

Année 2003 -2004

D.E.A. Art, Science, Technologies

Grenoble

Technologie et Processus de la création musicale

Alberto Novello

Introduction

On peut diviser chronologiquement l'ensemble de la production musicale humaine en trois grandes périodes : la préhistoire¹, l'époque primitive et moderne (concernant ce que l'on définira comme étant « *nouvelles technologies* »).

La proposition de ce traité est d'analyser attentivement tous les cotés de la relation entre l'homme et la technologie employée dans le processus de création sonore dans chaque époque de son histoire musicale ; on s'intéressera aussi bien au niveau poïétique qu'esthétique des phénomènes sonores. Chez l'être humain les cotés sensibles de production et de perception ne sont pas superposables ; cette différence donne lieu à d'intéressantes conséquences telle la nécessité de créer des instruments de musique pour communiquer en époque préhistorique.

En suite on examinera la relation entre notation - théorie musicale et production musicale et on étudiera comment cet outil puisse étendre le processus mnésique d'un phénomène sonore, comment pouvoir concentrer les caractéristiques d'un son dans la représentation la plus synthétique possible (stocker le maximum d'information dans la façon plus simple et efficace).

Mais la notation reste encore un tentative limitée de représentation d'un phénomène sonore : on se trouve encore à un niveau subjectif (interprétation de la partition). Une grande révolution technologique dans le monde sonore est la naissance des appareils d'enregistrement avec lesquels l'homme peut passer d'un niveau subjectif à objectif de mémorisation sonore. Cette conquête technologique correspond à la naissance d'une nouvelle forme d'art musicale : la musique concrète. La différence principale par rapport au passé est qu'à travers un enregistrement on perd la perception directe du son produit par le jeu instrumental : l'enregistrement analogique transmet un phénomène sonore que n'est plus ni joué, ni produit directement ; c'est un son decontextualisé, produit par aucune action directement identifiable.

Mais le son conserve encore sa matérialité : la bande magnétique qui le conserve est un support physique que on peut modifier, couper, traiter avec nos mains. Très différente est ce qu'apportera l'arrivée d'un nouveau système de production sonore : le numérique. Le son perd totalement son côté matériel devenant un signal, une suite de chiffres binaires. Toutes ses caractéristiques physiques que l'on pouvait contrôler, même indirectement avec la bande magnétique, nous échappent. On perd le contact physique parce que ce sont les électrons qui transportent l'information physique du son. L'information matérielle est cachée à un niveau électronique de microcircuit, un univers impalpable par la sensibilité humaine.

La production musicale se transforme en production des signaux avec toutes les problèmes annexes : le temps de calcul, la générativité des algorithmes. On retrouve alors le même problème que l'on avait rencontré à la naissance de la notation musicale. Le problème d'économie et efficacité, l'histoire humaine est une spirale qui se boucle sur soi même, mais au niveau plus haut. Pour conclure on réfléchira sur les principaux concepts de psycho-acoustique pour mettre en évidence les limites de l'approche signal dans la description de la complexité du système auditif humain et on montrera aussi qu'il est indispensable, pour la production musicale moderne, de reculer d'un pas logique dans la chaîne de création sonore. Conscients de ce qu'on a appris grâce à la musique synthétique, on se concentrera sur la reconstruction numérique des objets physiques, des instruments musicaux (reconstituer donc la « matérialité » que on avait perdu) pour recréer des sons artificiels intéressants et cohérents. On ne s'intéressera plus aux signaux mais aux objets qui les produisent.

¹ Concevant la naissance de l'histoire humaine avec l'écriture, on peut aussi nommer préhistoire la partie de l'histoire sonore humaine précédente à la notation musicale.

1 Conditions matérielles pour la création artistique

La musique comme tous les arts est liée à la technologie : son développement a commencé quand l'homme s'est libéré de sa posture « à quatre pattes » et a exploité la station debout. Avant ce changement physique substantiel, l'homme employait sa bouche pour la préhension alimentaire et ses mains non seulement pour la locomotion mais aussi pour la préhension, comme tous les autres animaux. La station debout a donné à l'espèce humaine l'opportunité d'utiliser ses mains et sa bouche pour des différentes fonctions : les mains peuvent être employées pour interagir avec les objets (naissance de l'art-isanat) et la bouche pour exercer sa fonction vocale (naissance du langage). On peut donc identifier avec cette étape fondamentale de l'histoire humaine la naissance de la plus primitive technologie. La station debout (le pouvoir d'utiliser les arts supérieurs et la bouche) a augmenté le spectre des possibilités humaines sans attendre une mutation biologique, par exemple la chasse était faite avec des outils plus efficaces. On peut donc voir la technologie comme une forme humaine d'accélération de son évolution.

1.1 Un media ultime: le phénomène sensible

Le définition d'art n'est pas simple en général : on peut appliquer différentes définitions selon les cas examinés et les phénomènes artistiques auxquels on s'intéresse.

On peut, pour l'instant, approcher le problème d'une façon élémentaire et très générale en affirmant que l'art est lié à la perception: pour que l'on puisse parler de « phénomène artistique », il faut qu'on puisse percevoir un phénomène sensible produit par quelqu'un. Il s'agit donc d'une relation entre un ou plusieurs émetteurs et un ou plusieurs récepteurs. La condition nécessaire pour qu'il se réalise une connexion entre les deux parties de ce simple système est la présence d'un média destiné à transporter l'information à travers le temps et l'espace. Ce moyen, en général appelé « canal », selon ses caractéristiques physiques peut modifier l'information pendant sa transmission.

