



**HAL**  
open science

# Analyse, représentation et traitement du geste instrumental : application aux instruments à clavier

Christophe Ramstein

## ► To cite this version:

Christophe Ramstein. Analyse, représentation et traitement du geste instrumental : application aux instruments à clavier. Modélisation et simulation. Institut National Polytechnique de Grenoble - INPG, 1991. Français. NNT : . tel-00340367

**HAL Id: tel-00340367**

**<https://theses.hal.science/tel-00340367v1>**

Submitted on 20 Nov 2008

**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.

# **THESE**

présentée par

**Christophe RAMSTEIN**

pour obtenir le grade de **DOCTEUR**  
**de l'Institut National Polytechnique de Grenoble**  
(arrêté ministériel du 23 novembre 1988)

Spécialité : Informatique

## **Analyse, représentation et traitement du geste instrumental**

- Application aux instruments à clavier -

Date de soutenance : 18 décembre 1991

Composition du jury :

P. Jorrand	Président du jury
J.B. Barrière	Rapporteurs
C. Bellissant	
C.Cadoz	Examineurs
B. Guérin	
J. Haury	





**institut national polytechnique  
de Grenoble**

46, avenue Félix-Viallet  
38031 GRENOBLE CEDEX  
FRANCE

## **CERTIFICAT DE SOUTENANCE**

d'une THÈSE DE DOCTORAT  
(Diplôme national institué par l'arrêté ministériel du 23 novembre 1988)

Le Président de l'Institut National Polytechnique de Grenoble certifie que

**Monsieur RAMSTEIN Christophe , Bernard , Marie**

né(e) le 3 janvier 1966 à MULHOUSE ( Haut-Rhin )

inscrit(e) à l'Institut National Polytechnique de Grenoble sous le numéro DED8800171

a soutenu le 18 décembre 1991 à GRENOBLE ( Isère )

conformément aux règlements, devant un jury de l'Institut National Polytechnique de Grenoble,  
une thèse présentée en vue de l'obtention du diplôme de

**DOCTEUR de L'INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE**

spécialité : " Informatique "

Après délibération, le jury décerne ce diplôme à

**Monsieur RAMSTEIN Christophe**

avec la mention " Très Honorable "

Le présent diplôme est délivré à l'intéressé(e) pour lui permettre de jouir de son grade avec les  
droits et prérogatives qui y sont attachés par les lois, décrets et règlements.

Fait à Grenoble, le 8 janvier 1992

**Georges LESPINARD**

Président

de l'Institut National Polytechnique  
de Grenoble







**INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE**  
**46, AVENUE FELIX VIALLET 38031 GRENOBLE CEDEX**  
**TEL : 76.57.45.00**

**PRESIDENT DE L'INSTITUT**  
**MONSIEUR G. LESPINARD**

**ANNEE 1990-1991**

**PROFESSEURS DES UNIVERSITES**

- |            |             |               |
|------------|-------------|---------------|
| ENSERG     | BARIBAUD    | MICHEL        |
| ENSIEG     | BARRAUD     | ALAIN         |
| ENSPG      | BAUDELET    | BERNARD       |
| UFR PGP    | BAUDIN      | GERARD        |
| ENSIEG ILL | BEAUFILS    | JEAN-PIERRE   |
| ENSERG     | BLIMAN      | SAMUEL        |
| ENSHMG     | BOIS        | PHILIPPE      |
| ENSPG      | BONNET      | GUY           |
| ENSEEG     | BONNETAIN   | LUCIEN        |
| ENSIEG     | BRISSONNEAU | PIERRE        |
| CUEFA      | BRUNET      | YVES          |
| ENSHMG     | CAILLERIE   | DENIS         |
| ENSPG      | CAVAIGNAC   | JEAN-FRANCOIS |
| ENSPG      | CHARTIER    | GERMAIN       |
| ENSERG     | CHENEVIER   | PIERRE        |
| UFR PGP    | CHERADAME   | HERVE         |
| ENSIEG     | CHERUY      | ARLETTE       |
| ENSERG     | CHOVET      | ALAIN         |
| ENSHMG     | COGNET      | GERARD        |
| ENSEEG     | COLINET     | CATHERINE     |
| ENSIEG     | CORNUT      | BRUNO         |
| ENSIEG     | COULOMB     | JEAN-LOUIS    |
| ENSIMAG    | CROWLEY     | JAMES         |
| ENSEEG     | DALARD      | FRANCIS       |
| ENSHMG     | DARVE       | FELIX         |
| ENSIMAG    | DELLA DORA  | JEAN          |
| ENSERG     | DEPEY       | MAURICE       |
| ENSPG      | DEPORTES    | JACQUES       |
| ENSEEG     | DEROO       | DANIEL        |
| ENSEEG     | DESRE       | PIERRE        |
| ENSEEG     | DIARD       | JEAN-PAUL     |
| ENSERG     | DOLMAZON    | JEAN-MARC     |
| ENSEEG     | DURAND      | FRANCIS       |
| ENSPG      | DURAND      | JEAN-LOUIS    |
| ENSHMG     | FAUTRELLE   | YVES          |
| ENSIEG     | FOGGIA      | ALBERT        |
| ENSIMAG    | FONLUPT     | JEAN          |
| ENSIEG     | FOULARD     | CLAUDE        |

UFR PGP	GANDINI	ALESSANDRO
ENSPG	GAUBERT	CLAUDE
ENSERG	GENTIL	PIERRE
ENSIEG	GENTIL	SYLVIANE
ENSERG	GUERIN	BERNARD
ENSEEG	GUYOT	PIERRE
ENSIEG	IVANES	MARCEL
ENSERG	JANOT	MARIE-THERESE
ENSIEG	JAUSSAUD	PIERRE
ENSPG	JOST	REMY
ENSPG	JOUBERT	JEAN-CLAUDE
ENSIEG	JOURDAIN	GENVIEVE
UFR PGP	LACHENAL	DOMINIQUE
ENISEG	LACOUME	JEAN-LOUIS
ENSIEG	LADET	PIERRE
ENSIEG	LIENARD	JOEL
ENSHMG	LESIEUR	MARCEL
ENSHMG	LESPINARD	GEORGES
ENSPG	LONGUEQUEUE	JEAN-PIERRE
ENSHMG	LORET	BENJAMIN
ENSEEG	LOUCHET	FRANCOIS
ENSEEG	LUCAZEAU	GUY
ENSIMAG	LUX	AUGUSTIN
ENSIEG	MASSE	PHILIPPE
ENSIEG	MASSELOT	CHRISTIAN
ENSIMAG	MAZARE	GUY
ENSIMAG	MOHR	ROGER
ENSHMG	MOREAU	RENE
ENSIEG	MORET	ROGER
ENSIMAG	MOSSIERE	JACQUES
ENSHMG	OBLED	CHARLES
ENSERG	PANANAKAKIS	GEORGES
ENSEEG	PAULEAU	YVES
ENSIEG	PERRET	ROBERT
ENSHMG	PIAU	JEAN-MICHEL
ENSERG	PIC	ETIENNE
ENSIMAG	PLATEAU	BRIGITTE
ENSERG	POUPOT	CHRISTIAN
ENSEEG	RAMEAU	JEAN-JACQUES
ENSPG	REINISCH	RAYMOND
UFR PGP	RENAUD	MAURICE
ENSIMAG	ROBERT	FRANCOIS
ENSIEG	ROYE	DANIEL
ENSIEG	SABONNADIERE	JEAN-CLAUDE
ENSERG	SAGUET	PIERRE
ENSIMAG	SAUCIER	GABRIELE
ENSPG	SCHLENKER	CLAIRE
ENSPG	SCHLENKER	MICHEL
UFR PGP	SILVY	JACQUES
ENSHMG	SIRIEYS	PIERRE
ENSEEG	SOHM	JEAN-CLAUDE
ENSIMAG	SOLER	JEAN-LOUIS
ENSEEG	SOUQUET	JEAN-LOUIS
ENSHMG	TICKIEWITCH	SERGE
ENSHMG	TROMPETTE	PHILIPPE
ENSIMAG	VERJUS	JEAN-PIERRE
ENSPG	VINCENT	HENRI
ENSERG	ZADWORNY	FRANCOIS

PERSONNES AYANT OBTENU LE DIPLOME

D'HABILITATION A DIRIGER DES RECHERCHES

BALESTRA  
BALME  
BECKER  
BIGEON  
BINDER  
BOE  
BOUVIER  
CHASSERY  
CHOLLET  
COEY  
COMMAULT  
CORNUJOLS  
COURNIL  
DALLERY  
DESCOTES-GENON  
DUGARD  
DURAND  
FERRIEUX  
GAUTHIER  
GHIBAUDO  
HAMAR  
HAMAR  
HORAUD  
KUENY  
LATOMBE  
LE HUY  
LE GORREC  
LOZANO-LEAL  
MAHEY  
MEUNIER  
MICHEL  
MONMUSSON-PICQ  
MULLER  
MULLER  
NGUYEN TRONG  
NIEZ  
PASTUREL  
PERRIER  
PLA  
RECHENMANN  
ROGNON  
ROUGER  
TCHUENT  
TRYSTRAM

FRANCIS  
LOUIS  
MONIQUE  
JEAN  
ZDENECK  
LOUIS-JEAN  
GERARD  
JEAN-MARC  
JEAN-PIERRE  
JEAN-PIERRE  
CHRISTIAN  
GERARD  
MICHEL  
YVES  
BERNARD  
LUC  
MADELEINE  
JEAN-PAUL  
JEAN-PAUL  
GERARD  
SYLVIANE  
ROGER  
PATRICE  
JEAN-LOUIS  
CLAUDINE  
HOANG  
BERNARD  
ROGELIO  
PHILIPPE  
GERARD  
GERARD  
GEORGETTE  
JEAN  
JEAN-MICHEL  
BERNADETTE  
JEAN-JACQUES  
ALAIN  
PASCAL  
FERNAND  
FRANCOIS  
JEAN-PIERRE  
JEAN  
MAURICE  
DENIS

SITUATION PARTICULIERE
------------------------

PROFESSEURS D'UNIVERSITE
--------------------------

**DETACHEMENT**

ENSPG	BLOCH	DANIEL.....	RECTEUR	21.12.1993
ENSIMAG	LATOMBE	J.CLAUDE.....	DETACHEMENT	01.05.1993
ENSHMG	PIERRARD	J.MARIE.....	DETACHEMENT	01.05.1991
ENSIMAG	VEILLON	GERARD.....	DISPONIBLE	01.10.1993

**SURNOMBRE**

ENSHMG	BOUVARD	MAURICE .....		30.09.1990
--------	---------	---------------	--	------------

DIRECTEURS DE RECHERCHE DU CNRS

ALEMANY  
 ALLIBERT  
 ALLIBERT  
 ANSARA  
 ARMAND  
 AUDIER  
 BERNARD  
 BINDER  
 BONNET  
 BORNARD  
 CAILLET  
 CARRE  
 CHATILLON  
 CLERMONT  
 COURTOIS  
 CRISTOLOVEANU  
 DAVID  
 DION  
 DRIOLE  
 DURAND  
 EUSCUDIER  
 EUSTATHOPOULOS  
 FINON  
 FRUCHARD  
 GARNIER  
 GIROD  
 GLANGEAUD  
 GUELIN  
 HOPFINGER  
 JORRAND  
 JOUD  
 KAMARINOS  
 KLEITZ  
 KOFMAN  
 LANDAU  
 LEJEUNE  
 LEPROVOST  
 MADAR  
 MARTIN  
 MERMET  
 MICHEL  
 NAYROLLES  
 PASTUREL  
 PEUZIN  
 PHAM  
 PIAU  
 RENOARD  
 SENATEUR  
 SIFAKIS  
 SIMON

ANTOINE  
 COLETTE  
 MICHEL  
 IBRAHIM  
 MICHEL  
 MARC  
 CLAUDE  
 GILBERT  
 ROLAND  
 GUY  
 MARCEL  
 RENE  
 CHRISTIAN  
 JEAN-ROBERT  
 BERNARD  
 SORIN  
 RENE  
 JEAN-MICHEL  
 JEAN  
 ROBERT  
 PIERRE  
 NICOLAS  
 DOMINIQUE  
 ROBERT  
 MARCEL  
 JACQUES  
 FRANCOIS  
 PIERRE  
 EMIL  
 PHILIPPE  
 JEAN-CHARLES  
 GEORGES  
 MICHEL  
 WALTER  
 IOAN  
 GERARD  
 CHRISTIAN  
 ROLAND  
 JEAN-MARIE  
 JEAN  
 JEAN-MARIE  
 BERNARD  
 ALAIN  
 JEAN-CLAUDE  
 ANTOINE  
 MONIQUE  
 DOMINIQUE  
 JEAN-PIERRE  
 JOSEPH  
 JEAN-PAUL

SUERY  
 TEODOSIU  
 VACHAUD  
 VAUCLIN  
 WACK  
 YAVARI  
 YONNET

MICHEL  
 CHRISTIAN  
 GEORGES  
 MICHEL  
 BERNARD  
 ALI-REZA  
 JEAN-PAUL



à Joris et Katia ...





---

## REMERCIEMENTS

---

Merci à Monsieur Claude Cadoz, directeur de l'ACROE (Association pour la Création et la Recherche sur les Outils d'Expression), pour m'avoir accepté dans son équipe. Plus qu'un directeur de thèse, il m'a communiqué une partie de sa foi et de son attachement à la recherche,

Merci à Monsieur Jean-Baptiste Barrière, responsable de la pédagogie à l'IRCAM (Paris), chercheur en informatique musicale et compositeur, à Monsieur Camille Bellissant, professeur à l'IMAG (Grenoble), ainsi qu' à Monsieur Jean Haury, médecin et chercheur en organologie. Tous trois ont lu ce mémoire avec beaucoup d'attention et de sens critique,

Merci à Monsieur Philippe Jorrand, directeur du LIFIA (Grenoble) qui a accepté de présider ce jury et Monsieur Bernard Guérin, chercheur à l'ICP (Grenoble),

Merci à tous les membres de l'ACROE avec qui je partage la responsabilité de l'aboutissement de ces travaux.

Je tiens plus particulièrement à remercier Olivier Raoult, ingénieur système, pour sa compétence, son dévouement et sa participation conceptuelle et pratique à ce projet. Je remercie également Jean-Loup Florens pour son ouverture d'esprit et sa grande disponibilité. J'exprime ma vive reconnaissance à Annie Luciani pour son enthousiasme et son intervention dans différentes situations délicates.

Merci à Mademoiselle Claire Bécaud, bibliothécaire du Conservatoire de musique de Grenoble, pour m'avoir accueilli sans réserve et m'avoir facilité l'accès à ses ouvrages,

Je remercie enfin mes parents, mes frères, mes amis, ...



<b><u>Introduction générale</u></b> .....	<b>1</b>
I. Cadre général.....	2
II. Objectif de la thèse .....	3
III. Organisation du mémoire .....	3
<b><u>Chapitre I :</u></b>	
<b>De l'édition des sons à celle du geste instrumental ....</b>	<b>5</b>
II. Dialogue homme-ordinateur et création musicale .....	7
III. Le système CORDIS.....	11
IV. Geste et musique.....	21
V. Eléments de typologie du geste instrumental.....	29
VI. Capture du geste instrumental.....	37
VII. Conclusion : l'édition du geste instrumental .....	42
<b><u>Chapitre II :</u></b>	
<b>Segmentation du geste instrumental .....</b>	<b>47</b>
II. Un exemple de référence : la parole .....	49
III. Segments gestuels et clavier .....	52
IV. Segmentation "manuelle".....	62
V. Segmentation automatique.....	64
<b><u>Chapitre III :</u></b>	
<b>Représentation et traitement neutres .....</b>	<b>75</b>
II. Attributs des segments gestuels.....	77
III. Représentation neutre des segments gestuels .....	80
IV. La formule gestuelle .....	95
V. Traitements neutres élémentaires .....	101
VI. Conclusion.....	111

## **Chapitre IV :**

### **Edition en référence à l'action .....113**

II. Approche cohérente du geste instrumental.....	115
III. Un algorithme de validation d'une articulation gestuelle.....	124

## **Chapitre V :**

### **Expériences et validation .....137**

II. Gestes percussifs et signaux gestuels.....	138
III. Exemples de segmentation.....	149
IV. Estimation de l'algorithme de validation .....	166
V. Validation d'une articulation .....	174

### **Conclusion générale ..... 179**

I. Objectif initial.....	180
II. Etat d'avancement pratique .....	181
III. Etat d'avancement théorique .....	190
IV. Perspectives.....	193

### **Bibliographie ..... 195**

---

# **Introduction générale**

---

**I. CADRE GENERAL**

**II. OBJECTIF DE LA THESE**

**III. ORGANISATION DU MEMOIRE**

---

## I. CADRE GENERAL

Les travaux qui font l'objet de cette thèse concernent l'utilisation de l'ordinateur pour la création musicale.

Ce projet s'intègre dans un programme de recherche mené conjointement par l'ACROE (Association pour la Création et la Recherche sur les Outils d'Expression) et le LIFIA (Laboratoire d'Informatique Fondamentale et d'Intelligence Artificielle) et vise la conception et le développement d'un système complet de création artistique, musicale et visuelle, à l'aide de l'ordinateur.

L'ordinateur a fait son apparition dans le domaine musical en 1956 avec les expériences de composition automatique réalisées par Hiller et Isaacson [HI58]. Les nombres engendrés aléatoirement par l'ordinateur et qui satisfont des règles choisies constituent une partition codée, transcrite ensuite en notation traditionnelle avant d'être exécutée par des interprètes.

L'ordinateur a ensuite été utilisé pour la synthèse directe des sons, c'est-à-dire comme moyen de représentation et de traitement de l'univers sonore.

Le programme MUSIC V réalisé par M.Mathews [Mat69], premier système de synthèse directe de sons, propose un ensemble de modules élémentaires (oscillateurs, enveloppes, additionneurs, générateurs de nombres aléatoires, ...) connectables de maintes façons. La définition d'une séquence d'événements sonores repose sur deux phases distinctes : 1/ la détermination d'une "structure-instrument" par assemblage de ces modules, et 2/ celle d'une séquence de "notes" jouées sur cet instrument, soit pour chaque "note", une liste de paramètres où sont précisés les instants d'intervention de "l'instrument" et des paramètres tels que la fréquence, l'amplitude du vibrato ... qui "activent l'instrument".

Plus récemment, C.Cadoz, A. Luciani et J.L Florens [CLF81] ont introduit une démarche dans laquelle l'ordinateur est considéré comme moyen de représentation et de traitement de l'univers, non pas sonore, mais instrumental. Le principe fondamental à la base de cette démarche est d'aborder la création musicale à partir de l'expérience instrumentale.

L'instrument, support de cette expérience, est simulé à l'aide de l'ordinateur. Les phénomènes produits par cet instrument simulé sont restitués par des interfaces matérielles adaptées au canal acoustique et gestuel. L'instrument, bien que totalement virtuel, apparaît alors à l'utilisateur du système comme réel.

Dans cette approche, le geste intervient de façon essentielle dans le déroulement du son et le signal acoustique est entièrement déterminé 1/ par l'instrument et 2/ par les gestes qui s'y appliquent.

## **II. OBJECTIF DE LA THESE**

Dans le champ des techniques informatiques, les possibilités de mémorisation objective introduisent des latitudes radicalement nouvelles. Avec le système CORDIS [CLF90] - système général d'analyse, de modélisation et de simulation de l'univers instrumental -, le geste est capté et mémorisé sous la forme d'un ensemble de signaux échantillonnés. Le geste, qui jusqu'alors n'avait de réalité que dans son temps propre, est devenu objet, susceptible d'être représenté et traité.

Notre objectif est alors précisément de concevoir et de réaliser un "éditeur de geste", pour analyser, représenter et traiter l'objet gestuel mémorisé. Nous visons par cet ensemble matériel et logiciel un outil complet pour la création musicale qui abandonne (provisoirement) l'ordinateur en tant que moyen de représentation de l'univers sonore mais l'aborde comme moyen de représentation et de traitement de ses causes (gestuelles et instrumentales).

## **III. ORGANISATION DU MEMOIRE**

Ce rapport comporte cinq chapitres.

Le premier chapitre présente le système général d'analyse, de modélisation et de simulation CORDIS, et introduit le problème de l'édition du geste instrumental. Une étude préliminaire sur le geste et sur la morphologie des transducteurs nous a conduits à proposer une typologie du geste et à dégager des éléments de structuration.

Dans le deuxième chapitre nous abordons le problème de la décomposition du geste instrumental. Ce dernier, capté et mémorisé sous la forme de signaux échantillonnés, assimilé à une séquence d'événements gestuels, est décomposé par une opération de segmentation.

Les troisième et quatrième chapitres traitent deux aspects majeurs de l'édition du geste instrumental. Il s'agit plus précisément d'élaborer des représentations du geste instrumental, pertinentes pour le compositeur et bien adaptées à des contraintes



informatiques, et d'autre part des traitements qui permettent d'organiser le geste instrumental selon différentes logiques et à chaque niveau de ses représentations.

Le dernier chapitre propose des exemples et des expériences pour tester et valider les fonctions proposées dans ce mémoire.

## De l'édition des sons à celle du geste instrumental

---

### I. INTRODUCTION

### II. DIALOGUE HOMME-ORDINATEUR ET CREATION MUSICALE

- II.1. les débuts avec Music V
- II.2. la tendance actuelle avec Esquisse

### III. LE SYSTEME CORDIS

#### III.1. mécanismes instrumentaux

- 1. analyse
- 2. modélisation et modularité
  - a. modules physiques - b. modules fonctionnels -
  - c. exemple
- 3. algorithmes
- 4. langage

#### III.2. les transducteurs gestuels

- 1. les dispositifs non rétroactifs
- 2. les dispositifs à retour d'effort
  - a. les débuts - b. actuellement

#### III.3. les modes de fonctionnement

- 1. mode compositionnel
- 2. mode instrumental
  - a. enregistrement du geste - b. restitution du geste

### IV. GESTE ET MUSIQUE

#### IV.1. gestes et création musicale

- 1. l'instrumentiste - 2. le compositeur
- 3. le chef d'orchestre - 4. conclusion

#### IV.2. geste instrumental et claviers

- 1. le clavecin
- 2. le piano

3. clavier de contrôle

- a. l'orgue classique - b. le synthétiseur à clavier -
- c. CORDIS et clavier rétroactif modulaire

5. conclusion

**V. ELEMENTS DE TYPOLOGIE DU GESTE INSTRUMENTAL**

V.1. analyse descriptive

- 1. critère cinématique - 2. critère spatial - 3. critère fréquentiel

V.2. analyse fonctionnelle

- 1. critère effecteur
- 2. critère de saisie
  - a. contact temporaire/permanent
  - b. avec saisie - c. sans saisie

V.3. analyse intrinsèque

V.4. conclusion

**VI. CAPTURE DU GESTE INSTRUMENTAL**

VI.1. quelques définitions

VI.2. représentation discrète du geste instrumental

VI.3. capture du geste instrumental

- 1. sans réduction
- 2. sur l'instrument
  - a. réduction 1 - b. réduction 2
- 3. fréquence d'échantillonnage
- 4. exemples de capture avec clavier
  - a. le système CORDIS - b. le système 290 SE -
  - c. le système Moog MTS - d. récapitulatif

**VII. CONCLUSION**

VII.1. édition du geste instrumental

VII.2. notre façon de traiter ce problème

- 1. les signaux gestuels - 2. structuration des variables gestuelles
-

## **I. INTRODUCTION**

Dans ce premier chapitre, nous souhaitons mettre en évidence que le geste instrumental a un rôle déterminant dans le processus de création musicale, et que grâce aux technologies actuelles, ce geste peut devenir objet, constituant une base particulièrement féconde pour aborder la composition musicale à l'aide de l'ordinateur.

## **II. DIALOGUE HOMME-ORDINATEUR ET CREATION MUSICALE**

Depuis l'apparition de l'ordinateur dans le domaine de la création musicale, de nombreux logiciels ont été réalisés. Ils permettent de représenter et de traiter les objets musicaux à différents niveaux et selon différentes démarches.

Cependant le dialogue entre le musicien-compositeur et l'ordinateur s'avère souvent difficile. Et cela d'autant plus que ce dialogue est supporté par des langages techniques et peu pertinents du point de vue musical.

Pour mieux situer ce problème, on évoquera successivement deux systèmes : Music V (op.cit.), premier système de synthèse acoustique, et Esquisse, l'une des dernières réalisations significatives en la matière de logiciels d'aide à la composition.

### **II.1. les débuts avec Music V**

Music V est un programme général de synthèse acoustique.

Le langage que propose Music V partage la détermination d'un événement sonore en deux phases :

- 1 - la définition d'une structure dite "instrument" qui décrit l'algorithme de synthèse par combinaison de modules élémentaires (oscillateurs, générateur d'enveloppe, additionneurs, ...). L'oscillateur, par exemple, reproduit périodiquement une fonction prédéfinie (Fn). L'amplitude et la fréquence du signal produit sont déterminées par deux paramètres d'entrées,

- 2 - la définition de "notes" qui précisent les paramètres des "instruments".

L'ensemble des événements sonores constitue une "partition" qui conduit au calcul en temps différé des échantillons sonores.

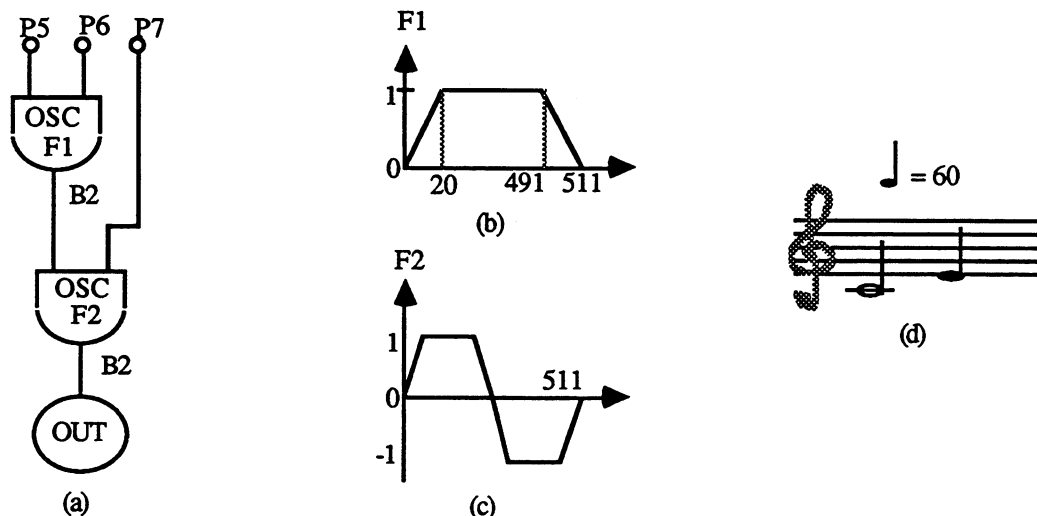


Figure 1. Exemple de notation Music V. La figure (a) décrit le graphe de l'"instrument" et la figure (d) la notation conventionnelle des deux "notes" souhaitées. (Figure extraite de [Mat69]).

La "partition" Music V est en fait un codage intermédiaire entre le programme qui calcule les échantillons sonores et les représentations conventionnelles de l'"instrument" (sous la forme d'un diagramme) et des "notes" (en partition traditionnelle).

Ce codage consiste initialement en l'énoncé d'une succession d'instructions. Pour l'"instrument", les instructions permettent de spécifier des modules élémentaires (oscillateurs, enveloppes, ...), les connexions avec d'autres modules et des valeurs initiales de certains paramètres. Pour chaque "note", les instructions précisent une durée, une fréquence, une intensité, ...

Partition "instrument"

- 1 INS 0 1;
- 2 OSC P5 P6 B2 F1 P30;
- 3 OSC B2 P7 B2 F2 P29;
- 4 OUT B2 B1;
- 5 END;

Partition "notes"

- 6 GEN 0 1 1 0 0 .99 20 .99 491 0 511;
- 7 GEN 0 1 2 0 0 .99 50 .99 205 -.99 306 -.99 461 0 511;
- 8 NOT 0 1 2 1000 .0128 6.7;
- 9 NOT 2 1 1 1000 .0256 8.44;
- 10 TER 3;

### description de la partition :

La ligne 1 indique le numéro de l'"instrument" ; la ligne 5 la fin de la définition de l'"instrument" ; les lignes 2 et 3 définissent deux oscillateurs distincts de fonction F1 et F2. La forme de ces fonctions est déterminée par les lignes 6 et 7. Enfin, les lignes 8 et 9 spécifient les deux notes à jouer avec dans l'ordre d'apparition des paramètres : la date d'intervention, le numéro de l'instrument, la durée, l'entrée d'amplitude du premier oscillateur (P5), la fréquence (P6) et l'entrée de fréquence du second oscillateur (P7). La ligne 10 précise que la séquence se termine au bout de 3 secondes.

Les problèmes posés par ce type de représentation ne sont pas élémentaires. Outre les aspects purement informatiques, la difficulté essentielle réside dans la forme des instructions pour un utilisateur du système non informaticien. L'usage d'un code alpha-numérique en ligne pour décrire les "notes" est l'un des aspects qui rend la "partition" fastidieuse à la lecture pour un musicien.

### II.2. la tendance actuelle avec Esquisse

"Aujourd'hui, le champ privilégié d'application des recherches musicales avec les systèmes personnels est de fait l'aide à la composition plutôt que la synthèse : le développement de la première est en effet favorisé par la mise à la disposition sur les mini-ordinateurs de certains langages et concepts de l'intelligence artificielle, qui facilitent par rapport aux langages de programmation traditionnels, le type de manipulations symboliques propres à la composition ; la synthèse demeure en revanche mal représentée sur ces systèmes, car elle exige des ressources de traitement numérique du signal trop spécifiques et importantes pour qu'une machine générale puisse être efficace. " [Bar90]

Le système d'aide à la composition Esquisse, développé à l'IRCAM (Institut de recherche et Coordination Acoustique / Musique à Paris) permet au compositeur de décrire ses processus ou opérations compositionnels à l'aide d'un langage graphique adapté.

Il devient alors possible de penser en continuité l'utilisation de l'ordinateur

- 1 - pour définir des structures à partir d'un système de composition,
- 2 - pour éditer le matériau produit.

Le travail du compositeur se réalise ainsi dans une boucle permettant de créer le matériau sonore selon certaines règles tout en conservant la possibilité de le modifier "à la main" et de revenir de la sorte à une situation plus traditionnelle.

Dans Esquisse, le compositeur manipule des objets musicaux de complexité variable. Il peut créer ou modifier leur structure par des fonctions, représentées sous la forme de modules graphiques, et les chaîner selon l'idée musicale visée. Les principales opérations réalisables sur les objets musicaux se répartissent en quatre groupes :

- création de nouveaux objets (par des données extérieures, à partir d'un éditeur de notation traditionnelle, par transformation d'objets existants),
- archivage,
- visualisation des représentations, ou écoute sur un synthétiseur numérique périphérique,
- impression.

" La fonction du système est donc de créer un matériau musical à un stade «pré-compositionnel», de le représenter, l'écouter et le modifier «manuellement» ou par fonction, et enfin d'en conserver une forme imprimée, ou bien de transférer les résultats sous forme de fichiers vers d'autres machines. " [Bar90]

### **III. LE SYSTEME CORDIS**

On peut qualifier de "synthèse acoustique" les démarches qui découlent du procédé de synthèse directe des sons. Ces démarches ont comme objet central le signal sonore qu'elles décrivent en termes d'acoustique. Le système Music V illustre cette approche.

C.Cadoz propose, en amont de cette démarche, de considérer l'ordinateur de façon plus générale, comme moyen de représentation de l'univers physique. Il définit le cadre conceptuel d'un système pour la synthèse sonore par simulation de mécanismes instrumentaux en temps réel.

Dans ce système, baptisé CORDIS [Cad79], il ne s'agit pas seulement de synthétiser les phénomènes produits par l'objet instrumental, mais de synthétiser cet objet lui-même.

#### **III.1. mécanismes instrumentaux**

Pour construire des instruments à l'aide de l'ordinateur, il faut considérer quatre types de problèmes :

- l'analyse des mécanismes instrumentaux
- la modélisation des mécanismes instrumentaux
- la traduction de cette modélisation en algorithmes de calcul
- la constitution d'un langage de description de haut niveau qui permet de définir un instrument selon ses caractéristiques mécaniques.

##### **1. analyse**

Un instrument est considéré comme décomposable en trois éléments :

- " la structure vibrante dont les comportements mécaniques sont essentiellement acoustiques
- la structure excitatrice dont les comportements mécaniques sont lents et dont la fonction est de transformer le geste instrumental en excitation de la structure vibrante
- la jonction entre l'excitateur et la structure vibrante, qui détermine la manière dont l'énergie mécanique est transmise de l'excitateur à la structure vibrante" [Cad83].

L'analyse des instruments consiste alors en l'étude de l'excitateur, de la structure vibrante et de leur relation. Elle conduit à classer et inventorier les instruments selon leurs constituants et les différentes jonctions possibles entre eux.



Parmi les structures vibrantes, on rencontre les cordes, les colonnes d'air, les lames, les membranes et les volumes solides ou fluides. Les excitateurs sont par exemple la mailloche d'une percussion, le mécanisme d'excitation du piano, la main du musicien qui percute une membrane ou pince une corde.

La relation entre l'excitateur et la structure vibrante peut être caractérisée par la rapidité de l'échange d'énergie effectué : instantanée ou étalée. Le premier correspond à la percussion ou au pincement et le second au contrôle lent tel que le frottement d'archet.

## **2. modélisation et modularité**

CORDIS est un système de modélisation et de simulation. Il permet d'élaborer le modèle d'un objet physique et de traduire cette connaissance en des programmes qui le simulent.

Le principe de modélisation de l'univers instrumental repose sur une analyse des objets physiques en constituants élémentaires. Tout objet physique est vu comme un assemblage d'éléments. ( Ces éléments, tout en étant ceux de la modélisation, sont aussi ceux de la simulation).

CORDIS propose deux types de modules de simulation : les modules physiques et les modules fonctionnels.

### **a. modules physiques**

" La plus petite décomposition donnant lieu à un système générique est celle où l'on a deux types d'atomes distincts : l'un correspondant à un point matériel, l'autre à un élément de liaison. " [Cad91]. Le point matériel (traduit par un module masse) et l'élément de liaison (traduit par un module ressort-frottement) sont les constituants de base de l'univers instrumental de CORDIS.

Cependant avec ces deux éléments, les possibilités de modélisation sont encore limitées. L'agencement des modules masse et ressort constitue des structures monolithiques, non dissociables au cours du temps. Pour rompre, ou créer temporairement des structures, CORDIS propose la fonction de liaison conditionnelle : "c'est une liaison ressort-frottement dont l'état peut changer en fonction des conditions sur ses variables d'entrée." [Cad91]

Cette liaison est essentielle pour décrire la plupart des jonctions excitateur / structure vibrante.

La communication entre les modules met en cause deux variables duales, l'une position, l'autre force. Un module matériel reçoit une force en entrée et fournit une

position en sortie. Un module liaison reçoit deux positions en entrée et fournit deux forces en sortie.

Deux autres types de modules prennent en compte des considérations pratiques :

- les modules dégénérés sont incomplets du point de vue des communications. Ils permettent par exemple de décrire des données constantes telles qu'une position fixe du sol ou la pesanteur.
- les modules intégrés sont des macro-modules à base de modules élémentaires destinés à diminuer la complexité matérielle et conceptuelle. L'oscillateur élémentaire, constitué d'une masse couplée au sol par un ressort, en est un exemple.

### b. modules fonctionnels

" Un module fonctionnel ne manipule pas exclusivement des variables physiques et ne répond pas exclusivement à un modèle physique. "[Cad91]

Les modules de contrôle structurel possèdent un ou plusieurs points de connexion mais aussi des sorties logiques qui dépendent des valeurs des variables d'entrée. Associés à des connexions, ils en contrôlent l'établissement ou la rupture.

Les modules de contrôle paramétrique agissent sur les paramètres d'un algorithme. Les sorties sont donc non-homogènes aux points de communication des modules physiques mais correspondent aux formats des paramètres contrôlés.

### c. exemple

La figure suivante donne un exemple simple de structure physique décrite dans le formalisme CORDIS.

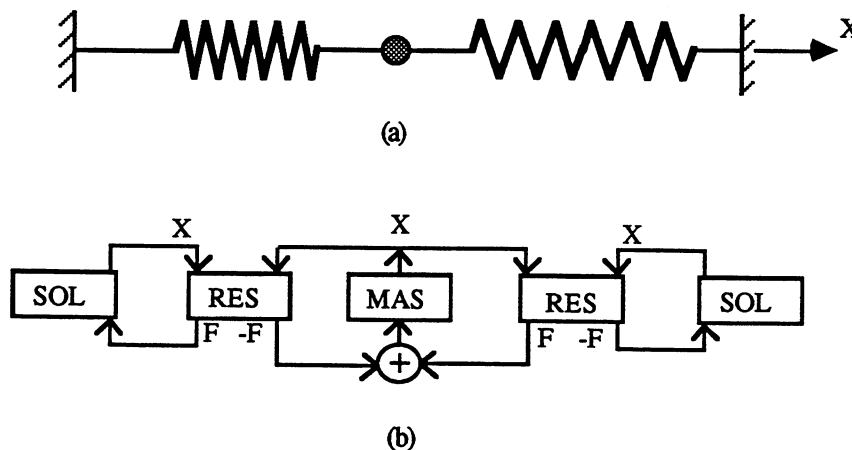


Figure 2. Exemple de structure physique. Une masse attachée au sol par deux liaisons élastiques.

La figure 2 (a) représente une masse couplée au sol par deux liaisons ressort. La figure 2 (b) décrit cette structure dans le langage modulaire de CORDIS. A chaque élément physique correspond un module : module SOL, module RES(sort) et module MAS(se). Les flèches traduisent les communications entre les modules. A noter que les sorties en force des modules RES sont sommées à l'entrée du module MAS.

**3. algorithmes**

Les algorithmes élémentaires du système CORDIS sont la traduction en des processus numériques de chacun des modules définis. Ils doivent être fidèles aux modèles et optimaux en temps de calcul.

L'algorithme d'une structure n'en est pas la traduction numérique monolithique, mais est constitué de l'ensemble, agencé séquentiellement (ou parallèlement d'ailleurs), des algorithmes des modules la constituant.

Les algorithmes choisis sont décrits sous la forme d'équations récurrentes.

Par exemple, l'algorithme de la masse est calculable de la manière suivante :

On pose

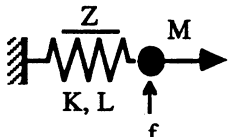
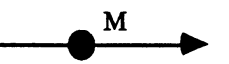
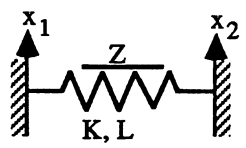
$$x'(n) = x(n) - x(n-1)$$

$$x''(n) = x(n) - 2x(n-1) + x(n-2)$$

et d'après le principe fondamental de la dynamique  $f = M.g$

$$x(n) = f(n)/M + 2x(n-1) - x(n-2)$$

On retrouve dans le tableau suivant les algorithmes de trois modules : la cellule, la masse et la liaison.

Identificateur	Diagramme	Algorithme
Cellule		$x(n) = \frac{f(n)}{M} + (2 - \frac{K-Z}{M}).x(n-1) + (\frac{Z}{M} - 1).x(n-2) + \frac{K.L}{M}$
Masse		$x(n) = \frac{f(n)}{M} + 2.x(n-1) - x(n-2)$
Identificateur	Diagramme	Algorithme
Ressort		$f(n) = (K+Z).[x_1(n) - x_2(n)] - Z.[x_1(n-1) - x_2(n-1)] + K.L$

#### **4. langage**

" En tant qu'ensemble d'éléments et règles de combinaisons, CORDIS est un langage. Il permet de décrire des objets du monde physique, mais également de les reconstruire dans l'espace virtuel de la simulation. " [Cad91]

Sachant qu'un module matériel reçoit une seule force en entrée et fournit une seule position en sortie, qu'un module de liaison reçoit deux positions et fournit deux forces, les règles d'assemblage des modules physiques vérifient un principe simple :

- un module matériel se connecte à un (ou plusieurs) modules (s) de liaison, mais pas à un autre module matériel. La force entrant dans le module matériel est la somme des forces sortant des modules de liaison qui y sont connectés. L'unique position sortant du module matériel est distribuée à chaque module de liaison connecté.

- un module de liaison se connecte à deux modules matériels, mais pas à d'autres modules de liaison. Les deux positions entrant dans le module de liaison correspondent aux positions des deux modules matériels connectés. Les deux forces sortant du module de liaison, de même norme, mais de direction opposée, sont fournies aux deux modules matériels.

Il est donc possible de réaliser tout graphe connexe dont les nœuds sont des modules matériels et les arcs des modules de liaison.

#### **III.2. les transducteurs gestuels**

Pour restituer les phénomènes sensibles de l'instrument simulé, il faut des interfaces adaptées aux canaux sensoriels du musicien. Mais si la technologie des transducteurs acoustiques (le haut-parleur) et des transducteurs visuels (l'écran cathodique) peut être considérée à un stade très avancé, il n'en est pas de même de celle des transducteurs qui doivent traiter le canal gestuel.

Le transducteur gestuel est un dispositif matériel qui réalise la correspondance entre l'espace sensori-moteur du geste et l'espace numérique de la simulation. Cependant, J.L.Florens fait remarquer que "le transducteur gestuel universel, pour des raisons technologiques, est aujourd'hui irréalisable. Aussi, le transducteur gestuel, sous sa forme générale se présentera nécessairement comme un ensemble de dispositifs complémentaires dans leurs fonctions." [Flo84]

Une étude préliminaire sur la typologie des gestes [CLF81] fait apparaître deux classes importantes de gestes, dont l'une met en évidence les transducteurs gestuels à retour

d'effort et l'autre les transducteurs gestuels sans retour d'effort. La rétroaction permet à l'opérateur de "sentir" par le toucher l'objet physique qu'il manipule.

### **1. les dispositifs non rétroactifs**

Dans certains cas, le retour d'effort est considéré comme non nécessaire. La fonction d'une main gauche de violoniste est par exemple de sélectionner une longueur de corde particulière. Ce type de geste implique généralement une dépense d'énergie faible, et par conséquent une faible importance du retour tactile.

Deux types de transducteurs gestuels sans retour d'effort sont à distinguer : ceux qui captent un déplacement (à effort nul) tel que les "joysticks" et ceux qui captent une force (à déplacement nul) tels que les systèmes à jauge de contrainte.

### **2. les dispositifs à retour d'effort**

Depuis sa création, plusieurs prototypes de transducteurs rétroactifs ont été réalisés dans le laboratoire.

#### **a. les débuts**

Le premier prototype de transducteur gestuel rétroactif a été construit par JL Florens [Flo78]. Il a les fonctions d'un archet et est dédié aux gestes d'ambitus important (50cm), en trajectoire rectiligne horizontale et avec saisie. Un second prototype a vu le jour en 1980. Il correspond à une touche unique dont les dimensions ont été inspirées de la touche de piano. La touche est dédiée aux gestes à caractère digital (5cm) en trajectoire rectiligne verticale. Un capteur de position permet de mesurer le déplacement en amplitude de la touche. Elle dispose également d'un capteur de force. Le retour d'effort est assuré par un moteur dont la force, ramenée à l'extrémité du bras d'une touche, est de l'ordre de  $\pm 1\text{KgForce}$ .

#### **b. actuellement**

Le dernier prototype en date est le clavier rétroactif modulaire [CLF90]. Il est prévu pour des gestes à caractère digital (5cm par touche) en trajectoire rectiligne verticale et de sélection latérale. La forme de base est celle d'un clavier dont les dimensions ont été inspirées de celles du piano, mais au contraire du piano, les touches qu'il comporte sont de forme identique (il n'y a pas de touches noires et de touches blanches).

Le clavier est modulaire. Cela signifie que le nombre de ses touches n'est pas définitivement fixé, mais peut être augmenté ou diminué.

Cette propriété est directement liée à sa constitution. Chaque touche est en fait dépendante d'un module capteur / moteur. Pour étendre le clavier, il suffit d'ajouter un module, ou un groupe de modules.

Une caractéristique intéressante des moteurs est qu'ils sont en "tranche". Une "tranche" est constituée d'un bloc aimant / bobine et se sert de l'aimant de la "tranche" voisine pour "fermer" le moteur.

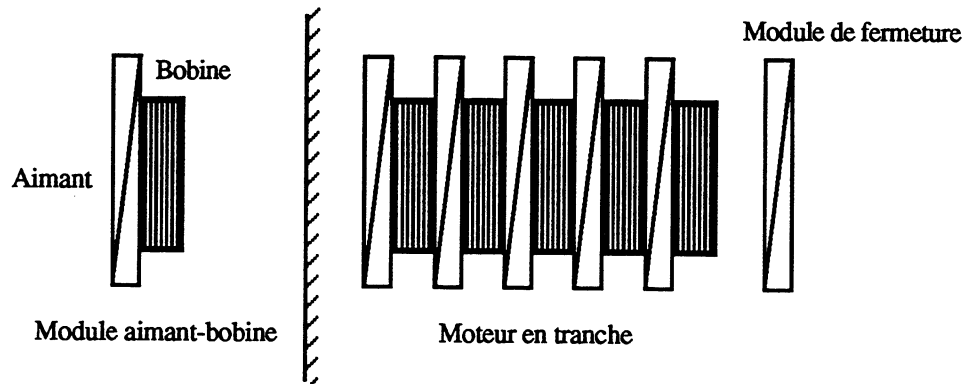


Figure 3. - *Principe du moteur en tranche.* Comporte plusieurs "tranches" aimant-bobine et un module de fermeture.

Le clavier peut être "habillé". Cela signifie que l'on peut greffer au clavier de base une mécanique supplémentaire. Pour cela, l'extrémité de chaque bras de touche est creux et permet la fixation d'éléments mécaniques extérieurs.

Deux "habillages" existent actuellement : le clavier de type piano et un stick qui permet de balayer une surface 2D.

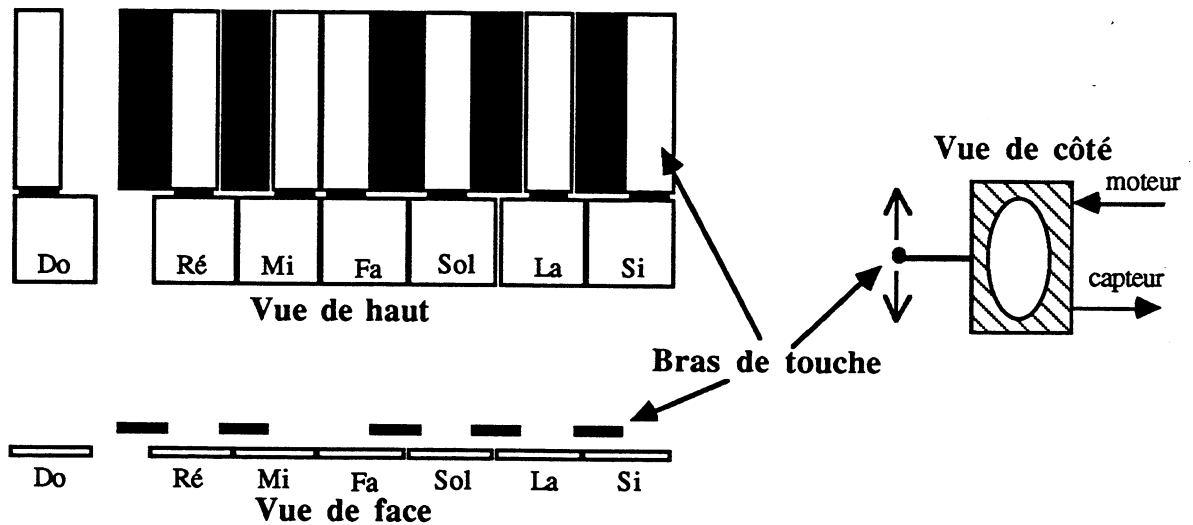


Figure 4. - *"Habillage" du clavier modulaire.* Le clavier modulaire peut être habillé comme un clavier de piano avec l'alternance et le décalage en amplitude de touches noires-blanches d'un clavier de piano.

### **III.3. les modes de fonctionnement**

Dans l'activité de création coexistent en permanence deux modes particuliers : le "mode compositionnel" et le "mode instrumental".

" Le mode compositionnel correspond avant tout à une activité dégagée des contingences matérielles, et en particulier temporelles propres à l'objet final, c'est à dire au phénomène sonore. Il se développe en temps différé.

[...] Le mode instrumental est son dual dans la mesure où, par essence, il se développe dans le temps du phénomène final, en temps réel, et donc en prise directe (gestuelle entre autre) avec les processus producteurs. " [Cad88]

#### **1. mode compositionnel**

Dans CORDIS, le compositeur a une activité de luthier. Il définit les mécanismes instrumentaux, c'est à dire la structure physique de l'excitateur, de la structure vibrante et de leur jonction.

Actuellement, un éditeur rudimentaire permet de construire de telles structures. Il offre au compositeur une bibliothèque de modules prédéfinis d'excitateurs, de structures vibrantes et de jonctions d'une part, et d'autre part, des modules de transducteurs gestuels et acoustiques.

Une interface graphique fait correspondre à chacun de ces modules un objet graphique, que l'utilisateur peut connecter "manuellement" à d'autres modules-graphiques pour former ainsi un réseau CORDIS. Le système contrôle automatiquement que les connexions souhaitées soient autorisées par le langage CORDIS.

Les principales opérations réalisables dans cet éditeur se répartissent en trois catégories :

- création et modification d'un réseau CORDIS,
- sauvegarde / restitution des réseaux,
- visualisation détaillée du réseau.

#### **2. mode instrumental**

Lorsque l'instrument a été construit, le compositeur laisse la place à l'interprète.

Si la complexité des algorithmes de simulation le permet, l'interprète peut contrôler et percevoir l'instrument, en temps réel, par le biais des différents transducteurs. L'instrument simulé est alors totalement réel.

### a. enregistrement du geste

Lors d'un jeu instrumental en temps réel, le système mémorise des signaux échantillonnés représentant le comportement mécanique de l'instrument. Ils décrivent, dans le cas du clavier rétroactif modulaire actuel, l'évolution en position de chacune de ses touches.

Nous verrons dans le paragraphe VI "Capture du geste instrumental", que ces signaux sont suffisants pour représenter le geste instrumental.

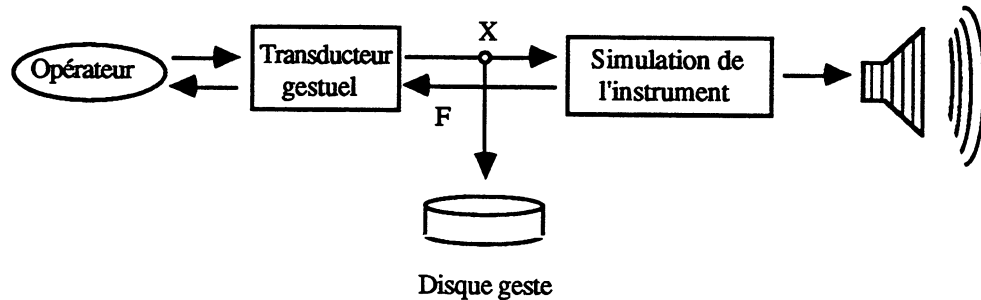


Figure 5. - *Jeu en temps réel*. Lors du mode de jeu temps réel, le geste instrumental est capté et mémorisé sous la forme du signal de commande de l'instrument. Remarque : on ne mémorise pas le son.

### b. restitution du geste

Le système autorise également la lecture (ou restitution) des signaux mémorisés. Cette lecture se réalise selon deux cas de figure :

- l'instrument est simulable en temps réel,
- l'instrument est simulé en temps différé.

Dans le premier cas, le système simule les mécanismes instrumentaux en temps réel mais les signaux d'entrée de la simulation, au lieu de provenir des transducteurs gestuels, sont simplement relus en mémoire. Pendant cette lecture, l'instrumentiste ne peut donc pas contrôler l'instrument. Il regarde et écoute l'instrument "jouer" de lui-même, comme le ferait un piano mécanique.

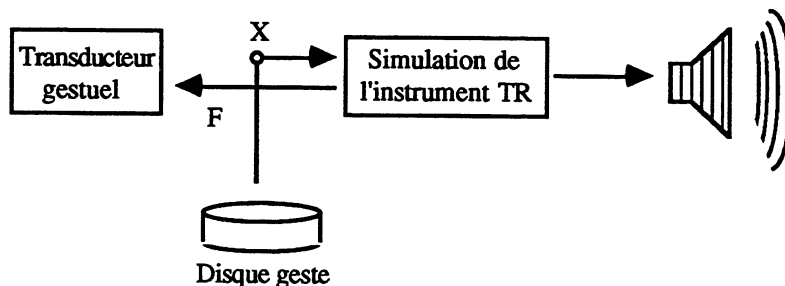


Figure 6. - *Restitution du GI. Instrument temps réel*. Si la simulation de l'instrument peut se réaliser en temps réel, la restitution du geste instrumental se fait en temps réel.



Si l'instrument n'est pas simulable en temps réel, deux étapes sont nécessaires. En premier lieu, les échantillons sonores sont calculés et mémorisés à partir des algorithmes de simulation recevant en entrée les signaux mémorisés.

Ils sont ensuite restitués en temps réel sur les hauts-parleurs. On peut concevoir dans cette deuxième étape de lire simultanément aux échantillons sonores, les échantillons gestuels. Comme dans le cas précédent, l'instrumentiste regarde et écoute l'instrument "jouer" de lui-même, comme le ferait un piano mécanique.

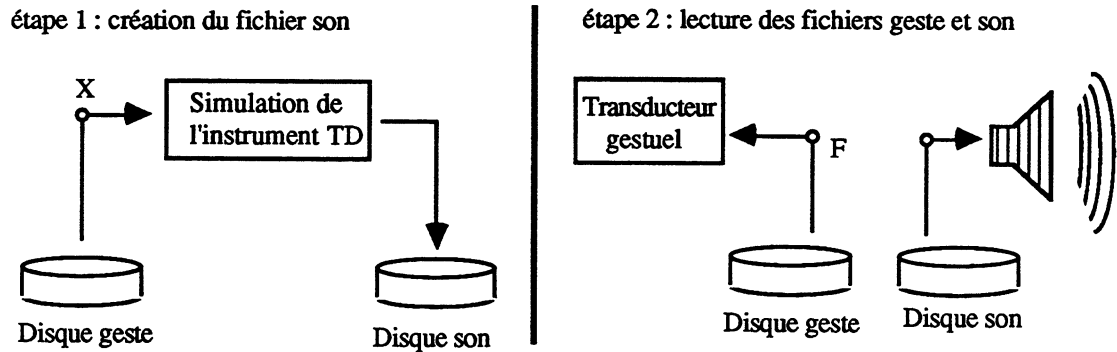


Figure 7. - *Restitution du GI. Instrument temps différé.* Si la simulation de l'instrument se réalise uniquement en temps différé, la restitution du geste instrumental doit se faire en deux étapes : 1/ fabrication du fichier son et 2/ lecture des fichiers sons et gestes sur les transducteurs correspondants.

## **IV. GESTE ET MUSIQUE**

L'être humain a des facultés motrices et est capable d'un certain nombre de mouvements, que l'on peut caractériser par des trajectoires, des vitesses, des forces, etc..

Cependant, au delà de ces aspects purement phénoménologiques, les actions motrices réalisées servent différentes fonctions. Dans la musique plus particulièrement, elles prennent un sens précis et déterminant.

### **IV.1. gestes et création musicale**

Que ce soit le geste du luthier qui façonne de ses mains un violon, celui du violoniste qui en fait vibrer les cordes, du chef d'orchestre qui le dirige, ou encore celui du compositeur, il s'agit dans chaque cas, à des niveaux certes différents, de gestes pour la musique.

#### **1. l'instrumentiste**

Les gestes de l'instrumentiste s'intègrent dans un ensemble de processus qui lui permettent d'animer et de faire vibrer le corps sonore. Il travaille et perfectionne une connaissance générale de l'instrument ; par un processus continuellement rétroactif, il sollicite par des gestes l'instrument qui lui répond par diverses manifestations acoustiques, gestuelles, visuelles, ...

Ces gestes ne le limitent pas aux seules parties susceptibles d'être en contact avec l'instrument. C'est d'ailleurs le plus souvent le corps entier qui est impliqué. Les gestes du pianiste par exemple, font intervenir ses mains, mais aussi ses coudes et ses épaules dans un mouvement d'accompagnement.

