

Historique de la synthèse sonore par modèles physiques

Introduction

Le lien entre science et musique a été significativement renforcé au cours du vingtième siècle, selon deux processus duaux. D'une part, les progrès technologiques ont permis une étude plus approfondie du son et de la perception sonore, avec l'épanouissement de l'acoustique, de l'analyse du signal, de l'informatique et de la psychoacoustique. Réciproquement, la meilleure compréhension des phénomènes sonores, ainsi que les développements technologiques et tout particulièrement l'informatique, ont engendré un développement spectaculaire de la création musicale.

La synthèse par modèles physiques, qui vise à modéliser les structures qui produisent le son, est une approche originale qui conjugue ces deux aspects. Si la formalisation des lois de l'acoustique et de l'évolution du son permet la modélisation des systèmes vibrants, la simulation à partir de cette modélisation permet réciproquement d'approfondir, de valider ou d'invalidier ces lois. En plus d'imiter des systèmes réels, la synthèse sonore par modèles physiques possède un intérêt musical majeur : l'agencement des éléments de base de la modélisation dans un ordre arbitraire permet la création d'instruments virtuels qui produisent des sons inédits et « authentiques » - ce en quoi cette synthèse se distingue des autres. Les interactions entre les différentes parties de l'instrument, comme les cordes d'une guitare qui vibrent par sympathie, le dotent d'un réalisme tout particulier.

La possibilité de contrôler ces instruments de manière intuitive et **expressive** à l'aide d'interfaces multisensorielles, comme le Transducteur Gestuel Rétroactif de l'ACROE, est un autre atout qui suscite un intérêt croissant chez les compositeurs.

La commercialisation récente de synthétiseurs basés sur la modélisation physique, comme le synthétiseur VL1 de Yamaha, montre que la synthèse par modèles physique est promise à un bel avenir.

Ce rapport récapitule et développe certaines étapes fondatrices de la synthèse par modèles physiques, avant de conclure sur les avantages et les limites de cette synthèse.

I. Les origines

Les origines de la synthèse par modèles physiques datent des années 1960, aux Etats-Unis. En 1963, **M. Mathews, J. Kelly** et **C. Lochbaum** parviennent à synthétiser par ordinateur une chanson à partir d'une simulation du conduit vocal. En 1967, **P. Ruiz**, à l'université de l'Illinois et en collaboration avec les laboratoires Bell, effectue une thèse sur la simulation des cordes frottées à partir de la résolution numérique d'équations différentielles. Ses travaux sont ensuite repris et approfondis par **L. Hiller** et **J. Beauchamp**. Parallèlement, **E. Ferreti**, de l'université d'Utah, fait une tentative infructueuse de commercialisation de musique assistée par ordinateur basée sur une modélisation physique.

La lenteur et le coût de l'informatique à l'époque ont causé l'échec relatif de ces tentatives et il faudra attendre 1978 pour que **C. Cadoz, A. Luciani** et **J.L. Florens** de l'ACROE mettent au point le formalisme CORDIS-ANIMA, qui consiste à judicieusement représenter une structure vibrante par un ensemble de composants mécaniques de base –des éléments matériels et des liaisons– et insuffle à la synthèse par modèles physiques une nouvelle jeunesse.

J.M. Adrien, de l'IRCAM, en 1985, a posé les bases de la synthèse modale qui a donné lieu aux logiciels MOSAÏC et MODALYS.

Une forme totalement différente de modélisation physique apparaît en 1983, lorsque **K. Karplus** et **A. Strong** inventent la synthèse par guides d'ondes, qui permet une économie de calculs suffisante pour simuler

simultanément plusieurs sons de corde pincée en temps réel sur un processeur de l'époque.

Ces procédés sont détaillés dans les sections suivantes.

II. L'approche de Pierre Ruiz

Les travaux de Pierre Ruiz, repris ensuite par L. Hiller et J. Beauchamp, sont à l'origine du courant de la modélisation physique le plus répandu.