Pour qualifier un objet d'« artistique » il faut que le signal porté soit un *phénomène sensible* et produit par une *intention du sujet* (on pourrait dire que le degré 0 de l'art est l'**intention** de concentrer l'attention d'un sujet sur un phénomène physique déjà existant).

Une fois cette première définition établie, il est intéressant d'examiner les possibilités que le corps humain a développé pour communiquer le phénomène artistique.

Le sujet émetteur a la possibilité de produire ou modifier le phénomène sensible.

Moyens Humains d'Expression	PHENOMENES	EFFECTEURS HUMAIN
La voix	Ondes acoustiques	Organe vocal
Geste et expression corporelle	*	Le corps, les mains

On voit clairement dans le tableau précédent que l'on peut pas émettre d'information exploitant tous nos sens ; on a pas encore développé, par exemple, l'émission d'odeurs ou la sécrétion des substances avec différents goûts pour envoyer des signaux comme certains insectes et animaux à travers l'odorat et le goût ; on n'a pas non plus la capacité d'émettre des signaux lumineux comme les lucioles.

Le sujet récepteur doit être capable de percevoir les phénomènes produits par l'émetteur.

Moyens Humains de CAPTATION	phenomenes	capteurs humains
Audition	Ondes acoustiques	Tympan/cochlée
Vision	Ondes lumineuses	Rétine
sens du «toucher » (ou mieux sens TPK) ²	Mécaniques/thermiques	Peau, muscles, etc.
Goût	Chimiques	Papilles
Odorat	Chimiques	*

En analysant les différentes voies de communication on peut déjà remarquer une première dissymétrie entre la communication visuelle et celle sonore: on peut recevoir des phénomènes visuels mais on ne peut pas en émettre : on a besoin d'une source extérieure et d'une relation entre nous et la source (par exemple le soleil). Par contre, on peut émettre et recevoir des ondes sonores. Il y a une deuxième différence : les phénomènes mécaniques (gestuels et corporels) ne se propagent pas comme des ondes sonores ou lumineuses dans l'espace. L'expression corporelle exploite le canal visuel, grâce à une source lumineuse, mais le geste ne se propage pas; ce n'est qu'à travers le contact physique que l'on transmet des gestes. On n'a besoin d'aucun autre sens.

En réfléchissant sur la relation entre émetteur et récepteur on peut comprendre d'importantes caractéristiques de la perception humaine. Il faut d'abord remarquer que tous processus perceptifs nécessitent une *action* à la base. Pour avoir une perception il faut un mouvement (par exemple on sait par expérience que si l'œil ne bouge pas en se concentrant sur un point fixe, après quelque second, on voit tout noir). De la même façon l'audition est active. Sans la vibration des récepteurs (poils) dans nos oreilles, la transmission nerveuse n'aurait pas lieu. Même les animaux tournent les oreilles pour capter mieux les sons. Si on examine encore la perception on peut s'apercevoir que chaque mouvement que on produit, nous fournit des informations : on voit, on entend, on peut avoir la sensation de la chaleur ou du froid, ou humide ou sec, rugueux ou lisse. On peut donc conclure en affirmant que toute action est plus ou moins perceptible et que toute perception est plus ou moins active.

Les études scientifiques modernes sur les appareils du sens nous révèlent une autre importante caractéristique de la perception humaine: toutes les perceptions sont multi-sensorielles. On peut se rendre compte de cette propriété sans être scientifique : en regardant une glace, on a aussi des « sensation » sur son goût ou sur sa température. Ça n'a donc pas de sens de parler de « sens » parce que notre perception est plutôt une représentation mentale totale, une expérience qui regroupe tous les sens.

A partir de la multi-sensorialité on peut s'apercevoir que les sens sont intrinsèquement liés. Les études inter sensorielles se concentrent sur cette caractéristique d'interaction entre nos sens.

Ces considérations analytiques donnent une vision plus complète et profonde des phénomènes perceptifs. En examinant les cas décrits (la boucle action-perception, la multi-sensorialité et l'interaction entre les différents sens) on peut définir tous les processus perceptifs humains comme des *boucles multi sensori-motrices*.

1.2 Production et perception du phénomène sensible.

La forme la plus simple d'art que l'on peut exprimer avec notre corps est l' « **art naturel** », c'est-à-dire l'art sans l'aide d'aucune technologie ; dans cette catégorie on peut regrouper :

² On parle ici de perception TPK, ou tactilo-proprio-kinesthésique, pour indiquer les informations que l'être humain peut capter par les cellules de la peau, des muscles et des jointures. On expliquera dans la suite la raison d'une telle appellation pour distinguer les complexes fonctions du sens TPK de celles du simple « toucher ».

- Les arts vocaux
- Les arts chorographiques
- Le mime
- L'art dramatique

Si on considère que le sujet et le percevant sont plongés dans un environnement physique que contient des objets, alors la conception d'art change. De plus, la perception du sujet récepteur dépend aussi de l'interaction du phénomène sensible produit par l'artiste avec l'environnement. On définit alors les « **objets sensibles** » comme tout ce qui est accessible d'un de nos sens.

Les objets peuvent être à l'origine des phénomènes sensibles sans interagir avec les actions d'un sujet (*spontanément*) (par exemple une lumière projetée dans l'air) ; mais peuvent aussi être des sources sensibles sous l'action humaine (*interaction*), par exemple la même lumière sur un plateau théâtral ; selon le sens exploité on peut distinguer entre une *interaction tactile, audible* (un marteau battu sur une table), du *visuelle*³ :

	organes	phenomenes
La voix ⁴	*	*
Expression corporelle	*	*
Le geste ⁵	Les mains, les membres, le corps	visibles
		audibles
		tactiles (TPK)

L'intervention d'objets sensibles complique les relations entre émetteur et perceuteur:

Interaction cote émetteur		Interaction cote récepteur	
Action gestuelle	Perception acoustique	présence	Perception acoustique
	Perception visuelle		Perception visuelle
	Perception tactile TPK		*

On remarque que il y a encore une dissymétrie profonde dans le coté émetteur, qui manque de la possibilité d'une perception tactile, par rapport au récepteur.