Par ailleurs, la connaissance de l'instrumentiste participe à deux types de jeu : l'interprétation et l'improvisation.

L'interprète est avant tout un traducteur. Il donne une réalité acoustique à une représentation non sonore. Mais plus exactement, il transforme les symboles de la partition ou ceux du chef d'orchestre en un geste instrumental. Celui-ci doit donner lieu à des sons dont les attributs (durée, hauteur, ...) sont décrits par ces symboles. Le geste instrumental, configuré par l'instrument et paramétré par les signes musicaux, contient en soi déjà toute l'expression du message sonore.

En revanche, lors de l'improvisation, l'instrumentiste ne soumet pas son jeu à des consignes extérieures. Il compose sur le champ, de lui-même, les séquences de gestes

susceptibles d'avoir une signification musicale. Cette aptitude naît d'un processus assez complexe, issu de son expérience et de son apprentissage de l'instrument. Les schémas moteurs qu'il met en place sont en partie prédéfinis par cet apprentissage, et leur variété dépend aussi bien du thème du morceau que de l'inspiration du musicien pour les composer.

## **2. le compositeur**

La pensée du compositeur élabore des formes et des organisations musicales. Elle s'opère sur des représentations atemporelles des phénomènes en cause et différents systèmes de notation permettent de la matérialiser. Le geste du compositeur est donc essentiellement un geste d'écriture : sur papier, à l'aide du clavier ou de la souris d'un ordinateur, ...

En outre, le geste intervient à un autre niveau de la composition, plus subtil et profond que le premier. Il est une base à partir de laquelle se constitue les représentations mentales en cause dans la pensée compositionnelle.

On pourrait penser au premier abord que l'acte de composition s'appuie sur des représentations mentales exclusivement "sonores". Mais ce serait omettre que le compositeur a aussi une connaissance instrumentale, acquise au fil du temps, qui est devenue indissociable de son pouvoir d'abstraction et de conceptualisation.

Cette connaissance n'est pas uniquement une connaissance d'instrumentiste. Bien plus large, elle naît d'une expérience du monde physique, aussi bien du contact avec des objets en apparence anodins, que du contact effectif avec les instruments de musique.

Des études récentes en psycho-acoustique [MD89] ont insisté sur ce point. Les différents mécanismes impliqués dans les chaînes d'apprentissage et de reconnaissance des sons sembleraient autant relever de propriétés observables dans le signal sonore que des propriétés intrinsèques aux objets qui les produisent.

L'acte de composition est donc inévitablement imprégné de toute une expérience physique du monde environnant, et en particulier, d'une expérience gestuelle de ce monde.

## **3. le chef d'orchestre**

Le chef d'orchestre, comme le musicien-interprète, lit une partition. Mais au lieu de la traduire en un geste instrumental, il la traduit en des gestes qui expriment la naissance, la vie et la mort du son, le tempo et la pulsation, le phrasé et le mode d'enchaînement des sons, la nuance, etc... Le geste du chef d'orchestre a donc une fonction d'expression en soi.

Il s'élabore sur la base d'un véritable vocabulaire, convenu avec les musiciens, et permet de décrire toutes les situations de jeu. La majorité des informations de base se trouvent dans la partition. Elles sont décrites explicitement (le rythme, les nuances, le phrasé, etc.), ou implicitement (mode du phrasé, graduation des nuances, etc).

" Il en découle que d'un langage imparfait (l'écriture musicale), et par l'intermédiaire d'un rudiment de langage (la gestique), il est possible d'obtenir une exécution d'une précision quasi absolue. " [Fan88]

On peut ramener la gestique de base, au tableau suivant[Fan88] (Figure 8) :

ANALYSE DE BASE	CARACTERISTIQUES	BRAS	CARACTERES DU GESTE
tempo	métrique	droit	métrique
nuance	pulsation	droit-gauche	redondant
	qualité de l'attaque	droit-gauche	amplitude-vitesse
	intensité expressive	droit-gauche	poids-amplitude
conduite du son	legato/non legato/staccato	droit	conduite du geste
spatialisation	du collectif à l'individuel	regard	amplitude
vie du son	du plus court au plus long	droit-gauche	sec-étiré
mort du son	terminaisons/silences/ suspensifs	droit-gauche	mode d'arrêt du geste
phrasé	attaques/ respirations/ cadences/ appuis	droit-gauche	élans-appels-levées
style	ex.sf chez MOZART	droit-gauche	amplitude-vitesse- poids-sécheresse

#### 4. conclusion

Selon G.Mentel, " si le compositeur va du geste à la composition, l'interprète doit faire, lui, le chemin inverse, c'est à dire aller de la composition, de la partition, au geste" [Men88].

Mais dans la pratique, le musicien n'est jamais uniquement instrumentiste, ni exclusivement compositeur. Chacun combine plus ou moins son activité principale avec les autres. Par exemple, le geste de l'instrumentiste, et même s'il sert d'abord à sélectionner, à modifier et exciter l'instrument, conserve également une fonction liée à l'expression. Delalande [Del88] illustre cela en décrivant les gestes du pianiste G.Gould :

"dès que sa main gauche est libre, GOULD en profite pour «diriger un orchestre imaginaire» [...] Sa conduite paraît divisée en deux orientations, l'une «esthésique» , tournée vers la réception, et l'autre «poïétique», vers la production".

Ainsi, les gestes pour la musique sont de différentes natures. Mais quels qu'ils soient, utilisés de façon directe ou dérivée, comme langage ou pour manipuler et s'informer, les gestes interviennent de façon déterminante dans le processus de création musicale traditionnelle.

#### IV.2. geste instrumental et claviers

Le geste instrumental participe directement à l'élaboration des sons et implique un contact direct entre l'instrument et le musicien.

Les facteurs de sa complexité sont multiples : il y a le musicien, bien sûr, mais aussi l'instrument, caractérisé par sa morphologie et par ses mécanismes excitateurs et vibratoires. Les différences du geste instrumental exécuté sur un violon et sur un piano apparaissent de soi parce que ces instruments ont une morphologie différente. En revanche, les nuances de geste instrumental entre les instruments de même apparence sont plus subtiles. Prenons en exemple les instruments à clavier.

##### 1. le clavecin

L'idée à la base du clavecin est très ancienne. Elle consiste à tendre un grand nombre de cordes en parallèle sur une caisse de résonance et à les exciter à l'aide d'un plectre (cithare, guitare, etc.). Le son est d'autant plus sec que le matériau est plus raide et dur. Sur la cithare, et étant donnée la proximité des cordes, il est très difficile de les pincer sans se tromper. L'idée est donc apparue d'une mécanique qui ferait cette opération, par l'intermédiaire d'un clavier. C'est le principe du clavecin.

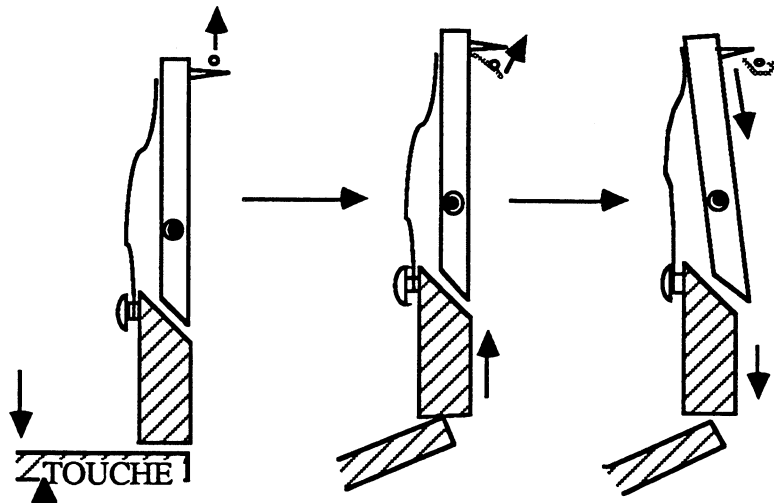


Figure 9. - *Le mécanisme du clavecin. Le sautereau.* Les 3 figures montrent comment la corde est pincée une seule fois lorsqu'on frappe une note : le sautereau «échappe» en effet lors du retour de la touche. Figure reproduite à partir de [Lei71].

Cette mécanique comporte des touches noires pour les notes naturelles et blanches pour les altérations (contrairement au piano). Le dispositif qui permet de pincer la corde est le sautereau : lorsqu'on appuie sur la touche, une pointe souple vient en contact avec la corde, l'accroche, la soulève et glisse le long de la corde.

Le clavecin sous sa forme actuelle se présente de diverses façons, avec des mécanismes de pincement multiple, un jeu de luth, etc...

Le clavecin présente musicalement des caractéristiques propres, parfois qualifiées de "peu expressives" : le pincement de la corde se ferait de façon identique, quoi qu'on fasse.

En réalité, il y a plusieurs manières de pincer la corde, selon que la touche est enfoncée rapidement ou lentement. La pédale également permet de modifier l'intensité et l'amortissement des cordes, et d'influencer le son.

Contrairement aux autres instruments à corde (guitare, cithare), le clavecin permet un jeu sûr, léger et rapide grâce au clavier, mais moins expressif que le permet la guitare ou la cithare.

## **2. le piano**

Au XVII<sup>e</sup> apparaît le clavicorde, dont le principe simple repose sur la frappe d'une corde par l'intermédiaire d'une tangente fixée à l'extrémité de la touche. Au XVIII<sup>e</sup> Cristofori invente la mécanique de piano. Elle comporte les éléments essentiels actuels : marteau articulé, système d'échappement et étouffoir. Au début du XIX<sup>e</sup> apparaît le mécanisme à "double échappement" qui permet de rejouer la note sans que l'étouffoir ne soit plaqué contre la corde.

Pour ce qui est du jeu pianistique, le seul paramètre concernant la modification du timbre est la vitesse avec laquelle le marteau touche la corde. Avec le mécanisme d'échappement, dès que le marteau a échappé, tout mouvement de la touche devient "inutile".

En outre, le musicien dispose d'un moyen d'action sur le transitoire d'extinction des sons par le biais des pédales. La pédale de droite lève tous les étouffoirs et permet donc à toutes les cordes de vibrer. La pédale de gauche rapproche les têtes de marteau et diminue donc l'intensité. Une troisième pédale permet de ne pas étouffer les notes jouées.

Le pianiste peut donc moduler le timbre dans une certaine mesure, en déployant plus ou moins d'énergie à l'attaque du son et en jouant sur son extinction.

### **3. clavier de contrôle**

#### **a. l'orgue classique**

L'orgue classique est une machine très élaborée dont la mise au point a duré des siècles. Le principe de base repose sur un mécanisme d'excitation très sophistiqué d'un ou de plusieurs tuyaux par un jet d'air adapté.

Dans l'orgue mécanique, on peut enfoncer plus ou moins vite la touche, et donc ouvrir plus ou moins vite la soupape par laquelle va se dégager le jet d'air. On joue donc sur les transitoires d'attaque et le musicien peut ainsi agir sur le timbre et le moduler dans une certaine mesure.

Avec l'invention de l'électro-aimant, au lieu de toutes les complications mécaniques, un simple contact sous la touche suffit. Lorsqu'on abaisse celle-ci, un moteur tire sur la soupape. Mais avec lui, la nuance sur l'attaque s'est perdue : le mouvement d'ouverture de la soupape est stéréotypé.

#### **b. le synthétiseur à clavier**

L'ordinateur utilisé pour la synthèse du son est resté, jusqu'à il y a encore quelques années seulement, à la disposition exclusive des laboratoires de recherche et de certains compositeurs. En revanche, les retombées des concepts de synthèse ont permis la réalisation de dispositifs grand-public baptisés rapidement "synthétiseurs".

Un synthétiseur est constitué d'une unité de génération des sons, et d'un système de contrôle. Habituellement, les synthétiseurs possèdent un clavier à "toucher léger", parfois équipé d'un "aftertouch" qui permet de contrôler certains paramètres du son selon la pression exercée sur les touches. D'autres contrôleurs permettent d'accroître le potentiel d'expressivité : pédale de soutien, molette de vibrato, etc...

Le clavier du synthétiseur ne dispose pas de mécanisme d'échappement comme le piano, ou de pincement comme celui du clavecin. Généralement, le comportement de chaque touche est déterminé par un simple ressort qui ramène la touche à une position d'équilibre après une excitation.

Outre les intérêts purement commerciaux, une raison majeure justifie cette mécanique. Avec un synthétiseur classique, il est théoriquement possible de créer tous les sons. Pour décrire les sons "naturels", il est d'usage de les qualifier du nom de la référence instrumentale. On parle de "sons de hautbois", de "sons de piano" par exemple.

Cependant, la restitution de ces phénomènes acoustiques n'implique en rien la restitution des comportements mécaniques de leur référence instrumentale. Il se pose

alors un problème de cohérence entre les sons ayant une référence physique et le moyen de les contrôler.

La solution adoptée a consisté à rendre l'interface gestuelle la plus neutre possible. La réalisation de claviers de type piano, dont la mécanique, rudimentaire, reste très éloignée de celle des instruments à clavier traditionnels, remplit en quelques sortes ce contrat de neutralité.

### **c. CORDIS et clavier rétroactif modulaire**

Le clavier rétroactif modulaire est un prototype conçu dans les "ateliers" de l'ACROE (Association pour la Création et la Recherche sur les Outils d'Expression à Grenoble).

Ce clavier est un transducteur gestuel à retour d'effort : il sert d'interface pour la simulation en temps réel d'objets physiques, et permet donc de restituer les comportements de ces objets et de les contrôler.

Les bras de touche de ce clavier sont tous identiques mais il est possible de programmer l'amplitude de repos de chacun d'eux pour obtenir l'alternance blanche-noire d'un clavier de piano et d'y ajouter une plaquette qui détermine l'avancée des touches blanches du piano. Bien entendu, cet "habillage" est un cas particulier : il est préférable de considérer le clavier rétroactif modulaire comme un ensemble de N degrés de liberté, chaque degré de liberté étant associé à un bras de touche.

Dans ce contexte, le musicien peut retrouver toute la subtilité et toute la richesse du jeu instrumental propre à l'instrument simulé. Si l'instrument simulé est un clavecin, le jeu sera sûr, léger, etc...

### **5. conclusion**

A de rares exceptions près, il faut admettre que par le clavier, le musicien n'a de pouvoir que sur l'apparition et la disparition des sons.

"Le jeu au clavier repose avant tout sur le maniement des touches. Si l'on élimine toutes les possibilités d'accès au son qui ne procèdent pas du clavier comme les registrateurs, les accouplements, les pédales, les potentiomètres, jouer du clavier se résume à un contrôle temporel du mouvement des touches et, dans certains cas, à la maîtrise de la force avec laquelle on frappe ces touches"[Hau87]

La famille des instruments à clavier réunit des instruments très différents dans leur mécanisme producteur de son : corde percutée, cordes pincées, tuyau excité par un jet d'air... Ils n'ont de commun que le clavier qui aujourd'hui se présente comme un



ensemble de touches à mouvement vertical, diatonique majeur, chromatique à douze sons, reproductibles à l'octave. Il a fallu 16 siècles aux générations de facteurs pour le produire.

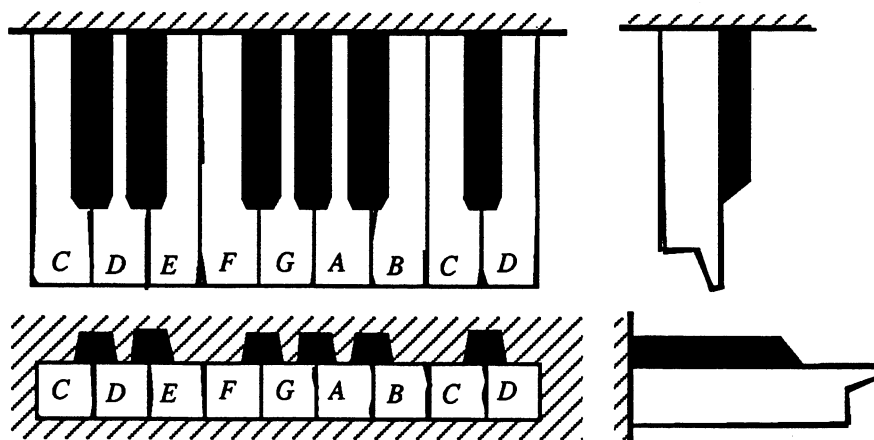


Figure 10. - *Coupes de piano*. Vue de profil, vue de face et vue du dessus d'une touche de clavier de piano du Do(C) au Ré(D). "Vu par dessus, le clavier présente une succession de formes identiques : l'octave, adapté à l'empan de la main, qu'il reproduit plus ou moins de fois selon la tessiture. A l'avant du clavier l'octave est diatonique à 7 touches ; à l'arrière elle est chromatique à 12 touches"[Hau87]

J.Haury propose un classement des claviers selon la forme de l'enveloppe et le contrôle de l'intensité du son :

- \* claviers à son droit : orgue à tuyau, orgue électrique, harmonium, accordéon
- \* claviers à son décroissant : clavicorde, clavecin, piano forte, piano
- \* clavier «non-sensible» sans contrôle de l'intensité : la registration, les potentiomètres, l'accouplement des jeux font office de variateur d'intensité. C'est l'ensemble du clavier qui est affecté : orgue à tuyau, orgue électrique, harmonium
- \* clavier «sensible» permettant le contrôle de l'intensité séparément sur chaque touche : clavicorde, pianoforte, piano
- \* le synthétiseur peut imiter toutes les formes de son et tous les types de contrôle et d'intensité.

A cette liste il faut à présent ajouter le clavier rétroactif modulaire couplé au système de simulation CORDIS. L'ensemble permet de synthétiser les instruments eux-mêmes, de les contrôler directement par le clavier, et par ce biais, de contrôler (indirectement) l'intensité des sons. Tous les types de contrôle d'instrument à clavier sont donc reproductibles.

## **V. ELEMENTS DE TYPOLOGIE DU GESTE INSTRUMENTAL**

Les nombreuses tentatives de classification des mouvements, faites le plus souvent dans une perspective ergonomique, ont permis d'insister sur la complexité et le nombre élevé des critères de classement.

Dans ce paragraphe, nous énumérons un ensemble de critères, qui permettent de décrire et caractériser au mieux le geste instrumental. Trois approches sont proposées :

- une approche phénoménologique (analyse descriptive),
- à partir des fonctions que le geste exerce sur l'instrument (analyse fonctionnelle),
- du point de vue du musicien (analyse intrinsèque)

### **V.1. analyse descriptive**

#### **1. critère cinématique**

Parmi les critères de classement proposés dans les sciences biologiques, le critère de vitesse d'exécution semble être le plus généralement admis. Ainsi Bailey et Presgrave [BP58] distinguaient les mouvements rapides, balistiques, dont l'arrêt est provoqué par l'impact sur une surface, les mouvements peu précis mais dont la moindre rapidité est liée à la nécessité d'un freinage actif, enfin les mouvements plus lents impliquant un ajustement précis sur un objectif.

#### **2. critère spatial**

Dans les situations instrumentales habituelles, les mouvements de l'instrumentiste se développent dans des espaces aux dimensions variables : les cas limites que nous considérons sont d'une part les mouvements où l'implication du corps est grande, et d'autre part les mouvements plus fins, impliquant essentiellement la main.

Pour ce qui concerne les mouvements manuels, D.Beaubaton et J.Paillard [Bea80][PB80] expliquent l'existence d'un double contrôle : celui du bras, responsable du transport de la main vers un objectif spatialement défini et celui de la main, au service des opérations de préhension et de manipulation. Les mouvements de grande amplitude sont pris en charge par le bras, alors que la main ne réalise que des mouvements de faible amplitude.

Par ailleurs, JL.Florens [Flo84] considère que "le comportement gestuel se développe selon des axes principaux en nombre limité, à partir desquels des variations sont

possibles". Cette analyse, qualitative et globale, peut simplifier le classement par le fait que ces variations n'ont pas toujours un rôle fondamental.

Un mouvement de frottement d'archet par exemple, peut s'inscrire dans la famille des mouvements de grande amplitude ( $\approx 50\text{cm}$ ) en trajectoire rectiligne ; le mouvement de frappe d'une touche de piano dans la famille des mouvements de faible amplitude ( $\approx 5\text{cm}$ ) en trajectoire rectiligne verticale.

### **3. critère fréquentiel**

A. Berthoz [Ber75] décrit une technique de décomposition des mouvements en leur contenu fréquentiel. La description fréquentielle des activités motrices, valable au sens strict pour des mouvements périodiques a été étendue à des mouvements uniques (transitoire), et à des mouvements aléatoires. L'auteur remarque que la plupart des mouvements humains se situent dans une bande de fréquence comprise entre 0 (immobilité) et 10Hz. Le tremblement physiologique des doigts, activité non contrôlée (et donc de référence à ce titre) a la bande de fréquence maximale (10Hz).

Type de mouvement	Bandes de fréquence Hertz (Hz)	
Oscillations posturales .....	0 - 3	
Respiration .....	0 - 2	
Course (compétition) .....	1.5	(1500m)
	1.3	(100m)
Natation (Crawl) .....	1 - 3	(bras)
	3 - 5	(jambes)
Poursuite manuelle d'une cible .....	0 - 2	
Poursuite oculaire d'une cible .....	0 - 2	
Mouvements de la tête		
- au cours de la marche .....	0 - 4	
- évoqués par impulsion de forces .....	0 - 3	
Saut (rythme spontané) .....	2	
Musique de danse .....	2.2	
Ecriture .....	4 - 6	
Tremblement Parkinsonien .....	5 - 8	
Mouvements rythmiques		
(avant-bras, fréquence maximale) .....	4 - 6	
Frappe des doigts (fréquence maximale) .	8 - 10	
Tremblement physiologique (doigts) .....	10	

Figure 12. Bandes de fréquence de mouvements humains (valeurs indicatives). Tableau extrait de [Ber75]

## V.2. analyse fonctionnelle

Le geste qu'implique le jeu instrumental s'analyse a priori, selon F.Delalande [Del88], en au moins trois niveaux, qui vont du purement fonctionnel au purement symbolique.

Le **geste effecteur** - frotter, souffler, appuyer sur la touche... - est nécessaire pour produire mécaniquement le son. Ensuite le **geste accompagnateur** engage le corps entier. L'instrumentiste associe aux mouvements strictement indispensables d'autres qui le sont apparemment moins : geste du buste, des épaules, mimiques, respiration pour un pianiste, etc. L'auditeur pour sa part perçoit un **geste figuré** : un balancement dans une mélodie - ce qui ne suppose pas que le pianiste se soit balancé.

### 1. critère effecteur

C.Cadoz [Cad87][Cad88] précise et complète la notion de geste effecteur donnée par F.Delalande et met ainsi en évidence trois catégories de gestes effecteurs, qui font intervenir des processus énergétiques différents :

\* le geste d'excitation est celui par lequel l'instrumentiste communique à l'instrument l'énergie qui se transformera en énergie vibratoire. Il s'agit par exemple de la manipulation de l'archet d'un violon, du pincement d'une corde de guitare, ou encore le geste d'enfoncement d'une touche de piano.

\* le geste de modulation a pour fonction de modifier certaines caractéristiques qualitatives ou quantitatives de la structure de l'instrument.

Il s'agit soit d'une modification structurelle - le musicien introduit ou supprime des éléments dans le système vibrant -, soit d'une modification paramétrique - il n'y a pas de changement qualitatif de la structure -.

\* le geste de sélection consiste à désigner un ou plusieurs éléments d'un instrument et se manifeste surtout sur les instruments polyphoniques. On peut citer comme exemple la sélection d'une touche ou d'un ensemble de touches sur un clavier de piano.

Ce type de geste correspond pour l'essentiel au transport de l'organe de préhension et de manipulation (main ou pied) et précède la phase d'excitation. Il n'y a donc pas d'échange énergétique entre le musicien et l'instrument à ce moment-là.

### le geste d'excitation

Lors d'un geste d'excitation, le musicien communique de l'énergie à la structure vibrante d'un instrument, soit directement avec une partie de son corps (percussion manuelle d'une membrane ou pincement digital d'une corde par exemple), soit par le biais d'un mécanisme excitateur. Dans ce deuxième cas, trois situations sont à distinguer :

1- le musicien prolonge son corps d'un dispositif mécanique, indépendant de l'instrument : un archet, une mailloche, un plectre...

2- le musicien excite la structure vibrante par l'intermédiaire d'un dispositif excitateur couplé à l'instrument, mais indépendant mécaniquement du corps de l'homme. C'est le cas du piano où la structure vibrante n'est accessible (sauf détournement de la fonction initiale de l'instrument) que par le biais des touches du clavier.

3- l'excitateur se divise en deux parties mécaniquement indépendantes : l'une couplée au musicien, l'autre à l'instrument. C'est le cas des cymbales qui servent la double fonction d'excitateur et de vibrateur : l'homme tient dans une main une cymbale excitateur de la seconde, qui est elle même excitateur de la première.

En résumé, la répartition de la structure excitatrice complète ( $E^*--E^*$ ) entre le musicien (H) et la structure vibrante (SV) de l'instrument peut se décrire par la relation suivante :  $H.E^*--E^*.SV$  où  $E^*$  signifie la présence ou non d'une partie de l'excitateur, sur le musicien ( $H.E^*$ ) et / ou sur la structure vibrante ( $E^*.SV$ ).

### **2. critère de saisie**

Le critère de saisie permet de distinguer deux types de gestes manuels : celui avec préhension de l'objet (manipulation d'un archet ou d'une baguette de percussion) et celui sans (percussion d'une touche de piano par exemple).

#### **a. contact temporaire/permanent**

Pour une expérience se réalisant sur un intervalle de temps  $\Delta T$  le contact entre le musicien et l'instrument peut être soit "permanent", soit "temporaire".

Le contact permanent signifie que le couplage musicien-instrument est maintenu pendant toute la durée  $\Delta T$ . Lors de ce mode de manipulation, l'objet suivra toutes les évolutions du mouvement de la main du musicien.

En revanche, le contact temporaire signifie que le couplage a été rompu une fois au moins pendant cette période. L'objet n'est alors entraîné que dans un espace en général limité et dans un sens privilégié du geste.

#### **b. avec saisie**

Le geste préhensile consiste à saisir et manipuler ou déplacer un objet. Pour ce type de geste qui suppose un contact durable avec l'objet, et qui met à contribution une zone sensorielle large, la proprioception revêt une importance prépondérante par rapport à la stabilité de maintien de l'objet, et aux corrections apportées à l'action motrice elle-même.

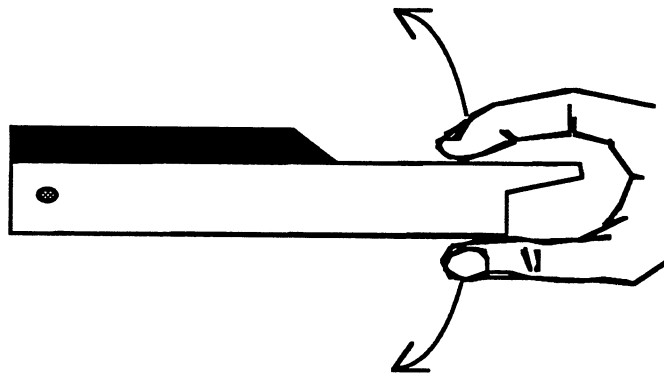


Figure 13. - *Geste préhensile*. Consiste à saisir et manipuler, ou déplacer un objet. Il faut distinguer la manipulation et le transport de l'objet.

Deux fonctions de préhension sont à distinguer :

- la déformation de l'objet saisi : c'est le cas de nombreux gestes de modification structurelle et paramétrique : tourner la molette pour accorder une guitare par exemple.
- le transport de l'objet saisi : c'est le cas des jeux percussifs avec mailloche, ou des gestes de mouvement d'archet.

#### **c. sans saisie**

1/ la percussion consiste à taper un objet ou à le pousser. La percussion manuelle provoque généralement l'entraînement de la main dans un mouvement de transport du bras. Elle suppose la mobilité de tout ou d'une partie du membre supérieur sans faire participer les récepteurs tactiles localisés au niveau des doigts.

2/ le geste de frottement est un geste qui met à contribution une zone sensorielle large, et se caractérise par un temps de contact durable.

3/ le geste de butée s'applique à des objets en mouvement pour contraindre leur espace d'évolution. A la différence du geste percussif, il est relativement statique, quasi stationnaire. L'instrument, par son mouvement propre, vient buter sur les doigts ou la main du musicien.

Etouffer l'oscillation d'une corde de guitare en vibration est un exemple de geste de butée.

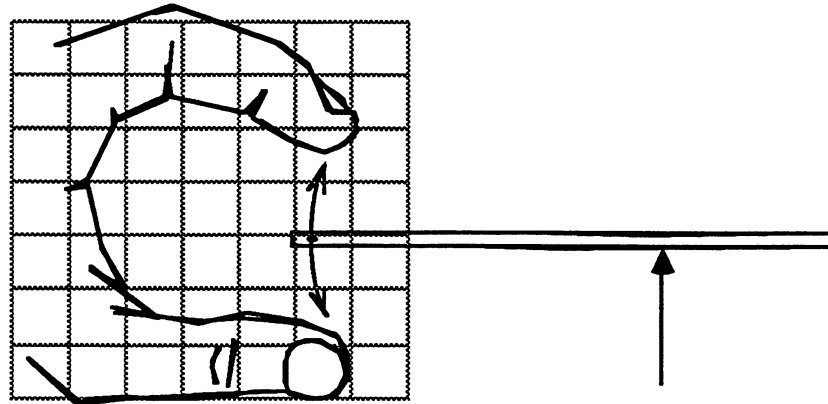


Figure 14. - *Geste de butée*. Exemple de geste de butée dans le cas d'une touche oscillante. Les deux doigts se substituent aux butées naturelles de l'instrument.

### **V.3. analyse intrinsèque**

On s'intéresse ici aux conditions de production du geste par l'instrumentiste.

Les instruments traditionnels accordent une grande place au geste manuel, certainement parce qu'il constitue l'une des formes les plus perfectionnées du système gestuel humain. La main, organe de saisie et de manipulation, est extraordinaire de possibilités et de vivacité. Elle comporte un grand nombre de degrés de liberté par la multiplicité et la disposition des attaches articulaires. Elle est dotée de plus d'une très forte densité de récepteurs nerveux, ce qui lui confère une excellente capacité de discrimination sensorielle.

Le pied, plus rigide, est utilisé pour des mouvements plus lents et plus statiques, tels que ceux qui actionnent les pédales d'un piano ou d'un orgue.

Le reste du corps sert à leur transport (le bras et la jambe), à un accompagnement (le coude, l'épaule ou la tête par exemple qui décrivent le rythme), mais aussi à la stabilité de l'instrument (maintien d'une harpe ou d'un violon).

contrôle et programme moteurs

Dans certains cas, on s'intéresse aux conditions physiologiques de la genèse du mouvement dans le système humain. Le problème majeur est de comprendre l'organisation du mouvement, de clarifier les interactions entre les commandes issues des centres nerveux et les informations sensorielles.

Parmi les nombreuses hypothèses faites pour expliquer l'organisation comportementale on trouve la notion de programme moteur, dérivée de la terminologie informatique.

Selon D.Beaubaton "les deux modalités fondamentales du contrôle moteur, impliquant d'une part une régulation proactive et d'autre part des ajustements rétroactifs, peuvent être assimilées aux mécanismes de base d'un programme informatique. Ce programme ne peut être conçu que comme un dispositif synthétique, comportant un ensemble de structures d'instructions et de boucles conditionnelles, mettant en relation de manière dynamique l'unité centrale et les périphériques." [Bea83]

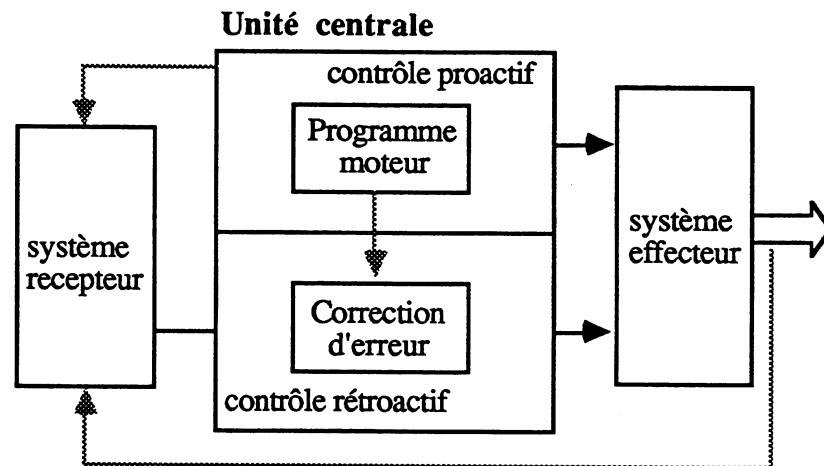


Figure 15. - Représentation d'un système de contrôle de la motricité. Le programme central peut commander directement les systèmes effecteurs ou intervenir au niveau d'un «comparateur». Dans ce cas, c'est après un processus de détection d'erreur que les effecteurs vont être commandés. Les systèmes effecteurs reçoivent des informations d'origine périphérique mais sont également contrôlés par l'unité centrale. Figure inspirée de [Cha65]

V.4. conclusion

Dans l'activité musicale traditionnelle, les gestes digitaux, rapides ou lents, de grande ou faible amplitude, temporaires ou permanents, en saisie ou non, pour l'excitation ou la modulation, constituent une base importante du geste de l'instrumentiste. Les critères de classification que nous avons proposés, loin de décrire toute la variété du geste instrumental, considèrent au moins ces éléments de base.



## **VI. CAPTURE DU GESTE INSTRUMENTAL**

L'analyse du geste est un problème difficile et les nombreux travaux entrepris à ce sujet l'ont montré. La majorité des études réalisées, notamment dans les sciences biologiques, cible en priorité les processus internes qui activent le mouvement, depuis la prise de décision, jusqu'à la commande chimique des muscles. Les travaux de Beaubaton [Bea83] de Cooke [Coo79] et de Hatwell le montrent [Hat86].

Pour analyser objectivement le geste, il faut pouvoir l'observer hors de son temps, selon des moyens adaptés à nos sens et à nos modes de représentation. Cela signifie que l'analyse du geste ne peut être menée rigoureusement que conjointement à l'élaboration de dispositifs aptes à le capter et à le mémoriser.

### **VI.1. quelques définitions**

La situation instrumentale met en relation un opérateur et un instrument. L'opérateur, c'est d'un point de vue purement physique, un système mécanique actif. L'instrument est ici pris dans un sens très général. C'est un système mécanique manipulable susceptible de déplacements et de déformations.

Le jeu instrumental est l'histoire de cette relation. L'opérateur et l'instrument se caractérisent chacun par des "zones de contact" entre lesquelles des "liaisons" (ou contacts) peuvent s'établir. Le jeu instrumental se décrit alors comme une séquence de liaisons mécaniques qui s'établissent ou se défont entre l'opérateur et l'instrument.

Le geste instrumental est le phénomène produit du jeu instrumental. C'est un ensemble de mouvements et de forces exercés par l'opérateur qui s'exécutent soit librement, soit en couplage avec l'instrument. Le geste instrumental consiste en la totalité de ces comportements, y compris ceux qui conduisent à des changements de contact.

### **VI.2. représentation discrète du geste instrumental**

L'espace naturel dans lequel évolue le geste instrumental est un espace continu. Lorsque la structure mécanique du musicien entre en contact avec l'instrument, ce sont des lignes, et plus généralement des surfaces qui se rencontrent, s'effleurent, se percutent ou se repoussent. En revanche, dans le cadre de la représentation avec l'ordinateur, toutes les données manipulées sont par essence discrètes. La notion d'infini et d'espace continu lui est inconnue.

Pour ces raisons techniques, on substitue aux zones de contact des points théoriques de contact qui peuvent avoir chacun plusieurs degrés de liberté. On aura alors des points de contact opérateur et des points de contact instrument.

Un contact (ou liaison) entre l'homme et l'instrument consiste en un couple (h, i) où h et i sont respectivement des points de contact opérateur et des points de contact instrument.

Le geste instrumental sous forme discrète est entièrement caractérisé par :

- 1 - le comportement mécanique de chaque point de contact (h) de l'opérateur,
- 2 - l'histoire des contacts (h, i) établis.

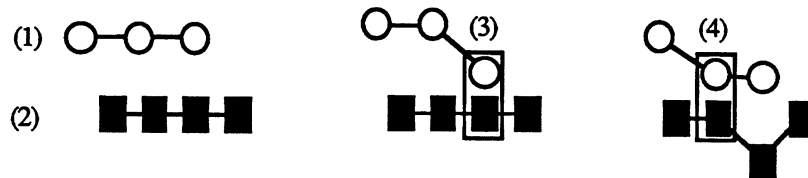


Figure 16. - Représentation discrète du geste instrumental. Les figures (1) et (2) sont des représentations discrètes respectivement de l'opérateur et de l'instrument. Les figures (3) et (4) montrent deux contacts possibles.

### VI.3. capture du geste instrumental

Pendant le jeu instrumental, les comportements mécaniques des points de contact, aussi bien opérateurs qu'instruments, sont entièrement caractérisables par deux catégories de variables mécaniques duales telles que forces et déplacements.

"Capturer" le geste instrumental revient alors d'une part à capter l'évolution des variables mécaniques au niveau même des points de contact, (c'est à dire à convertir le phénomène en un signal électrique puis numérique) et d'autre part à mémoriser ce signal.

#### 1. sans réduction

En premier lieu, ce sont les points de contact-opérateur qu'il faut considérer. Les capteurs sont disposés sur l'opérateur lui-même.

Le gant numérique [Fol87] illustre cette approche. Il détecte certains mouvements de la main et des doigts, et les code sous forme de signaux électriques. Mais la grande complexité articulaire des membres de l'opérateur, de même que la nécessité de gêner au minimum ses déplacements et son confort, rendent la réalisation de ces dispositifs difficiles.

En second lieu, il faut déterminer l'histoire des liaisons établies et pour cela, considérer les points de contact instrument. Une liaison entre un point de l'opérateur et un point de l'instrument est établi lorsque leurs variables mécaniques respectives indiquent une même valeur de position par exemple.

## **2. sur l'instrument**

Une démarche moins complète, mais néanmoins plus simple, consiste à ne placer des capteurs que sur l'instrument et laisser l'homme libre de ses mouvements.

Les difficultés techniques sont quelque peu réduites, mais en même temps, il n'est plus possible de capter le geste instrumental de façon exhaustive.

### **a. réduction 1**

On perd deux types d'informations gestuelles :

- le comportement mécanique des points de contact-opérateur en dehors des contacts effectifs,
- quel point de l'opérateur est en contact avec quel point de l'instrument.

### **b. réduction 2**

A la réduction précédente, incontournable par principe, vient s'ajouter une seconde. Les capteurs ne sont pas toujours placés strictement au niveau des points de contact. Il y a alors un décalage entre les points de contact et les "points de capture". Ceci est lié, soit à une impossibilité matérielle, soit à une optimisation technique et/ou conceptuelle.

Par exemple, on peut considérer que l'aspect pertinent du geste instrumental excitant une touche de piano n'est pas tant l'endroit précis de la touche où se réalise le contact, que son évolution en amplitude. Dès lors un capteur monté en bout de la touche et mesurant son amplitude est suffisant.

En toute rigueur, il est alors impropre de parler de "variables gestuelles" à propos des signaux issus des capteurs placés sur l'instrument. Ce que l'on capte, ce sont des signaux qui en fait caractérisent les comportements de l'instrument au niveau des points de capture.

Mais cette information demeure significative dans la mesure où elle contient de manière exhaustive la participation de l'opérateur aux échanges énergétiques, lors des phases de couplage avec l'instrument.

On peut par ailleurs envisager des processus d'identification pour compléter l'information dont on dispose avec une information dont on ne dispose pas. Par exemple en essayant de retrouver les phases de contact opérateur-instrument et d'estimer le comportement de l'opérateur en dehors de ces phases.

### **3. fréquence d'échantillonnage**

La fréquence d'échantillonnage des capteurs est également un facteur susceptible de réduire le geste instrumental. D'après le théorème de l'échantillonnage, si cette fréquence est inférieure au double de la fréquence maximale dans la bande passante du geste instrumental, l'information exhaustive initiale est perdue.

De nombreuses études ont été faites pour déterminer la bande de fréquence de certains mouvements. A.Berthoz [Ber75] (cf. chapitre I, §V.1.3) récapitule les résultats importants obtenus. La fréquence maximale est de  $F_H=10\text{Hz}$  et concerne le tremblement physiologique des doigts. Cependant, les mouvements cités ont un point commun. Il s'agit de mouvements dont la composante instrumentale reste très faible, voire inexistante. Le geste instrumental n'est donc pas représenté de façon significative.

Pour capter les mouvements sans composante instrumentale importante (respiration, natation, écriture, tremblement, etc.) et sans perte d'information, il suffit de les échantillonner à une fréquence supérieure à  $2 * F_H = 20\text{Hz}$ . En revanche, pour les mouvements ayant une composante instrumentale importante, il faut tenir compte, a priori, de la bande de fréquence de l'instrument ( $F_I$ ), et cela d'autant plus que l'instrument est actif.

On peut donc définir la fréquence d'échantillonnage minimale  $F_e$  du geste instrumental par la relation suivante :

$$F_e \geq 2 * \text{Max}(F_H, F_I)$$

### **4. exemples de capture avec clavier**

#### **a. le système CORDIS**

Avec le système CORDIS et le clavier modulaire à retour d'effort (cf. chapitre I, §III), le geste instrumental est capté et mémorisé sous la forme d'un ensemble de signaux échantillonnés à une fréquence, variant selon la complexité des algorithmes simulant l'instrument, entre 300Hz et 1kHz. Chaque signal est issu d'un capteur de position, monté en bout de chaque touche du clavier.

Les données captées sont stockées sur un disque selon différents codages. Le plus simple, appelé codage brut, conserve les signaux échantillonnés tels qu'ils sont captés. D'autres méthodes de codage, efficace d'un point de vue de la compression des données, ont été proposées [Cad87].

**b. le système 290 SE**

Le système 290 SE [MR90] est monté sur un piano Bösendorfer et mesure le geste instrumental au travers des mouvements des touches et des pédales du piano. Les données sont stockées sur un disque ou bande magnétique pour un traitement et un "playback" éventuel.

La capture du geste instrumental se fait par mesure de la vitesse de déplacement des touches et plus exactement des marteaux qui viennent percuter les cordes. Il s'agit de capteurs optiques constitués chacun d'un émetteur de lumière (LED) et d'un photo-transistor. Lors de l'activation d'un marteau, une fine plaquette d'aluminium monte ou descend et vient obstruer plus ou moins la zone d'éclairage de la diode. La fréquence d'échantillonnage d'un capteur est de l'ordre de 800Hz.

Pour réaliser le "playback" d'une pièce enregistrée, des moteurs linéaires sont montés sous chaque touche du piano.

Il faut cependant remarquer que seuls les comportements des marteaux sont captés et mémorisés. Les étouffoirs ne sont pas analysés. Il y a donc une perte d'information par rapport au jeu instrumental qui sera particulièrement sensible lors du "playback".

Les pédales disposent également de capteurs et de moteurs. Les capteurs sont de nature différente des capteurs optiques et mesurent la position des pédales avec une fréquence d'échantillonnage variant entre 1Hz et 256Hz. Les moteurs nécessitent une plus grande puissance que ceux utilisés pour contrôler les touches du clavier. Les moteurs des pédales sont échantillonnés entre 50Hz et 100Hz.

**c. le système Moog MTS**

Le système Moog MTS [MR90] est conçu uniquement pour mesurer le geste instrumental sur les instruments à clavier. Aucun moyen de "playback" n'est envisagé. Une attention particulière a été accordée sur la richesse de la mesure. Il s'agit en l'occurrence de mesurer 1/ l'enfoncement et le relâchement des touches et 2/ la position des doigts sur les touches.

Le principe technologique utilisé est celui du film résistif qui tapisse chacune des touches du clavier. Un circuit électrique calcule en permanence la différence de potentiel entre les deux surfaces du film. Lorsqu'un contact est réalisé entre un doigt du musicien et le clavier, le courant change à l'endroit de couplage.

**d. récapitulatif**

Le système 290 SE et le système CORDIS envisagent la capture du geste instrumental avec une réduction du second type : ni les contacts, ni les points de ces contacts ne sont mesurés. C'est le comportement des touches qui est capté, soit en position, soit en vitesse, soit en force.

Le système Moog MTS capte en revanche le geste instrumental selon 3 axes : latéral, frontal et en profondeur. Plus complet que les deux précédents du point de vue de la capture, il réalise une réduction du premier type : l'information caractéristique en dehors des contacts, de même que la manière spécifique et dynamique dont s'établissent les contacts restent inconnus.

## **VII. CONCLUSION**

Considérons que le geste instrumental intervient de façon signifiante dans le processus de création musicale.

A un premier niveau, il permet de modifier les propriétés mécaniques d'un instrument et lui fournit de l'énergie pour le faire vibrer. Il participe de ce fait directement à l'élaboration des phénomènes acoustiques. A un deuxième niveau, il intervient dans le processus de composition. Les représentations mentales du compositeur sont influencées et imprégnées par son expérience du monde physique.

Considérons également que la technologie actuelle nous offre des moyens suffisants d'une part pour extraire d'un geste instrumental naturel les données représentatives, et d'autre part, pour restituer ce "geste", mémorisé, et faire (re)naître les phénomènes sensibles de l'objet déformable auquel il s'applique.

La combinaison de ces considérations permet d'envisager un nouveau problème, qui est susceptible de compléter et même de généraliser la notion de partition et de notation musicale.

Il s'agit de se saisir du geste comme d'un objet, en dehors de son écoulement temporel, c'est à dire d'opérer sur ses mémorisations et de produire, par différentes formes d'interventions, de nouveaux objets, mémorisables à leur tour et utilisables sans distinction particulière en lecture.

### **VII.1. édition du geste instrumental**

En isolant le phénomène gestuel des phénomènes sonores et visuels, on s'intéresse à une analyse et à un traitement de celui-ci, qui suppose que l'on soit capable de le décomposer en éléments, puis de traiter, modifier, recomposer, etc.. ces éléments en leur conservant le statut d'événements gestuels, et enfin de les recomposer en "gestes" complexes.

Le traitement du geste revient alors à :

- 1 - déterminer des critères pour décomposer le geste instrumental en entités élémentaires et caractériser les attributs de ces entités. Les entités de la décomposition doivent être des "événements gestuels" élémentaires et leurs attributs des caractéristiques quantitatives, telles que des modifications de leur valeur conservent à l'élément son statut d'événement gestuel,

2 - définir des représentations graphiques des éléments décomposés et de leur articulation telles que le compositeur puisse percevoir et interpréter de manière directe les différentes valeurs de leurs attributs, ainsi que l'organisation des événements,

3 - déterminer des possibilités de traitement et de transformation utilisant ces représentations. Le traitement proprement dit est une intervention effective permettant d'une part de modifier les attributs et d'autre part de combiner les entités résultant de la décomposition précédente,

4 - déterminer et mettre en œuvre des modèles de composition du geste. La construction de structures élaborées peut être manuelle, à l'aide de fonctions de traitement de base. Mais elle peut également être prise en charge par des "modèles", c'est-à-dire des règles préalablement définies par le créateur, et opérant à un niveau élaboré, sur les éléments de la représentation et à partir des primitives de traitement..

L'éditeur de geste présenté ici fait la synthèse de ces fonctionnalités.

#### les approches envisageables

Une approche complètement générale de l'édition du geste devrait envisager trois démarches complémentaires : l'une neutre, l'autre faisant référence à l'action, la dernière à la perception.

1- approche neutre : les signaux captés sont les signaux de commande du processus numérique qui réalise la synthèse (à partir de simulations d'objets physiques ou de modèles fonctionnels plus généraux). A ce titre, ils présentent en eux-mêmes une certaine pertinence, indépendamment de leur signification gestuelle. Ils peuvent être traités et étudiés pour eux-mêmes, sans référence sensori-motrice particulière. Leur traitement ne doit veiller qu'à conserver leur cohérence de signaux de contrôle.

2- référence à l'action : il faut caractériser l'"espace gestuel" en lui-même en fonction de la manière dont l'opérateur le conçoit et se le représente. D'abord, on peut supposer qu'il existe pour l'opérateur une manière de se représenter ses comportements gestuels, et une sorte d'espace gestuel subjectif avec, comme pour l'espace sonore, des attributs. Ensuite, pour le son, on sait qu'il n'y a pas a priori d'isomorphie entre les paramètres perceptuels et les paramètres physiques



(exemple : différence entre hauteur et fréquence). Il en va de même pour les gestes : les variables physiques (force, déplacement, ...), même parfaitement captées, donnant une "mesure" objective du geste intrinsèque, ne sont pas directement explicites quant aux attributs du geste. Enfin, il est possible, du point de vue des représentations que l'on peut en avoir, de parler du geste sans mettre en jeu un objet (physique) auquel il s'applique.

3- référence à la perception : on peut s'intéresser aux conséquences perceptuelles (acoustique et/ou visuelle) des signaux de contrôle et souhaiter maîtriser celles-ci directement. Cela suppose que l'instrument soit défini, de façon à ce qu'il y ait une bijection entre les commandes et les résultats en fin de chaîne. De plus, la représentation et les traitements des signaux devraient faire appel à une connaissance de cet instrument, de façon à établir le lien entre leur caractéristiques et les attributs perceptuels en cause.

## **VII.2. notre façon de traiter ce problème**

Nous restreindrons l'étude aux situations possibles dans le cadre du système de simulation CORDIS actuellement opérationnel et du clavier rétroactif modulaire. Le geste instrumental au clavier aura donc valeur d'exemple.

Compte tenue de la finalité "opérative" de ce travail, nous nous efforcerons d'aborder le problème de l'édition comme un tout, et tenterons donc de maintenir un équilibre entre les différents problèmes qu'elle soulève (décomposition, représentation et traitement). Le principe consistera alors à sédimer progressivement des couches de solutions plutôt que d'approfondir certains aspects particuliers au dépens de la progression de l'ensemble.

L'approche neutre constituera dans cette idée la première couche à laquelle succédera l'approche en référence à la cause.

### **1. les signaux gestuels**

Les signaux que l'on va traiter ne sont jamais que les "signaux de commande" de l'instrument. C'est un parti pris que de les traiter avec une référence au geste instrumental.

Un problème de principe tient donc au fait que l'on va effectuer ces traitements sur l'information dont on dispose en faisant référence à une information dont on ne dispose pas intégralement.

Les signaux sont captés et mémorisés à partir d'un clavier rétroactif modulaire de 16 touches. Chaque signal est une suite d'échantillons numériques, traduisant le comportement mécanique d'une touche. Le déplacement du point de manipulation est de  $CP = \pm 15$  mm. Le rayon au point de manipulation est de 200mm. L'épaisseur d'une tranche est de 13.7 mm (la tranche est constituée d'un moteur, d'un capteur et d'un bras de touche).

Le signal est défini comme suit :

$$S : N \rightarrow Df$$

$$\text{avec } Df = [-CP .. +CP]$$

course en position du TGR :  $\pm CP$

codage de la position sur 12 bits.

La fréquence d'échantillonnage est connue et identique pour tous les signaux dans le cadre d'une expérience donnée. Elle est comprise entre 300Hz et 1kHz.

## 2. structuration des variables gestuelles

L'instrument présente un certain nombre de points de contact mono ou multi dimensionnels, prédéfinis et invariants pour l'expérience considérée. A chaque dimension d'un point de contact est associé un capteur de position, représenté par une variable gestuelle.

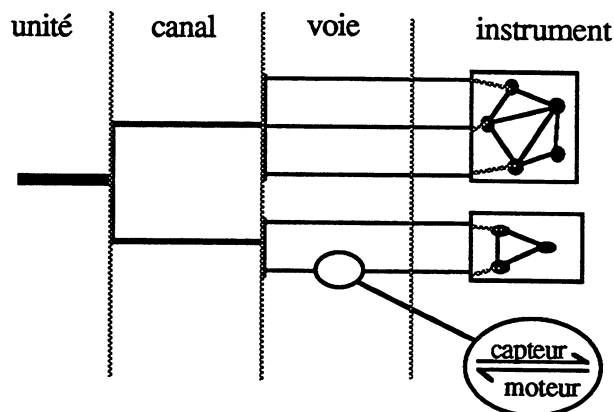


Figure 1. - *Structuration logique*. Les informations provenant des instruments sont structurées en voies, canaux et unités.

Pour structurer logiquement ces variables, et rendre compte au mieux de l'instrument et plus particulièrement des transducteurs gestuels, nous faisons apparaître une structure hiérarchisée à trois niveaux : la voie, le canal et l'unité.

On attache à chaque double ligne capteur/moteur du transducteur gestuel une voie, dédoublée en une voie d'entrée (capteur) et une voie de sortie (moteur). Ces voies sont structurées entre elles en canaux qui mettent en évidence la dimension du point de contact. Enfin, les canaux sont regroupés en des unités, de manière également prédéfinie. A l'intérieur d'une unité, les canaux sont de même nature (même dimension et même catégorie), mais il peut s'y ajouter des canaux particuliers pour le contrôle des paramètres.

## Segmentation du geste instrumental

---

### I. INTRODUCTION

### II. UN EXEMPLE DE REFERENCE - LA PAROLE

- II.1. critères de segmentation - II.2. méthodes de segmentation -
- II.3. en résumé

### III. SEGMENTS GESTUELS ET CLAVIER

#### III.1. événement gestuel

- 1. les trois phases - 2. geste balistique -
- 3. geste lent - 4. remarque

#### III.2. syntaxe élémentaire

- 1. segments gestuels - 2. grammaire élémentaire

#### III.3. la phase d'activité

- 1. segment-actif simple
  - a. segment et événement gestuel - b. syntaxe
  - c. les trois phases
- 2. segments-actifs consécutifs

#### III.4. le segment-pause

- 1. état de repos et point de contact
  - a. repos contraint - b. repos naturel
- 2. définition du segment-pause

### IV. SEGMENTATION MANUELLE

### V. SEGMENTATION AUTOMATIQUE

#### V.1. algorithme de segmentation élémentaire

- 1. critère de pause
- 2. gestion informatique des segments-actifs
- 3. identification des segments
  - a. grammaire élémentaire réduite -
  - b. processus d'identification
- 4. position de repos naturel

V.2. algorithmes de sursegmentation

1. segment-actif simple

a. critère d'extremum global - b. algorithme

2. segments-actifs consécutifs

a. critère d'amortissement - b. algorithme

V.3 conclusion

---

## **I. INTRODUCTION**

Nous introduisons dans ce chapitre le problème de la décomposition du geste instrumental. Il s'agit de déterminer des critères pour le décomposer en entités élémentaires et caractériser les attributs de ces entités.

Sur le clavier rétroactif modulaire (op.cit.), l'information gestuelle captée est explicitement décomposée en N canaux mono-voie et donc N signaux amplitude/temps.

Le problème que nous soulevons alors est un problème de segmentation, qui consiste à fragmenter chaque signal en "segments gestuels" suivant un critère donné, en présupposant que ce signal est effectivement structuré et assimilable à une simple juxtaposition séquentielle de segments.

## **II. UN EXEMPLE DE REFERENCE - LA PAROLE**

Le problème de la segmentation n'est pas nouveau.

Dans le domaine de la reconnaissance automatique de la parole par exemple, de nombreuses solutions ont été proposées à ce problème, aussi bien pour délimiter et étiqueter des intervalles spécifiques associés à des portions de sons connus, que pour localiser et identifier des unités phonétiques (phonème, syllabes, phones, etc.) en situation de reconnaissance.

Nous allons considérer cet exemple mono-dimensionnel en guise d'introduction à la segmentation du geste instrumental.

### **II.1. critères de segmentation**

Les critères utilisés pour découper le signal de la parole sont très variés et dépendent en général des représentations du signal choisies. Ils peuvent être regroupés selon P.F. Marteau [Mar88] en deux classes :

- les critères indépendants de processus de reconnaissance, qui sont basés sur les propriétés statistiques du signal. La détection de non stationnarité est interprétée comme une frontière entre deux segments.

- les critères dépendants de processus de reconnaissance sont basés sur une mesure de ressemblance entre des modèles de segments, ou une suite de ces modèles, et une séquence du signal à segmenter. Les modèles sont des prototypes de segment, c'est-

à-dire des références qui représentent l'ensemble des segments possibles. La segmentation est alors induite par le processus de reconnaissance et s'effectue simultanément à celui-ci.