L'objet vibrant est défini comme un ensemble de masses reliées par des ressorts. Ces systèmes peuvent constituer une approximation des systèmes réels, leur précision dépendant largement de leur complexité. Le modèle est censé reproduire les caractéristiques de l'objet vibrant : dimensions, caractéristiques physiques, conditions initiales et conditions aux limites, ainsi que son comportement en régime transitoire. L'objet vibrant peut être représenté par un modèle qui reproduit *grosso modo* son apparence géométrique et/ou ses propriétés topologiques : modèle linéaire pour une corde, modèle en étoile pour une membrane ou modèle en treillis pour un volume. Ceci n'est cependant pas une nécessité, comme nous le verrons dans la partie consacrée à Cordis-Anima.

Les vibrations du modèle sont caractérisées par un ensemble d'**équations différentielles** et de **contraintes**, qui sera résolu numériquement lorsque les caractéristiques de l'excitation seront connues. La résolution numérique se fait par **approximations successives**, et nécessite l'introduction *a priori* d'une solution supposée proche de la solution réelle.

L'approche de P. Ruiz est cependant trop coûteuse en temps de calcul, insuffisamment précise – en particulier pour les transitoires – et ses modèles sont trop difficile à contrôler pour que sa méthode puisse satisfaire des attentes musicales.

Ces exigences nécessitent de repenser, d'une part, les caractéristiques des modèles et d'autre part la méthode de résolution des équations

différentielles. De ce point de vue, l'apparition de Cordis-Anima en 1978 fut un progrès incontestable dans l'histoire des modèles physiques.

III. Cordis-Anima

Développé par Claude Cadoz, Annie Luciani et Jean-Loup Florens, de l'ACROE, depuis 1978, Cordis-Anima est un formalisme permettant la simulation éventuellement temps réel d'objets physiques. Une des forces de Cordis-Anima est qu'il est intrinsèquement axé sur les **échanges** entre ses éléments, ce qui le rend adapté à un contrôle multisensoriel. Il se décompose en Cordis, dédié aux structures sonores, qui a donné naissance au logiciel Genesis, et Anima, dédié aux images animées, dont est issu Mimesis.

III.1. Structure des objets vibrants

Les modèles sont constitués d'éléments de matière (MAT) reliés par des éléments de liaisons (LIA). A l'inverse de l'approche de P. Ruiz, les algorithmes de calcul sont **intrinsèques** à ces éléments, ce qui minimise les temps de calcul et facilite la simulation en temps réel. Plus exactement, les réponses des éléments sont calculées à chaque pas d'échantillonnage, ce qui produit des résultats suffisamment fins pour donner l'impression de phénomènes continus.

Un élément **MAT** prend en paramètre une force et renvoie une position ; pour un élément **LIA**, c'est l'inverse.



L'entrée d'un point MAT peut être reliée à la sortie de plusieurs LIA, il reçoit alors la somme des forces d'entrée. En revanche, l'entrée d'une LIA ne peut être reliée qu'à une sortie de MAT, car elle ne peut être située à

plusieurs positions différentes en même temps. De même, deux points MAT et deux points LIA ne peuvent être reliés entre eux.

Les MAT et les LIA peuvent d'autre part être regroupées en sous-ensembles interconnectés, autrement dit les structures sont **modulaires**. Ceci permet entre autres de contrôler le modèle à différents niveaux.

Le modèle obtenu peut être multidimensionnel ; c'est le cas avec Anima, où l'apparence visuelle des images animées est fondamentale. Dans le cas de Genesis au contraire, le choix a été fait d'une topologie unidimensionnelle.