Lien entre la technologie et le processus artistique

L'introduction des objets sensibles modifie et étends les possibilités communicationnelles des moyens naturels humains de perception et de production; les objets ont en général une portée plus grande.

La portée des processus communicationnels peut être spatiale, nos informations peuvent alors arriver plus loin, et temporelle, on peut conserver un objet artistique (garder une partie du message à travers le temps). La portée plus grande des objets est la raison de la naissance de l'art non naturelle.

³ C'est évident que les interaction homme-objet ne peuvent naître qu'à travers les moyennes d'expression humaine que l'on a déjà traités : donc la voix, l'expression corporelle et les gestes. Dans ce sens on ne peut pas penser d'interagir avec un objet à travers la vision, parce que l'homme ne peut pas émettre des phénomènes visuels.

⁴ La voix peut en effet interagir avec les objets, mais le résultat ne change pas totalement le message transporté comme dans le cas gestuel ; on considère l'interaction matérielle vocale comme négligeable.

⁵ Pour la première fois on introduit le terme « *geste* » qui aura une grande importance dans le reste de notre disquisition. Pour l'instant on se contentera de définir comme geste toute action produite par nos mains.

L'instrument artistique

A ce stade de l'évolution humaine apparaît *l'instrument*: un objet physique qui augmente les possibilités de communication artistiques par la voie qu'il exploite.

Le même raisonnement peut être appliqué à la peinture et à la sculpture : l'instrument visuel permet d'utiliser nos gestes pour créer une chose importante visuellement (*objet artistique visuel*).

Dans le cas de la musique, l'importance artistique de l'instrument musical concerne surtout le champ auditif : l'émission des phénomènes sonores est aussi un *objet artistique sonore*^{6 7}.

Dans les arts visuelles l'objet artistique a une portée spatiale et temporelle énorme : la toile peut être transportée sur toute la planète, exposée et résiste bien à travers le temps.

La peinture et la sculpture sont deux arts avec une caractéristique très particulière : l'objet artistique est en même temps l'instrument du processus artistique : la toile et la pierre sculptée sont soit instrument de création, soit oeuvre d'art finale. Il y a ici une dissymétrie entre musique et peinture selon ce point d'observation :

	musique	peinture
L'instrument	existe	Instrument et objet sont la même chose
Ouvre d'art	Pas d'objet transmissible	

Organologie

C'est l'étude humaine des différents instruments de musique, pour les classer selon leurs composants, structure, matériel, production sonore, etc.

1.3 Modes d'expression artistique (« sans nouvelles technologies »)

On se propose ici d'analyser les différents types d'expression artistique humaine, avant l'introduction dans l'histoire moderne des « nouvelles technologies », et de les classer selon différents critères.

Se concentrant sur la nécessité d'un instrument pour l'expression artistique on peut diviser tous les arts humains en deux catégories : « non instrumentaux » et « instrumentaux ».

Les arts non instrumentaux

Dans le premier groupe on retrouve les arts vocaux, la danse, le mime, le théâtre. On peut ultérieurement les subdiviser analysant, pour chacun les moyens de production et de perception.

Par exemple dans le cas du mime, il y a une production totalement gestuelle et une perception seulement visuelle. La danse est une production par des moyens d'expressions corporelles, qui, comme le mime, n'est perçue que par la vision. Les arts vocaux sont basés sur l'exploitation de la voix et donc perçus par l'audition. Le théâtre est probablement le plus complexe des arts non instrumentaux parce que produit par des moyens vocaux, gestuels et corporels en même temps. Il est perçu par la vision et l'audition. Ci-dessous la table de classification des arts non instrumentaux:

Moyens de production	Moyens de perception
Voix	Vision
Expression corporelle	Audition
Geste	Perception TPK

⁶ Mais il y a aussi une indéniable communication gestuelle : les gestes produits par un instrumentiste, peuvent être considérés comme phénomènes objectifs ; on peut penser de les enregistrer en créant un *objet gestuel*.

⁷ A propos de l'objet sonore on traitera dans le chapitre relatif à la musique concrète et à la conception sonore de Pierre Schaeffer.

Les arts instrumentaux

Les arts instrumentaux sont la musique, la sculpture et l'animation. Pour tous les trois, le moyen de production est le geste appliqué sur l'instrument. Cependant la perception change : auditive pour la première et visuelle pour les autres deux (par fois, l'animation est accompagnée par un son, une musique, des paroles. Cela implique donc une perception auditive et visuelle au même temps).

Les arts du temps et arts de l'espace

On peut appliquer un autre critère de classification à l'ensemble des arts humains en divisant entre « arts du temps » et « arts de l'espace » selon la façon dans laquelle se développe l'oeuvre d'art et où est transmise l'information artistique :

	Arts du temps	Arts de l'espace
Non instrumentaux	Arts vocaux, mime	*
	Danse, art dramatique	
Instrumentaux	Musique	Arts picturaux Sculpture
	Animation	

Il est évident que la musique est un art du temps : on ne peut pas avoir une perception d'une oeuvre musicale en dehors du temps : on doit attendre son exécution, ou reproduction, pour pouvoir la percevoir complètement. C'est différent dans le cas de la peinture : la vision d'un tableau est totale immédiatement. L'oeuvre d'art ne se cache pas derrière le temps ; elle se montre en face de nous sans nécessiter une attente pour la « recevoir » par nos sens. L'animation, le théâtre et la danse se révèlent comme plus complexes sous ce point de vue. Ils sont immédiatement visibles, mais on doit attendre le développement temporel de l'oeuvre entière pour avoir une perception totale.