## **II.2. méthodes de segmentation**

Parmi les méthodes utilisant des critères indépendants des processus de reconnaissance, il faut distinguer trois approches :

- 1/ la détection de discontinuités : ces méthodes opèrent sur des représentations paramétriques du signal et se différencient les unes des autres par le critère d'instabilité et par l'importance donnée au contexte lors de leur évaluation,
- 2/ la détection de ruptures : le signal est ici appréhendé comme la juxtaposition de segments homogènes d'un point de vue statistique. La segmentation consiste donc à retrouver ces instants de juxtaposition (rupture ou interface entre deux segments) à partir de critères statistiques,
- 3/ à partir de la courbe d'énergie : ce type de segmentation est basé sur l'analyse des extrema de la courbe d'amplitude du signal. Le principe des algorithmes développés repose sur la détection de formes respectant certaines contraintes. Les contraintes sont basées sur des critères de durée et d'amplitude entre les pics et les vallées constituant la forme.

Les méthodes utilisant des critères dépendants des processus de reconnaissance peuvent se répartir selon D.Vicard et L.Miclet en trois catégories [VM88] :

- 1/ méthodes directes : une portion du signal inconnu est directement comparée, grâce à une mesure de distorsion, à des formes contenues dans un dictionnaire,
- 2/ méthodes basées sur les modèles markoviens : on essaie d'identifier un modèle, en évaluant la probabilité que la séquence d'événements observés en soit produite,
- 3/ méthodes par règle : les déductions multiples du phonéticien et les stratégies qu'il met en œuvre sont transcrites sous forme de règles logiques. Les propriétés acoustiques sont alors des faits qui dirigent les processus algorithmiques dont la finalité (appelée but) est la confirmation ou l'infirmité des hypothèses.

## **II.3. en résumé**

La segmentation est un problème central qui a suscité de nombreuses études notamment dans le domaine de la parole. Les méthodes conduisent toutes à des solutions différentes, conditionnées par le domaine d'application.

Selon P.F.Marteau [Mar88], les méthodes qui ne font pas appel à un processus de reconnaissance ont le mérite d'être simples et efficaces d'un point de vue "temps réel". Mais les notions de parties stables et de parties instables du signal, bien souvent associées à des seuils, nécessitent un ajustement en fonction du locuteur et du niveau d'enregistrement.

D'autre part, les méthodes qui nécessitent une certaine reconnaissance, un certain apprentissage, - soit sous forme d'expertise, soit sous forme de modèles -, aboutissent à une segmentation que l'on peut considérer comme le sous produit d'un processus de reconnaissance. Mais l'acquisition des connaissances indispensables à l'apprentissage est très coûteuse en temps et place mémoire, surtout dans un contexte multi-locuteur.

L'exemple de la segmentation du signal de la parole est intéressant à considérer dans la mesure où il y aura probablement des similitudes et nécessairement des différences avec notre problème de la segmentation du geste instrumental.



### **III. SEGMENTS GESTUELS ET CLAVIER**

#### **III.1. événement gestuel**

On considère a priori qu'une séquence gestuelle est décomposable en un ensemble d'entités élémentaires que l'on appellera "événements gestuels".

Sur les instruments à clavier, un événement gestuel consiste à manipuler une touche. Il peut être décrit selon plusieurs critères. C'est habituellement un geste digital de faible amplitude et en trajectoire verticale. Sa vitesse et son amplitude sont suffisantes pour décrire la dynamique et l'articulation musicale du son qu'il permet de produire.

#### **1. les trois phases**

J.Haury [Hau87] souligne l'importance des trois phases du va-et-vient des touches dans les mécanismes producteurs de son (enfoncement, maintien, relâchement). Il distingue pour insister sur ce point deux familles de claviers : ceux à sons soutenus et les autres.

Sur les instruments à "sons soutenus" (orgue électrique par exemple), le son débute et est maintenu dès lors que le contact est établi. Il s'éteint au moment où ce contact est rompu. L'attaque et l'extinction du son sont stéréotypées. Il n'y a pas de moyen de les modifier par le mouvement de la touche sélectionnée.

En revanche, sur les instruments à "sons non soutenus" (piano, ...) il est possible d'intervenir sur les transitoires d'attaque et dans certains cas sur les transitoires d'extinction du son. Généralement, la vitesse d'enfoncement de la touche détermine l'intensité du son ; le moment du relâchement de la touche permet éventuellement d'intervenir sur l'amortissement du son.

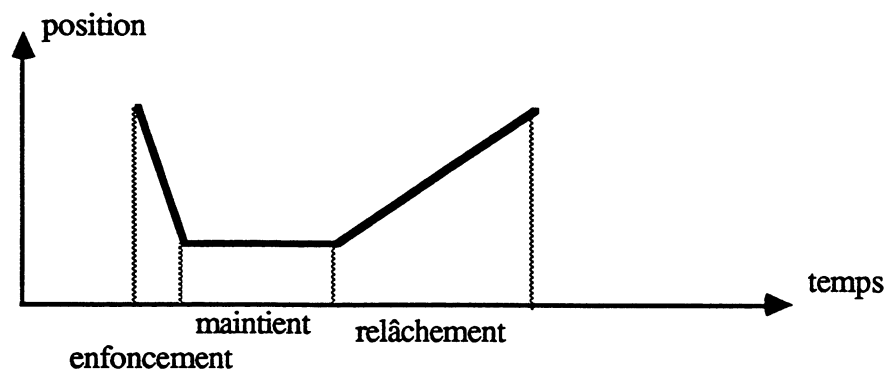


Figure 1. *Événement gestuel*. L'événement gestuel sur un clavier consiste à enfoncer la touche, éventuellement la maintenir, et enfin la relâcher.

On peut observer cette forme dans des circonstances très variées mais ses attributs restent déterminés, pour l'essentiel, par l'amplitude et la vitesse d'exécution du geste.

### **2. geste balistique**

Le geste balistique résulte d'une brève contraction initiale d'un muscle, suivie de sa relaxation. Le mouvement est stoppé mécaniquement.

En toute rigueur, le mouvement balistique exclut tout arrêt intentionnel et donc tout freinage volontaire. Par extension, la notion de phase balistique du mouvement a été introduite pour décrire la portion initiale de projection du membre, précédant la phase de freinage et d'ajustement terminal.

Une phase préparatoire permet habituellement de positionner le membre et définir, en fonction de la cible, les paramètres de vitesse, de direction et d'amplitude.

Selon D. Beaubaton [Beau83], la phase balistique a une durée inférieure à quelques dizaines de msec. Cette durée correspond à un seuil en-dessous duquel aucun contrôle n'est, semble-t-il, plus possible.

### **3. geste lent**

L'exécution de certains mouvements nécessite un freinage actif, ou plus encore, un ajustement précis sur un objectif.

A l'inverse du geste balistique, les informations issues des récepteurs tactiles, visuels et autres, interviennent pendant le mouvement dans les processus de contrôle moteur. Cette prise en compte d'informations d'origine périphérique ajoute un temps supplémentaire au temps de "programmation" des différents muscles. On parle donc de geste lent.

On peut considérer comme cas singulier de mouvement lent celui qui consiste à garder la partie manipulée dans une même position. Il s'agit par exemple du maintien d'une touche d'orgue en butée.

Un tel geste ne sert pas à déplacer ni à déformer l'instrument, mais au contraire à conserver une certaine configuration physique. Il suppose un temps de contact durable et met à contribution une zone sensorielle large. La perception par le toucher a une importance prépondérante pour la stabilité de l'objet.

**4. remarque**

L'événement gestuel défini par la succession enfoncement-[maintien-] relâchement, bien qu'à la base du jeu pianistique, ne représente qu'une partie des cas possibles de maniement d'une touche.

Un cas plus général serait celui construit sur la base de deux événements terminaux A et B:

S ::= AS | BS | A | B  
 A ::= enfoncement [-maintien]  
 B ::= relâchement [-maintien]

Il serait alors possible de décrire les notes liées sur un violon, le double échappement, ou l'"aftertouch" sur un clavier de synthétiseur. Par exemple, la séquence S=Enfoncement-Maintien-Enfoncement-Relâchement se traduirait par S=A.A.B.

**III.2. syntaxe élémentaire****1. segments gestuels**

Nous faisons une distinction entre "événement gestuel" et "segment gestuel" pour séparer respectivement la référence gestuelle et une représentation pratique de cette référence. L'événement gestuel correspond au mouvement intrinsèque de l'instrumentiste, alors que le segment-gestuel est une portion des signaux captés et mémorisés sur un canal donné.

Chaque segment est constitué d'un certain nombre de signaux, en fonction de la dimension du canal dont il est extrait. Par exemple, les segments issus des canaux bi-dimensionnels (par exemple un "joystick") seront constitués de deux signaux, etc... Pour un segment donné, la longueur des signaux est identique.

Les signaux gestuels décrivent le comportement mécanique des points de contact de l'instrument et traduisent implicitement des phases de couplage homme / instrument et des phases où l'instrument évolue librement (absence de contact).

En toute rigueur, il faudrait distinguer dans le signal observé des phases pendant lesquelles il y a effectivement contact entre l'opérateur et l'instrument, de celles où l'instrument évolue librement. De là pourrait se déduire un critère de définition des segments.

Mais en fait cette séparation n'est pas observable de façon simple et immédiate sur le signal. En conséquence de quoi dans un premier temps, nous adoptons une définition des segments qui correspond à l'activité et la non-activité de l'instrument lui-même.

On parlera de "segment-actif" et de "segment-pause". Le segment-pause décrit une phase de repos naturel de l'instrument. Le segment-actif représente une phase d'activité de l'instrument, dans laquelle apparaît au moins une phase de contact opérateur / instrument.

## 2. grammaire élémentaire

La grammaire élémentaire présente les règles d'agencements simples des segments-actifs et des segments-pauses sur un canal donné. Elle met en évidence deux caractéristiques importantes du comportement d'un point de contact de l'instrument :

- 1 - une pause ne correspond à aucune activité gestuelle significative pour l'instrument et n'est donc pas décomposable en sous-pauses,
- 2 - l'activité du point de contact est décomposable en une succession de segments-actifs, lesquels font référence à des événements gestuels.

Les signaux gestuels débutent soit par un segment-pause, soit par un segment-actif. Si la pause n'est pas le dernier segment du signal, elle peut être suivie uniquement par un segment-actif. En revanche, si l'actif n'est pas le dernier segment du signal, il peut être suivi, soit par un autre segment-actif, soit par une pause.

Les symboles terminaux sont indiqués en gras : **A** désigne un segment-actif et **P** un segment-pause.

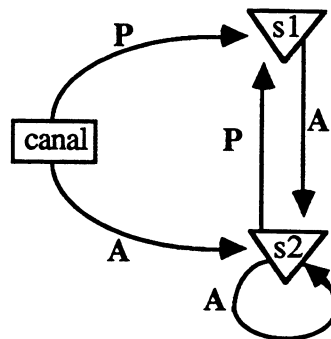


Figure 2. *Automate*. canal est l'état initial, s1 et s2 sont les états finals. L'arc joignant deux états indique le symbole permettant de passer de l'un à l'autre.

La grammaire correspondante s'écrit :

(G1) canal ::= **P.s1** | **P**  
 canal ::= **A.s2** | **A**  
 s1 ::= **A.s2** | **A**  
 s2 ::= **A.s2** | **P.s1** | **P**

### III.3. la phase d'activité

Nous avons décrit un événement gestuel comme une succession de plusieurs phases : une phase d'enfoncement, une phase de maintien et une phase de relâchement. Cet événement, je le rappelle, doit être vu comme le représentant des gestes percussifs : il ne tient pas compte par exemple du double échappement ou de l'"aftertouch".

La portion des signaux gestuels représentant le comportement de la touche durant cet événement est appelée "segment-actif". Lorsque plusieurs événements gestuels se succèdent, les signaux correspondant contiennent une succession de tels segments.

#### 1. segment-actif simple

##### **a. segment et événement gestuel**

Le segment-actif simple comporte, de façon similaire à l'événement gestuel, trois phases. Une phase d'enfoncement (notée E), éventuellement une phase de maintien (M), et enfin une phase d'amortissement (A).

Les phases d'enfoncement et de maintien sont des phases de couplage musicien-touche. Pour l'amortissement, on peut distinguer deux cas limites : soit le musicien et la touche sont en couplage, soit ils ne le sont pas.

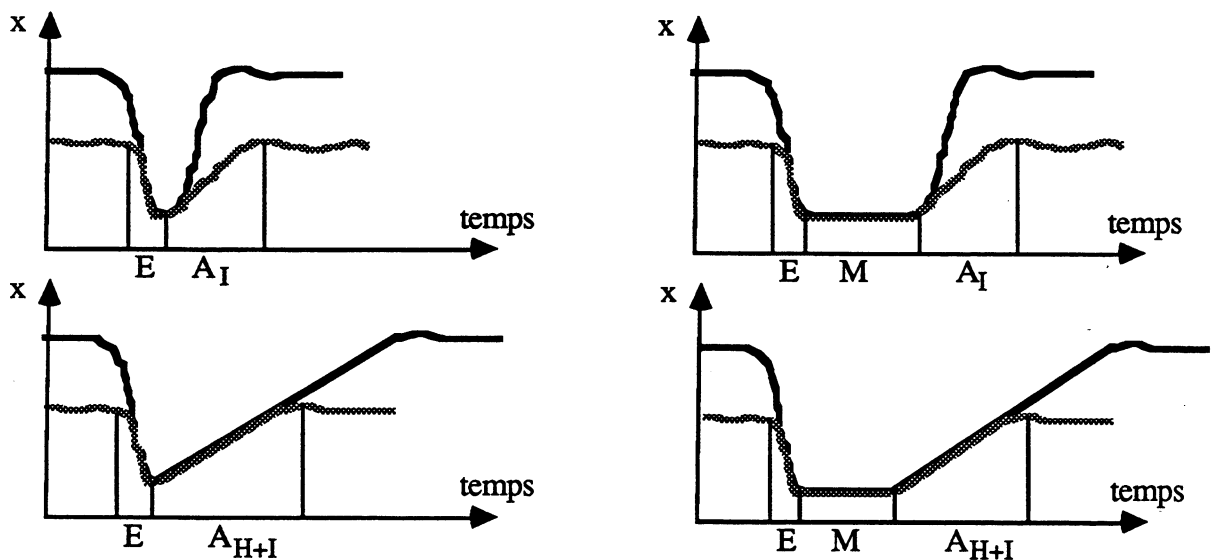


Figure 5. *Segments-actifs*. En traits gras sont représentées les évolutions du point de contact-opérateur lors d'un événement gestuel, en grisés, les évolutions du point de contact-instrument. Il faut remarquer deux cas limites : l'amortissement se réalise en contact avec l'opérateur ( $A_{H+I}$ ) ou librement ( $A_I$ ).

Un geste percussif (sans phase de maintien), donne naissance à un segment-actif particulier. Il est constitué de deux phases seulement : une phase d'enfoncement et une phase d'amortissement.

### b. syntaxe

Nous décrivons les différentes phases du segment-actif selon le couplage instrument / opérateur. On se place du point de vue de l'instrument.

Dans la syntaxe suivante, **E** désigne la phase d'enfoncement, **M** la phase de maintien, **A** la phase d'amortissement.

Pour l'amortissement, l'indice H+I indique qu'il y a couplage permanent. C'est le cas lorsque la vitesse de relâchement est inférieure à la vitesse d'amortissement de l'instrument. D'autre part, l'indice I indique que la touche est en amortissement libre, sans couplage. C'est le cas lorsque la vitesse de relâchement est supérieure à la vitesse d'amortissement de la touche.

$$\text{Seg-Actif} ::= \mathbf{E} \cdot \mathbf{A}_{\mathbf{H+I}} \quad | \quad \mathbf{E} \cdot \mathbf{A}_{\mathbf{I}} \quad |$$

$$\mathbf{E} \cdot \mathbf{M} \cdot \mathbf{A}_{\mathbf{H+I}} \quad | \quad \mathbf{E} \cdot \mathbf{M} \cdot \mathbf{A}_{\mathbf{I}}$$

### c. les trois phases

Dans la situation réelle, les trois phases de maniement d'une touche ne sont pas de simples segments de droite, ni même des portions de signaux ayant chacun un sens d'évolution constant. L'opérateur et l'instrument introduisent tous deux des oscillations dans le système, localisées en fin de phase d'enfoncement et d'amortissement, comme l'illustre la figure suivante.

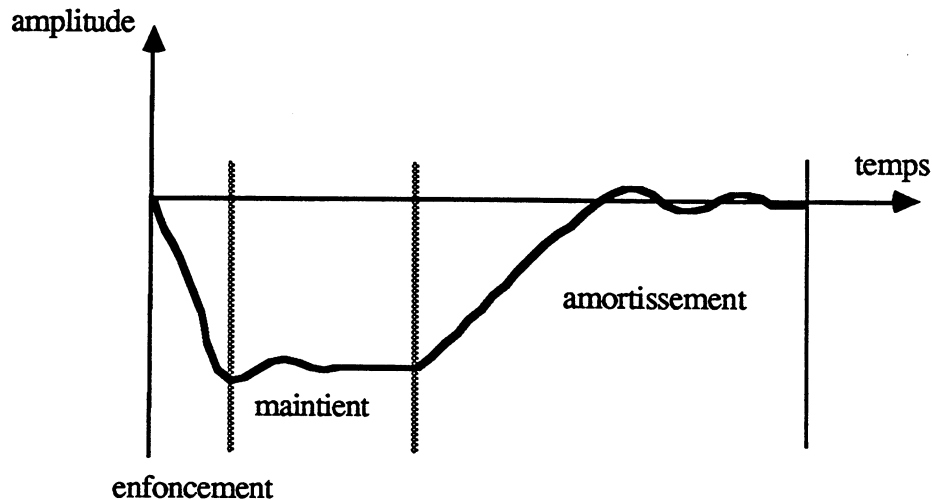


Figure 6. *Traitement des oscillations.* La figure donne un exemple de segment-actif avec en fin de phase d'enfoncement et d'amortissement de légères oscillations.

La phase d'enfoncement peut être lente ou balistique. Si l'on oriente le sens positif du bas vers le haut, la vitesse est négative (contraire de l'amortissement). En fin de course

apparaît une légère oscillation autour d'une position visée. Cette oscillation est essentiellement déterminée par l'ajustement terminal du mouvement de l'opérateur.

La phase de maintien implique un couplage homme / instrument. La vitesse du point de contact est nulle durant cette phase et les signaux correspondant sont stationnaires. La position de repos (contraint) se situe soit en butée, soit ailleurs. On considère que du point de vue de l'activité gestuelle, le maintien débute à partir du premier extremum reconstré dans le segment.

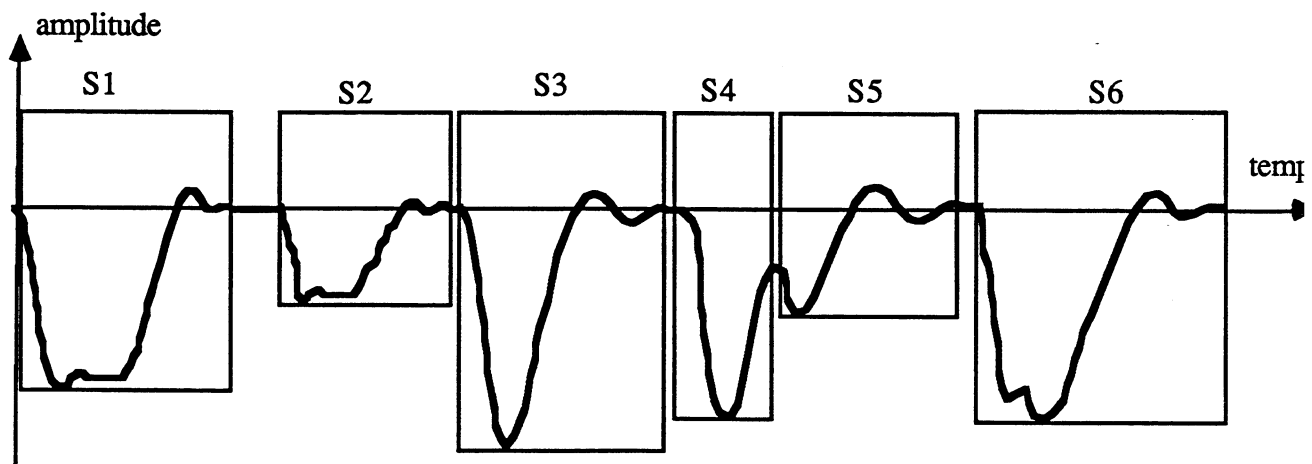
La phase d'amortissement peut être lente ou balistique. Si l'on oriente le sens positif du bas vers le haut, la vitesse est positive. En fin de course apparaît une légère oscillation autour de la position de repos naturel. Cette oscillation est déterminée par la force de rappel de l'instrument.

### 2. segments-actifs consécutifs

Lors d'un jeu instrumental, les événements gestuels se succèdent sur les touches du clavier selon un certain rythme, une certaine polyphonie, etc.

Lorsque deux ou plusieurs événements gestuels se suivent sur une même touche, il arrive que les segments-actifs correspondants ne soient pas séparés par une pause.

La figure suivante (figure 7) présente une phase d'activité constituée de six segments-actifs successifs sur une même touche, dont la position de repos naturel est unique. Elle décrit quelques situations de transitions types. Cet ensemble, loin d'être exhaustif, servira néanmoins de référence aux traitements ultérieurs.



S1 et S2 sont des segments-actifs de la forme E.M.A (enfoncement-maintien-amortissement) alors que S3, S4, S5 et S6 sont des percussions décrites par la chaîne E.A (enfoncement-amortissement).

S1-S2 : une petite pause sépare les deux segments. Sa durée, inférieure au seuil de pause, ne la rend pas détectable en tant que pause lors de la segmentation élémentaire. Les bornes de cette courte pause déterminent la fin de S1 et le début de S2. Idem pour la transition entre S5 et S6.

S2-S3 : le signal n'est plus fixe après la phase d'amortissement de S2, mais un changement de pente important en début de S3 indique qu'il y a apparition d'un nouveau segment. Idem pour la transition entre S3-S4.

S4-S5 : la phase d'amortissement de la courbe ne passe pas par la position de repos naturel. Une nouvelle percussion a lieu et un nouveau segment apparaît à partir de l'extremum correspondant.

S6 : S6 a une double bosse en fin d'enfoncement, liée très certainement à une action articulaire complexe. Il faut remarquer que :

- la durée séparant les deux extrema est faible. De l'ordre de .02sec
- les rebonds se font près des positions limites de l'instrument (butée).

En résumé, le double rebond ressemble à une oscillation de faible amplitude, de faible période et proche des butées. Il est difficile de la confondre avec une nouvelle percussion.

### III.4. le segment-pause

#### 1. état de repos et point de contact

Un point de contact est dans un "état de repos" à l'instant T s'il conserve la même position entre l'instant T et l'instant T+ $\Delta T$ .

L'intervalle de temps  $\Delta T$  doit permettre de distinguer une phase de repos stable d'une phase de repos instable. Cette situation a lieu par exemple lors d'un changement de sens d'une touche.

Un instrument est dit dans un "état de repos" si l'ensemble de ses points de contact sont tous dans un état stable.

De plus, selon le couplage avec l'opérateur, nous distinguons deux types de repos : le repos naturel et le repos contraint.



### a. repos contraint

Le repos est dit "contraint" si le point de contact est couplé avec l'opérateur.

Diverses situations de contraintes sont envisageables.

Par exemple, celle du violoniste qui maintien son archet : il oppose aux forces en présence (pesanteur, rappel, ...) d'autres forces qui maintiennent le point de contact dans une même position. Pour autre exemple, la situation du pianiste qui fait intervenir les butées mécaniques de son instrument. Il exerce sur la touche une force afin de la bloquer en bout de course.

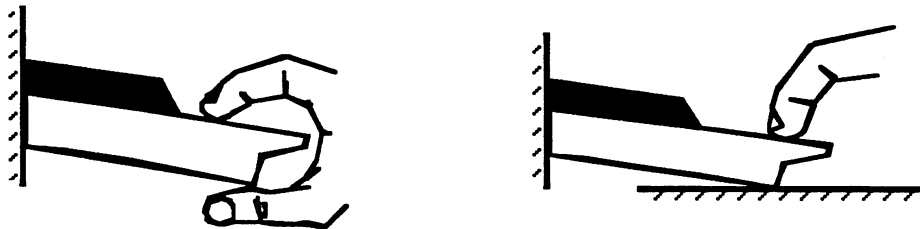


Figure 3. *Repos contraint*. On distingue deux cas : celui où l'instrument est au repos contraint uniquement par le musicien et celui où interviennent les butées mécaniques.

### b. repos naturel

L'état de repos est dit "naturel" si le point de contact est isolé mécaniquement de l'opérateur.

Dans le cas du piano, cette position est unique. Mais plus généralement, le nombre de positions de repos naturel varie selon le point de contact considéré. Nous retenons trois cas importants : celui où la position est unique, celui où elle est multiple et enfin celui où aucune position de repos naturel particulière n'est prédéfinie.

1 - position de repos unique : quelque soit la force extérieure appliquée à un point (dans les limites de ruptures de l'objet), il existe une force de rappel qui le ramène invariablement à une même position. C'est le cas d'une touche de piano.

2 - position de repos multiple : quelque soit la force extérieure appliquée à un point de contact, il existe une force de rappel qui le ramène à une position prédéfinie, mais non unique. L'ensemble des positions de repos est fini pour le point considéré. C'est le cas d'un interrupteur électrique qui présente deux positions de repos naturel.

3 - position de repos indéterminée : il n'existe pas de force de rappel. Le point conserve la position qu'il avait juste avant d'être lâché. C'est le cas d'un potentiomètre

à déplacement continu : toute position du périmètre balayé peut devenir une position de repos naturel.

## 2. définition du segment-pause

Dans la mesure où nous souhaitons conserver toute l'information faisant référence à une activité gestuelle, nous devons tenir compte de la distinction entre un repos naturel et un repos contraint.

Lors d'un repos contraint, il y a une intervention gestuelle, même si celle-ci ne se traduit pas par le déplacement du point considéré. Un effort est fourni par l'opérateur pour maintenir une touche de piano en butée. Ce geste a un sens du point de vue de l'instrument. En revanche, lors d'un repos naturel, l'opérateur n'est pas en contact avec l'instrument et ses mouvements n'ont pas d'incidence sur lui.

Un segment-pause décrit donc une phase de repos naturel d'un point de contact de l'instrument.

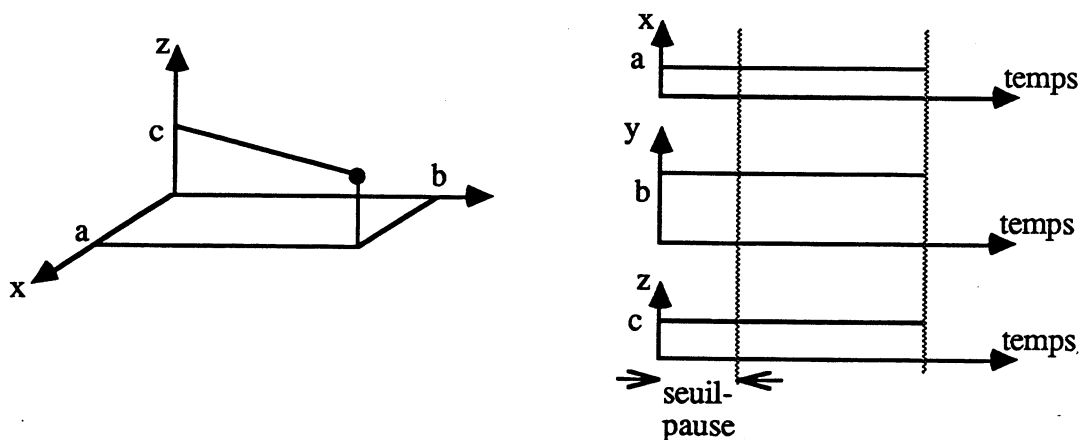


Figure 4. *Le segment-pause*. Le segment pause décrit une phase de repos naturel d'un point de contact. La durée de cette phase est supérieure à un seuil donné. A chaque dimension correspond un signal stationnaire dont l'amplitude est donnée par les coordonnées de la position de repos. Ici, a, b et c.

## **IV. SEGMENTATION MANUELLE**

La segmentation manuelle d'un signal continu est l'opération qui consiste en une fragmentation par l'opérateur lui-même. Ce processus repose sur l'existence d'un outil interactif qui permet à l'opérateur d'observer les signaux gestuels, de sélectionner des parties de ces signaux et de les mémoriser. Les caractéristiques de cet outil se décrivent selon trois aspects :

- le type et la forme de la représentation des signaux :
  - \* le type : graphique, lexicographique, ...
  - \* la forme : l'objet lui-même, transformée de Fourier, un équivalent paramétrique, ....
  
- les moyens pour observer et sélectionner interactivement une partie ou l'intégralité des représentations :
  - \* pour observer : agrandissement, focalisation, transformation ....
  - \* pour sélectionner : un curseur, un pointeur,....
  
- les fonctions de base pour la gestion des segments
  - \* création d'un segment à partir d'une portion de l'objet sélectionné,
  - \* suppression d'un segment existant,
  - \* archivage et restitution des segments.

La segmentation manuelle des signaux gestuels est une fonction essentielle de l'outil de création que nous développons.

Elle permet à l'utilisateur de se familiariser avec les représentations objectives du geste instrumental et d'apprendre à traduire ces signaux en des actions gestuelles, et inversement.

Ce processus d'analyse se réalise dans une boucle où l'opérateur observe un objet dont il connaît la provenance, extrait de cet objet des sous-parties pour éventuellement le faire renaître indépendamment de l'objet initial.

La fonction première de la segmentation manuelle est de décomposer l'information initiale en un ensemble d'éléments pour le traitement. Les frontières définissant ces éléments sont entièrement déterminées par l'opérateur. Il n'est contraint que par les représentations et les possibilités d'observation et de sélection de ces représentations.

Une autre de ses fonctions est de proposer des éléments de comparaison. En effet, elle permet de former des experts en matière de segmentation des signaux gestuels. L'expert est un utilisateur qui a, par la pratique et l'expérience, acquis une grande habilité pour délimiter des "segments" sur un signal. Ses choix deviennent la référence.

Il s'agit alors pour le concepteur d'un processus de segmentation automatique, de définir des critères capables d'identifier automatiquement, toutes ou en partie, les frontières de segmentation posées par l'expert.

Cependant, et même lorsque des processus de segmentation automatique ont été définis, la fonction manuelle reste essentielle. Elle permet précisément de ne pas se laisser enfermer dans une expertise et de s'échapper des processus automatiques existants. Cette fonction constitue donc bien une fonction de base de la segmentation.

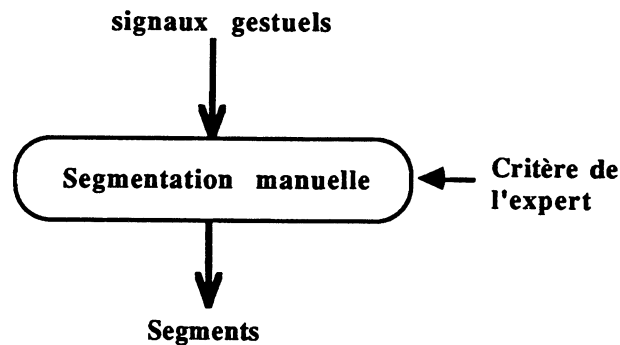


Figure 8. Segmentation manuelle. Processus qui consiste à extraire des signaux gestuels un ensemble de segments. Les critères de segmentation sont décidés par l'utilisateur-expert.

## V. SEGMENTATION AUTOMATIQUE

Nous proposons dans ce paragraphe une méthode de segmentation automatique. Elle se réalise en deux étapes et opère sur des signaux gestuels issu d'un jeu instrumental sur le clavier rétroactif modulaire.

Dans un premier temps, le signal est décomposé en deux ensembles : d'un côté les phases d'activité et de l'autre les segments-pauses.

Dans un deuxième temps, les phases d'activité sont décomposées en une succession de segments-actifs et chacun de ces segments est étiqueté en fonction de la vitesse et de l'amplitude des différentes phases le constituant. Nous appelons cette deuxième étape "sursegmentation".

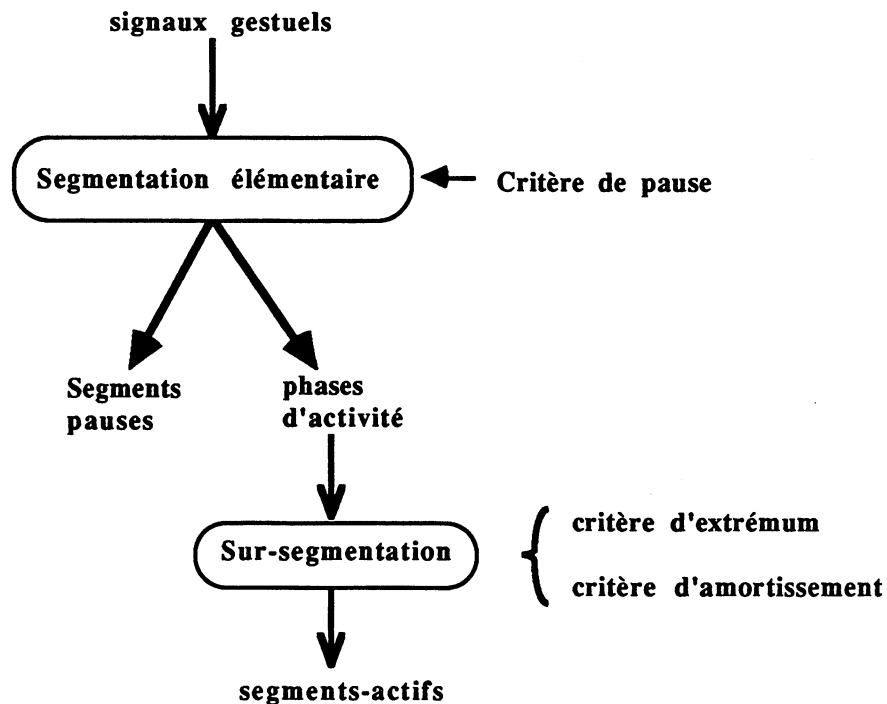


Figure 9. *Segmentation automatique*. Deux étapes successives sont envisagées pour extraire à partir des signaux gestuels les entités destinées au traitement et à la représentation.

Dans la mesure où il y a dépendance pour la segmentation de caractéristiques comportementales de l'instrument, on sera obligé de faire des hypothèses générales sur cet instrument et même plus, d'introduire une possibilité de modifier ces paramètres en fonction de l'instrument.

### V.1. algorithme de segmentation élémentaire

L'algorithme de segmentation élémentaire a pour fonction de séparer les phases d'activités et les phases de repos naturel représentées dans un signal de position noté S. Le critère de segmentation utilisé s'appuie sur les propriétés statistiques des signaux et consiste à identifier les phases de repos naturel.

L'algorithme de segmentation est écrit en langage C. Les hypothèses que nous faisons sur les données de l'algorithme sont les suivantes :

- le signal S est mémorisé, de longueur TMAX,
- il est supposé idéal : sans bruit, sans dérive, ....,
- la position de repos R est connue et unique,
- le seuil minimal de pause ( $\Delta P$ ) est prédéfini.

#### 1. critère de pause

D'après le paragraphe III.3.2, lors d'une phase de repos naturel, toutes les valeurs du signal durant cette phase sont identiques et égales à R.

On considère qu'un segment-pause est identifié dès lors que le signal conserve la position de repos R sur  $\Delta P$  échantillons.

La fonction PositRN(t) teste si la valeur du signal S à l'instant t est une position de repos naturel. La fonction CP(t) teste si le signal S traduit une pause à partir de l'instant t.

```
boolean CP(t) int t ;
/* ----- */
Critère de Pause : une pause est reconnue à
l'instant t si S(t) conserve la position de
repos naturel entre t et t+  $\Delta P$ .
optimisation : la boucle est interrompue dès
que le test devient faux.
/* ----- */
{ int i = 0, test = vrai ;
  while ( (i <  $\Delta P$ ) && test )
    { test = PositRN(i+t) ; i++; }
  return (test) ;
}
```

```
boolean PositRN(t) int t ;
/* ----- */
Comme la position de repos est unique, le
test consiste à comparer S(t) et R
/* ----- */
{ return ( S(t) == R ) ;
}
```

#### 2. gestion informatique des segments-actifs

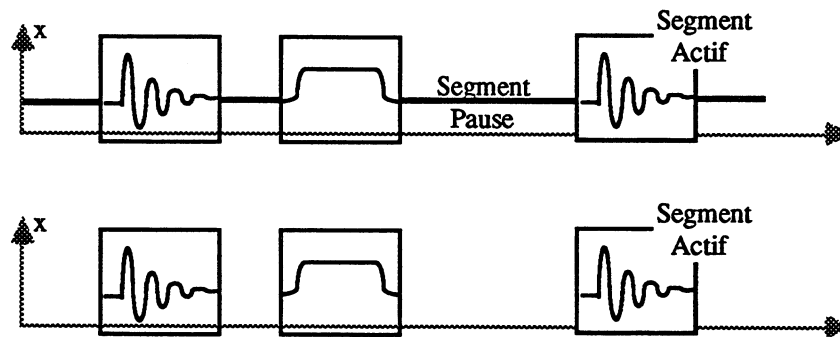
Nous définissons trois fonctions pour gérer les segments-actifs : l'une pour l'initialisation, la seconde pour la complétion et la troisième pour la mémorisation.

- l'initialisation d'un segment consiste en l'allocation d'une structure informatique et l'initialisation de ses paramètres.

- la complétion d'un segment a pour but d'ajouter en fin de segment une valeur du signal  $S(t)$ . Il n'est donc pas nécessaire de gérer explicitement des indices.
- la mémorisation d'un segment se charge de fermer le segment en calculant le cas échéant quelques paramètres, et l'ajoute à l'ensemble de segments déjà existants.

Pour optimiser l'espace mémoire utilisé, on peut envisager un codage de l'information (compression), soit lors de la complétion, soit lors de la mémorisation.

De plus, il n'est pas nécessaire de mémoriser les segments-pauses. Ceux-ci sont directement déductibles des segments-actifs. En effet, un segment-pause contient un signal stationnaire dont l'amplitude est déterminée soit par la dernière valeur du segment-actif précédant, soit par la première valeur du segment-actif suivant. Rappelons que la pause, repos naturel de l'instrument, ne doit pas être confondue avec le repos contraint, lequel, bien que représenté par un signal fixe, appartient au segment-actif.



### 3. identification des segments

#### **a. grammaire élémentaire réduite**

Nous basons l'extraction des phases d'activité sur la grammaire élémentaire  $G1$ , légèrement réduite. Elle décrit l'alternance de phases d'activité (**SA**) et de segments-pauses (**P**) sur un canal. Elle ne précise pas le contenu d'une phase d'activité.

$$\begin{aligned}
 (G2) \quad \text{canal} &::= && \mathbf{P.s1} \mid \mathbf{SA.s2} \mid \mathbf{P} \mid \mathbf{SA} \\
 s1 &::= && \mathbf{SA.s2} \mid \mathbf{SA} \\
 s2 &::= && \mathbf{P.s1} \mid \mathbf{P}
 \end{aligned}$$

Les chaînes possibles se présentent sous la forme  $\mathbf{P^*.(SA.P)^n.SA^*}$  avec  $n \geq 0$ . Le symbole  $X^*$  signifie soit  $X$  soit rien.

### b. processus d'identification

Pour identifier des chaînes de ce type dans les signaux gestuels, nous réalisons trois fonctions.

La première, `RFPause()`, recherche la fin du signal correspondant à une pause. Elle base son calcul sur la comparaison de la valeur du signal avec le repos naturel `R`.

La seconde fonction, `RActif()`, recherche la fin de la phase d'activité. Cette procédure initialise, complète et mémorise toute cette phase. La boucle principale se termine à l'identification d'une phase de repos.

La troisième, `IdentifSeg()`, traduit la grammaire réduite. La boucle principale attend une chaîne de la forme  $(SA.P)^n$ . La dernière pause n'est pas systématique. Dans la boucle, un test en tient compte. D'autre part, la chaîne peut commencer par une pause. Un premier test le vérifie.

```

int RActif(t)      int t;
/* ----- */
Recherche la fin de l'activité
/* ----- */
{ InitSeg() ;
  while ( (t<TMAX) && !CP(t) )
    { CompletSeg(t) ; t++ ; }
  MemoSeg() ;
  return (t) ;
}

int RFPause ( t ) int t ;
/* ----- */
Recherche la fin de la pause
/* ----- */
{ while ( (t<TMAX) && PositRN(t) ) t++ ;
  return (t) ;
}

void IdentifSeg()
/* ----- */
Identification de la séquence P*(SA.P)n.SA*
optimisation : la procédure RFPause() est
appelée pour l'instant t+ΔP car le critère de
pause CP() a déjà effectué le test de repos
naturel sur les ΔP premiers instants.
/* ----- */
{ int t=0 ;

  if ( CP(t) ) /* le premier segment est une pause */
    t = RFPause(t+ΔP) ;
  while ( t < TMAX ) /* identification de la chaîne (SA.P)n.P* */
    { t = RActif(t) ;
      if(t+ΔP<TMAX) t = RFPause(t+ΔP) ; }
}

```

### 4. position de repos naturel

Nous avons supposé que la position de repos `R` est connue et unique. En règle générale, ce n'est pas le cas.

Selon la connaissance que l'on a du type de repos (unique, multiple, non-prédéfini) et de la nature du couplage, le test de repos naturel `ReposRN()` introduit dans le paragraphe V.1.1 se complexifie.



La combinaison des paramètres de repos et de couplage donne lieu à quatre tests différents. Les algorithmes correspondant sont écrits en langage C.

cas 1 - pas d'information, ni sur le repos naturel, ni sur le couplage

```

ReposRN1(t) int t;
/* ----- */
Par défaut, toujours faux. On conserve toute l'information.
/* ----- */
{ return (faux) ; }
    
```

cas 2 - pas d'information sur le type de repos

```

ReposRN2(t) int t ;
/* ----- */
En l'absence de couplage, la fonction PenteNulle(t)
teste la stationnarité du signal à l'instant t.
/* ----- */
{ if (! couplage) return(PenteNulle(t)); else return(faux) ; }
    
```

cas 3 - pas d'information sur le couplage

```

ReposRN3(t) int t;
/* ----- */
Toujours faux sauf si le repos est unique ou multiple.
/* ----- */
{ if (repos_unique) return( TestReposUnique(t) ) ;
  else if (repos_multiple) return( TestReposMultiple(t) ) ;
  else return(faux) ;
}
    
```

cas 4 - l'algorithme connaît le type de repos et le type de couplage

```

ReposRN4(t) int t;
/* ----- */
Conditions idéales.
/* ----- */
{ if (repos_unique) return( TestReposUnique(t) ) ;
  else if (repos_multiple) return( TestReposMultiple(t) ) ;
  else if(couplage) return(faux);
  else return( TestReposNonPrédéfini(t) ) ;
}
    
```

remarque :

Lors d'un couplage permanent, tout repos est a priori contraint. Mais dans le cas où la position mesurée  $S(t)$  est à la position de repos naturel (R) de l'instrument, tout permet de supposer que l'opérateur voulait atteindre cette position. On considère donc que la position vérifie un repos naturel.

Cette hypothèse va dans le sens d'une simplification des cas de figure distingués par le type de couplage. Elle concerne le test `ReposRN4()`.

## V.2. algorithmes de sursegmentation

La phase d'activité que nous proposons de sursegmenter est constituée d'un ou de plusieurs segments-actifs.

Le premier problème consiste à séparer les phases d'enfoncement, d'amortissement, éventuellement de maintien dans un segment-actif simple. Le second problème consiste à séparer deux segments-actifs se succédant en considérant les quelques cas de figure proposés au paragraphe III.4.2.

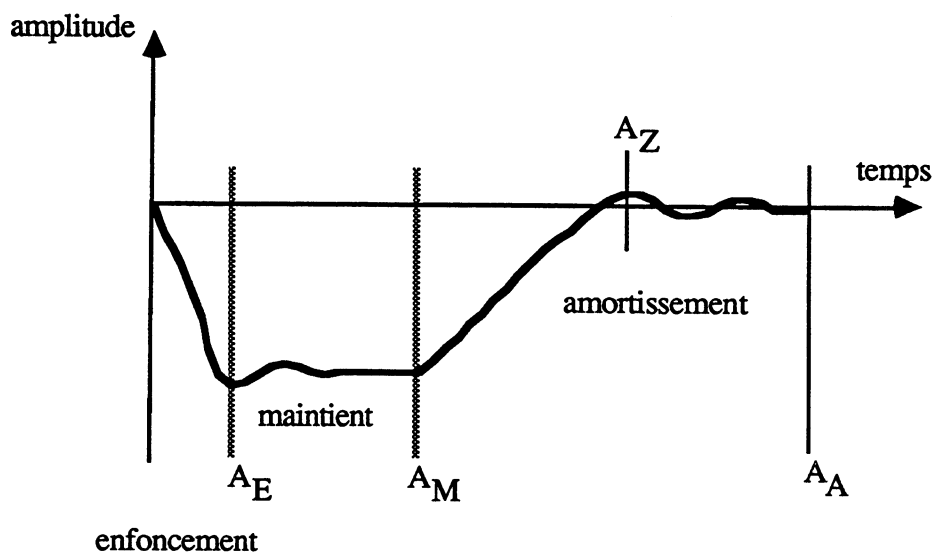
### 1. segment-actif simple

#### **a. critère d'extremum global**

Dans un segment-actif simple, la phase d'enfoncement est située entre le début du signal et le premier extremum rencontré ; la phase de maintien entre cet extremum et la fin de la stationnarité ; la phase d'amortissement entre ce point et la fin du signal.

De plus, le signal associé à chaque phase n'a pas un sens d'évolution constant. Il n'est pas monotone. Il comporte également des oscillations. Pour l'enfoncement, cette oscillation est attribuée au mouvement d'ajustement terminal. Pour la phase d'amortissement, elle coïncide avec l'amortissement libre de l'instrument.

La méthode de sursegmentation que nous proposons repose sur l'identification de variations significatives du sens de la vitesse. Du point de vue du signal de position, la phase d'enfoncement a globalement une pente de signe négatif, la phase de maintien une pente nulle et la phase d'amortissement un signe positif.



Dans la mesure où les oscillations sont petites devant l'ambitus des différentes phases de maniement de la touche, on introduit un seuil d'amplitude. Une nouvelle frontière est détectée si l'écart en amplitude entre deux changements de signe dépasse le seuil fixé.

De la sorte, on identifie seulement 1/ la fin de l'enfoncement correspondant au premier extremum significatif ( $A_E$ ) et 2/ le premier extremum ( $A_Z$ ) de l'amortissement. La fin de la phase de maintien ( $A_M$ ), si celle-ci existe, correspond au dernier extremum rencontré précédant ( $A_Z$ ). La fin de la phase d'amortissement ( $A_A$ ) correspond à la fin du signal.

### b. algorithme

L'algorithme de sursegmentation `SurSegA()` identifie la fin de la phase d'enfoncement et la fin de la phase de maintien dans un segment actif simple.

La procédure `SurSegA()` fait appel à trois fonctions :

- 1 - `NextAExtr(A, segment, t)` retourne l'instant  $T_e$  de l'extremum suivant l'instant  $t$  tel que  $|S(T_e) - A| < \Delta A$ .  $\Delta A$  représente un écart d'amplitude et  $A$  une amplitude de référence. L'amplitude de cet extremum est conservée dans une variable globale  $A_e$ . L'instant de l'extremum précédent  $A_e$  est stocké dans la variable globale  $T_p$ . Sa détection est indépendante de la relation  $|S(T_p) - A| < \Delta A$ .
- 2 - `MarquerFA(segment, t)` pose une frontière sur le segment à l'instant déterminant la fin de la phase d'enfoncement,
- 3 - `MarquerFM(segment, t)` pose une frontière sur le segment à l'instant déterminant la fin de la phase de maintien.

```
int SurSegA(A, segment, t) int A; TypeSeg segment ; int t ;
/* ----- */
Pose les marques de fin d'enfoncement et de fin de
maintien. La fin de l'amortissement est la fin du
segment.
/* ----- */
{
    t = NextAExtr(A, segment, t) ;           // détection de fin d'enfoncement
    MarquerFA(segment, t) ; A = Ae ;         // marquage de AE
    t = NextAExtr(A, segment, t) ;           // détection de AZ
    MarquerFM(segment, Tp) ;                 // marquage de AM
    return (t);
}
```

## 2. segments-actifs consécutifs

L'identification des phases se complique si l'on s'intéresse à un signal contenant plusieurs segments-actifs successifs. Nous ne connaissons plus, a priori, l'instant de fin de l'amortissement.

Nous proposons dans ce paragraphe une technique pour scinder une suite de segments-actifs. Nous définissons pour cela un critère d'amortissement : un nouveau segment-actif est détecté lorsque le signal mesuré ne vérifie plus ce critère.

### **b. critère d'amortissement**

- Considérons le comportement d'un point de contact. Soit  $x(t)$  la fonction décrivant ce comportement et  $Ex(t)$  l'enveloppe tangente à  $x(t)$ . La courbe  $Ex(t)$  est un exemple de dépendance des paramètres de la segmentation automatique avec l'instrument.

On a 
$$Ex-(t) < x(t) \leq Ex+(t) \quad (1)$$

avec  $Ex-(t)$  la courbe tangente minorante et  $Ex+(t)$  la courbe tangente majorante de  $x(t)$ .

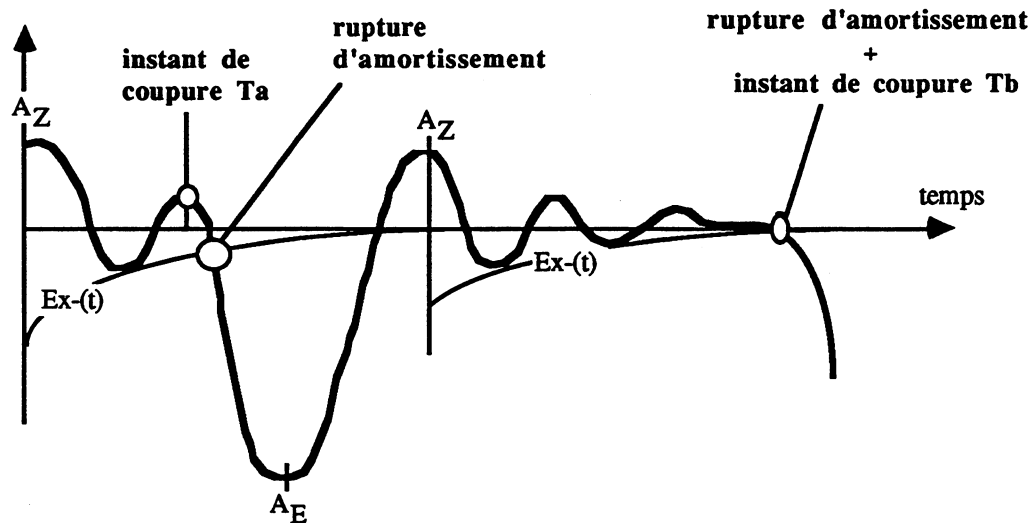
- Sur un signal mesuré  $S(t)$  provenant d'un événement gestuel ou d'une série d'événements, tous les intervalles  $[ta..tb]$  correspondant à un amortissement libre (sans contact opérateur) vérifient la relation (1)

$$Ex-(t) < S(t) \leq Ex+(t)$$

avec  $t$  dans  $[ta..tb]$ .

- Si l'on cherche l'instant à partir duquel une portion du signal  $S(t)$  n'est plus en amortissement libre, il suffit de trouver l'instant à partir duquel il ne vérifie plus la condition (1). Si la courbe d'amortissement sort de la courbe enveloppe, cela signifie que ce n'est plus l'amortissement libre de l'instrument et qu'un nouvel événement gestuel a eu lieu.

Sur la figure représentant un segment-actif, le critère d'amortissement s'applique à partir de l'extremum AZ. Il a pour fonction d'identifier l'instant à partir duquel la courbe mesurée ne vérifie plus l'amortissement libre de l'instrument. Cet instant correspond au segment suivant.



Le calcul du critère repose sur les deux points suivants :

cas 1 : l'extremum AZ est la valeur d'amplitude de référence. On suppose qu'à partir de là la courbe traduit un amortissement libre.

cas 2 : l'extremum qui suit vérifie ou non la condition (1).

S'il la vérifie, c'est que l'amortissement est libre. Dans le cas contraire, c'est **une rupture d'amortissement**. Une nouvelle perturbation a été introduite.

On considère que l'**instant de coupure** a lieu soit au dernier extremum rencontré (voir Ta sur la figure), soit en fin de la phase stationnaire si la courbe a atteint la position de repos (voir Tb sur la figure).

### c. algorithme

L'algorithme de sursegmentation  $\text{SurSegA}_n()$  donne les instants de séparation entre les segments-actifs d'une série, et à l'intérieur de chaque phase, précise l'instant de fin d'enfoncement et de maintien.

Elle fait appel à trois fonctions :

1 -  $\text{SurSegA}(A, \text{segment}, t)$  identifie et marque les fins d'enfoncement et de maintien dans un segment-actif simple. Elle retourne l'instant du premier extremum AZ de la phase d'amortissement. A est l'amplitude de référence prise en compte pour détecter un nouvel extremum global.

2 -  $\text{FinAmort}(\text{segment}, t)$  retourne l'instant **Tfa** de la fin d'amortissement (instant de coupure). Celle-ci est détectée lorsque le signal mesuré passe sous la courbe enveloppe  $\text{Ex}(t)$ , ou lorsque le signal se termine. L'instant Tfa dépend de la position de cette rupture dans l'amortissement :

- si la rupture est détectée en cours d'oscillation, Tfa correspond au dernier extremum rencontré avant l'instant de rupture,
- si la rupture est détectée en phase stationnaire de repos naturel, Tfa correspond à l'instant de rupture.

L'amplitude de l'instant de rupture choisi est appelée **Afa**.

3 - MarquerFA(segment, t) pose une frontière sur le segment à l'instant de fin de la phase d'amortissement. Cet instant correspond également au début du segment-actif suivant. Il n'y a donc pas une marque pour la fin d'amortissement et une pour le début de l'enfoncement suivant, mais une frontière commune pour les deux.

```

void SurSegAn(segment) TypeSeg segment ;
/* ----- */
Identification de la séquence (E.M*A)n.
La valeur d'amplitude A passé en paramètre à SurSegA0
est initialisée à la valeur du repos naturel, puis mis à
jour par l'amplitude de l'instant de fin d'amortissement.
/* ----- */
{ int A = R; int t = 0 ;

  while (t < TSEG )
  { t = SurSegA(A, segment, t) ;
    t = FinAmort(segment, t) ; A = Afa ;
    MarquerFA(segment, t);
  }
}

```

### V.3 conclusion

Nous avons tenté d'aborder la segmentation automatique des signaux gestuels à partir de considérations neutres, mais très vite, nous avons dû faire intervenir des connaissances gestuelles et instrumentales. Il semble en fait que plus l'analyse des signaux est fine, plus les processus de segmentation nécessitent la prise en compte de telles connaissances.

Citons deux exemples :

- 1- pour identifier les phases de repos naturel d'un signal, nous avons défini le critère de pause, paramétré par la position de repos naturel R du point de contact-instrument correspondant. Cette position peut varier d'un point à un autre.
- 2- pour identifier les instants de coupure d'une succession de segments-actifs contigus, nous avons défini un critère d'amortissement, paramétré par la courbe enveloppe de l'amortissement. Nous aurions également pu définir un critère de coupure réalisant ses comparaisons avec le signal issu des algorithmes de simulation. Il aurait été probablement plus juste et précis mais néanmoins plus exigeant vis-à-vis de la qualité

des signaux et certainement plus difficile à mettre en œuvre dans une situation pratique.

Il semble donc qu'on ne puisse pas supprimer la nécessité d'une connaissance minimale de l'instrument simulé. Plus loin, c'est en toute rigueur des propriétés mécaniques propres à l'instrument et à l'instrumentiste lui-même (de ses organes moteurs) qu'il faudrait tenir compte pour établir les critères de segmentation.

C'est une manière de dire qu'il n'existe pas d'algorithme de segmentation universel et général.