III.2. Structure des modèles dans Genesis

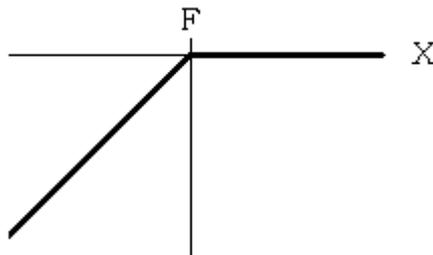
Les éléments MAT et LIA se déclinent dans Genesis en plusieurs éléments, décrits ci-dessous. Dans ce qui suit, n représente l'instant courant, F une force et X un déplacement

Pour les MAT :

- la masse ponctuelle (**MAS**) : localisée dans l'espace, sans dimensions, elle a pour algorithmme $X_n = \frac{1}{M} F(n-1) + X(n-1) + X(n-2)$
- le solide (**SOL**) : sa réponse est un déplacement nul, il permet de donner une référence au modèle
- la cellule (**CEL**) : elle est composée d'une masse reliée à un point fixe par un ressort et un frottement, c'est un des éléments les plus utiles des modèles Genesis.

Les LIA se décomposent en :

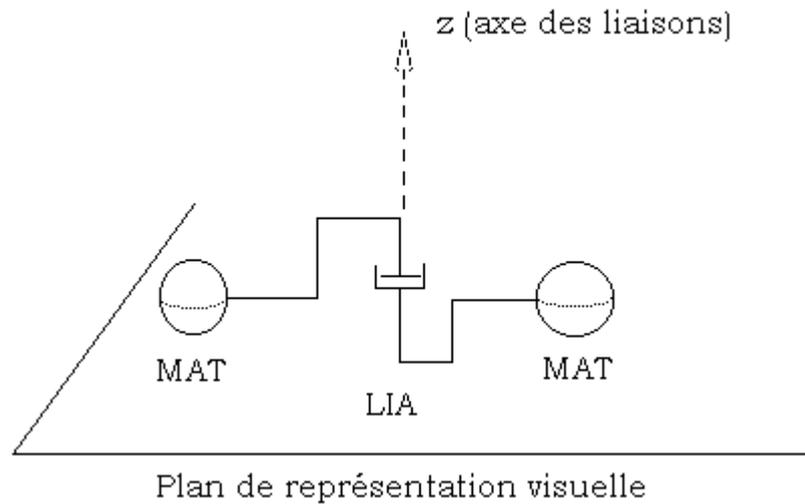
- le ressort (**RES**) d'algorithme $F = K(X_1(n) - X_2(n))$
- le frottement (**FRO**) qui constitue un apport fondamental par rapport à l'approche de Ruiz
- le ressort + frottement (**REF**), constitué d'un ressort et d'un frottement
- la butée (**BUT**) : la butée permet de prendre en compte les chocs physiques entre particules ; sa réponse en force par rapport au déplacement obéit à une loi non linéaire du type :



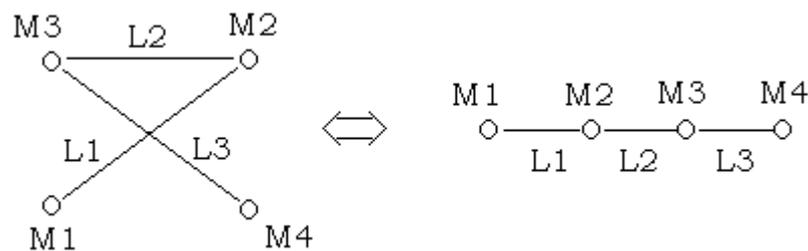
- la liaison non linéaire (**LNL**) : cet élément permet de prendre en compte les non-linéarités du modèle en dessinant « à la main » l'allure de la courbe F en fonction de X .
- la liaison conditionnelle (**LI Cond**) : cette liaison prend en compte ses états aux instants précédents
- le microphone (**SoX**) : le microphone est un capteur de déplacement, sa sortie traduit les vibrations en un point donné.

Les paramètres de ces éléments (caractéristiques physiques, vitesse et position initiales) sont modifiables hors de la simulation par l'utilisateur.

Dans Genesis, ces liaisons se déroulent selon **un seul axe z**, invisible sur la représentation visuelle de construction des modèles. Cette astuce allège considérablement les calculs, en évitant des calculs de racine carrée.



La topologie de Genesis rend ainsi les deux représentations visuelles suivantes équivalentes :



L'unidimensionnalité des modèles est cependant compatible avec une **spatialisation** des sons produits, comme le montrent des travaux récents menés par O. Meunier.

