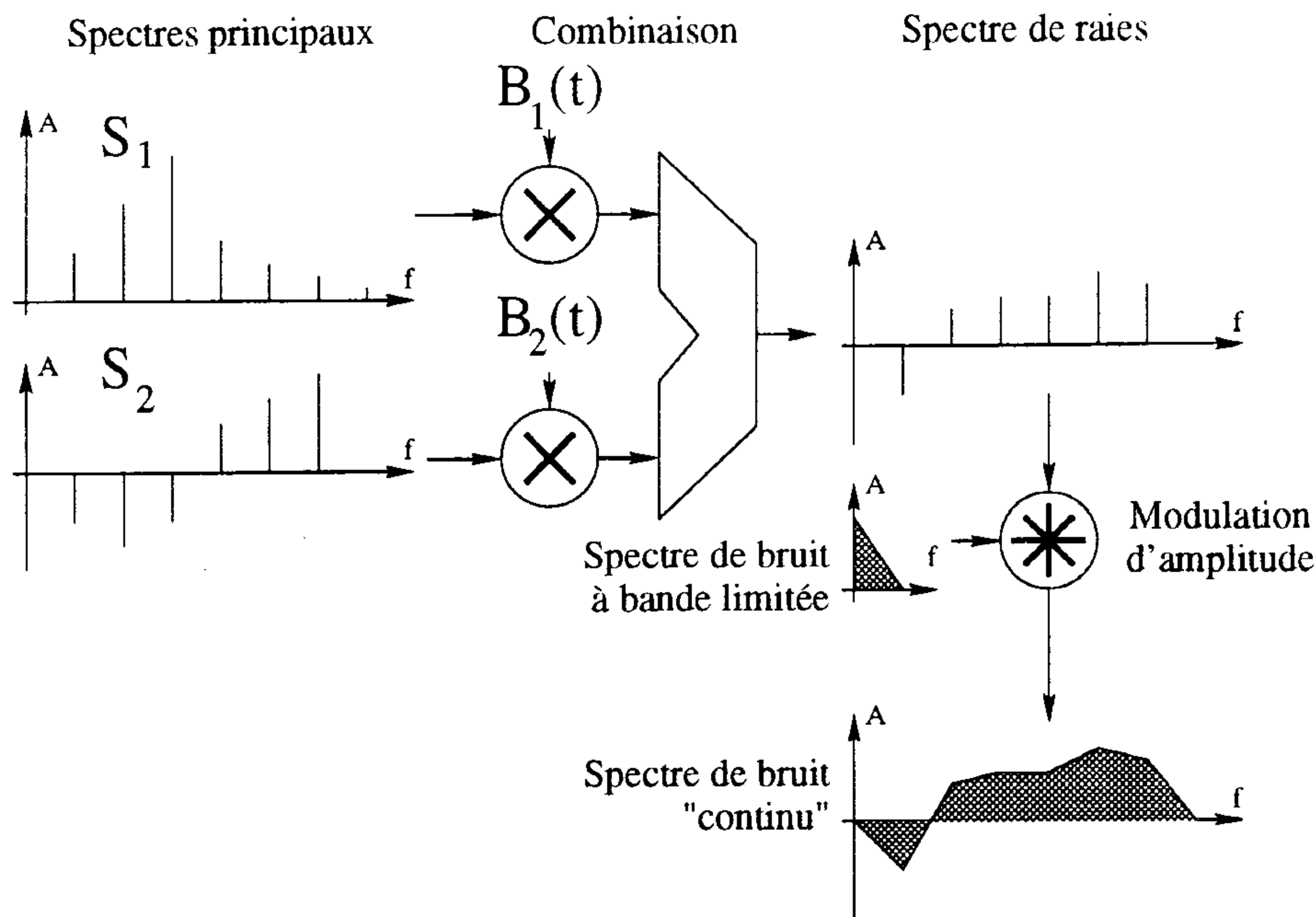


FIG. 3.18 – Synthèse de la composante bruitée par combinaison de tables et modulation d'amplitude

Le procédé proposé est décrit en figure 3.18 : les raies spectrales sont "étalées" par convolution selon le spectre du bruit à bande limitée. On obtient ainsi l'équivalent d'un spectre de bruit continu avec des tables d'ondes de faible longueur.

Hélas, cette méthode provoque des effets indésirables comparables à un effet de "phasing"<sup>9</sup> employé sur un signal bruité. La superposition des spectres latéraux engendre des interférences qui se traduisent par un filtrage en peigne caractéristique de l'effet de "phasing".

### Synthèse de la composante bruitée par combinaison de grains à temps aléatoire

Les deux solutions précédentes ne donnent pas entière satisfaction, car l'une nécessite des tables de taille considérable, l'autre produit des effets indésirables. La solution que nous envisageons ici est une extension de la synthèse par FOF à temps aléatoire développée par Christophe d'Alessandro [RdG92, RdG93]. Nous n'avons, jusqu'ici, procédé qu'à un nombre très limité d'essais qui nous ont semblés prometteurs.

Rappelons, dans un premier temps, le principe de la synthèse par FOF à temps aléatoire. La synthèse par FOF décrite en section 1.2.6, page 38, est étendue à un signal d'excitation virtuel d'impulsions à temps aléatoire. Ce signal d'excitation virtuel est assimilable à un "bruit de grenaille"<sup>10</sup>. Selon la densité du bruit de grenaille, il est possible d'obtenir des bruits à caractère impulsionnel aléatoire (bruit d'applaudissements ou d'ondées) ou des bruits "continus". L'obtention de bruits continus nécessite une densité

9. Le "phasing" est un effet de modulation de phase réalisé à l'aide d'une série de filtres passe-tout modulés par un oscillateur basse fréquence.

10. Aussi dénommé "shot-noise", ce type de processus aléatoires est rencontré, par exemple, dans l'étude de flux d'électrons dans des tubes à vide.

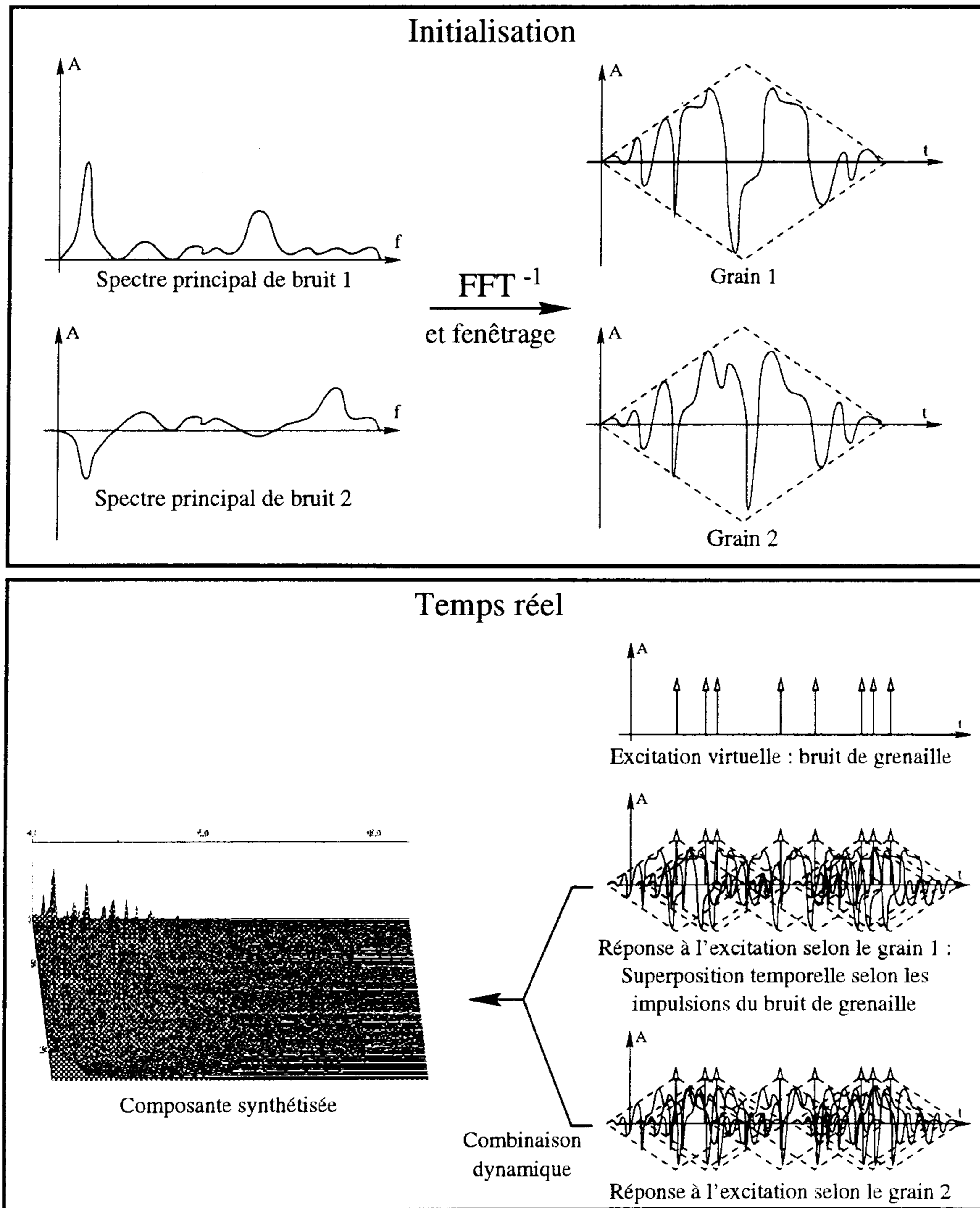


FIG. 3.19 – Synthèse de la composante bruitée par combinaison de grains à temps aléatoire

supérieure à 10 impulsions par période équivalente à la FOF, c'est-à-dire à la durée de la réponse impulsionnelle du filtre correspondant<sup>11</sup>. Le signal résultant de la superposition des FOF peut alors être assimilé à un bruit filtré par le filtre correspondant à la FOF.

Nous proposons ici d'étendre ce procédé dans deux directions :

- remplacer la FOF par un "grain" de forme plus élaborée ;
- combiner plusieurs grains afin d'obtenir un spectre de bruit évolutif.

Le spectre résultant de ce type de synthèse est le produit du spectre du grain par le spectre du bruit de grenaille. Si la densité du bruit de grenaille est suffisante, son spectre peut être considéré comme uniforme. Le profil spectral du son résultant est alors déterminé par la transformée de Fourier inverse du grain utilisé.

Nous allons exploiter cette propriété en utilisant pour grain une combinaison des tables de bruit évoquées précédemment.

Ce procédé est résumé en figure 3.19. Les quelques essais que nous avons menés grâce au logiciel CSOUND[Ver92] nous ont conduits à des résultats de bonne qualité, ne montrant pas d'aspects périodiques ni d'effet de *phasing*.

Il est difficile de conclure sur un procédé que nous ne faisons qu'aborder. Il semble cependant qu'à l'inverse des autres méthodes, celle-ci soit capable de restituer fidèlement des bruits d'impacts de faible durée. Cette capacité peut être expliquée par le fait que le grain est assimilé à une réponse impulsionnelle de nature proche des bruits d'impacts<sup>12</sup>. Une analyse, tenant partiellement compte des phases — ou plus exactement de leur synchronicité au moment de l'impact — permettrait d'exploiter pleinement cette forme de synthèse.

En dernier lieu, il convient d'observer les problèmes éventuels d'implantation et de coût de calcul. En effet, cette forme de synthèse nécessite la superposition — c'est-à-dire la lecture simultanée — d'un nombre de grains variable. Deux problèmes d'implantations se posent :

- la gestion d'un nombre de superpositions variable induisant un coût de calcul variable ;
- un nombre considérable (supérieur à dix) de superpositions nécessaire à la synthèse de bruits à caractère continu.

### 3.2.4 Conclusions

Les trois méthodes de synthèse que nous proposons ne sont pas encore satisfaisantes :

- la synthèse par combinaison de tables nécessite des tables de longueur conséquente ;
- la synthèse par combinaison de tables et modulation d'amplitude produit des effets indésirables de *phasing* ;

---

11. Le signal d'excitation virtuelle peut alors être assimilé à un bruit gaussien [Rd96].

12. Dans le cas du piano, par exemple, le bruit d'impact pourrait être assimilé à la réponse impulsionnelle des éléments autres que la corde, cette réponse impulsionnelle est sans doute très complexe mais déterministe.

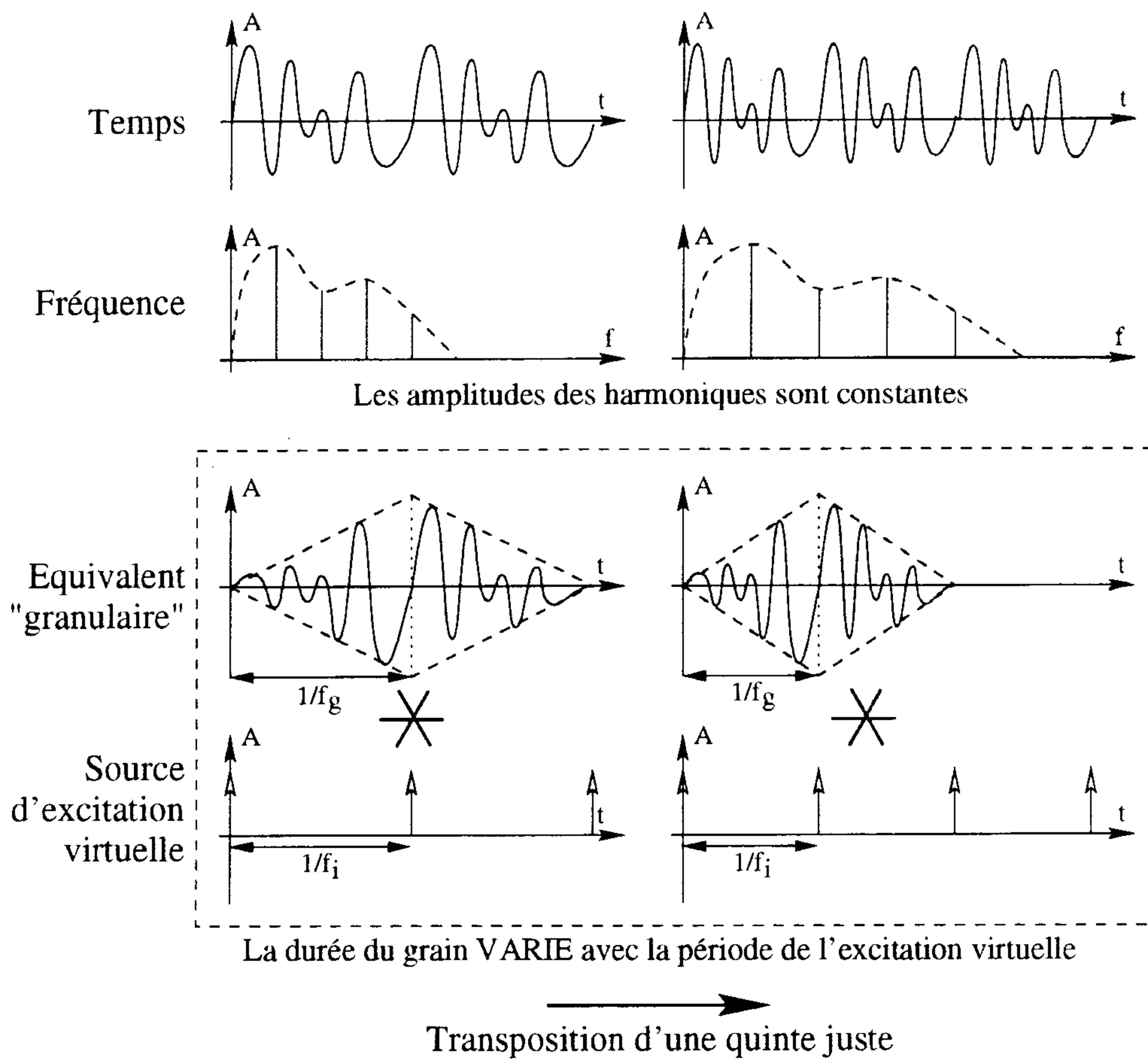


FIG. 3.20 – *Interprétation granulaire de la transposition d'une table d'onde*

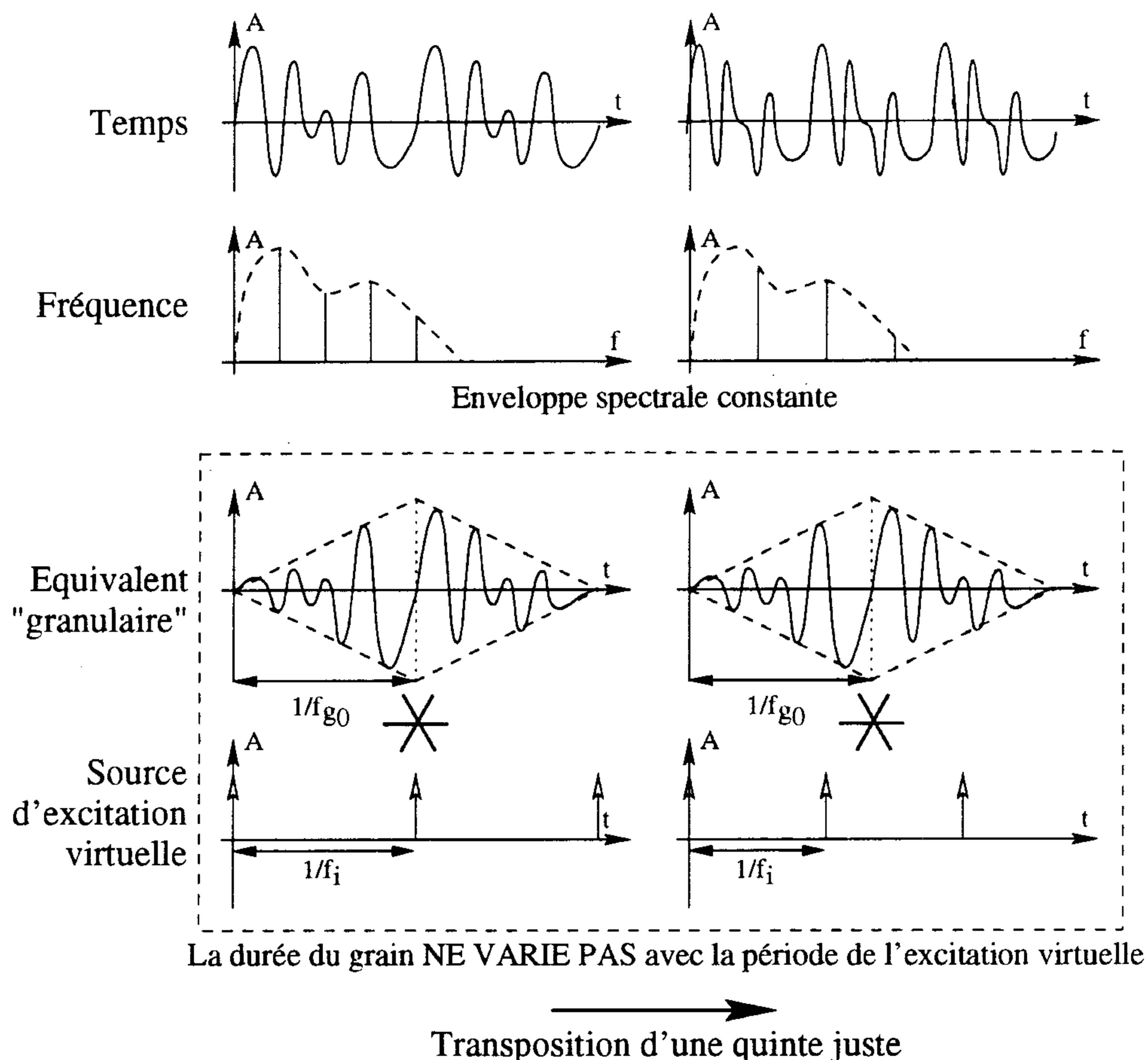
- la synthèse par combinaison de grains à temps aléatoire est trop coûteuse en calculs pour les bruits continus.

De nouvelles études — combinant plusieurs techniques et/ou incluant d'autres techniques (telles que la lecture aléatoire de tables [DD82]) — devront être menées dans l'avenir afin d'obtenir une méthode satisfaisante en termes de qualité et de coût de calcul.

### 3.3 Synthèse par combinaison de grains généralisée

Le procédé que nous présentons dans cette section est la synthèse de plusieurs procédés : synthèse par combinaison de tables d'ondes, synthèse granulaire, et, synthèse par FOF à temps aléatoire.

La section 3.3.1 propose un lien entre forme d'onde, grains et réponse impulsionnelle. La section 3.3.2 discute et généralise différents procédés de transposition. La section 3.3.3 généralise la notion de combinaison de tables d'ondes. Enfin, la section 3.3.4 établit une relation entre ce nouveau modèle et la synthèse de bruit par combinaison de grains à temps aléatoire.

FIG. 3.21 - *Transposition granulaire d'une "table d'onde"*

### 3.3.1 Dualité grain/table d'onde

La figure 3.20 représente la lecture de table d'onde comme un cas particulier de synthèse granulaire. Le grain considéré est composé de deux périodes de la table dotées d'une fenêtre triangulaire. La lecture de table d'onde est assimilée à la superposition de tels grains. L'enveloppe triangulaire assure un enchaînement correct des tables d'ondes. Notons qu'une pondération de chaque "grain" reviendrait — grâce au fenêtrage triangulaire — à faire varier l'amplitude du signal de façon linéaire<sup>13</sup>. La fréquence de lecture des grains  $f_g$  est égale à la fréquence des impulsions d'excitation virtuelles  $f_i$ . Le respect de cette égalité lors de la transposition se traduit par une invariance des amplitudes des harmoniques.

### 3.3.2 Transposition et enveloppe spectrale

La figure 3.21 montre une transposition "granulaire". La fréquence de lecture du grain  $f_g$  reste constante à une valeur de référence  $f_{g0}$  quelle que soit la fréquence des impulsions d'excitation virtuelles  $f_i$ . Le signal synthétisé par ce procédé est égal au produit de convolution du signal d'excitation et du grain (considéré comme la réponse impulsionnelle d'un filtre). Dans le domaine spectral, on obtient le produit d'un spectre de raies et de

13. Ce procédé a été évoqué au §1.2.1, page 33.

l'enveloppe spectrale correspondant au grain. Etant donné que la fréquence de lecture du grain est constante, l'enveloppe spectrale qui lui est associée<sup>14</sup> est restée constante. Cette forme de transposition permet donc de conserver l'enveloppe spectrale globale d'un son. Cette méthode a déjà avantageusement été utilisée pour transposer des sons musicaux en limitant les distorsions de timbre [Len89].

Nous allons maintenant généraliser les deux types de transpositions exposées ci-dessus en introduisant un facteur de transposition d'enveloppe spectrale  $F_T$  :

- $F_T = 1$ ,  $f_g = f_i$ , transposition de type "tables d'ondes", l'enveloppe spectrale "suit" les harmoniques ;
- $F_T = 0$ ,  $f_g = f_{g0}$ , transposition de type "granulaire", l'enveloppe spectrale reste constante.

La fréquence de lecture du grain peut être définie par :

$$f_g = F_T(f_i - f_{g0}) + f_{g0} \quad (3.4)$$

Des valeurs intermédiaires de  $F_T$  correspondent à des suivis d'enveloppes spectrales intermédiaires.  $F_T$  n'est pas sans rappeler le paramètre "filter tracking" disponible sur de nombreux synthétiseurs à synthèse soustractive.

### 3.3.3 Combinaison de grains/tables d'onde

Nous allons maintenant exploiter les notions présentées dans les paragraphes précédents dans le cadre de la synthèse par combinaison de tables d'ondes.

Pour cela, il suffit de remplacer le grain obtenu par fenêtrage de deux périodes d'une table d'onde par un grain obtenu par combinaison de plusieurs tables. Le cas  $F_T = 1$  nous ramène trivialement au modèle que nous exploitons depuis le chapitre 2.

Le cas  $F_T = 0$ , quant à lui, nous conduit à un nouveau modèle : la combinaison ne correspond plus à des variations du profil harmonique mais à des variations de l'enveloppe spectrale globale. Ainsi, pour exploiter pleinement ce nouveau modèle, il faudrait établir une nouvelle procédure d'analyse et de réduction de données ayant pour objet l'enveloppe spectrale globale du son par opposition aux amplitudes des harmoniques. Une telle méthode n'a pas encore été expérimentée mais nous pensons qu'elle peut nous conduire à des représentations plus fidèles pour certaines catégories de sons musicaux.

Les cas intermédiaires sont plus délicats à interpréter. Néanmoins, nous pensons qu'il serait digne d'intérêt d'établir un procédé permettant de déterminer une valeur  $F_T$  optimale pour un son considéré.

Cette extension est un exemple de la fécondité des méthodes d'analyse-synthèse : dans un premier temps, une méthode d'analyse nous a conduit à une méthode de synthèse efficace, dans un second temps, la forme que prend la synthèse nous suggère un nouveau modèle plus complet nécessitant à son tour de nouveaux outils d'analyse.

---

14. Elle est égale au produit de convolution du spectre de raies par la transformée inverse de la fenêtre, dans notre cas un sinus cardinal au carré.

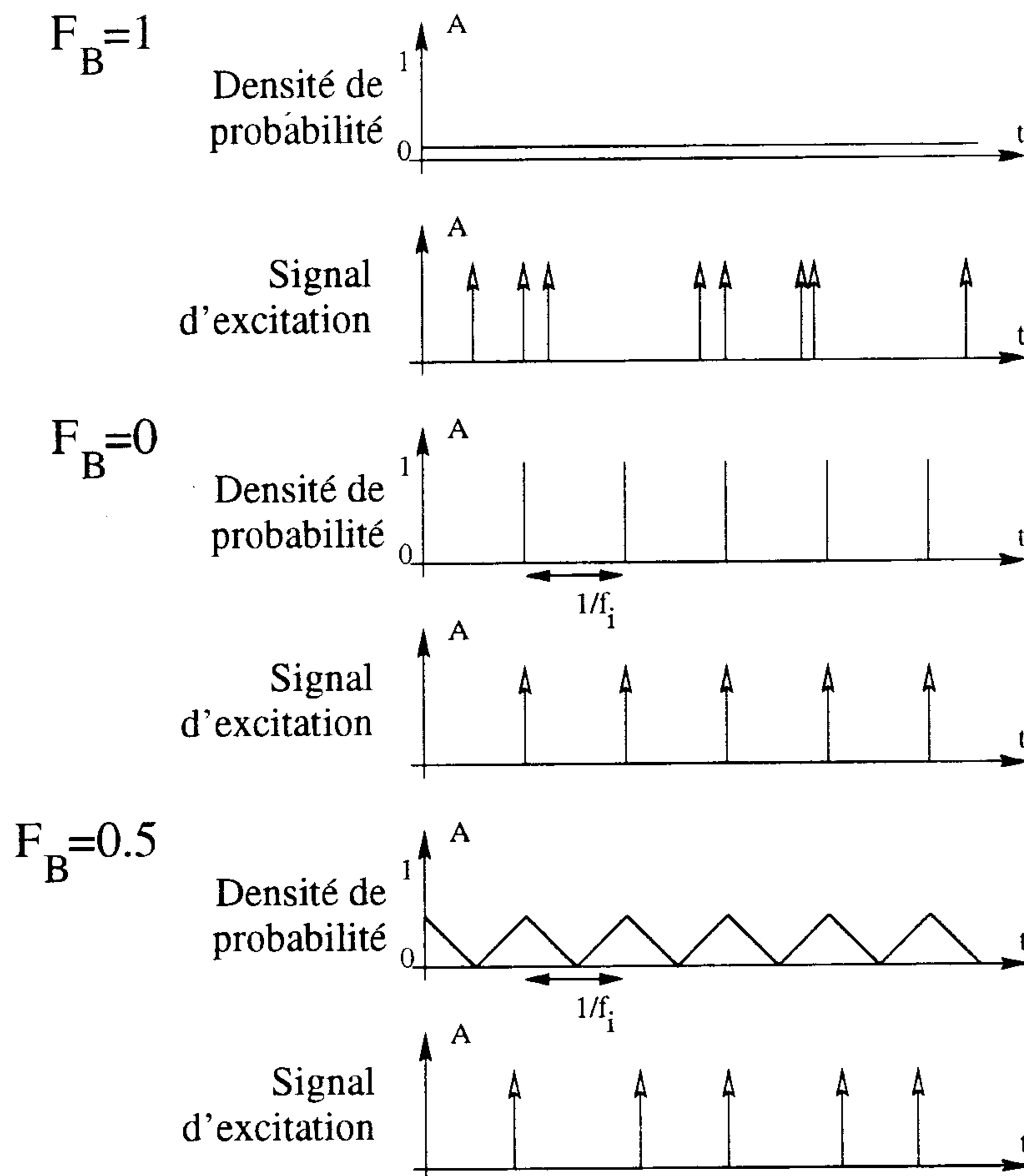


FIG. 3.22 – Généralisation de la synthèse par grains à temps aléatoire

### 3.3.4 Extension, conclusions et perspectives

Nous poussons le principe de généralisation de la méthode de synthèse un petit peu plus loin en y ajoutant un facteur de bruit  $F_T$ . Nous avons déjà évoqué en section 3.2.3 la synthèse de composante bruitée par combinaison de grains à temps aléatoire. Il nous semble naturel de profiter du lien que nous venons d'établir entre tables et grains pour y inclure les grains à temps aléatoire.

Pour ce faire, il suffit de considérer le signal d'excitation déclenchant la lecture des tables/grains de la façon suivante :

- $F_B = 0$ , signal d'excitation déterministe dont la fréquence est fixée par  $f_i$  ;
- $F_B = 1$ , la probabilité d'excitation d'un grain est uniforme.

La figure 3.22 suggère une implantation possible permettant de gérer  $F_B$  mettant en jeu une forme de densité de probabilité périodique dont la forme varie selon  $F_B$ . Dans les applications pratiques, nous pensons que  $F_B$  doit être considéré comme variable au cours du temps afin de restituer les bruits présents durant certains transitoires.

Cette dernière section propose de fusionner la synthèse par lecture de tables, la synthèse granulaire, et, la synthèse de la composante bruitée. Cette nouvelle forme de

synthèse n'est encore qu'un projet et fait l'objet de diverses expérimentations visant à en observer les propriétés. Nous pensons cependant que les outils d'analyse et de synthèse que nous avons développé au cours de cette thèse pourront être étendus et exploités dans ce nouveau cadre.

### 3.4 Conclusions et perspectives

Les trois extensions que nous avons présenté dans ce chapitre sont trois nouvelles voies à explorer.