## Représentation et traitement neutres

---

### I. INTRODUCTION

### II. ATTRIBUTS DES SEGMENTS GESTUELS

#### II.1. attributs intrinsèques

1. segment-actif - 2. segment-pause

#### II.2. attributs extrinsèques et attributs externes

### III. REPRESENTATION NEUTRE DES SEGMENTS GESTUELS

#### III.1. représentation avec l'ordinateur

1. objets filaires et perspective 3D -
2. interfaces gestuelles et temps-réel

#### III.2. représentation exhaustive

1. segments gestuels - 2. articulation des segments gestuels

#### III.3. représentation des attributs

##### 1. support de visualisation

- a. tablature et notation musicale traditionnelle -
- b. tablature-étendue
- c. les représentations du temps - d. les commentaires

##### 2. représentation visuelle des segments gestuels

- a. origine et durée - b. segment-actif simplifié -
- c. segment-actif codé

##### 3. représentation visuelle des segments-pauses

### IV. LA FORMULE GESTUELLE

#### IV.1. définition

#### IV.2. attributs de la formule gestuelle

1. attributs intrinsèques - 2. attributs extrinsèques

#### IV.3. structure

1. exemple - 2. optimisation structurelle

#### IV.4. représentation visuelle de la formule



## **V. TRAITEMENTS NEUTRES ELEMENTAIRES**

### **V.1. segment-actif**

1. tracer "manuel" - 2. exemples de transformations

### **V.2. traitements sur la tablature-étendue**

1. pointeur de la partition gestuelle
2. sélection sur la partition gestuelle
3. traitements "compositionnels" élémentaires
  - a. implémentation informatique -
  - b. remarque sur la taille mémoire utilisée
4. exemples de transformations
  - a. affinité - b. translation - c. inversion

### **V.3. structuration de la formule**

1. grouper - 2. dégrouper

## **VI. CONCLUSION**

---

## **I. INTRODUCTION**

Nous développons dans ce chapitre un ensemble de considérations spécifiques à une approche neutre de la représentation et du traitement du geste instrumental.

Une séquence de jeu instrumental sur clavier est faite d'un assemblage d'événements gestuels. Les segments dont on vient de parler représentent, à un premier niveau, la substance de ces événements, et à un deuxième niveau, la façon même dont ils s'enchaînent les uns par rapport aux autres. On appellera articulation, l'organisation des segments gestuels dans l'espace et le temps.

Le problème se pose de la façon suivante :

Il s'agit de définir des représentations graphiques des segments gestuels et de leur articulation telles que le compositeur puisse percevoir et interpréter de manière directe les différentes valeurs de leurs attributs, ainsi que leur articulation.

Il s'agit ensuite de déterminer des possibilités de traitement et de transformation utilisant ces représentations. Le traitement proprement dit doit permettre d'une part de modifier les attributs, directement ou non, et d'autre part de combiner les segments gestuels.

Nous abordons ce problème en définissant a priori un ensemble d'attributs caractérisant de façon pertinente les segments gestuels. Ils sont directement suggérés par les signaux les constituant.

## **II. ATTRIBUTS DES SEGMENTS GESTUELS**

Les attributs d'un segment doivent permettre d'en décrire certaines caractéristiques propres et de le représenter de façon plus économique et plus pertinente, bien que moins exhaustive, que la donnée complète des signaux qui le constituent.

On parle d'attributs intrinsèques pour décrire ce qui est à l'intérieur des segments et d'attributs extrinsèques pour ce qui les caractérise de l'extérieur.

Ils sont calculés lors de la segmentation ou lors des traitements.

### **II.1. attributs intrinsèques**

La définition des attributs intrinsèques doit se faire uniquement à partir des signaux constituant les segments gestuels.

### 1. segment-actif

De façon très générale, un segment-actif comporte un ensemble de  $N$  signaux, qu'il est possible de caractériser de maintes façons : par des longueurs, des ambitus, des moyennes, etc...

Cependant, au lieu de rester à un niveau de considérations totalement neutre et général, nous appliquons notre étude aux segments-actifs définis au chapitre précédent. Ils se décomposent en au plus trois phases : une phase d'enfoncement (E), une phase éventuelle de maintien (M) et une phase d'amortissement (A).

Ces trois phases successives sont significatives d'un geste de manipulation d'une touche de clavier.

Dans ce cas particulier, nous faisons la proposition de trois attributs : la longueur en nombre d'échantillons de chaque phase, leur ambitus, et leur pente c'est à dire le rapport entre l'ambitus et la longueur.

Ces trois attributs nous semblent pertinents. La pente de chaque phase traduit une intensité du geste. L'ambitus traduit une amplitude du geste et la longueur des signaux une durée objective.

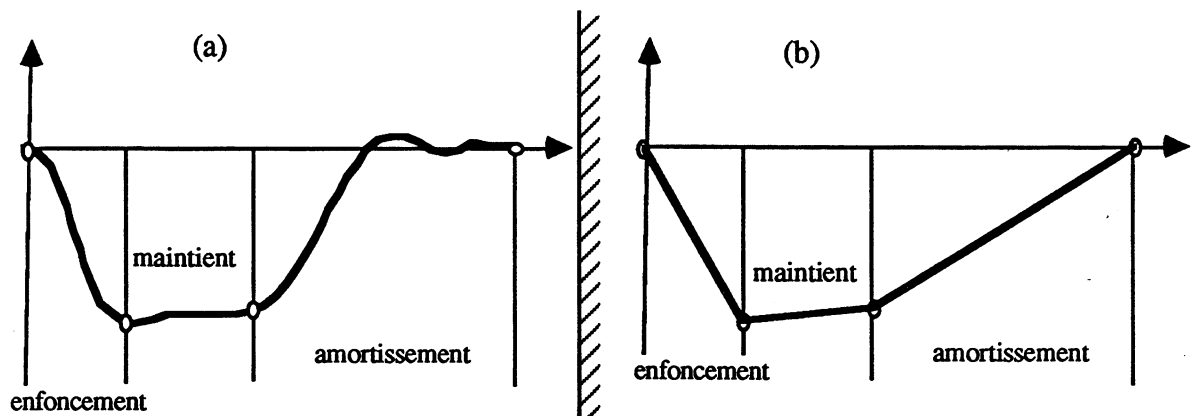


Figure 1. *Les attributs intrinsèques d'un segment-actif.* Un segment-actif caractéristique du maniement d'une touche se décompose en trois phases. Chaque phase peut se représenter par une longueur, un ambitus et une pente.

La modification de ces attributs permet alors d'intervenir sur la force, l'amplitude et la durée de l'événement gestuel correspondant.

On remarquera qu'il n'y a pas bijection entre les signaux gestuels et les attributs intrinsèques les représentant. Les attributs se calculent de façon déterministe à partir des signaux bruts. A l'inverse, il n'existe pas qu'une seule manière de reconstruire le

signal à partir des attributs. Il faudra choisir arbitrairement le processus de reconstruction.

C'est une façon de rappeler que l'information gestuelle la plus exhaustive reste la donnée des signaux gestuels.

## **2. segment-pause**

Dans le cas général, le segment-pause est constitué d'un ensemble de signaux stationnaires dont les amplitudes correspondent aux positions de repos naturel du point de contact considéré. Cela signifie que théoriquement, lors d'une pause, l'opérateur n'est pas en couplage avec l'instrument. Il n'y a donc pas d'information significative d'une action gestuelle à ce moment-là.

Néanmoins d'un point de vue du signal, on peut déterminer deux attributs caractéristiques d'une stationnarité : la longueur des signaux et leur amplitude.

### **II.2. attributs extrinsèques et attributs externes**

La définition des attributs extrinsèques s'appuie sur toute autre information que celles contenues dans les signaux. Ils sont indépendants des attributs intrinsèques.

C'est par exemple un instant d'apparition et un numéro de touche de clavier. Plus généralement, on appellera "origine" d'un segment, le couple (C, T) où C représente le numéro de canal et T son instant d'apparition.

Mais c'est également la dimension du segment (segment 1D pour un jeu pianistique, segment 2D pour un frottement d'archet, ...), le type de segment (segment-actif, segment-pause, ...), etc...

Nous distinguerons dans la terminologie les attributs extrinsèques et les attributs externes.

Ces derniers décrivent exclusivement les aspects de la représentation des segments : pictogramme, couleur, coefficient de dilatation visuelle, etc...

On peut étendre la notion de représentation visuelle à d'autres représentations.

### **III. REPRESENTATION NEUTRE DES SEGMENTS GESTUELS**

Le problème de la représentation se pose à plusieurs niveaux : au niveau intrinsèque des segments, au niveau de leur articulation, et à un niveau que nous avons baptisé "formule", parce qu'elle permet de structurer les segments entre eux de façon à obtenir des formes composites indissociables.

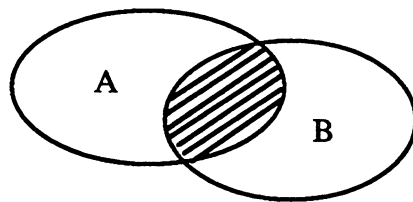
Il s'agit pour chaque niveau de proposer des représentations graphiques des entités gestuelles exprimant le plus directement possible leurs attributs.

Dans la mesure où notre étude a dû aborder un ensemble de problèmes différents, les propositions faites doivent être considérées comme introductrices. Elles demanderaient à être approfondies et davantage justifiées.

#### **III.1. représentation avec l'ordinateur**

De façon générale, l'ordinateur offre la possibilité de visualiser des objets graphiques sous forme filaire, éventuellement surfacique, monochromes ou en couleur, dans des plans 2D ou en perspective 3D, etc..., et surtout de les manipuler et de les transformer en temps-réel.

Mais comme le rappelle C.Cadoz [Cad79], tout système de représentation peut ajouter à l'information extraite de l'objet de référence, une information supplémentaire, qui n'appartient pas à ce dernier.



A : objet de référence  
 B : système de représentation  
 $A \cap B$  : réduction de l'objet représenté  
 $B - A \cap B$  : apport propre au système

Cet aspect de la représentation est particulièrement important dans la mesure où lors du processus de représentation, il devient difficile de séparer complètement l'objet représentant et le moyen de représentation.

Nous décrivons donc dans ce paragraphe quelques caractéristiques du poste graphique (Evans&Sutherland PS300) sur lequel a été réalisée l'application logicielle de la thèse. Elles expliquent dans une certaine mesure les orientations que nous avons prises pour représenter les éléments du geste instrumental.

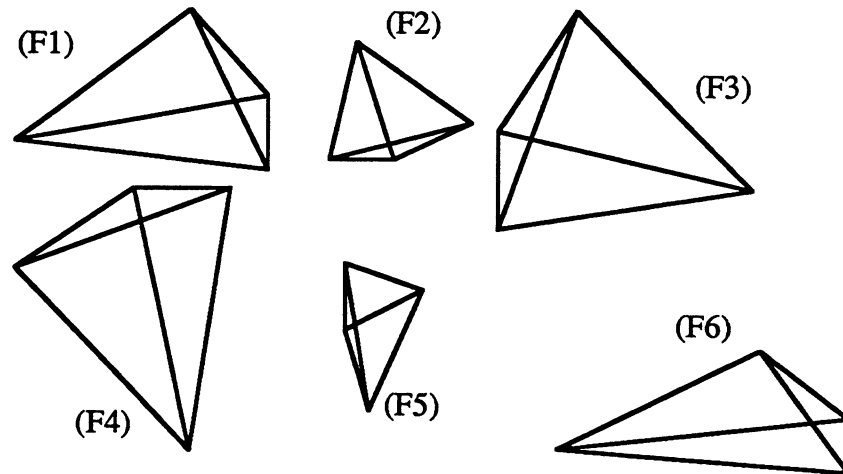
### 1. objets filaires et perspective 3D

Un programmeur, travaillant avec le PS300 pour désigner, orienter et manipuler des objets, commence par créer une base de données contenant les informations mathématiques qui les définissent.

Ces objets sont représentés à l'écran sous forme filaire ; le PS300 ne permet pas de visualiser des surfaces pleines.

Dans un tel système, le programmeur peut définir des objets complexes constitués de primitives et de noeuds de transformation.

Dans la figure suivante, le prisme (F1) est la primitive, définie par ses six arêtes. Les autres occurrences de cette primitive sont obtenues par application de noeuds de translation, de rotation et d'affinité.



La perspective est gérée par le système et paramétrée par la donnée d'un cône de visualisation (point de vue, point de visée, ...).

### 2. interfaces gestuelles et temps-réel

Le PS300 offre la possibilité d'alimenter des noeuds de transformation (translation, rotation, affinité ...) en temps-réel par le biais d'un ensemble d'interfaces de contrôles : boîte à boutons, souris, ou touches clavier. Chaque interface a un équivalent fonctionnel dans le système, susceptible d'être connecté aux noeuds de transformation.

La souris "absolue" (resp. "relative") retourne un déplacement absolu (resp. relatif) sur l'écran

et comporte quelques boutons, sélectionnables séparément ou simultanément.

La boîte à boutons comporte 8 potentiomètres qui retournent pour chaque déplacement  $\Delta a$ , une valeur  $k.\Delta a$  ( $k$  est défini par le programmeur).

Une suite de petites "leds" programmables accompagne chaque bouton.

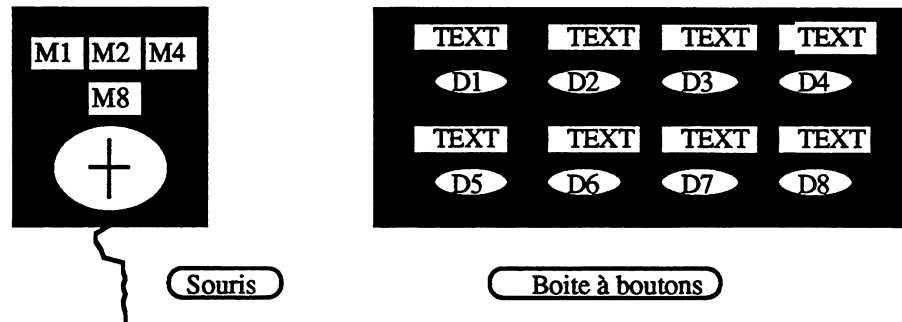


Figure 2. - *Les interfaces*. La souris retourne 1/ les coordonnées (absolues ou relatives) d'un curseur écran et 2/ un nombre codant les boutons sélectionnés. La boîte à boutons délivre des valeurs déterminées par le déplacement des potentiomètres. Chaque bouton peut être commenté par une chaîne de 8 caractères.

### III.2. représentation exhaustive

La représentation la plus exhaustive des segments gestuels au niveau intrinsèque et extrinsèque est celle des signaux les constituant et de leur localisation dans un espace approprié. Elle ne fait pas apparaître directement les attributs pertinents et demande à l'opérateur un décodage explicite en général difficile à réaliser.

En outre, tout processus de représentation introduit des informations supplémentaires, n'appartenant pas à l'objet de référence, et en supprime nécessairement d'autres. La représentation exhaustive est le seul moyen qui permette une lecture exhaustive de l'information gestuelle captée et mémorisée avec le système CORDIS.

#### 1. segments gestuels

Dans le système CORDIS, les signaux gestuels représentent des comportements en amplitude. Pour visualiser les signaux des segments gestuels de façon exhaustive, il suffit d'un, ou de plusieurs repères cartésiens temps-amplitude.

Afin d'en faciliter la lecture, on peut ajouter aux repères de base plusieurs moyens d'informations :

- commentaires sur le segment : attributs intrinsèques : durée du segment, phases de maniements (enfoncement ...) ; attributs extrinsèques : type (si d'autre formes que enfoncement-maintien-amortissement ont été définies), nombre de voies (si

les segments désignent des points de contact N-D), etc ; attributs de visualisation : voie en cours de visualisation, etc,

- fonctions de mise au point : elles concernent uniquement la visualisation des signaux et du repère cartésien. Elles permettent d'augmenter la visibilité de certaines parties, de changer d'échelle d'observation, etc,
- fonctions d'analyse : mesure, marquage, traitement du signal, etc. Elles permettent par exemple de calculer la valeur de la durée, de la pente ou de la courbure de certaines portions des signaux.

Pour désigner les portions du signal, on peut utiliser des règles (l'une horizontale se déplaçant verticalement et une deuxième verticale se déplaçant horizontalement), un pointeur associé à une souris, etc.

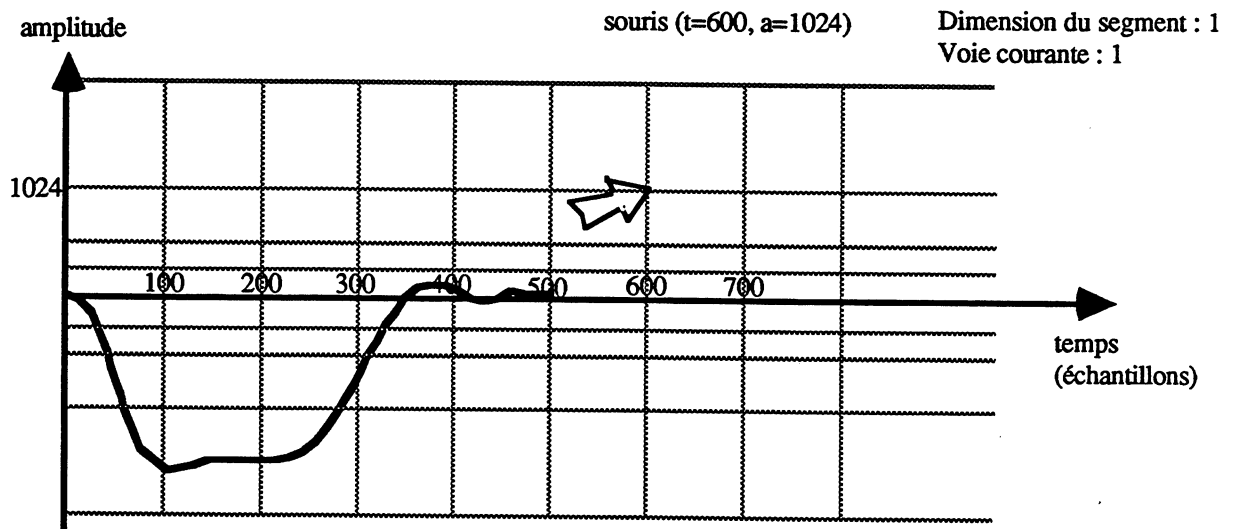


Figure 3. *Support de visualisation du segment.* La figure donne un exemple de visualisation d'un segment mono-dimensionnel, sur un repère amplitude / temps. L'unité de l'amplitude est une subdivision du format de codage des échantillons et celle du temps est donnée en nombre d'échantillons. Un pointeur de souris, du commentaire et des graduations favorisent une observation plus poussée.

## 2. articulation des segments gestuels

Un segment comporte des signaux temps-amplitude et une origine temps-canal. Un repère 3D canal-temps-amplitude est donc suffisant pour représenter de façon exhaustive l'articulation d'un ensemble de segments.

Chaque segment y est représenté sous la forme de signaux.

On peut espérer de cette manière mettre en évidence des formes d'une logique immanente au jeu instrumental.



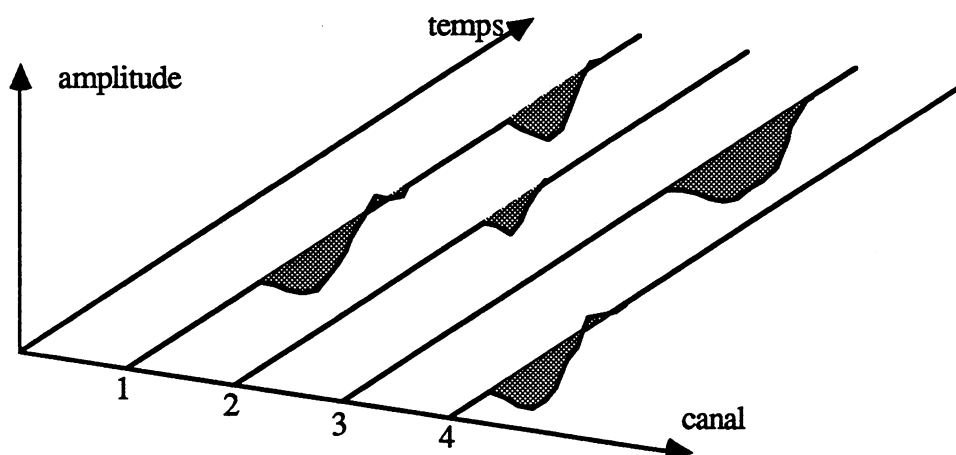


Figure 4. *Représentation exhaustive de l'articulation*. Les segments gestuels sont représentés sous forme de signaux dans un repère 3D : canal / temps / amplitude.

### **III.3. représentation des attributs**

La représentation exhaustive, bien que nécessaire, a des limites : sa complexité graphique rend difficile l'interprétation directe et rapide des différents attributs. Pour réduire ces limites, et augmenter de ce fait la lisibilité de l'information gestuelle, le problème consiste donc à définir un ensemble de représentations aptes à mettre en évidence plus directement les attributs intrinsèques et extrinsèques des segments gestuels.

Les attributs que nous souhaitons mettre en évidence sont :

- les attributs intrinsèques : durée du segment, durée, ambitus et pente des différentes phases le constituant,
- attributs extrinsèques : une origine spatiale (canal, amplitude), une origine temporelle, dimension du segment et son type.

#### **1. support de visualisation**

En premier lieu, nous nous intéressons à la définition d'un support de représentation plus économique et plus pertinent que le repère 3D canal-temps-amplitude.

Etant donné les attributs à représenter, il peut être intéressant de réaliser un support sur lequel apparaissent explicitement une dimension temporelle et la structuration en canaux et voies de l'instrument. Les attributs extrinsèques sont alors traduits directement. Les attributs intrinsèques doivent être codés.

Certains aspects de ce support sont à rapprocher de celles d'une tablature telle que la notation musicale traditionnelle l'a défini.

**a. tablature et notation musicale traditionnelle**

Il faut distinguer deux formes de tablature. La première est une simplification de la notation usuelle et décrit donc des hauteurs, des durées, etc... ; c'est le cas de la tablature pour orgue. La seconde représente les notes directement sur l'instrument et n'indiquent donc pas explicitement des hauteurs fréquentielles ; c'est le cas de la tablature pour luth.

P.MacNeillage [Mac71] définit la tablature comme "une représentation des sons musicaux propres à tels instrument ou à telle catégorie d'instrument. Cette définition est plus générale, mais plus exacte que celle que l'on rencontre dans certains traités et qui définit la tablature comme un système de signes déterminant la place de chaque note sur l'instrument."

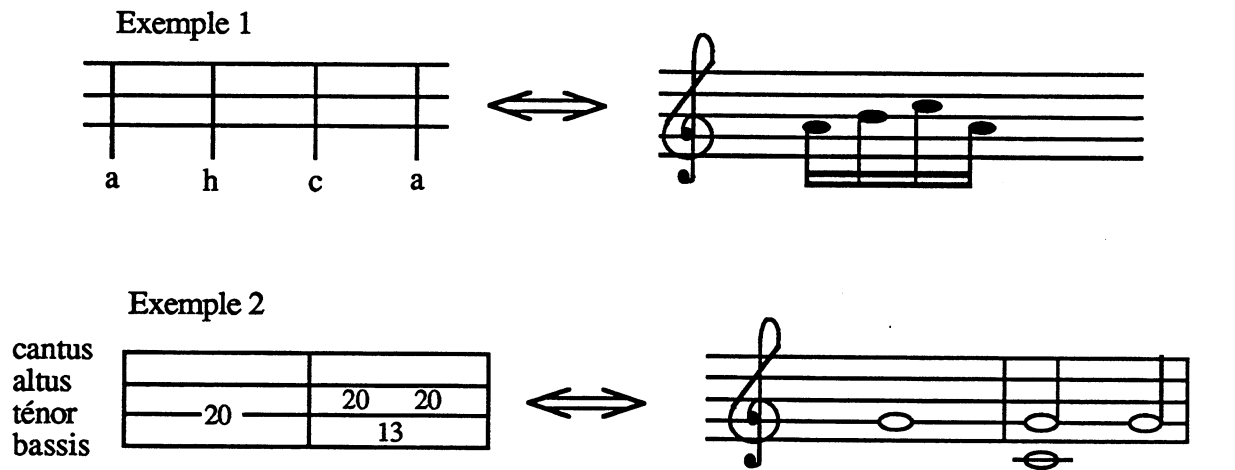


Figure 5. - *Tablature pour orgue*. Dans l'exemple 1 (Allemagne), les notes sont alphabétiques et leur valeur précisée par un, deux ou trois traits horizontaux. Dans l'exemple 2 (Espagne), les notes sont désignées par des chiffres s'échelonnant de 1(=Do) à 42(=La5).

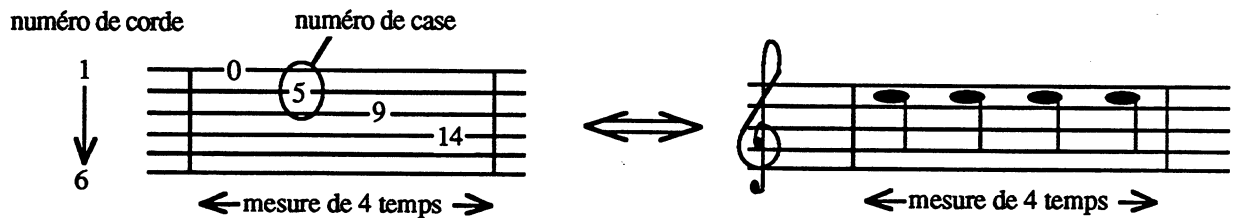


Figure 6. - *Tablature pour luth*. Les cordes du luth sont représentées par des lignes horizontales sur lesquelles sont spécifiés les numéros de cases à sélectionner en même temps qu'elles sont pincées. La tablature pour luth fixe sans équivoque la position de la note à jouer.

### **b. tablature-étendue**

La tablature pour les instruments à corde a persisté jusqu'à nos jours, et ce pour une raison assez évidente ; elle représente directement l'instrument. Elle guide explicitement les doigts de l'exécutant et leur position sur l'instrument. Son principe diffère donc complètement du principe d'une portée, et se dit très efficace.

Nous proposons comme support de représentation un repère canal-temps. On l'appellera "tablature-étendue" parce qu'elle étend le principe de la tablature classique pour luth selon deux directions :

- la tablature-étendue représente tout instrument structuré en canaux et voies, quelque soit le nombre de ces canaux et de leur dimension,
- les événements représentés sur la tablature-étendue sont des segments-gestuels (segments-actifs et segments-pauses).

#### représentation des canaux-gestuels

Pour représenter les canaux de la tablature, nous avons choisi un objet graphique plan constitué d'une juxtaposition d'interlignes de même largeur. Chaque canal correspond à un interligne, dans lequel les segments gestuels sont directement visualisés.

La juxtaposition de droites a des qualités certaines pour la perception visuelle. Il suffit pour s'en convaincre de considérer l'efficacité de la tablature et de la portée classiques. Remarquons que la représentation d'un canal par un interligne (plutôt que par une ligne) favorise la visualisation de pictogrammes simples tels que des segments-droite. Rappelons que notre poste graphique ne permet de représenter que des objets filaires.

#### tablature-perspective

Notre poste graphique offre la possibilité de définir des objets tri-dimensionnels et de les représenter en perspective.

Nous avons introduit une représentation de la tablature dans un espace 3D en perspective. Les canaux sont numérotés de la gauche vers la droite et le temps se profile dans la profondeur de l'écran.

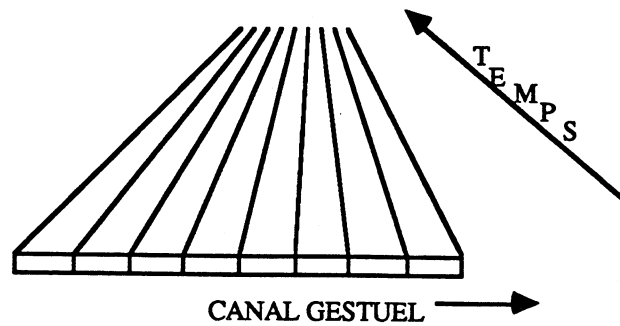


Figure 7. - *Tablature dans un espace 3D.* Avec l'ordinateur, on a la possibilité de représenter et manipuler rapidement des objets graphiques dans des scènes bi- ou tri-dimensionnelles.

Les avantages de cette forme de visualisation sont multiples :

- visualisation multiple : perspective / plane 2D verticale / plane 2D horizontale,
- elle traduit de façon originale certains aspects du temps,
- le déplacement sur la tablature se fait à l'image d'un marcheur sur une route [CR90] et traduit des propriétés spatiales et temporelles très intéressantes,
- elle permet d'avoir devant soi toute l'information future. Celle-ci n'est pas masquée par des pages à tourner, ni tronquée par la largeur de ces pages. La lecture est facilitée par les dispositions graphiques (point de vue, le point de visée, zoom, etc.).

Néanmoins, la tablature, telle qu'elle a été définie, présente aussi quelques inconvénients par rapport à une représentation classique sur papier :

- la façon dont on se positionne détermine nécessairement un instant présent et donc une frontière entre le futur, vers laquelle le marcheur se tourne, et une zone passée, qui elle est complètement masquée,
- elle ne permet pas d'avoir un "coup d'œil global" comme le permet par exemple une partition papier traditionnelle,
- la complexité en nombre de canaux peut rapidement devenir un facteur d'illisibilité.

#### paramètres de la visualisation

La représentation en perspective n'a probablement que peu d'intérêt dans un contexte de représentation sur papier. En revanche, l'ordinateur permet de la manipuler, en temps-réel, comme un objet graphique quelconque. On peut réaliser des changements de point de vue, des transformations mathématiques telles que des affinités, des translations et rotations, etc...

Mais la définition des transformations graphiques ne doit pas se faire de manière abusive. Elle doit tenir compte de facteurs ergonomiques, tels que l'opérateur puisse obtenir des points de vues pertinents, rapidement et simplement.

Nous avons proposé comme solution à ce problème l'analogie au marcheur. Au lieu de déplacer et transformer directement l'objet graphique, c'est l'opérateur qui se déplace sur l'objet : il avance, il recule et change l'angle et la distance de visée.

Les changements du point de vue sont déterminés par 4 paramètres :

- un **plan de visée** orthogonal à la partition-gestuelle, orienté soit dans la direction temporelle, soit par le côté. Ce plan contient l'axe de visée.
- la distance **R** de l'axe de visée entre un œil et un point de la tablature. La modification de ce paramètre permet de s'approcher ou de s'éloigner du point visé,
- l'angle  $\alpha$  formé par l'axe de visée et le plan de la partition-gestuelle. Ce paramètre, variant entre 0 et 90°, donne à l'utilisateur la possibilité de voir la tablature dans le sens du temps,
- l'angle  $\beta$  modifie l'échelle temporelle. Ce paramètre permet de dilater ou de compresser la partition-gestuelle dans le temps.

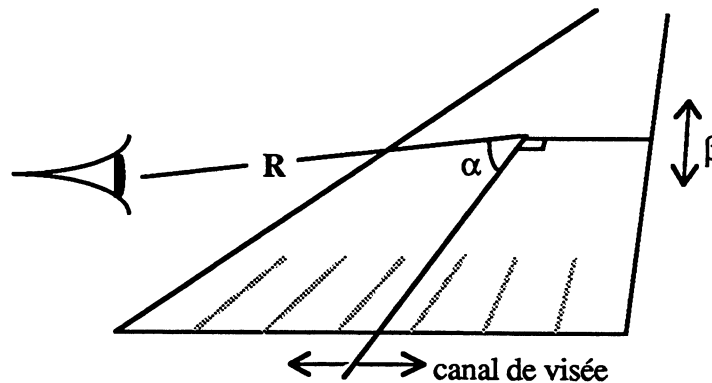


Figure 8. *Mise au point sur la route.* Pour une représentation perspective de la route, l'axe de visée principal se situe dans un plan orthogonal au plan de la tablature dans le prolongement d'un canal de visée, à une distance R du point visé.

### c. les représentations du temps

Dans le cadre de la représentation de la tablature-étendue, qu'elle soit en "vue 2D" ou en perspective (voir figure suivante), le temps est représenté par une dimension spatiale. Néanmoins, la façon d'inscrire le temps sur cette dimension spatiale peut être envisagée de plusieurs façons.

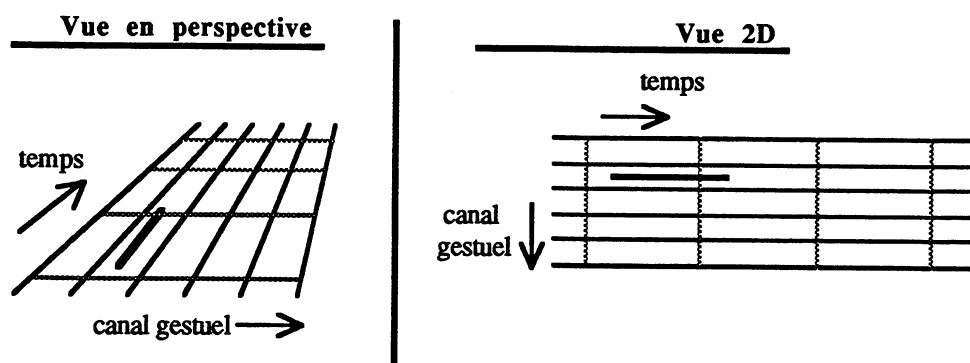


Figure 9. *Visualisations de la tablature.* La représentation en perspective intègre également la visualisation classique 2D horizontales.

### représentation absolue

La représentation la plus immédiate associe directement une longueur à une durée. Chaque événement est repéré temporellement par l'intervalle de temps qui le sépare d'une origine absolue. Cette origine est commune à tous les canaux.

Soit alors  $F_e$  la fréquence d'échantillonnage et  $n$  la position sur la tablature. L'instant  $T$  correspondant est donné par :  $T = n / F_e$ . De même, une durée  $\Delta T$  sera donnée par la relation  $\Delta T = \Delta n / F_e$ .

Dans cet esprit, les attributs de durée et de date d'apparition d'un segment devront être traduits par des informations spatiales : une position  $n$  sur la tablature et la longueur  $\Delta n$  du pictogramme.

Pour aider la lecture du temps absolu, on peut graduer la tablature dans la dimension temporelle. Par exemple, l'espace séparant deux graduations correspond à une seconde.

Remarquons que dans le cadre d'une visualisation s'appuyant sur le paradigme du marcheur, la représentation spatiale du temps a une propriété singulière intéressante : plus l'événement est éloigné, plus son encombrement graphique est faible.

Ceci correspond à une situation humaine très naturelle : les phénomènes lointains ont moins d'importance que les phénomènes proches. Cette constatation vaut autant pour les représentations spatiales que pour les représentations temporelles.

### la mesure

Dans la réalité, le processus de représentation mentale du temps n'est pas aussi simple. Il est facile de constater que le musicien ne perçoit pas le temps de façon homogène. Il est par exemple des situations telles que deux événements comparables, apparaissant à deux instants différents, ne seront pas perçus de la même manière.

On peut alors s'intéresser à cette non-homogénéité et en proposer des représentations.

Considérons par exemple que la distribution du temps le long de l'axe temporel n'est pas linéaire, mais soumise à des facteurs d'échelle locaux. Nous traduisons cette distribution par la donnée d'une succession de mesures.

Nous définissons ainsi la mesure comme un couple de donnée  $(\Delta T, \alpha)$  :  $\Delta T$  décrit la durée d'un intervalle séparé par deux barres de mesure et  $\alpha$  est un coefficient de dilatation visuelle.  $\alpha$  ne modifie pas la durée  $\Delta T$ .

Les mesures sont séparées par des barres de mesure.

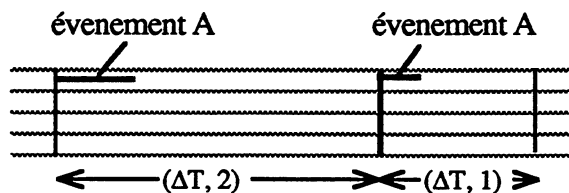


Figure 10. - *La mesure*. Sur la figure sont représentées deux occurrences d'un même événement A. Les deux mesures ont une durée  $\Delta T$ . La première mesure est visualisée avec un coefficient de dilatation 2 et la seconde avec un coefficient 1. La durée de chaque événement n'est pas modifiée.

#### représentation relative

On peut rapprocher la mesure telle que nous l'avons définie avec celle de la notation traditionnelle.

Dans ce système de notation, les durées des événements (notes et silences) ne sont pas traduits par des distances mais par un code graphique : le graphisme de la noire correspond à un temps, celui d'une croche correspond alors à un demi-temps.

La durée  $d$  (en minutes) d'une note  $n$  est calculée en fonction du tempo (nombre de noires à la minute) :  $d = n / 60 \text{ mn}$ .

Remarquons néanmoins que le graphisme d'une note est d'autant plus complexe que la durée de la note est courte. En ajoutant un crochet à la hampe, on divise sa durée par deux. Ceci tend à montrer que le graphisme d'une note sur une portée traditionnelle n'est pas une durée mais un rapport de durée, soit une fréquence.

#### **d. les commentaires**

D'après B.Gagnepain [Gag87], le rôle essentiel de la mesure est de mettre en évidence certaines zones temporelles et de permettre une plus grande facilité pour le lecteur à reconnaître visuellement les rythmes. Ceux-ci peuvent d'ailleurs se manifester sous la forme la plus parfaite en l'absence complète de mesure.

On peut généraliser la mesure et la barre de mesure à la notion de commentaire.

Définissons le commentaire comme un signe qui n'a aucune incidence sur l'information gestuelle. Il permet au compositeur de mettre en évidence, graphiquement, certains aspects tels que des intervalles de temps, des zones spatiales particulières, de les commenter avec du texte, ...

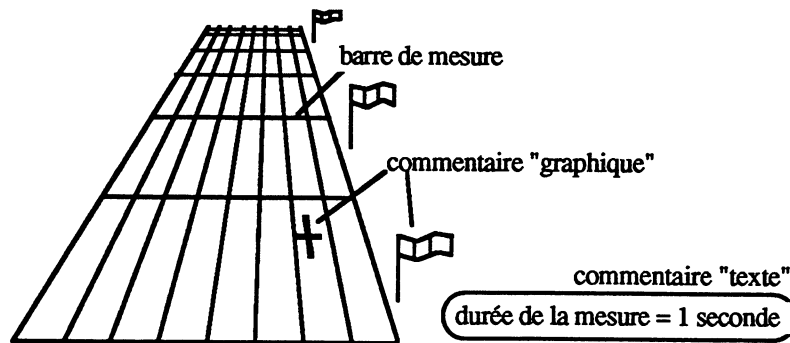


Figure 11. - *Les commentaires*. Toute l'information accompagnant la représentation du geste sur la tablature est appelée commentaire.

## **2. représentation visuelle des segments gestuels**

Il s'agit dans cette sous-section de définir des pictogrammes qui représentent plus directement les valeurs des attributs extrinsèques et intrinsèques. Ils seront visualisés sur la tablature-étendue.

### **a. origine et durée**

Pour décrire l'origine spatio-temporelle d'un segment gestuel, il suffit dans une échelle temporelle absolue, d'une information graphique ponctuelle.

Ce point peut être isolé ou correspondre à une partie prédéfinie d'un objet graphique plus complexe : l'extrémité d'un trait ou le centre d'une croix par exemple.

Pour ce qui concerne l'attribut de durée d'un segment, il faut à nouveau poser un problème de représentation qui n'est pas nécessairement simple, car la représentation mentale du temps n'est pas forcément linéaire.

Dans la mesure où nous n'avons pas cherché à résoudre ce problème, nous avons adopté une représentation qui associe à la durée un trait de longueur proportionnelle.

### **b. segment-actif simplifié**

Nous nous intéressons ici aux attributs intrinsèques d'un segment actif. Il s'agit de la durée, de l'ambitus et de la pente de chacune des trois phases le constituant : une phase d'enfoncement (E), une phase éventuelle de maintien (M) et une phase d'amortissement (A).



Pour les mettre en évidence plus directement, une solution consiste à construire un signal plus simple à partir de ces attributs.

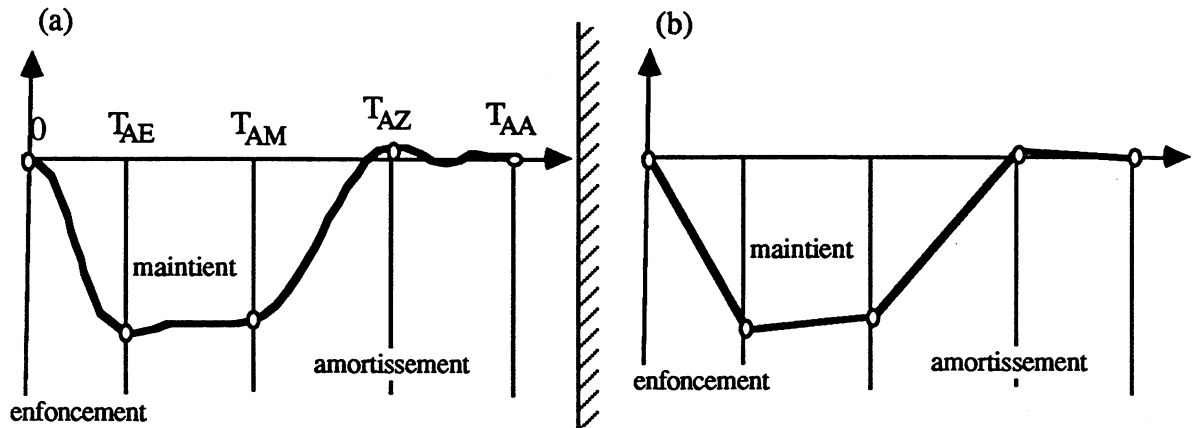


Figure 12. Représentation simplifiée. Les attributs intrinsèques du segment (a) sont mis en évidence par une simplification du graphisme (b).

#### phase d'enfoncement

Elle se caractérise par une pente de signe négatif. Nous représentons cette phase par un segment de droite unique défini par deux points : le premier point du signal ( $0, A_0$ ) et le premier extremum rencontré ( $T_{AE}, AE$ ).

#### phase de maintien

Elle a une pente faiblement évolutive. Quelques oscillations liées à un ajustement terminal de l'opérateur peuvent apparaître en son début. Pour la représenter, nous réduisons cette phase à un segment de droite reliant ( $AE$ ) et la fin de la stationnarité ( $AM$ ).

#### phase d'amortissement :

Deux cas importants d'amortissement sont à distinguer :

- la vitesse de relâchement est supérieure à celle de l'amortissement libre de l'instrument. Cela signifie que la phase de contact s'arrête au plus tard au premier extremum global rencontré ( $AZ$ ),
- la vitesse de relâchement est inférieure à celle de l'amortissement. La phase de contact se termine en  $AZ$ .

Pour décrire au mieux cette distinction, nous séparons la phase d'amortissement en deux parties : la phase  $AM-AZ$  et la phase  $AZ-AA$ , représentées chacune par un segment-de droite.

### c. segment-actif codé

La représentation de signaux amplitude-temps sur la tablature se rapproche fortement d'une représentation exhaustive sur un repère canal-temps-amplitude. On retrouve dans le même temps les inconvénients que cela comporte.

Une manière d'éviter cette "complexification" consiste à définir des graphismes originaux, dont les formes ne sont pas forcément en rapport direct avec celles du signal, et qui codent les valeurs des attributs.

Prenons un exemple.

La phase d'enfoncement (de l'origine au premier extremum AE) et la phase d'amortissement (de AM à AZ) sont représentées par des triangles. La largeur de la base du triangle est fonction de l'ambitus de la phase. Plus cet ambitus est important, plus la base du triangle est large.

Pour la phase d'enfoncement, la pointe du triangle va dans le sens du temps. Pour l'amortissement, c'est le sens contraire.

La hauteur du triangle correspond à la durée de la phase associée.

La phase de maintien (AE-AM) et l'amortissement libre (AZ-AA) sont représentées simplement par un trait décrivant leur durée.

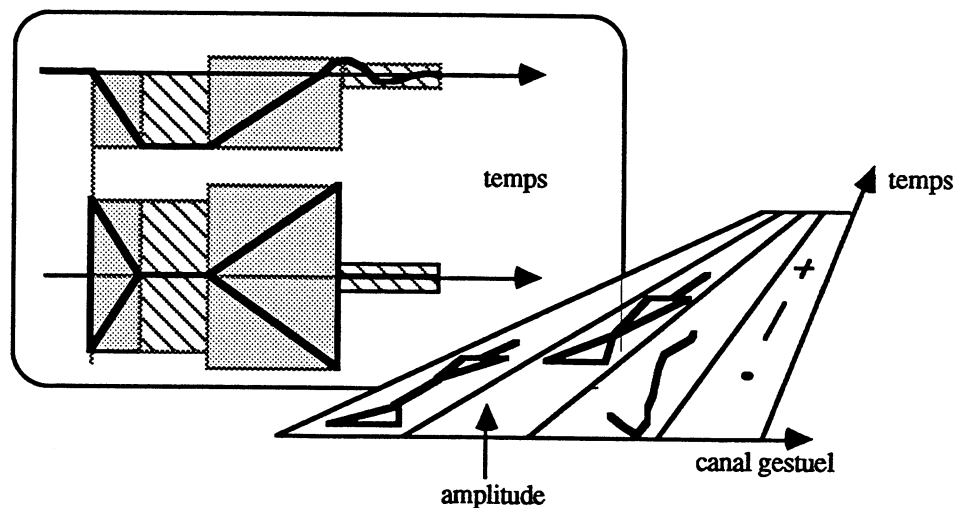


Figure 13. *Représentation du segment-actif.* Le pictogramme point (ou croix,...) renseigne sur l'origine du segment, le "trait" complète cette information par la durée totale et le "sablier" évoque les différentes phases d'un segment-actif réalisé à partir d'un clavier.

### 3. représentation visuelle des segments-pauses

Un segment-pause se caractérise par deux attributs intrinsèques : une durée et une amplitude de repos naturel. Rappelons également, que lors d'un segment-pause, l'opérateur n'est pas en couplage avec l'instrument. Il n'y a donc pas d'information significative d'une action gestuelle à ce moment là.

On peut traiter le problème de la représentation de différentes manières :

- on peut l'envisager de la même façon que pour les segments-actifs, c'est à dire que l'on cherche à définir des pictogrammes particuliers décrivant les attributs des segments-pauses,

- on met en évidence les segments-pauses en ne les visualisant pas. Une portion de canal vide de graphisme signifie que le point de contact correspondant est dans un état de repos naturel. On privilégie de la sorte l'information purement gestuelle.

Rappelons alors que sur un canal donné, les bornes temporelles des segments-pauses sont entièrement déterminées par celles des segments-actifs. Un problème de lecture peut donc se poser lorsque la durée effective des segments-actifs n'est pas mise en évidence graphiquement.

Nous avons retenu cette option dans les exemples précédents de représentation visuelle des segments-actifs.

## **IV. LA FORMULE GESTUELLE**

Nous avons défini en introduction de ce chapitre la notion d'articulation gestuelle : il s'agit de la façon dont s'organisent les segments gestuels entre eux, et par conséquent, la façon dont s'enchaînent les événements gestuels auxquels ils font référence.

Pour décrire la structure de cette information, nous proposons ici le concept de "formule gestuelle" et posons les problèmes de représentation et de traitement de la même façon que pour les segments gestuels.

### **IV.1. définition**

La formule gestuelle est une entité composée, servant uniquement à structurer les segments-gestuels entre eux. Elle n'introduit pas d'autre information logique que celle concernant les liens existants entre ces segments.

Nous définissons la structure d'une formule sur le principe simple d'une arborescence. Les feuilles d'une formule sont les segments gestuels et les nœuds de l'arbre sont appelés "sous-formules". Chaque sous-formule vérifie donc le même principe qu'une formule : ses feuilles sont des segments et les nœuds la constituant sont également des sous-formules.

### **IV.2. attributs de la formule gestuelle**

Pour représenter et traiter la formule, nous définissons, comme nous l'avons fait pour les segments gestuels, des attributs intrinsèques et extrinsèques.

#### **1. attributs intrinsèques**

La définition d'attributs intrinsèques pertinents de la formule structurant une articulation au clavier peut se rapprocher de celle d'attributs intrinsèques pour le segment gestuel provenant d'une touche de clavier.

Les déplacements d'une touche de clavier sont limités dans la direction verticale à trois mouvements : descente, maintien, montée. Tout mouvement d'une touche peut donc se décrire par un cycle de trois phases, chaque phase étant caractérisée par un sens d'évolution constant. Dans la première partie de ce document, nous avons proposé un cas particulier de cette généralité (enfoncement-maintien-amortissement) mais néanmoins très caractéristique du maniement d'une touche. Nous en avons ensuite déduit des attributs pertinents.

Pour l'articulation, on peut envisager le problème sous le même angle.

Les déplacements sur un clavier, du point de vue de la sélection de ses touches, sont limités dans la direction horizontale à trois sens : croissante (de la gauche vers la droite), stationnaire et décroissante. Toute articulation peut donc être représentée à un niveau élémentaire comme une combinaison de groupes d'événements, chaque groupe étant caractérisé par un sens d'évolution constant.

Remarquons une première difficulté qui vient du fait que l'articulation peut être polyphonique<sup>1</sup>. Remarquons également qu'il peut exister plusieurs manières de grouper un ensemble d'entités gestuelles tel que chaque groupe se caractérise par un sens d'évolution constant.

Dans cette approche neutre, nous proposons trois attributs intrinsèques caractérisant un groupe d'entités gestuelles : la durée, la largeur et la surface occupée par l'ensemble des segments du groupe sur la tablature.

Cette surface peut être représentée par un polygone convexe ou plus simplement par un rectangle englobant l'ensemble des entités désignées. Ces trois attributs sont à rapprocher dans le principe des attributs de durée, d'ambitus et de pente d'un segment-actif.

Rappelons cependant bien la nuance des deux objets (formule et segment) : le segment est constitué d'un signal échantillonné à fréquence constante (chaque échantillon est séparé du précédent par une durée constante) ; une formule est constituée de segments combinés de manière arbitraire dans l'espace géométrique de l'instrument et le temps (chaque segment est séparé des autres segments de manière arbitraire).

Le pouvoir de représentation des attributs de durée, de largeur spatiale et de surface est donc moins fort que celui des attributs du segment-actif.

## **2. attributs extrinsèques**

Les attributs extrinsèques d'une formule sont ceux qui la décrivent de l'extérieur, dans le contexte de représentation de la tablature-étendue.

Comme pour le segment, on appellera "origine" d'une formule (ou d'une sous-formule) un couple (C, T) tel que C représente le numéro de canal et T son instant d'apparition. Cette origine est relative à l'origine de la formule dans laquelle elle se situe.

---

<sup>1</sup> l'utilisation (abusives) du terme "polyphonique" est un raccourci pour évoquer que plusieurs événements gestuels peuvent apparaître simultanément sur des canaux différents

Pour la formule, et contrairement au segment, la détermination de l'origine n'est pas immédiate. Il faut en fait choisir un point (C, T) et attribuer à ce point le statut de "origine de la formule". Plusieurs possibilités, plus ou moins naturelles, sont envisageables.

Par exemple, pour conserver la cohérence avec l'origine des segments, on peut décider que l'origine de la formule sera l'origine de l'un de ses segments.

On peut également décider que l'origine de la formule correspond à un point du polygone convexe englobant le groupe ou, s'il s'agit plus simplement d'un rectangle, un de ses sommets.

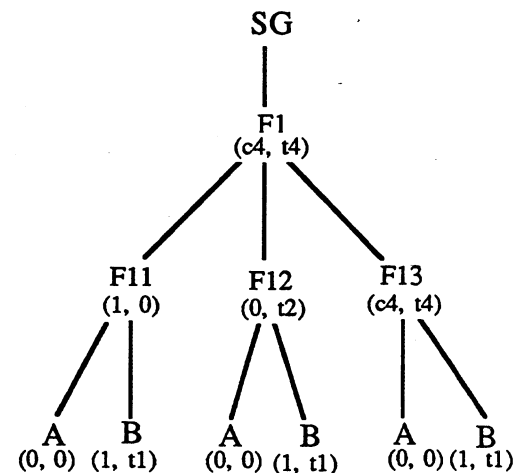
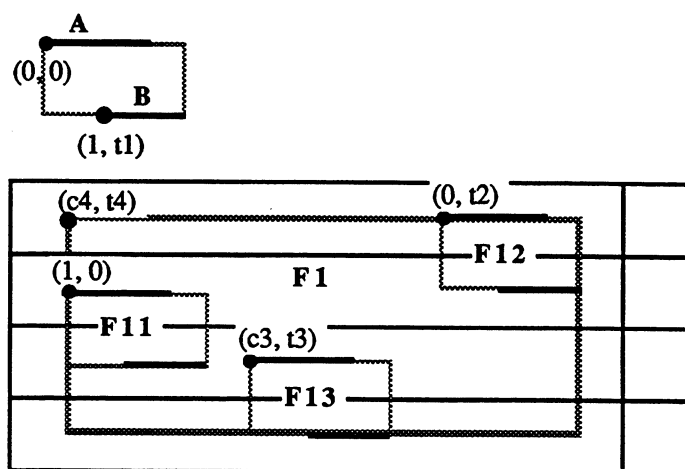
**IV.3. structure**

La grammaire suivante est une interprétation possible de la notion de formule. Fp désigne une formule placée, F une liste de formules placées ou un segment S et Ofml l'origine relative de la formule .

Fp ::= Ofml . F  
 F ::= List\_Fp | S  
 List\_Fp ::= Fp . List\_Fp | Fp  
 S ::= segments-gestuels

**1. exemple**

Nous donnons dans ce paragraphe un exemple de formule. On la représente graphiquement sur une tablature-étendue et sous la forme logique d'une arborescence.



La formule F1, localisée en (c4, t4) est constituée des trois sous-formules F11, F12, F13 de coordonnées respectivement en (1, 0), (0, t2), (c3, t3). Chacune d'elle contient deux segments A et B de coordonnées relatives (0, 0) et (1, t1).

L'arbre correspondant à cette formule est constitué de l'ensemble des sommets  $\{F1, F11, F12, F13\}$  et des segments A et B.

Les origines (relatives à F1) des sommets F11, F12, F13 sont respectivement  $(C1, T1)$ ,  $(C2, T3)$  et  $(C3, T2)$ . On les inscrit sur les nœuds de l'arbre.

SG est l'origine absolue. Les coordonnées de F1  $(c4, t4)$  sont relatives à cette origine.

## **2. optimisation structurelle**

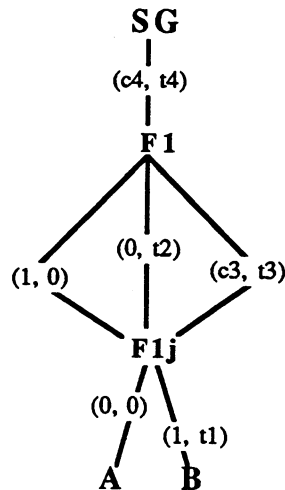
Nous proposons dans ce paragraphe de définir une structure informatique pour la formule qui tienne compte du fait qu'une formule peut avoir plusieurs occurrences d'une même sous-formule ou d'un même segment.

Si l'on crée plusieurs occurrences d'un même objet, comme c'est le cas des segments A et B de l'exemple précédent, ils ne doivent pas être dupliqués en mémoire.

En supprimant la contrainte sur le cycle d'une arborescence, on réalise une structure dans laquelle les occurrences multiples d'un même objet se traduisent par un seul sommet et autant d'arcs que d'occurrences. Il n'y a plus de duplication systématique de l'information pour plusieurs occurrences d'une même donnée.

Cette notion impose de séparer l'information interne et l'information externe d'une formule. Ce qui identifie deux occurrences d'un même objet, c'est le chemin qui les relie à la racine. L'information externe (l'origine relative), n'est pas attachée au sommet, mais à l'arc reliant un sommet à son fils. L'information interne reste associée au nœud.

La figure suivante reprend l'exemple précédent en tenant compte de la redondance d'information. Il ne reste plus qu'une seule représentation du couple (A,B) et des segments A et B. En revanche, il y a 3 arcs reliant F1 à F1j ( $j=1, 2, 3$ ). Chaque arc est porteur des coordonnées relatives d'une occurrence de F1j :  $(1, 0)$ ,  $(0, t2)$  et  $(c3, t3)$ .



#### IV.4. représentation visuelle de la formule

La visualisation d'une formule doit mettre en évidence ses différents attributs et ceux de ses sous-formules : origine, encombrement spatial, structure arborescente, etc...

Pour décrire l'origine d'une formule gestuelle, il suffit dans une échelle temporelle absolue, comme pour le segment, d'une information graphique ponctuelle.

Ce point correspond par exemple à l'origine de l'un de ses segments ou à un point du rectangle les englobant.

Pour mettre en évidence les attributs de durée, de largeur (en nombre de canaux) et de surface, il est suffisant de définir une représentation qui englobe les constituants de la formule (sous-formules et segments) : un polygone convexe ou un rectangle par exemple.

Mais le problème n'est pas résolu de façon rigoureuse dans la mesure où la perception du temps, mais aussi celle de l'espace (largeur et surface) ne repose pas nécessairement sur une loi de proportionnalité.