Il pourrait sembler à première vue que cette représentation rende délicate la modélisation d'objets multidimensionnels. Or il n'en est rien : le théorème de Lanczos, mis en application par des travaux récents de l'ACROE, montrent que pour tout ensemble de modes - fréquence, amplitude, amortissement - il existe un modèle **linéaire** de type masses et ressorts amortis qui reconstitue ces modes. Autrement dit, tout son percussif peut être modélisé par l'excitation d'une corde inhomogène !

Un certain nombre de **structures « standard »** sont cependant présentes dans Genesis, qui tiennent compte de la structure géométrique et topologique des objets de référence : structures en anneaux pour les cloches, en cercles concentriques pour les membranes, en spirale ou en structures irrégulières pour des matières inhomogènes etc.

Il est à noter que pour minimiser les temps de calcul, les algorithmes choisis reviennent à une méthode d'Euler explicite, d'où un son **légèrement différent de celui qu'aurait produit l'objet de référence**. Ceci se justifie dans l'optique de la création musicale, où la proximité aux sons réels n'est pas d'une importance capitale.

III.3. « Multisensorialité »

La modularité qui caractérise Cordis-Anima s'étend à la possibilité d'intégrer le modèle simulé dans un environnement **multisensoriel** : ainsi un objet sonore peut-il être simultanément vu, entendu et certains de ses paramètres contrôlés par le geste.

C'est le but du Transducteur Gestuel Rétroactif, ou **TGR**, de Jean-Loup Florens, un clavier dont chaque touche peut contrôler un paramètre du modèle, et qui réciproquement renvoie des informations à l'utilisateur grâce à la présence de moteurs magnétiques sous les touches. Dans Cordis-Anima, le TGR est considéré comme un ensemble de LIA ; l'entrée de chaque LIA est la force envoyée du modèle aux touches et la sortie correspond à la position de la touche activée par l'utilisateur.

Appliqué à un instrument de musique virtuel, le TGR permet d'ajouter à la chaleur propre aux sons produits par modèles physiques une grande expressivité, et de contrôler très intuitivement le modèle.

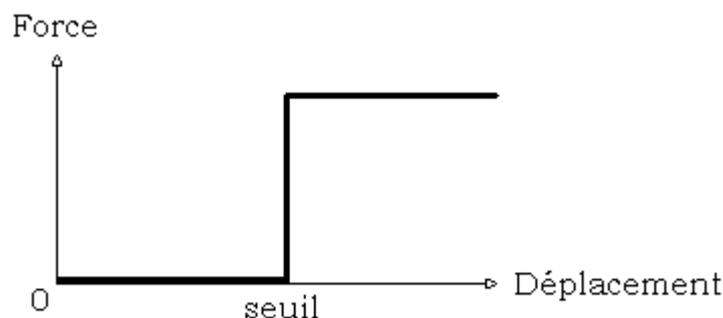
III.4. Que peut-on ajouter à Genesis ?

D'importants efforts ont été récemment consacrés à l'**interface** de Genesis et à l'intégration de fonctions de haut niveau dans un **éditeur évolué**. Ces fonctionnalités sont en effet nécessaires à l'exportation de Genesis hors de l'ACROE, car elles le rendent plus utile et plus sympathique.

Il est par ailleurs possible d'étendre les algorithmes des points MAT et LIA à **d'autres synthèses** par modèles physiques, comme les algorithmes de Karplus-Strong, grâce aux modules « ad-hoc ».

Néanmoins le formalisme Cordis-Anima est suffisamment bien conçu pour permettre l'adaptation aux modèles actuels d'une extension non mécanique qui démultiplierait leurs possibilités et pourraient leur conférer un net avantage sur les autres procédés de synthèses, même par modèle physique. Comment ? En transformant les structures vibrantes en structures pensantes !¹

La connexion d'éléments de matière par des liaisons n'est en effet pas sans faire penser à des **réseaux de neurones** reliés par des axones. Ce procédé informatique a fait ses preuves dans l'intelligence artificielle, et plus particulièrement dans l'apprentissage et la classification de modèles. Chaque neurone peut recevoir la « force » de plusieurs axones afférents, et calcule grâce à un algorithme intrinsèque un « déplacement », qu'il transmet à l'axone sortant. Les axones disposent eux aussi d'algorithmes, qui se limitent souvent à une porte binaire.