1.4 Le canal gestuel

Le geste est l'ensemble des comportements corporels associés à notre activité musculaire. Nos mains et nos doigts ne sont que les extrémités plus actives et les plus performantes d'un système complet et complexe : le corps dans son entier.

Le geste peut être alors *direct*, quand on entre en relation avec l'environnement physique ou humain qui nous entoure. Il peut être *indirect*, quand on interagit avec l'environnement par un objet intermédiaire.

Le canal gestuel lie donc le sujet avec son environnement. Il permet aussi bien des relations communicationnelles que des relations interactionnelles. Il nous permet d'interagir avec le monde physique extérieur et il permet de communiquer avec nos semblables. De plus, se concentrant sur sa fonctionnalité communicationnelle, on peut s'apercevoir que le canal gestuel permet d'émettre des informations mais aussi d'en recevoir. Cette émission-réception se réalise à travers le sens du «toucher».

Le canal gestuel possède trois différentes fonctions :

- *épistémique*, lorsqu'il permet la connaissance de ce qui nous entoure (la direction de l'information est de l'environnement vers le sujet : on reçoit),
- *sémiotique*, lorsqu'il permet d'émettre de l'information, pour communiquer, faire connaître (unidirectionnelle du sujet vers l'environnement : on envoie),

- *ergotique*, lorsqu'il permet de transformer physiquement le milieu environnant en dépensant de l'énergie puisée dans le corps humain, interaction avec les objets (contact bidirectionnel).⁸

En analysant la relation entre nos sens et l'environnement extérieur, on peut presque instinctivement affirmer que la vue, l'ouïe, l'odorat et le goût servent à l'homme comme fonctionnalités pour connaître et explorer le monde autour de lui ; ils ont donc une fonction épistémique. La voix par contre est utilisée pour communiquer et a donc une fonction sémiotique. Le toucher sert soit à explorer l'entourage, soit à interagir avec les objets. Le geste sert à communiquer et à interagir avec les objets ; geste et toucher ont une fonction ergotique.

L'activité artistique est en premier lieu une activité de communication : elle crée un contact entre deux individus, donc principalement épistémique et sémiotique à sa base, même si elle est souvent basée sur un certain nombre de phénomènes purement matériels et techniques. Ainsi, quand l'homme produit de l'art, il exploite aussi sa fonction ergotique.

La main organe de perception: la fonction épistémique du geste

Même si le sens du toucher regarde toute l'enveloppe corporelle, la main est l'organe privilège pour la fonctionnalité épistémique. On parle de *fovéa tactile* du corps humain⁹.

La fonction TPK (tactilo-proprio-kinestésique) se divise en:

1) **Informations tactilo-kinestésiques ou haptiques.** La main est un organe extrêmement élaboré. Quand on entre en contact tactile avec un objet par le moyen de la palpation, la main fournit à notre cerveau une image définie de l'objet et de l'information sur sa forme, orientation, distance et grandeur.

Dans certains cas, il est même possible avec le seul contact tactile de percevoir des propriétés 3D de l'objet touché. Par exemple, le seul contact avec une boîte fermée, peut me faire comprendre qu'elle est vide parce que son poids est trop faible.

2) **Fonction proprioceptive :** exerce un contrôle direct sur notre corps. Des récepteurs très sensibles envoient à notre cerveau les informations qui viennent de l'intérieur de notre corps et renseignent notre cerveau sur notre position, et sur la position de chaque segment mobile de notre squelette. Le système vestibulaire de l'oreille interne est un capteur de gravité fondamental pour repérer la position verticale ; il nous fournit d'importantes informations sur notre poids, nos trajectoires, nos mouvements.

Le sens TPK donne des informations essentielles sur l'environnement qui nous entoure et sur nous-mêmes. Il donne des informations sur la température, la plasticité, la pression, l'état de surface, la durée/la mollesse, la forme, l'orientation, la distance, la taille, le poids, les structures articulaires, les mouvements.

La main organe d'action: la fonction ergotique du geste

L'interaction gestuelle, à la différence de l'interaction visuelle et sonore, nous permet de déformer l'objet physique avec lequel on entre en contact. Avec nos mains, on opère une transformation physique et matérielle de l'environnement. Mais la main, avec ses 23 degrés de liberté, est la partie privilégiée de notre système moteur. Elle nous permet d'interagir avec les objets physiques mais

⁸ On parlera dans la suite de *relation* pour indiquer un échange entre l'homme et son environnement, *communication* entre deux êtres humains, et *interaction* entre deux ou plusieurs objets physiques.

⁹ La fovéa est une place du corps humain où la densité des récepteurs est très grande et on trouve, donc, une grande précision perceptive dans un espace limité. Le palme de la main avec son pouvoir séparateur de 2,3 mm, contre 42, par exemple, sur son dos, est la fovéa du corps humain pour le sens tactile.

n'est qu'une extrémité du système moteur complexe qu'est le corps humain. Le corps possède 792 muscles, 110 articulations, 100 degrés de liberté. Analysant les mouvements que notre corps exprime, une catégorisation fait apparaître des mouvements réflexes, inévitables, instinctifs et des mouvements contrôlés. Une autre subdivision peut distinguer les mouvements pour garder la posture et les mouvements pour l'exécution des gestes.