L'étude et la synthèse des sons par analyse en composantes principales complexes nous apporte des éléments d'informations nouveaux sur les phases des spectres principaux. L'interprétation de cette information de phase est difficile dès lors qu'on analyse des sons relativement compliqués. Elle mérite, à notre avis, d'être étudiée de façon plus systématique et plus approfondie. Ses applications à la synthèse pourrait être importantes si l'on s'accorde sur l'importance "perceptive" des phases relatives — et de leurs variations — pour certaines catégories de sons.

L'analyse et la synthèse de la composante bruitée sont elles aussi des techniques à développer. La composante "bruitée" — trop longtemps considérée comme un élément parasite et non musical — joue un rôle plus ou moins important selon les instruments. Pour une catégorie d'instruments, la composante bruitée correspond à un bruit d'impact, très localisé dans le temps. Pour une autre catégorie d'instruments, elle correspond à une source continue d'excitation (archet, souffle). Une méthode d'analyse-synthèse efficace devra pouvoir tenir compte de ces deux aspects et être relativement économique en calculs. La synthèse par combinaison de grains à temps aléatoire nous semble très prometteuse car elle permet de synthétiser convenablement les bruits "impulsionnels" et les bruits "continus".

La synthèse par combinaison de grains — que nous avons présenté en fin de chapitre — regroupe plusieurs types de synthèses : synthèse granulaire, synthèse par tables d'ondes, synthèse par combinaison de tables et synthèse par grains à temps aléatoires. Par rapport à la synthèse par combinaison de tables d'ondes du chapitre 2, elle fait apparaître deux nouveaux paramètres :

- le facteur de transposition  $F_T$  qui permet de passer continuellement de la synthèse par tables à la synthèse par grains ;
- le facteur de bruit  $F_B$  qui permet de passer continuellement du temps déterministe au temps aléatoire.

A l'heure actuelle, nous ne disposons pas de procédé automatique permettant d'extraire ces deux nouveaux paramètres de synthèse. L'étude d'un tel procédé d'analyse permettrait d'utiliser un unique procédé de synthèse pour synthétiser efficacement une très grande variété de sons musicaux.



## Conclusions et perspectives

Le travail présenté s'est attaché dans un premier temps à exploiter l'analyse en composantes principales (ACP) dans le contexte de l'analyse-synthèse des sons musicaux.

Cette première étude ne considère que la composante pseudo-périodique des sons musicaux accordés. Cette composante est décomposée en une somme de signaux sinusoïdaux harmoniques d'amplitudes — variables au cours du temps — grâce à un procédé d'analyse spectrale fondée sur l'interpolation de transformée de Fourier à temps discret. La masse importante de données spectro-temporelles est ensuite réduite par ACP. Le résultat de l'ACP est une représentation optimale — au sens de la consommation de variance des axes ordonnés — du son musical considéré.

L'ACP, parce qu'elle réduit les informations, est très sensible à la “qualité” des données qui lui sont fournies. Nous pensons que les conclusions [HBH93]<sup>15</sup> des travaux menés par Andrew Horner, James Beauchamp et Lippold Haken sur l'efficacité de l'ACP sont dues pour l'essentiel à un “simple” problème de normalisation des données spectrales. Ils ont notamment constaté des variations importantes de l'amplitude du signal d'erreur<sup>16</sup>. En fait, ces artefacts, produisant des effets indésirables pendant les transitoires, disparaissent lorsque les profils spectraux sont normalisés : l'ACP porte alors son “attention” aux variations spectrales et non aux amplitudes absolues des harmoniques.

Outre cette constatation, l'originalité de nos travaux réside dans le souci d'interprétation que nous avons porté aux données réduites. En effet, l'équipe précédemment citée n'a pas poussé plus avant ses investigations<sup>17</sup>, son objectif essentiel étant de réduire le coût de la synthèse par combinaison de tables d'ondes. L'interprétation des données réduites que nous avons menée s'articule autour de quatre méthodes :

- l'interprétation géométrique de l'ACP, qui nous donne une idée de l'opération réalisée ;
- l'étude de données réduites de sons synthétiques, qui nous permet de vérifier et de compléter les assertions faites précédemment ;
- l'étude de données réduites de sons acoustiques et leur confrontation aux résultats de l'expérience psychoacoustique de John M. Grey, qui nous permet d'envisager notre procédé dans le cadre de l'étude des sons musicaux ;

15. [HBH93, page 353] : *“In conclusion, though statistically optimal, PCA should by no means be regarded as the best matching method.”*

16. [HBH93, page 348] : *“Also note that the error is consistently low in the sustain portion of the tone, but is much higher during the attack and decay.”*

17. [HBH93, page 342] : *“On the other hand, these spectra may have a rather artificial relation to the original tone, since generally no basis spectrum will resemble any of the actual analysis spectra.”*

- enfin, l'utilisation d'un procédé de synthèse par combinaison de tables d'ondes permettant de valider l'efficacité de la représentation en matière de synthèse et d'édition du son (synthèse croisée).

Les interprétations des données réduites nous ont permis de surpasser certaines difficultés et d'envisager leur utilisation pour la synthèse de sons expressifs, c'est-à-dire tenant compte des nuances de jeu. A ce sujet, nous suggérons la possibilité d'établir un lien entre des "méta-paramètres" de jeu et les données réduites. De telles possibilités ont été expérimentées et ont conduit à des résultats intéressants.

Cette première série de travaux nous a permis de dépasser les conclusions de Andrew Horner, James Beauchamp et Lippold Haken. Nous disposons donc d'une méthode efficace pour analyser et synthétiser des sons musicaux harmoniques. Cette représentation souffre cependant de plusieurs défauts :

- elle est incapable de représenter des sons inharmoniques ;
- elle ne tient pas compte de la composante bruitée des sons dont l'importance est primordiale notamment en ce qui concerne les transitoires ;
- enfin, certaines classes de sons, tels que les sons d'instruments à cordes frottées, sont fort mal représentées par notre méthode.

Le second volet de nos travaux est une série d'extensions visant à tenir compte de ces éléments.

Constatant que notre étude préliminaire et que l'étude de Andrew Horner, James Beauchamp et Lippold Haken éludaient le problème des phases<sup>18</sup>, nous avons poussé notre étude en la généralisant aux nombres complexes. Cette nouvelle forme d'analyse, fondée sur la notion d'amplitude harmonique complexe<sup>19</sup>, permet de tenir compte de légères inharmonicités sous forme de rotation de phase. De plus, cette représentation permet une resynthèse tenant strictement compte de la phase de chaque composante ; elle rend donc possible la soustraction directe, dans le domaine temporel<sup>20</sup>, du son synthétique au son original afin d'en extraire la composante bruitée.

L'étape suivante est une évaluation de méthodes d'analyse et de synthèse permettant de synthétiser la composante bruitée. Dans un premier temps, nous procédons à une analyse spectrale fondée sur une série de transformées de Fourier à court terme. Dans un second temps, nous utilisons l'ACP pour réduire ces données spectrales. Enfin, nous devons opérer la resynthèse. Cette resynthèse pose quelques problèmes liés à la nature d'un spectre de bruit ; par opposition aux spectres de raies de la composante harmonique, il présente un spectre "continu". Nous avons établi que la resynthèse d'un tel spectre ne peut être convenablement réalisée par combinaison de tables d'ondes modulées ou non. Il nous semble qu'une méthode inspirée par la synthèse granulaire et la synthèse par FOF à

---

18. [HBH93, page 338]: "Note that the phases of the corresponding harmonics of multiple wavetables must be the same to avoid inadvertent phase cancellation."

19. Le module et l'argument d'une amplitude harmonique complexe correspondent à l'amplitude et à la phase relative de l'harmonique sous-jacent.

20. D'autres méthodes procèdent dans le domaine spectral[SS90, SS91].

temps aléatoire offre une fidélité de resynthèse exempt d'artefacts. Cette méthode souffre cependant d'un coût de calcul qui pourrait en limiter les applications temps-réel.

La dernière étape propose deux nouveaux paramètres de synthèse offrant de nouvelles possibilités.

Le premier paramètre, le facteur de transposition d'enveloppe spectrale  $F_T$ , affecte le comportement spectral du son lors de transpositions. Lorsque  $F_T = 1$ , on a affaire à une transposition de tables d'ondes : les amplitudes des harmoniques restent constantes. Lorsque  $F_T = 0$ , on obtient une transposition "granulaire" : l'enveloppe spectrale globale du son reste constante. Les valeurs intermédiaires permettent de rendre compte de types de transposition intermédiaires. Nous pensons que la transposition "granulaire" est préférable pour les sons dans lesquels les phénomènes "formantiques" sont déterminant tels que les instruments à cordes frottés. Cependant, nous n'avons pas encore mis au point de procédé automatique permettant d'extraire ce nouveau paramètre. Nous suggérons, à ce propos, l'utilisation d'échantillons plus nombreux et enregistrés dans des conditions précises.

Le second paramètre, le facteur de bruit  $F_B$ , affecte la régularité de déclenchement de lecture des grains. Il rend possible la transition entre un son harmonique  $F_B = 0$  à un son bruité  $F_B = 1$ . Ce facteur permettrait de fusionner le procédé de synthèse évoqué ici et la synthèse de bruit par grains proposée précédemment. Bien que nous ne disposions pas de méthode d'extraction de ce paramètre, il nous semble évident qu'il devra être considéré comme variable au cours du temps permettant ainsi de restituer les transitoires.

D'une part, nos travaux se situent dans le cadre des méthodes d'extraction de paramètres de synthèse. Les enjeux de cette famille de méthodes prendront de l'importance dans l'avenir : un nombre croissant de synthétiseurs — basés sur des circuits puissants et souples — proposent de multiples méthodes de synthèse. Utiliser un tel synthétiseur suppose des connaissances croissantes des divers types de synthèse. Le problème que se pose un musicien soucieux d'obtenir un son donné est de choisir l'architecture<sup>21</sup>, et les paramètres permettant de le synthétiser convenablement.

Les méthodes d'extraction de paramètres de synthèse peuvent assister le musicien dans cette démarche et lui proposer diverses solutions. Notre méthode en est un bon exemple : elle propose un ensemble de tables d'ondes et une façon de les combiner afin d'approcher un son analysé. Comme nous l'avons suggéré à propos de la synthèse par modulation de fréquence, elle peut aussi servir d'intermédiaire dans un procédé d'extraction plus général : il suffirait pour ce faire de réaliser un "catalogue" de sous-espaces caractéristiques de diverses méthodes de synthèse et de projeter le son analysé dans des sous-espaces candidats.

D'autre part, nos travaux se situent dans le cadre des méthodes de synthèse par combinaison de tables d'ondes et/ou de grains dans sa version étendue. Les capacités propres à ce type de synthèse méritent d'être étudiées plus avant. Ses avantages potentiels étant :

- la généralité de la représentation ;
- un coût de calcul raisonnable ;

---

21. Rappelons que le DX7 en proposait déjà 32 en 1983.

- une simplicité d'utilisation liée à l'interprétabilité des paramètres de synthèse<sup>22</sup> ;
- la disponibilité d'outils d'analyse permettant de les extraire ;
- la possibilité d'établir des modèles de génération de paramètres dynamiques expressifs.

Estimer toutes ces capacités suppose un travail conséquent, l'utilisation et/ou le développement d'outils adaptés (carte de développement audio-numérique pour l'implantation temps-réel, langage de modélisation dédié), et, des collaborations avec d'autres laboratoires disposant d'équipements complémentaires aux nôtres (piano MIDI et dispositifs d'enregistrements).

---

22. Par opposition, par exemple, à la synthèse par modulation de fréquence.

## **Annexe B**

### **Fiches de données réduites de 42 sons**

Figure	Instrument	Note	Référence McGill	Page
B.1	Clarinete en si $\flat$	Do 4	ClarinetBbC4	153
B.2	Clarinete en mi $\flat$	Do $\sharp$ 4	ClarinetEbC#4	154
B.3	Hautbois	Ré 4	OboeD4	155
B.4	Hautbois baroque	Ré 4	OboebaroqueD4	156
B.5	Orgue baroque registre plenum	Mi 4	OrgBaroqplenumE4	157
B.6	Orgue baroque registre prinzipal	Ré 4	OrgBaroqprinzipalD4	158
B.7	Orgue baroque registre flûtes	Mi 4	OrgBaroqfluteE4	159
B.8	Flûte à bec basse	Ré 4	RecrdrgreatbassD4	160
B.9	Flûte traversière sans vibrato	Ré 4	FluteWovibD4	161
B.10	Flûte traversière avec vibrato	Ré 4	FluteWvibD4	162
B.11	Flûte alto sans vibrato	Ré 4	FlutealtoD4	163
B.12	Flûte alto avec vibrato	Do $\sharp$ 4	FlutealtoWvibC#4	164
B.13	Flûte jouée "flutter"	Ré 4	FlutflutterD4	165
B.14	Basson	Ré 4	BassoonD4	166
B.15	Cor Anglais	Do $\sharp$ 4	EnglishHornC#4	167
B.16	Cor	Ré 4	FrenchHornD4	168
B.17	Cor avec sourdine	Ré 4	FrenchHornmutedD4	169
B.18	Cornet	Do $\sharp$ 4	CornetttrebleC#4	170
B.19	Trombone	Do $\sharp$ 4	TrombonetenorC#4	171
B.20	Trombone avec sourdine	Do $\sharp$ 4	TrombonemutedC#4	172
B.21	Trompette baroque	Do $\sharp$ 4	TrumpetBachC#4	173
B.22	Trompette en Do	Ré 4	TrumpetD4	174
B.23	Tuba	Do $\sharp$ 4	TubaC#4	175
B.24	Celesta	Ré 4	CelestaD4	176
B.25	Marimba	Ré 4	MarimbaD4	177
B.26	Piano joué mezzo forte	Ré 4	PianomfD4	178
B.27	Piano joué fortissimo	Ré 4	PianoffD4	179
B.28	Guitare classique	Ré 4	ClassGtrD3	180
B.29	Guitare classique jouée sul tasto	Do $\sharp$ 4	ClassGtrsultastoC#3	181
B.30	Clavecin	Ré 4	Harpsichord4D3	182
B.31	Harpe	Mi 4	HarpE3	183
B.32	Alto	Ré 4	ViolabowedD4	184
B.33	Violon	Do $\sharp$ 4	ViolinbowedC#4	185
B.34	Violoncelle joué martele	Ré 4	CellomarteleD4	186
B.35	Alto joué martele	Ré 4	ViolamarteleD4	187
B.36	Violon joué martele	Do $\sharp$ 4	ViolinmarteleC#4	188
B.37	Violoncelle joué avec vibrato	Do $\sharp$ 4	CelloWvibC#4	189
B.38	Alto joué avec vibrato	Ré 4	ViolaWvibD4	190
B.39	Violon joué avec vibrato	Ré 4	ViolinWvibD4	191
B.40	Violoncelle joué pizzicato	Do 4	CellopizzC4	192
B.41	Alto joué pizzicato	Ré $\sharp$ 4	ViolapizzD#4	193
B.42	Violon joué pizzicato	Do $\sharp$ 4	ViolinpizzC#4	194

TAB. B.1 - Table des sons analysés

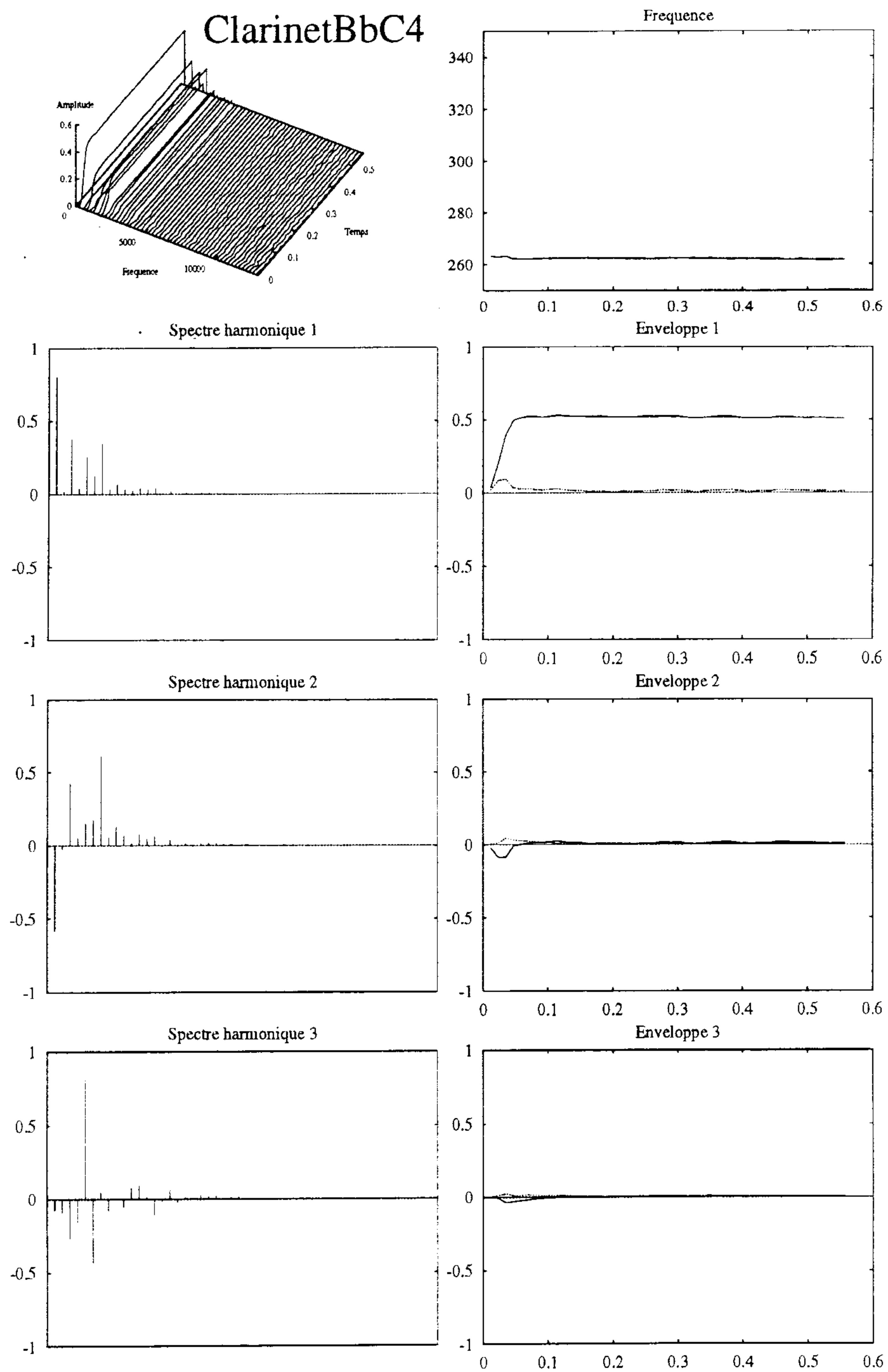


FIG. B.1 – Clarinette en si b

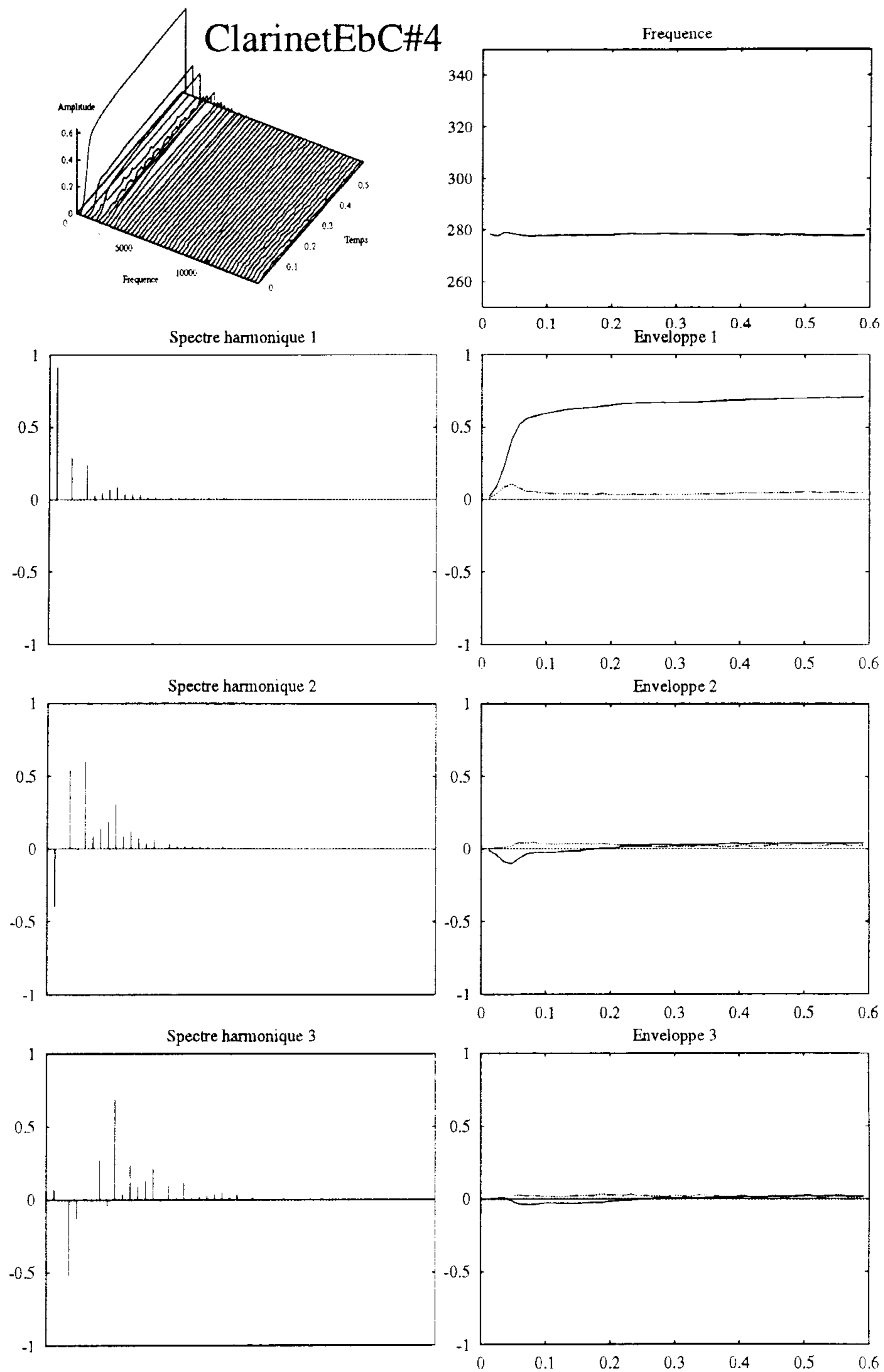
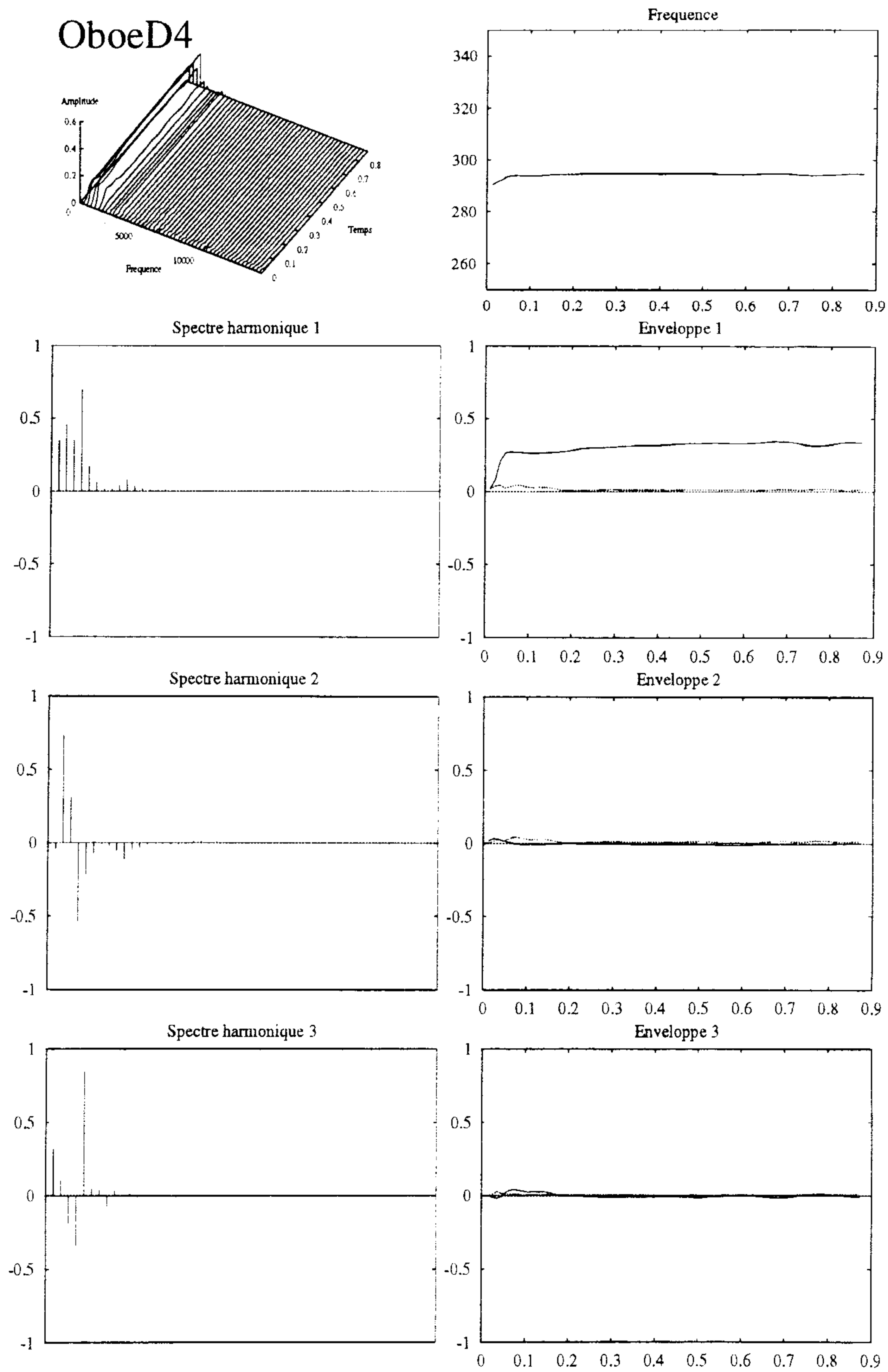