Pour évoquer la structure arborescente d'une formule, les solutions sont nombreuses. Mais dans notre situation, il s'agit de la représenter à même la tablature-étendue.

On peut pour cela définir deux formes de visualisations : une visualisation détaillée dans laquelle apparaissent tous les événements constituant la formule (ou sous formule) et une visualisation simplifiée qui ne fait apparaître que la surface de l'encombrement. Cette solution a l'avantage d'être simple dans le principe, mais par ailleurs, elle a l'inconvénient de réduire fortement la représentation de l'arborescence à deux niveaux.



On peut également utiliser les couleurs en considérant que plus le segment est enfoncé dans l'arborescence, plus sa couleur doit être sombre. Mais en faisant cela, on met uniquement en évidence les couches de l'arbre et pas sa structure complète.

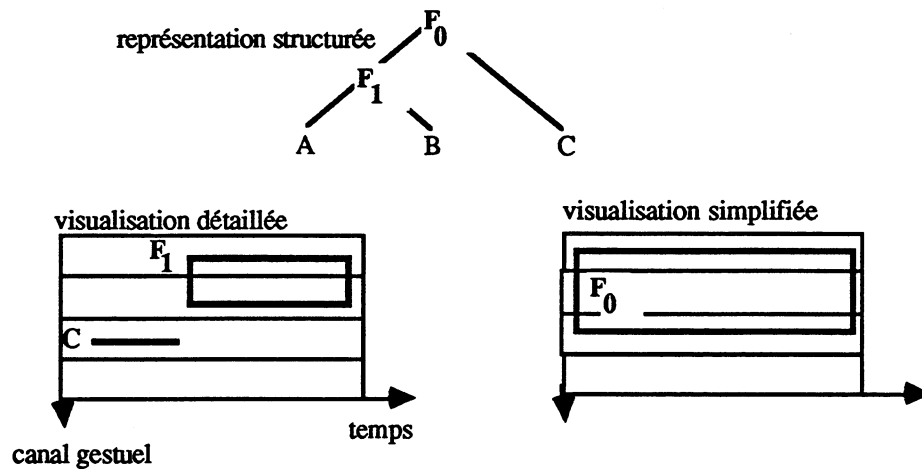


Figure 14. *La formule gestuelle*. La formule permet de lier des événements entre eux. La visualisation simplifiée consiste à traduire deux attributs de la formule : l'encombrant rectangle et l'origine. La visualisation détaillée renseigne sur les descendants immédiats.

## **V. TRAITEMENTS NEUTRES ELEMENTAIRES**

Les traitements neutres doivent permettre de modifier les attributs extrinsèques et intrinsèques des entités gestuelles (segments et formules gestuels). Ils ont cela de caractéristique qu'ils ne conservent pas a priori la logique gestuelle et instrumentale inhérente à l'entité gestuelle manipulée. La seule contrainte qui leur est imposée est de conserver le format des objets manipulés : un signal échantillonné doit rester un signal échantillonné.

Deux approches sont envisageables :

- les traitements s'appliquent directement sur les entités gestuelles. Les attributs sont calculés dans un second temps à partir des entités transformées,
- les traitements s'appliquent directement sur les attributs. Les entités sont reconstruites à partir de ces attributs.

Dans la mesure où notre étude a dû aborder plusieurs problèmes différents, nous avons considéré seulement la première approche : les traitements s'appliquent directement sur les entités.

Nous proposons ici un ensemble de fonctions élémentaires susceptibles de constituer un noyau de traitement de base. Au niveau intrinsèque, l'utilisateur doit pouvoir intervenir directement sur les signaux constituant les segments-actifs, les créer et les transformer. Au niveau de l'articulation, il doit avoir la possibilité d'organiser les entités gestuelles et de les transformer. Enfin, il doit pouvoir réaliser la structure arborescente d'une formule à partir d'une articulation.

### **V.1. segment-actif**

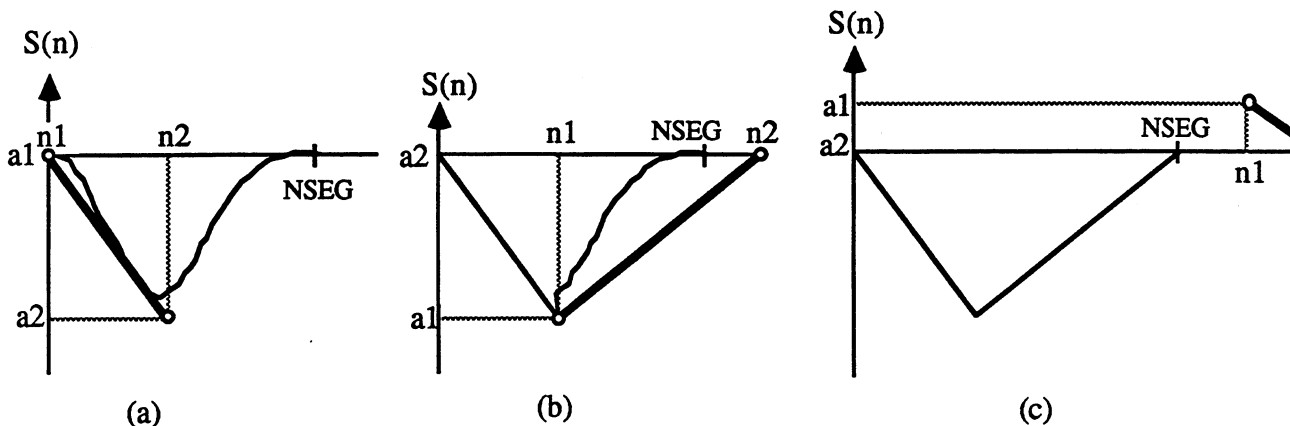
Nous proposons comme solution introductive au problème du traitement du segment-actif une fonction de tracé "manuel" et quelques exemples de transformations pertinentes lesquels permettent d'agir sur les attributs intrinsèques du segment gestuel.

#### **1. tracé "manuel"**

Quelque soit le système, le problème de tracé "manuel" d'un signal amplitude-temps revient à définir une succession de segments de droites déterminés chacun par deux points  $P1=(n1, a1)$  et  $P2=(n2, a2)$ , avec  $n2 \geq n1 \geq 1$ .

Appelons S le signal constituant un segment-actif 1D de longueur NSEG.

D'un point de vue des traitements informatiques, trois cas de tracé significatifs sont à distinguer :



- cas (a) : P1P2 est situé dans l'intervalle du signal existant ( $n1 \leq NSEG$  et  $n2 \leq NSEG$ ). Les échantillons de  $S(n)$  sont substitués par le trait P1P2.
- cas (b) : P1P2 est en intersection avec la fin du signal ( $n1 \leq NSEG$  et  $n2 \geq NSEG$ ).
- cas (c) : pas d'intersection entre P1P2 et le signal existant ( $n1 > NSEG$  et  $n2 > NSEG$ ). Un premier trait relie le dernier échantillon de  $S(n)$  avec P1, un second relie P1 à P2.

Dans la mesure où un segment gestuel est représenté par un signal de position, la fonction de tracé permet de réaliser toute forme de segments gestuels. Rien ne garantit donc que le signal obtenu soit constitué des trois phases : enfoncement-maintien-amortissement.

## 2. exemples de transformations

Contrairement à la fonction de tracé, les transformations sont des opérations qui s'appliquent nécessairement à un signal existant. Elles permettent d'en modifier l'allure dans le temps ou en amplitude. Certaines fonctions conservent au signal transformé une cohérence partielle ou totale.

Considérons ici 4 exemples : l'affinité et l'inversion en amplitude et dans le temps.

- L'affinité en amplitude consiste à multiplier chaque échantillon du signal par un coefficient  $\alpha_a$ . Son intérêt majeur est qu'elle traduit une augmentation ou une diminution de l'intensité de l'événement gestuel et par extension celle du son produit. Si l'algorithme de simulation de l'instrument vérifie une loi affine, l'affinité de la phase d'amortissement continuera d'être cohérente avec celle produite par l'instrument.

- Pour une affinité dans le temps, le traitement consiste soit en un sur-échantillonnage, soit en un sous-échantillonnage. Appliqué aux phases d'enfoncement et de maintien, cela se traduit par un ralentissement ou une accélération relativement cohérente de l'événement gestuel. En revanche, la cohérence du signal produit à partir de la phase d'amortissement est rompue.
- La fonction-miroir en amplitude consiste à inverser chaque échantillon du signal  $S(n) \rightarrow -S(n)$ . Cette fonction traduit un événement gestuel dont l'enfoncement aurait eu lieu, non pas du haut vers le bas mais du bas vers le haut (ou inversement). Pour peu que l'instrument ait un comportement en amplitude symétrique, cette transformation conserve à l'objet transformé une grande cohérence.
- Pour un miroir dans le temps, les échantillons sont permutés. Le signal produit n'a probablement aucune cohérence ni en phase d'enfoncement (qui a été à l'origine phase d'amortissement), ni en phase d'amortissement (initialement enfoncement).

## **V.2. traitements sur la tablature-étendue**

Le traitement sur la tablature présuppose l'existence d'un pointeur pour sélectionner les entités gestuelles (segments et formules) et autres éléments. Sans revenir sur les raisons qui ont motivé la représentation en perspective de la tablature-étendue, remarquons simplement que la question du pointage et de la sélection n'est pas habituelle.

Ensuite, en considérant les segments comme des formules particulières, les fonctions de traitement effectif doivent permettre d'intervenir directement sur les attributs extrinsèques et intrinsèques des formules.

### **1. pointeur de la partition gestuelle**

Le pointeur, généralement contrôlé par une interface gestuelle (souris, ...), se déplace dans la surface de l'écran : tantôt flèche pour des zones de dialogue (menus, fenêtres, etc.), tantôt barre verticale pour des zones de texte, ou encore petite horloge pour spécifier une attente. Il fournit à l'utilisateur le moyen de désigner une position particulière de l'objet, susceptible d'être utilisée par le système.

D'un point de vue informatique, le pointage d'un objet 2D visualisé dans un espace 2D ne soulève pas de problème particulier. A chaque point de l'objet correspond bijectivement un point de l'écran. Ceci n'est plus le cas lorsque l'objet à désigner est représenté dans un espace de dimension 3D. Plusieurs points de l'objet peuvent correspondre à un seul point écran.

Nous avons envisagé deux façons de pointer la tablature :

- Connaissant à tout instant les paramètres de visualisation ( $R$ ,  $\alpha$ , ...) de la partition-gestuelle, il est possible de calculer facilement la correspondance entre un point désigné à l'écran et le point de la partition-gestuelle visé,
- On utilise un pointeur qui ne se déplace pas sur la surface 2D de l'écran, mais directement sur la partition. Ce type de pointeur perd sa polyvalence, dans la mesure où il ne peut désigner que des points de la tablature. En revanche, il augmente les possibilités de visualisation puisqu'il n'est plus nécessaire de faire des calculs de projection d'un point écran sur la partition.

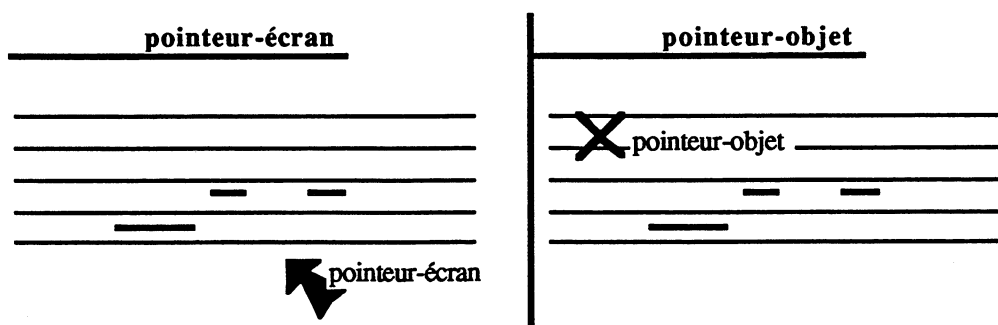


Figure 15. - *Pointeurs-tablature*. Le pointeur-écran se déplace sur le plan 2D de l'écran et le point désigné sur la partition-gestuelle est calculé par des fonctions de projections appropriées. En revanche, le pointeur-objet se déplace à même la partition-gestuelle et les coordonnées du point sont obtenues immédiatement.

## 2. sélection sur la partition gestuelle

L'opération de sélection doit permettre à l'utilisateur de désigner, tous ou en partie, les événements de la tablature et de les intégrer ou de les enlever de la "liste de sélection".

En faisant référence aux systèmes informatiques interactifs classiques, remarquons que la désignation des événements se fait soit par une boîte rectangulaire les englobant, soit par un point les marquant. On y distingue deux modes :

- mode absolu : la liste de sélection est mise à zéro et les événements désignés sont intégrés dans cette liste,
- mode complémentaire : les événements désignés qui ne sont pas déjà sélectionnés le deviennent et inversement.

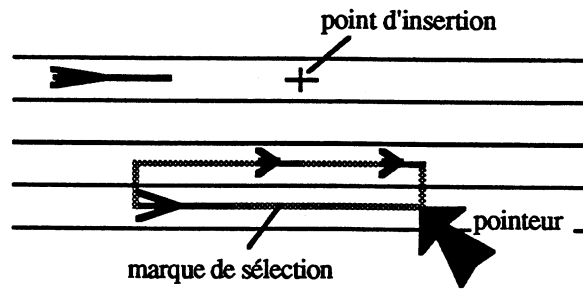
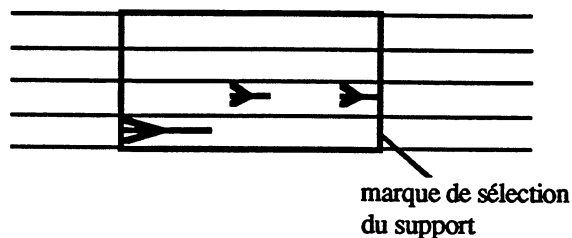


Figure 16. - *La sélection sur la partition.* Le pointeur permet de désigner un point de la partition-gestuelle. La marque de sélection indique à l'utilisateur les objets sélectionnés. Le point d'insertion (ou curseur) indique l'origine de la prochaine insertion.

#### alternative de sélection

Dans notre situation, une alternative de sélection peut s'envisager. Elle consiste à sélectionner, non pas les événements représentés sur la tablature, mais la tablature elle-même. Le traitement s'applique sur la zone du support sélectionnée et par extension, aux événements qu'il contient.

La désignation s'opère sous la forme d'un rectangle de la largeur de la tablature. Les paramètres de désignation sont l'instant de début et l'instant de fin du rectangle.



### **3. traitements "compositionnels" élémentaires**

Les traitements compositionnels élémentaires doivent permettre d'organiser les entités gestuelles (segments et formules gestuels) entre elles selon toutes les configurations possibles. Nous proposons pour cela les fonctions élémentaires classiques : couper / copier / coller / substituer.

La fonction couper supprime les entités sélectionnées de la tablature et les mémorise dans un buffer. La fonction copier duplique les entités sélectionnées et mémorise ces copies dans le buffer. La fonction coller place un duplicata des entités mémorisées à partir du point d'insertion. La fonction substituer remplace chaque entité sélectionnée par un duplicata des entités mémorisées.

**a. mise en œuvre informatique**

Nous avons considéré les segments comme des formules particulières. Appelons **Lf** une liste qui contient les formules représentées sur la tablature.

Pour pouvoir appliquer les traitements élémentaires aux formules, nous définissons deux listes supplémentaires :

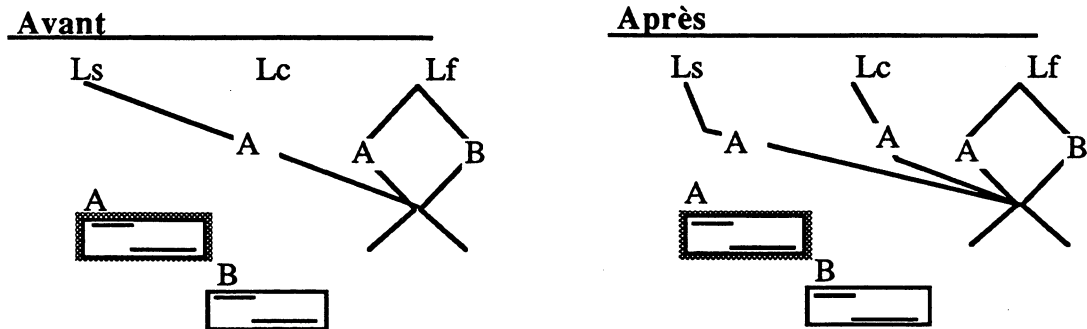
- une liste notée **Ls** qui pointe les formules sélectionnées sur la partition,
- une liste **Lc** qui fait office de "buffer de mémorisation" pour les formules.

**copier**

Après avoir détruit la liste de copie Lc, chaque formule sélectionnée de Ls est dupliquée, puis insérée dans Lc. En prévision des traitements suivants, l'origine commune à l'ensemble des formules copiées est mise à zéro.

Concrètement : il n'est pas nécessaire de dupliquer les descendants des formules sélectionnées. Un simple lien est créé entre la formule sélectionnée et la liste de copie Lc.

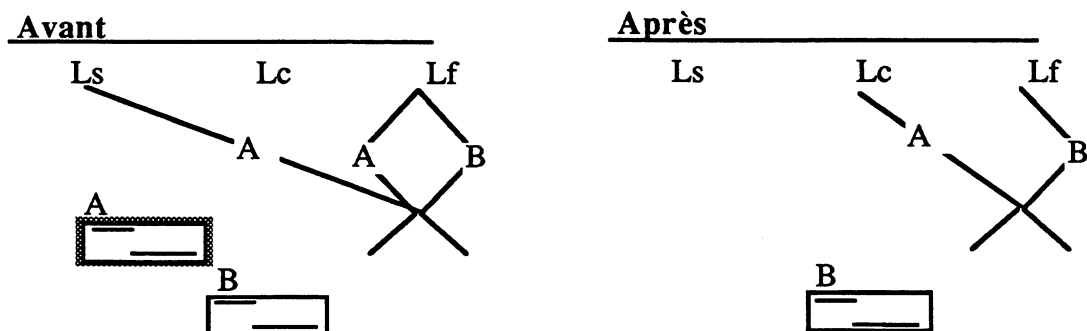
Les formules copiées restent sélectionnées.



**couper**

Chaque formule sélectionnée est supprimée de Lf et est insérée dans la liste de copie Lc. L'origine commune à l'ensemble des formules coupées est initialisée à zéro. Il est possible de couper simultanément plusieurs formules, tout en conservant leur position relative.

Concrètement : il suffit de réaliser les deux opérations suivantes : 1/ les liens des formules avec la liste Lf sont supprimés (les formules coupées ne sont pas détruites) et 2/ la liste de sélection est simplement transformée en la liste de copie.

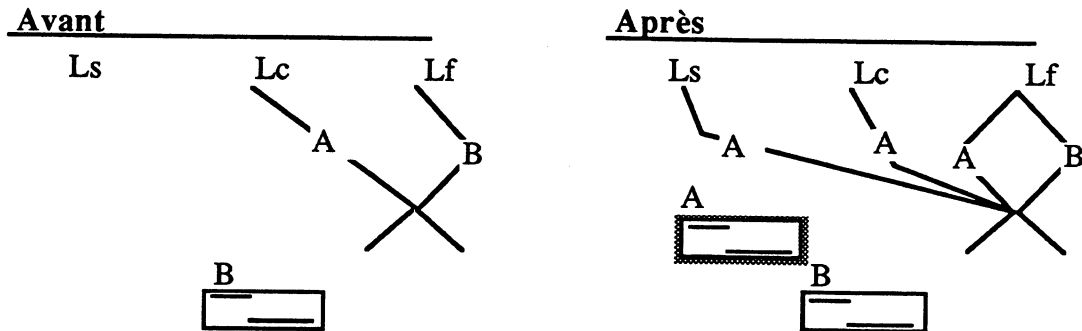


**coller**

Chaque formule de Lc est dupliquée puis insérée dans Lf. L'origine de chaque duplication est mise à la position courante. Si les formules sont à plusieurs niveaux, il n'y a pas de duplication des descendants des formules copiées. Un simple lien est créé.

Les formules collées sont sélectionnées.

Concrètement : il suffit de créer un lien entre l'événement de Lc et la liste Lf.

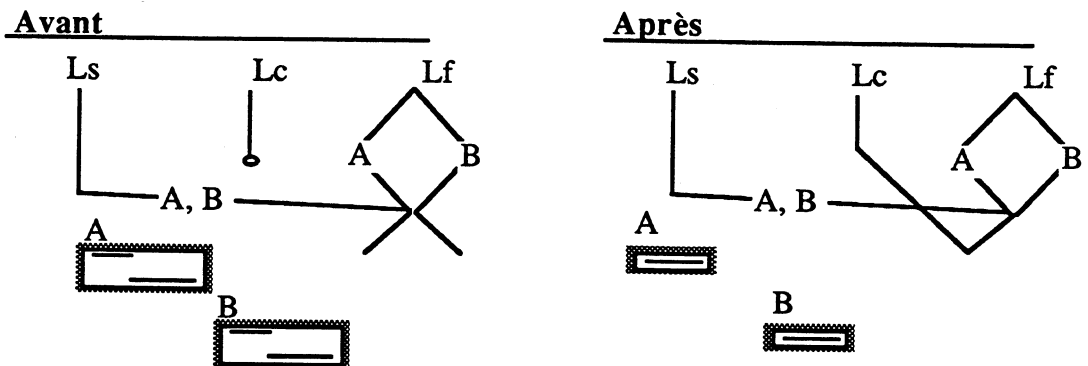


**substituer**

Chaque formule sélectionnée est remplacée par un duplicata des formules de Lc dont l'origine est mise à l'origine de la formule substituée.

Concrètement : au lieu de recréer des formules dupliquées (création et intialisation), on conserve les structures de donnée des formules sélectionnées, qui sont simplement ré-initialisées.

Les événements substitués restent sélectionnés.



**b. remarque sur la taille mémoire utilisée**

Lors des traitements compositionnels, l'espace mémoire requis par les structures-formules évolue uniquement en fonction du nombre de liens créés et des suppressions effectuées dans les trois listes (Lf, Ls, Lc). En effet, les N occurrences d'une même formule adressent cette seule formule par N arcs distincts sur lesquels sont précisés une origine et des attributs externes. Lorsqu'une occurrence supplémentaire est



ajoutée (resp. enlevée), c'est un simple lien qui est créé (resp. supprimé). Remarquons que lors d'une suppression, en plus du lien, c'est toute la structure qui est supprimée si c'était la dernière occurrence.

#### **4. exemples de transformations**

Les fonctions de transformation se distinguent des fonctions de composition élémentaire par le fait qu'elles ne s'appliquent que sur un ensemble d'entités gestuelles existantes. Elles n'ajoutent pas de segments ni n'en suppriment.

Nous proposons ici trois fonctions intéressantes en exemple. Elles permettent d'intervenir sur les attributs intrinsèques (encombrement) et extrinsèques (origine) des formules.

Notons A une articulation constituée de N segments-actifs, localisés chacun par une origine spatiale ( $O_c, O_t$ ).

##### **a. affinité**

L'origine de chaque événement de A est multipliée par un doublet de coefficients réels :

$$(O_c, O_t) \rightarrow (\alpha O_c, \beta O_t) \quad \text{avec } \alpha, \beta \text{ réels.}$$

Lorsque  $\alpha=1$ , il s'agit d'une affinité temporelle pure. Suivant la valeur de  $\beta$  cette transformation se traduit par un ralentissement (dilatation) ou une accélération (compression) de l'ensemble A. On modifie par ce biais le tempo de la pièce. Le rapport des durées reste constant.

Lorsque  $\beta=1$ , il s'agit d'une affinité spatiale pure. Cette transformation se traduit par un écartement ou un rapprochement des événements gestuels dans l'espace.

##### **b. translation**

L'origine de chaque événement de A est translatée par un doublet de coefficients réels :

$$(O_c, O_t) \rightarrow (\alpha + O_c, \beta + O_t) \quad \text{avec } \alpha, \beta \text{ réels.}$$

Lorsque  $\alpha=0$ , il s'agit d'une translation temporelle pure. Cela se traduit par un retard ou un avancement de l'articulation considérée.

Lorsque  $\beta=0$ , il s'agit d'une translation spatiale pure. D'un point de vue musical, cette transformation est équivalente à une transposition. Si deux canaux successifs  $c_1$  et  $c_2$  correspondent respectivement à deux hauteurs  $h_1$  et  $h_2$ , un événement gestuel, translaté de  $c_1$  vers  $c_2$  sera transposé dans l'espace des hauteurs de  $h_1$  vers  $h_2$ .

##### **c. inversion**

L'idée d'inverser une structure n'est pas nouvelle en musique. Selon S.Martin, "inverser les rapports entre les sons apparaît comme une loi fondamentale du langage musical". [Mar78]

Nous avons défini deux fonctions d'inversion : l'une dans le temps et l'autre dans l'espace des canaux. Il s'agit de symétries par rapport à l'axe médian de la dimension concernée. Par exemple, pour l'inversion temporelle, on détermine le milieu  $T_m$  entre  $O_{tmax}$  et  $O_{tmin}$ , instants extrêmes des origines de A :

$$T_m = (O_{tmax} - O_{tmin}) / 2 + O_{tmin}$$

La nouvelle origine de chaque événement de A est donnée par la relation suivante :

$$(O_c, O_t) \rightarrow (O_c, 2 \cdot T_m - O_t)$$

### **V.3. structuration de la formule**

La formule gestuelle est une entité composée permettant de structurer (ou de formuler) les segments en arborescence. On doit pouvoir ainsi mettre en évidence certaines formes d'articulation intéressantes. Nous avons proposé en guise d'introduction à cet exercice des formes neutres telles que les événements la constituant soient organisés dans un même sens (croissant, stationnaire, décroissant).

Le problème de la structuration de l'articulation se rapproche d'un problème de segmentation dans la mesure où il s'agit d'identifier dans une information des sous-groupes. Distinguons cependant bien ces deux opérations : la segmentation opère sur des signaux et consiste à placer des frontières temporelles alors que la structuration, s'applique initialement à un ensemble de segments et consiste à placer des frontières spatio-temporelles.

Ce problème est complexe et nous ne l'avons pas traité dans le sens d'une automatisation, comme nous l'avons fait pour la segmentation.

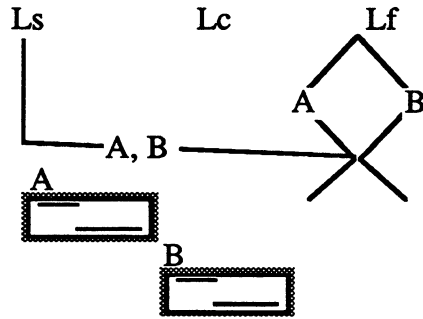
Nous proposons ici deux fonctions qui permettent de structurer une articulation "manuellement". Il s'agit d'une part de la fonction "grouper" qui associe un ensemble d'événements (sous-formules et/ou segments) en une nouvelle formule et d'autre part de la fonction "dégrouper" qui désolidarise une formule en ses constituants. Ces deux fonctions ont un rôle similaire aux fonctions "associer" et "dissocier" utilisées sous Mac-Draw.

#### **1. grouper**

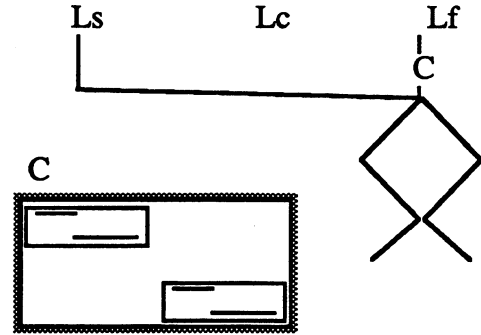
Une formule (Fg) est créée. Son origine est initialisée à l'origine minimale des formules sélectionnées et son encombrement à l'encombrement maximal des formules sélectionnées. Les formules sélectionnées sont enlevées de (Lf) et placées dans (Fg).

La formule (Fg) est insérée dans (Lf) et sélectionnée.

**Avant**



**Après**

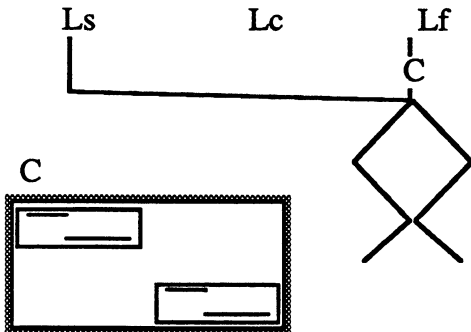


**2. dégrouper**

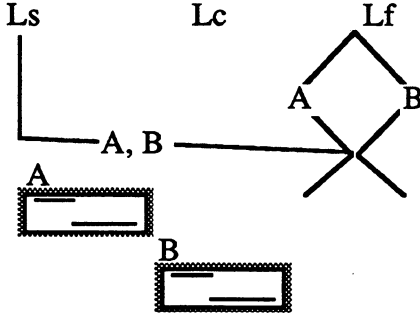
Chaque formule sélectionnée est éclatée en ses sous-formules directes dans Lf. Leurs origines sont mises à jour. Les formules dégroupées sont sélectionnées.

Concrètement : il suffit de supprimer les liens des formules sélectionnées avec Lf ainsi que leurs attributs externes ( C dans la figure) et insérer les fils de cette ex-formule dans Lf.

**Avant**



**Après**



## **VI. CONCLUSION**

Les signaux gestuels captés et mesurés comportent une combinatoire variée de segments.

Nous avons fait de nombreux enregistrements de gestes sur clavier, que nous avons segmentés et traités : nous savons maintenant que les niveaux de représentation et de traitement multiples sont complémentaires dans une approche neutre. Ils assurent le moyen d'analyser finement les actions à différents niveaux de détails et selon différentes échelles.

Le compositeur peut ainsi faire toutes sortes d'expériences, en créant et modifiant les entités gestuelles, sans limites autres que celles imposées par le système, en les restituant sur un instrument donné, et en produisant par là-même les effets acoustiques et visuels correspondants.

Remarquons que nous avons tenu compte de connaissances instrumentales et gestuelles : la tablature est une représentation de l'instrument structuré en unités/canaux/ voies et pas un simple repère spatio-temporel Oxyz ; la forme du segment-actif est déduite d'un événement gestuel caractéristique du jeu sur clavier.

Remarquons également que pour les traitements, et dans la mesure où ils doivent conserver au moins le format des entités manipulées, celles-ci ne peuvent pas enfreindre ni la structure instrumentale représentée par la tablature, ni celle du support amplitude/temps.

Notre approche neutre n'est donc pas restée complètement neutre.

Rappelons enfin que nous avons considéré uniquement la première approche citée en introduction du traitement neutre : les traitements s'appliquent directement sur les entités gestuelles ; les attributs en sont déduits.

Pour compléter cette étude, il sera nécessaire d'aborder le traitement en proposant des fonctions s'appliquant directement, non pas sur les entités gestuelles, mais sur leurs attributs. Remarquons que cette approche soulève au moins deux catégories de problèmes nouveaux. La première concerne l'espace de représentation de ces attributs et la seconde les processus de reconstruction qu'il faut mettre en place pour passer des attributs aux informations exhaustives.



## Edition en référence à l'action

---

### I. INTRODUCTION

### II. APPROCHE COHERENTE DU GESTE INSTRUMENTAL

#### II.1. objectivation et séparation des causes

1. représentations - 2. traitements

#### II.2. un exemple de générateur de mouvements simples

1. modèle mécanique d'un mouvement simple
2. commandes du modèle mécanique
  - a. commandes du mouvement - b. domaine de validité
3. application à quelques mouvements élémentaires
  - a. mouvement balistique simple - b. mouvement balistique avec retour à la position initiale -c. mouvement contrôlé

#### II.3. un exemple de fonction d'interaction

### III. UN ALGORITHME POUR VALIDER UNE ARTICULATION GESTUELLE

#### III.1. le critère de validation

1. doigté et notation traditionnelle
  - a. définition - b. spécification - c. densité
2. doigté et articulation de segments
  - a. "doigté possible" et "meilleur doigté" -
  - b. quantification d'une transition simple -
  - c. quantification du doigté

#### III.2. l'algorithme de validation

1. algorithme
2. quelques solutions au problème de la complexité

#### III.3. exemple d'une fonction de facilité

1. facilité dans le temps - FT
  - a. forme de FT - b. paramètres de FT
2. facilité en amplitude - FA
3. facilité dans l'espace des canaux - FC

a. forme de FC - b. paramètres de FC -

c. matrice de facilité

III.4. conclusion

---

## **I. INTRODUCTION**

L'approche neutre ne donne pas une image du geste qui traduise de manière pertinente la substance de l'information émise. Et si certaines manifestations du geste peuvent s'exprimer par des expressions telles que "tendre les doigts", ou "assouplir le poignet", il n'est pas trivial d'établir une corrélation entre le geste effectué et les formes caractéristiques des informations captées et codées.

Il faudrait maintenant, dans le cadre d'une approche causale, pouvoir établir de telles correspondances et offrir à l'utilisateur une manière de traiter et modifier l'information gestuelle de façon cohérente avec la cause instrumentiste-instrument. Il faudrait également pouvoir inférer des modèles de l'instrumentiste lui-même et de la mécanique de ses organes à partir de l'information dont on dispose qui est l'enregistrement des variables gestuelles (rappelons à ce sujet les travaux réalisés par S.Gibet [Gib87]), de façon à transformer le traitement des signaux gestuels en un traitement des conditions objectives de leurs productions.

Une telle approche est particulièrement difficile, voir impossible dans certaines conditions. Pour être menées complètement, elles demanderaient à être approfondies à partir des bases présentées ici.

Dans ce chapitre, nous allons présenter deux études spécifiques qui ont pu être envisagées grâce à ce contexte et qui correspondent à deux problèmes particuliers qu'il m'a paru intéressant de faire figurer ici.

Ces deux problèmes n'ont pas la prétention de régler le problème de la modélisation.

## **II. APPROCHE COHERENTE DU GESTE INSTRUMENTAL**

### **II.1. objectivation et séparation des causes**

La situation du jeu instrumental met en présence deux systèmes mécaniques : l'opérateur et l'instrument. L'opérateur est d'un point de vue de l'interaction gestuelle un système mécanique actif alors que l'instrument est un objet physique manipulable susceptible de déplacements et de déformations. L'interaction entre les deux systèmes peut alors se décrire comme un couplage mécanique.

Nous proposons d'introduire l'idée d'une simulation de l'opérateur de façon à permettre au créateur de pouvoir intervenir sur les mécanismes de production des



phénomènes sensibles. Il s'agit en quelque sorte de mettre en oeuvre les moyens d'une "situation instrumentale simulée".

Mais en fait, ce que nous présentons réellement est un substitut de l'approche véritablement causale : au lieu d'inférer le modèle de l'opérateur à partir des signaux existants, on s'autorise à construire des modèles de l'opérateur indépendants des signaux gestuels captés et à faire agir ces modèles synthétiques et simplifiés sur les mêmes simulations d'instrument que celles utilisées pour un jeu instrumental réel. L'intérêt de cette approche est d'avoir la possibilité de produire, ou de retrouver, des types de gestes donnés, qu'ils soient arbitraires ou réels.

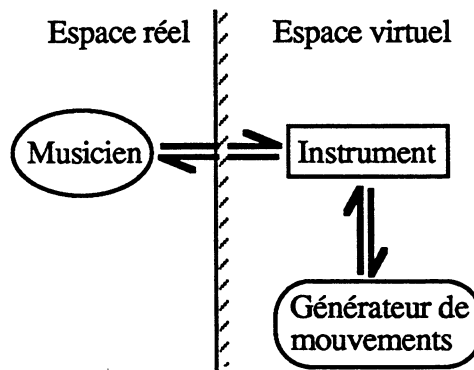


Figure 1. Dans la situation instrumentale initiale, l'instrument est en couplage mécanique avec un musicien réel. Dans la situation instrumentale simulée, les fonctions d'excitation sont prises en charge par un "générateur de mouvements".

Cette approche implique une distanciation par rapport à la situation de jeu instrumental direct. Le geste instrumental n'est plus représenté par un signal de commande de l'instrument mais par un ensemble de commandes applicables au générateur de mouvements.

### 1. représentations

Dans le prolongement de la construction d'objets instrumentaux avec le système CORDIS, l'outil de composition doit permettre de construire des générateurs de mouvements de toute pièce. A cet effet, il est intéressant de rendre similaire la représentation du générateur et des objets instrumentaux. Les deux structures (instrument et générateur de mouvements) sont alors définies comme des objets physiques au sens CORDIS et sont donc vues comme des assemblages de modules physiques et de modules fonctionnels. Dans le contexte de la simulation mécanique, la "synthèse des gestes" peut alors être englobée dans le système de synthèse des objets instrumentaux.

Remarquons que cela n'apporte rien à la base du point de vue de la modélisation de l'instrument puisqu'il faut alors considérer l'ensemble opérateur-simulé / instrument-simulé, ni plus ni moins que comme une simulation d'instrument dont une partie joue un rôle particulier.

Pour permettre des interactions cohérentes avec celles d'une situation instrumentale réelle, chaque structure doit disposer de zones de contact : des zones de contact sur l'instrument et des zones de contact sur le générateur de mouvements. L'interaction mécanique entre les deux structures n'est faite que lors du couplage de ces zones de contact. La forme de l'interaction doit donc également être étudiée et définie.

Ainsi, lorsque le générateur est mis en mouvement, par application de commandes paramétriques ou structurelles, l'instrument est susceptible d'être déplacé et déformé et par conséquent, de produire les différentes manifestations sensibles qui lui sont propres.

Remarquons enfin que deux formes de représentation opposées mais complémentaires sont prises en compte : la représentation structurelle (à temps implicite) et la représentation événementielle (à temps explicite).

Une représentation à temps explicite associe à chaque instant un état ou une valeur unique et prédéterminée. Il s'agit par exemple d'un signal de commande amplitude / temps.

En revanche, une représentation à temps implicite ne prédétermine pas l'état ou les valeurs du système. Par exemple, l'évolution d'un ressort modélisé sous la forme d'une équation récurrente

$$x(n) = a \cdot x(n-1) + b \cdot x(n-2) + Fe(n)$$

dépend des paramètres  $a$ ,  $b$ ,  $c$ , de l'état initial  $x(n-1)$  et  $x(n-2)$  et des perturbations extérieures  $Fe(n)$ .

## **2. traitements**

Le créateur doit pouvoir intervenir à plusieurs niveaux :

- l'instrument est vu comme un objet physique, représenté par un assemblage de différents modules, traduits sous la forme d'algorithmes numériques. Les traitements doivent donc permettre de créer et de modifier ces réseaux de modules. On retrouve la problématique posée dans l'activité de lutherie du système CORDIS,
- de façon similaire, le générateur de mouvements sera vu d'abord comme un objet physique, que les traitements doivent permettre de créer et modifier. Ces

traitements peuvent être arbitraires, ou s'intégrer dans une logique de modélisation de l'opérateur humain,

- le geste instrumental, au sens de la simulation, se décrit comme l'ensemble des mouvements du générateur. La maîtrise de ses mouvements au cours du temps implique un contrôle possible d'une part au niveau de sa structure mécanique et d'autre part au niveau de ses paramètres.

Ainsi, des modifications sur la micro-structure des mécanismes producteurs pourront influencer sur la nature micro-structurale du son produit, et agir au niveau même du timbre.

Les différents traitements se répartissent donc en deux groupes :

- 1- les traitements événementiels qui consistent à spécifier explicitement le temps dans les commandes,
- 2- les traitements structurels qui opèrent soit directement sur les structures et en dehors du jeu, soit par le biais de commandes, dont le déclenchement est assuré par un traitement événementiel.

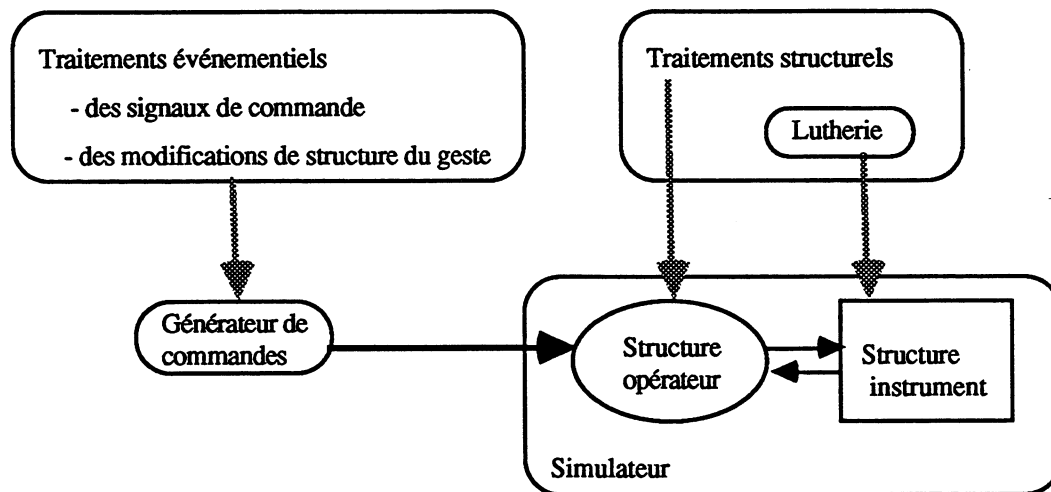


Figure 2. - *Traitements des causes*. Les causes gestuelles sont représentées structurellement et les traitements sont soit événementiels (à temps explicite) soit structurels (à temps implicite).

## II.2. un exemple de générateur de mouvements simples

Pour illustrer le principe de l'approche cohérente, nous proposons ici un exemple de générateur de mouvements simples (un mouvement simple est un mouvement

monoarticulaire). Il s'appuie sur les travaux de P.Perrier & Co [PAK89] qui ont tenté de modéliser le mouvement du point le plus haut de la langue sous la forme d'un système de masses et de ressorts.

Bien que ce modèle soit très simple, voir simpliste en comparaison des possibilités de modélisation offertes par le système CORDIS, il constitue en soi un intérêt dans la mesure où les dits travaux sont orientés dans le sens d'une modélisation de l'opérateur et de ses comportements.

### 1. modèle mécanique d'un mouvement simple

Le modèle proposé est un modèle du second ordre distribué, qui autorise une représentation symétrique des rôles des muscles agonistes et des muscles antagonistes. Seuls les mouvements dans la direction horizontale sont considérés.

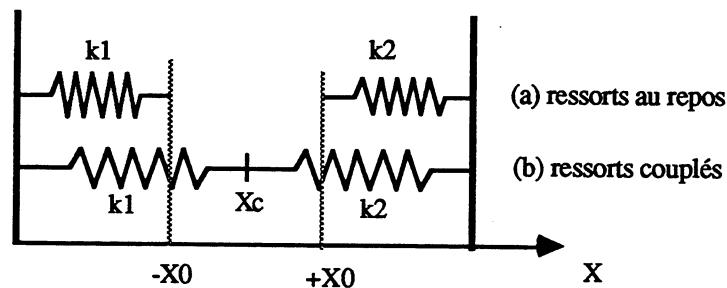


Figure 3. Modèle du second ordre distribué.

L'équation du mouvement sur l'axe Ox est la suivante :

$$m \cdot \frac{d^2 x}{dt^2} = -(k_1 + k_2) \cdot x - (z_1 + z_2) \cdot \frac{dx}{dt} + (k_2 - k_1) \cdot x_0 + f_e \quad (1)$$

où  $k_1$  et  $k_2$  (raideur),  $z_1$  et  $z_2$  (viscosité) et  $x$  (déplacement) sont des fonctions du temps.  $X_0$  est la longueur initiale des ressorts sur l'axe X.

Les choix de la modélisation portent sur :

- le statut du temps dans les programmes moteurs (implicite ou explicite),
- la nature des cibles (spatial ou équilibre),
- les paramètres de contrôle : force, longueur, raideur, viscosité,
- le contrôle de ces paramètres :
  - 1/ [Coo79] propose une action sur la raideur pour passer d'une cible à l'autre,
  - 2/ [Fel80] propose une action directe sur le seuil d'activation et sur la raideur.

Les orientations clés de la modélisation sont les suivantes :

- contrôle de la composante temporelle,

- spécification des cibles-équilibres,
- contrôle du paramètre de raideur.

## **2. commandes du modèle mécanique**

Le modèle étant posé, le mouvement se décrit comme le résultat d'une succession de commandes au niveau des cibles-équilibres, exécutées avec des contraintes mécaniques (raideur) et temporelles.

Une cible est définie comme un état d'équilibre entre les forces agonistes et antagonistes. Une position d'équilibre donnée  $X_c$  est donc atteinte lorsque la vitesse et l'accélération sont nulles dans l'équation (1). Le rapport des raideurs  $k_1/k_2$  vérifie alors la relation :

$$L = \frac{k_1}{k_2} = \frac{(x_0 - x_c)}{(x_0 + x_c)} \quad \text{lorsque} \quad \frac{d^2x}{dt^2} = \frac{dx}{dt} = 0 \quad (2)$$

Le passage d'une cible à une autre peut s'effectuer dans 4 conditions différentes :

- $k_1$  constant,  $k_2$  variable,
- $k_1$  variable,  $k_2$  constant,
- $k_1$  et  $k_2$  variables,  $k_1+k_2$  constant ( $k_1+k_2$ ) est la raideur de co-contraction,
- $k_1$  et  $k_2$  variables,  $k_1+k_2$  variable.

### **a. commandes du mouvement**

La priorité a été donnée à la composante agoniste ( $k_1$ ), qui dans la modélisation impose la raideur de co-contraction ( $K_{co}=k_1+k_2$ ) finale.

Dans ces conditions, la définition d'une cible est faite en donnant ses caractéristiques mécaniques et temporelles :

1- caractéristiques mécaniques :

- $L=k_1/k_2$  (dédit de la position d'équilibre souhaitée  $X_c$ ),
- raideur agoniste  $k_1$ ,
- raideur de co-contraction  $K_{co} = k_1+k_2$ .

2- caractéristiques temporelles :

- temps de transition entre cibles,
- temps de tenue de la cible.

### **b. domaine de validité**

Pour la simplicité des calcul, on pose  $z=z_1=z_2$ .

Le domaine de validité des coefficients  $K_{co}$  et  $z$  se définit de la même manière que pour un oscillateur élémentaire [Rao91].

Posons  $K = K_{co}/(n.Fe^2)$  et  $Z = 2z/(m.Fe)$  où  $Fe$  représente la fréquence d'échantillonnage.

Les solutions convergentes de l'équation (1) sont données par  $K < 1$  et  $Z < 1$ .

- lorsque  $z$  augmente, l'enveloppe d'amortissement diminue,
- lorsque  $K_{co}$  augmente, la période d'oscillation augmente.

### **3. application à quelques mouvements élémentaires**

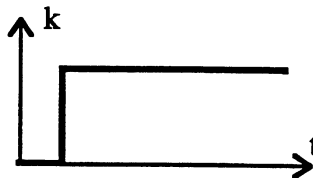
Le mouvement se décrit donc comme le résultat d'une succession de commandes au niveau des cibles-équilibres, exécutées avec des contraintes mécaniques (raideur) et temporelles.

Nous considérons ici trois exemples de mouvements élémentaires que nous nous proposons de retrouver à partir du modèle du second ordre.

#### **a. mouvement balistique simple**

Le mouvement balistique simple se caractérise par une position cible que la structure génératrice doit atteindre d'un seul tenant, sans qu'aucun contrôle dans le même temps ne modifie sa trajectoire.

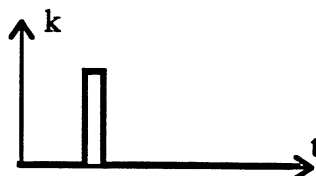
Pour réaliser ce mouvement avec notre modèle, il faut spécifier la cible-équilibre  $X_c$ , puis calculer la raideur agoniste  $k_1$ . La variation instantanée de la raideur amène naturellement la masse du système à la position souhaitée.



#### **b. mouvement balistique avec retour à la position initiale**

Le mouvement balistique avec retour à la position initiale se caractérise par une position cible, que la structure génératrice doit tenter d'atteindre d'un seul tenant (comme le mouvement balistique simple). Après un certain temps, elle revient à la position de départ.

Pour réaliser ce mouvement avec notre modèle, on calcule  $k_1$  avec la position d'équilibre souhaitée. La commande se réalise en deux temps : d'abord la valeur agoniste est mise à jour et quelques instants après, l'ancienne valeur de  $k_1$  est restaurée.



### c. mouvement contrôlé

Le mouvement contrôlé se caractérise par une position cible, que la structure génératrice doit tenter d'atteindre dans un certain temps donné.

Pour réaliser ce mouvement avec le modèle il faut définir la cible  $X_c$  et calculer la raideur agoniste cible  $k_{cible}$ . La cible  $X_c$  est atteinte après une durée de transition  $t_t$  ( $t_t \geq \text{SeuilLent}$ ). On peut entre autres réaliser une interpolation linéaire entre les deux coefficients  $k_1$  et  $k_{cible}$  :

$$k = \frac{(k_1 - k_{cible})}{t_t} \cdot t + k_1$$



### II.3. un exemple de fonction d'interaction

Dans le système CORDIS, l'interaction entre deux structures mécaniques doit être pris en compte par une fonction d'interaction. Il en va ainsi pour le générateur de mouvement et l'instrument.

La fonction d'interaction est en fait un cas particulier de liaison temporaire. On peut traiter ce problème en terme de collision : cela revient à ajouter entre les deux objets (instrument et générateur de mouvements) un élément de liaison (ressort par exemple). Il s'agit donc implicitement d'une modification structurelle de la scène instrumentale.

Nous donnons à l'interaction mécanique trois caractéristiques :

- un effet élastique,
- la non pénétrabilité de la matière : il existe une force, fonction de la distance entre un point et l'instrument, qui s'oppose à la pénétration de l'un dans l'autre. Cette force est d'autant plus forte que la distance de pénétration augmente,
- le mouvement du point se réalise au-dessus de l'instrument.

Citons pour exemple une fonction de répulsion : tant que la distance ( $\Delta d$ ) signée entre la masse du générateur et l'instrument est positive, la force exercée égale 0. Dès que cette distance devient négative, la force exercée est  $F = -k \cdot \Delta d$  sur le générateur et  $-F$  sur l'instrument.

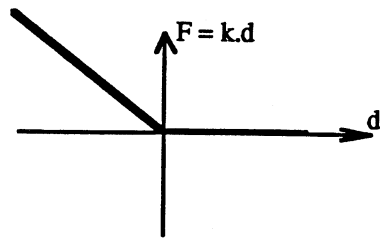


Figure 4. - Fonction d'interaction.



### **III. UN ALGORITHME POUR VALIDER UNE ARTICULATION GESTUELLE**

Une approche complètement causale qui permette à l'utilisateur de traiter et modifier l'information gestuelle de façon cohérente avec la cause instrumentiste-instrument est difficile à mettre en œuvre, voir impossible.

Mais entre une telle approche et une approche neutre, il existe un intermédiaire sur lequel on peut s'arrêter et qui consiste à donner une certaine "validation" des entités gestuelles issues d'un traitement neutre par des considérations sur la cause.

Nous nous intéressons ici tout particulièrement aux articulations gestuelles correspondant à un jeu instrumental sur clavier. Il s'agit de proposer à l'utilisateur qui travaille dans un contexte d'édition neutre, un algorithme pour évaluer la cohérence de l'articulation créée ou modifiée, selon certains critères prédéfinis, avec des causes gestuelles et instrumentales.

#### **III.1. le critère de validation**

La démarche que nous adoptons repose sur un aspect particulier du geste instrumental et consiste à définir un critère pour choisir les doigts capables de réaliser l'enchaînement des événements gestuels sur un clavier de piano.

##### **1. doigté et notation traditionnelle**

###### **a. définition**

Dans la notation traditionnelle, le doigté est défini comme le choix des doigts à employer pour l'exécution d'un morceau sur un instrument à clavier, à cordes frottées ou pincées, à trous, à clés ou à pistons.

Ce choix, dont l'importance est considérable pour faciliter le déchiffrage est souvent suggéré à l'instrumentiste par des annotations chiffrées. Pour les instruments à clavier, chaque doigt de la main porte un numéro : 1 pour le pouce, 2 pour l'index, 3 pour le majeur, 4 pour l'annulaire et 5 pour le petit doigt. Ces indications sont placées au-dessus ou en-dessous des notes sur une partition.

###### **b. spécification**

"Pour la plupart des instruments, c'est leur conformation qui détermine la manière dont il faut placer les doigts ; mais s'agissant des claviers, elle semble tout à fait arbitraire, puisque l'emplacement des touches est ainsi conçu qu'elles peuvent se jouer

par n'importe quel doigt" [Bac79]. Malgré cela, il semblerait qu'il n'existe selon le même auteur "qu'une manière de bien utiliser ses doigts sur un clavier", ce qui tend à montrer que pour spécifier un doigté, il faut tenir compte de considérations multiples, et en particulier de celles propres à la morphologie de l'instrument et à celles du musicien.

Par exemple, sur un clavier de type piano, "nous ne pouvons jouer avec nos cinq doigts que cinq notes de suite ; on a donc remarqué deux moyens qui permettent de se servir aussitôt de tous les doigts dont on a besoin ; le passage d'un doigt par-dessus ou par-dessous un autre".

### **c. densité**

On comprend que des partitions destinées à des débutants ou à des amateurs soient beaucoup plus chargées de semblables indications que celles qui s'adressent à des professionnels en pleine possession de leur technique instrumentale.

Mais quel que soit le niveau de virtuosité, et dans la mesure où le doigté vient compléter une information suffisamment exhaustive, son indication sur une partition n'est pas nécessaire. Chaque interprète peut le trouver de lui-même : tout le processus d'apprentissage de l'instrument consiste précisément à acquérir la connaissance permettant de décoder une information relative au son en une information gestuelle en rapport avec l'instrument.

La densité des indications de doigté dépend aussi de la nature de l'instrument : faible dans le cas des instruments à vent qui ne laissent à l'exécutant qu'un choix très limité, elle est nécessairement plus forte et plus utile pour les instruments à cordes qui offrent maintes façons d'émettre la même note, et, bien entendu, pour le piano.

## **2. doigté et articulation de segments**

Nous avons défini une articulation comme un ensemble de segments-actifs localisés chacun par une origine spatio-temporelle (C, T).

Considérons plus particulièrement l'ensemble A des origines de ces segments, ordonnés dans le temps et enrichies de l'amplitude initiale.

$$A = \{ m_i = (a_i, c_i, t_i) / (m_i < m_j) \text{ si } (t_i < t_j) \}$$

Considérons par ailleurs les doigts comme un ensemble ordonné dans l'espace et numérotés de 1 à 5 :

$$D = \{ d_i / (d_i < d_j) \text{ si } (X_{d_i} < X_{d_j}) \text{ pour } i, j = 1..5 \}$$

### a. "doigté possible" et "meilleur doigté"

Nous appelons doigté d'une articulation, une séquence  $S_q$  de doigts associés dans l'ordre aux événements triés temporellement de A.

$$S_q ::= S_q.(d_i.m_j) \mid (d_i.m_j)$$

Pour distinguer parmi les doigtés d'une même articulation ceux qui sont "possibles" et les doigtés "non-possibles", nous définissons un critère de facilité telle que la séquence est possible si  $\text{Facilité}(S_q) > 0$  et impossible dans le cas contraire.

Les doigtés dit "possibles" d'une articulation A (non vide) pourront ainsi être classés selon ce critère :

$$\text{doigtés possibles} = \{ S_q / S_q \leq S_p \text{ si } 0 < \text{Facilité}(S_q) \leq \text{Facilité}(S_p) \}$$

Définir le doigté d'une articulation consiste donc à choisir parmi l'ensemble des "doigtés possibles" ce que nous appellerons le "meilleur doigté" : celui de facilité la plus grande par exemple.

### b. quantification d'une transition simple

Un problème intermédiaire à celui du choix du doigté d'une articulation consiste à déterminer une fonction qui indique la facilité avec laquelle une transition entre deux couples  $m_1.d_{m1}$  et  $m_2.d_{m2}$  se réalise.

Cette fonction doit tenir compte de la morphologie de l'instrument, des caractéristiques mécaniques d'une main et des attributs extrinsèques (éventuellement intrinsèques) des segments.

Pour exprimer la facilité d'une transition entre deux événements  $m_1$  et  $m_2$  par deux doigts  $d_{m1}$  et  $d_{m2}$ , nous définissons une fonction  $F(m_1, d_{m1}, m_2, d_{m2})$ , dite "fonction de facilité" telle que :

$$1 - 0 \leq F(m_1, d_{m1}, m_2, d_{m2}) \leq 1$$

2 - plus la valeur de F est proche de 0, plus la transition entre les doigts  $d_{m1} \rightarrow d_{m2}$  est difficile à réaliser ; plus la valeur est proche de 1 plus la transition entre les deux doigts est facile.