La main organe d'expression : la fonction sémiotique du geste

La main a aussi une fonction expressive : elle peut servir à produire les gestes dans la transmission des messages informationnels vers l'extérieur.

1.5 Typologie du geste

La division principale entre les différents types du geste dépend du fait d'utiliser ou non un intermédiaire matériel : il y a donc des gestes dits « équipés » dans le premier cas et des gestes dit « à nu » dans le deuxième.

Les gestes « à nu »

Représente la typologie la plus simple du geste, sa fonction est totalement sémiotique : faire connaître. Quelques exemples des gestes à nu sont les gestes qui accompagnent la parole, le langage des signes, la gestique du chef d'orchestre, le mime et les arts chorographiques.

Les gestes « équipés »

Les gestes équipés, au-delà de la fonction purement expressive, peuvent avoir aussi des fonctions ergotique et épistémique.

Le geste équipé peut être subdivisé entre *gestes ordinaires* (qu'on accomplit tous les jours) et *gestes experts* (qui nécessitent de notre attention), *gestes utiles* (appliqués sur un outil), ou *gestes instrumentaux* : l'ensemble des comportements gestuels qui établissent un contact physique avec un instrument. C'est à cette dernière catégorie qu'on s'intéressera particulièrement.

Le geste non instrumental

Un geste n'est pas instrumental quand l'énergie intrinsèquement consommée par le phénomène résultant n'est pas de source humaine. Il y a alors trois composants dans le processus de création:

- l'énergie humaine : toute action gestuelle entraîne une dissipation d'énergie, soit dans le corps seul, soit à la fois dans le corps et dans un objet physique qui le prolonge. Il y a donc toujours un circuit énergétique humain,
- l'énergie d'une source extérieure : dans le cas du geste non instrumental l'énergie n'est pas que humaine. Pour la production d'un objet final on a besoin d'un circuit énergétique non humain,
- l'intention du sujet humain : cela crée la dépendance entre les deux circuits énergétiques et garantit que le résultat est bien lié à une volonté humaine.

Dans ce cas on parle de systèmes à relais : il y a deux circuits énergétiques comprenant une source et une destination chacun : le circuit engendré par une source énergétique externe (lumière, électricité, énergie mécanique, etc.) et l'énergie humaine. Les deux circuits sont *énergétiquement* indépendants mais il y a un rapport de dépendance du deuxième sur le premier parce que le sujet humain peut modifier l'émission énergétique du premier circuit. On peut le contrôler et exploiter selon notre désir. La direction d'émission est orientée. Ainsi, le sujet peut avoir un contrôle sur la source, mais aussi l'inverse : la source à la base du phénomène ne peut pas changer modifier l'émission de l'autre source. La caractéristique plus importante est que si on examine la relation entre les énergies en jeu dans notre système on notera une discontinuité ; si on augmente l'énergie

du circuit humain l'effet final ne change pas nécessairement (par exemple, si on projet des images chinoise et que l'on veut placer un écran entre la source et la toile de projection l'effet final ne change pas si on utilise un grand panneau de carton ou un piano, mais notre effort sera différent).

Typologie du geste non instrumental

Gestes à nu manuels :

- le geste qui accompagne la parole. Il est intéressant de remarquer que, dans ce cas particulier, les gestes partagent avec la parole certaines étapes de traitement. En conséquence, ils font partie de la même structure communicative ou expressive,
- la langue des signes : dans ce cas le geste remplace la parole ;

Gestes à nu corporels :

- le mime
- la danse

Gestes à main équipée :

- la gestique du chef d'orchestre : le geste se base sur des éléments universellement acceptés mais aussi avec un ensemble de conventions entre le chef et l'orchestre.

Si on examine les exemples précédents, on peut se rendre compte que le geste non instrumental peut être que communicationnel. Dans sa vie quotidienne l'homme nécessite aussi d'interagir avec l'extérieur. On parle alors du geste non instrumental qui peut être communicationnel ou interactionnel.

Le geste instrumental

Le geste instrumental peut avoir les trois fonctionnalités décrites : sémiotique (pour communiquer un message, artistique ou non), ergotique (interaction neutre avec un instrument), ou épistémique (pour connaître, investiguer l'environnement). Dans certains cas, on peut appliquer plusieurs fonctions du geste instrumental simultanément : par exemple jouer (ergotique) pour exprimer notre tristesse (sémiotique).

Tout processus de perception consomme de l'énergie, aussi infime soit-elle. Toute transformation physique ou matérielle de notre environnement consomme également de l'énergie, mais d'un ordre de grandeur en général beaucoup plus élevée. On définira comme *instrumentale* toute activité qui fait appel à un dispositif matériel (l'instrument) extérieur au corps humain et telle que :

- sa finalité nécessite une consommation d'énergie en dehors des frontières du corps humain et de l'instrument,
- une partie au moins de cette énergie provient du corps humain.

Selon cette définition, on peut déduire une définition du *geste instrumentale* : c'est l'ensemble des comportements gestuels appliqués à l'instrument et dont une partie produira l'énergie nécessaire à la finalité de la tâche.

Le geste instrumental peut être divisé en geste avec *continuum énergétique* et avec *relais énergétique*.

Dans le premier cas, l'énergie que je produit, arrive proportionnellement à l'instrument physique employé (ex. un violon).

Dans le deuxième cas, il y a une discontinuité dans la fonction énergétique ; c'est le cas des systèmes à relais. Un circuit énergétique, dit primaire, est contrôlé par un autre circuit dit secondaire. L'énergie en sortie dépend de l'interaction des deux.