FIG. B.2 - Clarinette en mi b





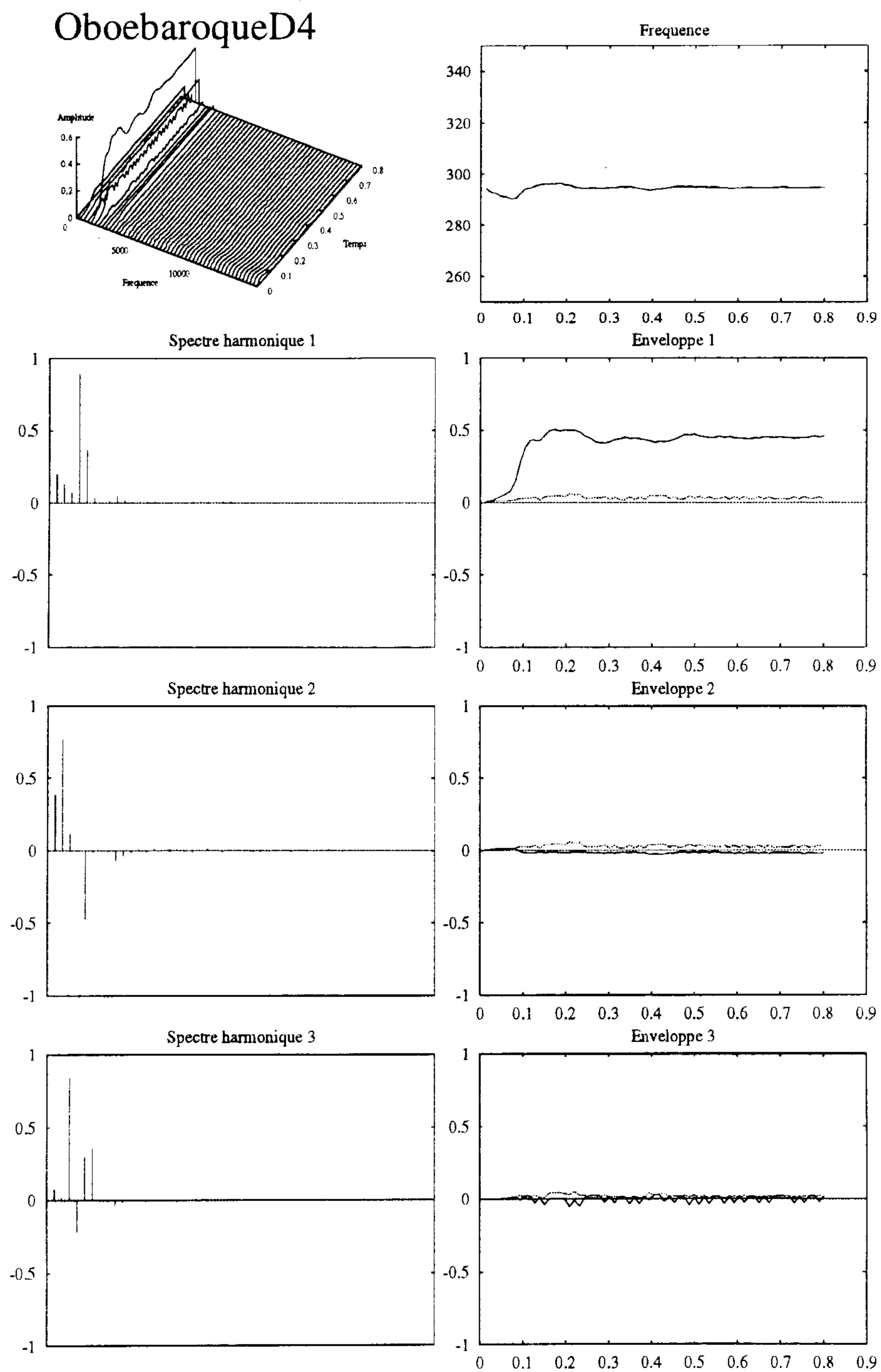


FIG. B.4 - Hautbois baroque

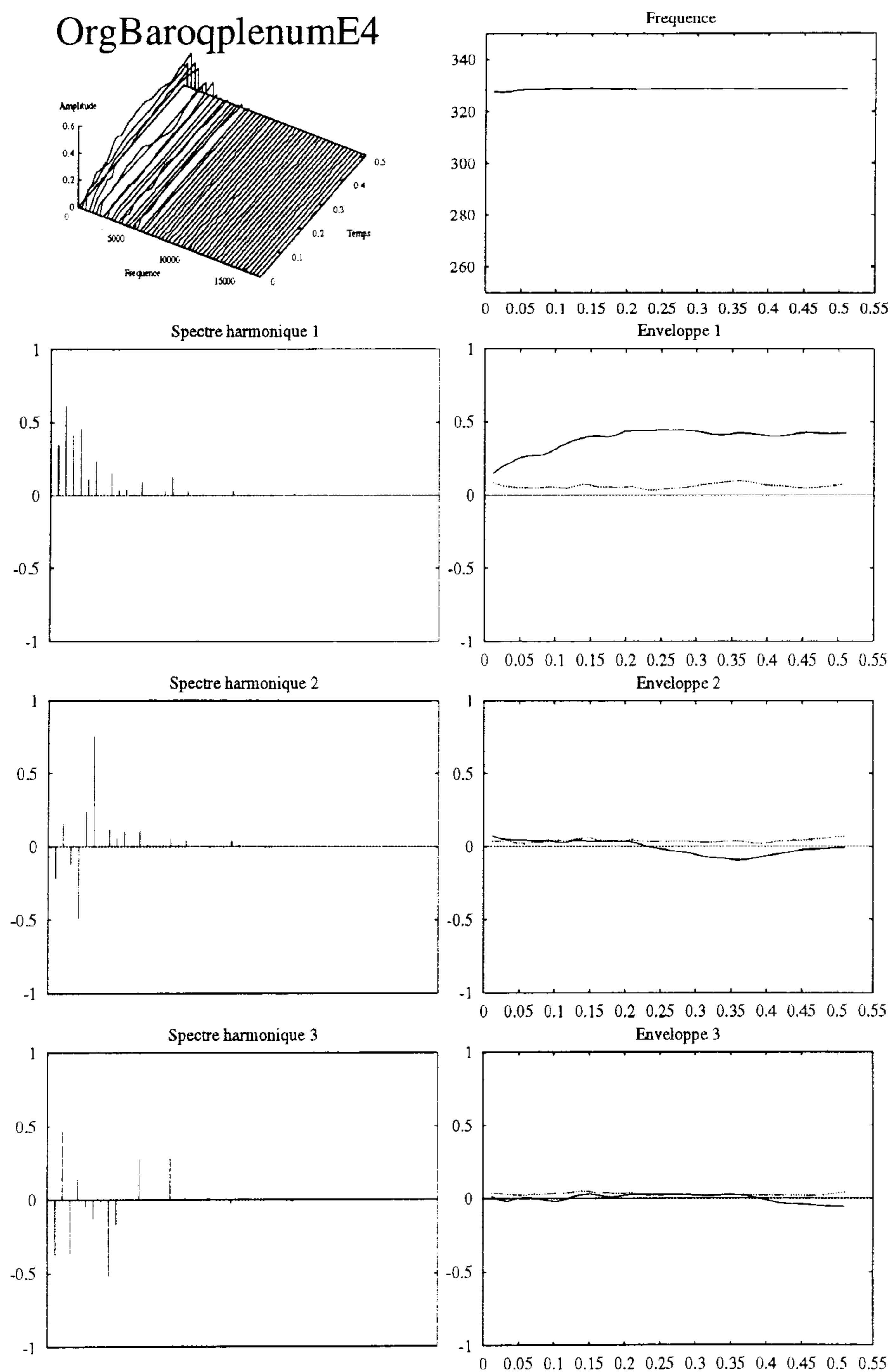


FIG. B.5 – Orgue baroque registre plenum

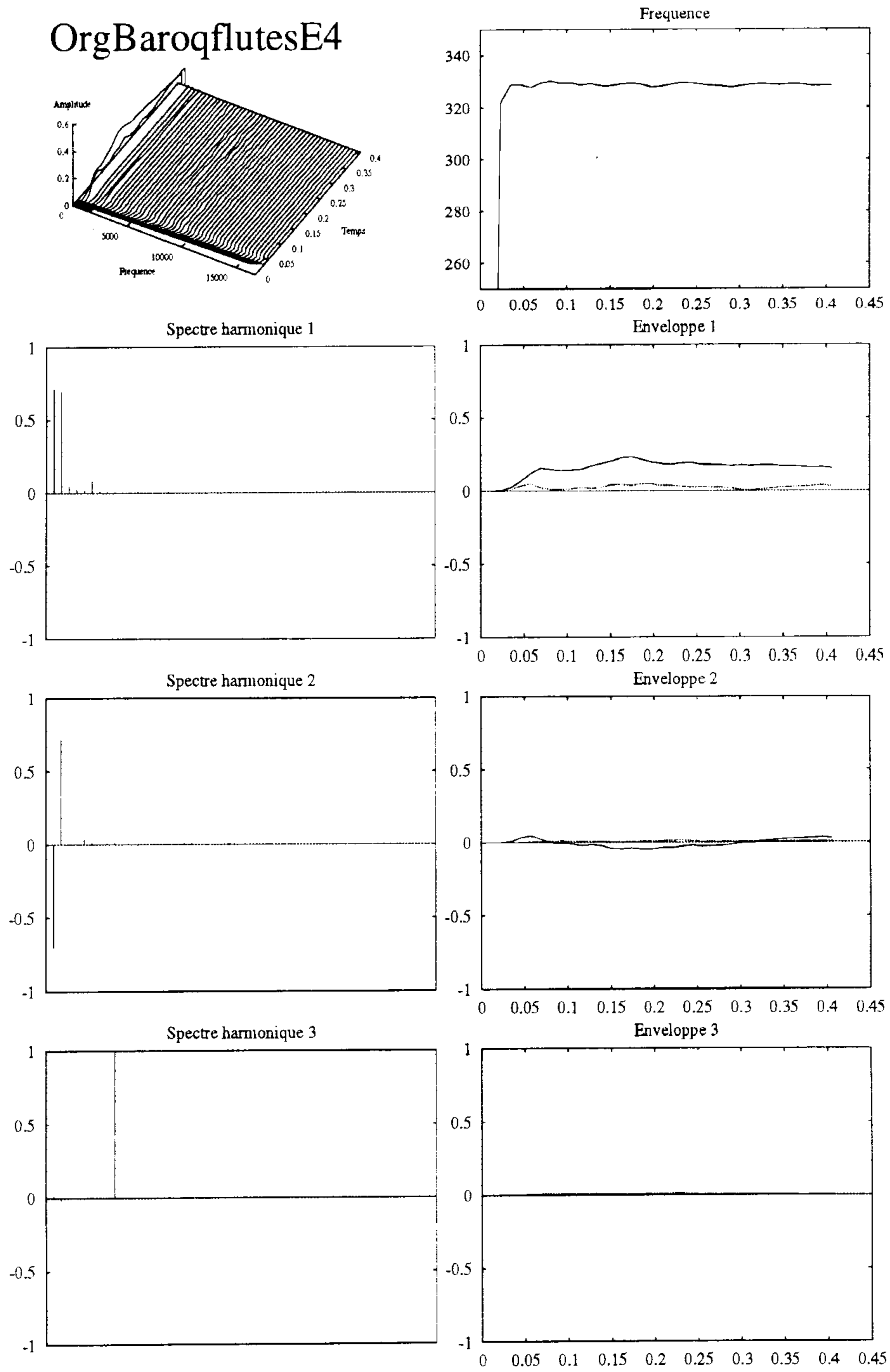


FIG. B.7 – Orgue baroque registre flûtes

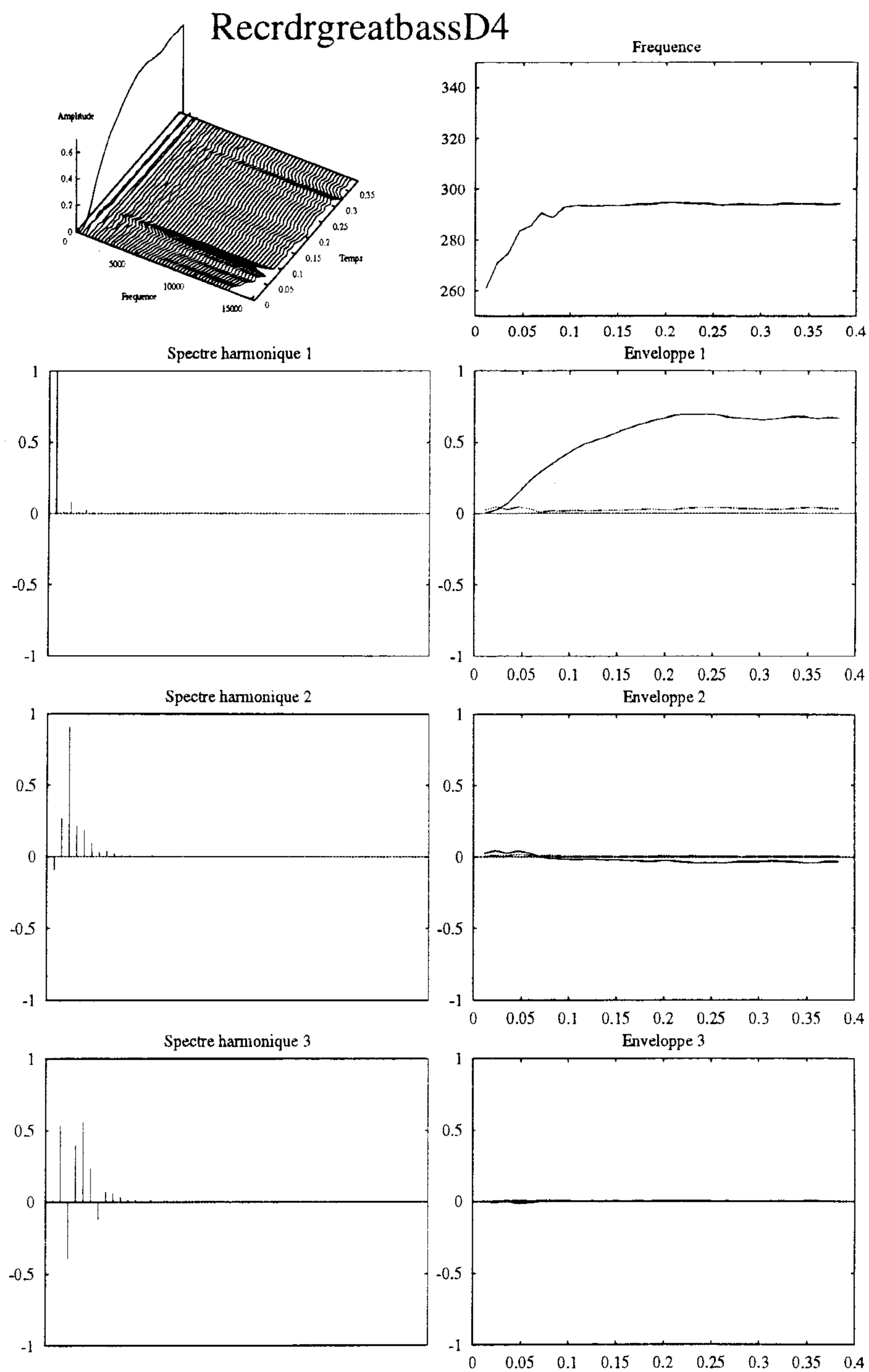


FIG. B.8 – Flûte à bec basse

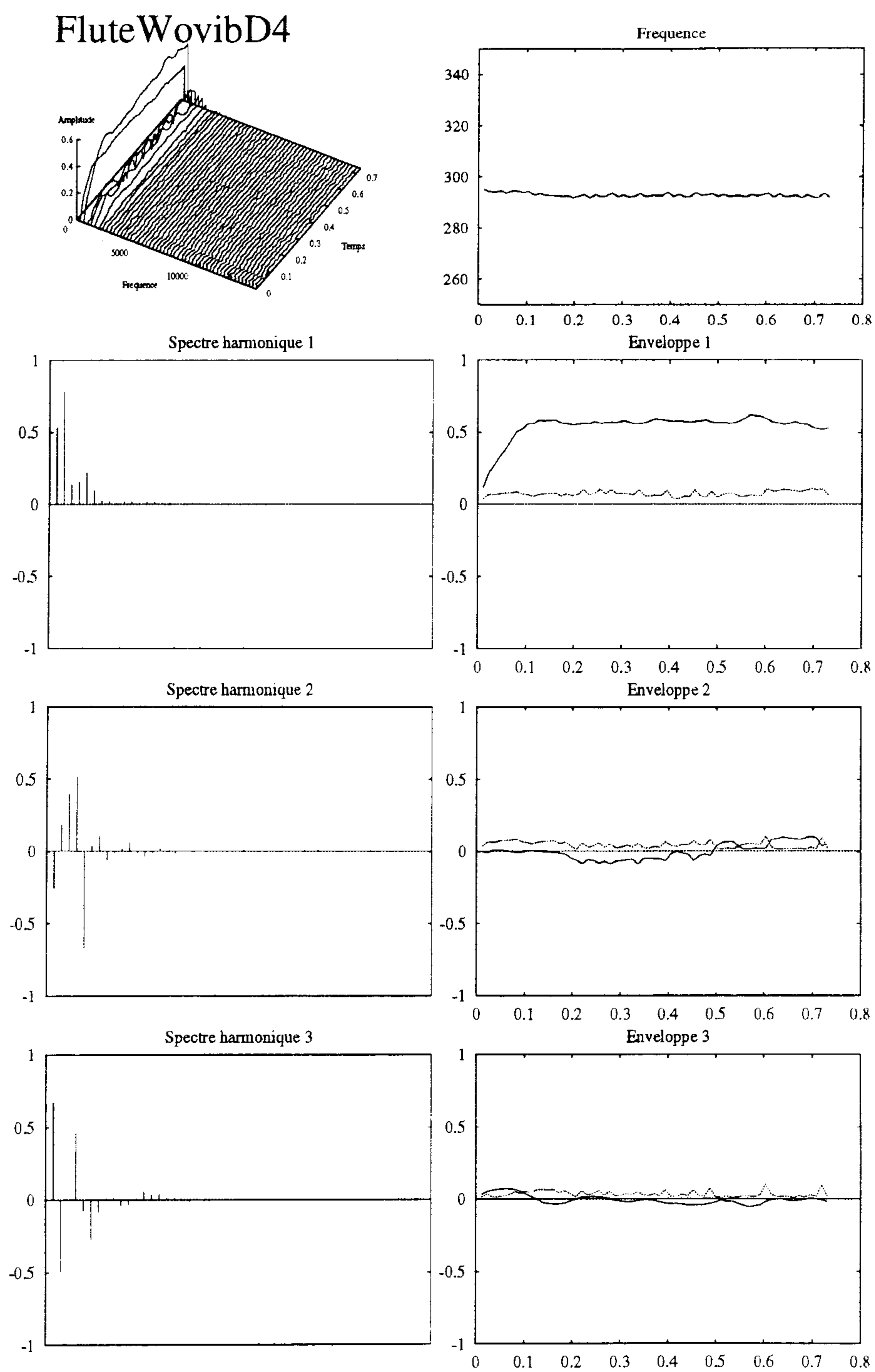


FIG. B.9 – *Flûte traversière sans vibrato*

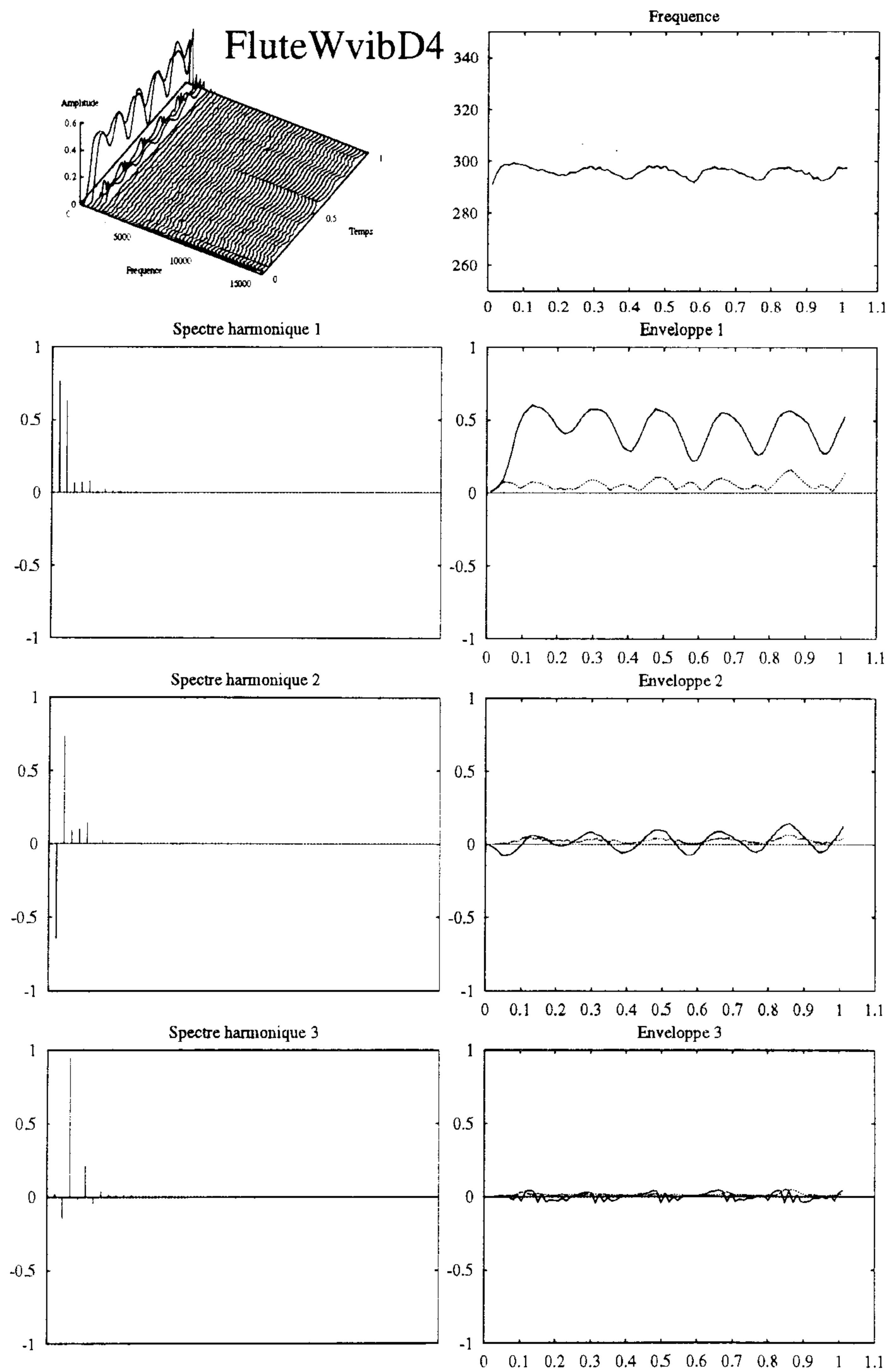


FIG. B.10 – Flûte traversière avec vibrato

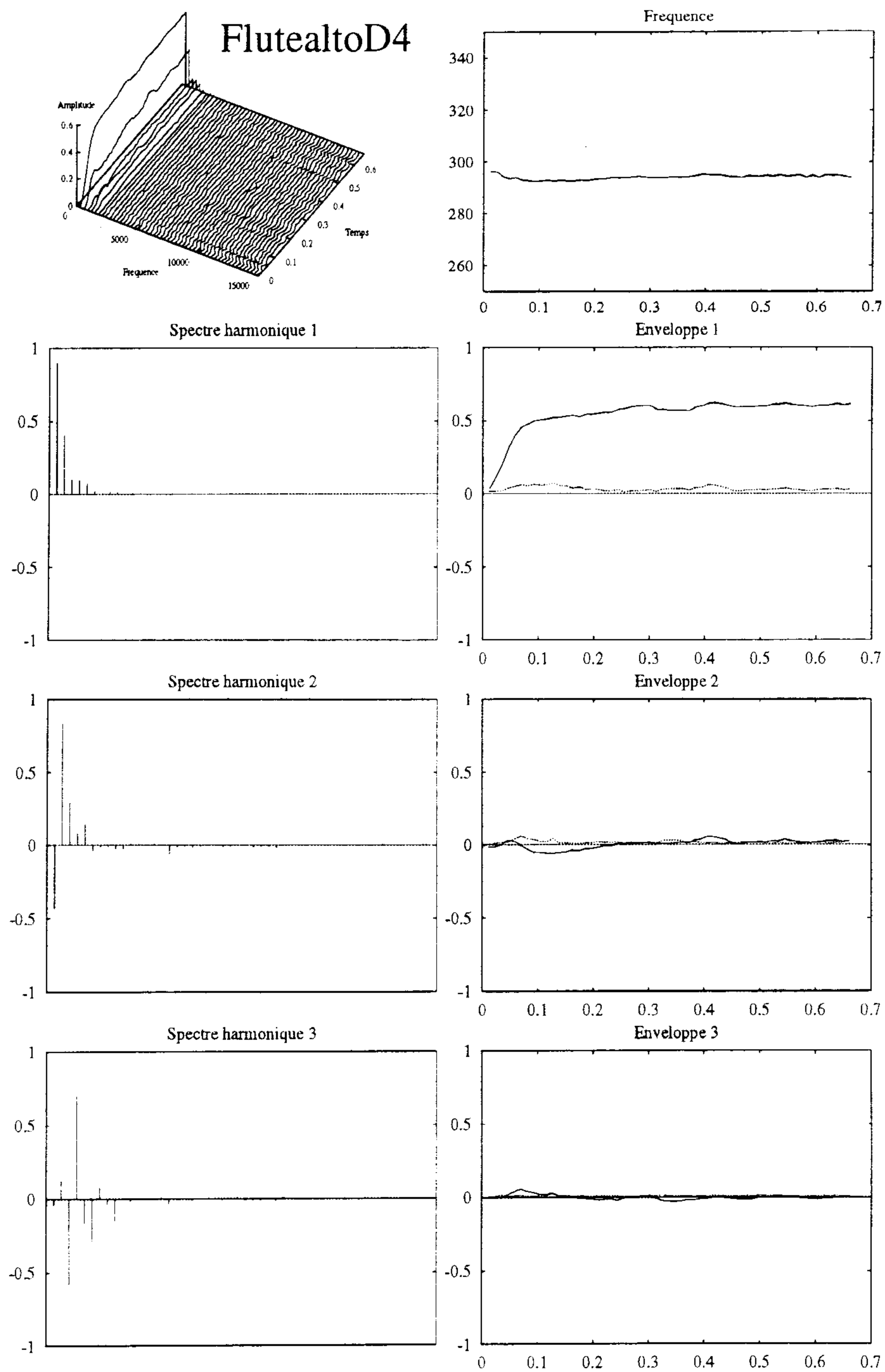


FIG. B.11 – *Flûte alto sans vibrato*



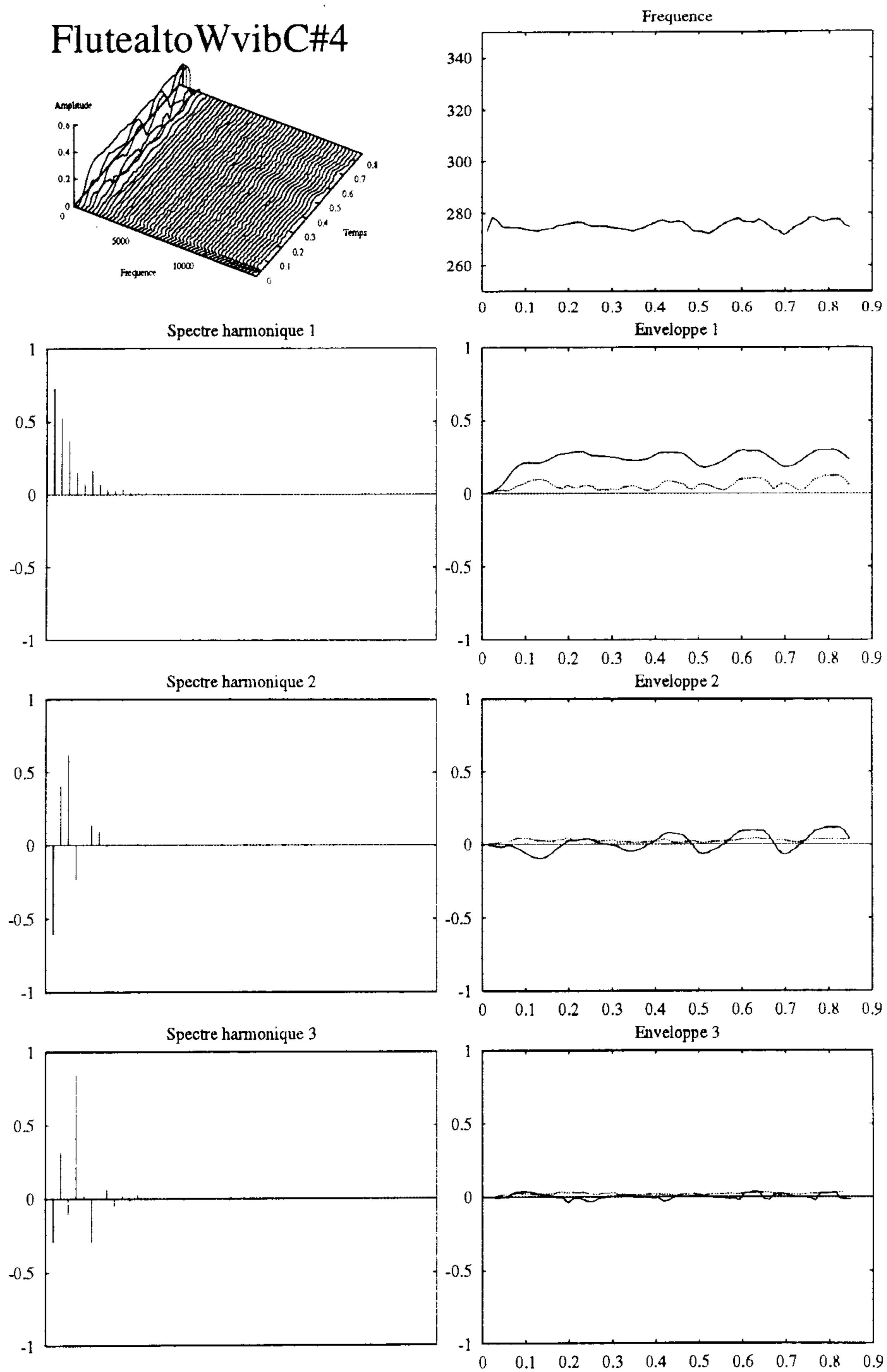


FIG. B.12 – Flûte alto avec vibrato

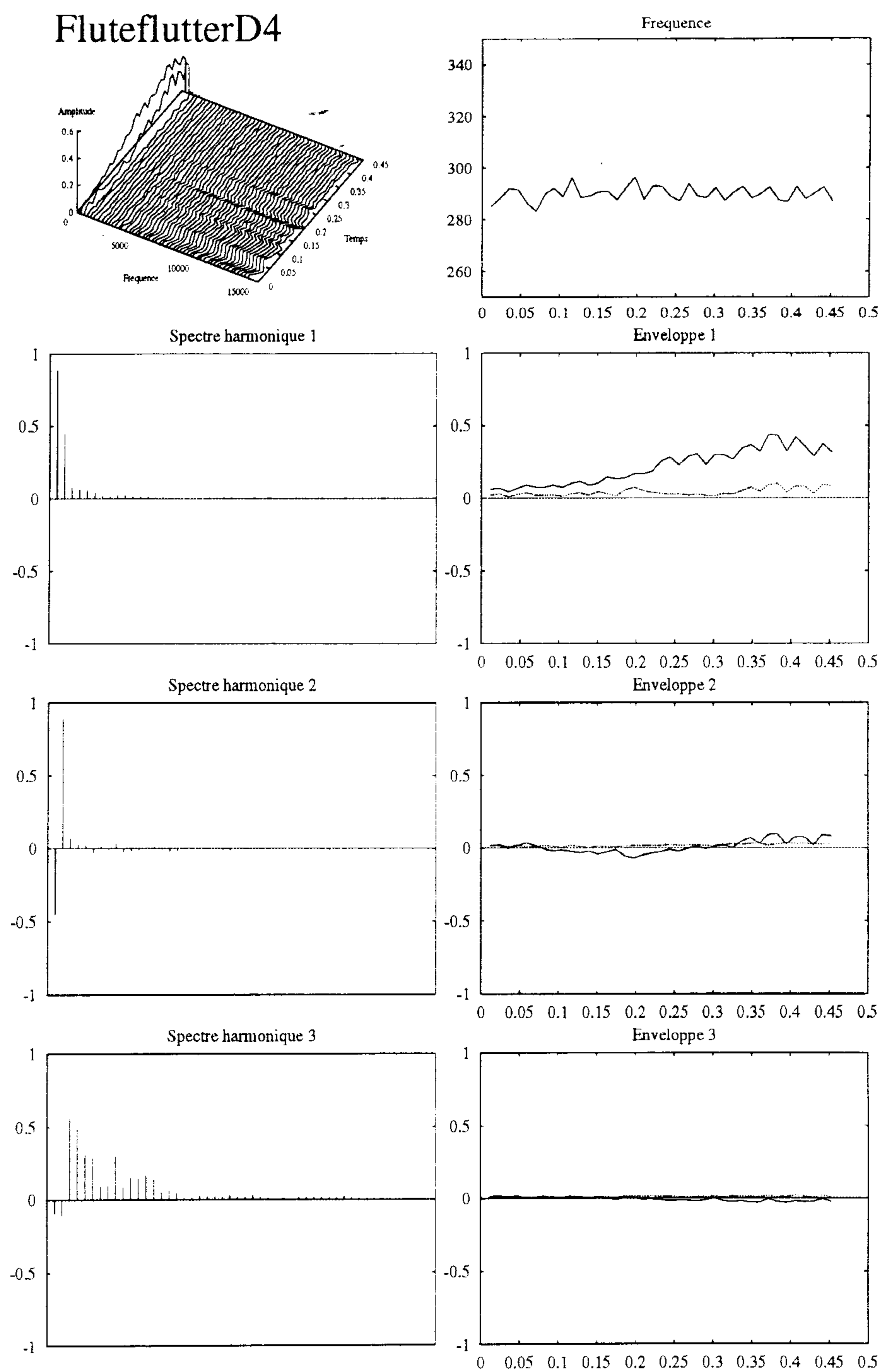


FIG. B.13 – Flûte jouée “flutter”