### c. quantification du doigté

Pour exprimer la facilité d'une articulation de N événements nous proposons une moyenne et un écart type :

1 - la moyenne  $E(A)$  et le moment d'ordre 2  $E(A^2)$  de l'articulation A

$$E(A) = \sum_{i=1}^{i=N-1} F(m_i, d_{m_i}, m_{i+1}, d_{m_{i+1}}) \cdot p_i \quad \text{avec } p_i = 1 / N-1$$

$$E(A^2) = \sum_{i=1}^{i=N-1} F^2(m_i, d_{m_i}, m_{i+1}, d_{m_{i+1}}) \cdot p_i \quad \text{avec } p_i = 1 / N-1$$

2 - l'écart type  $V(A)$

$$V(A) = \sqrt{E(A^2) - [E(A)]^2}$$

La moyenne et l'écart type représentent globalement le doigté. Cela signifie donc qu'ils ajoutent de l'information à l'objet de référence et en suppriment. Ils traduisent bien ce que l'on appellera la "facilité moyenne" mais décrivent en revanche mal les difficultés fortes locales suffisantes pour rendre un doigté non-possible. Un doigté non-possible est un doigté comportant au moins une transition de facilité nulle ; elle peut être maximum partout ailleurs.

### III.2. l'algorithme de validation

L'algorithme de validation d'une articulation A a pour fonction de décider si elle est réalisable, i.e s'il existe pour tout couple d'événements consécutifs  $m_i$  et  $m_{i+1}$  une transition de doigts possible  $d_{m_i} \rightarrow d_{m_{i+1}}$  telle que la facilité de transition entre ces doigts est non nulle.

1 - si A est réalisable, alors on a fabriqué au moins un doigté, tel que

$$\text{Card}(\text{doigté}(A, D)) = \text{Card}(A) = N.$$

Il existe donc au moins un chemin reliant le premier événement au dernier tel que chaque transition admette un coefficient de facilité non-nulle. Chaque chemin se caractérise par une moyenne  $E(A)$  et un écart type  $V(A)$ . Le nombre de transitions égale  $\text{Card}(A)-1$ .

2 - soit l'ensemble des moyennes de facilité obtenues pour les doigtés possibles de A et  $E_{\max}$  la facilité maximale de cet ensemble. Nous appellerons "meilleurs doigtés" ceux dont la facilité moyenne est  $E_{\max}$ .

### 1. algorithme

L'algorithme de validation fabrique l'arbre de tous les doigtés possibles d'une articulation à partir d'une fonction de facilité prédéfinie F. Sur un chemin de l'arbre, la recherche est stoppée lorsque le coefficient de transition est nul (le doigté est non-possible).

La procédure retourne VRAI lorsque la fin de la liste est atteinte, c'est-à-dire lorsqu'un doigt a été attribué à chaque événement de l'articulation (le doigté est possible).

"Emax" est la moyenne maximale des coefficients de transitions sur un chemin complet.

La procédure suivant(list\_M) retourne l'événement suivant de la liste list\_M. Cette liste est initialement triée temporellement et spatialement.

```

int Validation(      typlist      list_M ;      // liste des événements triés
                   typM  Mpred ;      // événement précédent
                   typD  Dpred ;      // doigt associé à Mpred
                   int   som ;      // somme des coefficients de facilité
                   int   N ;      // nombre d'événements rencontrés
                   )
/*-----*\
\*-----*/
{ int p ;           // valeur du coefficient de transition
  int dcour ;      // doigt courant

  if (list_M=∅)           // Emax est mis à jour en fin de liste
    { Emax=max(Emax, som/N); return(1) }

  for(dcour=1; dcour< NBDOIGT ; dcour++)      // développement du sous-arbre
    { if ( (p = F(Mpred, Dpred, list_M.m, dcour)) != 0) // test du coefficient de transition.
      return ( Validation( suivant(list_M), list_M.m, dcour, p+som, N+1) )
    }
  return(0) ;
}

```

### 2. quelques solutions au problème de la complexité

A est l'articulation constituée de N événements.

La complexité de l'algorithme en nombre d'appels de la procédure dépend

- 1 - du nombre de doigts q,
- 2 - du nombre d'événements de A. La longueur de tous les chemins reliant le premier événement au dernier événement de A est égale à Card(A) - 1,
- 3 - des valeurs de facilité F non-nulles.

Dans le pire des cas (F≠0), le coût de cet algorithme est donc de l'ordre de  $O(q^A)$ .

Par exemple, pour une articulation de n=10 segments, un ensemble de q=5 doigts et une fonction de facilité non-nulle partout, le coût sera de l'ordre de  $10^6$  appels. On

comprend donc que la complexité en nombre d'appels, et par suite en temps d'utilisation CPU et attente utilisateur devient très vite exorbitant.

Nous proposons deux idées de solution pour réduire la complexité de l'algorithme :

- une première solution consiste à limiter le nombre de transitions dont la facilité est non-nulle. De ce fait le nombre d'appels possible est réduit.

Une variante à cette solution est de fixer des seuils de facilité en-dessous desquels une transition est considérée comme non avenante. Ces seuils sont utilisés dans l'algorithme et n'imposent donc pas de simplifier le modèle comme le suggère la première partie de la solution.

- une deuxième solution consiste à fragmenter l'articulation en une succession de groupes qui sont validés séparément. Pour garantir la continuité dans le doigté, les groupes sont reliés comme des maillons de chaîne par les événements des extrémités. Le premier événement de chaque groupe correspond au dernier événement du groupe précédent.

On peut également ne pas préfixer le nombre d'événements de chaque groupe, mais clore un groupe dynamiquement, en fonction du doigt et de la position de l'événement par exemple. Nous avons pour cela quelques indications proposées dans la musique qui concernent par exemple l'utilisation du petit doigt, ou celle du pouce.

Plusieurs méthodes sont donc envisageables pour faire chuter la complexité en nombre d'appels de procédure, et donc en temps d'exécution raisonnablement interactif. Les quelques solutions proposées sont complémentaires dans l'approche de ce problème.

### **III.3. exemple d'une fonction de facilité**

Nous donnons dans cette sous-section un exemple de fonction de facilité. Elle concerne initialement la main droite et opère exclusivement sur des articulations monophoniques. Elle indique pour toute paire de couple  $m_1.d_{m1}$  et  $m_2.d_{m2}$  un coefficient compris entre 0 et 1 qui exprime la facilité avec laquelle la transition se réalise.

Pour diminuer la complexité de cette fonction, nous séparons, la fonction de facilité  $F$  en une composante statique  $F_{\Omega}$  et une composante dynamique  $F_T$  telles que :  $F = F_{\Omega} * F_T$ .

De plus, la morphologie de la main et les comportements possibles des doigts lors d'un jeu sur clavier nous suggèrent de séparer la composante statique  $F_{\Omega}$  en deux composantes : une composante en amplitude  $F_a$  (enfoncement des touches) et l'autre dans la dimension des canaux  $F_c$  (sélection horizontale par écartement des doigts et déplacement latéral de la main).

En définitive,

$$F(m_1, d_{m1}, m_2, d_{m2}) \\ = \\ F_T(t_1, d_{m1}, t_2, d_{m2}) * F_C(c_1, d_{m1}, c_2, d_{m2}) * F_A(a_1, d_{m1}, a_2, d_{m2}).$$

### 1. facilité dans le temps - $F_T$

La composante de facilité dans le temps représente la dynamique d'une transition entre deux doigts et deux instants, quelques soient les positions géométriques de  $m_1$  et  $m_2$ .

Dans une situation instrumentale réelle, le musicien peut accélérer ou ralentir le tempo d'une articulation, c'est à dire, jouer sur la vitesse de transition entre chaque note.

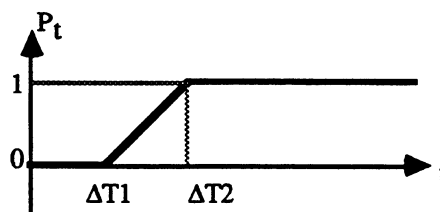
Ceci peut se faire dans une certaine plage dont les bornes sont fonction entre autres de la virtuosité du musicien, de la complexité de la transition et de la morphologie de l'instrument.

On pourrait ainsi montrer qu'il existe pour chaque triplet musicien/ instrument/ transition une limite temporelle inférieure en-dessous de laquelle la transition serait non-possible.

On imagine symétriquement à cette situation une seconde borne, non limitative comme la première, à partir de laquelle la transition soulève une autre difficulté physiologique.

#### **a. forme de $F_T$**

La fonction de facilité temporelle est donc constituée de trois segments de droite : le premier sur l'intervalle  $[0.. \Delta T_1]$ , le second sur  $[\Delta T_1.. \Delta T_2]$  et le dernier sur  $[\Delta T_2.. \infty]$ .



• quels que soient  $a_1, c_1, a_2, c_2$ , il existe un seuil  $\Delta T_1$  en-dessous duquel la transition de doigt de  $m_1$  vers  $m_2$  ne peut être réalisée :

$$F_T(t_1, d_{m1}, t_2, d_{m2}) = 0 \quad \text{pour } t_2 - t_1 < \Delta T_1$$

• quels que soient  $a_1, c_1, a_2, c_2$ , il existe un seuil  $\Delta T_2$  au-dessus duquel, la transition de doigt de  $m_1$  vers  $m_2$  est toujours réalisable :

$$F_T(t_1, d_{m1}, t_2, d_{m2}) = 1 \quad \text{pour } t_2 - t_1 \geq \Delta T_2$$

• s'il est un seuil  $\Delta T_1$  en-dessous duquel la transition n'est plus réalisable et un seuil  $\Delta T_2$  au-dessus duquel la transition est réalisable quels que soient les doigts et positions impliqués, alors il existe très certainement une fonction continue reliant les valeurs correspondant à ces deux seuils. Nous retenons pour des raisons de simplicité de calcul une fonction linéaire.

$$F_T(t_1, d_{m1}, t_2, d_{m2}) = (\Delta t - \Delta T_1) / (\Delta T_2 - \Delta T_1) \\ \text{pour } \Delta T_1 \leq (\Delta t = t_2 - t_1) \leq \Delta T_2$$

### **b. paramètres de $F_T$**

Pour simplifier l'exemple, on ne donne qu'une seule fonction pour la totalité des transitions possibles.

•  $\Delta T_1$  correspond à la durée en-dessous de laquelle, quel que soit le couple de doigts, la transition est non-possible. Cette valeur est le minimum des seuils d'impossibilité. On choisit comme indication de base la bande de fréquence (BHz) de tremblement physiologique des doigts donné [Ber75] : BHz $\approx$ 10Hz.  $\Delta T_1=0.1$ sec.

•  $\Delta T_2$  correspond à la durée à partir de laquelle, quel que soit le couple de doigt, la transition a la facilité maximale. Sa valeur correspond au maximum des seuils vérifiant cette propriété.

On choisit pour  $\Delta T_2$  une valeur arbitraire qui correspond à un temps moyen pour parcourir toute la largeur du clavier de piano. Par exemple  $\Delta T_2 \approx 1$ sec.

### **2. facilité en amplitude - $F_A$**

La composante en amplitude doit montrer la facilité de transition entre deux doigts et deux amplitudes, quel que soit l'écart temporel et l'intervalle en nombre de canaux.



En séparant la composante statique  $F_{\Omega}$  en deux composantes  $F_A$  et  $F_C$ , nous avons pris le risque d'en réduire fortement les possibilités de description exhaustive.

Mais, nous avons fait remarquer plus haut que nous ne pouvons jouer avec nos cinq doigts que sur cinq touches de suite. Qu'elles correspondent à des touches blanches ou à des touches noires, la transition de l'une à l'autre ne suscite pas en première analyse de difficultés incontournables.

Le seul cas qui semble soulever une difficulté est celui du passage d'un doigt par-dessus ou par-dessous un autre, en montant ou en descendant.

Sur un clavier de type piano, quatre cas de transitions sont possibles suivant que l'on passe d'une touche blanche à une blanche, d'une blanche à une noire, etc.

Un test pratique sur un clavier permet de classer la difficulté de chacun de ces cas. Il nous est apparu que, dans le sens de la montée, le plus simple était la transition noire→blanche suivie de blanche→blanche, noire→noire et loin derrière blanche→noire.

Pour mettre en relief la transition la plus difficile, à la fois dans le sens de la montée (blanche→noire) et dans le sens de la descente (noire→blanche), nous facilitons au maximum les trois autres transitions et rendons impossible celle décrivant cette situation.

montée			descente		
	blanche	noire		blanche	noire
blanche	1	0	blanche	1	1
noire	1	1	noire	0	1

### 3. facilité dans l'espace des canaux - $F_C$

La composante dans la dimension des canaux doit montrer la facilité de transition entre deux doigts ( $d_i, d_j$ ) et un intervalle de touches  $\Delta c = (c_2 - c_1)$ , quel que soit l'écart temporel des événements et l'intervalle d'amplitude des touches.

Cette fonction est décisive pour l'évaluation du doigté car elle permet de décrire les différents écarts, les déplacements et les passages que les doigts d'une main sont susceptibles de réaliser.

Nous ne nous intéressons ici qu'aux écarts et aux passages de doigts. Les possibilités de déplacements latéraux (mouvement du bras entraînant la main) ne sont pas traitées.

### a. forme de $F_C$

La conformation de la main rend les possibilités d'écartement entre deux doigts limitées. Les possibilités d'écartements sont représentables par quelques attributs :

- l'écart le plus facile à réaliser tel que  $d_i \leq d_j$ ,
- l'intervalle d'écart maximum tel que  $d_i \leq d_j$ ,
- l'intervalle d'écart minimum tel que  $d_i \leq d_j$ .

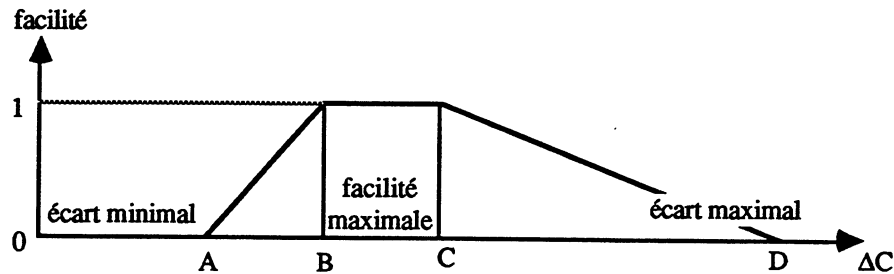


Figure 5. - *Ecarts de doigt*. A chaque couple de doigt est associé une fonction en trois morceaux : OA est l'écart minimal réalisable, OC-OB correspond aux écarts de facilité maximale et OD à l'écart le plus grand réalisable.

### b. paramètres de $F_C$

J'ai pris en référence ma propre main. Les valeurs des paramètres sont évaluées par des mesures statiques directement sur le clavier. Faire des mesures statiques signifie ramener l'intervalle de temps séparant les deux événements à 0. L'évaluation des facilités des écarts et passages de doigts sont faites subjectivement et par comparaisons relatives.

1 - les transitions  $d_i \rightarrow d_i$  (un doigt vers lui-même) ne sont possibles que si les canaux de départ et d'arrivée sont identiques. La facilité est maximum pour  $\Delta c=0$  et nulle ailleurs.

Remarquons que par une mesure statique, on ne peut pas décrire les transitions consistant à sélectionner successivement deux touches différentes avec un même doigt.

2 - pour les transitions  $d_i \rightarrow d_j$  telles que  $i < j$ , nous mesurons l'écart le plus facile à réaliser et les écarts extrémaux.

Le tableau suivant résume les valeurs mesurées des écarts extrémaux et facilités maximales en nombres de canaux.

di->dj	écart min (0A)	facilité max(0B-0C)	écart max (0D)
1->2	0	1-3	12
1->3	0	3-5	13
1->4	0	4-6	14
1->5	0	6-8	15
2->3	0	1-2	5
2->4	0	3-4	6
2->5	2	4-5	9
3->4	0	1-3	4
3->5	2	3-4	8
4->5	0	1-3	4

3 - pour les passages de doigt, Bach remarquait que "le passage par-dessus le pouce se fait soit avec le deuxième, soit avec le deuxième et le troisième, soit avec le deuxième, le troisième et le quatrième doigt ; il se fait également par dessus le quatrième avec le troisième. Nous trouvons également une petite exception, que l'on peut s'autoriser sous certaines conditions, où l'on passe le quatrième doigt par dessus le cinquième."

Nous ne retenons que les passages de doigts où intervient le pouce. Dans le sens de la descente, nous proposons pour chaque couple de doigt (1→3), (1→2) et (1→4) un coefficient constant, indépendant de l'intervalle Δc.

Le tableau suivante résume ces coefficients et leur classement :

passage de doigt	(1→3)	>	(1→2)	>	(1→4)
coefficient de facilité	0.6		0.5		0.4

**c. matrice de facilité**

L'ensemble de toutes les fonctions de facilité dans la dimension des canaux est appelé **matrice de facilité**. Elle est constituée de 5\*5=25 fonctions de facilité.

Nous considérons néanmoins que FC est symétrique :

$$FC(c_1, d_{m1}, c_2, d_{m2}) = FC(c_2, d_{m2}, c_1, d_{m1}).$$

Cela signifie que le coefficient de transition de montée d'un couple  $m_1.d_{m1}$  vers  $m_2.d_{m2}$  est identique à celui de la descente du second  $m_2.d_{m2}$  vers le premier  $m_1.d_{m1}$ . Nous ne décrivons donc que 15 de ces fonctions :  $d_i \rightarrow d_j$  avec  $d_i \leq d_j$ . Les 10 autres fonctions ( $d_i > d_j$ ) sont le symétrique des fonctions  $d_i \rightarrow d_j$ .

Les tableaux suivants représentent l'ensemble des 15 fonctions de facilité dans la dimension des canaux. Chaque ligne décrit une transition  $d_i \rightarrow d_j$  et chaque colonne l'intervalle  $\Delta c = c_2 - c_1$  tel que  $c_1$  est le canal de départ (associé à  $d_i$ ) et  $c_2$  est le canal d'arrivée (associé à  $d_j$ ).

	-1	0	1
1→1	-	1.0	-

	-3	-2	-1	0	1	2	3	4	5	6
1→2	0.6	0.6	0.6	0.6	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	0.89	0.78	0.677
1→3	0.5	0.5	0.5	0.5	0.667	0.83	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	0.875
1→4	0.4	0.4	0.4	0.4	0.625	0.75	0.875	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>
1→5	-	-	-	-	0.167	0.333	0.5	0.667	0.83	<b>1.0</b>

	7	8	9	10	11	12	13	14
1→2	0.57	0.44	0.33	0.22	0.11	-	-	-
1→3	0.75	0.625	0.5	0.375	0.25	0.125	-	-
1→4	0.875	0.75	0.625	0.5	0.375	0.25	0.125	-
1→5	<b>1.0</b>	<b>1.0</b>	0.875	0.714	0.571	0.429	0.286	0.143

	0	1	2	3	4	5	6	7	8	
2→3	-	-	<b>1.0</b>	<b>1.0</b>	0.667	0.333	-	-	-	
2→4	-	-	0.333	0.667	<b>1.0</b>	<b>1.0</b>	0.5	-	-	
2→5	-	-	-	-	0.5	<b>1.0</b>	<b>1.0</b>	0.75	0.5	0.25
3→4	-	-	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	-	-	-	-	
3→5	-	-	-	-	<b>1.0</b>	<b>1.0</b>	0.667	0.333	-	
4→5	-	-	<b>1.0</b>	<b>1.0</b>	<b>1.0</b>	-	-	-	-	

**III.4. conclusion**

Nous verrons dans le chapitre V une évaluation des performances de l'algorithme de validation d'une articulation en comparant les résultats obtenus avec des doigts de référence.

Mais avant cela, il faut rappeler que nous avons réduit le problème initial à différents endroits. Le problème initial consistait à définir un algorithme pour valider une articulation. Nous avons en guise de solution introductive à ce problème, proposé un algorithme qui fait appel à une "fonction de facilité", sorte de modèle de générateur de doigtés. Les réductions faites sur la généralité du doigté sont nombreuses. Les plus importantes sont que :

- 1- l'algorithme ne traite que des articulations monophoniques et le doigté est recherché avec une seule main. Un cas plus complexe serait celui des articulations polyphoniques jouées avec plusieurs mains,
- 2- la validation se ramène à une évaluation de doigté en tenant compte de la position de repos naturel du canal et des attributs extrinsèques des segments (canal et temps). On ne tient pas compte de leurs attributs intrinsèques,
- 3- la fonction de facilité est décomposée en un produit de trois fonctions  $F_A$ ,  $F_C$ ,  $F_T$ . Considérer qu'elles sont indépendantes est un facteur de réduction : par exemple, il n'est pas possible de traduire l'idée simple qu'un même doigt peut sélectionner successivement deux touches différentes.

## Expériences et validation

---

### I. INTRODUCTION

### II. GESTE PERCUSSIF ET SIGNAUX GESTUELS

- II.1. matériel de l'expérience - II.2. exemples de gestes percussifs
- II.3. interprétation des signaux - II.4. conclusion

### III. EXEMPLES DE SEGMENTATION

- III.1. situation pratique
  - 1. critère de pause adapté
  - 2. facteur d'échelle
- III.2. exemples de segmentation
  - 1. percussion simple
  - 2. percussion successives

### IV. ESTIMATION DE L'ALGORITHME DE VALIDATION D'UNE ARTICULATION MONOPHONIQUE

- IV.1. doigtés proposés par [Bac79]
- IV.2. doigtés obtenus avec la matrice de facilité
  - 1. gamme de DO majeur montante - 2. gamme de DO dièse majeur montante - 3. gamme de FA majeur montante - 4. gamme de FA dièse majeur montante
- IV.3. conclusion

### V. VALIDATION D'UNE ARTICULATION

- V.1. décomposition de l'articulation - V.2. mesures 1 à 4 -
  - V.3. mesures 5 à 8 - V.4. mesures 9 à 12 - V.5. conclusion
-

## I. INTRODUCTION

Nous présentons dans ce dernier chapitre quelques expériences réalisées dans le contexte expérimental du système CORDIS et de l'éditeur de geste en cours de développement. Il s'agit en particulier de mettre en évidence certaines caractéristiques des signaux gestuels captés et mémorisés, d'évaluer les processus de segmentation proposés au chapitre II et l'algorithme de validation d'une articulation défini dans le chapitre IV.

Le geste percussif sur clavier aura valeur d'exemple. Le segment-gestuel correspondant est constitué d'une phase d'enfoncement (E) et d'une phase d'amortissement (A).

## II. GESTE PERCUSSIF ET SIGNAUX GESTUELS

Nous proposons dans ce paragraphe un montage matériel qui permet de comparer le mouvement de l'opérateur et celui d'une touche lors de son excitation, et d'en déduire les phases de contact.

### II.1. matériel de l'expérience

Pour comparer les mouvements de l'opérateur et ceux de la touche, il faut pouvoir mesurer simultanément le point de contact-opérateur et le point de contact-touche impliqués dans le geste instrumental.

Comme nous ne disposons pas de moyen pour instrumenter l'homme, nous utilisons le clavier rétroactif modulaire.

Le principe du montage est simple : on choisit deux touches juxtaposées. Celle de gauche est la touche-instrument, celle de droite est saisie par deux doigts de l'opérateur. Il s'arrange pour que le doigt situé au-dessus de la touche déborde légèrement à gauche. De cette façon, lors du geste d'enfoncement (mesuré par la touche saisie), la touche-instrument est entraînée avec le doigt qui déborde légèrement à gauche.

Les mouvements opérateur et instrument sont donc simultanément captés et mémorisés.

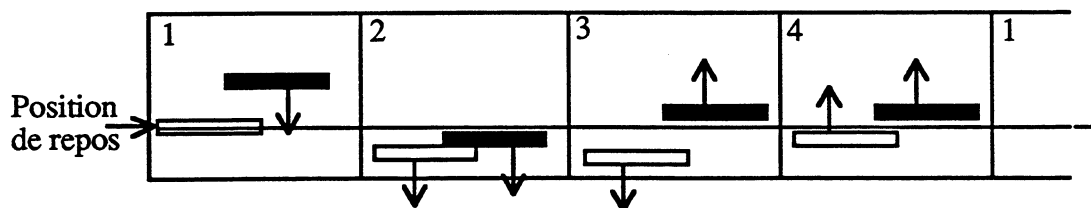


Figure 1. - *Conditions de l'expérience*. La touche-opérateur (noire) en couplage permanent avec un opérateur et toujours au-dessus de la touche-instrument (blanche), permet de comparer de manière synchrone le mouvement de l'opérateur et le mouvement de l'instrument.

## II.2. exemples de gestes percussifs

### figure 2 :

La figure 2 montre une séquence de 3 percussions simples bien distinctes suivies d'une quadruple-percussion. La courbe du haut (voie 14) représente le mouvement de la touche-opérateur et la courbe du bas (voie 13) le mouvement de la touche excitée. La fréquence d'échantillonnage des signaux est de 300Hz. La durée séparant deux graduations est donc de 1 sec.

### figure 3:

La figure 3 présente plus en détail le premier événement percussif de la figure 2. La courbe du haut correspond au mouvement de l'instrument (voie 13), la courbe du milieu à celui du musicien (voie 14). Le repère du bas superpose les deux premières.

L'événement-instrument est constitué de deux phases : une phase d'enfoncement et une phase d'amortissement. L'événement opérateur est constitué de trois phases : une phase de préparation, une phase de projection (enfoncement) et une phase de relâchement (ou amortissement ou freinage).

Le tableau suivant indique les durées et amplitudes des phases constituant ces événements. La course en amplitude du point manipulable est de 2.8 cm.

	événement voie 13	événement voie 14
durée totale :	0.132 s	0.590 s
ambitus :	1.2 cm	1.850 cm
phase enfoncement		
durée :	0.03 s	0.090 s
amplitude :	0.775 cm	1.850 cm
phase amortissement		
durée :	0.102 s	0.333 s
amplitude :	1.2 cm	1.600 cm

La zone de couplage a lieu entre l'instant où la touche de l'instrument (voie 13) quitte sa position de repos, et un instant précédant de peu le deuxième extrémum global de cette même courbe. La durée du couplage est approximativement de 0.07 s.

### figure 4 et 5 :

La figure 4 et la figure 5 sont des agrandissements des événements de la courbe de superposition de la figure 3.



Remarquons que le mouvement de la touche-musicien passe sous la touche-instrument. Cela se comprend par le fait que lors de l'expérience, le choc entre le doigt et la touche-instrument est mou. En considérant que le point du doigt mesuré est situé à l'extrémité digitale, lors du couplage, l'épiderme se comprime et le point mesuré passe sous la touche de l'instrument. Le doigt, lui, reste bien entendu toujours au-dessus.

figure 6 :

La figure 6 détaille les trois premières percussions. Pour chaque mouvement balistique du musicien, on retrouve une phase de préparation, une phase de projection et une phase de freinage. Il faut remarquer d'une part le bruit au repos de l'instrument (courbe du bas) et les phases stationnaires séparant deux gestes balistiques (courbe du haut). Le tableau suivant indique la durée de ces phases de repos.

	<u>courbe opérateur</u>	<u>courbe instrument</u>
repos 1	0.133s	0.722s
repos 2	0.189s	0.700s

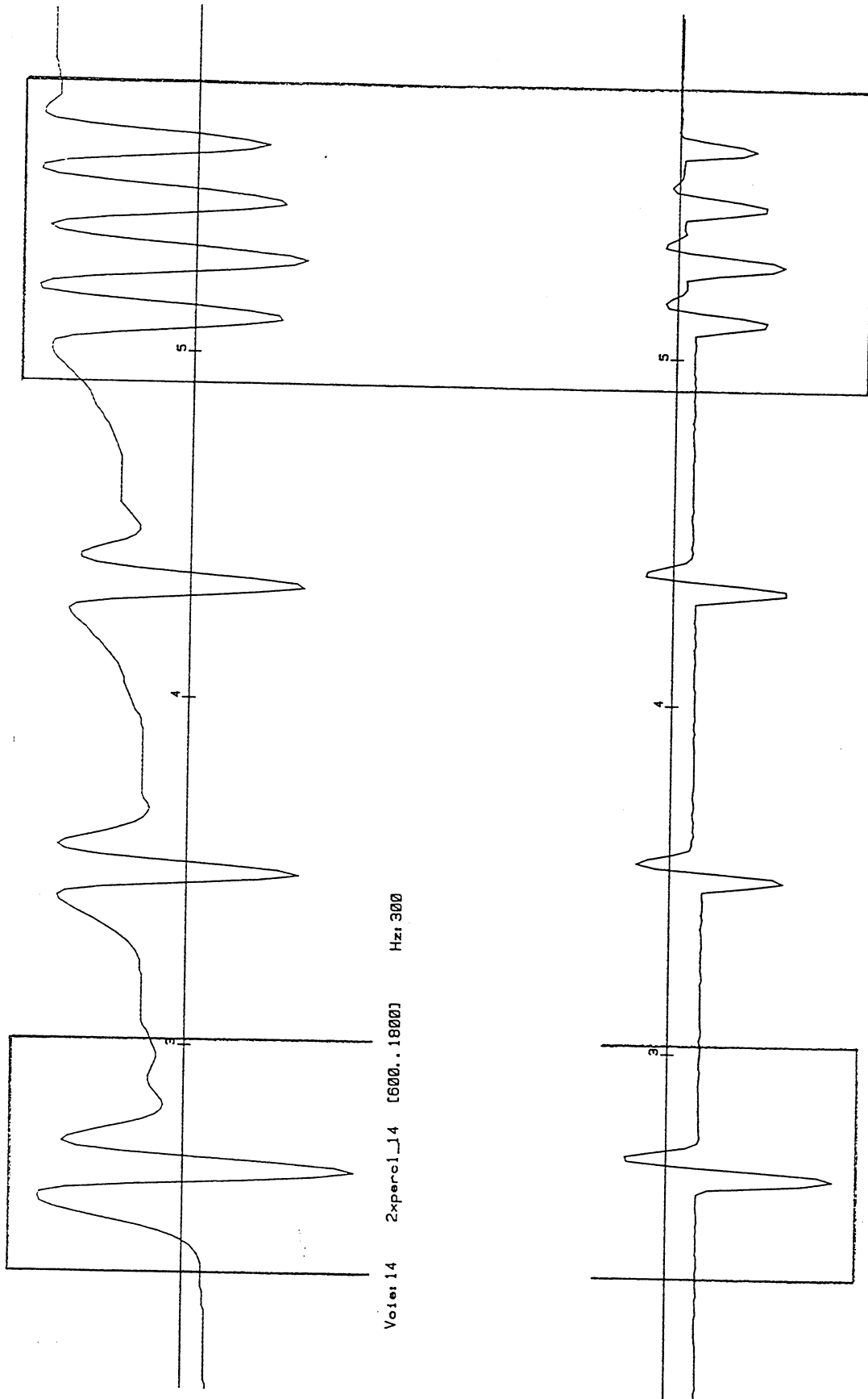
figure 7 :

La figure 7 montre les 4 percussions-successives de la fin de la figure 2. La courbe-opérateur (signal qui part du haut) est constituée de 4 oscillations, d'amplitudes et de périodes presque identiques.

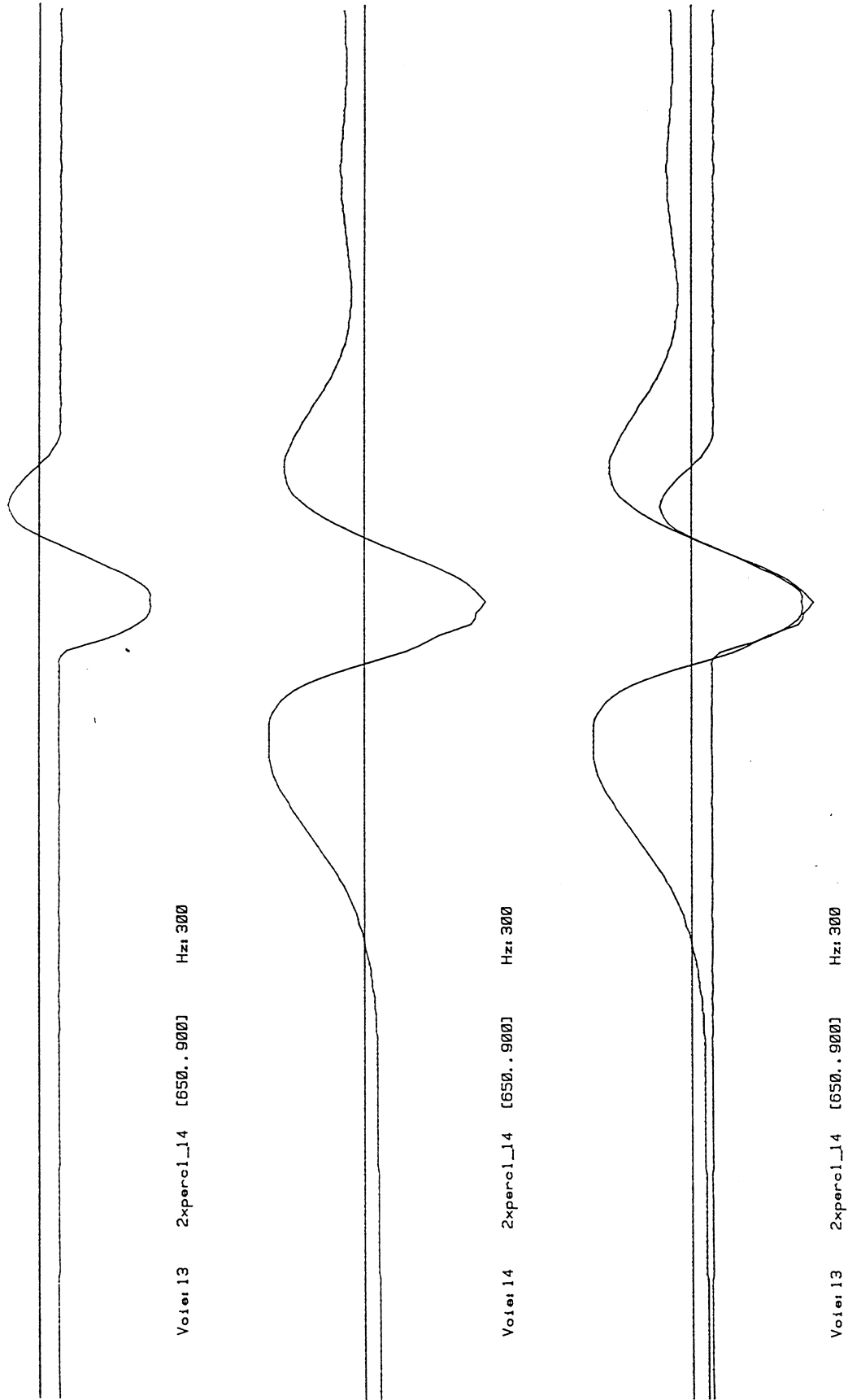
La courbe-instrument met en évidence de petites phases de stationnarité séparant deux percussions successives. Elles n'ont pas une durée suffisante pour être considérées comme des phases de repos. Le tableau suivant récapitule les durées de ces phases.

	<u>courbe opérateur</u>	<u>courbe instrument</u>
repos 1	0s	0.038s
repos 2	0s	0.041s
repos 3	0s	0.048s

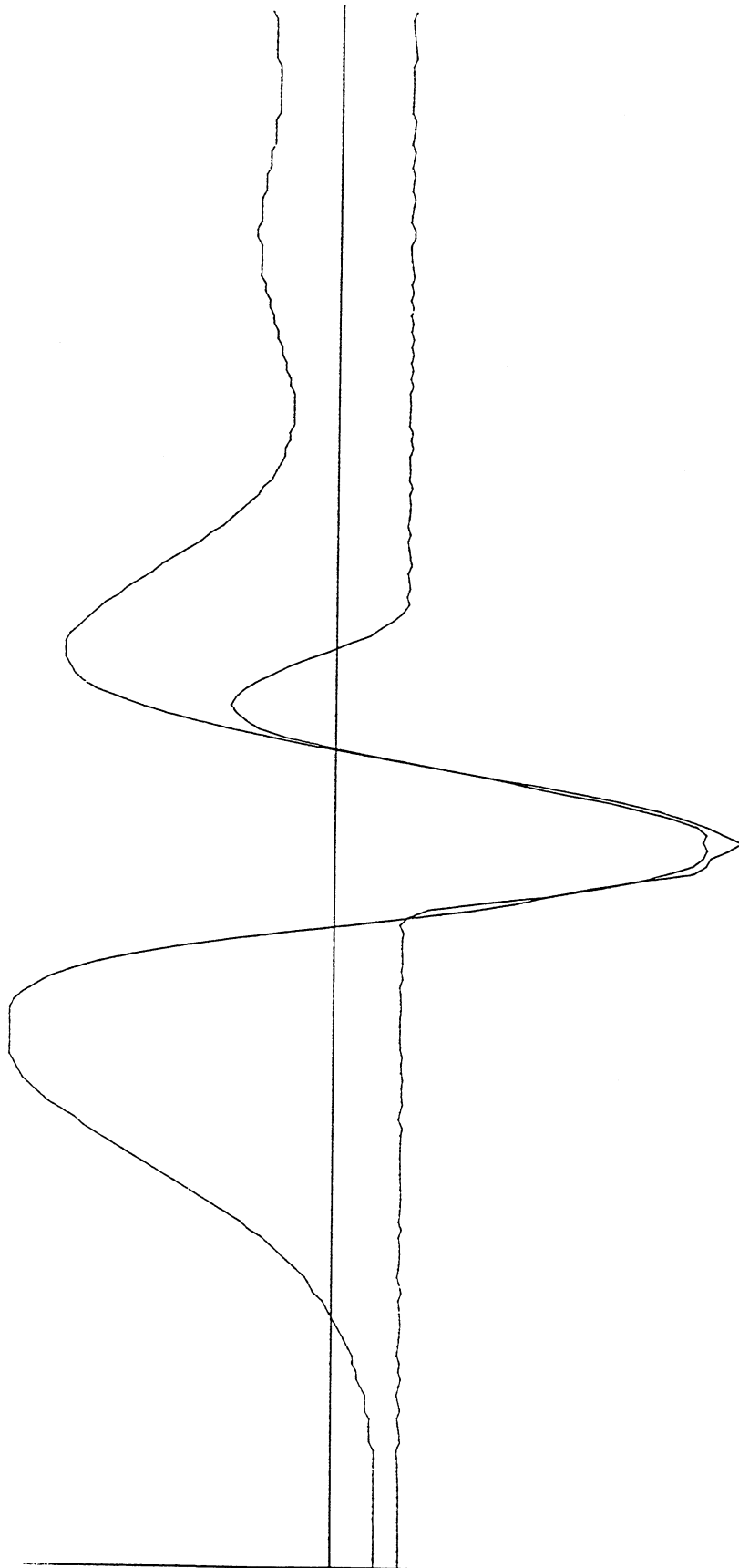
**Figure 2**



**Figure 3**



**Figure 4**



2xperi1\_14 [700..900] Hz: 300

**Figure 5**

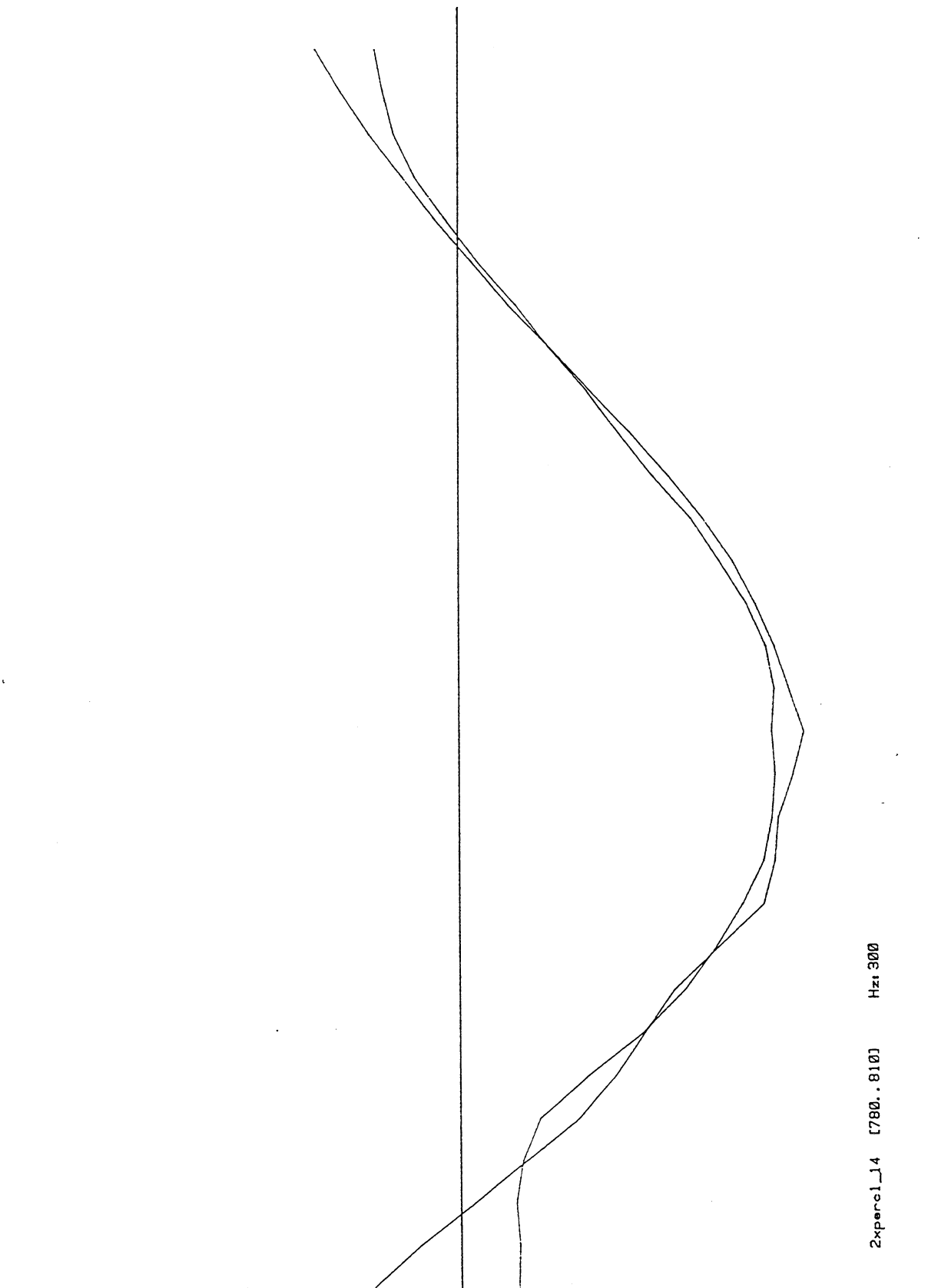
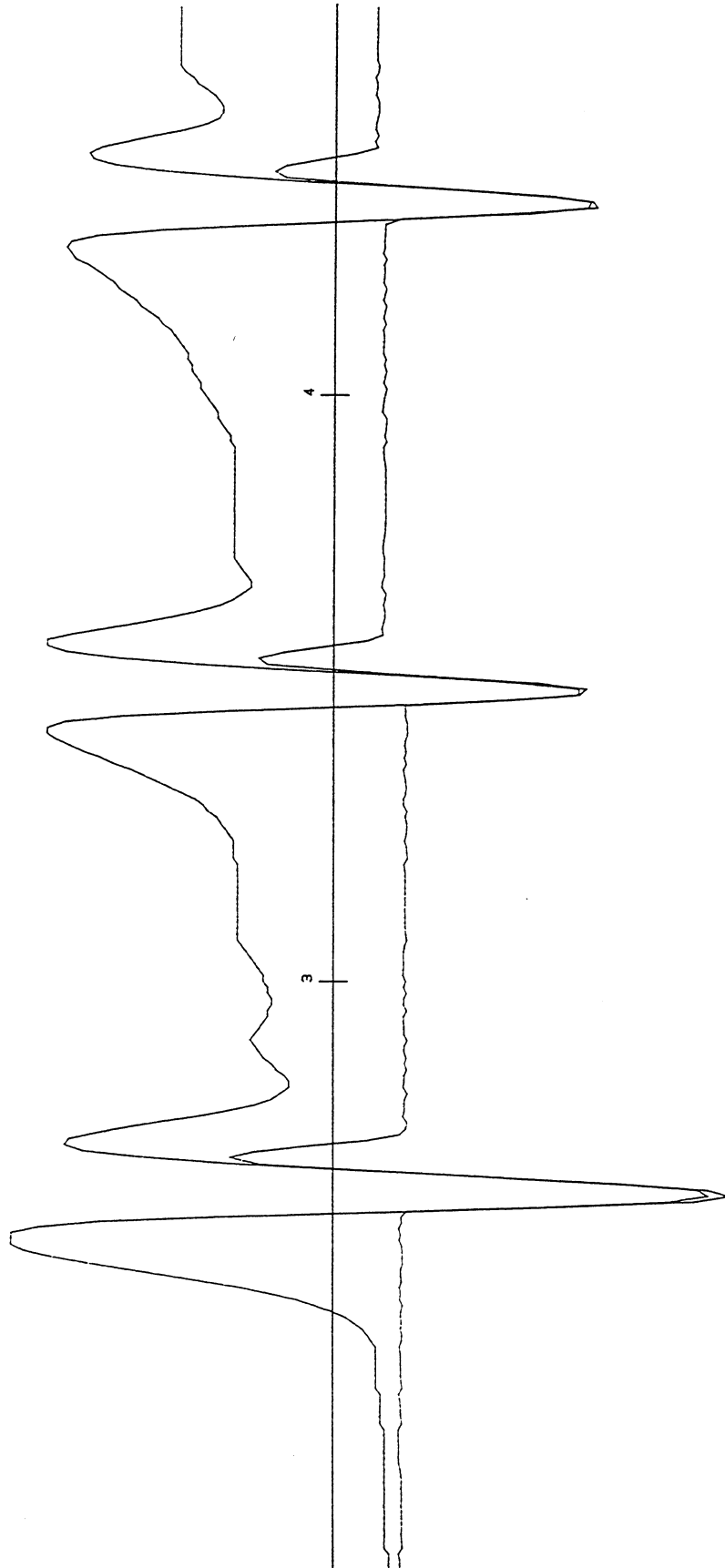
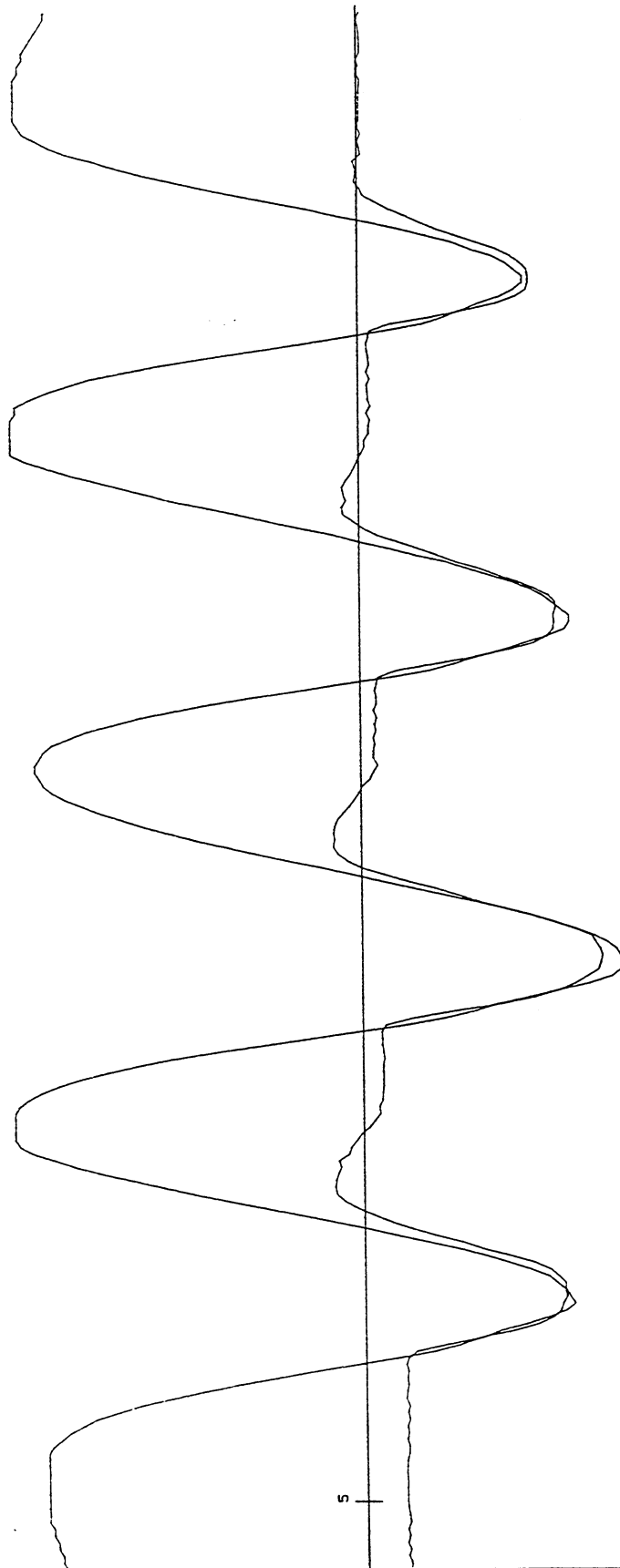


Figure 6



2xper01\_14 [600..1400] Hz: 3000

**Figure 7**



2xperc1\_14 [1490..1720] Hzi 300

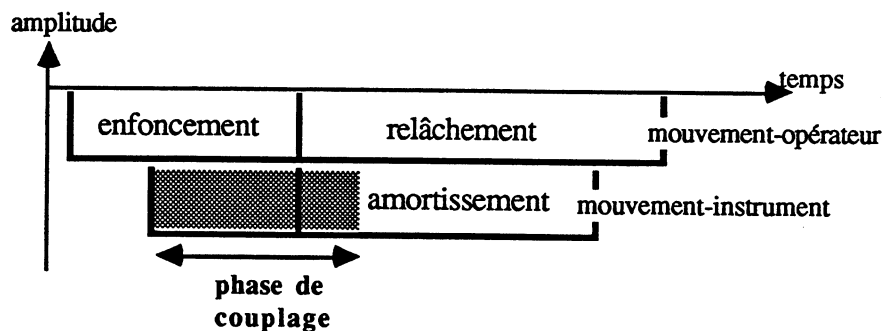
### II.3. interprétation des signaux

Nous distinguons deux cas de figure selon le rapport entre la vitesse d'amortissement opérateur ( $V_H$ ) et celle de l'instrument ( $V_I$ ). Nous considérons que le mouvement du doigt est mono-dimensionnel et vertical.

Dans le cas où  $V_I \geq V_H$ , l'amortissement de l'instrument se réalise plus rapidement que le freinage du geste. Dans le cas où  $V_I < V_H$ , c'est l'inverse ; l'amortissement de l'instrument se réalise moins rapidement que celui du geste.

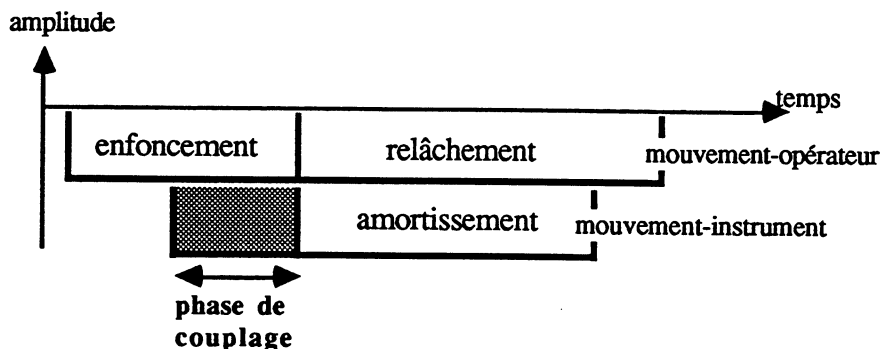
cas  $V_I \geq V_H$  : pour l'opérateur, la phase de couplage débute en fin d'enfoncement et se poursuit sur la phase de relâchement. Pour le mouvement de l'instrument, elle se réalise depuis le début de la percussion, passe par le premier extrémum global et se termine avant le deuxième extrémum.

Ensuite, chaque système se comporte librement.



cas  $V_I < V_H$  : pour l'opérateur, la phase de couplage se réalise dans la phase finale de l'enfoncement. Pour le mouvement de l'instrument, elle se réalise durant la phase d'enfoncement (depuis le début de l'activité jusqu'au voisinage du premier extrémum).

En dehors du couplage, chaque système se comporte librement.





#### **II.4. conclusion**

Sur l'exemple simple d'une touche, initialement au repos et percutée par un doigt, nous avons constaté que la phase de couplage est bornée par l'instant où la touche quitte sa position d'équilibre et un instant proche précédant le second extremum global.

Cependant, dans une situation plus générale, cette constatation n'est encore qu'un cas particulier de couplage. Lors d'une percussion simple, il peut en effet y avoir plusieurs phases de couplage, non forcément contiguës.

Pour approfondir cette question, nous avons introduit dans [Ram89a] l'idée d'une méthode générale permettant d'identifier les intervalles de couplage. Elle repose sur le principe suivant : soit  $F_e$  une force extérieure appliquée à un point de contact-instrument M et S le signal capté sur ce point.

En l'absence de couplage-opérateur ( $F_e=0$ ), S vérifie la courbe issue des algorithmes de simulation du point M. Aussi, pour identifier les phases de couplage ( $F_e \neq 0$ ), il suffit d'identifier les phases où S ne dérive plus uniquement des algorithmes de simulation.

Le problème reste bien entendu les différentes perturbations introduites par le dispositif réel de capture.

### **III. EXEMPLES DE SEGMENTATION**

Cette sous-section fait suite au chapitre II.

Elle a pour but de montrer les performances des algorithmes de segmentation et de sur-segmentation dans le cas de gestes percussifs simples et successifs.

Mais avant cela, nous faisons quelques remarques sur la qualité des signaux captés et mémorisés dans le cadre du système CORDIS avec le clavier rétroactif modulaire.

#### **III.1. situation pratique**

Dans la situation idéale, les hypothèses sont nombreuses : les signaux sont supposés être sans bruit et décrire exactement les positions de repos naturel de l'instrument lorsqu'elles ont lieu. Dans la situation pratique, ces deux hypothèses sont généralement fausses. En effet, les signaux captés sont bruités et montrent que le dispositif de mesure ne décrit pas exactement les positions de repos naturel de l'instrument. Cela provient d'une part de la difficulté à initialiser correctement les positions à zéro des transducteurs gestuels et d'autre part d'une dérive, liée très certainement à un frottement sec.

En conséquence de quoi, le critère de pause, très sensible aux perturbations, devient impropre. Rappelons que dans le cas d'un repos naturel unique, une pause est détectée sur un signal  $S(t)$  si pendant un intervalle de temps  $\Delta T$  le test  $S(t)=R$  reste vrai, avec  $R$  position de repos naturel de l'instrument et  $\Delta T > \Delta \text{SeuilPause}$ .

#### **1. critère de pause adapté**

##### **problème du bruit**

Le problème du bruit peut être traité avec des techniques de lissage. Cela implique un pré-traitement initial qui transforme le signal mesuré bruité en un second signal, image non bruitée du premier.

Cependant, afin de préserver l'information gestuelle d'origine, nous avons initialement exclu toute forme de lissage en entrée du système de segmentation.

##### **problème des zéros**

Le second problème concerne les positions à zéro des transducteurs gestuels.

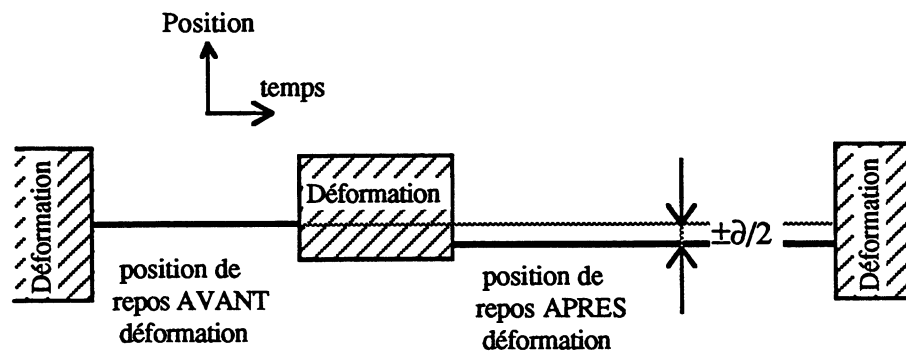
Par exemple, si une entrée en force est supposée mettre le TGR à la position nulle, la sortie correspondante (capteur en position) devra mesurer une valeur nulle. Actuellement, cette adéquation est difficilement réalisable.