Geste instrumental musical

Pour parler du caractère « musical » d'un geste, il faut d'abord définir ce que on appelle musique. Dans la suite de ce texte, ce sera une activité artistique qui privilège le sens de l'ouïe, et qui prend sa place entre la main et l'oreille. On sait bien que la quantité des phénomènes sonores que l'on peut recevoir par l'oreille est beaucoup plus grande que celle que l'on peut produire avec notre voix. C'est pour produire la musique qu'il entend avec ses oreilles que l'homme a développé le geste musical.

Mais nos mains nues sont inaptes à produire une grande variété de sons. On a besoin d'un instrument musical pour réussir à faire vibrer les tympans et donc à stimuler l'ouïe. C'est la naissance du geste instrumental ; le geste que l'on applique à un instrument de musique.

Chaîne instrumentale musicale

L'oreille offre, par sa nature, la possibilité d'entendre des sons dont la fréquence est comprise entre 20 Hz et 20 KHz. Cela n'est pas la seule contrainte physique : pour l'audition on nécessite d'une source qui puisse mettre en vibration mécanique les tympans par le moyen d'une vibration de pression de l'air. Le processus d'audition lui même dépense donc de l'énergie mécanique.¹⁰ L'intensité sonore de la source doit être suffisamment élevée pour déterminer une vibration dans le tympan de l'ordre de l'atome. Par contre la nature macroscopique qui nous entoure est caractérisée par des grandeurs tout à fait différents : nos ne main bougent qu'avec une fréquence comprise entre 0 et 20 Hz, soit beaucoup plus petite de celle que on aurait besoin pour des sons perceptibles par notre oreille. De plus l'intensité d'un événement sonore décroît proportionnellement en $1/r^2$. Un calcul peut montrer qu'à seulement un mètre d'une source, le tympan récupère moins d'un millionième de l'énergie rayonnée par la source.

Ainsi, on doit introduire une discontinuité matérielle pour adapter la fréquence de vibration entre le mouvement des mains et les fréquences audibles par nos oreilles. L'instrument musical devra donc être constitué d'au moins deux parties physiques distinctes en relation « non linéaire ». De plus, on doit trouver une façon d'amplifier suffisamment les sons produits par le contact de nos mains sur l'instrument.

Une chaîne instrumentale générale est donc composée d'au moins trois parties : l'environnement aérien qui transfère le son, un excitateur qui produit la vibration et une structure vibrante pour la recevoir. Il est important de remarquer que cette composition n'est pas un choix humain, mais l'effet des lois physiques qui règlent notre univers. On retrouvera donc ces deux éléments dans chaque chaîne instrumentale musicale et en général dans chaque instrument de musique.

Le quatrième élément, qui s'ajoute presque systématiquement à cette chaîne, est une caisse de résonance qui sert à diriger le rayonnement de la musique vers une direction précise, en essayant de conserver le plus possible de son intensité initiale.

Le cinquième élément, qu'influence avec ses caractéristiques physiques la nature du phénomène sonore produit, est le milieu global. Il contient tous les autres.

L'instrument musical

L'instrument est la matière conditionnée pour assurer la transmission d'une certaine énergie des muscles aux tympans : il opère la transformation de l'énergie musculaire (bio - mécanique) en énergie vibratoire acoustique.

Le geste instrumental musical est alors le geste que l'on utilise pour entrer en contact matériel avec un instrument. Par intermittence ou par permanence, on s'en sert pour produire ou moduler l'énergie destinée aux tympans.

Les composantes du geste instrumental

Le geste instrumental musical peut être divisé en :

¹⁰ En général la communication a son économie : on informe pas sans dépenser de l'énergie.

- *Geste d'excitation*, il transporte l'énergie qui produit les vibrations mécaniques dans l'instrument et qui se traduisent comme ondes sonores dans l'air. Il existe quatre façons de communiquer de l'énergie à une structure vibrante :
 - Excitation instantanée par pincement* : le son commence quand le geste est fini. Les gestes sont instantanés et impliquent une mise en vibration instantanée. La phase contact main-corps vibrant est silencieuse,
 - Excitation instantanée par percussion* : le son commence quand le geste est fini. Les gestes sont instantanés et impliquent une mise en vibration instantanée. La phase de contact main-corps vibrant est violente et fugitive,
 - Excitation continue* : le son et le geste coexistent. Les gestes sont continus et impliquent une mise en vibration continue,
 - Excitation entretenue* : les gestes sont continus et impliquent une mise en vibration instantanée;

- *Geste de modification* : il est appliqué sur la chaîne instrumentale pour changer les caractéristiques physiques de l'instrument. Il a comme finalité non seulement le changement du son émis mais aussi la relation du geste d'excitation au son. Il y a trois catégories du geste de modification :
 - Modification structurelle* : ajout ou suppression d'un élément,
 - Modification paramétriques* : portant sur un paramètre,
 - Modulation de modulation* : continues ou discrètes ;

- *Geste de sélection* : c'est l'ensemble des comportements gestuels qui permettent de sélectionner la ou les structures vibrantes qui seront excitées (dans le cas d'une chaîne multi-instrumentale : un instrument qui comporte plusieurs structures vibrantes) ;

Dans le geste d'excitation l'artiste dépense de l'énergie qui est transformée en énergie vibratoire mécanique et qu'on entend sous forme d'intensité sonore ; le geste a une fonction expressive immédiate.

Dans le geste de modification l'énergie musculaire de l'instrumentiste n'est pas transformée en énergie acoustique, elle ne change pas l'intensité finale du son émis, mais les propriétés de l'instrument ; il est intéressant de remarquer que l'énergie employée dans le geste d'excitation a une relation physique avec l'intensité produite à la fin de la chaîne instrumentale ; dans le cas du geste de modification il n'existe pas de lien particulier entre cette dépense d'énergie et le résultat : c'est un système à relais .