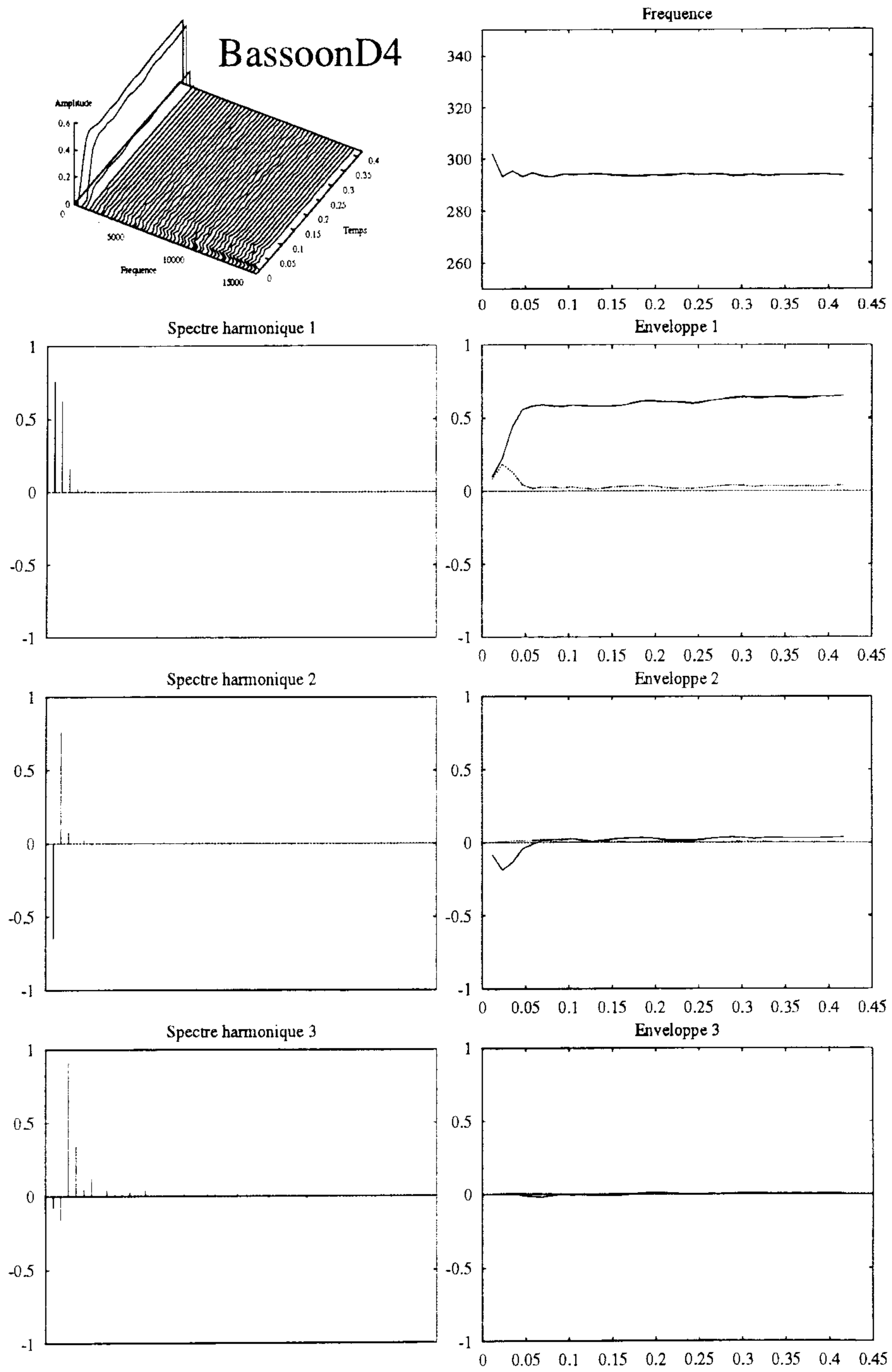
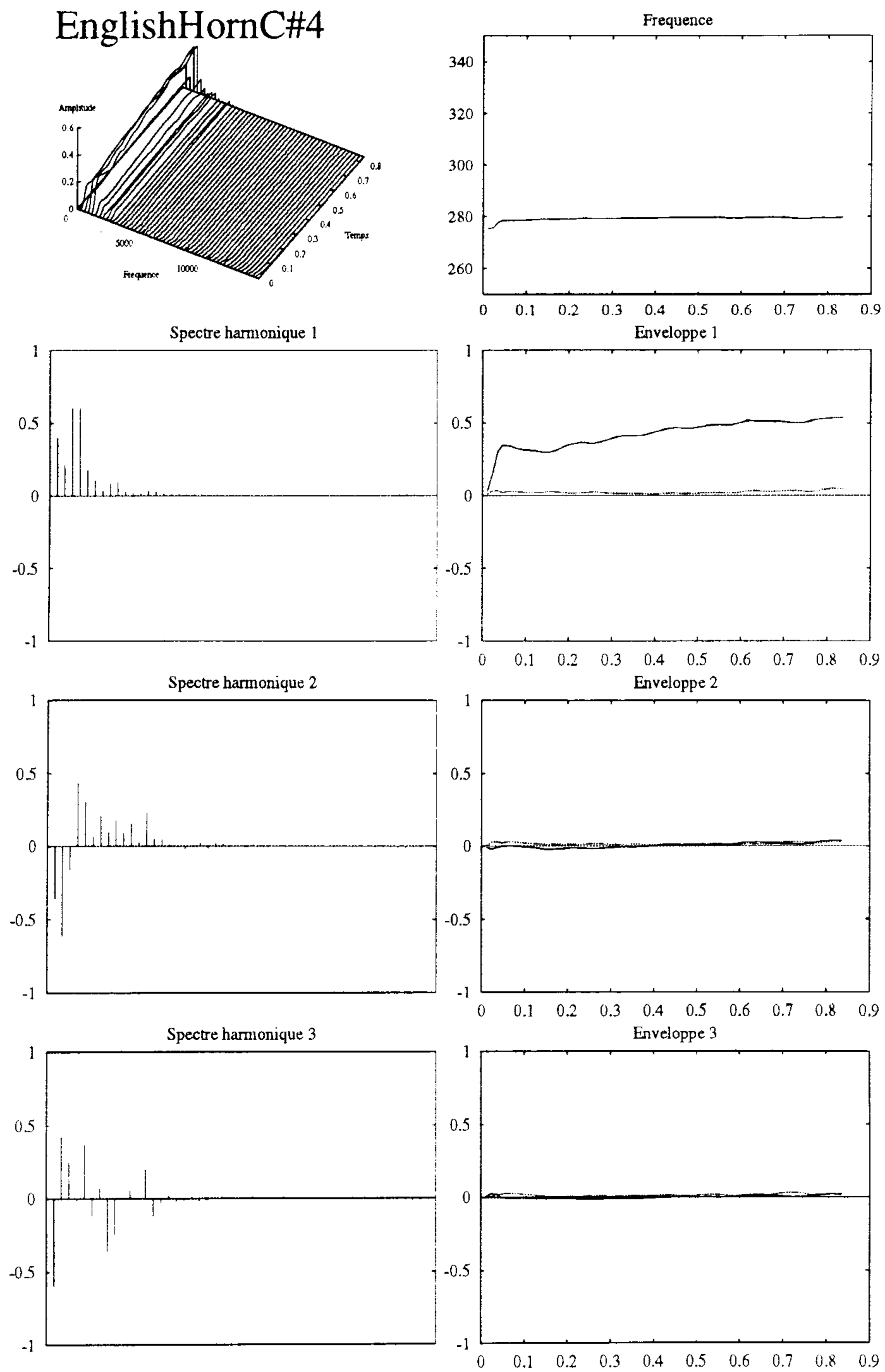


FIG. B.14 - Basson

FIG. B.15 – *Cor Anglais*

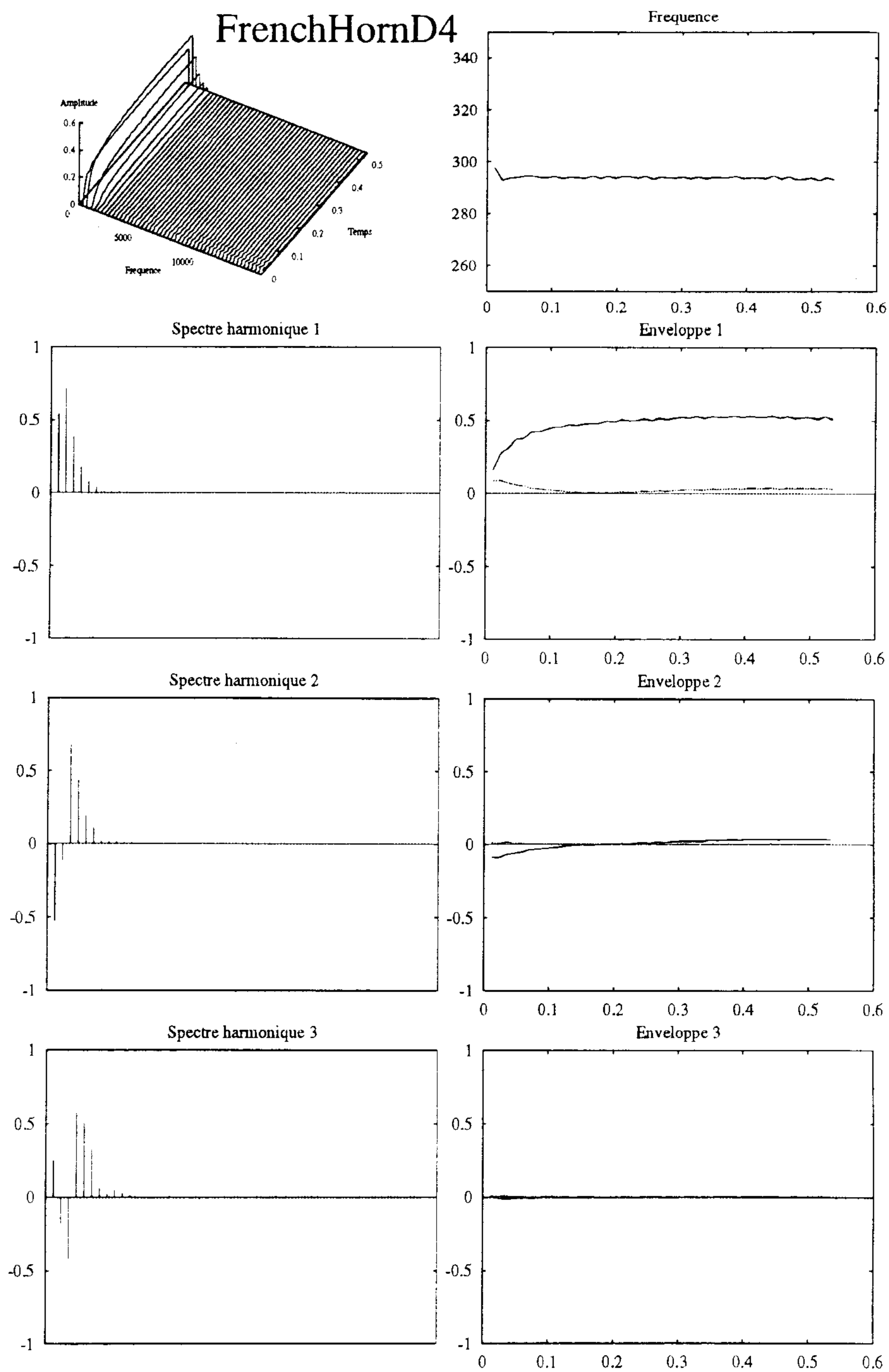
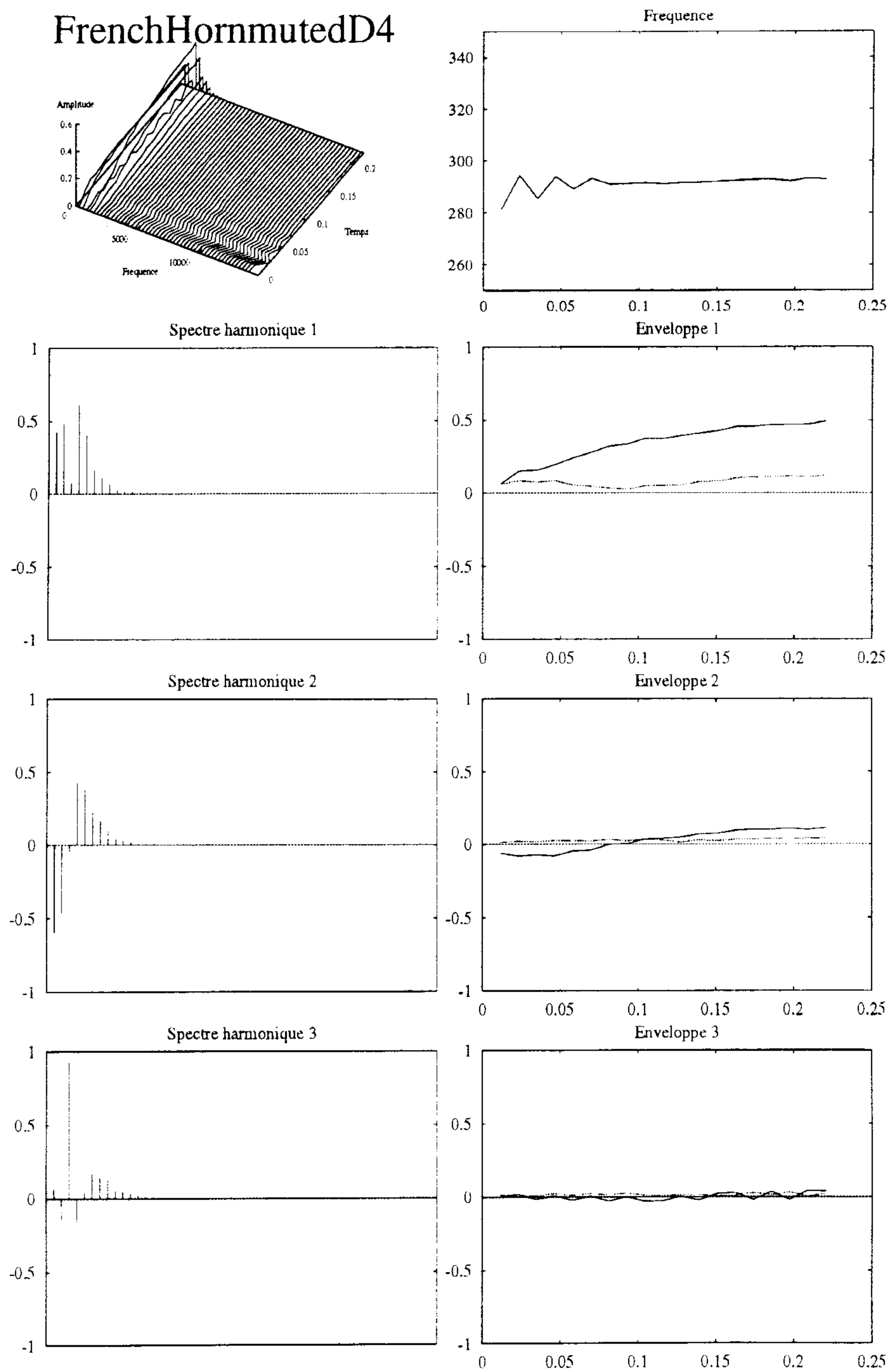
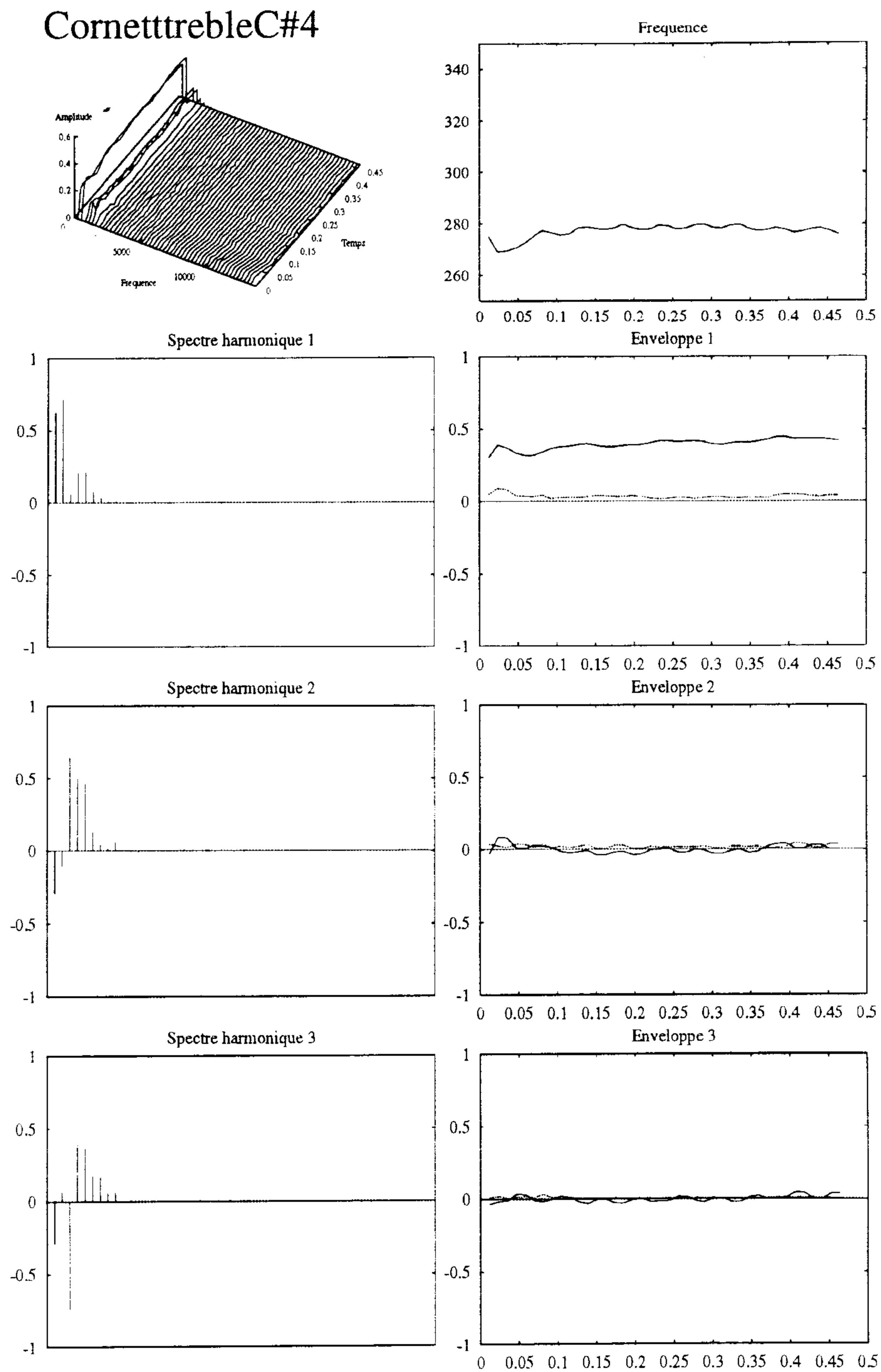
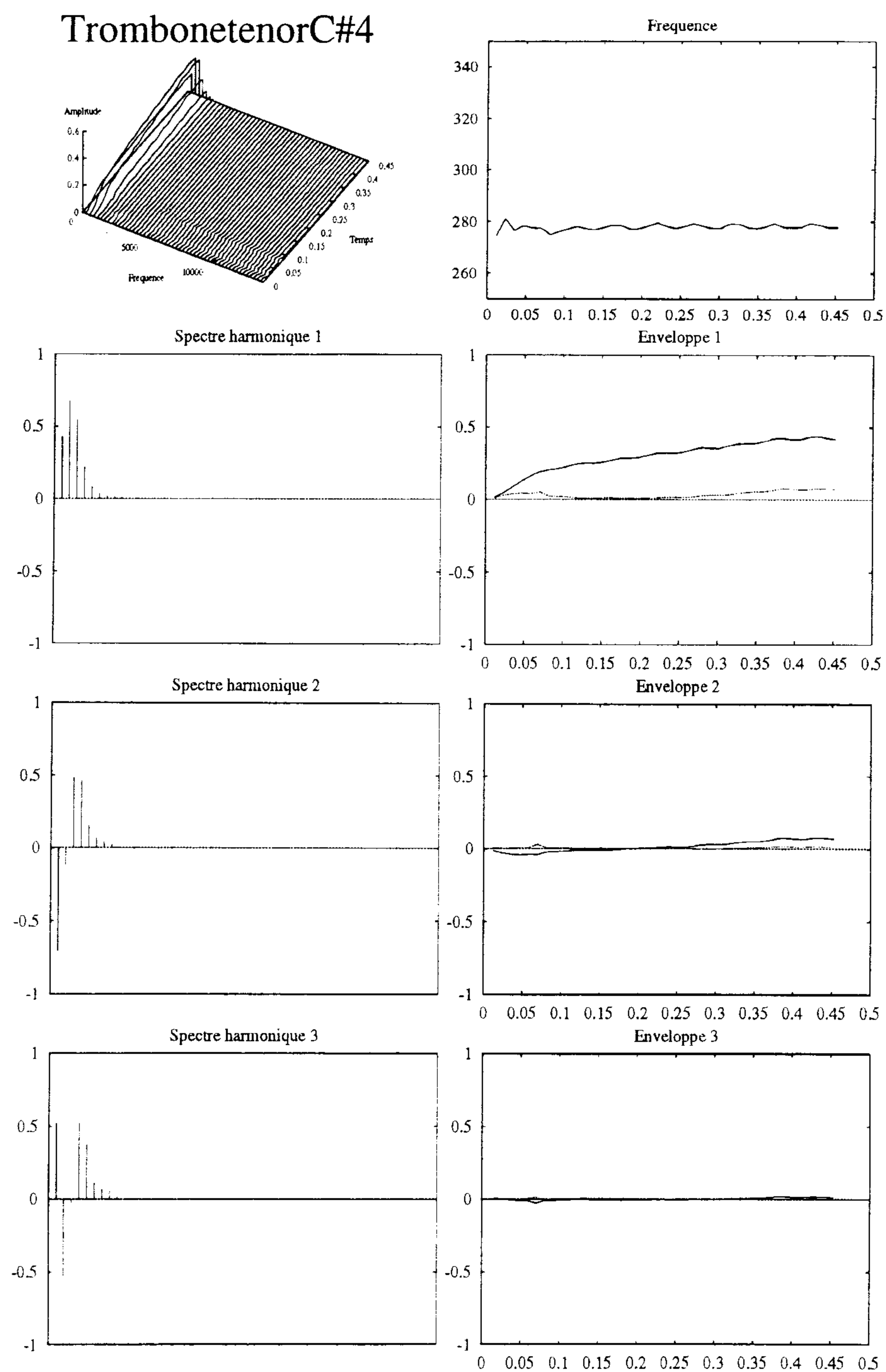