Pour s'absoudre de la contrainte du réglage, nous associons à la position de repos naturel théorique (notée R), une seconde valeur (appelé R-zéro) indiquant l'amplitude réelle d'une phase de repos naturel. Cette valeur est mesurée directement sur les signaux.

### problème de la dérive

Le phénomène de la dérive est observable sur le signal enregistré. Les phases de repos se traduisent par un signal stationnaire, mais leurs amplitudes ne sont pas identiques. Elles varient autour de la position de repos unique théorique.

Dans la figure suivante, l'amplitude du repos naturel avant déformation est l'origine. Après la déformation, elle est décalée de  $-\delta/2$ .



Sur le clavier rétroactif le point extrême de chaque touche a une course de 2.8cm. La dérive maximale observée est de l'ordre de 0.1cm.

### critère de pause adapté

Plusieurs compromis sont envisageables pour identifier une pause dans le signal perturbé.

L'un d'eux consiste à comparer chaque position du signal  $S(n)$  par rapport à la position du zéro-relatif :

$$| S(n) - R\text{-zéro} | < \alpha \quad (\text{test A})$$

Une pause est détectée si le test est vérifié sur une durée minimale de  $\Delta\text{SeuilPause}$ .

Un autre critère cible, non pas les repos naturels de l'instrument, mais toute forme de stationnarité de l'instrument (naturels ou contraints). Dans ce cas, il n'y a plus à comparer la valeur du signal avec une valeur de repos théorique, mais uniquement à tester la pente du signal. Le test aura la forme suivante :

$$| S(n) - S(n-1) | < \alpha \quad (\text{test B})$$

Une pause est détectée si ce test est vérifié sur une durée minimale de  $\Delta\text{SeuilPause}$ .

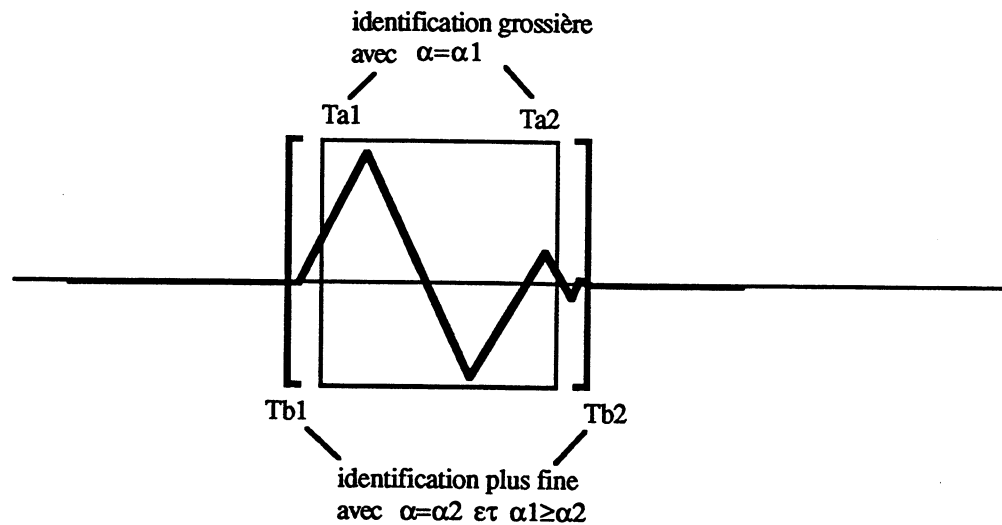
## 2. facteur d'échelle

Une pause est détectée avec une précision qui dépend de la valeur du coefficient  $\alpha$ . Plus  $\alpha$  est petit, plus la précision des bornes d'une pause est grande. Mais si  $\alpha$  est trop petit, le bruit du signal sera identifié comme une information gestuelle significative.

Une solution intéressante au problème du choix de la valeur de  $\alpha$  consiste à réaliser l'identification en deux temps. Le premier temps permet de localiser grossièrement le segment ( $\alpha = \alpha_1$  suffisamment grand). On obtient l'intervalle  $\Delta T_a = [T_{a1}..T_{a2}]$ .

Dans un deuxième temps, il s'agit d'affiner les bornes du segment en choisissant  $\alpha = \alpha_2$  tel que  $\alpha_2 < \alpha_1$ . On obtient l'intervalle  $\Delta T_b = [T_{b1}..T_{b2}]$ .

Théoriquement, si  $\alpha_2 < \alpha_1$ ,  $T_{b1} \leq T_{a1} \leq T_{a2} \leq T_{b2}$ .



### III.2. exemples de segmentation

Pour évaluer nos critères et algorithmes de segmentation automatique, nous comparons les segments-gestuels issus du processus de segmentation automatique avec ceux issus d'une segmentation manuelle (pris comme référence).

Les signaux sur lesquels s'appliquent la segmentation se répartissent en deux groupes :

- 1- les signaux réalisés par synthèse,
- 2- les signaux issus d'un enregistrement en temps-réel avec le système CORDIS sur le clavier rétroactif.

La segmentation de ces signaux se réalise en deux étapes :

- 1- une première étape pour séparer les phases d'activité (segments-actifs) et d'inactivité de l'instrument (segments-pauses). C'est la segmentation élémentaire. Le critère de pause s'appuie sur le test B (op.cit.) :  $|S(n) - S(n-1)| < \alpha$ . Dans le cas de

gestes percussifs, cela ne fait pas de différence avec les autres critères : il n'y a pas de phases de repos contraints.

2- une seconde étape pour identifier le geste instrumental dans les segments-actifs. On parle de sur-segmentation.

Les exemples sont visualisés en notation "tablature-étendue" :

1- les signaux gestuels à segmenter (ou à sursegmenter) sont visualisés sous forme de signaux de position,

2- les segments issus de l'étape 1 (segmentation élémentaire) sont représentés sous la forme de traits qui indiquent les portions du signal segmentées,

3- les segments issus de l'étape 2 (sur-segmentation) sont représentés sous la forme de triangles prolongés par un trait : l'hypothénuse du triangle représente la pente de la droite passant par le début de la percussion et par le premier extremum.

### **1. percussion simple**

#### **a. figure 8**

Sur la figure 8 apparaît une tablature constituée de trois canaux. La fréquence d'échantillonnage des signaux est de 300Hz et la durée séparant deux graduations est de 1 seconde.

Le canal du haut (canal 1) contient le signal brut, le canal du milieu (canal 2) le résultat de la segmentation manuelle et le canal du bas (canal 3) le résultat de la segmentation automatique élémentaire, toutes deux appliquées au signal brut du canal 1.

Le signal brut est un signal de synthèse contenant 3 événements-percussifs ayant chacun la même pente d'attaque mais des amortissements différents. Pour le premier, l'amortissement est un trait rejoignant directement l'origine, pour le second, il y a trois traits dont l'ensemble représente une oscillation schématique, et pour le dernier, l'amortissement vérifie une oscillation élémentaire de la forme  $x(n)=a.x(n-1)+b.x(n-2)$ .

Les paramètres de la segmentation élémentaires sont  $\alpha_1$ ,  $\alpha_2$  et  $\Delta\text{SeuilPause}$ .  $\alpha_1$  permet d'identifier le segment grossièrement,  $\alpha_2$  d'en affiner les bornes et  $\Delta\text{SeuilPause}$  indique la taille minimale d'une pause.

Les valeurs de ces paramètres sont les suivantes :

$$\alpha_1=10$$

$$\alpha_2=2$$

$$\Delta\text{SeuilPause}=15 (=0.05 \text{ sec à } 300\text{Hz})$$

En comparant les segments du canal 2 et du canal 3, on remarque que les différences de segmentation se situent en fin de segment et sont liées au paramètre  $\alpha_2$ .

#### **b. figure 9**

La tablature de la figure 9 est constituée de 2 canaux ; celui du haut contient les segments-actifs issus de la segmentation manuelle du signal précédent (figure 1, canal 2) ; celui du bas, le résultat de la sur-segmentation automatique.

Le critère de sur-segmentation d'un geste percussif simple consiste à identifier le premier extrémum global du signal.

Les unités de temps sont les mêmes que sur la figure 8.

#### **c. figure 10**

Sur la figure 10 apparaît une tablature constituée de trois canaux.

Comme pour la figure 8, le canal du haut (canal 1) contient le signal brut, le canal du milieu (canal 2) le résultat de la segmentation manuelle et le canal du bas (canal 3) le résultat de la segmentation élémentaire, toute deux appliquées au signal brut du haut.

Le signal brut est un signal enregistré issu d'une série de 13 percussions qui s'étalent sur environ 14 secondes.

Les paramètres de la segmentation élémentaires sont les mêmes que ceux définis à la figure 8.

L'absence de certains événements sur le canal 3 (par rapport au canal 2), de même que la réduction de certaines phases d'enfoncement proviennent du coefficient  $\alpha_1$ . Rappelons que la valeur de ce coefficient a été déterminée en fonction du taux de bruit. En la diminuant, les événements insignifiants sont ressuscités, mais dans le même temps, les parties bruitées réapparaissent.

#### **d. figure 11**

Sur la figure 11 sont représentées 4 voies du clavier modulaire : la voie 0, 2, 4 et 6 (de la gauche vers la droite). Le geste percussif réalisé ne se joue plus sur un seul canal, mais sur 4. C'est un exemple d'articulation percussive. Elle consiste notamment en un arpège cyclique débutant sur la voie 2 avec l'index, puis sur la voie 4 avec le majeur, puis sur la voie 6 avec l'annulaire, et se termine sur la voie 0 avec le pouce.

#### **e. figure 12**

Nous représentons sur la figure 12 les segments extraits des signaux de la figure 11.

Les canaux pairs contiennent le résultat de la segmentation manuelle, les canaux impairs celui de la segmentation automatique. Les canaux (1, 3, 5, 7) et (2, 4, 6, 8) correspondent respectivement aux voies (0, 2, 4, 6) de la figure 11.

Les valeurs des paramètres de la segmentation automatique sont les mêmes que ceux de la figure 8.

On remarque quelques différences avec la référence manuelle. En particulier, sur le canal 4, juste après la graduation 6 (voir voie 2, figure 4, même instant), le segment est légèrement tronqué. Cette différence provient de la valeur seuil  $\alpha_1$ .

En outre, ces événements d'amplitude plus faible que les percussions sont beaucoup plus importantes que du bruit. Il s'agit d'événements parasites, non pas introduits par le dispositif de mesure, mais liés à la précision du jeu sur le clavier : lors de la frappe d'une touche, la touche voisine est effleurée en même temps et esquisse un petit mouvement. C'est le cas des événements localisés sur la voie 4, avant les percussions de la graduation 5, 6 et 7. Ces événements correspondent aux percussions de la voie 2.

Il est également possible de voir dans ces petits segments un couplage du 3ème doigt qui effleure la touche.

Voici quelques mesures sur les 4 arpèges.

1- les durées moyennes séparant les événements de chaque arpège  $a_i$  ( $i=1..4$ ) sont

$a_1 : 0.199s$     $a_2 : 0.195s$     $a_3 : 0.195s$     $a_4 : 0.203s$ .

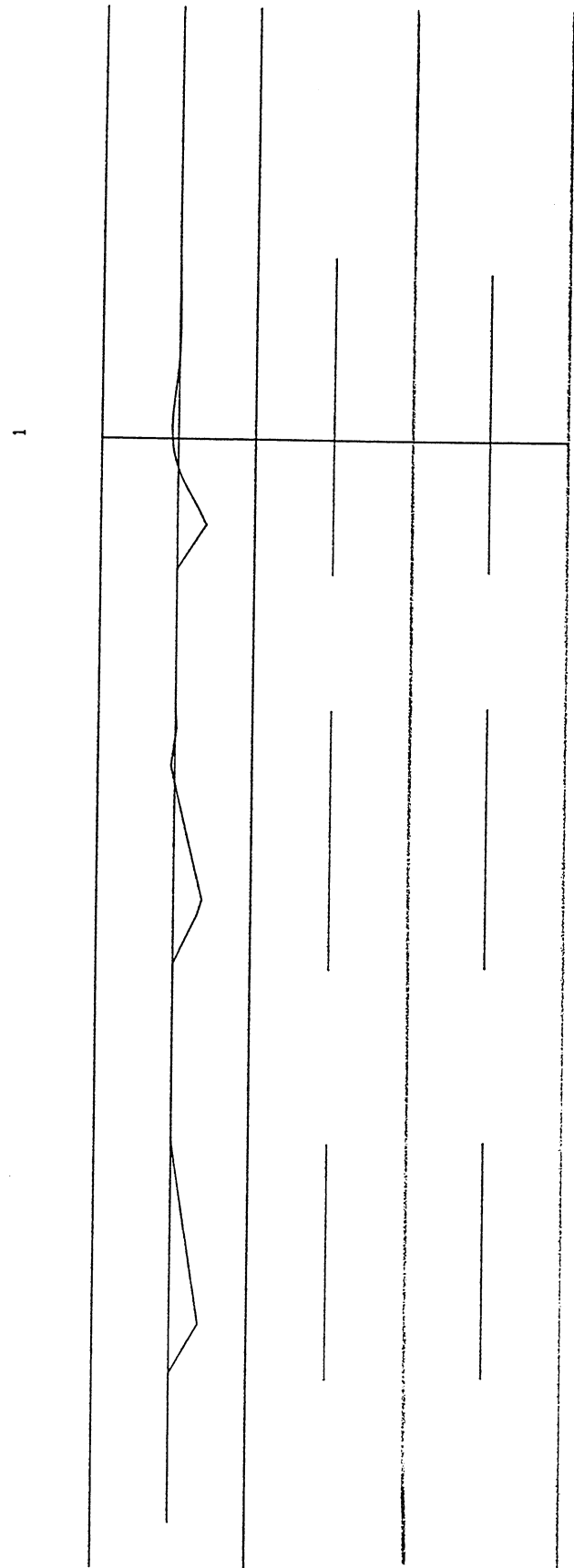
La durée moyenne est de  $0.198 \pm 0.005s$ .

2- les durées moyennes séparant les événements sur une même voie  $v_i$  ( $i=1..4$ ) sont

$v_1 : 0.803s$     $v_2 : 0.803s$     $v_3 : 0.803s$     $v_4 : 0.808s$

La durée moyenne est de  $0.804 \pm 0.004s$

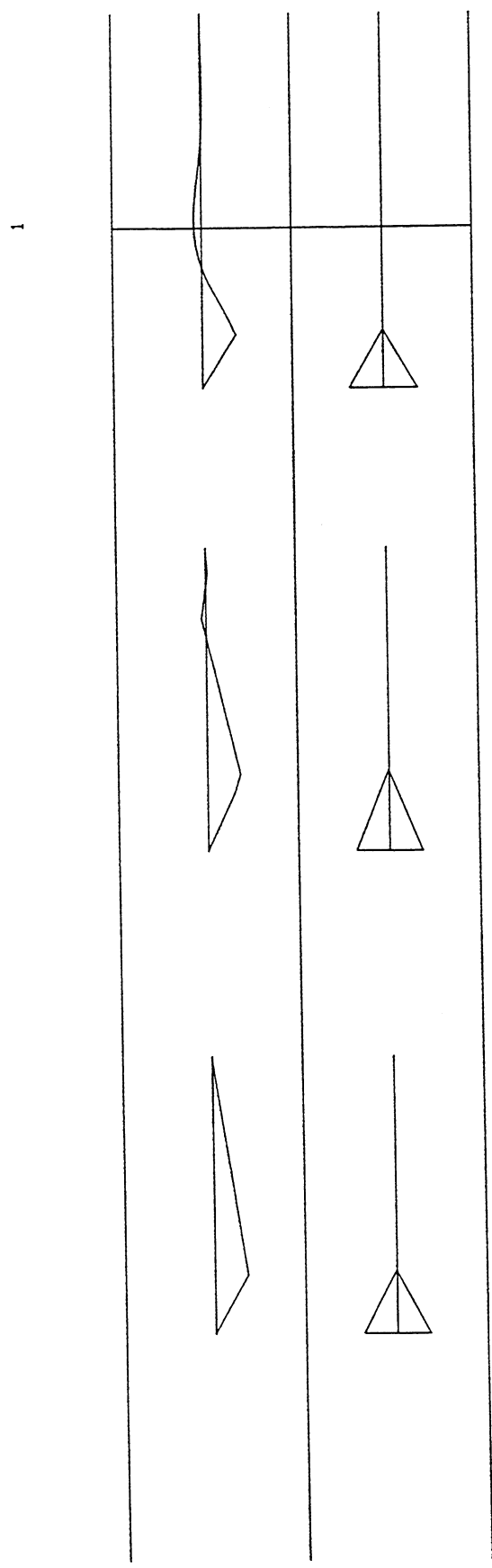
Figure 8



eegCP.get 0.414 Hzi 300

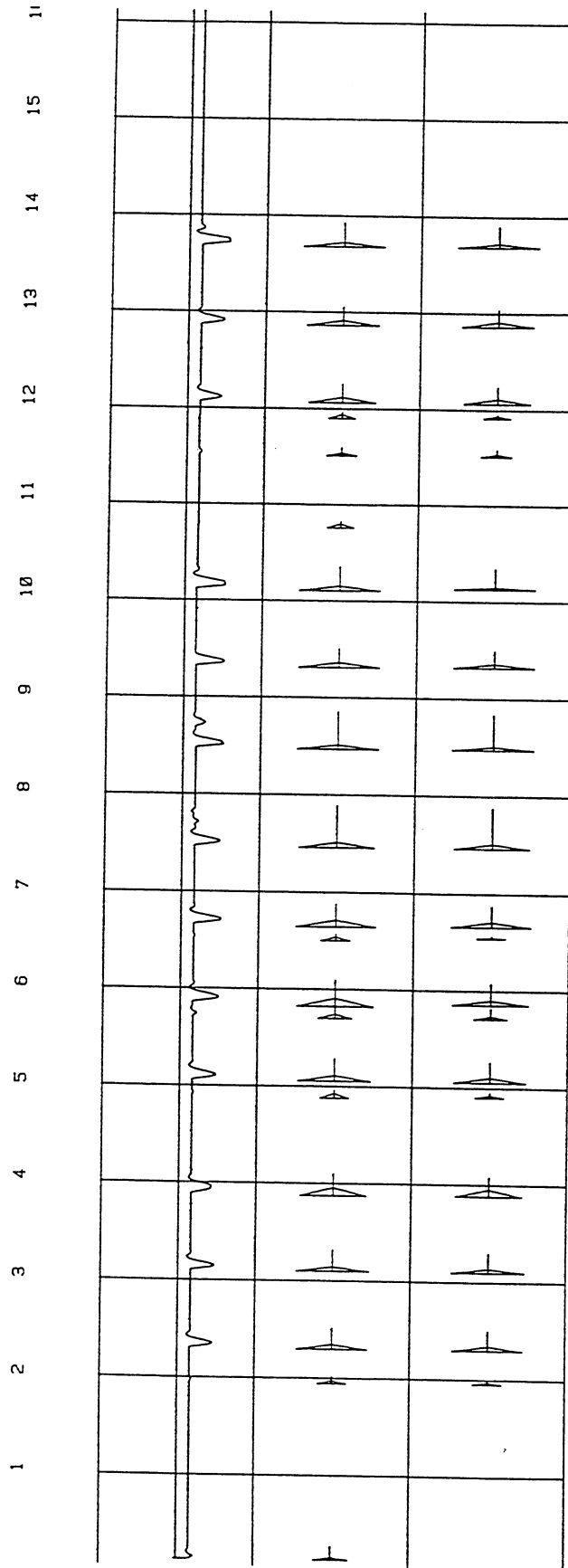


Figure 9



eegCI\_GP.get 0.348 Hz: 300

Figure 10



segPreal.get 0..4834 Hz: 300

Figure 11

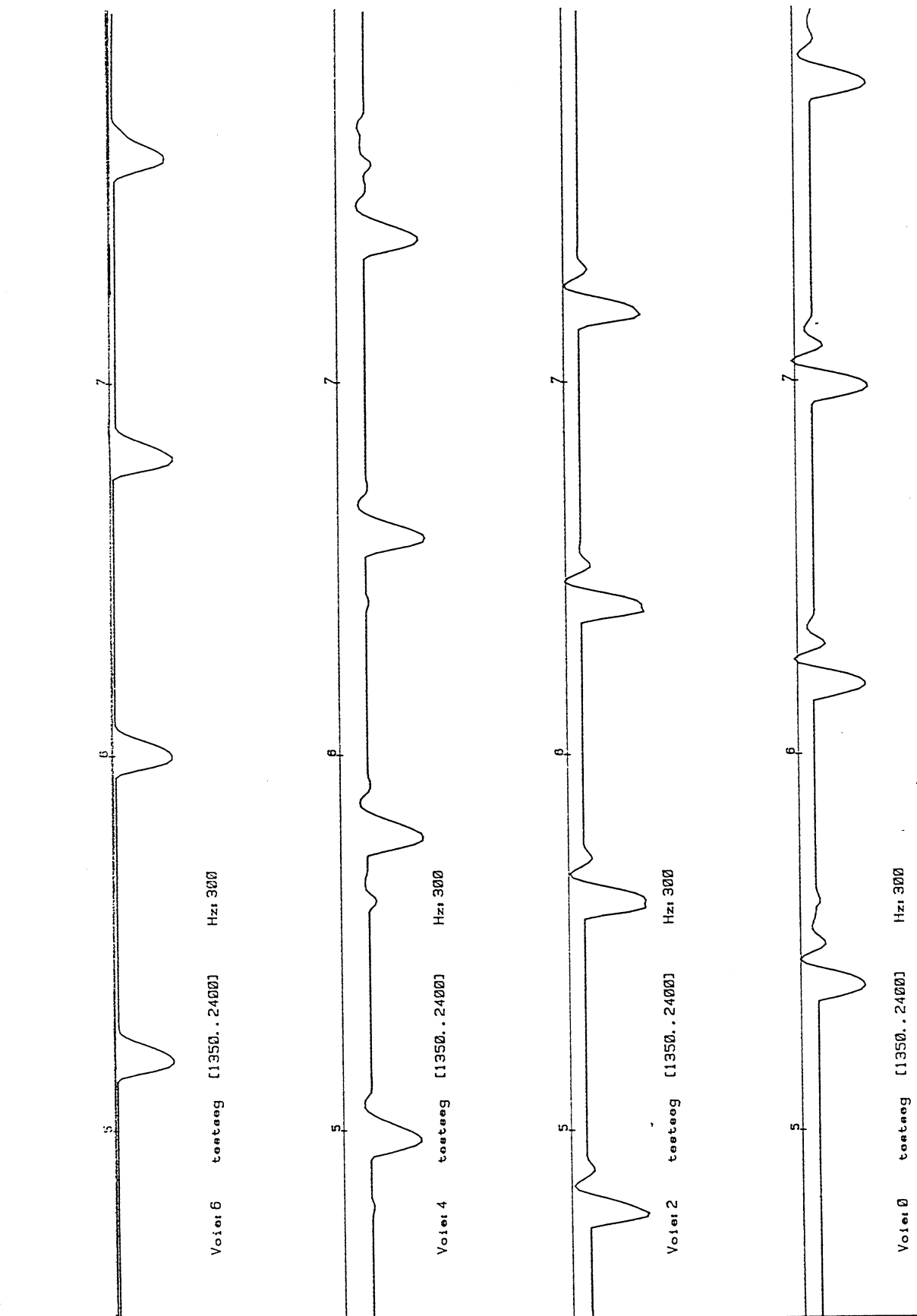
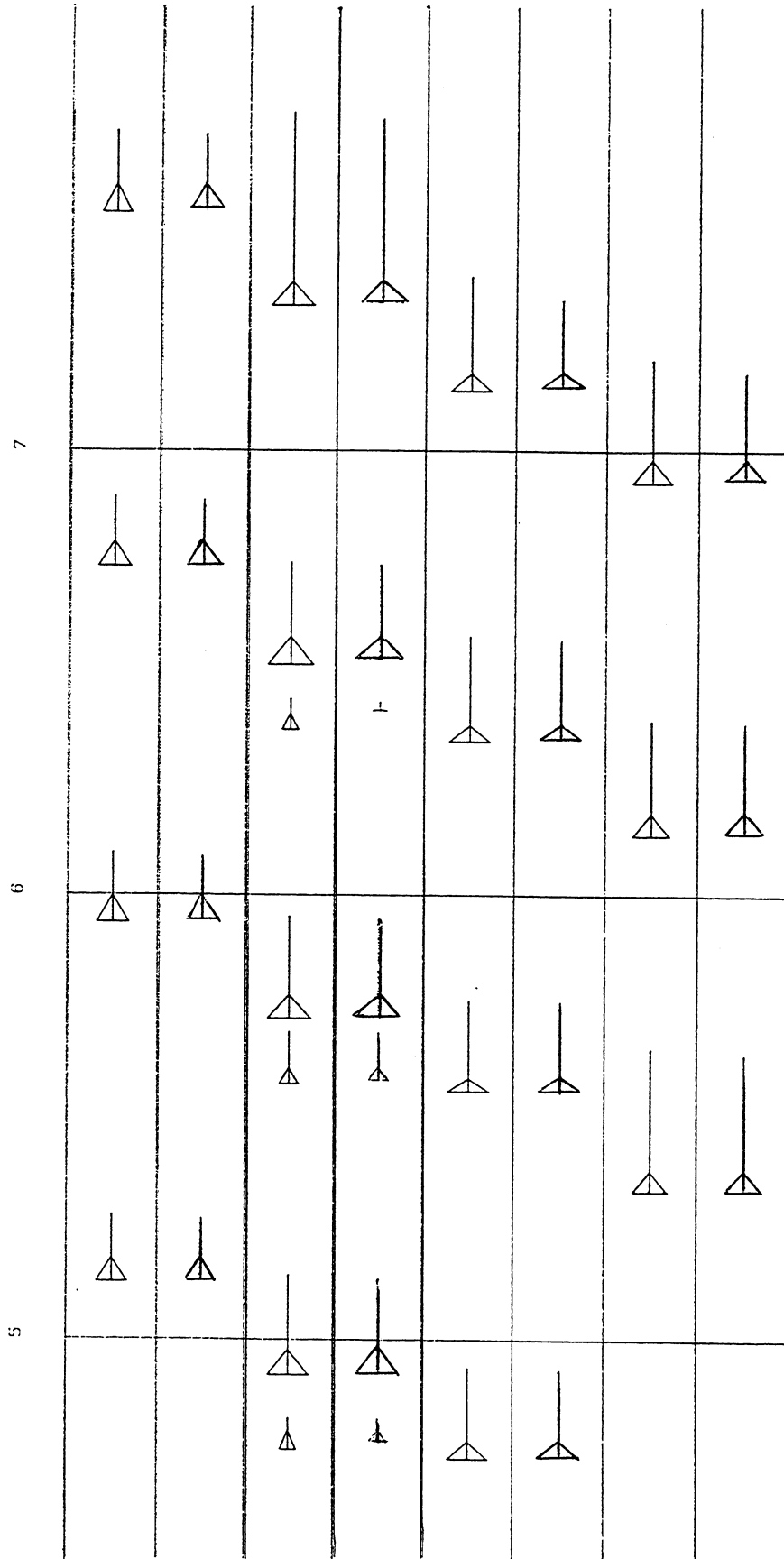


Figure 12



4d\_pegman. gat 1350. 2400 Hz: 300

## **2. percussion successives**

### **a. figure 13**

La signification des canaux de la figure 13 est la même que celle de la figure 8. Le canal du haut (canal 1) contient le signal brut, le canal du milieu (canal 2) le résultat de la segmentation manuelle et le canal du bas (canal 3) le résultat de la segmentation automatique élémentaire, toutes deux appliquées au signal brut du canal 1.

Le signal brut du canal 1 est un signal de synthèse contenant 4 événements percussifs doublés. La pente d'attaque de la première percussion de chacun de ces événements est identique. L'amortissement de chaque percussion vérifie une oscillation élémentaire décrit par une équation aux différences du type  $x(n)=a.x(n-1)+b.x(n-2)$ .

La différence entre ces 4 événements réside en l'instant de début de la deuxième percussion. Dans le premier cas, il y a retour à l'équilibre (mais sans être suffisamment long pour être une pause). Dans le second cas, on devine un point d'inflexion. Dans le troisième cas, la courbure est continue. Dans le 4ème cas enfin, la courbe ne passe pas par l'origine.

Les paramètres de la segmentation sont les mêmes que ceux de la figure 8 :

$$\alpha_1=10$$

$$\alpha_2=2$$

$$\Delta\text{SeuilPause}=15 \text{ (=0.05 sec à 300Hz)}$$

Les segments du canal 3 sont légèrement plus courts que les segments du canal 2. Cette différence peut se corriger en diminuant légèrement le coefficient  $\alpha_2$ .

### **b. figure 14**

Les segments du canal 2 de la figure 14 sont sursegmentés. La sursegmentation consiste d'une part à séparer les phases percussives et d'autre part à marquer la fin de la phase d'enfoncement et d'amortissement de chacune d'elle.

Les segments sont visualisées sous forme brute sur le canal du haut (canal 1), le canal 2 en donne la sur-segmentation manuelle et le canal 3 le résultat de la sur-segmentation automatique.

Les paramètres de la sur-segmentation élémentaires sont  $\alpha$  et  $\psi$ .  $\alpha$  est un seuil de bruit et  $\psi$  est le coefficient de l'exponentielle  $e^{-\psi\tau}$  décrivant l'enveloppe d'amortissement de l'instrument.

Les valeurs de ces paramètres sont les suivantes :

$$\alpha=5$$

$$\psi=.03$$

Les segments du canal 3 sont identiques à ceux du canal 2.

### c. figure 15

Comme pour la figure 8, le canal du haut (canal 1) contient le signal brut, le canal du milieu (canal 2) le résultat de la segmentation manuelle et le canal du bas (canal 3) le résultat de la segmentation élémentaire, toutes deux appliquées au signal brut du haut.

Le signal brut est un signal enregistré issu d'une série de 6 événements constitués de 4 quadruples percussions et 2 percussions simples. Le signal est très bruité et la dérive particulièrement importante (7 bits de codage).

Les paramètres de la segmentation élémentaire sont les mêmes que ceux définis à la figure 8.

La surabondance de segments du canal 3 par rapport au canal 2 vient du fait que le signal brut est très bruité. le coefficient  $\alpha_1$  est trop faible. Pour supprimer ce bruit, il faudrait l'augmenter.

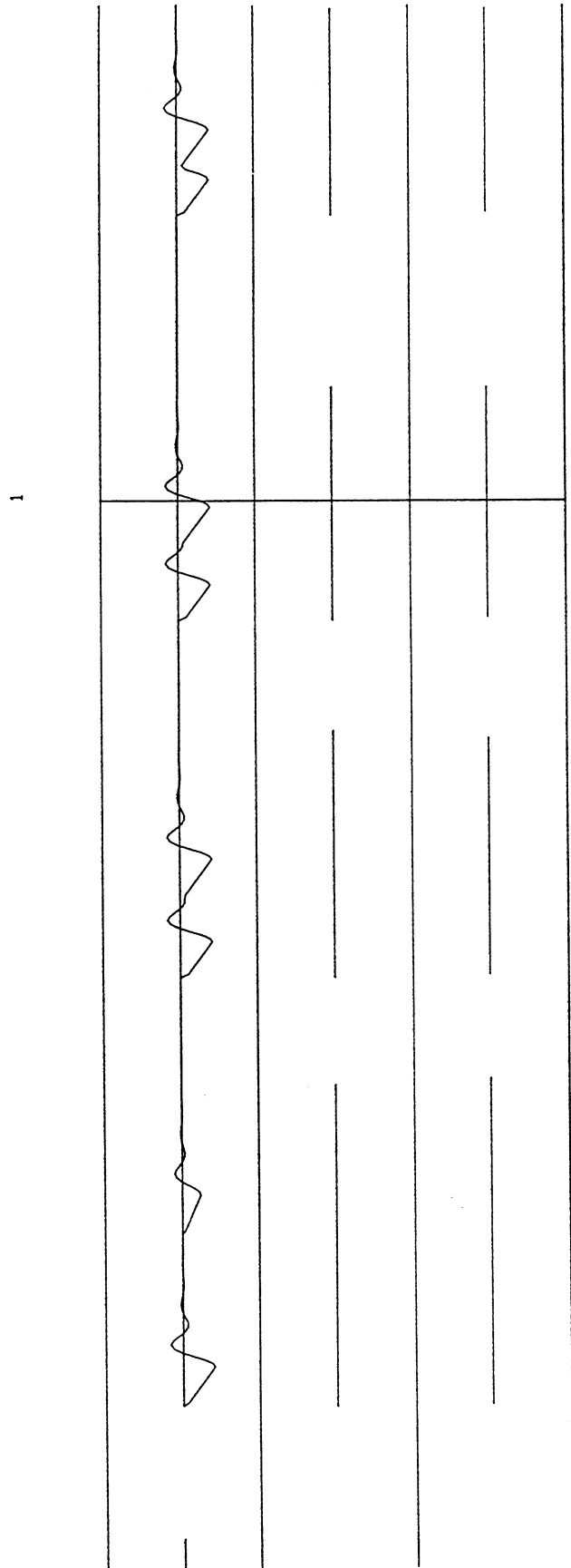
### d. figure 16

Les segments du canal 2 de la figure 16 sont sursegmentés. Ils sont visualisés sous forme brute sur le canal 1, le canal 2 en donne la sur-segmentation manuelle et le canal 3 le résultat de la sur-segmentation automatique.

Les paramètres de la sur-segmentation élémentaire sont les mêmes que ceux de la figure 13.

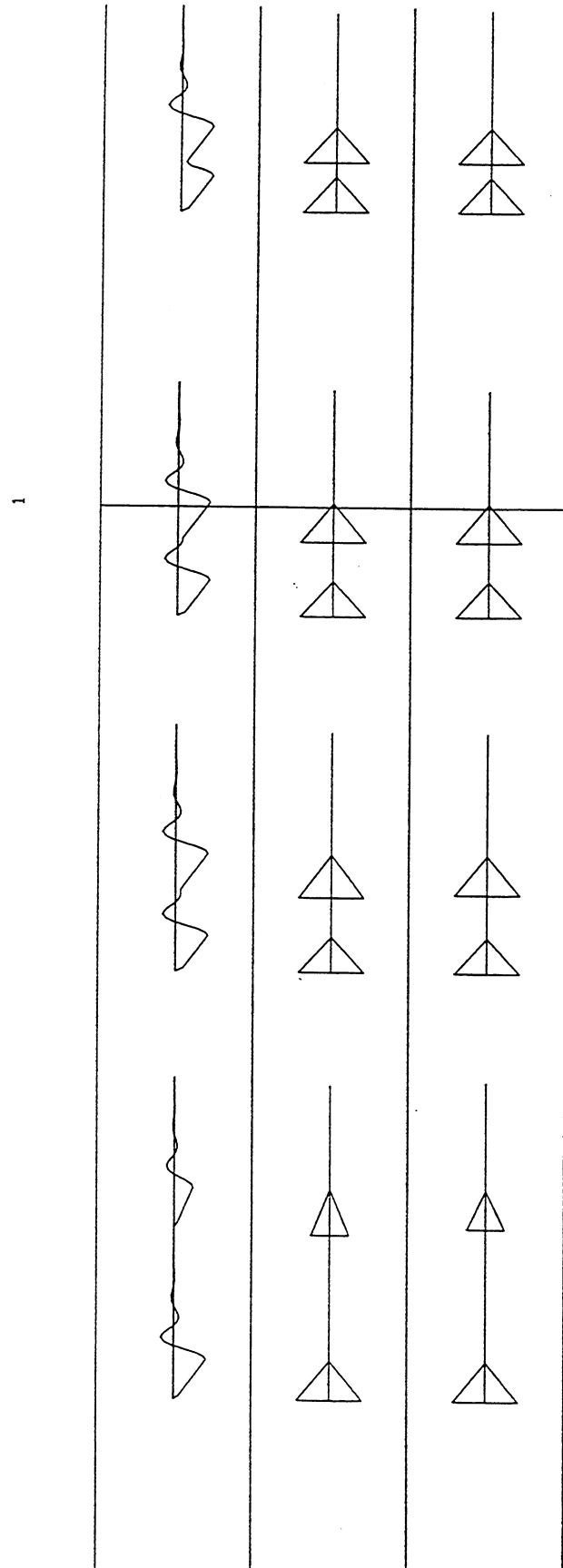
On n'observe pas de différence notable entre les segments des canaux 2 et 3.

Figure 13



segCPR.gst Ø. 439 Hz: 300

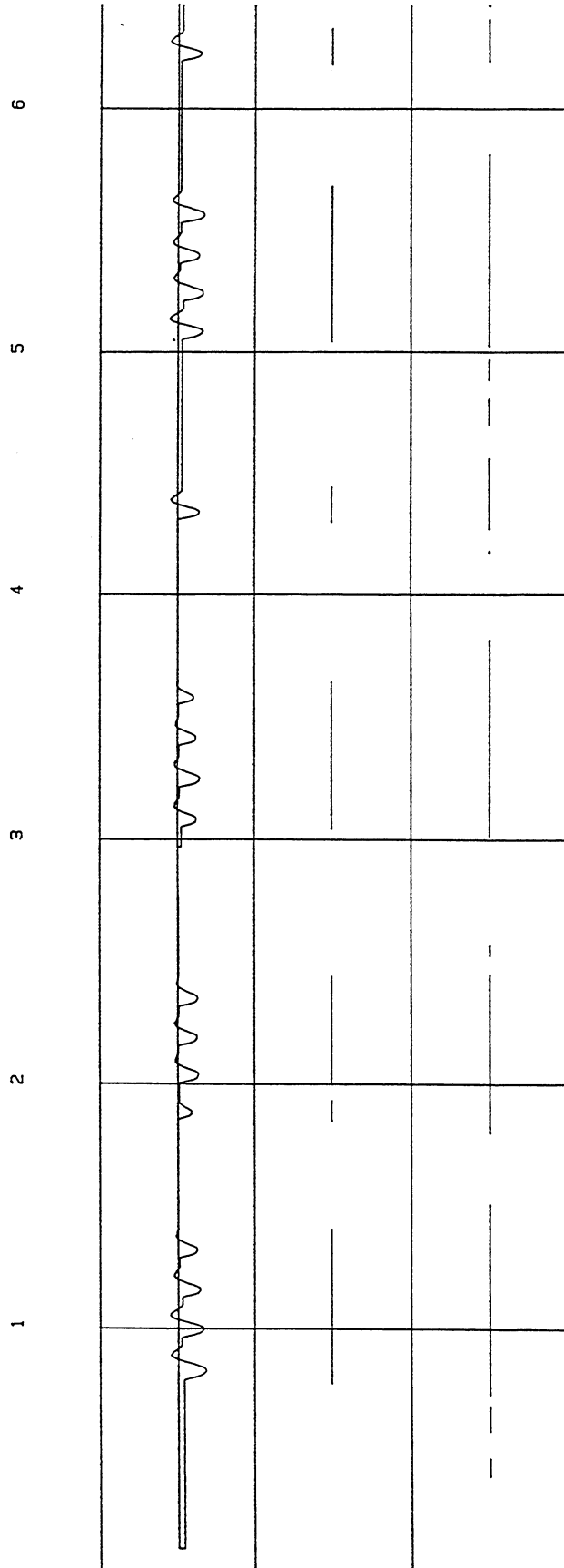
Figure 14



segCI\_CPR.get 0..441 Hz: 300

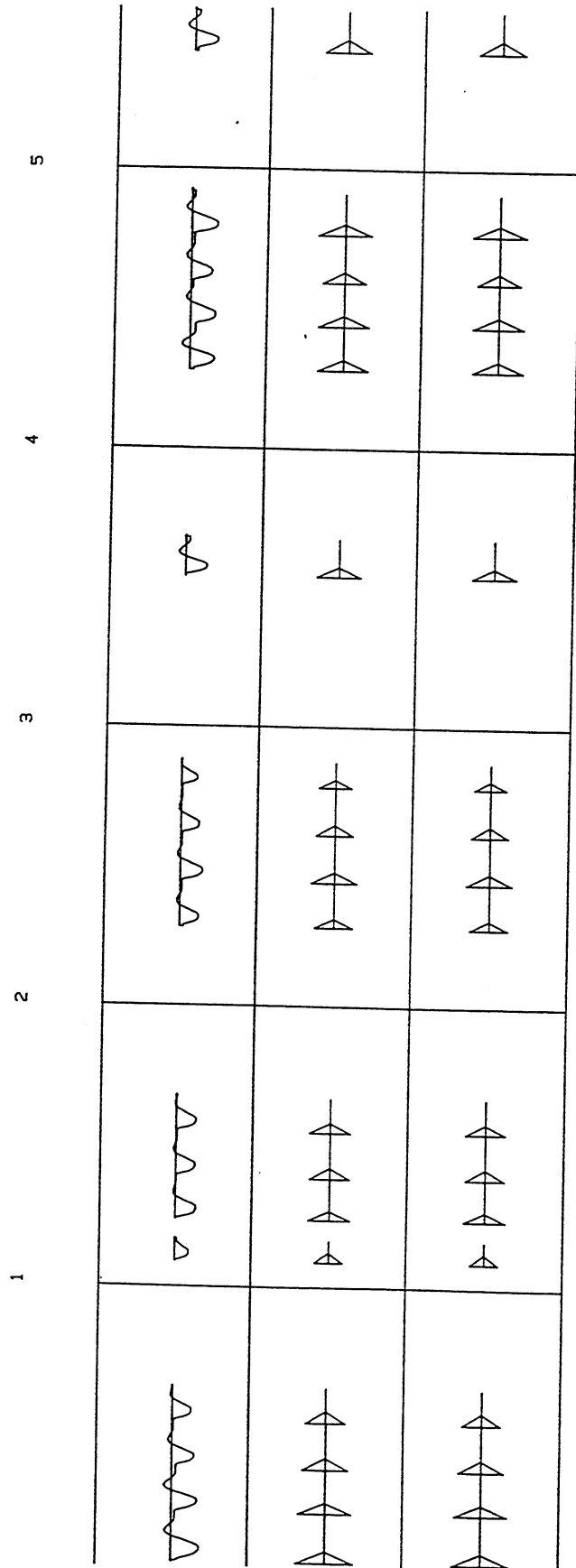


Figure 15



segCPR-recl.get 0..1927 Hzi 300

Figure 16



seg01\_CFRRecal.get 0..1672 Hz: 300

## **IV. ESTIMATION DE L'ALGORITHME DE VALIDATION D'UNE ARTICULATION MONOPHONIQUE**

Une gamme majeure est constituée de 7 notes qui se reproduisent à l'octave. Chaque note de la gamme est séparée de la suivante soit par un ton soit par un demi-ton, selon les écarts suivants : 1, 1, 1/2, 1, 1, 1, 1/2. Par exemple, pour la gamme de do majeur, on a Do(C), Ré(D), Mi(E), Fa(F), Sol(G), La(A), Si(B).

C.P.E Bach [Bac79] a proposé un ensemble de doigtés de référence pour ces gammes.

Pour estimer l'algorithme de validation d'une articulation monophonique défini au chapitre IV, nous considérons des séquences de 8, 9 et 10 notes basées sur quatre gammes majeures dans le sens de la montée : la gamme de do majeur (CM), de do dièse majeur (C#M), de fa majeur (FM) et de fa dièse majeur (F#M).

Pour chaque séquence (de 8, 9 et 10 notes), nous indiquons les doigtés les plus faciles proposés par notre algorithme, et donnons le coefficient de facilité des doigtés proposés par Bach.

### **IV.1. doigtés proposés par [Bac75]**

C.P.E Bach [Bac79] a donné pour les 12 gammes majeures et les 12 gammes mineures de l'octave, aussi bien en montée qu'en descente, un ensemble d'exemples de doigté.

"Il serait possible de trouver beaucoup d'autres doigtés, si l'on voulait élaborer des exemples. Je les ai cités, bien qu'ils soient moins naturels que les premiers, davantage pour mettre en garde le lecteur que pour qu'il les imite, et parce que je sais qu'ici et là ils sont à la mode".

CM / main droite	:	1231234-1	1234343-4	1234123-1
CM / main gauche	:	5432132-1	4321432-1	4321212-1
C#M / main droite	:	2312341-2		
C#M / main gauche	:	3214321-3		
DM / main droite	:	1231234-1		
DM / main gauche	:	2143213-2	4321212-1	4321432-1
D#M / main droite	:	2123412-3		
D#M / main gauche	:	3214321-3		

EM / main droite	:	1231234-1		
EM / main gauche	:	5432132-1		
FM / main droite	:	1234123-1		
FM / main gauche	:	3214321-3	5432132-1	321321212-1
F#M / main droite	:	234123-1		
F#M / main gauche	:	432132-1		
GM / main droite	:	1234123-1	1234343-1	1231234-1
GM / main gauche	:	4321432-1	5432132-1	4321212-1
G#M / main droite	:	2312312-3		
G#M / main gauche	:	3214321-3		
AM / main droite	:	1231234-1		
AM / main gauche	:	2132143-2	5432132-1	
A#M / main droite	:	2123123-4		
A#M / main gauche	:	3214321-3		
BM / main droite	:	1231234-1		
BM / main gauche	:	4321432-1		

#### **IV.2. doigtés obtenus avec la matrice de facilité**

Nous donnons pour chaque séquence de 8 notes les doigtés les plus faciles proposés par notre algorithme, d'une part pour la main gauche et d'autre part pour la main droite. Les doigtés pour la main gauche sont déterminés à partir de la fonction de facilité de la main droite. Etant données les propriétés de symétrie de la main droite et de la main gauche, nous supposons que leur facilité de jeu est identique, mais en sens inverse. Il suffit donc dans l'algorithme de validation de l'articulation de considérer l'égalité suivante :

$$F_{\text{gauche}}(m_1=(a_1, \mathbf{c1}, t_1), d_{m1}, m_2=(a_2, \mathbf{c2}, t_2), d_{m2}) \\ = \\ F_{\text{droite}}(m_1=(a_1, \mathbf{c2}, t_1), d_{m1}, m_2=(a_2, \mathbf{c1}, t_2), d_{m2})$$

Pour chaque doigté, nous en indiquons le coefficient de facilité. Les doigtés soulignés correspondent aux doigtés donnés par [Bac79]. Pour les séquences de 9 et 10 notes, nous

indiquons seulement le doigté le plus facile avec son coefficient moyen, les doigtés proposés par Bach (en souligné), et le nombre de doigtés intermédiaires.

### 1. gamme de DO majeur montante

8 notes : C D E F G A B - C

#### doigté main droite

0.943 :	1231234-5					// doigté le plus facile
0.914 :	1234123-4	2341234-5				
0.890 :	1234134-5	1341234-5				
0.886 :	1231231-2					
0.871 :	1212312-3	1231212-3				
0.867 :	1234124-5	1241234-5				
0.862 :	1231312-3	1312312-3				
0.857 :	1212123-4	<u>1231234-1</u>	<u>1234123-1</u>	2312341-2	2341231-2	
	3123412-3	3412312-3	4123123-4			

#### doigté main gauche

0.943 :	<u>5432132-1</u>					// doigté le plus facile
0.914 :	<u>4321432-1</u>	5432143-2				
0.886 :	2132132-1	3213213-2	4321321-3			
0.871 :	3213132-1	3213212-1	4321321-2	5432121-3		
0.867 :	5421432-1	5431432-1	5432142-1	5432143-1		
0.862 :	3132132-1	3213132-1	4313213-2	5432131-3		
0.857 :	1321432-1	1432132-1	2132143-2	2143213-2	3213214-3	
	3214321-3	<u>4321212-1</u>	4321321-4	5432121-2		

9 notes : C D E F G A B - C D

#### doigté main droite

0.925 :	1234123-45					// doigté le plus facile
0.875 :	<u>1231234-12</u>	<u>1234123-12</u>				// 10 doigtés intermédiaires

#### doigté main gauche

0.925 :	5432143-21					// doigté le plus facile
0.900 :	<u>5432132-14</u>					// 11 doigtés intermédiaires
0.850 :	<u>4321432-14</u>					// 41 doigtés intermédiaires
0.812 :	<u>4321212-12</u>					// 153 doigtés intermédiaires

10 notes : C D E F G A B - C D E

#### doigté main droite

0.911 :	1231231-234					// doigté le plus facile
---------	-------------	--	--	--	--	--------------------------

0.889 :	<u>1231234-123</u> <u>1234123-123</u>	// 5 doigtés intermédiaires
	<u>doigté main gauche</u>	
0.911 :	5432132-132	// doigté le plus facile
0.889 :	<u>5432132-143</u>	// 5 doigtés intermédiaires
0.867 :	<u>4321432-143</u>	// 25 doigtés intermédiaires
0.833 :	<u>4321212-121</u>	// 110 doigtés intermédiaires

### 2. gamme de DO dièse majeur montante

8 notes : C# D# E F# G# A# C - C#

doigté main droite

0.857 : 23123412 // doigté le plus facile

doigté main gauche

0.857 : 3214321-3 // doigté le plus facile

0.843 : 3214321-2

9 notes : C# D# E F# G# A# C - C# D#

doigté main droite

0.875 : 2312341-23 // doigté le plus facile

doigté main gauche

0.875 : 3214321-32 // doigté le plus facile

10 notes : C# D# E F# G# A# C - C# D# E

doigté main droite

0.889 : 2312341-231 // doigté le plus facile

doigté main gauche

0.889 : 3214321-321 // doigté le plus facile

### 3. gamme de FA majeur montante

8 notes : F G A A# C D E - F

doigté main droite

0.914 : 1234123-4 // doigté le plus facile

0.890 : 1234134-5

0.886 : 3123123-4

0.871 : 1212312-3 2123123-4 2312123-4

0.867 : 1234124-5

0.862 : 1312312-3 3123134-5 :

0.857 : 1212123-4 1234123-1 3123412-3 3412312-3 4123123-4

doigté main gauche

0.943 : 5432132-1 // doigté le plus facile  
 0.914 : 5432143-2  
 0.886 : 2132132-1 3213213-2  
 0.871 : 3212132-1 3213212-1 5432121-3  
 0.867 : 5432142-1 5432143-1  
 0.862 : 3132132-1 3213132-1 4313213-2 5432131-3  
 0.857 : 1432132-1 2132143-2 2143213-2 3213214-3 3214321-3  
 5432121-2

9 notes : F G A A# C D E - F G

doigté main droite

0.925 : 1234123-45 // doigté le plus facile  
 0.875 : 1234123-12 // 7 doigtés intermédiaires

doigté main gauche

0.925 : 5432143-21 // doigté le plus facile  
 0.875 : 3214321-32 5432132-14 // 6 doigtés intermédiaires  
 0.825 : 3213212-12 // 65 doigtés intermédiaires

10 notes : F G A A# C D E - F G A

doigté main droite

0.911 : 2312312-345 // doigté le plus facile  
 0.889 : 1234123-123 // 2 doigtés intermédiaires

doigté main gauche

0.911 : 5432132-132 // doigté le plus facile  
 0.889 : 3214321-321 5432132-143 // 5 doigtés intermédiaires  
 0.844 : 3213212-121 // 44 doigtés intermédiaires

**4. gamme de FA dièse majeur montante**

8 notes : F# G# A# C C# D# E - F#

doigté main droite

0.943 : 1231234-5 // doigté le plus facile  
 0.914 : 2341234-5  
 0.890 : 1341234-5  
 0.886 : 1231231-2  
 0.867 : 1241234-5  
 0.857 : 2341231-2

doigté main gauche

0.914 : 4321432-1 // doigté le plus facile

9 notes : F# G# A# C C# D# E - F# G#

doigté main droite

0.900 : 1231231-23 // doigté le plus facile

0.875 : 2341231-23 // 0 doigté intermédiaire

doigté main gauche

0.900 : 4321321-32 // doigté le plus facile

0.875 : 4321321-43 // 2 doigtés intermédiaire

10 notes : F# G# A# C C# D# E - F# G# A#

doigté main droite

0.911 : 1231231-234 // doigté le plus facile

0.889 : 2341231-234 // 0 doigté intermédiaire

doigté main gauche

0.911 : 4321321-321 // doigté le plus facile

0.889 : 4321321-432 // 0 doigté intermédiaire

**IV.3. conclusion**

Nous faisons plusieurs constatations :

1/ les doigtés proposés par [Bac79] se retrouvent tous avec notre algorithme sauf ceux qui contiennent un passage de doigt autre que  $1 \rightarrow i$  ou  $i \rightarrow 1$ , tel que  $i \in [2..5]$  :  $3 \rightarrow 4$  par exemple.

Si dans l'algorithme, nous avons décrit de tels passages de doigts, nous les aurions retrouvés dans les exemples.

2/ les coefficients de facilité des doigtés de Bach ne sont pas systématiquement maximaux avec notre algorithme.

Cela s'explique par le fait qu'en ajoutant ou supprimant une note à une séquence de N notes, on ne conserve pas forcément le doigté optimal obtenu sur la séquence de N notes. Dans le cas de la gamme de DO majeur réalisée avec la main droite sur 8 notes, le doigté optimal est : 12312345. Toutes les transitions sont maximales sauf pour le passage de doigt  $3 \rightarrow 1$ . Avec 9 notes, le doigté optimal est : 123412345. Toutes les transitions sont maximales sauf pour le passage de doigt  $4 \rightarrow 1$ . Avec 10 notes, le doigté optimal est : 1231231234. Ainsi, bien que les 8 premières notes de ces trois séquences soient identiques, le doigté optimal change.



Le doigté proposé par Bach pour une gamme donné relève d'une volonté d'égaliser le jeu, problème nouveau avec l'apparition du piano-forte.

3/ selon Bach,

"chaque gamme a sept degrés, et, lorsqu'on la répète à l'octave, le début doit en être identique pour qu'on puisse conserver le même doigté." [Bach79]

Pour la gamme de DO majeur, les doigtés à la main droite proposés par Bach sont 1231234-1 ou 1234123-1. On trouve le pouce sur la note fondamentale et à l'octave. Remarquons cependant que ceci n'est pas général. Par exemple, le doigté le plus facile de la gamme de DO jouée à la main gauche est 5432132-14. La période se trouve à partir de la seconde à l'octave.

Avec notre algorithme, pour la gamme de DO majeur réalisée à la main droite sur une séquence de 8 notes, il existe plusieurs doigtés dont la période se retrouve à partir de l'octave. Pour une facilité supérieure à 0.875, on a 7 doigtés périodiques : 1231234-1 1234123-1 2312341-2 2341231-2 3123412-3 3412312-3 4123123-4. Avec 9 notes, ce nombre diminue et avec 10 notes il ne reste plus que 3 doigtés ayant une facilité moyenne supérieure à 0.889 : 1231234-123 1234123-123 2341231-234.

Il semble donc que les doigtés de référence proposés par Bach aient une facilité moyenne limite : plus le nombre de notes augmente, plus on tend vers ces doigtés.

4/ Bach fait une constatation importante que nous pouvons vérifier avec nos exemples de doigté :

"les gammes qui ont peu de dièses et de bémols ou qui n'en n'ont pas du tout sont celles qui permettent le plus grand nombre de modifications dans le doigté, lorsqu'il s'agit de passer les doigts aussi bien par-dessus que par-dessous les autres. Les autres ne permettent que quelques possibilités de doigté. C'est pourquoi les tons prétendus faciles [...] sont en fait bien plus fourbes et difficiles que les tons prétendument tels, dans la mesure où ceux-ci n'admettent qu'un seul doigté, et où le pouce, par l'exercice, apprend sans peine à prendre la place qui lui revient."

5/ pour évaluer les performances de notre algorithme, nous venons de comparer les doigtés qu'il propose pour certaines gammes particulières avec les doigtés utilisés habituellement sur un clavier de piano. Il va de soi que cette évaluation est loin d'être complète, et peut-être même, trop particulière. J'insiste donc encore sur le fait que les doigtés proposés par Bach permettent de contrôler le legato et la dynamique sur un piano-forte. Ce problème ne se pose pas pour un clavecin et le passage de pouce par exemple n'y est pas vital.

## **V. VALIDATION D'UNE ARTICULATION**

Pour compléter la dernière remarque de la conclusion précédente, nous proposons ici de valider une articulation constituée d'intervalles plus grand que trois demi-tons. Nous prenons à titre d'exemple les mesures 19 à 30 de la pièce "SPAIN" (main droite), composée par Chick Corea.