Une autre différence consiste dans la mineure complexité à adapter l'application du geste de modification sur l'instrument. Dans le geste d'excitation le nombre des contraintes physiques et mécaniques de la relation limite fortement les choix technologiques pour la construction de l'instrument et pour la façon de le jouer ; dans le cas du geste de modification la choix est beaucoup plus simple, le nombre des contraintes est plus petit.

Dans le cas du geste de sélection, l'énergie est dépense dans le mouvement des membres de l'artiste pour atteindre les parties choisies de l'instrument. L'énergie n'est pas transformée en énergie acoustique comme dans le cas de l'excitation et n'est pas transférée à l'instrument comme dans le cas de la modification. A la différence des deux autres pratiques gestuelles, le geste de sélection ne pose pas de problème particulier, dans la conception de l'instrument, les contraintes dérivent de la facilité de le jouer mais sont complètement arbitraires. Remarquons que la sélection est, par essence discrète, mais les autres deux typologies gestuelles sont en général continues.

Réflexions à posteriori

Après cette analyse détaillée des sens et gestes humains on peut tirer des conclusions sur les gestes instrumentaux (s'adressant surtout au champ musical, mais aussi valable pour les autres arts).

Le geste instrumental communicationnel est par essence ergotique, sémiotique et épistémique tout à la fois. L'aspect sémiotique est implicite dans la définition de l'art que nous avons donné tout au début : le geste musical sert à exprimer quelque chose. Il est aussi ergotique parce que il nécessite l'interaction humaine avec un instrument donc un objet matériel extérieur : l'instrumentiste est lui-même la source d'une l'énergie qui se transmet jusqu'aux sens du destinataire, c'est beaucoup plus qu'une simple transmission des données ou de codes, cette continuité énergétique entre instrumentiste et sujet récepteur met en jeu une infinité d'éléments qui dépassent les représentations conscientes et les formalisations explicites que l'on peut en donner.

Pour attribuer la caractéristique de instrumentalité à un geste, il est nécessaire qu'il soit ergotique et que la chaîne instrumentale ait une continuité énergétique.

La relation instrumentale est généralement sensori - motrice et multi – sensorielle. Comme nous l'avons déjà expliqué, il y a une forte liaison entre action et perception (une boucle implicite) : toute action est perceptive, toute perception est basée sur une action ; il y a multisensorialité parce que nos effecteurs et récepteurs sensibles sont sollicités en association, il est artificiel de les séparer.

2 Le processus de création

2.1 Le jeu instrumental

Dans ce chapitre, on essaiera de décrire le développement de la relation homme-instrument de musique dans le cours de l'histoire. Cela sera un chemin conduit suivant une approche logique dans le temps, plus que strictement chronologique, même si, bien sûr, les deux sont intimement liés.

Dans le dernier chapitre, on avait introduit l'environnement et les objets sensibles avec lesquels l'homme entre en contact par le moyen des gestes. On s'était concentré sur le geste instrumental, fondamental pour les arts musicaux.

Quand le sujet entre pour la première fois en interaction avec l'instrument, il ne sait pas comment l'utiliser. Le sujet commence alors par une interaction directe, expérimentale qui peut sembler être basé sur le hasard (gestes désordonnés qui produisent des effets inattendus). En réalité, le sujet est mené à interagir avec l'instrument employant seulement des actions appartenants à un champ très limité de sa motricité. La cause de cette limite se trouve soit dans les contraintes physiques de l'instrument soit dans les propriétés et contraintes physiques du corps humain.

Ces deux causes générales limitent notre capacité exploratrice.

De plus, dans la relation homme-instrument s'ajoute une cause personnelle qui dépend du sujet. Son patrimoine cognitif (schémas mentaux, moteurs, instrumentaux qui dérivent de son passé particulier) restreint encore le champ d'action. On sait aujourd'hui que la mémoire humaine combine des « programmes moteurs » de perception TPK avec les éventuelles perceptions auditives et visuelles associées à ces programmes. Nous sommes donc influencés dans nos actions par les schémas cognitifs qui appartiennent à nos expériences passées, à des parcours logiques déjà essayés et que l'on essaie de re-proposer.

La première étape de l'interaction homme-instrument est une exploration des actions possibles guidées par l'observation des phénomènes sensibles produits. Voici une analyse des relations cause - effet de l'instrument : un particulier phénomène sonore est conséquence directe d'une action précise du sujet, choisie parmi les actions possibles sur l'instrument étudié.

L'effet d'une nouvelle interaction avec l'instrument produit des nouveaux systèmes sonores. Dans la tête du sujet sont stockées les informations qui donneront forme aux représentations mentales et qui sont relatives à deux causalités :

- physique structurelle (liée à l'instrument) : le sujet achève des connaissances à propos des propriétés de l'objet (cette partie ne dépend pas du temps mais est inscrite directement dans la composition physique de l'instrument),

- humaine et temporelle (liée aux gestes effectués) : le sujet comprend quels sont les effets liés à ses gestes (cet partie dépend fortement du temps, ça peut se développer ou se perdre).

A ce point de l'interaction homme-instrument, on peut déjà commencer à parler de *processus de création musicale*. C'est le « niveau 1 » de la production d'un phénomène artistique: le sujet explorant l'instrument a réussi à produire des nouveaux phénomènes sonores.

Un effet intéressant du contact avec un nouvel instrument est la découverte d'un espace d'effets éventuellement structurés (des réflexions, des reconnaissances avec des anciens phénomènes connus dans le passé). Le sujet réalise inconsciemment des associations entre phénomènes sonores produits, leurs différentes propriétés, actes gestuels employés et perceptions TPK reçues. Avec ces associations *active-cognitives* le sujet acquiert de nouvelles représentations cognitives (sensorimotrices/multi sensorielles).