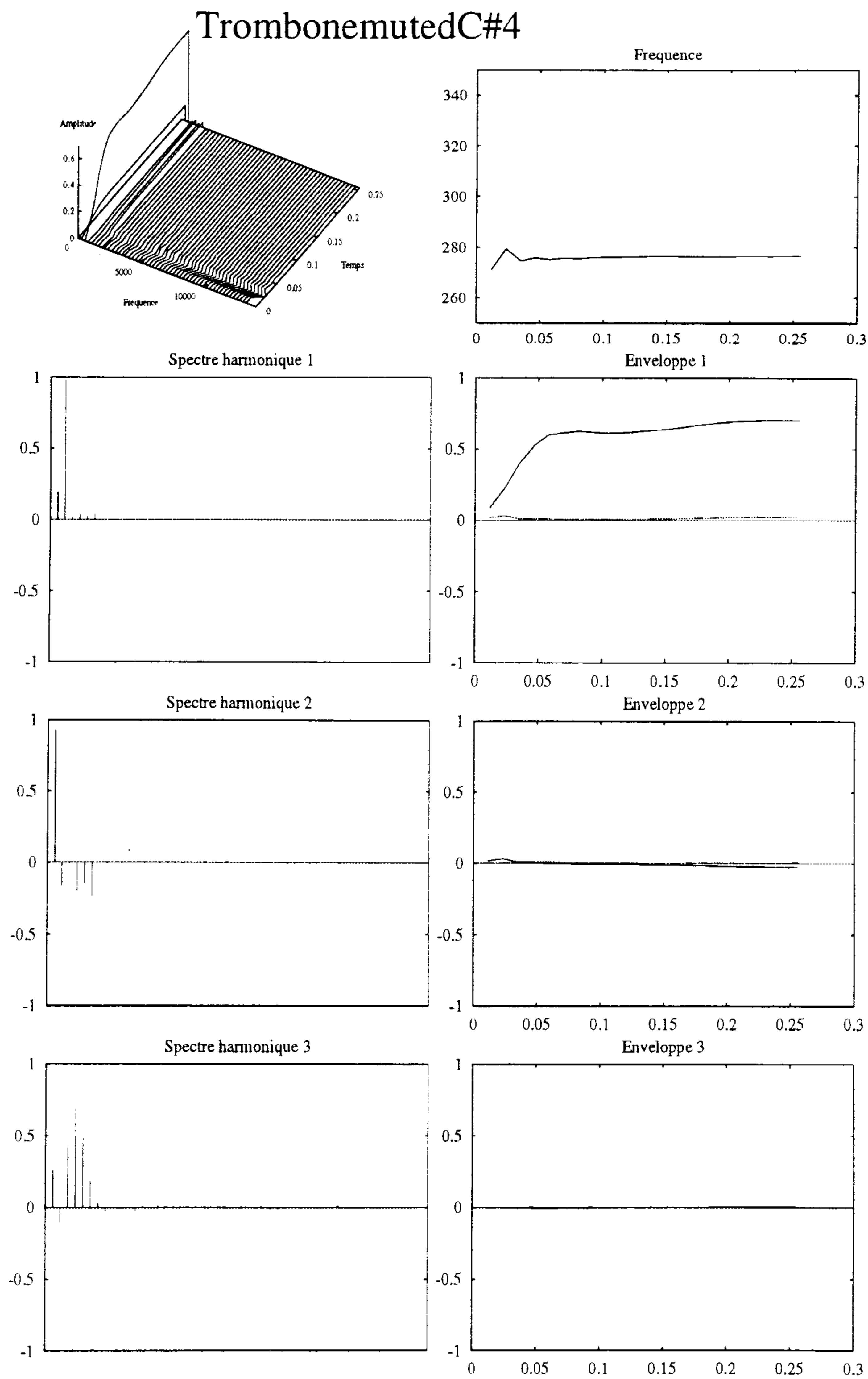
FIG. B.16 - Cor

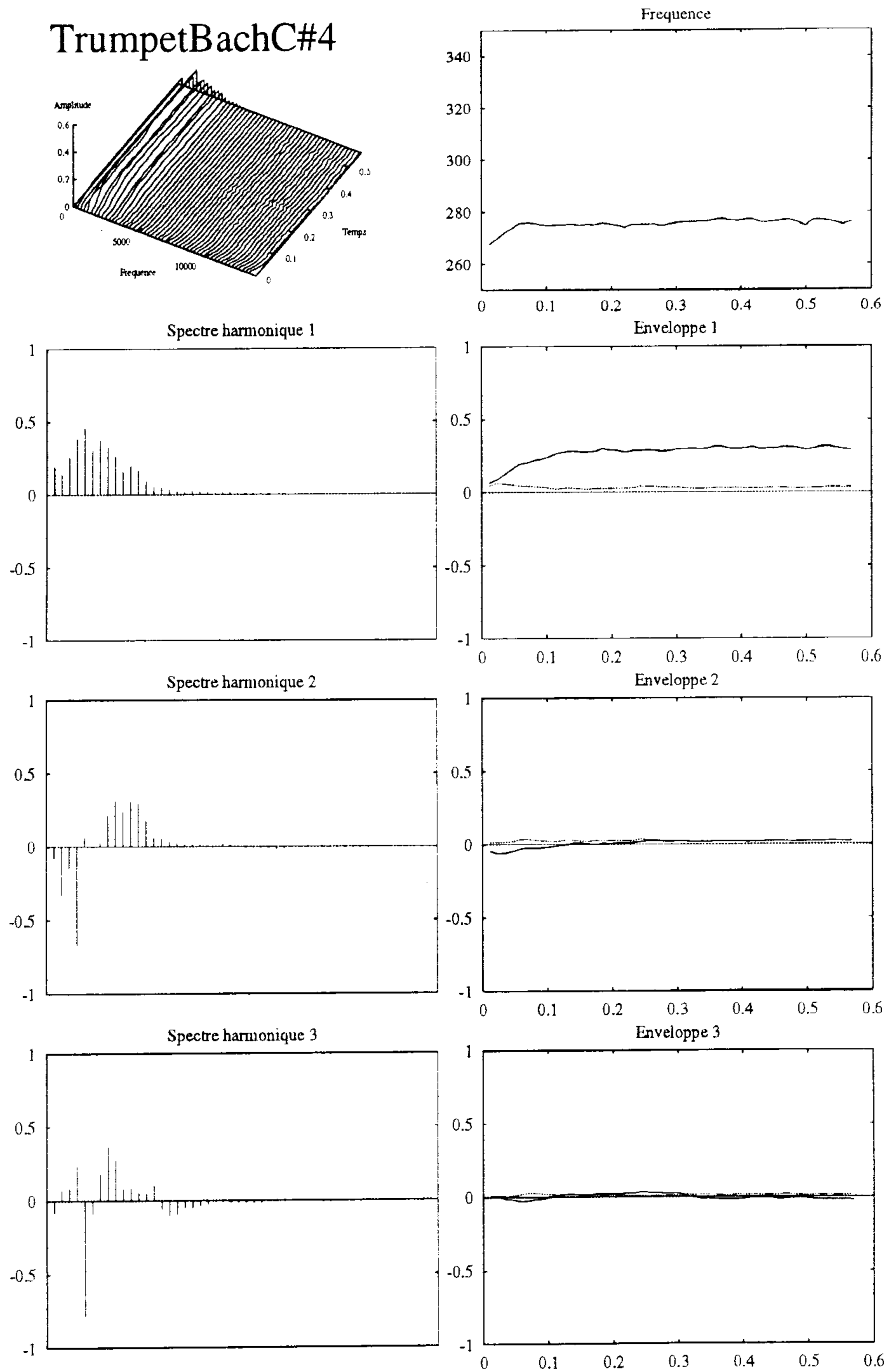
FIG. B.17 – *Cor avec sourdine*

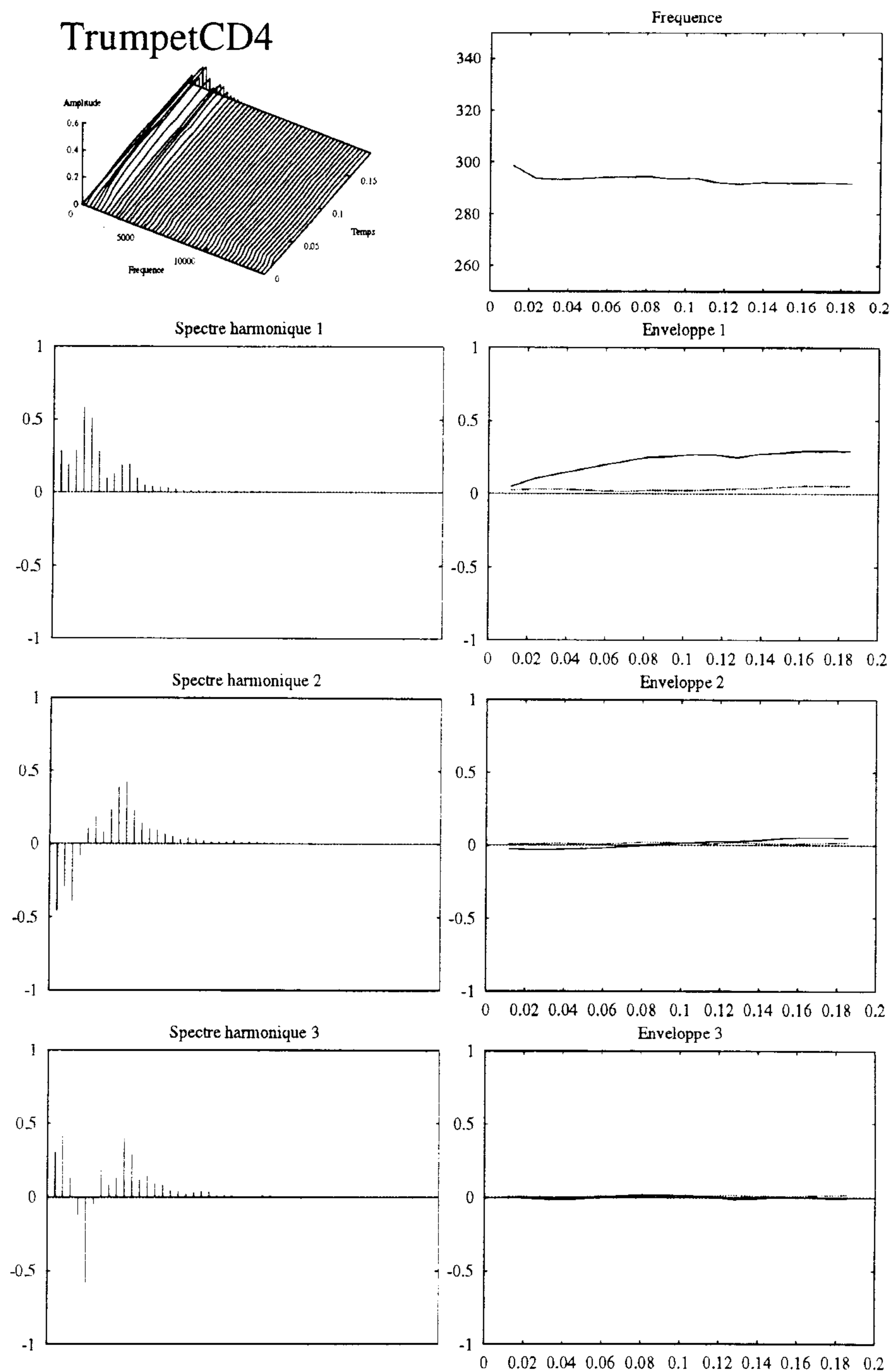
FIG. B.18 – *Cornet*

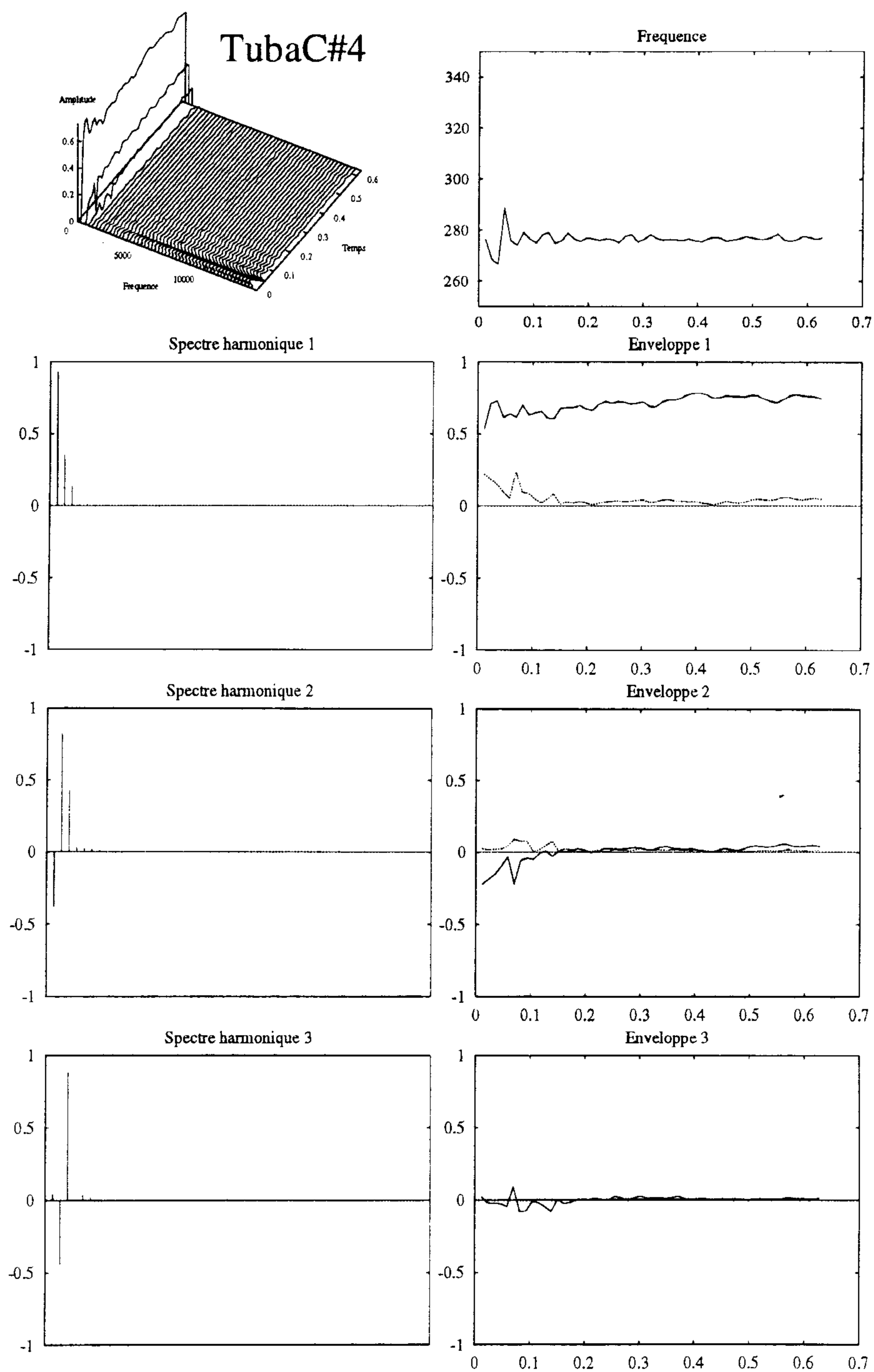
FIG. B.19 – *Trombone*



FIG. B.20 – *Trombone avec sourdine*

FIG. B.21 – *Trompette baroque*

FIG. B.22 - *Trompette en Do*

FIG. B.23 – *Tuba*

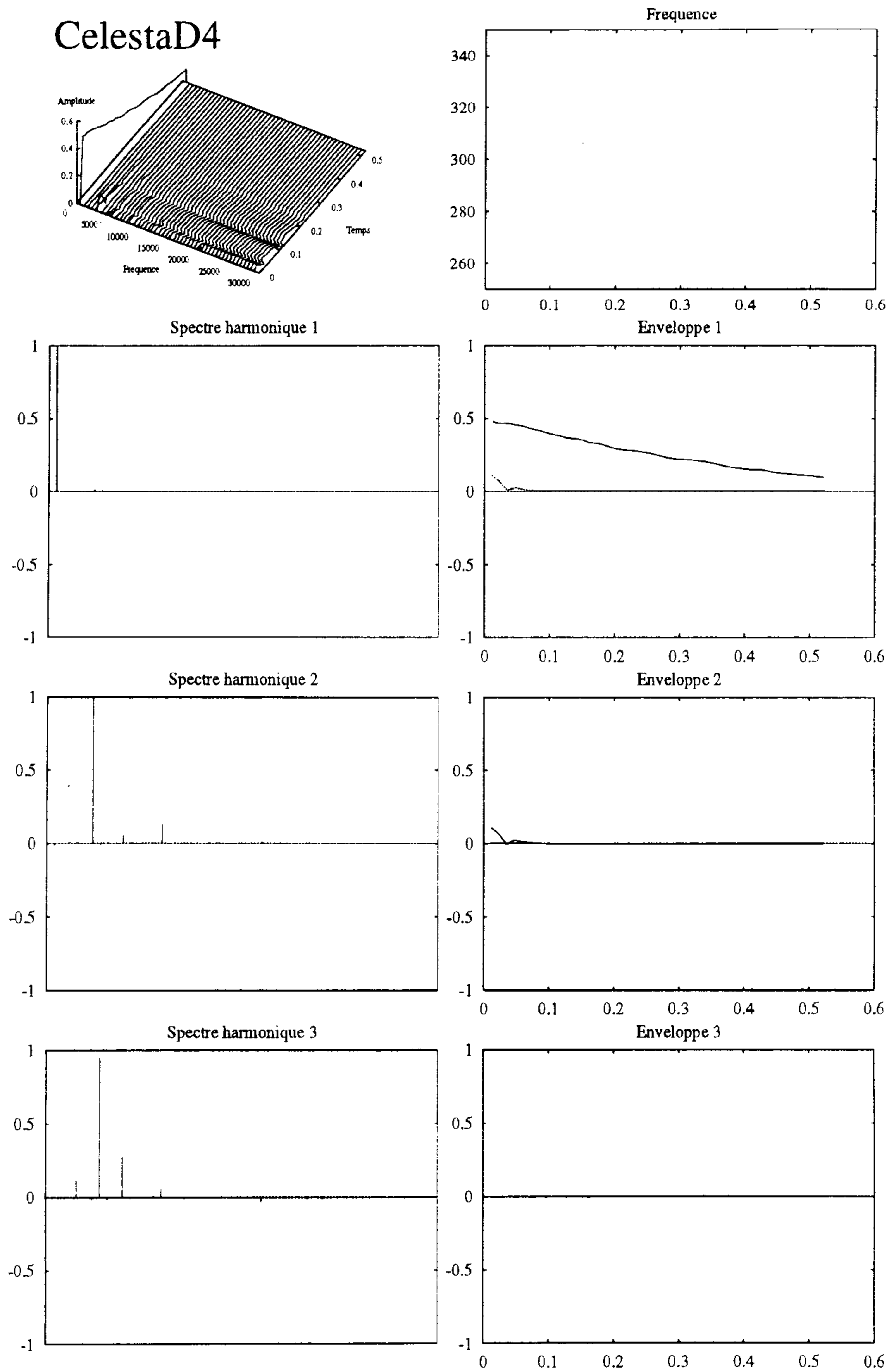
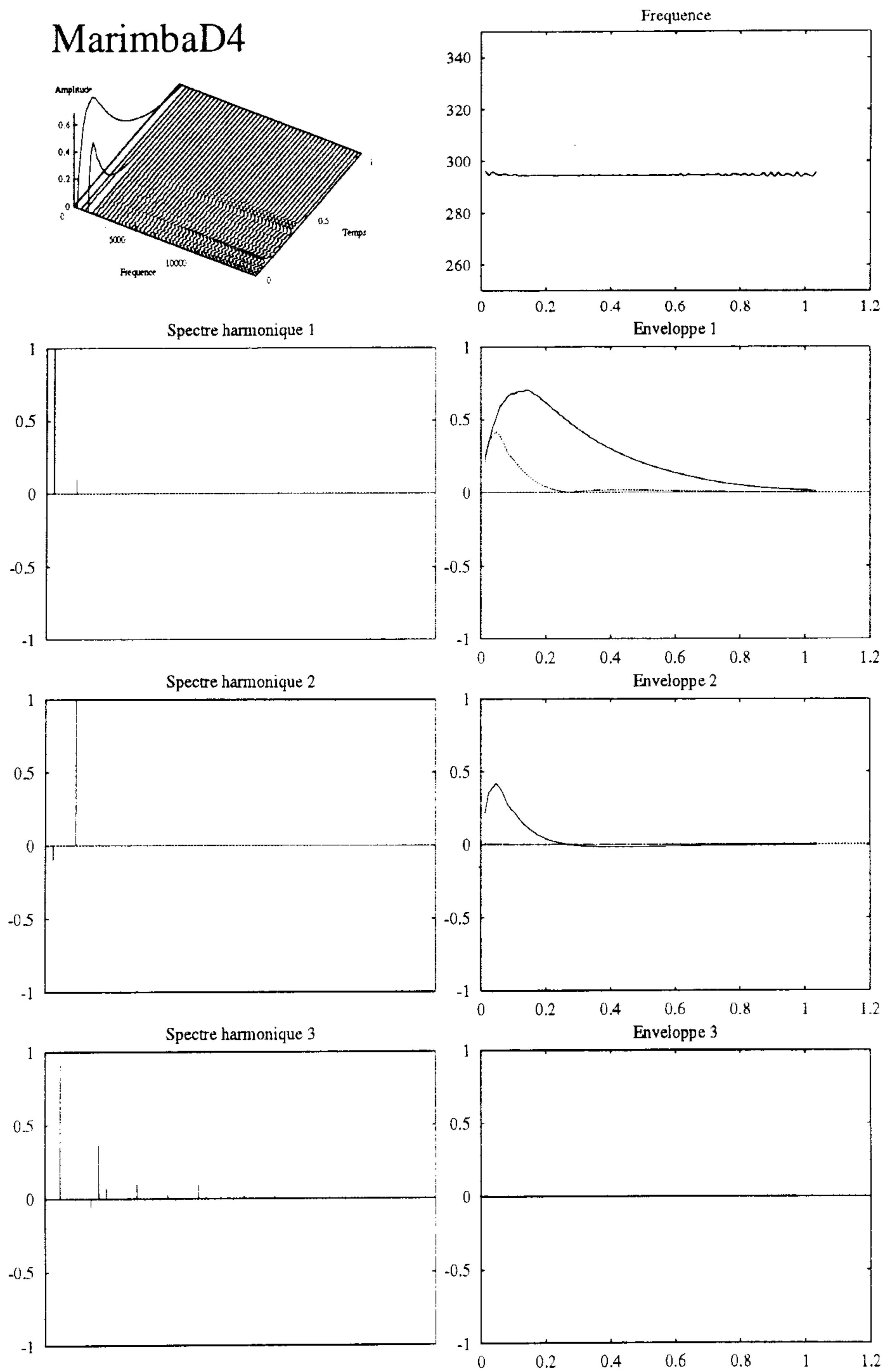
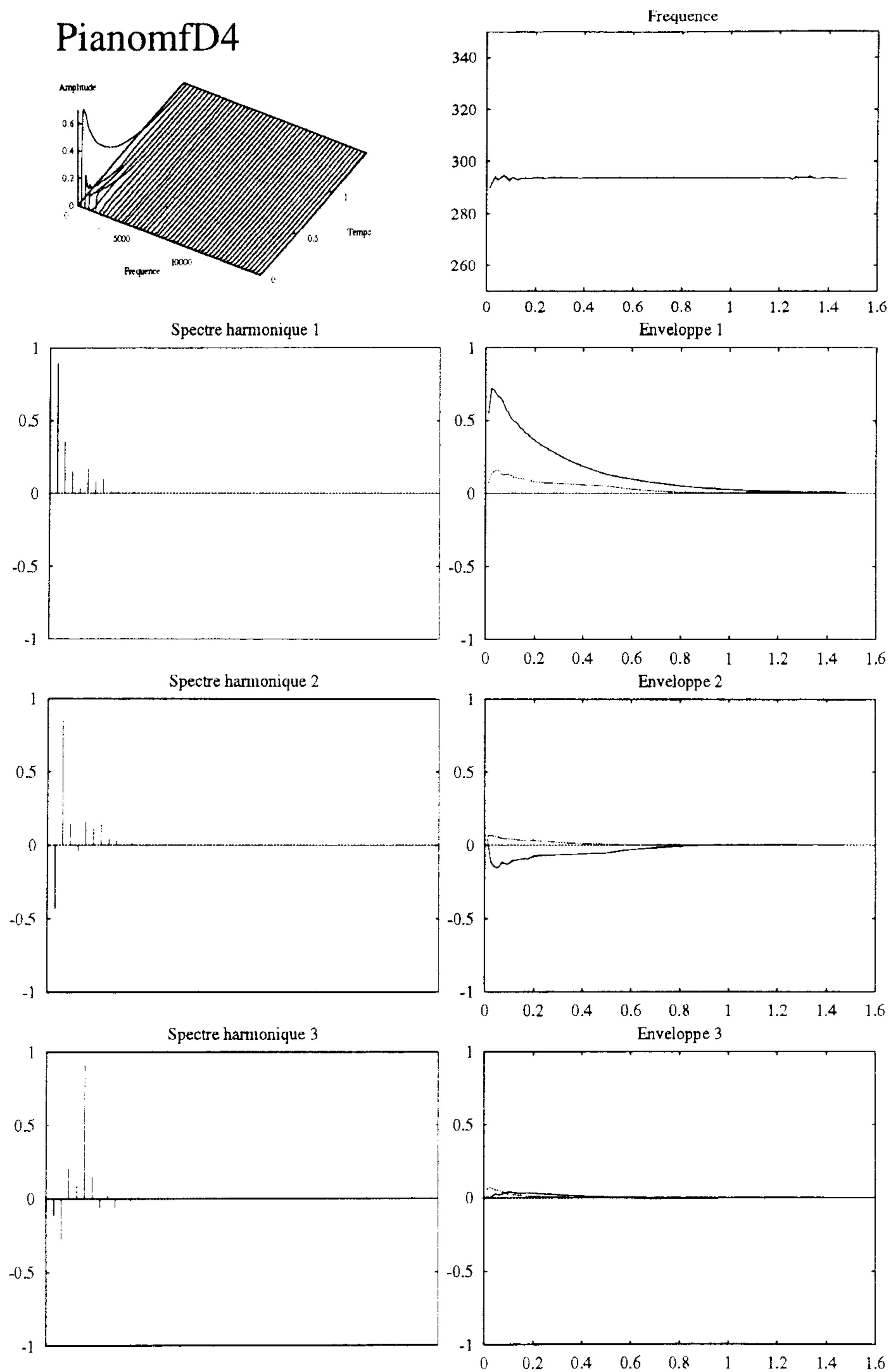


FIG. B.24 - Celesta

FIG. B.25 – *Marimba*

FIG. B.26 – *Piano joué mezzo forte*

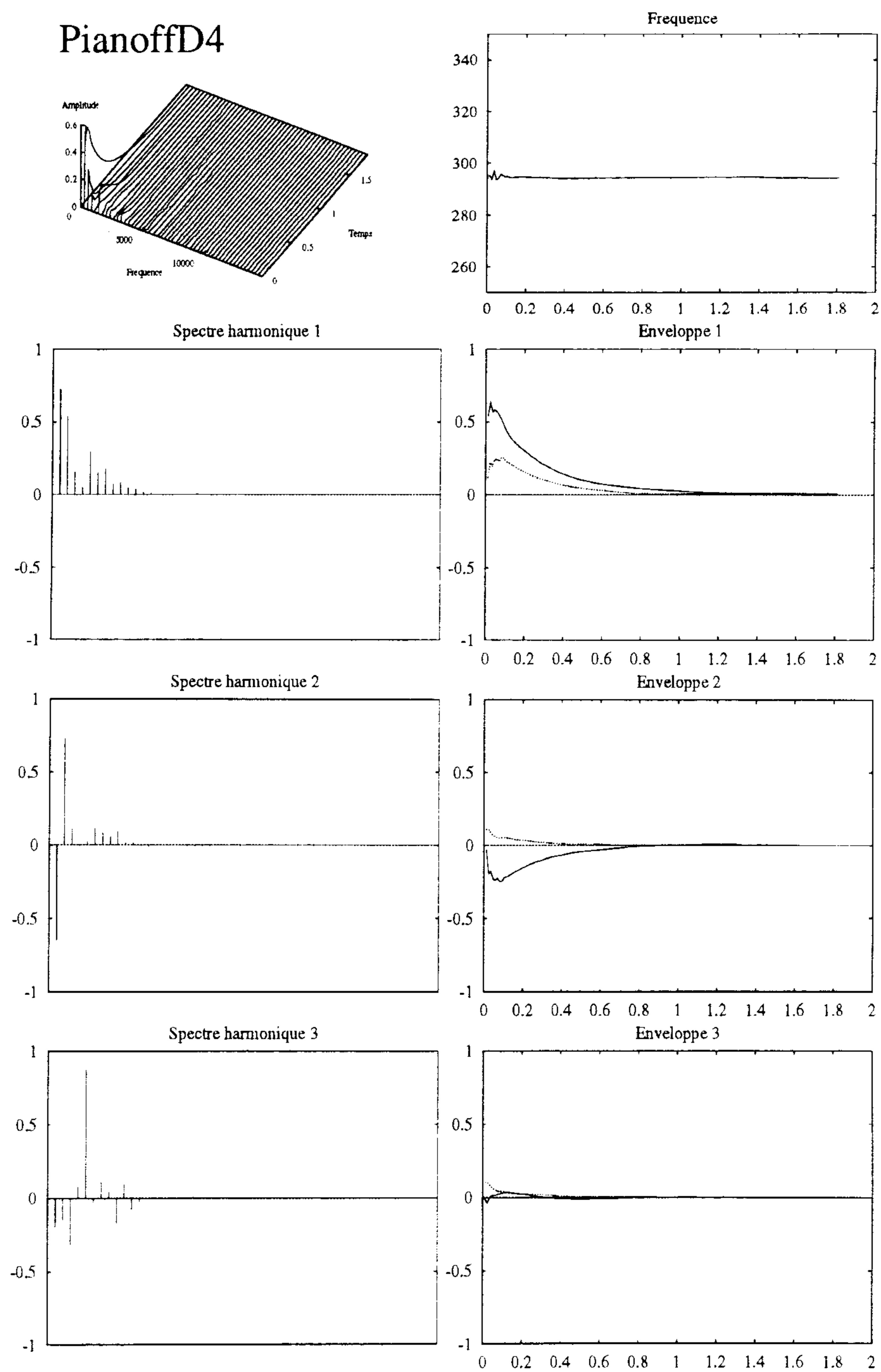


FIG. B.27 – *Piano joué fortissimo*



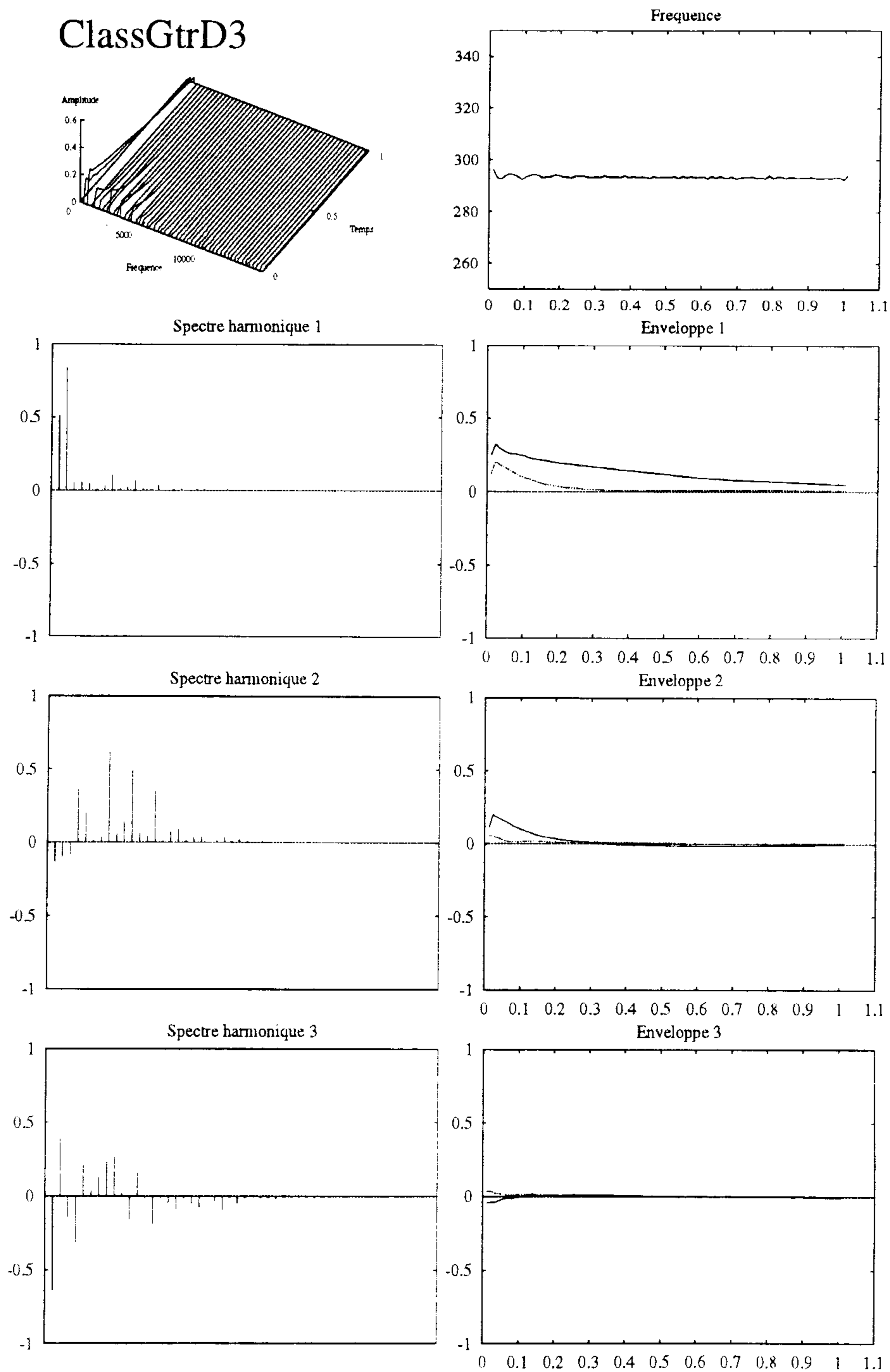


FIG. B.28 – Guitare classique

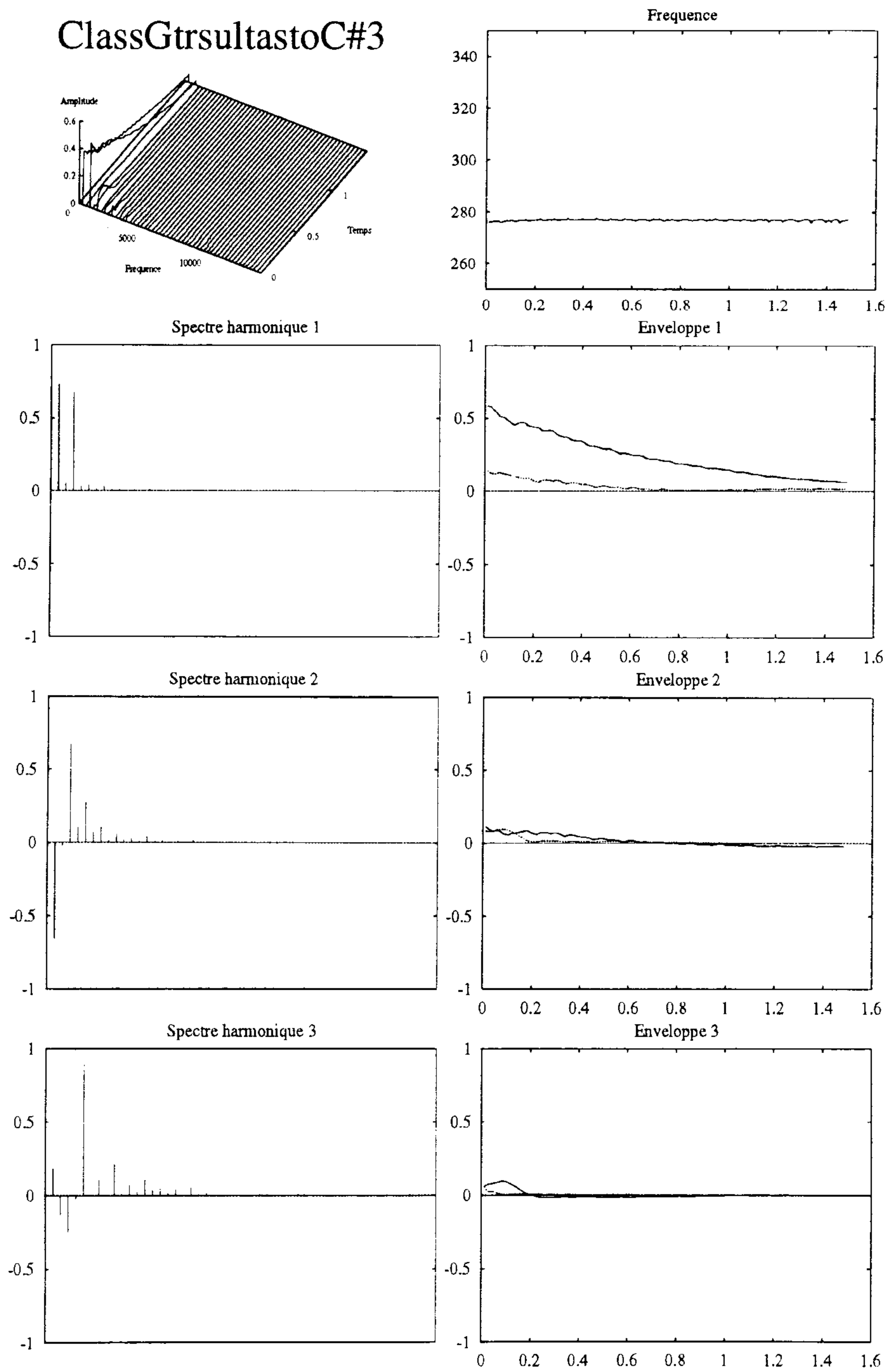
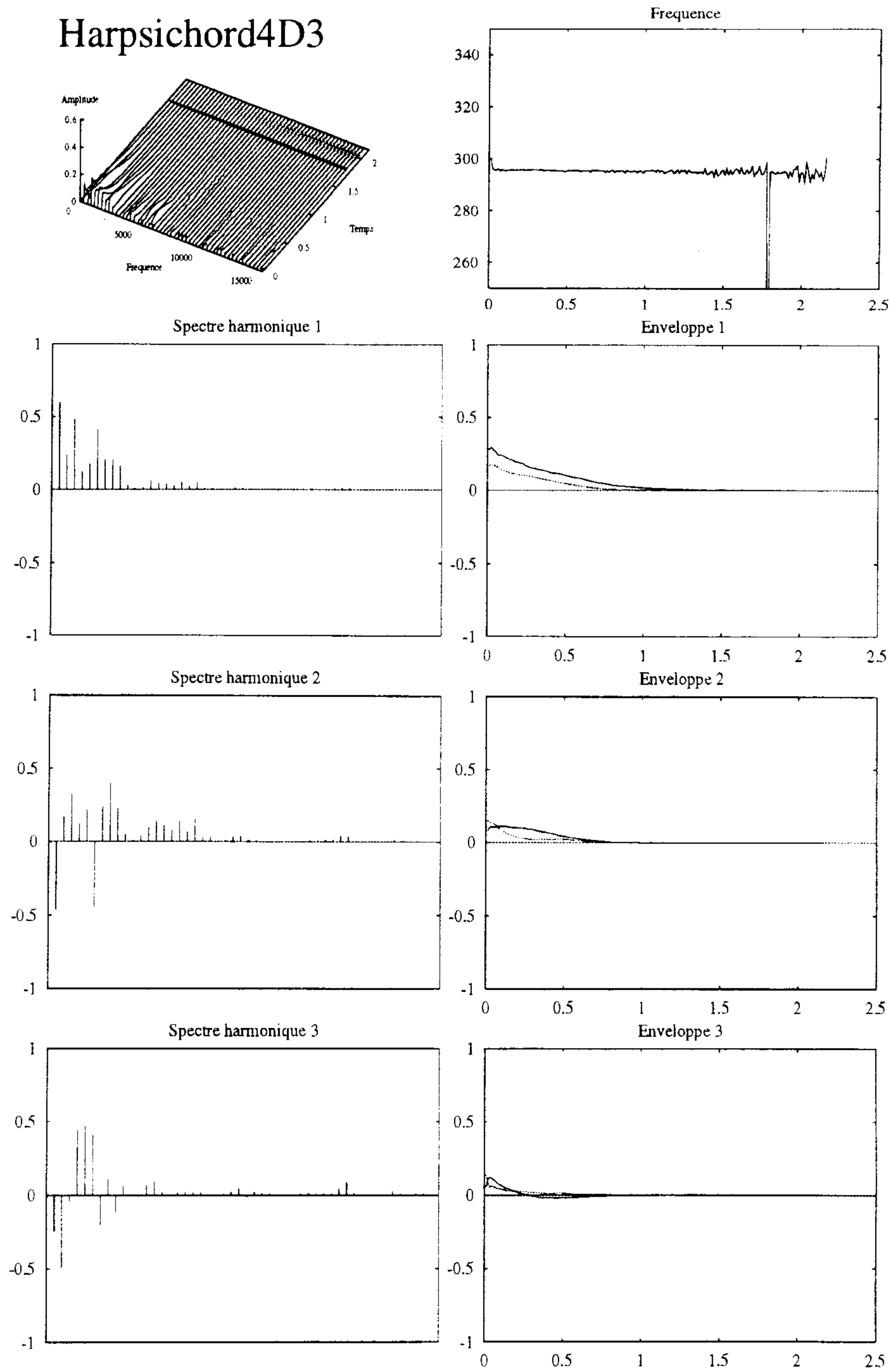