### **V.1. décomposition de l'articulation**

L'algorithme d'évaluation du doigté a une complexité de l'ordre de  $O(q^n)$  avec  $q$ =nombre de doigts et  $n$  le nombre de segments. Au delà de 15 segments, les temps d'attente deviennent très importants.

Pour traiter ce problème, nous avons décomposé manuellement les 12 mesures en trois parties : mesures 1-4, mesures 5-7 et mesures 8-12. D'une séquence de 47 notes, nous sommes donc passés à trois séquences ayant respectivement 15, 15 et 17 événements.

J'ai défini un doigté que j'ai pris comme référence.

Voici le doigté de référence (DR) sur les 47 événements :

135 | 4215 | 143 | 12113 | 43542 | 121543 | 2124 | 5432 | 1213143 | 155 | 43 | 1 |

### **V.2. mesures 1 à 4**

Le doigté de référence (DR) 135 | 4215 | 143 | 12113 | (moyenne de facilité : 0.962)

Les doigtés les plus faciles calculés par l'algorithme sont :

135 | 4214 | 143 | 12115 | (moyenne de facilité : 0.983)

145 | 4214 | 143 | 12115 | (moyenne de facilité : 0.983)

Les différences sont indiquées en souligné. Elles concernent l'utilisation du 4 et 5ème doigts au lieu du 3ème et 5ème doigt.

Remarquons que l'algorithme générateur de doigté effectue ses calculs sur 15 notes. Pas une de plus. Pour garantir une certaine cohérence avec la totalité des notes et le doigté de référence, il faut rechercher dans notre liste de doigté ceux de facilité supérieure à 0.962 et terminant par le doigt 3.

On trouve les 3 doigtés suivants :

135 | 4214 | 143 | 12113 | avec une moyenne de 0.965

145 | 4214 | 143 | 12113 | avec une moyenne de 0.965

145 | 4215 | 143 | 12113 | avec une moyenne de 0.962

Les différences avec la référence se situent respectivement sur la seconde et la septième note. Le doigt 4 est conseillé au lieu du doigt 5.

Nous convenons que le meilleur doigté est celui qui se rapproche le plus de notre référence, à savoir 135 | 4214 | 143 | 12113 | avec une moyenne de 0.965.

### **V.3. mesures 5 à 8**

Notre doigté de référence est 43542 | 121543 | 2124 | (infaisable avec notre algorithme)

Les doigtés les plus faciles calculés avec notre algorithme sont

33542 | 142543 | 2125 | avec une moyenne de 0.937

44542 | 142543 | 2125 | avec une moyenne de 0.937

33542 | 121543 | 2125 | avec une moyenne de 0.934

33542 | 131543 | 2125 | avec une moyenne de 0.934

44542 | 121543 | 2125 | avec une moyenne de 0.934

44542 | 131543 | 2125 | avec une moyenne de 0.934

Les différences avec le doigté de référence sont indiquées en souligné.

Remarquons en premier lieu que le changement de doigt sur les deux premières notes (34 dans le doigté de référence) n'est contenu dans aucun des doigtés proposés par notre algorithme. Ceux-ci commencent soit par 33 soit par 44. Cette différence était prévisible car dans la matrice de facilité, les transitions d'un doigt vers un autre (différent du premier) sur une même touche n'est possible que si l'un d'eux est le pouce. Ainsi le doigté 34 est impossible à réaliser sur une même touche. Cela explique que le doigté de référence sur les mesures 5 à 8 est infaisable.

Remarquons ensuite que les doigtés les plus faciles terminent par le doigt 5. Comme précédemment, dans le doigté de référence, nous avons fait usage de la connaissance que nous avons a priori de l'ensemble du doigté. Il est donc souhaitable que le doigté se termine par le doigt 4.

Enfin, le doigté de la 6ème à la 8ème note est 121 dans la référence. Pour l'algorithme, il s'agit dans l'ordre de facilité de 142, de 121 et de 131. Le simplicité du doigté 142 se justifie par les notes qui suivent.

### **V.4. mesures 9 à 12**

Notre doigté de référence est 5432 | 1213143 | 155 | 43 | 1 | avec une moyenne de 0.847.

Les doigtés les plus faciles calculés avec notre algorithme sont

5432 | 1315154 | 155 | 43 | 1 | avec une moyenne de 0.934

5432 | 1415154 | 155 | 43 | 1 | avec une moyenne de 0.934

Les différences de doigté se situent sur la mesure 27.

Dans le doigté de référence, on trouve une certaine logique arithmétique : Le Do est doigté avec le pouce et les autres notes avec successivement les doigts 2, 3 et 4.

L'algorithme pour sa part ne tient compte que de l'écart entre les touches, et pas d'une logique arithmétique. Cela explique que le doigté de référence, avec le passage 121314 n'est pas si "simple" à réaliser.

### **V.5. conclusion**

Pour récapituler, voici le doigté de référence complet :

135 | 4215 | 143 | 12113 | 43542 | 121543 | 2124 | 5432 | 1213143 | 155 | 43 | 1 |

et voici le doigté le plus facile calculé par notre algorithme :

135 | 4214 | 143 | 12113 | 44542 | 142543 | 2124 | 5432 | 1315154 | 155 | 43 | 1 |

Le doigté que propose notre algorithme sur des séquences réduites de l'ensemble des 12 mesures se rapproche sensiblement de celui de référence. Mais pour se rapprocher davantage de la référence, il faudra résoudre les problèmes concernant :

- 1- la transition entre un groupe de notes (paquets de 3 ou 4 mesures) et le groupe suivant,
- 2- les transitions impossibles à réaliser avec notre algorithme (changer de doigt sur une même touche si l'un des doigts n'est pas le pouce),
- 3- inclure dans la définition des doigtés une logique qui ne dépend pas forcément d'une facilité physiologique.

*partition traditionnelle de SPAIN*

# Spain

Chick Core

19 20 21

Musical notation for measures 19, 20, and 21. The piece is in 2/4 time with a key signature of two sharps (F# and C#). Measure 19 features a half note G4 in the treble and a half note G2 in the bass, with a fermata over the G4. Measure 20 has a quarter note A4 in the treble and a quarter note G2 in the bass. Measure 21 has a quarter note B4 in the treble and a quarter note F#2 in the bass.

22 23 24

Musical notation for measures 22, 23, and 24. Measure 22 has a quarter note C5 in the treble and a quarter note E2 in the bass. Measure 23 has a quarter note D5 in the treble and a quarter note F#2 in the bass. Measure 24 has a quarter note E5 in the treble and a quarter note G2 in the bass.

25 26 27

Musical notation for measures 25, 26, and 27. Measure 25 has a quarter note F#5 in the treble and a quarter note A2 in the bass. Measure 26 has a quarter note G5 in the treble and a quarter note B2 in the bass. Measure 27 has a quarter note A5 in the treble and a quarter note C#3 in the bass.

28 29 30

Musical notation for measures 28, 29, and 30. Measure 28 has a quarter note B5 in the treble and a quarter note D3 in the bass. Measure 29 has a quarter note C6 in the treble and a quarter note E3 in the bass. Measure 30 has a quarter note D6 in the treble and a quarter note F#3 in the bass.





---

# Conclusion générale

---

## I. OBJECTIF INITIAL

## II. ETAT D'AVANCEMENT PRATIQUE

II.1. le système GENESIS

II.2. les modes de GENESIS

II.3. les "menus" de GENESIS

1. mode "orchestre"

2. mode "lutherie"

3. mode "instrumental"

4. mode "édition"

a. "menu édition tablature" - b. "menu édition de segments" - c. "menu validation d'une articulation"

## III. ETAT D'AVANCEMENT THEORIQUE

III.1. approche neutre

1. segmentation - 2. représentation et traitement

III.2. référence à la cause

1. validation de l'articulation - 2. approche cohérente

## IV. PERSPECTIVES

IV.1. approfondissement et extensions - IV.2. portée de l'étude

---



## **I. OBJECTIF INITIAL**

L'objectif initial de ce travail a été de concevoir et de réaliser un "éditeur de gestes", pour analyser, représenter et traiter l'objet gestuel capté et mémorisé.

L'intégration de cet éditeur au système CORDIS (système de modélisation et de simulation de l'univers instrumental), entre dans l'objectif de la constitution d'un outil pour la création musicale. Cet outil abandonne (provisoirement) l'ordinateur en tant que moyen direct de représentation de l'univers sonore mais l'aborde comme moyen de représentation et de traitement de ses causes (gestuelles et instrumentales). L'ensemble doit permettre l'exploration de l'espace sonore corrélativement à l'espace gestuel par des transformations sur les objets gestuels mémorisés.

La recherche d'un espace de représentation du geste significatif pour le compositeur nous a alors amené à distinguer trois démarches complémentaires constituant une approche très générale de l'édition :

- approche neutre : les phénomènes gestuels sont étudiés et traités pour eux-mêmes,
- référence à la cause : les phénomènes gestuels sont étudiés et traités en référence à leurs processus générateurs,
- référence à la perception : les phénomènes sont étudiés et traités par rapport aux conséquences perceptuelles (acoustiques et visuelles).

## **II. ETAT D'AVANCEMENT PRATIQUE**

Nous allons situer ici l'état d'avancement pratique du travail par rapport à cet objectif initial et par rapport au système GENESIS.

### **II.1. le système GENESIS**

GENESIS est un ensemble matériel et logiciel. C'est la version concrète actuelle de l'outil pour la création musicale. La réalisation de l'ensemble logiciel revient à O.Raoult, ingénieur système et J.L. Florens, chercheur permanent à l'ACROE. Je suis intervenu pour la réalisation de l'éditeur de geste.

L'ensemble matériel est complexe : il s'agit de plusieurs machines interconnectées, adaptées d'une part aux différents besoins d'une simulation en temps réel et d'autre part à un outil d'aide à la composition.

L'unité superviseur est un VAX 730 : il est utilisé pour "lancer" les programmes dédiés à la simulation et donc gérer les dialogues entre les différentes machines, et pour les programmes dédiés à la composition.

Le terminal graphique est un Evans&Sutherland PS300 (cf. chapitre III, §III.1. représentation avec l'ordinateur). Les programmes sont écrits en langage C sur le VAX et dans un langage graphique orienté objet sur le PS.

### **II.2. les modes de GENESIS**

GENESIS propose à l'utilisateur quatre modes d'activité : le mode "orchestre", le mode "lutherie", le mode "composition" (édition du geste instrumental) et le mode "instrumental". Pour chaque mode, il a fallu réaliser des programmes propres à l'activité correspondante.

1- Pour le mode "orchestre", les programmes doivent permettre de définir la scène instrumentale, c'est à dire de définir, de modifier et de disposer des tablatures dans l'espace et dans la dimension temporelle.

2- Pour le mode "lutherie", les programmes doivent permettre à l'utilisateur de construire et de modifier un "instrument" , c'est-à-dire de construire et de modifier un réseau CORDIS.

3- Pour le mode "instrumental", les programmes doivent permettre à l'utilisateur de réaliser un jeu instrumental en temps réel sur l'"instrument" (défini en mode lutherie), de capter et de mémoriser simultanément les signaux gestuels, et inversement de restituer les signaux mémorisés.

4- Pour le mode "composition" (édition du geste), les programmes doivent permettre de décomposer (segmenter), de représenter et de traiter les signaux gestuels.

Le schéma suivant décrit les transitions entre ces différents modes dans le système GENESIS : le mode d'entrée est le mode "orchestre", à partir duquel l'utilisateur peut passer à une activité de luthier, d'instrumentiste ou encore de compositeur...

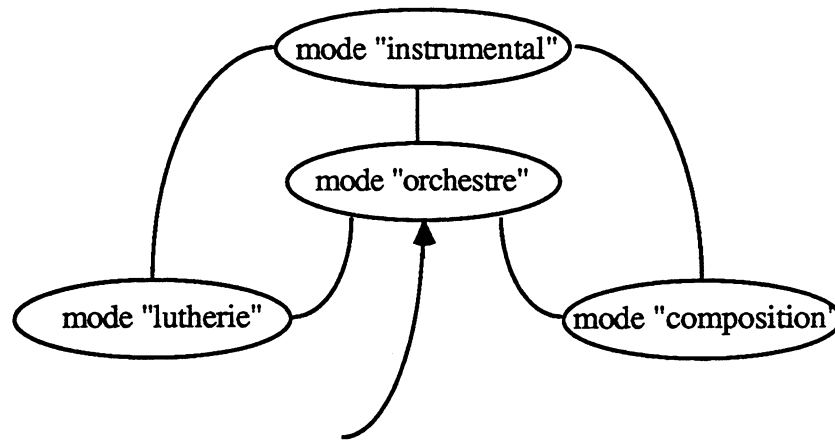


Figure 1. Organigramme de GENESIS. GENESIS propose quatre modes : le mode "orchestre", le mode "lutherie", le mode "composition" (édition du geste instrumental) et le mode "instrumental".

### II.3. les "menus" de GENESIS

O.Daclin [Dac91] a conçu un manuel de l'utilisateur de la version actuelle de GENESIS.

Il décrit les interfaces logicielles que nous avons été obligés de définir (faute de système de gestion de type Xwindow sur le PS) : les barres de menu, comment sélectionner une commande du menu, comment sélectionner un objet ou une partie de cet objet, comment utiliser les "boîtes" de dialogues...

Il introduit ensuite la notion de "menu" (un "menu" comporte un écran de visualisation et un ensemble de fonctions ramenées sous une "barre de menu") puis détaille les "menus" des quatre modes de GENESIS.

Dans ce qui suit nous ne précisons pas les commandes inscrites dans les "barres de menu" et suggérons au lecteur intéressé de se rapporter au manuel de l'utilisateur [Dac91].

La figure suivante représente l'organigramme des différents "menus" de GENESIS.

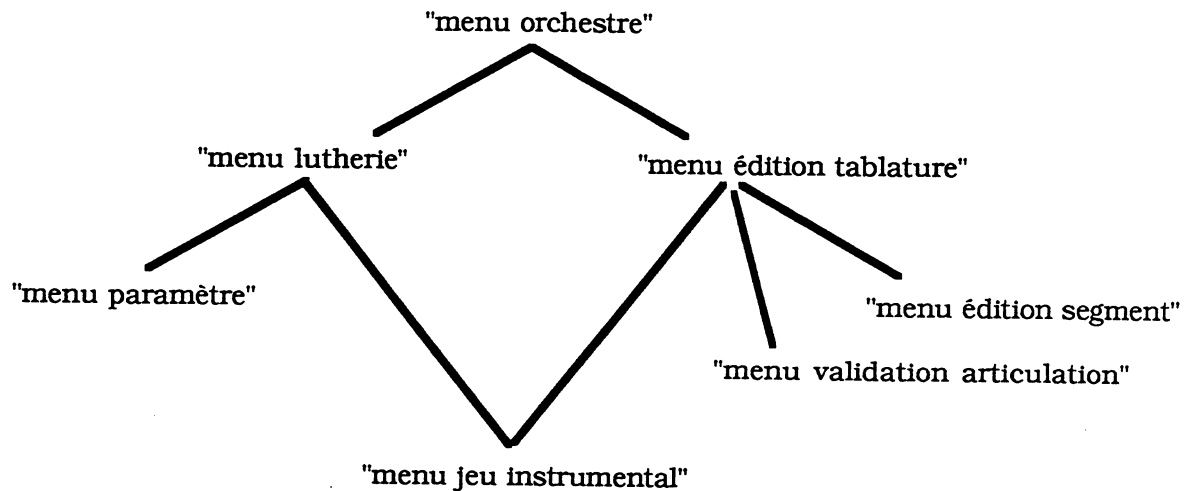


Figure 2. Organigramme des "menus" sous GENESIS. Le système GENESIS propose au moins un "menu" par mode. Par exemple, pour le mode composition, l'utilisateur a accès à trois "menus".

Les fonctions d'un "menu" se répartissent en 4 groupes :

- fonctions pour la visualisation de l'objet graphique (changement de point de vue, affinité, translation, ...),
- fonctions pour l'archivage : sauvegarder sous le nom courant ou sous un nouveau nom, ouvrir un fichier de sauvegarde et restituer l'objet sauvegardé,
- fonctions de transition d'un "menu" à un autre,
- fonctions de traitement utilisant l'objet représenté.

### 1. mode "orchestre"

Le mode "orchestre" a un seul "menu". Il permet de définir les paramètres d'une tablature, c'est-à-dire le nombre de canaux gestuels, la longueur (temporelle) de ces canaux, l'espacement des barres de mesure ainsi que des attributs pour la visualisation. L'utilisateur la positionne dans un espace 2D Oxy (le plan de l'écran). Le système peut gérer plusieurs tablatures sur l'écran, la tablature courante est pointée

par un triangle (pointe en bas). L'ensemble des tablatures est synchronisé à l'origine temporelle (le plan de l'écran).

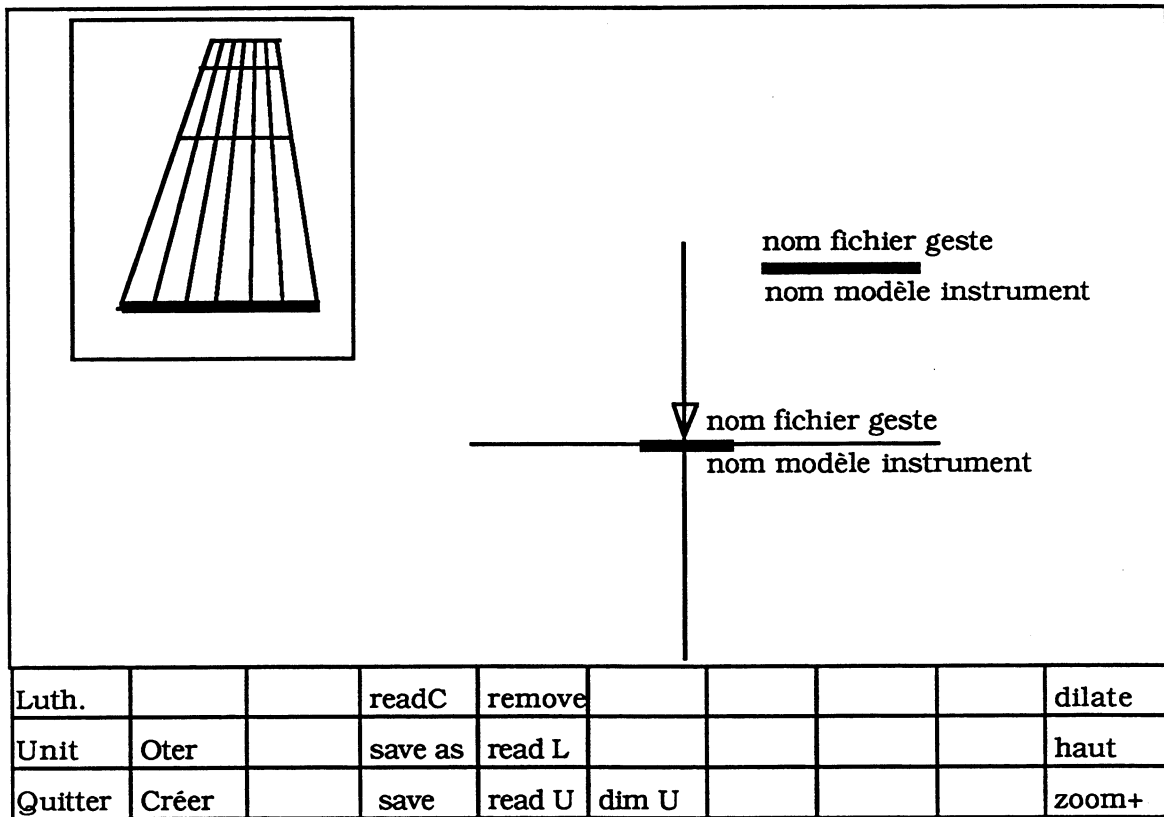


Figure 3. "Menu orchestre". Sur l'écran du "menu orchestre" apparaît un repère Oxy sur lequel sont positionnées les tablatures (2 sur la figure). La tablature courante est pointée par un triangle pointe en bas. La dimension temporelle se profile dans le fond de l'écran. Le rectangle à gauche en haut montre la tablature courante en perspective.

La sauvegarde de la tablature consistera systématiquement à sauvegarder les données suivantes : les paramètres déterminant la tablature, les données gestuelles, le modèle de l'"instrument" sous jacent et ses paramètres.

## 2. mode "lutherie"

Les programmes doivent permettre de construire l'"instrument" en définissant le réseau CORDIS. Ceci peut se faire en écrivant les modules physiques et leurs connexions dans un fichier, sous la forme de chaînes de caractères. Un éditeur de texte serait suffisant. Ceci peut également être fait par une interface graphique, qui permet de connecter des modules entre eux.

Le mode "lutherie" comporte deux "menus".

Le premier "menu" permet de définir un réseau CORDIS. L'utilisateur dispose d'une bibliothèque de modules physiques qu'il peut connecter selon toutes les configurations permises par le langage CORDIS.

Le second "menu" concerne la définition des paramètres de chacun des modules utilisés dans le réseau CORDIS.

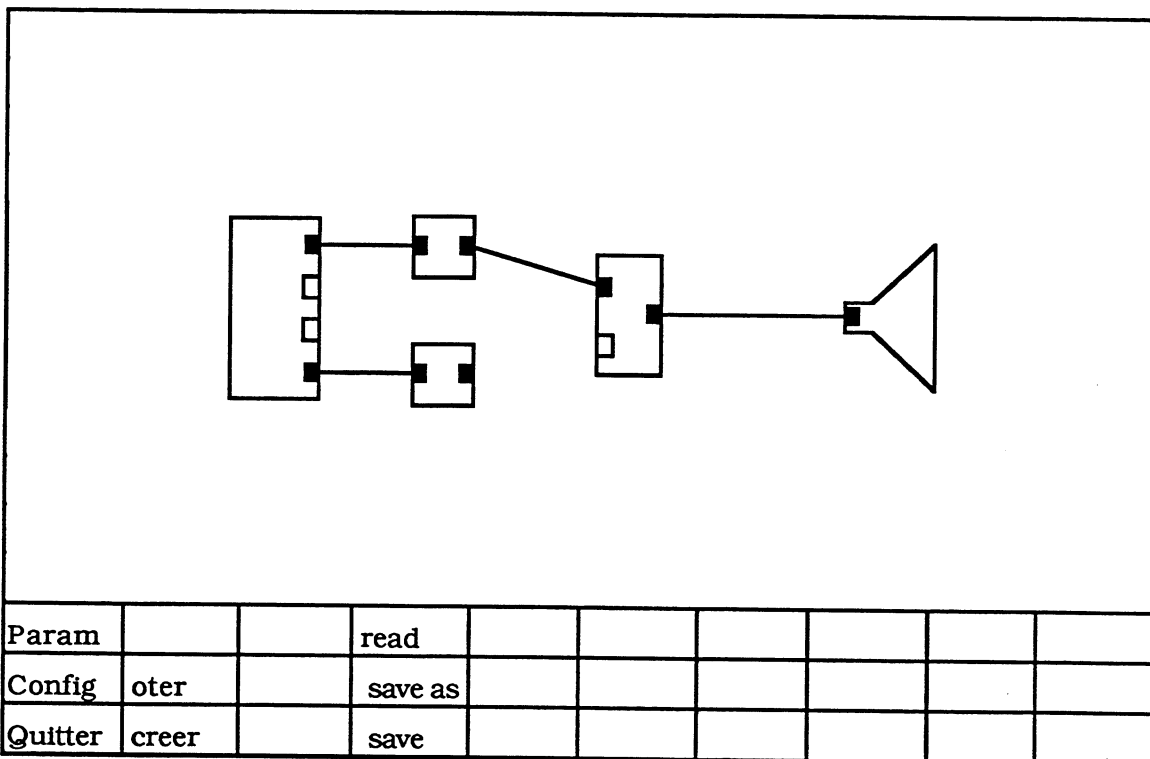


Figure 4. "Menu lutherie". Dans ce "menu" l'utilisateur définit la structure mécanique de l'"instrument", peut le sauvegarder et le restaurer ultérieurement.

### 3. mode "instrumental"

Le mode "instrumental" comporte un seul "menu".

Il donne accès aux fonctions pour le jeu instrumental en temps réel et la lecture en temps différé. Il faut bien entendu avoir défini un "instrument".

La lecture consiste à restituer les entités gestuelles (segments et formules) sélectionnées sur la tablature, c'est-à-dire à reconstruire les signaux gestuels à partir de ces entités puis à les injecter dans le processus de synthèse.

C'est également dans ce "menu" que sont paramétrés les algorithmes de segmentation utilisés lors du transfert des données gestuelles.

#### **4. mode "édition"**

L'éditeur de geste implanté dans le système GENESIS fait la synthèse des travaux présentés dans cette thèse.

Les signaux, captés et mémorisés à partir du clavier rétroactif modulaire, sont en premier lieu segmentés, puis représentés et traités de façon neutre. Un algorithme de validation permet d'évaluer la cohérence d'une articulation avec un doigté possible.

Le mode "édition" comporte trois "menus".

##### **a. "menu édition tablature"**

Le premier "menu" concerne la tablature et l'articulation des segments gestuels.

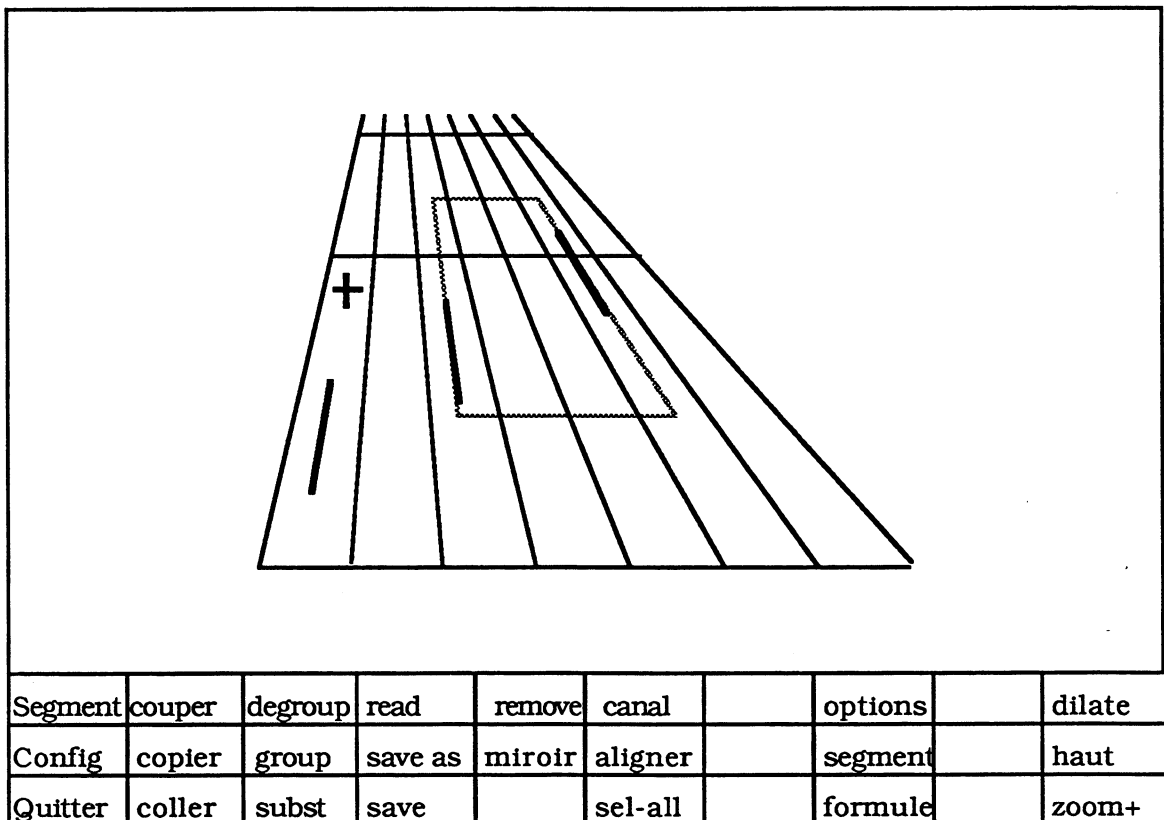


Figure 5. "Menu édition tablature". Sur l'écran du "menu édition tablature" apparaît la tablature courante. Ici elle est visualisée en perspective. L'utilisateur dispose de fonctions qui lui permettent de modifier le point de vue, le point de visée, la distance et l'angle de l'axe de visée, ...

Outre les fonctions de transition vers les autres "menus" et les fonctions d'archivage (sauvegarde et restitution), l'utilisateur dispose de moyens pour désigner et sélectionner des entités représentées sur la tablature, des fonctions classiques de composition (couper, copier, coller, substituer), des fonctions pour associer et dissocier les formules, de fonctions de transformation (déplacer, affiner et renverser) appliquées à un ensemble d'entités gestuelles sélectionnées .

Il dispose également de fonctions pour modifier la visualisation. D'une part des fonctions pour modifier la visualisation de la tablature et des événements qu'elle contient (changement de point de vue, translation temporelle, zoom, changement d'axe de visualisation) et d'autre part des fonctions pour changer la représentation des segments et des formules : couleur et forme (point, barre, triangle pour les segments ; forme simple ou forme détaillée pour les formules).

Une fonction particulière permet de définir la dimension des canaux gestuels de la tablature (initialement tous mono-dimensionnels) et la correspondance de chaque voie du canal avec une voie du transducteur gestuel (par exemple la première dimension du 3ème canal est connectée à la 4ème touche du clavier rétroactif modulaire). Le système vérifiera automatiquement la cohérence entre la dimension d'un segment et le canal sur lequel ce segment doit être placé : si leurs dimensions respectives ne sont pas identiques, le système affiche un message d'erreur et refuse le traitement.

#### b. "menu édition de segments"

Le second "menu" concerne les signaux amplitude/ temps constituant les segments gestuels.

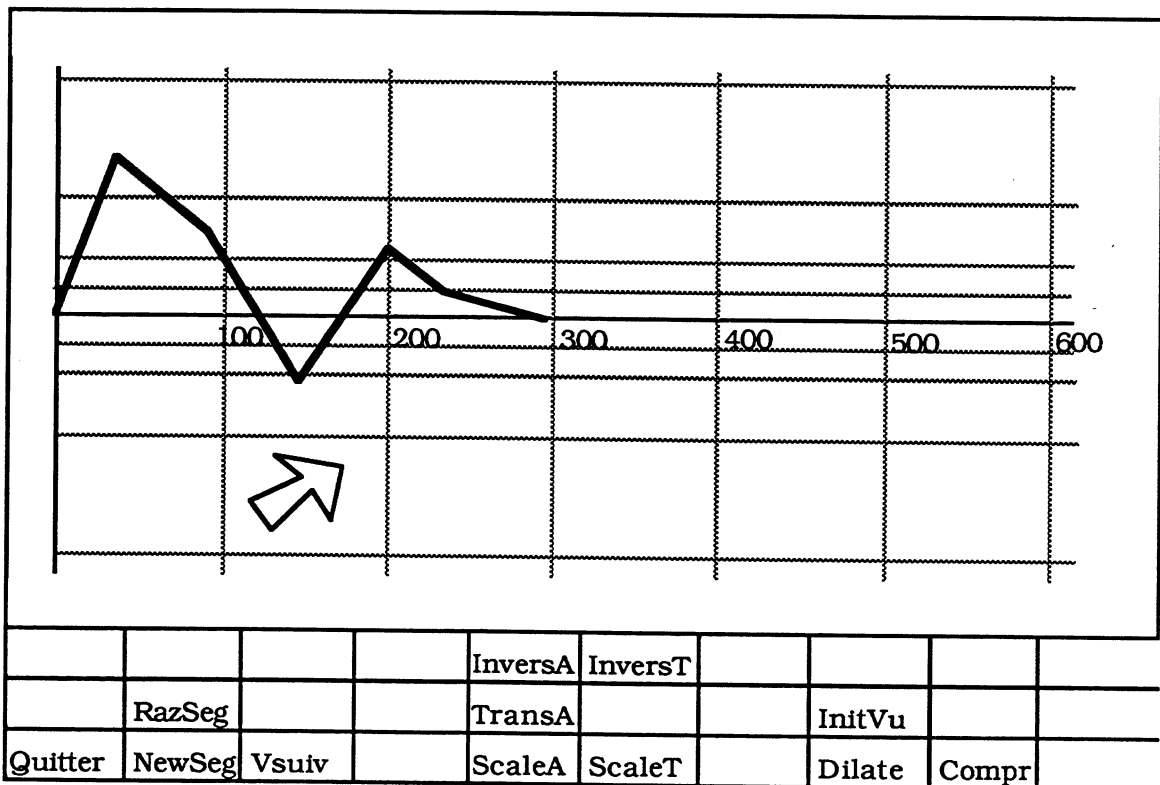


Figure 6. "Menu édition de segments". Sur le "menu édition de segments" apparaît un repère 2D amplitude/ temps. Les signaux constituant le segment sont visualisés sur ce repère.



Les signaux sont visualisés sur un repère amplitude/ temps. Si le segment est multi-dimensionnel (au maximum 3), les signaux sont superposés sur ce même repère. L'utilisateur traite les signaux séparément.

Des fonctions de visualisation permettent de compresser, de dilater et de translater le repère (et les signaux qu'il comporte) dans la dimension temporelle. La translation est contrôlée par un potentiomètre de la boîte à bouton du PS300 (cf. chapitre III, §III.1. représentation avec l'ordinateur).

L'utilisateur dispose de fonctions pour tracer le signal : il définit pour cela une succession de points à l'aide de la souris qui seront reliés par des segments de droite. Il dispose également de fonction pour le transformer : affinité, translation, renversement en amplitude et dans le temps.

Il n'y a pas de fonction d'archivage des segments.

### c. "menu validation d'une articulation"

Le troisième "menu" concerne la validation d'une articulation de segments gestuels.

**Boîte de dialogue  
"Validation de l'articulation"**

<input checked="" type="checkbox"/> main droite	<input type="checkbox"/> main gauche
<input type="checkbox"/> 2 premier doigt	<input type="checkbox"/> premier doigt

facilité minimale

0.85

facilité maximale

1.00

OK       ABORT

Figure 7. "Menu validation d'une articulation". Pour valider une articulation, l'utilisateur doit préciser la main (gauche ou droite), le premier doigt de la main (de 1 à 5) et enfin l'intervalle de coefficients de facilité (réels compris entre 0 et 1).

L'algorithme de validation a pour but de donner la facilité d'un ensemble de segments sélectionnés, et dans le cas où le coefficient de facilité moyen correspondant est

différent de zéro, il donne dans un fichier l'ensemble des doigtés de l'articulation et leur facilité.

L'utilisateur dispose de fonctions pour définir les paramètres de la fonction de facilité et de fonction d'archivage de cette fonction de facilité (sauvergarde et restauration).

Pour valider l'articulation, il doit déterminer un ensemble de paramètres : quelle main, quel premier doigt et éventuellement un intervalle de contraintes sur les coefficients de facilité.

## **III. ETAT D'AVANCEMENT THEORIQUE**

Nous allons situer ici l'état d'avancement théorique du travail par rapport à l'objectif initial.

### **III.1. approche neutre**

#### **1. segmentation**

Notre étude a porté exclusivement sur le jeu instrumental sur clavier. Nous avons défini, en analyse préliminaire, une syntaxe élémentaire décrivant la séquence d'événements et avons caractérisé certains de leur attributs intrinsèques et extrinsèques.

Nous avons alors développé une méthode de segmentation à partir de cette syntaxe. Les éléments produits ont été appelés "segments gestuels".

Pour évaluer les algorithmes de segmentation, la méthode a consisté à comparer les instants de coupure obtenus par le processus de segmentation automatique et ceux donnés par un expert. Le critère de comparaison s'appuie sur un calcul de distances entre les frontières obtenues par segmentation automatique et celles de l'expert. Son efficacité est liée à deux points :

- la spécificité et le nombre limité des formes ciblées,
- la décomposition de l'identification en plusieurs niveaux.

#### **2. représentation et traitement**

Les représentations et traitements proposés opèrent à trois niveaux des segments-gestuels :

- au niveau intrinsèque, les segments sont constitués d'un signal de position échantillonné,
- au niveau extrinsèque, les segments sont positionnés les uns par rapport aux autres,
- au niveau formule, les segments sont regroupés pour former des unités composées indissociables.

Pour chaque niveau nous avons introduit des connaissances gestuelles et instrumentales liées à un jeu sur clavier. L'événement gestuel de base est constitué

d'une succession de trois phases enfoncement-maintien-relâchement et le segment gestuel correspondant d'une succession enfoncement-maintien-amortissement.

D'une approche initialement neutre, nous sommes donc passés à une approche partiellement causale. Les raisons de cette mutation sont autant liées à la nécessité de pouvoir décrire les objets possibles du jeu instrumental sur clavier qu'à celle d'éditer un matériau référencé.

Il semble donc que l'approche neutre, bien que nécessaire, présente des intérêts qu'il faut relativiser en fonction de l'objectif. En ce qui nous concerne, la possibilité de représenter plus explicitement le geste dans sa substance a été déterminant.

### **III.2. référence à la cause**

#### **1. validation de l'articulation**

Nous nous sommes intéressés à l'articulation gestuelle monophonique, issue d'un traitement neutre, et avons proposé un algorithme pour la valider.

L'évaluation de cet algorithme n'est pas achevée.

Nous avons considéré un ensemble de comparaisons avec des doigtés de référence pris sur des gammes de 8, 9 et 10 notes. Une évaluation plus rigoureuse nécessiterait de passer en revue l'ensemble des possibilités offertes par un jeu pianistique monophonique à la main droite.

D'autre part, cet algorithme comporte un ensemble de réductions très important comparé à la complexité de l'organe manuel et aux possibilités motrices qu'elle autorise.

#### **2. approche cohérente**

Nous avons également introduit le principe d'une démarche en référence à la cause.

Notre objectif n'a pas visé à extraire une connaissance fondamentale des processus humains mais à définir l'idée d'une approche permettant d'intervenir directement sur les mécanismes de production.

Il s'agit en fait d'intégrer à l'instrument, et plus particulièrement à l'excitateur, des fonctions actives, qui peuvent être assimilées à un opérateur-simulé.

On peut concevoir aujourd'hui de pousser la référence à la causalité comme logique de production de structures sonores musicales, en considérant les gestes, non plus

seulement comme des traces d'une activité sensori-motrice, mais comme des éléments constitutifs du système de production.

Mais on peut également douter de la possibilité non seulement technique mais théorique d'aller jusqu'au bout d'une telle démarche et d'autre part de son intérêt pour la création même ; car peut-on réellement trouver la justification de toute finalité dans ses origines ? Il semblerait que non.

## **IV. PERSPECTIVES**

### **IV.1. approfondissement et extensions**

Les travaux présentés dans ce mémoire ne couvrent qu'une partie de l'objectif initial. Nous espérons pourtant avoir contribué à l'introduction d'une problématique.

La poursuite de cette étude pourrait s'envisager selon deux directions principales :

- la première devrait permettre d'approfondir chacun des points que nous avons abordés : syntaxe du geste sur clavier, segmentation (problème de bruit, de dérive...), traitement et représentation neutre (réflexion sur la représentation mentale du temps et de l'espace pour la définition des pictogrammes et des traitements plus évolués), validation des entités issues d'un traitement neutre (améliorer la fonction de facilité et diminuer la complexité de l'algorithme) et approche cohérente (outil pour représenter et traiter les intervenants de la situation instrumentale simulée), ...
- la seconde devrait permettre, à l'inverse, d'aller dans le sens d'une extension : type du geste instrumental (jeu sur violon, geste 2D...), type de segmentation (avec ou sans référence à des processus de reconnaissance), méthodes de validation des entités gestuelles résultant d'un traitement neutre, ...

Remarquons aussi que pour réaliser un éditeur de geste instrumental complet, intégrant les approches neutres, en référence à la cause et par rapport aux effets, il faudra les avoir étudiées pour elles-mêmes mais également s'être intéressé à la mise en correspondance de leurs représentations.

### **IV.2. portée de l'étude**

On peut souligner enfin que la portée de cette étude dépasse largement le cadre de son application à la création musicale ou image.

En effet, considérant la question du rapport homme-machine en général, il apparaît d'une part qu'une classe particulière de situations de dialogue peut s'apparenter à une relation de type instrumental - où des actions gestuelles mécaniques s'exercent sur des objets physiques ou des représentations d'objets physiques (exemple : télémanipulation, pilotage, aide à la conception mécanique, ...) - et d'autre part l'exploration de ce mode de dialogue semble prendre aujourd'hui un intérêt manifeste.

La communication par le geste "instrumental" (dans son sens large) peut devenir un mode complémentaire de la communication par la parole ou par l'image. Le canal gestuel, par sa spécificité et sa complémentarité vis-à-vis des autres canaux offre en effet des particularités probablement pertinentes et intéressantes à exploiter (il est plus facile de déplacer manuellement l'image d'un objet physique que de décrire verbalement ou selon tout autre langage son déplacement).

Ainsi, l'analyse du geste instrumental peut-elle apporter une connaissance d'ordre général sur les conditions dans lesquelles le canal de communication gestuelle est susceptible de transmettre le maximum d'information.

Et, plus loin, la représentation et le traitement (édition) des contrôles de type gestuel ouvre des perspectives au niveau de l'apprentissage (par un agent naturel : humain, ou par un agent artificiel : robot), ainsi qu'à celui de la description ou de la programmation de tâches de manipulation ou télémanipulation délicates.

---

## BIBLIOGRAPHIE

---

- [Aro85] AROM (S) - "De l'écoute à l'analyse des musiques centrafricaines" - Analyse Musicale n°1, 4ème trimestre 1985, p. 35-39.
- [Aro88] AROM (S) - "Du pied à la main": les fondements métriques des musiques traditionnelles d'Afrique Centrale" - Analyse Musicale n°10, 1er trimestre 1988, p.16-22.
- [Bac79] BACH (C.P.E) - "Essai sur la vraie manière de jouer des instruments à clavier" - Ed. M&M 1979.\*
- [Bar90] BARRIERE (J.B) - "Devenir de l'écriture musicale assistée par ordinateur : formes, formalismes, aide à la composition" - Analyse Musicale n°20, juin 1990, p.52-68.
- [BP58] BAYLEY (G.B) , PRESGRAVE (R) - "For motion time study" - New-York : MacGraw Hill 1958.\*
- [Bea80] BEAUBATON (D) - "Le contrôle de la main : planification du geste et assistance sensori-motrice" - Institut de Neurophysiologie et de Psychophysiologie de Marseille 1980.\*
- [Bea83] BEAUBATON (D) - "Contrôle proactif et rétroactif de la motricité" - THESE à la faculté des sciences de LUMINY (Aix-Marseille II) 1983.\*
- [Ber75] BERTHOZ (A) - "Rôle de la proprioception dans le contrôle de la posture et du geste - Contribution de quelques récepteurs à la détection des mouvements" 1975.\*
- [BM88] BULOT (R), MELONI (H) - "Reconnaissance des formes et localisation d'événements acoustiques et phonétiques" - J.Acoustique 1 (1988) p. 247-253.



- [Bro82] BROWN (F) - "La musique par ordinateur" - Que sais-je ? Presses universitaires de France 1982.
- [Cad79] CADOZ (C) - "Synthèse sonore par simulation de mécanismes vibratoires - Application aux sons musicaux" - THESE INPG (Institut National Polytechnique de Grenoble) 1979.\*
- [Cad83] CADOZ (C). - " Synthèse sonore par simulation mécanique de l'instrument ; le langage CORDIS " - p.1-4, Paris 1983. 11ème ICA.\*
- [Cad87] CADOZ (C) - "Codage du Geste Instrumental "- Rapport interne, Novembre 1987.\*
- [Cad88] CADOZ (C) - "Instrumental Gesture and Musical Composition" - International Computer Music Conference - Cologne 1988.\*
- [ Cad89a] CADOZ (C) - "Le Modèle Physique - Référence et Artifice dans la Création Musicale à l'Aide de l'Ordinateur" - Rapport de recherche juin 89.
- [ Cad89b] CADOZ (C) - "Le modèle physique dans le création musicale à l'aide de l'ordinateur " - In VIII Colloquio di Informatica Musicale - Cagliari (Italia), Oct. 1989.
- [Cad90a] CADOZ (C) - "Commande Instrumentale et Synthèse sonore - Simulation de l'Univers Instrumental" - Rapport de recherche Janvier 1990.
- [Cad90b] CADOZ (C) - "Réalité du timbre ? Virtualité de l'instrument !" - Analyse Musicale - 1er trimestre 1990.
- [Cad90c] CADOZ (C) - "Modèle physique - Création Musicale et Ordinateur" - Coll. Mod. Phys - Grenoble 1990.
- [Cad91] CADOZ (C) - "CORDIS-ANIMA : système de modélisation et de simulation d'instruments et d'objets physiques pour la création musicale et d'images animées " - Rapport ACROE 1991.\*
- [CLF81] CADOZ (C), LUCIANI (A), FLORENS (JL) - "Synthèse musicale par simulation des mécanismes instrumentaux transducteurs gestuels rétroactifs pour l'étude du jeu instrumental" - Revue d'acoustique N°59 - 1981.\*

- [CLF89] CADOZ (C), LUCIANI (A), FLORENS (JL) - "Responsive Input devices and Sound Synthesis by Simulation of Instrumental Mechanisms : The CORDIS System" - in Music Machine Curtis Road ED. MIT Press, Cambridge Mass 1989.
- [CLF90] CADOZ (C), LISOWSKI (L), FLORENS (JL)- "Clavier rétroactif modulaire" - CMJ (Computer Music Journal) Vol 14, N°2, Summer 1990.\*
- [CR90] CADOZ (C), RAMSTEIN (Ch) - "Capture, Representation and Composition of the Instrumental Gesture "- International Computer Music Conference - Glasgow 1990.\*
- [Cas67] CASTELLENGO (M)- "Peut-on apprécier objectivement le style des pianistes " - G.A.M Juin 1967, N°30 b.
- [Cha65] CHASE (R.A)- "An information-flow model of the organization of motor activity". Part II : Sampling, central processing, and utilization of sensory information. J. Nerv. Ment. Dis., 140, p.334-350 1965.\*
- [Coo79] COOKE (J.D) - "Dependance of human arm movements on limb mechanical properties" - Barin Research, 165 p.366-369 1979.\*
- [Cou76] COURT (R) - "Essai sur les fondements anthropologiques de l'art" - 1976.
- [Cou90] COUTAZ (J) - "Interface homme-ordinateur : conception et réalisation" - Dunod informatique 1990.
- [Dac91] DACLIN (O) - "Représentation et traitement du geste instrumental" - DEA INPG 1991.
- [Dan89] DANNENBERG (R.B) - "A Gesture Based User Interface Prototyping System" - ACM 1989.
- [DAS88] DELYON (B), ANDRE-ROBRECHT (R), SU (HY) - "Expériences en vue du décodage acoustico-phonétique à partir d'une recherche statistique d'événements acoustiques et d'un codage vectoriel" - J.Acoustique 1(1988) p.229-235.

- [Deh88] DEHOUX (V) - "Jazz et musique d'Afrique centrale : une lecture parallèle" - Analyse Musicale, 2ème trimestre 1988, p.38-43.
- [Del88] DELALANDE (F) - "Le geste, outil d'analyse - Quelques enseignements d'une recherche sur la gestique de Glenn Gould" - Analyse Musicale n°10, 1er trimestre 1988, p.43-49.\*
- [Del89] DELALANDE (F) - "La terrasse des audiences du clair de lune : essai d'analyse esthétique" - Analyse Musicale n°16, 3ème trimestre 1989, p.75-84.
- [Djo89] DJOHARIAN (P) - "Simulation de la structure vibrante : la synthèse modale" - Rapport interne, décembre 1989.
- [Dlg85] DELIEGE (I) - "Perception des formations élémentaires de la musique" - Analyse Musicale n°1, 4ème trimestre 1985, p.20-28.
- [Des82] DESCHAUSSEES (M) - "L'Homme et le Piano" - Ed. Van De Velde 1982.
- [Dor67] DORGEUILLE - "Essai d'Analyse Anatomico-physiologique de la technique de clavier" - G.A.M Juin 1967, N°30d.
- [EB88] EMERARD (F), BENOIT (C) - "Base de données prosodiques pour la synthèse de la parole" - J.Acoustique 1 (1988) p.303-307.
- [Fan88] FANTAPIE (HC) - "L'analyse de la partition dans la pratique du chef d'orchestre" - Revue d'Analyse Musicale n°10, 1er trimestre 1988, p.26-30.\*
- [FC84] FLORENS (C), CADOZ (C) - "Geste instrumental, ordinateur, transducteurs gestuels rétroactifs" - 10ème ICMC, p.59-61, IRCAM, 1984.
- [FC91] FLORENS (C), CADOZ (C) - "The physical Model, Modelisation ans Simulation Systems of the Instrumental Universe - In representation of Musical Signals". G. De Poli, A. picciali, C. Roads, Ed. MIT Press, 1991, à paraître.
- [Flo78] FLORENS (J.L) - "Coupleur gestuel rétroactif pour la commande et le contrôle de sons synthétisés" - THESE INPG.\*

- [Flo84] FLORENS (J.L) - "Geste instrumental, ordinateur, transducteurs gestuels rétroactifs " - In 10ème ICMC, p.59-61, IRCAM, Paris 1984.\*
- [Fol87] FOLEY (J) - "Les communications entre l'Homme et l'Ordinateur" - Pour la science, déc. 1987, p. 74-82.\*
- [Gag87] GAGNEPAIN (B) - "De la monodie à la polyphonie : quelques aspects de la conception du temps musical avant l'apparition de la barre de mesure" - Analyse Musicale n°6, 1er trimestre 1987, p.14-21.\*
- [GDP90] GILLE (J-CH), DECAULNE (P), PELEGRIN (M) - "Théorie et calcul des asservissements linéaires". Editions Dunod 1990.
- [GF88] GIBET (S), FLORENS (JL) - "Instrumental gesture modeling by identification with time varying mechanical models" - International Computer Music Conference - Cologne 1988.
- [Gib87] GIBET (S) - "Analyse, Représentation et Traitement du Geste Instrumental - application à la synthèse de sons musicaux par simulation de mécanismes vibratoires" - THESE INPG 1987.\*
- [Hat86] HATWELL (Y) - "Toucher l'Espace - La main et la perception tactile de l'espace " - Presses universitaires de Lille 1986.\*
- [Hag88] HAUPTMANN (A.G) - "Gesture Analysis for Graphic Manipulation" - Technical Report CMU-CS-88-198 , 1988.
- [Hau87] HAURY (J) - "La grammaire de l'exécution musicale au clavier et le mouvement des touches" - Analyse Musicale n°7, 2e trimestre 1987, p.20-26.\*
- [Hau88] HAURY (J) - "Le clavier " - Marsyas, N°5, Mars 1988.
- [Hau90] HAURY (J) - "Le jeu instrumental des androïdes musiciens" - Marsyas n°16, décembre 1990, p.10-19.
- [HI58] HILLER (L.A), ISSAACSON (L.M) - "Musical composition with a high speed digital computer" - J.A.E.S, vol. 6, n°3, 1958, p.56-69.\*

- [HJ75] HECAN (H), JEANNEROD (M) - "Du contrôle moteur à l'organisation du geste" - Ed. MASSON 1975.
- [Hon76] HONNEGER (M) - Science de la musique - Ed. Bordas 1976.
- [Iwa90] IWATA (H) - "Artificial reality with force-feedback : dev. of Desktop Virtual Space with Compact Master Manipulator" - Computer Graphics, vol.24, n°4, August 1990, p.165-170.
- [Lac85] LACORNERIE (P) - "Synthèse des sons par simulation de mécanismes vibratoires : logiciel pour le processeur Cordis-Temps-Réel" - THESE INPG 1985.
- [Lei71] LEIPP (E) - "Acoustique et Musique" - Ed. Masson & CIE 1971.\*
- [Lie67] LIENARD (JS) - "Mécanique et Toucher du Piano" - G.A.M Juin 1967, N°30 c.
- [Luc85] LUCIANI (A) - "Un outil informatique de création d'images animées: modèles d'objets, langage, contrôle gestuel en temps réel" - THESE INPG 1985.
- [LW88] LALLOUACHE (M.T), WORLEY (C) - "Saisie, édition et traitement d'images et de signaux articulatoires : lèvres et mâchoires" - J.Acoustique 1 (1988) 215-220.
- [Mac71] MACHABEY (A) - "La notation musicale - Que sais-je ?" Presses universitaires de France 1971.
- [Mac70] MACNEILAGE (P) - "The motor control of serial ordering in speech" - Psychol. Rev., 77, p.182-196 1970.
- [Mar88] MARTEAU (PF) - "Cibles et Trajectoires acoustiques - Application à la segmentation et à l'étiquetage automatique du signal de la parole" - THESE INPG 1988.\*
- [Mar78] MARTIN (S) - "Le langage musical" - Sémiotique des systèmes 1978.\*
- [Mat69] MATHEWS (M) - "The Technologie of Computer Music" - The MIT Press 1969.\*

- [MD89] MACDAMS (S), DELIEGE (I), - "La musique et les sciences cognitives" - Ed. Pierre Mardaga 1989.\*
- [Men88] MENDEL (G) - "Le jeu de l'instrumentiste à cordes : un mouvement corporel global, fonctionnel et expressif" - Analyse Musicale n°10, 1er trimestre 1988, p.36-41.\*
- [Mer72] MERTON (P.A) - "How we control the contraction of our muscles" - Sci. Amer, 226, p.30-37, 1972.
- [MR90] MOOG (R), RHEA (T) - "Evolution of the Keyboard Interface : the Bösendorfer 290 SE Recording Piano and the Moog Multiply-Touch-Sensitive Keyboards" - CMJ VOL.14, N°2, Summer 1990.\*
- [PAK89] PERRIER (P), ABRY (C), KELLER (E) - "Vers une Modélisation des Mouvements du Dos de la Langue" - J.Acoustique 2 69-77, 1989.\*
- [PB80] PAILLARD (J), BEAUBATON (D) - "De la coordination visuo-motrice à l'organisation de la saisie manuelle" - 1980.
- [Pic77] PICINBONO (B) - "Eléments de théorie du signal" - Dunod Université, 1977.
- [Pol83] DE POLI (J) - "A tutorial on Digital Sound Synthesis Techniques" - Computer Music Journal, Volume 7, Number 4, p.8-26, 1983.
- [Ram89a] RAMSTEIN (Ch) - "Analyse, représentation et traitement du geste instrumental" - DEA INPG 1989.\*
- [Ram89b] RAMSTEIN (Ch) - "Editeur d'objets filaires 3D corrélés" - Rapport de recherche, 1989.
- [Rao91] RAOULT (O) - "Etude d'un oscillateur élémentaire" - Rapport interne, 1991.\*
- [RC84] RODET (X), COINTE (P) - "FORMES : Composition ans Scheduling of Process"- Computer Music Journal 8(3) p.32-50, 1984.

- [RPB84]       RODET (X), POTARD (Y), BARRIERE (JB) - "The CHANT project : from the synthesis of the singing voice to synthesis in general" - Computer Music Journal p.15-31, 1984.
- [Ris86]       RISSET (JC) - "Aspects de la perception du temps en musique" - Bull. Audiophonol. Ann. Sc. Univ. Franche-Comté, Vol. 2NS, No4, 261-278, 1986.
- [Rou79]       ROUX (M) - "La musique" - Ed. Retz, 1979.
- [Rou88]       ROUTZ (D) - "La pédagogie du chef d'orchestre" - Analyse Musicale n°1à, 1er trimestre 1988, p.23-25.
- [Sad85]       SADAI (Y) - "Analyse musicale : par l'oeil ou par l'oreille" - Analyse Musicale n°1, 4ème trimestre 1985, p.13-19.
- [Sch66]       SCHAEFFER (P) - "Traiter des Objets Musicaux" - Essai interdisciplinaire - Ed. Seuil, 1966
- [Sco79]       SCOTT KELSO (J.A) - "On the Coordination of Two-Hand Movements" - Journal of Experimental Psychology Vol 5, N°2, p.229-238, 1979.
- [Smi72]       SMITH (S) - "A musician's approach to computer music" - J.A.E.S, vol.20, n°1, p.7-14, 1972.
- [Sto78]       STOIANOVA (I) - "Geste-texte-musique" - Union générale d'édition, 1978.
- [Str85]       STRAWN (J) - "Digital Audio Signal Processing" - William Kaufmann, INC, 1985.
- [VM88]       VICART (D), MICLET (L) - "Reconnaissance de parties transitoires dans le signal de parole continu" - J.Acoustique 1 317-320, 1988.\*
- [Xen63]       XENAKIS (I) - "Musiques formelles - Nouveaux principes formels de composition musicale" - Ed. Stock/ Musique 1963.