¹ Voir aussi l'article de N. Szilas et C. Cadoz, « Physical Models that learn », ICMA 1993

En incluant des modules de type « réseaux de neurones » parmi des structures vibrantes, par exemple en couplant neurones et MAT, axones et LIA, il devient possible de créer un instrument qui apprend ! L'apprentissage peut par exemple porter sur plusieurs enregistrements d'un même instrument, en optimisant à chaque enregistrement le modèle de manière à le faire ressembler le plus possible à l'instrument de référence. La prise en compte d'un enregistrement peut par exemple se faire à l'aide d'une analyse modale couplée à l'algorithme de Lanczos.

Une autre voie intéressante est l'utilisation **de chaînes de Markov cachées**, dont la structure fait penser aux modèles de Cordis-Anima. Ce procédé s'est imposé de manière spectaculaire dans la reconnaissance vocale et la reconnaissance de la parole, grâce à des algorithmes particulièrement performants. Il repose sur la théorie de l'information et les probabilités.

Des « états » sont reliés entre eux par des « transitions ». Chaque état et chaque transition sont constitués d'une fonction de densité de probabilités, respectivement appelés « probabilité d'émission » d'un état et « probabilité de transition » entre deux états. Cette représentation est modulaire, puisque deux ensembles d'états et de transitions peuvent être reliés entre eux à un niveau supérieur.

L'intérêt des chaînes de Markov cachées est qu'elles ne prennent pas de décisions sur des critères *a priori*, mais qu'elles élaborent elles-mêmes leurs critères de décision à partir de l'apprentissage qu'elles ont reçu.

Une application est par exemple la réalisation d'un modèle Genesis dont les différents modules représentent différentes matières – bois, métal, plastique... -. La création d'une chaîne de Markov à plusieurs modules auxquels sont donnés en apprentissage les enregistrements de sons issus des différentes matières répond à ces exigences. Celle-ci réalise en effet une classification automatique des matières en catégories selon les critères qui lui semblent les plus probables. La classification de tout nouvel enregistrement dans un des modules revient alors à un simple calcul de probabilités. Cela semble surnaturel, et pourtant cela fonctionne !

Certes, il existe des applications musicales ; le couplage des états avec des MAT et des transitions avec des LIA permet la synthèse de sons à partir de chaque catégorie. Mais l'intérêt scientifique n'est pas négligeable non plus, car les différentes classes séparent en réalité des invariants acoustiques de matière.

IV. Mosaïc, Modalys et la synthèse modale

Mosaïc est un logiciel de synthèse modale conçu par Jean-Marie Adrien et Joseph Morrison de l'IRCAM. Les objets de base sont représentés par leur **décomposition modale** – c'est à dire l'ensemble des caractéristiques de leurs partiels – ce qui permet une représentation d'assez haut niveau. Ainsi les cordes, les membranes, les barres ou les colonnes d'air sont des objets élémentaires, au même titre que les excitateurs – archet, marteau – dont il reste à préciser les caractéristiques physiques.

Ces éléments de base peuvent à leur tour être assemblés selon des règles dont les paramètres sont modifiables. La première méthode consiste à mélanger leurs caractéristiques, c'est l'hybridation. Une autre méthode est de les connecter tout en préservant leurs structures respectives, par des moyens aussi divers que le collage, le frottement, le pincement, le frappement ou d'autres procédés plus spécialisés, comme l'interaction « anche simple » pour les instruments à vents. Enfin, il est possible d'effectuer des actions sur les objets, comme l'action « touche » pour les instruments à corde, l'action « trou » qui simule la présence d'un trou pour les instruments à vent ou la vitesse d'un élément. La présence de contrôleurs dynamiques permet enfin la modification du modèle en cours de simulation.