FIG. B.29 – Guitare classique jouée sul tasto

FIG. B.30 – *Clavecin*

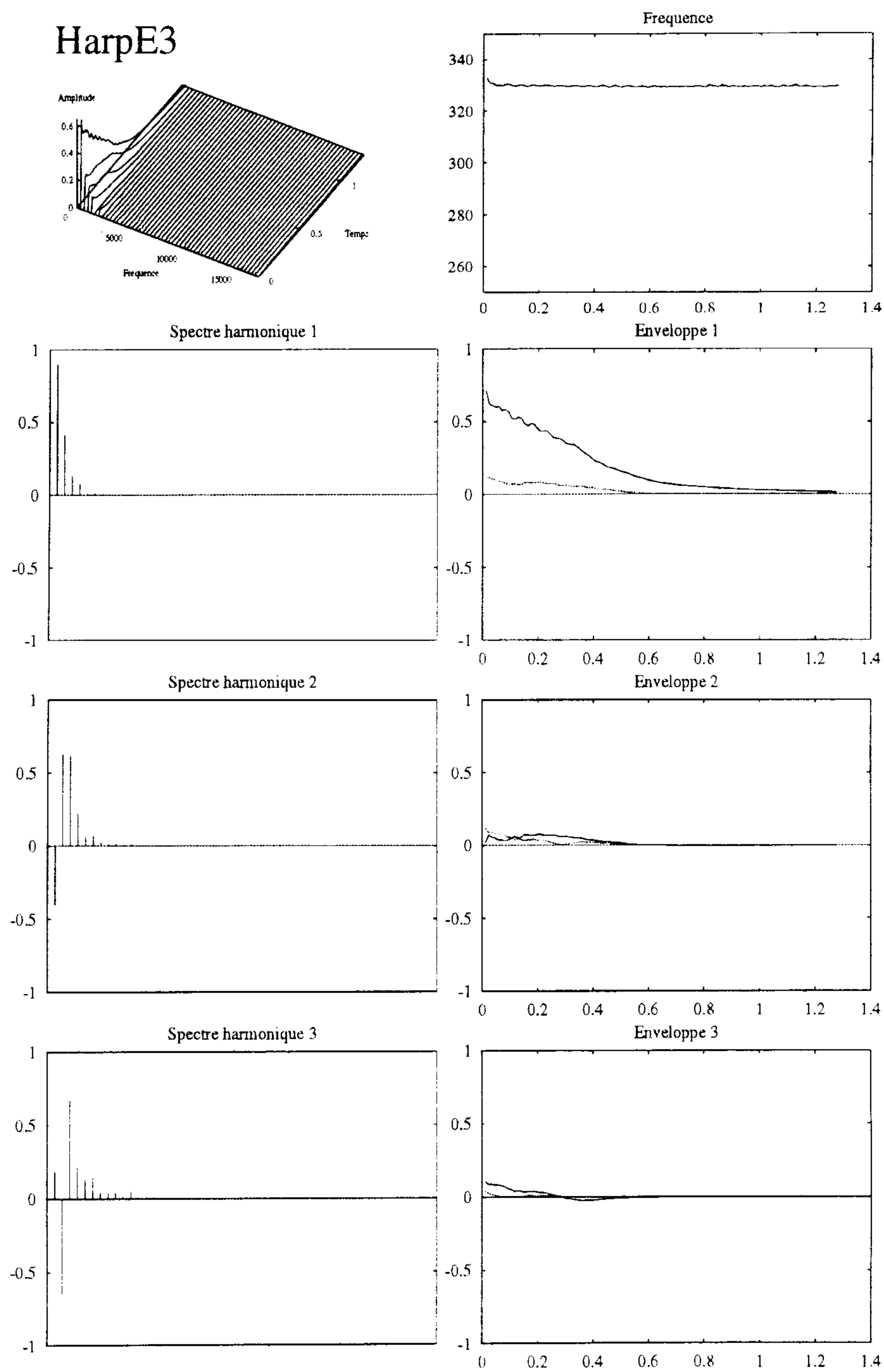


FIG. B.31 – Harpe

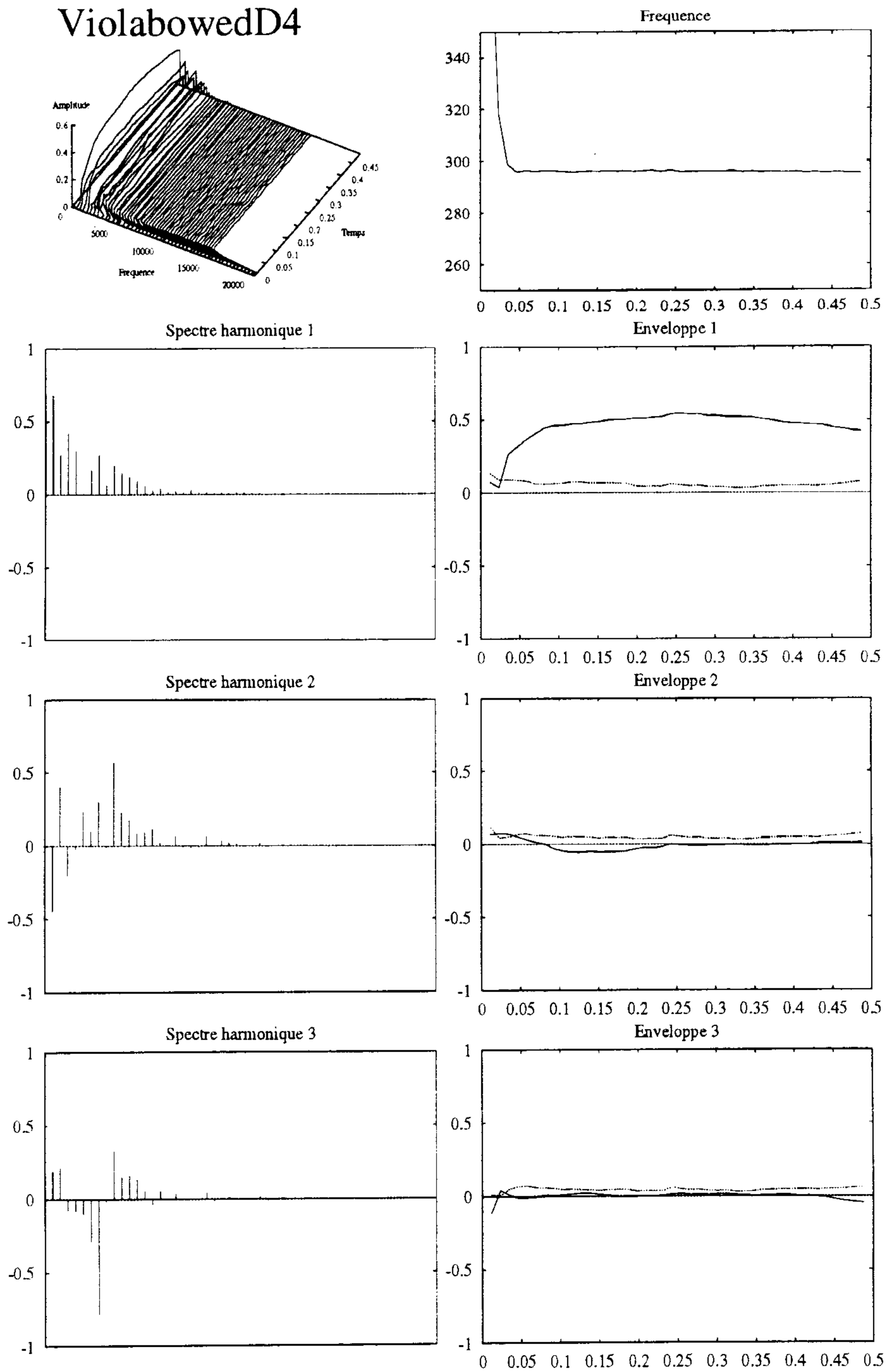
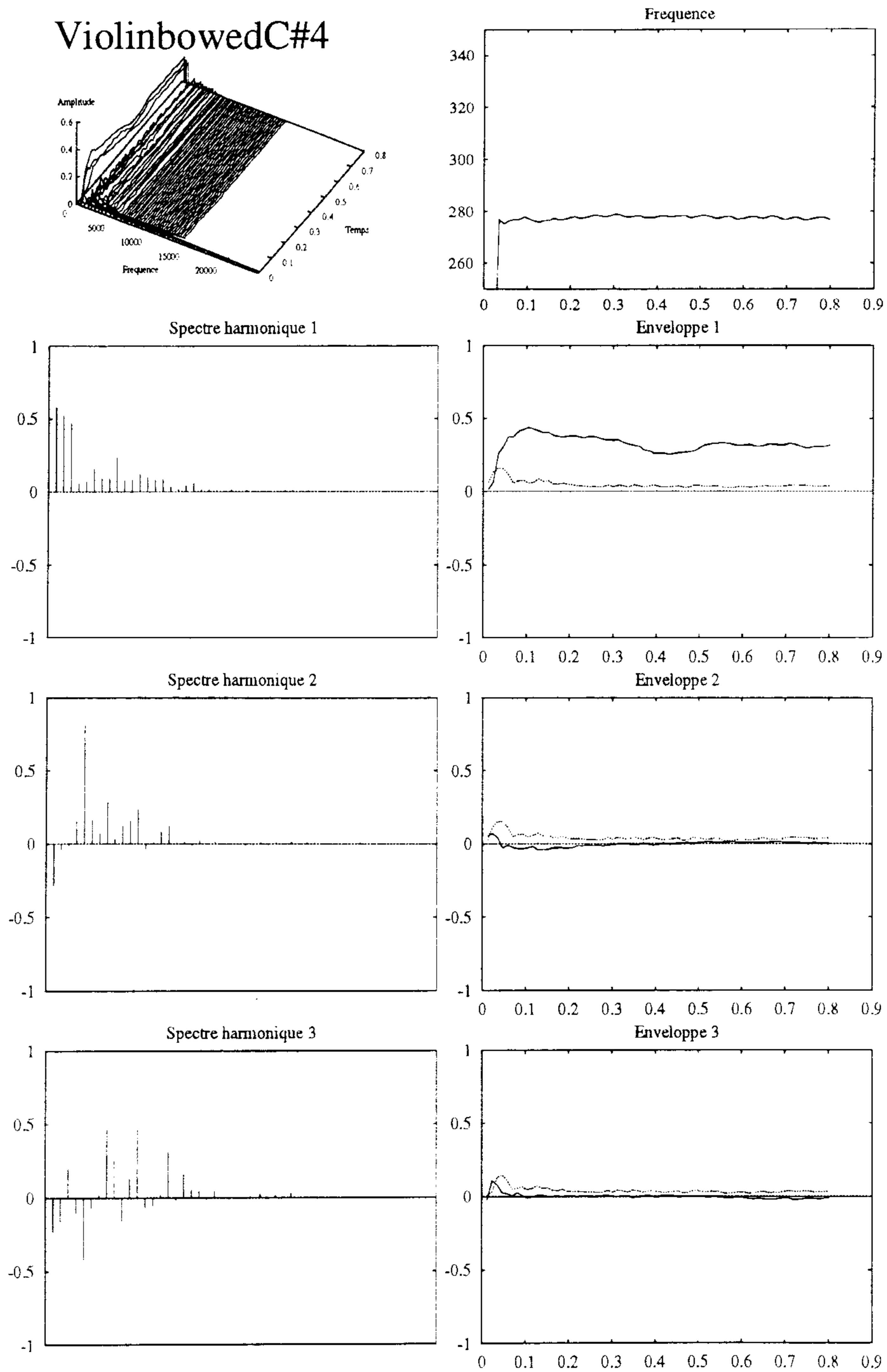
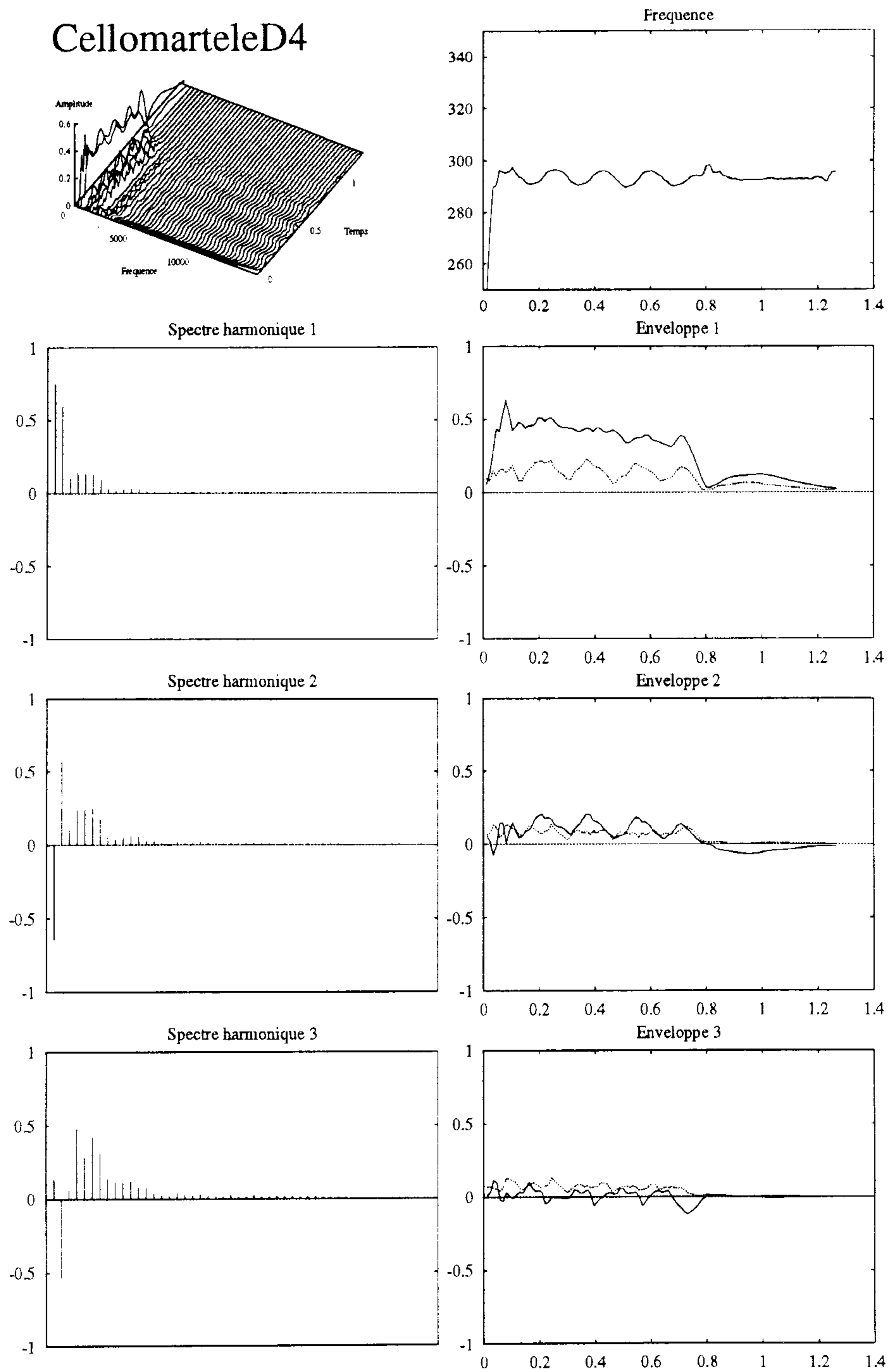
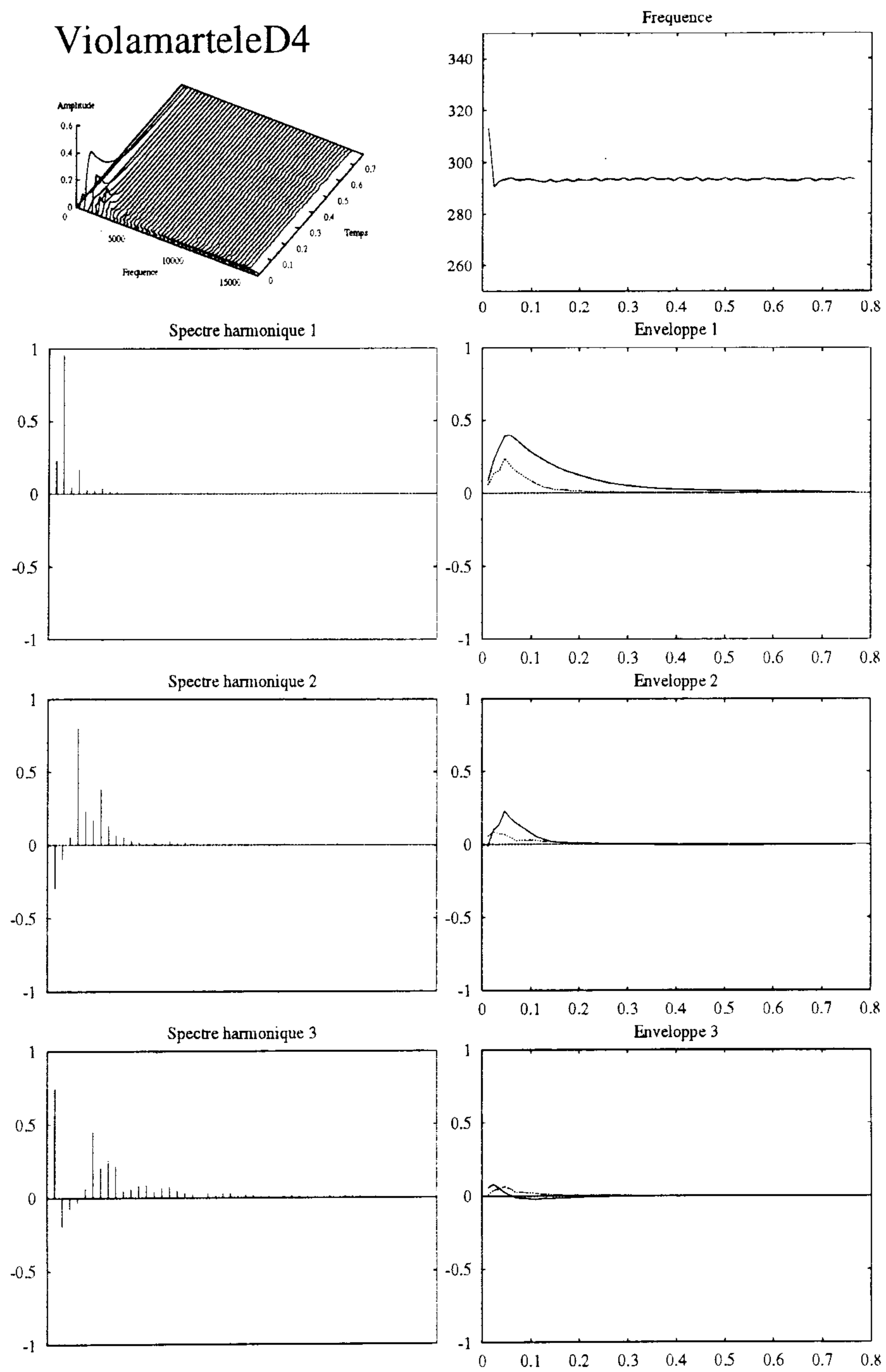


FIG. B.32 - *Alto*

FIG. B.33 – *Violon*

FIG. B.34 – *Violoncelle joué martelé*

FIG. B.35 – *Alto joué martele*



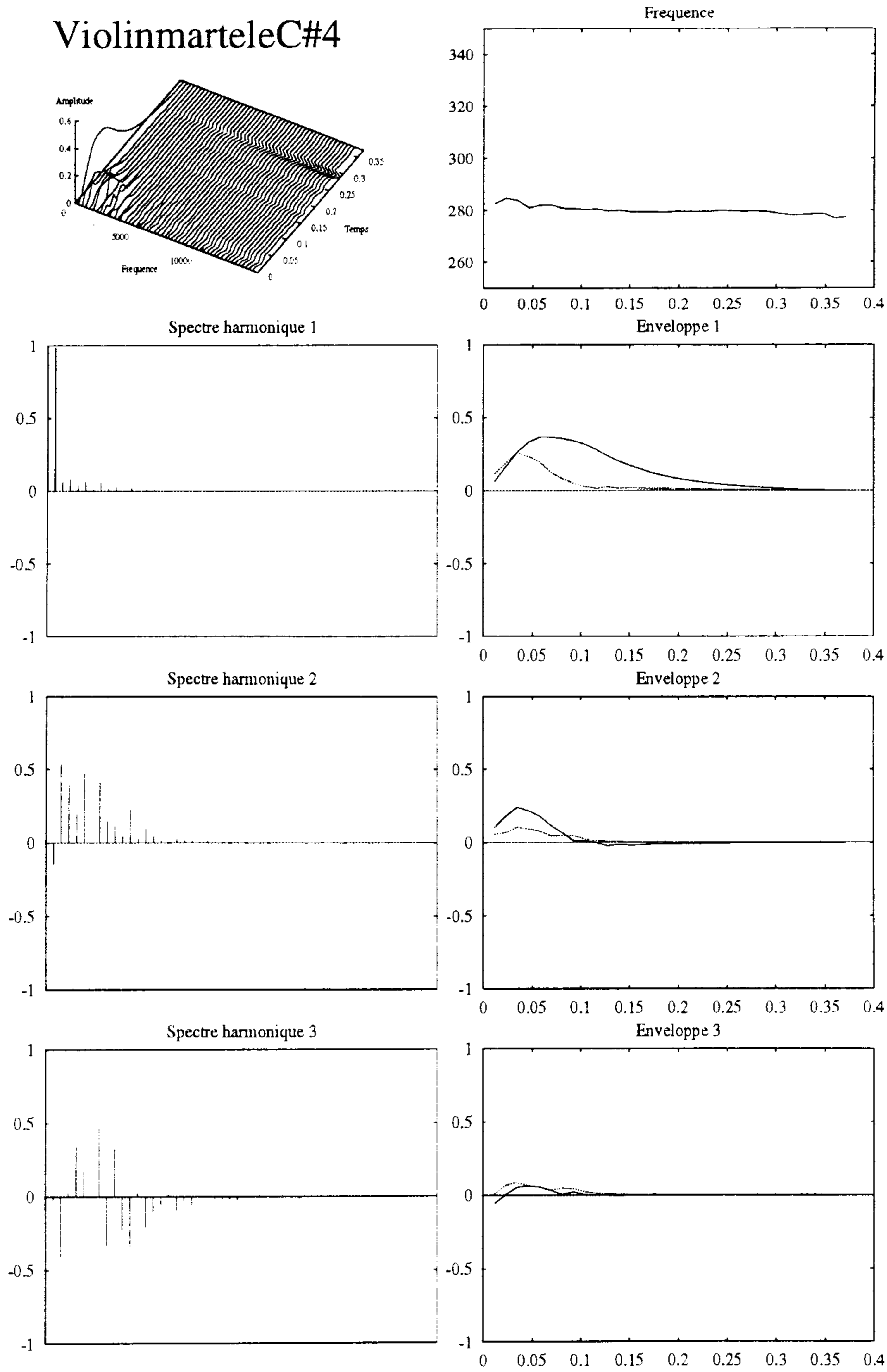


FIG. B.36 – Violon joué martele

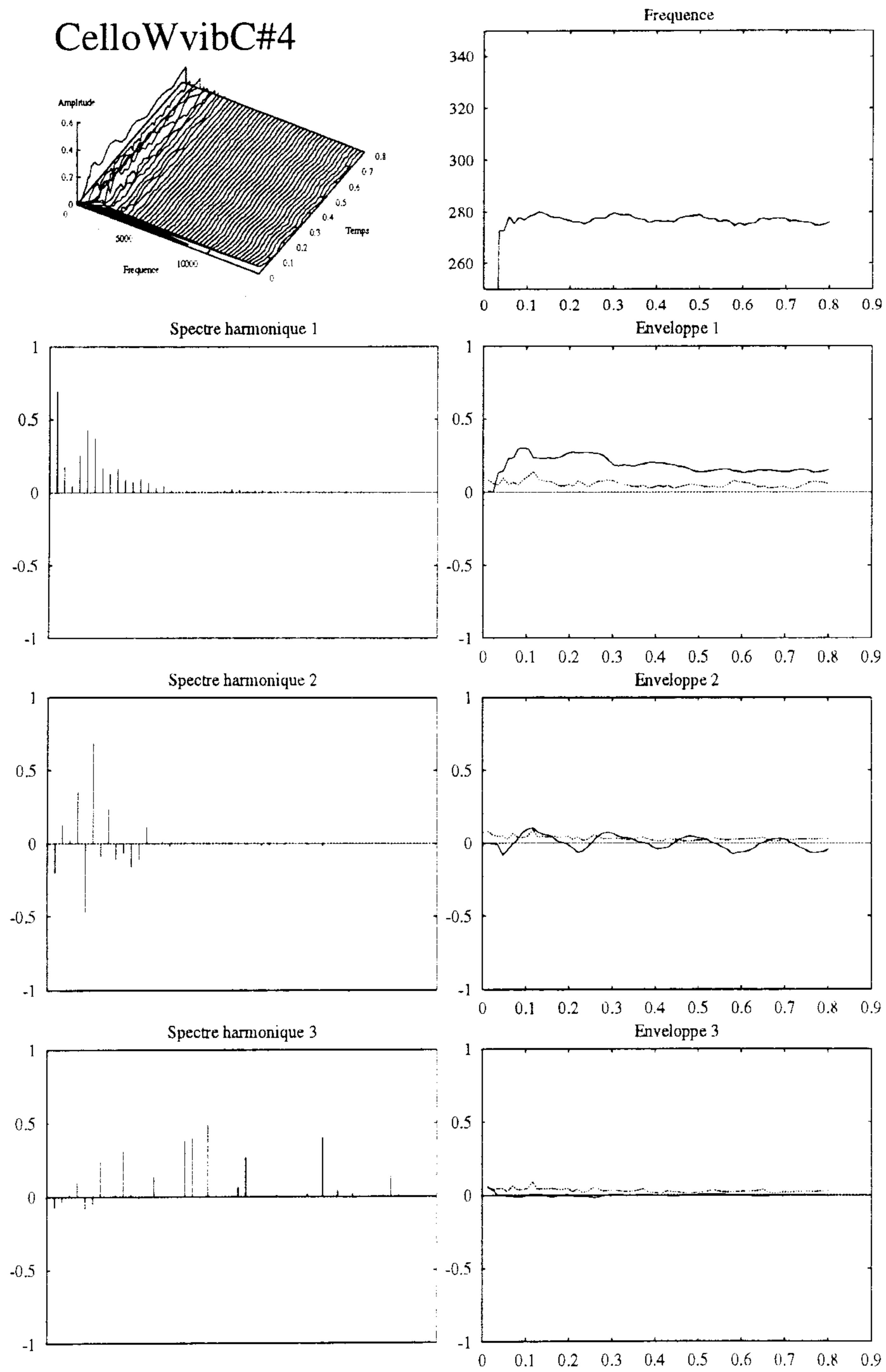
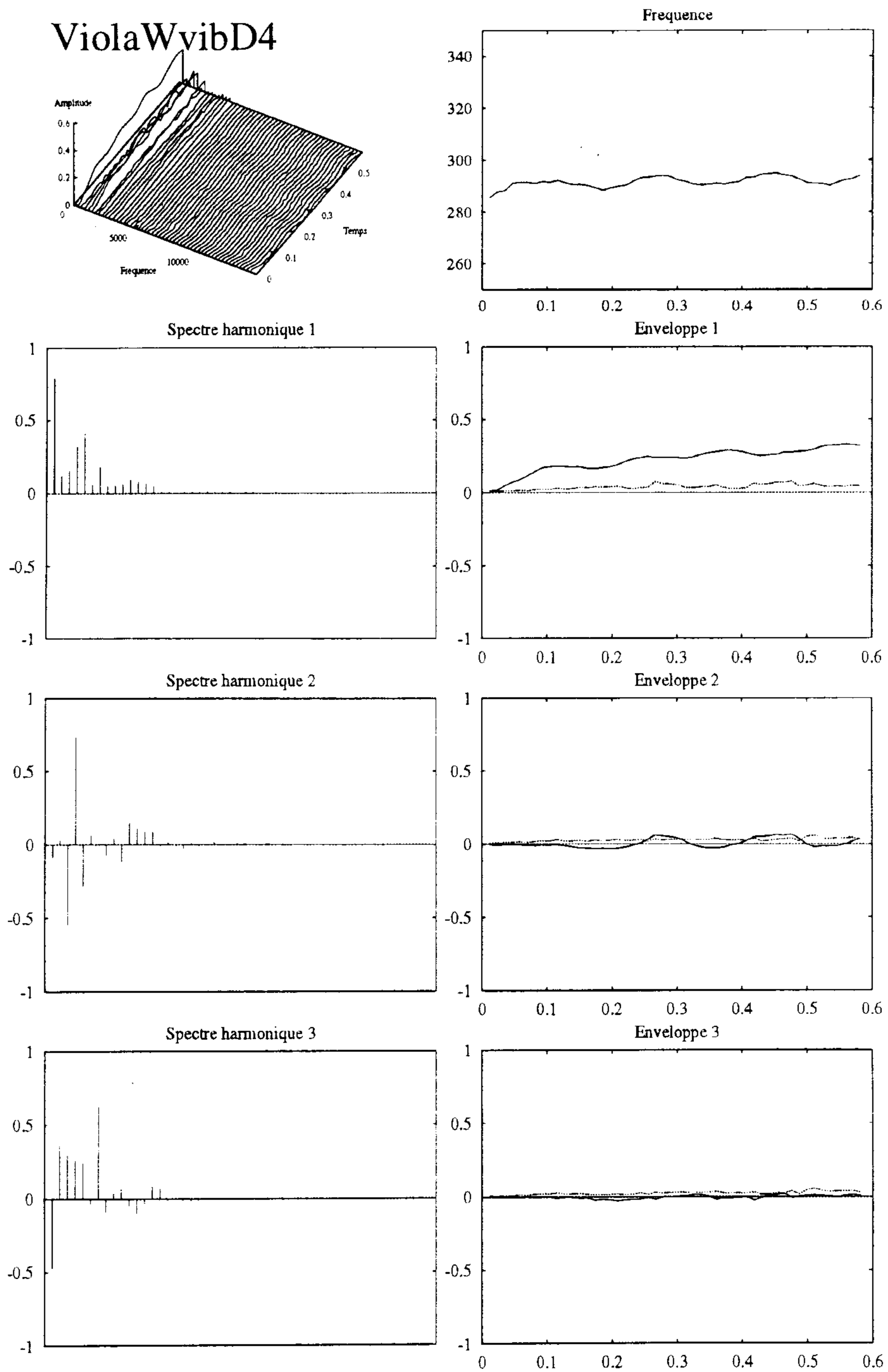