Un des intérêts de ce langage est la possibilité pour l'utilisateur de programmer ses propres actions et contrôleurs.

Modalys, successeur de Mosaïc, a été créé par l'équipe de René Caussé et Gerhard Eckel. Il permet de contrôler plus finement les paramètres

modaux des éléments et est davantage orienté temps réel, mais le principe de fonctionnement reste le même.

Malgré leurs qualités et les optimisations qu'ils ont reçu, les formalismes Cordis-Anima et Modalys sont exigeants en temps de calcul. La synthèse par guides d'ondes apporte une alternative plus économique.

V. La synthèse par guides d'ondes

La synthèse par guides d'ondes se situe à la limite de la synthèse par modèles physiques, car elle fonctionne par **filtrage de sources**. Elle a été découverte par Alex Strong en 1983. La simplicité de son fonctionnement lui confère une économie de calculs appréciable pour des applications musicales et en particulier temps réel, de plus les résultats sont réalistes.

Le premier son issu de cette synthèse a été un son de corde pincée. En initialisant une table d'onde avec des valeurs aléatoires pour simuler un bruit blanc, puis en effectuant à chaque étape p la moyenne M_p de la valeur courante de la table et de la valeur calculée précédente, le son obtenu est d'amplitude décroissante, et ressemble beaucoup à un son de corde pincée. La vitesse de lecture de la table définit la hauteur du son, et l'amplitude de ses composantes définit son intensité.

Kevin Karplus a apporté deux améliorations majeures à cet algorithme. D'une part, l'alternance de moyennes et d'opposées de moyennes selon une certaine probabilité permet de synthétiser des sons percussifs. D'autre part, le choix à chaque étape d'effectuer un calcul ou au contraire de préserver une valeur de la table d'onde selon une certaine probabilité permet d'introduire une notion de durée d'amortissement indépendante de la vitesse de lecture de la table d'onde.

David Jaffe et Julius Smith ont quant à eux résolu le problème de la justesse des notes inhérent aux lectures de tables d'ondes, ainsi que le problème de la dynamique des sons – modification du timbre selon l'intensité

– par l’introduction de deux filtres. Ils se sont penchés à d’autres problèmes encore, comme l’attaque du son ou la vibration de cordes par sympathie.

C’est Julius Smith qui a relié cette synthèse à la synthèse par modèles physiques, en en donnant une interprétation en termes de propagation d’ondes dans un ensemble de cylindres de section variable. Cette observation a donné lieu à des dispositifs à deux lignes de retard, qui permettent la simulation d’ondes se propageant en sens inverse comme dans une corde. La clarinette de S. Hirschman, P. Cook et J. Smith, réalisée entièrement par guides d’ondes, montre la capacité de cette synthèse de simuler des phénomènes complexes.

Perry Cook a créé quant à lui un instrument de contrôle – le « Meta-Wind Instrument Controller » – qui possède un intérêt musical majeur. Il a d’autre part étendu le principe des guides d’ondes à des processus stochastiques, qui permettent la modélisation de sons percussifs divers.

VI. Comparaison entre les guides d’ondes et les autres modèles physiques

Les guides d’ondes sont plus adaptés que les modèles physiques « purs » à la synthèse de sons simples, comme des sons de corde pincée, car ils sont à la fois convaincants, contrôlables et économiques.

En revanche, si cette synthèse est astucieuse, elle n’en reste pas moins limitée et manque de généralité, à l’inverse des autres synthèses par modèles physiques décrites ici. Notamment, la possibilité d’étendre les modèles vibrants à des procédés non mécaniques – voir le paragraphe III.4. – est une possibilité hors de portée des guides d’ondes.

Une solution est donc d’intégrer au sein des modèles physiques « classiques » des modules de guides d’ondes effectuant des opérations spécialisées. Cordis-Anima, en particulier, permet tout à fait ce type de manipulations.

Conclusion : avantages et limites des modèles physiques

Sur la nature des sons créés

Par nature, les modèles physiques créent des sons que notre perception accepte la plupart du temps comme des sons musicaux « naturels », même s'ils proviennent d'instruments virtuels, en raison des lois acoustiques qui les régissent. Sauf entraînement particulier, la reconnaissance des timbres est en effet un préalable nécessaire à l'écoute musicale. C'est le problème qui s'était posé à Pierre Schaeffer dans ses compositions de musique concrète ; les auditeurs ne pouvaient dissocier l'écoute de ces œuvres d'une recherche automatique de la causalité des sons. A l'inverse, ces lois acoustiques limitent l'éventail des possibilités de création sonore. Ainsi, les musiques spectrales ne sont clairement pas adaptées à ce type de synthèse. Cependant, la capacité des modèles physiques d'imiter et d'étendre les possibilités des instruments acoustiques traditionnels, ou de créer des « mutants » qui sonnent bien, est un atout apprécié de nombreux compositeurs, parmi lesquels F. Nicolas, K. Saariaho ou G. Gavazza.

Sur la possibilité de contrôler le modèle

Un autre atout de la synthèse physique est que les paramètres du modèle ont une signification palpable, ce qui permet une manipulation plus ou moins intuitive du modèle. Plus ou moins, car les paramètres physiques ne sont pas nécessairement la préoccupation des compositeurs. Il faut donc pouvoir contrôler le modèle à un niveau plus élevé, qui coordonne les autres paramètres. Ceci ne pose toutefois pas de problème théorique insurmontable, car les paramètres physiques du modèle sont étroitement liés à ceux du son. Cependant, quel que soit le mode de synthèse adopté, la présence d'une commande gestuelle intelligemment conçue est probablement suffisante à rendre au son musical son expressivité.

Bibliographie

ADRIEN, Jean-Marie, « The missing link : modal synthesis », dans *Representation of the musical signal*, MIT press 1991

CADOZ et INCERTI, « Synthèse musicale par modèles physiques : modélisation de structures vibrantes avec le langage CORDIS-ANIMA », dans *Recherches et applications musicales en informatique musicale*, ed. Hermès

CADOZ, LUCIANI et FLORENS, « CORDIS-ANIMA : a modeling and simulation system for sound and image synthesis – the general formalism », CMJ 1993

COLLECTIF, « Actes du colloque « Modèle physique, création musicale et ordinateur » 1990 » ed. Maisons des Sciences de l'Homme, collection Recherche, musique et Danse, Paris, 1994

COOK, Perry, « A meta-wind-instrument physical model, and a meta-controller for real-time performance control » dans *Proceedings of the International Computer Music Conference*, ICMA 1992

COOK, Perry, « Physically informed sonic modeling (PhISM) : synthesis of percussive sounds », CMJ 1997

JAFFE et SMITH, « Extensions of the Karplus-Strong plucked-string algorithm », CMJ 1983

HUFSCHMITT, Aline, « La synthèse par modèles physiques », mémoire de maîtrise de musicologie, Université de Paris Sorbonne, 2000

KARJALAINEN, VÄLIMÄKI et TOLONEN : « Plucked-string model : from the Karplus-Strong algorithm to digital waveguides and beyond »

KARPLUS et STRONG, « Digital synthesis of plucked string and drum timbres », CMJ 1983

SZILAS et CADOZ, « Physical models that learn », ICMA 1993

Table des matières

Introduction	1
I. Les origines.....	2
II. L'approche de Pierre Ruiz.....	3
III. Cordis-Anima.....	4
III.1. Structure des objets vibrants.....	4
III.2. Structure des modèles dans Genesis	5
III.3. « Multisensorialité »	8
III.4. Que peut-on ajouter à Genesis ?	9
IV. Mosaïc, Modalys et la synthèse modale.....	11
V. La synthèse par guides d'ondes	12
VI. Comparaison entre les guides d'ondes et les autres modèles physiques	13
Conclusion : avantages et limites des modèles physiques	14
Sur la nature des sons créés.....	14
Sur la possibilité de contrôler le modèle.....	14
Bibliographie	15
Table des matières	16