FIG. B.37 – Violoncelle joué avec vibrato

FIG. B.38 – *Alto joué avec vibrato*

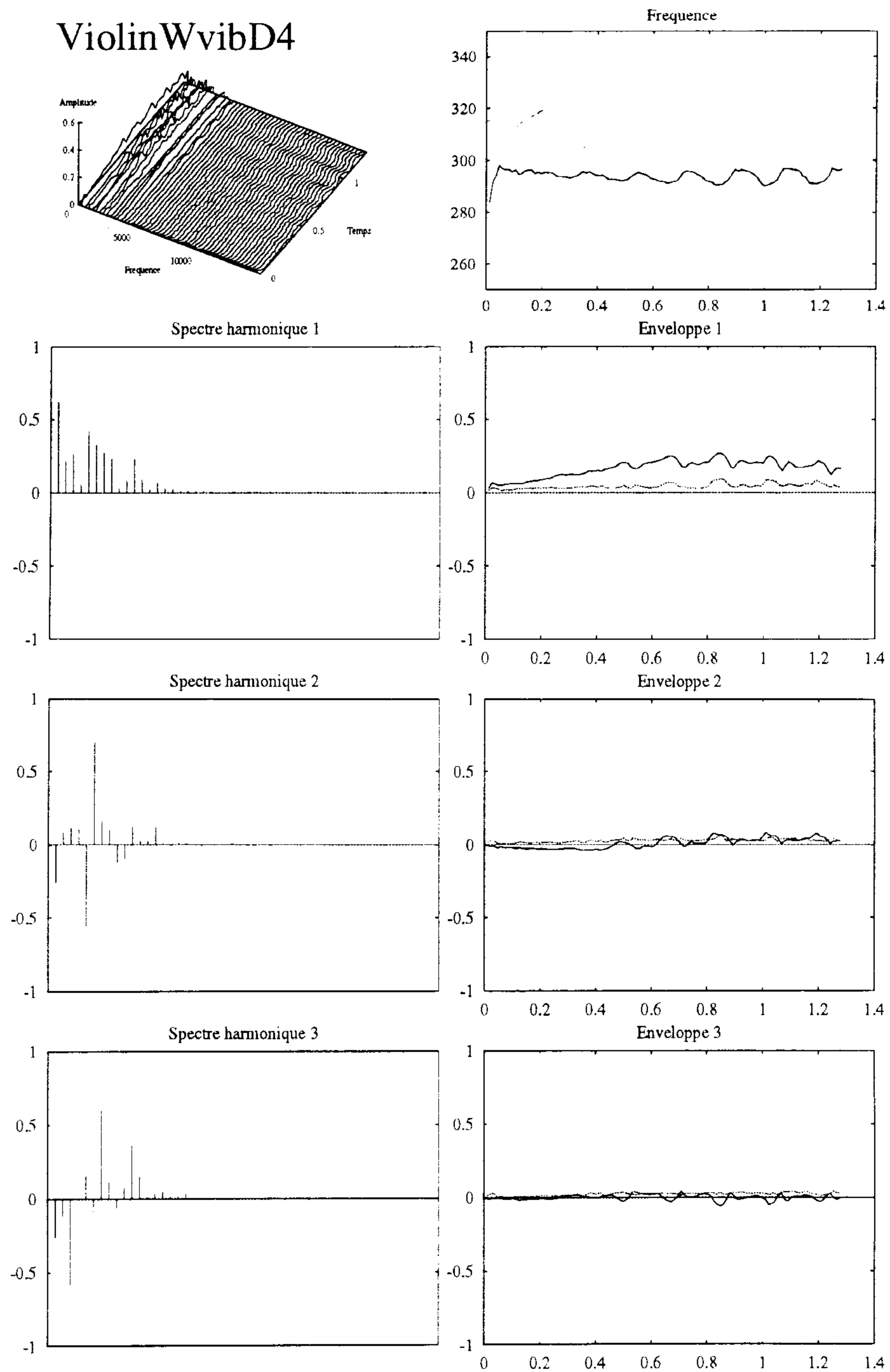


FIG. B.39 – *Violon joué avec vibrato*

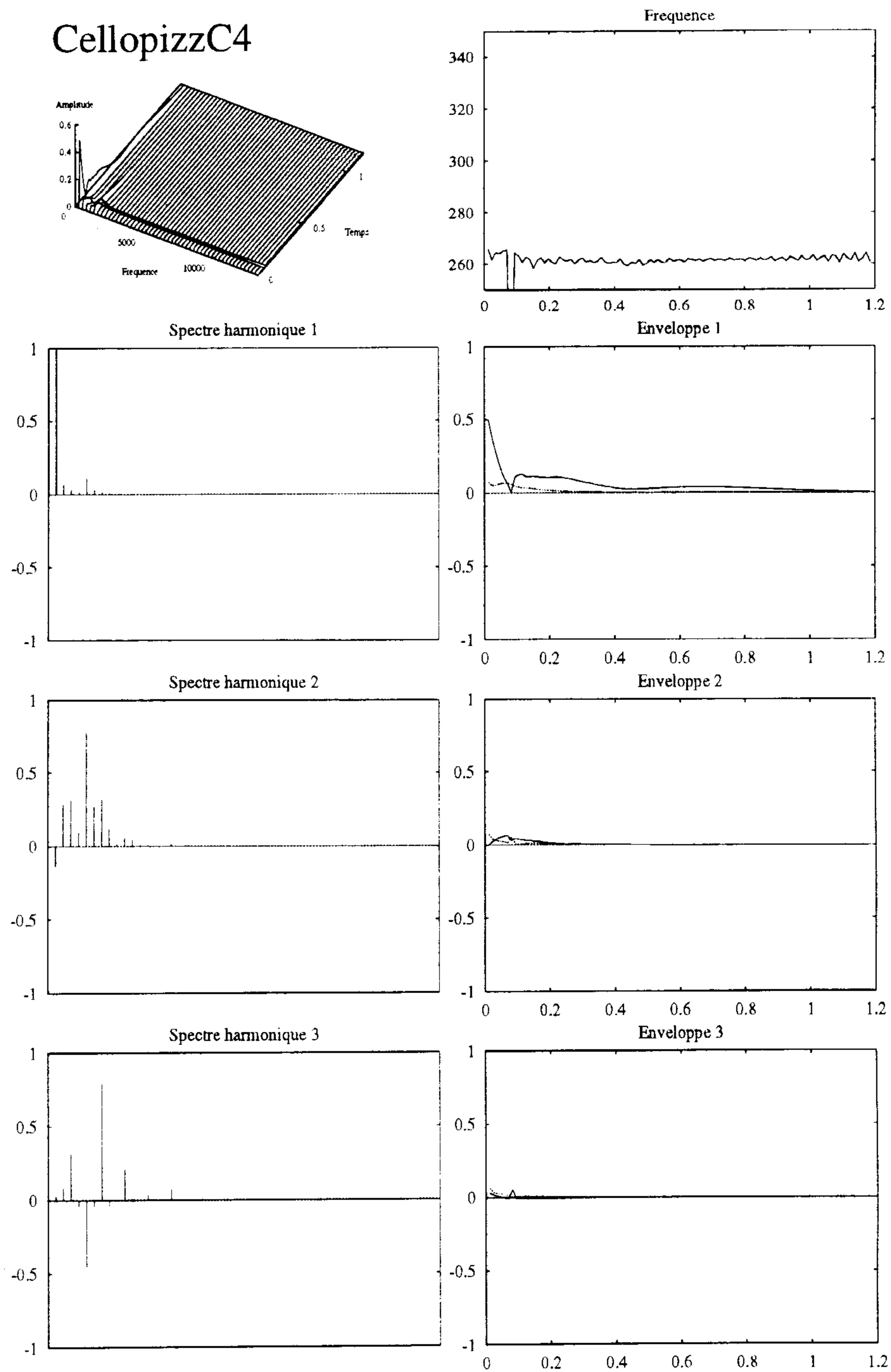
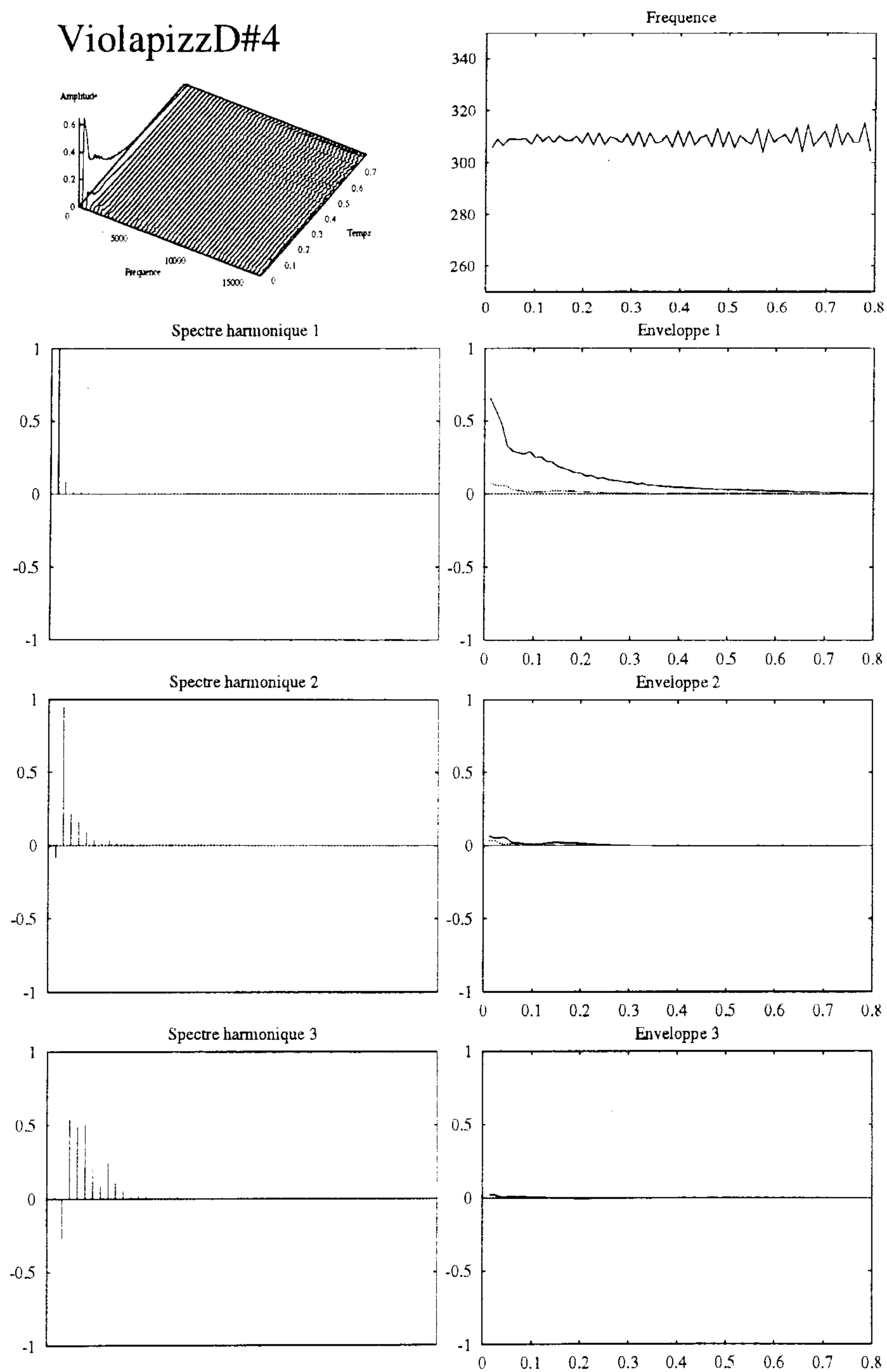


FIG. B.40 – Violoncelle joué pizzicato

FIG. B.41 – *Alto joué pizzicato*

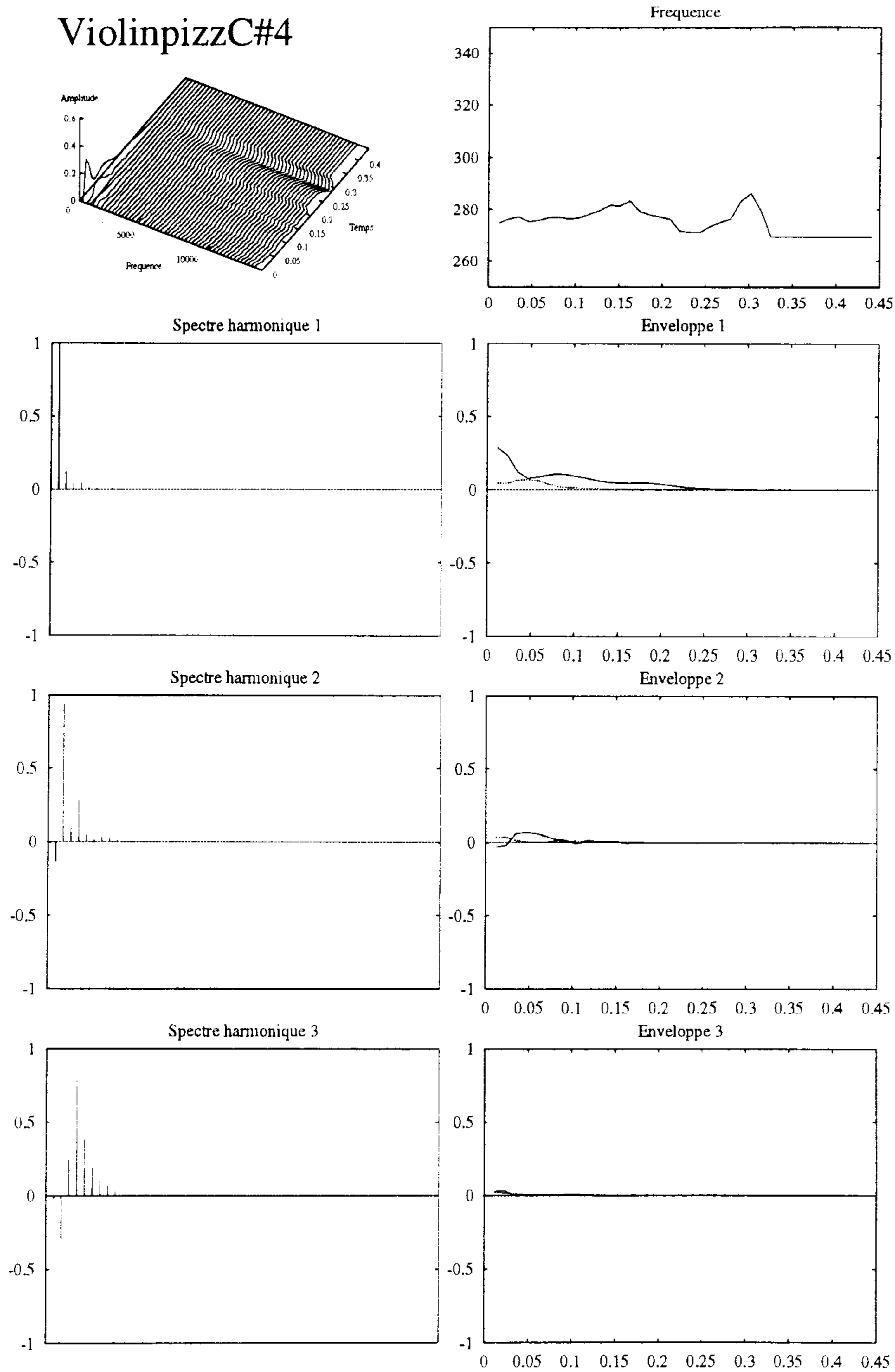


FIG. B.42 – Violon joué pizzicato

# Bibliographie

*La biblioteca si difende da sola,  
insondabile come la verità che ospita,  
ingannevole come la menzogna che custodisce.  
Labirinto spirituale, è anche labirinto terreno.  
Potreste entrare e potreste non uscire.*

Umberto Eco

*Il nome della rosa*



- [AA95] Abe (Mototsugu) et Ando (Shigeru). – Nonlinear time-frequency domain operators for decomposing sounds into loudness, pitch and timbre. *ICASSP-95 Conference Proceedings*, vol. 2, Mai 1995, pp. 1368–1371.
- [All77] Allen (J.B.). – Short term spectral analysis, synthesis, and modification by discrete fourier transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 25, n° 3, Juin 1977, pp. 235–238.
- [And93] Ando (Shigeo). – Statistical study of spectral parameters in musical instrument tones. *Journal of the Acoustical Society of America*, vol. 94, n° 1, Juillet 1993, pp. 37–45.
- [AR77] Allen (Jont B.) et Rabiner (Lawrence R.). – A unified approach to short-time fourier analysis and synthesis. *Proceedings of the IEEE*, vol. 65, n° 11, Novembre 1977, pp. 1558–1564.
- [ASS90] Asano (Futoshi), Suzuki (Yoiti) et Sone (Toshio). – Role of spectral cues in median plane localization. *Journal of the Acoustical Society of America*, vol. 88, n° 1, Juillet 1990, pp. 159–168.
- [BBC<sup>+</sup>97] Bernadas (Sal), Bian (Jieren), Cope (Bryan), Dong (Qiujie), Jaric (Melita), Jenkins (Michael), Munuswamy (Murali), Prihadi (Kartika) et Ramaswamy (Krishna). – A complete single-chip general midi (gm) wavetable synthesizer. – AES Preprint 4503 (O1), Mars 1997.
- [BCDS89] Botte (Marie-Claire), Canévet (Georges), Demany (Laurent) et Sorin (Christel). – *Psychoacoustique et perception auditive*. – INSERM/SFA/CNET, 1989.
- [BDC91] Bakamidis (S.), Dendrinos (M.) et Carayannis (G.). – Svd analysis by synthesis of harmonic signals. *IEEE Transactions Signal Processing*, vol. 39, n° 2, Février 1991, pp. 473–476.
- [Bea97a] Beaufils (Daniel). – Modulation d'amplitude et modulation de fréquence: les avantages du domaine sonore. *Bulletin de l'Union des Physiciens*, no793, Avril 1997, pp. 741–749.
- [Bea97b] Beaufils (Daniel). – Modulation de fréquence et synthèse f.m. *Bulletin de l'union des physiciens*, vol. 91, n° 794, Mai 1997, pp. 929–945.

- [Bel96] Bellanger (Maurice). – *Traitement numérique du signal*. – Masson, 1996, 5<sup>e</sup> édition.
- [Ber64] Berger (Kenneth W.). – Some factors in the recognition of timbre. *Journal of the Acoustical Society of America*, vol. 36, Octobre 1964, pp. 1888–1891.
- [Ber70] Beranek (L.L.). – Digital synthesis of speech and music. *IEEE Transactions on audio*, vol. 18, n° 4, Décembre 1970, pp. 427–433.
- [BH95] Beauchamp (James W.) et Horner (Andrew). – Wavetable interpolation synthesis based on time-variant spectral analysis of musical sounds. *Audio Engineering Society Preprint*, Février 1995.
- [Boa92a] Boashash (B.). – Estimating and interpreting the instantaneous frequency of a signal. part 1: Fundamentals. *Proceedings of the IEEE*, vol. 80, n° 4, Avril 1992, pp. 519–538.
- [Boa92b] Boashash (B.). – Estimating and interpreting the instantaneous frequency of a signal. part 2: Algorithms and applications. *Proceedings of the IEEE*, vol. 80, n° 4, Avril 1992, pp. 539–568.
- [Bol28] Boll (Marcel). – Vers une électrification de la musique. *La Science et La Vie*, no128, Février 1928, pp. 131–133.
- [Bol29] Boll (Marcel). – La production de musique par les courants alternatifs, application à l'électroacoustique. *La Science et La Vie*, no150, Décembre 1929, pp. 453–462.
- [BP93] Brown (Judith C.) et Puckette (Miller S.). – A high resolution fundamental frequency determination based on phase changes of the fourier transform. *Journal of the Acoustical Society of America*, vol. 94, n° 2, Août 1993, pp. 663–667.
- [Bra33] Brachet (Charles). – L'orgue des ondes du poste parisien et la synthèse de la musique. *La Science et La Vie*, no187, 1933, pp. 25–31.
- [Bro82] Brown (Franck). – *La musique par ordinateur*. – Presses Universitaires de France, 1982, *Que sais-je*.
- [BS80] Bouroche (Jean-Marie) et Saporta (Gilbert). – L'analyse des données. *In: Que sais-je?*, chap. 2. – 108, Bd Saint-Germain, 75006 Paris, Presses Universitaires de France, 1<sup>ère</sup> édition, 1980.
- [Cah97] Cahill (Thaddeus). – Art of and apparatus for generating and distributing music electrically. – US Patent Number 580,035, Avril 1897.
- [Cah14] Cahill (Thaddeus). – Music generating and music distributing apparatus. – US Patent Number 1,107,261, Août 1914.

- [Cah17] Cahill (Thaddeus). – Art of and apparatus for generating and distributing music electrically. – US Patent Number 1,213,804, Janvier 1917.
- [Cah19] Cahill (Thaddeus). – Art of and apparatus for generating and distributing music electrically. – US Patent Number 1,295,691, Février 1919.
- [Cam94] Campbell (Shari L.). – Uni- and multidimensional identification of rise time, spectral slope, and cutoff frequency. *Journal of the Acoustical Society of America*, vol. 96, n° 3, Septembre 1994, pp. 1380–1387.
- [CH96] Cheung (Ngai-Man) et Horner (Andrew B.). – Group synthesis with genetic algorithms. *Journal of the Audio Engineering Society*, vol. 44, n° 3, Mars 1996, pp. 130–147.
- [Cha81] Charbonneau (Gérard). – Timbre and the perceptual effects of three types of data reduction. *Computer Music Journal*, vol. 5, n° 2, 1981, pp. 10–19.
- [Che96] Chen (Wei). – Performance of cascade and parallel iir filters. *Journal of the Audio Engineering Society*, vol. 44, n° 3, Mars 1996, pp. 148–258.
- [Cho73] Chowning (John M.). – The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, vol. 21, n° 7, Septembre 1973, pp. 526–534.
- [Cho77] Chowning (John M.). – Method of synthesizing a musical sound. – US Patent Number 4.018,121, Avril 1977.
- [Chr90] Chrysafis (Andreas). – Digital sine-wave synthesis using the dsp56001/2. – Notice d'application APR1, 1990.
- [Col60] Collectif. – American standard acoustical terminology. – American Standards Association, 1960.
- [Col80] Collectif. – *Sons et musique*. – Belin/Pour la science, 1980.
- [Col95] Collectif. – *Les instruments de l'orchestre*. – Belin/Pour la science, 1995.
- [d'A89] d'Alessandro (Christophe). – *Représentation du signal de parole par une somme de fonctions élémentaires*. – Thèse de doctorat, Université Paris 6, 1989.
- [d'A90] d'Alessandro (Christophe). – Time-frequency speech transformation based on an elementary waveform representation. *Speech Communication*, vol. 9, 1990, pp. 419–431.
- [DD82] Deutsch (Ralph) et Deutsch (Leslie J.). – Digital noise generator for electronic musical instruments. – US Patent Number 4,327,419, Avril 1982.
- [DH95] Deforeit (Christian) et Heickroth (Jim). – A music synthesizer architecture that integrates a specialized dsp core and a 16-bit microprocessor on a single chip. – AES 98th convention Preprint 3958, Février 1995.

- [dL85] de Lagarde (Jean). – *Initiation à l'analyse des données*. – Dunod, 1985.
- [Dol60] Dolanský (Ladislav). – Choice of base signals in speech signal analysis. *IRE Transactions on audio*, vol. 18, Novembre/Décembre 1960.
- [Dol86] Dolson (Mark). – The phase vocoder: A tutorial. *Computer Music Journal*, vol. 10, n° 4, 1986, pp. 14–27.
- [Dou54] Douglas (Alan Lockhart Monteith). – *The electronic musical instrument manual; a guide to theory and design*. – London, Pitman, 1954, 2nd édition.
- [DQ97] Ding (Yinong) et Qian (Xiaoshu). – Processing of musical tones using combined quadratic polynomial-phase sinusoid and residual (quasar) signal model. *Journal of the Audio Engineering Society*, vol. 45, n° 7/8, Juillet Août 1997, pp. 571–584.
- [DR94] Depalle (Philippe) et Rodet (Xavier). – Sound synthesis process. – US Patent Number 5,401,897, 1994.
- [DR95] Ding (Yinong) et Rossum (David). – Filter morphing of parametric equalizers and shelving filters for audio signal processing. *Journal of the Audio Engineering Society*, vol. 43, n° 10, Octobre 1995, pp. 821–825.
- [dYD95] d'Alessandro (C.), Yegnanarayana (B.) et Darsinos (V.). – Decomposition of speech signals into deterministic and stochastic components. *ICASSP-95 Conference Proceedings*, vol. 1, Mai 1995, pp. 760–763.
- [Ell92] Ellis (Daniel Patrick Whittlesey). – *A Perceptual Representation of Audio*. – Thèse, Massachusetts Institute of Technology, Février 1992.
- [EP94] Ethington (Russ) et Punch (Bill). – Seawave: A system for musical timbre description. *Computer Music Journal*, vol. 18, n° 1, Printemps 1994, pp. 30–39.
- [EV94] Edwards (Brent W.) et Viemeister (Neal F.). – Frequency modulation versus amplitude modulation discrimination: Evidence for a second frequency modulation encoding mechanism. *Journal of the Acoustical Society of America*, vol. 96, n° 2, Août 1994, pp. 733–739.
- [Eva93] Evangelista (Gianpaolo). – Pitch-synchronous wavelet representations of speech and music signals. *IEEE Transactions on Signal Processing*, vol. 41, n° 12, Décembre 1993, pp. 3313–3330.
- [FH96] Fitz (Kelly) et Haken (Lippold). – Sinusoidal modeling and manipulation using lemur. *Computer Music Journal*, vol. 20, n° 4, 1996, pp. 44–59.
- [For06] Forest (Lee De). – The audion, a new receiver for wireless telegraphy. *Transactions of the Institute of Electrical Engineering*, vol. 25, Octobre 1906, p. 735.

- [For20] Forest (Lee De). – The audion, its action and some recent applications. *Journal of the Franklin Institute*, vol. 190, n° 1, Juillet 1920, pp. 1–38.
- [Fre67] Freedman (M. David). – Analysis of musical instrument tones. *Journal of the Acoustical Society of America*, vol. 41, n° 4, Août 1967, pp. 793–806.
- [Fuj96a] Fujita (Yoshio). – Filter apparatus for an electronic musical instrument. – US Patent 5,583,309, 1996.
- [Fuj96b] Fujita (Yoshio). – Waveform generating apparatus for musical instrument. – US Patent 5,553,011, 1996.
- [GB84] Goeddel (T.W.) et Bass (S.C.). – High quality synthesis of musical voices in discrete time. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 32, n° 3, Juin 1984, pp. 623–633.
- [GGPF78] Goldstein (J.L.), Gerson (A.), P.Srulovicz et Furst (M.). – Verification of the optimal probabilistic basis of aural processing in pitch of complex tones. *Journal of the Acoustical Society of America*, vol. 63, n° 2, Février 1978, pp. 486–497.
- [Giv29] Givelet (Armand). – Orgue électrique. – Brevet d'invention 669.433, Août 1929.
- [Giv34] Givelet (Armand). – L'orgue électronique. *l'onde électrique*, vol. 13, n° 148, 1934, pp. 157–171.
- [GL84] Griffin (Daniel W.) et Lim (Jae S.). – Signal estimation from modified short-time fourier transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 32, n° 2, Avril 1984, pp. 236–242.
- [GM77] Grey (John M.) et Moorer (James A.). – Perceptual evaluations of synthesized musical instrument tones. *Journal of the Acoustical Society of America*, vol. 62, n° 2, Août 1977, pp. 454–462.
- [GM94] Gardner (Bill) et Martin (Keith). – *HRTF Measurements of a KEMAR Dummy-Head Microphone*. – Rapport technique n° 280, MIT Media Lab Perceptual Computing, Mai 1994.
- [Gra76] Gray (Elisha). – Electro-harmonic telegraph. – US Patent Number 173,618, Février 1876.
- [Gra77] Gray (Elisha). – Transmission des sons musicaux. *Annales Télégraphiques*, vol. 3, n° 4, Mars Avril 1877, p. 97 107.
- [Gra78a] Gray (Elisha). – Art of transmitting rhythmical vibrations in an electric circuit. – US Patent Number 205,378, Juin 1878.

- [Gra78b] Gray (Elisha). – *Experimental Researches in Electro-Harmonic Telegraphy and Telephony 1867-1878*. – New York 17, 19, 21, 23 Rose Street, Russell Brothers, 1878.
- [Gra80] Gray (Elisha). – Telephonic telegraph. – US Patent Number 233,345, Octobre 1880.
- [Gre75] Grey (John M.). – *An Exploration of Musical Timbre*. – Rapport technique n° STAN M-2, Center for Computer Research in Music and Acoustics, Février 1975.
- [Gri94] Gribenski (André). – L'audition. *In: Que sais-je*, p. 3 15. – Presses Universitaires de France, 6ème édition, Février 1994.
- [GS92] George (E.B.) et Smith (M.J.T.). – Analysis-by-synthesis/overlap-add sinusoidal modeling applied to the analysis and synthesis of musical tones. *Journal of the Audio Engineering Society*, vol. 40, n° 6, Juin 1992, pp. 497–516.
- [GS94] George (E.B.) et Smith (M.J.T.). – Audio analysis/synthesis system. – U.S. Patent Number 5,327,518, 1994.
- [Hal91] Hall (Donald E.). – *Musical acoustics*. – Brooke/Cole Publishing Company, 1991, second édition.
- [Ham34] Hammond (Laurens). – Electrical musical instrument. – US Patent 1,956,350, avril 1934.
- [Ham38] Hammond (Laurens). – Electrical musical instruments. – US Patent 2,126,682, août 1938.
- [Har78] Harris (Frederic J.). – On the use of windows for harmonic analysis with discrete fourier transform. *Proceedings of the IEEE*, vol. 66, n° 1, Janvier 1978, pp. 51–83.
- [HB95] Horner (Andrew) et Beauchamp (James). – Synthesis of trumpet tones using a wavetable and a dynamic filter. *Journal of the Audio Engineering Society*, vol. 43, n° 10, Octobre 1995, pp. 799–812.
- [HBH93] Horner (A.), Beauchamp (J.) et Haken (L.). – Method for multiple wavetable synthesis of musical instruments tones. *Journal of the Audio Engineering Society*, vol. 41, n° 5, Mai 1993, pp. 336–356.
- [HCM97a] Hourdin (Christophe), Charbonneau (Gérard) et Moussa (Tarek). – A multidimensional scaling analysis of musical instruments' time-varying spectra. *Computer Music Journal*, vol. 21, n° 2, 1997, pp. 40–55.
- [HCM97b] Hourdin (Christophe), Charbonneau (Gérard) et Moussa (Tarek). – A sound-synthesis technique based on multidimensional scaling of spectra. *Computer Music Journal*, vol. 21, n° 2, 1997, pp. 56–68.

- [Hel54] Helmholtz (Hermann L. F.). – *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. – New-York, Dover, 1954. Traduction anglaise de A-J Ellis.
- [HFM95] Houghton (Andy D.), Fisher (Andy J.) et Malet (Thierry F.). – An ASIC for digital additive sine-wave synthesis. *Computer Music Journal*, vol. 19, n° 3, 1995, pp. 26–31.
- [Hig96] Higashi (Iwao). – Electronic tone generator system with CPU and DSP. – US Patent 5,567,900, Octobre 1996.
- [HNY81] Hiyoshi (Teruo), Nakada (Akira) et Yamada (Shigeru). – Musical tone forming device by F.M. technology. – US Patent 4,253,367, 1981.
- [Hor95] Horner (Andrew). – Wavetable matching synthesis of dynamic instruments with genetic algorithms. *Journal of the Audio Engineering Society*, vol. 43, n° 11, Novembre 1995, pp. 916–931.
- [Hor97] Horner (Andrew). – Wavetable modeling of muted brass tones. *Journal of the Audio Engineering Society*, vol. 45, n° 4, Avril 1997, pp. 241–252.
- [Hou95] Hourdin (Christophe). – *Etude psychophysique du timbre - Application au codage et à la synthèse des sons en musique*. – Thèse de doctorat, Université Paris XI, 1995.
- [IK93a] Irino (T.) et Kawahara (H.). – Signal reconstruction from modified auditory wavelet transform. *IEEE Transactions on Signal Processing*, vol. 41, n° 12, Décembre 1993, pp. 3549–3554.
- [IK93b] Irino (Toshio) et Kawahara (Hideki). – *Signal Reconstruction from Modified Auditory Wavelet Transform*. – Rapport technique n° ISRL-93-2, NTT Basic Research Laboratories, 3-9-11 Midori-cho Musashino-shi Tokyo, 180 JAPAN, Nippon Telegraph and Telephone Corporation, Juillet 1993.
- [IMS97] Ikegaya (Yuji), Muramatsu (Shinichi) et Shirayanagi (Toru). – Digital signal processing circuit. – US Patent 5,636,153, Juin 1997.
- [IP94] Irino (Toshio) et Patterson (Roy D.). – A theory of asymmetric intensity enhancement around acoustic transients. *International Conference on Spoken Language Processing*, Septembre 1994, pp. 1955–1958.
- [Jaf87a] Jaffe (David A.). – Spectrum analysis tutorial, part 1: The discrete Fourier transform. *Computer Music Journal*, vol. 11, n° 2, 1987, pp. 9–24.
- [Jaf87b] Jaffe (David A.). – Spectrum analysis tutorial, part 2: Properties and applications of the discrete Fourier transform. *Computer Music Journal*, vol. 11, n° 3, 1987, pp. 17–35.
- [Jaf95] Jaffe (David A.). – Ten criteria for evaluating synthesis techniques. *Computer Music Journal*, vol. 19, n° 1, 1995, pp. 76–87.

- [JP88] Jones (Douglas L.) et Parks (Thomas W.). – Generation and combination of grains for music synthesis. *Computer Music Journal*, vol. 12, n° 2, 1988, pp. 27–34.
- [Kee72a] Keeler (J.S.). – The attack transients of some organ pipes. *IEEE Transactions on audio and electroacoustics*, vol. 20, n° 5, Décembre 1972, pp. 378–391.
- [Kee72b] Keeler (J.S.). – Piecewise-periodic analysis of almost-periodic sounds and musical transients. *IEEE Transactions on audio and electroacoustics*, vol. 20, n° 5, Décembre 1972, pp. 339–344.
- [Kes34] Keszler (Pierre). – La science vient d'enrichir la musique de nouveaux instruments très expressifs. *La Science et La Vie*, no207, Septembre 1934, pp. 195–204.
- [Kes37] Keszler (Pierre). – Une révolution en musique polyphonique. *La Science et La Vie*, no241, Juillet 1937, pp. 28–33.
- [Kle89] Kleczkowski (Piotr). – Group additive synthesis. *Computer Music Journal*, vol. 13, n° 1, 1989, pp. 12–20.
- [KM88] Kronland-Martinet (Richard). – The wavelet transform for analysis, synthesis, and processing of speech and music sounds. *Computer Music Journal*, vol. 12, n° 4, 1988, pp. 11–20.
- [KR95] Karjalainen (Matti) et Rahkila (Martti). – Learning signal processing concepts and psychoacoustics in the quicksig dps environment. *ICASSP-95 Conference Proceedings*, vol. 2, Mai 1995, pp. 1125–1128.
- [Lar93] Laroche (J.). – The use of matrix pencil method for the spectrum analysis of musical signals. *Journal of the Acoustical Society of America*, vol. 94, n° 4, Octobre 1993, pp. 1958–1965.
- [Lau90] Laurendeau (Jean). – *Maurice Martenot, luthier de l'électronique*. – Dervy Livres, 1990.
- [LCW95] Lin (Gang-Janp), Chen (Sau-Gee) et Wu (Terry). – High quality and low complexity pitch modification of acoustic signals. *ICASSP-95 Conference Proceedings*, vol. 5, May 1995, pp. 2987–2990.
- [LD95] Lu (Shan) et Doerschuk (Peter C). – Modeling and processing speech with sums of am-fm formant models. *ICASSP-95*, vol. 1, Mai 1995, pp. 764–767.
- [Lei84] Leipp (Emile). – *Acoustique et musique*. – Paris, Seuil, 1984.
- [Len89] Lent (Keith). – An efficient method for pitch shifting digitally sampled sounds. *Computer Music Journal*, vol. 13, n° 4, 1989, pp. 65–71.
- [Loh85] Lohner (Henning). – Interview with robert moog. *Computer Music Journal*, vol. 9, n° 4, 1985, pp. 62–65.



- [lsr95] ller (Henrik Mø) et sørensen (Michael Friis). – Head-related transfer functions of human subjects. *Journal of the Audio Engineering Society*, vol. 43, n° 5, Mai 1995, pp. 300–321.
- [Mah90] Maher (Robert C.). – Evaluation of a method for separating digitized duet signals. *Journal of the Audio Engineering Society*, vol. 38, n° 12, Décembre 1990, pp. 956–979.
- [Mar30] Martenot (Maurice). – Electric musical instrument. – US Patent Number 1,853,630, Août 1930.
- [Mar95] Martin (Keith Dana). – *A Computational Model of Spatial Hearing*. – Thèse, Massachusetts Institute of Technology, Juin 1995.
- [MB94] Maher (Robert C.) et Beauchamp (James W.). – Fundamental frequency estimation of musical signals using a two-way mismatch procedure. *Journal of the Acoustical Society of America*, vol. 95, n° 4, Avril 1994, pp. 2254–2263.
- [McA84] McAulay (Robert J.). – Maximum likelihood spectral estimation and its application to narrow-band speech processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 32, n° 2, Avril 1984, pp. 243–251.
- [Mey94] Meyer (Y.). – *Les ondelettes, algorithmes et applications*. – Paris, Armand Colin, 1994, 2ème édition.
- [Moo65] Moog (Robert Arthur). – Voltage-controlled electronic music modules. *Journal of the Audio Engineering Society*, vol. 13, n° 3, Juin 1965, pp. 200–206.
- [Moo77a] Moore (F. Richard). – Table lookup noise for sinusoidal digital oscillators. *Computer Music Journal*, vol. 1, n° 2, 1977, pp. 26–29.
- [Moo77b] Moorer (James Anderson). – Signal processing aspects of computer music: A survey. *Proceedings of the IEEE*, vol. 65, n° 8, Août 1977, pp. 1108–1137.
- [MPV93] Monti (Carlo M.), Pierobon (Gianfranco L.) et Viaro (Umberto). – Synthesis of spectral densities using finite automata. *IEEE Transactions on Signal Processing*, vol. 41, n° 5, Mai 1993, pp. 1993–1996.
- [MQ86] McAulay (Robert J.) et Quatieri (Thomas F.). – Speech analysis/synthesis based on a sinusoidal representation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-34, n° 4, Août 1986, pp. 744–754.
- [MR93] Massie (Dana C.) et Rossum (David P.). – Digital sampling instrument. – US Patent 5,248,845, Septembre 1993.
- [MS72] Mullis (Clifford T.) et Steiglitz (Kenneth). – Circulant markov chains as digital signal sources. *IEEE Transactions on Audio and Electroacoustics*, vol. AU-20, n° 4, Octobre 1972, pp. 246–248.

- [MS77] Mitchell (Richard L.) et Stone (C.R.). – Table-lookup methods for generating arbitrary random numbers. *IEEE Transactions on Computers*, vol. C-26, n° 10, Octobre 1977, pp. 1006–1008.
- [MT97] Masahiro (Koyama) et Tetsuo (Nishimoto). – Musical tone signal synthesizer. – US Patent 5,610,354, Mars 1997.
- [MW93] Matsunaga (Kaoru) et Washiyama (Yutaka). – Filter coefficient generator for electronic musical instruments. – US Patenet 5,270,481, 1993.
- [NBD93] Nawab (Hamid), Beyerbach (Daniel M.) et Dorken (Erkan). – Principal decomposition of time-frequency distributions. *IEEE Transactions on Signal Processing*, vol. 41, n° 11, Novembre 1993, pp. 3182–3186.
- [NNO81] Nagai (Yohei), Nishimoto (Tetsuo) et Okamoto (Shimaji). – Electronic musical instrument. – US Patent 4,301,704, 1981.
- [NQ83] Nawab (S. Hamid) et Quatieri (Thomas F.). – Signal reconstruction from short-time fourier transform magnitude. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, n° 4, Août 1983, pp. 236–242.
- [OBT60] Olson (H.F.), Belar (H.) et Timmens (J.). – Electronic music synthesis. *Journal of the Acoustical Society of America*, vol. 32, n° 3, Mars 1960, pp. 311–319.
- [OIWH97] Okamoto (Seiji), Ishii (Katsushi), Washiyama (Yutaka) et Hirano (Sayoko). – Apparatus for and method of generating musical tones. – US Patent 5,641,929, Juin 1997.
- [Pap83] Papoulis (Athanasios). – Random modulation: A review. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, n° 1, Février 1983, pp. 97–105.
- [PB83] Petersen (Tracy L.) et Boll (Steven F.). – Critical band analysis-synthesis. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, n° 3, Juin 1983, pp. 656–663.
- [Pee96] Peers (Roi N.). – Method and integrated circuit for the flexible combination of four operators in sound synthesis. – US Patent 5,581,045, 1996.
- [Pol78] Pollack (Irwin). – Decoupling of auditory pitch and stimulus frequency: The shepard demonstration revisited. *Journal of the Acoustical Society of America*, vol. 63, n° 1, Janvier 1978, pp. 202–206.
- [PTVF92] Press (William H.), Teukolsky (Saul A.), Vetterling (William T.) et Flannery (Brian P.). – *Numerical recipes in C, the art of scientific computing*. – Cambridge University Press, 1992, second édition.

- [Puc95] Puckette (Miller). – Formant-based audio synthesis using nonlinear distortion. *Journal of the Audio Engineering Society*, vol. 43, n° 1/2, Janvier/Février 1995, pp. 40–47.
- [Puj90] Pujol (Rémy). – Le traitement du son dans l'oreille interne. *Pour la science*, no154, Août 1990, pp. 20–29.
- [RB93a] Roberts (B.) et Bailey (P.J.). – Spectral pattern and the perceptual fusion of harmonics: I. the role of temporal factors. *Journal of the Acoustical Society of America*, vol. 94, n° 6, Décembre 1993, pp. 3153–3165.
- [RB93b] Roberts (B.) et Bailey (P.J.). – Spectral pattern and the perceptual fusion of harmonics: II. a special status for added components? *Journal of the Acoustical Society of America*, vol. 94, n° 6, Décembre 1993, pp. 3165–3177.
- [RB98] Rochebois (Thierry) et Beaufils (Daniel). – L'histoire des synthétiseurs de musique : plus d'un siècle d'évolutions technologiques. *Bulletin de l'Union des Physiciens*, 1998. – A paraître.
- [RC97a] Rochebois (Thierry) et Charbonneau (Gérard). – Application de l'analyse en composantes principales à la synthèse de sons musicaux hybrides. *Systematische Musikwissenschaft*, vol. 4, n° 1-2, 1997, pp. 201–219.
- [RC97b] Rochebois (Thierry) et Charbonneau (Gérard). – Cross-synthesis using interverted principal harmonic sub-spaces. In: *Music, Gestalt, and Computing*, chap. 4, pp. 375–385. – Springer, 1997.
- [Rd96] Richard (Gaël) et d'Alessandro (Christophe). – Analysis/synthesis and modification of the speech aperiodic component. *Speech Communication*, vol. 19, 1996, pp. 221–244.
- [RdG92] Richard (Gaël), d'Alessandro (Christophe) et Grau (Sophie). – Unvoiced speech analysis and synthesis using poissonian random formant-wavefunctions. In: *Proceedings of EUSIPCO-92*, éd. par J.Vandewale, Boite (R.), M.Moonen et A.Oosterlinck.
- [RdG93] Richard (G.), d'Alessandro (C.) et Grau (S.). – Musical noises synthesis using random formant waveforms. In: *Proceedings of the Stockholm Music Acoustics Conference*. Royal Swedish Academy of Music.
- [Ris67] Risset (Jean-Claude). – *Sur l'analyse, la synthèse et la perception des sons, étudiées à l'aide de calculateurs électroniques*. – Thèse de doctorat, Faculté des Sciences d'Orsay, 1967.
- [RM69] Risset (Jean-Claude) et Mathews (Max V.). – Analysis of musical-instrument tones. *Physics Today*, vol. 22, n° 2, Février 1969, pp. 23–30.
- [Roa88] Roads (Curtis). – Introduction to granular synthesis. *Computer Music Journal*, vol. 12, n° 2, 1988, pp. 11–13.

- [Roa95] Roads (Curtis). – *The Computer Music Tutorial*. – Cambridge, Massachusetts, The MIT Press, 1995.
- [Rob80] Robertson (G.H.). – Signal characterization using spectral microanalysis. *Journal of the Acoustical Society of America*, vol. 67, n° 5, Mai 1980, pp. 1663–1669.
- [Roc97] Rochebois (Thierry). – Techno basics. – série d'articles de vulgarisation du magazine "Keyboards", Mars à Décembre 1997.
- [Ros92a] Rossum (David P.). – digital sampling instrument for audio data. – US Patent 5,111,727, Mai 1992.
- [Ros92b] Rossum (David P.). – dynamic digital iir audio filter and method which provides dynamic digital filtering for audio signals. – US Patent 5,170,369, Décembre 1992.
- [Rou28] Roussel (Joseph). – La production radioélectrique de la musique. *La Science et La Vie*, no134, Août 1928, pp. 108–110.
- [Sad84] Sadie (Stanley) (édité par). – *The New Grove Dictionary of Musical Instruments*. – Macmillan press limited, 1984.
- [Sai93] Sait (Tsutomu). – Tone signal processing apparatus for pcm waveform interpolation and filtering. – US Patent 5,218,155, 1993.
- [SB88] Stapleton (John C.) et Bass (Steven C.). – Synthesis of musical tones based on the karhunen-loève transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, n° 3, Mars 1988, pp. 305–319.
- [SC67] Strong (William) et Clark (Melville). – Synthesis of wind-instrument tones. *Journal of the Acoustical Society of America*, vol. 41, n° 1, 1967, pp. 39–52.
- [Sch86] Schroeder (Manfred R.). – Auditory paradox based on fractal waveform. *Journal of the Acoustical Society of America*, vol. 79, n° 1, Janvier 1986, pp. 186–189.
- [She64] Shepard (Roger N.). – Circularity in judgment of relative pitch. *Journal of the Acoustical Society of America*, vol. 36, n° 12, Décembre 1964, pp. 2346–2353.
- [Shi95] Shimizu (Masahiro). – Waveform synthesizing apparatus. – US Patent 5,451,710, Septembre 1995.
- [SK87] Strong (Alexander R.) et Karplus (Kevin J.). – Wavetable-modification instrument and method for generating musical sound. – US Patent Number 4,649,783, Mars 1987.
- [Sla94] Slaney (Malcolm). – *An Introduction to Auditory Model Inversion*. – Rapport technique n° IRC1994-014, 1801 Page Mill Road, Building C, Palo Alto, CA 94304, USA, Interval Research Corporation, 1994.

- [SLN95] Slaney (Malcolm), Lyon (Richard F.) et Naar (Daniel). – Sound analysis and resynthesis using correlograms. – US Patent Number 5,473,759, Décembre 1995.
- [SM95] Sandell (Gregory J.) et Martens (William L.). – Perceptual evaluation of principal-component-based synthesis of musical timbres. *Journal of the Audio Engineering Society*, vol. 43, n° 12, Décembre 1995, pp. 1013–1028.
- [Smi85] Smith III (Julius Orion). – Fundamentals of digital filter theory. *Computer Music Journal*, vol. 9, n° 3, 1985, pp. 13–23.
- [Sog97] Sogo (Akira). – Sine-wave generator circuit. – US Patent 5,631,586, Mai 1997.
- [SRD90] Serra (Marie-Helene), Rubine (Dean) et Dannenberg (Roger). – Analysis and synthesis of tones by spectral interpolation. *Journal of the Audio Engineering Society*, vol. 38, n° 3, Mars 1990, pp. 111–128.
- [SS90] Serra (Xavier) et Smith III (Julius Orion). – Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music Journal*, vol. 14, n° 4, 1990, pp. 12–24.
- [SS91] Serra (Xavier) et Smith (Julius). – Musical synthesizer combining deterministic and stochastic waveforms. – US Patent Number 5,029,509, Juillet 1991.
- [SS96a] Stilson (Tim) et Smith (Julius). – Alias-free digital synthesis of classic analog waveforms. In: *Proceedings of the International Computer Music Conference, Hong Kong*, p. 332 335.
- [SS96b] Stilson (Tim) et Smith (Julius). – Analysing the moog vcf with consideration for digital implementation. In: *Proceedings of the International Computer Music Conference, Hong Kong*, p. 398 401.
- [Str87a] Strawn (John). – Analysis and synthesis of musical transitions using the discrete short-time fourier transform. *Journal of the Audio Engineering Society*, vol. 35, n° 1/2, Janvier/Février 1987, pp. 3–13.
- [Str87b] Strawn (John). – Editing time-varying spectra. *Journal of the Audio Engineering Society*, vol. 35, n° 5, Mai 1987, pp. 337–352.
- [TB78] Tucker (Warren H.) et Bates (R.H.T.). – A pitch algorithm for speech and music. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, n° 6, Décembre 1978, pp. 597–604.
- [TG93] Tan (B.T.G.) et Gan (S.L.). – Real-time implementation of asymmetrical frequency-modulation synthesis. *Journal of the Audio Engineering Society*, vol. 41, n° 5, Mai 1993, pp. 357–363.
- [TGLT94] Tan (B.T.G.), Gan (S.L.), Lim (S.M.) et Tang (S.H.). – Real-time implementation of double frequency modulation (dfm) synthesis. *Journal of the Audio Engineering Society*, vol. 42, n° 11, Novembre 1994, pp. 918–926.

- [Thè28] Thèremin (Leo Ssergejewitsch). – Method and apparatus for the generation of sounds. – US Patent Number 1,661,058, Février 1928.
- [THH95] Tellman (Edwin), Haken (Lippold) et Holloway (Bryan). – Timbre morphing of sounds with unequal numbers of features. *Journal of the Audio Engineering Society*, vol. 43, n° 9, Septembre 1995, pp. 678–689.
- [TMM80] Teaney (Dale T.), Moruzzi (Victor L.) et Mintzer (Frederick C.). – The tempered fourier transform. *Journal of the Acoustical Society of America*, vol. 67, n° 6, Juin 1980, pp. 2063–2067.
- [Tom81] Tomisawa (Norio). – Tone production method for an electronic musical instrument. – US Patent 4,249,447, 1981.
- [TR95] Trussell (H. Joel) et Rajala (Sarah A.). – A dsp laboratory designed for teaching fundamental concepts. *ICASSP-95 Conference Proceedings*, vol. 2, Mai 1995, pp. 1117–1120.
- [Tra38] Trautwein (Friedrich). – Electrical musical instrument. – US Patent Number 2,141,231, Décembre 1938.
- [Tru88] Truax (Barry). – Real-time granular synthesis with a digital signal processor. *Computer Music Journal*, vol. 12, n° 2, 1988, pp. 14–26.
- [Ver92] Vercoe (Barry). – Csound. – disponible sur le site ftp: ftp.maths.bath.ac.uk/pub/dream, 1992. Langage dédié à la synthèse de sons musicaux.
- [VLU95] Vrhel (Michael), Lee (Chulhee) et Unser (Michael). – Fast continuous wavelet transform. *ICASSP-95 Conference Proceedings*, vol. 2, Mai 1995, pp. 1165–1168.
- [VMS97] Villasenor (John) et Mangione-Smith (William). – Les ordinateurs configurables. *Pour la science*, no238, Août 1997, pp. 62–68.
- [Wei31] Weiss (E.). – La musique de l'avenir sera-t-elle radioélectrique. *La Science et La Vie*, no163, Janvier 1931, pp. 71–74.
- [Wei95] Weidenaar (Reynold). – *Magic Music from the Telharmonium*. – The Scarecrow Press, 1995.
- [WEM79] Wittman (B.A.), Early (S.H.) et Messerschmitt (D.G.). – A hardware multitone frequency synthesizer. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, n° 6, Décembre 1979, pp. 804–809.
- [WK89a] Wightman (Frederic L.) et Kistler (Doris J.). – Headphone simulation of free-field listening. i: Stimulus synthesis. *Journal of the Acoustical Society of America*, vol. 85, n° 2, Février 1989, pp. 858–867.

- [WK89b] Wightman (Frederic L.) et Kistler (Doris J.). – Headphone simulation of free-field listening. ii: Psychophysical validation. *Journal of the Acoustical Society of America*, vol. 85, n° 2, Février 1989, pp. 868–878.
- [WV97] Wannamaker (Robert A.) et Vrscay (Edward R.). – Fractal wavelet compression of audio signals. *Journal of the Audio Engineering Society*, vol. 45, n° 7/8, Juiller Août 1997, pp. 540–553.
- [YH97] Yuen (Jennifer) et Horner (Andrew). – Hybrid sampling-wavetable synthesis with genetic algorithms. *Journal of the Audio Engineering Society*, vol. 45, n° 5, Mai 1997, pp. 316–330.
- [YWS92] Yang (Xiaowei), Wang (Kuansan) et Shamma (Shihab A.). – Auditory representation of acoustic signals. *IEEE Transactions on Information Theory*, vol. 38, n° 2, Mars 1992, pp. 824–839.
- [Zah81] Zahorian (Stephen A.). – Principal-components analysis for low-redundancy encoding of speech spectra. *Journal of the Acoustical Society of America*, vol. 69, n° 3, Mars 1981, pp. 832–845.
- [ZHL92] Zhong (Jicun), Han (Zhenxiang) et Lu (Weixue). – Design of windows with steerable sidelobe dips. *IEEE Transactions on Signal Processing*, vol. 40, n° 6, Juin 1992, pp. 1452–1459.

# Index

- A**  
 ACP ..... 74  
 ADSR ..... 24  
 analyse  
   des données ..... 72  
   en composantes principales ... 73, 74  
   factorielle ..... 73  
   factorielle des distances ..... 51  
   spectrale  
     par filtrage hétérodyne ..... 56  
     par la méthode de Prony ..... 59  
     par transformée de Fourier ..... 58  
     par transformée en ondelettes . 60,  
       61  
 audion ..... 21  
 audition ..... 46
- C**  
 Cahill (Thaddeus) ..... 20  
 Chowning (John) ..... 27, 38  
 Circuits dédiés ..... 45  
 cochlée ..... 47  
 composante  
   déterministe ..... 59, 67  
   harmonique ..... 71  
   principale complexe ..... 115  
   stochastique ..... 59  
 composantes principales ..... 73  
 Coupleux (Eloy) ..... 22
- D**  
 de Forest (Lee) ..... 21  
 données  
   harmoniques  
     brutes ..... 74  
     réduites ..... 77  
 DSP  
   généralistes ..... 44  
   spécialisés ..... 44
- E**  
 EG ..... 24  
 enchaînement ..... 35  
 erreur résiduelle ..... 90  
 Etherophone ..... 21
- F**  
 feedback ..... 38  
 fenêtre  
   gaussienne ..... 66  
 filtrage  
   formantique ..... 23  
   hétérodyne ..... 56, 114  
 filtres à Q constant ..... 60  
 flûte ..... 89  
 flûte ..... 123
- G**  
 Givelet (Armand) ..... 22  
 Gray (Elisha) ..... 18  
 Grey (John M.) ..... 50
- H**  
 Hammond  
   Laurens ..... 23  
   orgue ..... 23  
 harmonique ..... 64  
 hautbois ..... 88, 122  
 hybrides  
   sons ..... 110
- I**  
 individus ..... 73  
 inharmonicité ..... 114  
 interpolation  
   de la phase ..... 70  
   de transformée de Fourier ..... 64



- L**  
LFO .....24  
lobes secondaires ..... 67
- M**  
Martenot  
    Maurice ..... 21  
    Ondes musicales ..... 21  
Mathews (Max V.) ..... 26  
MIDI ..... 27  
modèle  
    expressif ..... 111  
modules ..... 24  
Moog (Robert Arthur) ..... 23  
multi-échantillonnage ..... 35
- N**  
normalisation ..... 79  
Novachord ..... 23
- O**  
oreille  
    externe ..... 46  
    interne ..... 47  
    moyenne ..... 46  
Orgue radioélectrique ..... 22
- P**  
phase  
    interpolation cubique ..... 70  
    renormalisation ..... 114  
piano ..... 88, 114, 122  
processeurs  
    DSP  
        généralistes ..... 44  
        spécialisés ..... 44  
        généralistes ..... 44  
programme  
    acp ..... 146  
    acpc ..... 147  
    cov ..... 145  
    covc ..... 146  
    fdc ..... 144  
    fftana ..... 143  
    fnd ..... 144  
    hrm ..... 145  
    prj ..... 147  
    prjc ..... 147  
    rec ..... 149  
    recc ..... 149  
    syn ..... 148  
    synd ..... 148  
    trk ..... 144  
    visuana ..... 143
- R**  
reconstruction ..... 104  
    par tables d'ondes complexes ... 115  
RM ..... 25
- S**  
sélection  
    de la composante harmonique .... 71  
    du fondamental ..... 71  
seuillage  
    absolu ..... 68  
    relatif ..... 68  
sous-espace  
    harmonique ..... 76, 90  
    harmonique complexe ..... 115  
spectres  
    harmoniques principaux ..... 76, 95  
stéréoscopie ..... 95  
synthétiseur  
    analogique ..... 25  
    modulaire ..... 23  
    polyphonique ..... 25  
synthèse  
    additive ..... 29  
        par banc d'oscillateurs ..... 30  
        par FFT inverse ..... 33  
    croisée ..... 110  
    de validation ..... 69  
    DSF ..... 42  
    FM ..... 27, 38, 43  
    FOF ..... 38  
    granulaire ..... 40  
    par distorsion ..... 41  
    par formule de sommation discrète 42  
    par groupes ..... 36  
    par modélisation ..... 29  
    par modélisation physique ..... 43

par tables d'ondes .....	28
multiples .....	35, 75
simples .....	34
re- .....	69
soustractive .....	36

**T**

télégraphe harmonique .....	19
tail-padding .....	66
telharmonium .....	20
Termen (Lev), dit Léon Thèremin ...	21
thereminvox .....	21
trajectoire	
harmonique .....	75, 95
transformée de Fourier .....	58
interpolée .....	64
transformée en ondelettes .....	61, 64

**V**

VCA .....	24
VCF .....	24
VCO .....	24
vecteurs propres .....	74
violoncelle .....	89, 125
vocodeur	
à canaux .....	110

**W**

wave-sequencing .....	35
wave-shaping .....	41

**Z**

zero-padding .....	67
--------------------	----