L'exploration de l'instrument commence à devenir ici un processus plus profond et pas directement contrôlable par le sujet. L'homme est guidé dans l'exploration de l'instrument par sa perception et sa cognition. Les effets de ce travail intellectuel consistent d'abord en une assimilation de l'instrument comme organe de production de phénomènes acoustiques. Il s'installe une forte symbiose entre l'instrument et le sujet. L'instrument est devenu une partie de l'instrumentiste qui à ce stade déplace son attention de l'interaction avec l'instrument vers le message émis. Les phénomènes corrélés aux différents gestes de l'instrumentiste ont été tellement intériorisés que l'intérêt se dirige maintenant vers le message émis plutôt que sur la technique instrumentale. Le sujet fait un seul corps avec son instrument. Ceci est la condition de base pour une bonne expressivité gestuelle et une haute efficacité de la chaîne communicationnelle¹¹.

L'instrumentiste a atteint à ce point, un niveau profond d'exploration avec l'instrument. Il arrive désormais à imaginer les effets sonores produits par une action précise sur l'instrument car il a dans une expérience suffisante pour généraliser le comportement de l'objet en de nombreuses situations. Le sujet peut imaginer des phénomènes qu'il n'a jamais expérimenté, telle une réaction à une action jamais essayée (travail cognitif profond, c'est le « degré 2 » de la création musicale : *la création mentale*).

2.2 Mémoire « subjective » et « transmission orale »

L'introduction dans le jeu musical d'un instrument a une importance très grande pour amplifier la portée spatiale et temporelle du phénomène sonore produit. Cependant la portée sonore a aussi une dimension sociale et historique à laquelle on veut s'intéresser. L'instrument peut alors être employé dans la communication sociale, et la pratique instrumentale peut être transmise entre les individus d'une société (portée spatiale) ou de maître à élève, de génération en génération (portée temporelle).

Un requis fondamental pour la transmission à un élève des connaissances sur un instrument est la permanence de l'expérience instrumentale du maître. Ceci est constituée par la permanence des données atemporelles (la mémoire objective des caractéristiques structurales de l'instrument), et par la permanence des données temporelles, dits aussi « événementielles ». Il faut donc que se soit possible de reproduire l'événement temporel exactement pareil.

La condition nécessaire pour conserver ces deux précieux éléments, et donc pour pouvoir reconstruire une expérience instrumentale, est la mémoire subjective du maître : la connaissance détaillée des gestes que il a pratiqués, les effets corrélés, la relation entre ses actions et les effets produits ; seulement une telle connaissance peut permettre de reproduire le même effet et la même expérience.

Un problème intrinsèque liée à la transmission subjective est que la mémoire subjective est sujette à pertes, modifications et en plus n'elle est pas transmissible parfaitement ; au mieux on peut arriver à

¹¹ C'est intéressant de remarquer ici comme pour "intérioriser" les parties instrumentales de notre corps on fait le processus contraire par rapport au cas instrumental : pour apprendre les arts vocaux par exemple on essaie de faire abstraction de notre gorge pensant la comme un instrument externe, pour après le explorer selon les façons décrites avant et après l'assimiler.

donner des approximation des nos représentations mentales de l'instrument mais : on arrivera jamais a transmettre entièrement ce que on a acquis pendant nos expériences. Ils se posent donc des obstacles à notre envie de donner une dimension sociale à notre expérience personnelle.

Une méthode que on peut appliquer pour transmettre l'expérience est le *mimétisme* : le maître reproduit les gestes instrumentaux plusieurs fois produisant des phénomènes sonores, ses disciples, qui ont observé les gestes et les action du maître, essayent en suite de l'imiter en cherchant de recréer les mêmes phénomènes sonores utilisant les mêmes gestes ; les disciples devraient, dans leur cerveau, recréer la même mémoire subjective de l'instrument que possède le maître ; l'expérience du maître dans ce cas est acquise par l'extérieur.

Un programme de confrontation direct avec le maître même permet de décider si le disciple produit le même phénomène sonore que essaye d'imiter.

Dans cet pratique aussi se posent des problèmes : les élèves se concentrent sur les effets mais pas sur les causes : même si le disciple imite exactement les mêmes gestes et produit les mêmes phénomènes sonores du maître, on peut jamais être surs que il ait acquise la même expérience et la même représentation mentale de l'instrument. On sait bien que dans notre univers deux effets pareils n'ont pas nécessairement la même cause, mais aussi que deux causes identiques mènent au même effet.

La *transmission orale* est une autre façon de transmettre la connaissance à travers le temps et l'espace ; avec cet approche, la pratique musicale est transmise par le moyen du langage entre les individus d'une société ; le résultat est un échange de connaissance dynamique, la naissance d'une représentation mentale collective et donc une culture musicale propre de cette société.

La musique de tradition orale est encore vivante aujourd'hui (par exemple la musique balinaise possède ce nature) ; le résultat d'une telle conception sonore collective est un processus de création musicale sociale ou il y a pas des individus qui dirigent mais un « sens » musical diffusé, qui dérive de la contribution de chaque instrumentiste.

2.3 Notation(s) et théorie(s) de la musique

L'écriture en musique, ou en général chaque forme de notation musicale, est la manifestation d'un désir humain d'étendre la mémoire subjective.

La représentation

On peut distinguer deux différents types de représentation de la réalité objective externe:

- La *représentation mentale* est l'ensemble d'idées que se forme dans notre cerveau quand on interagit avec un objet ; cet ensemble d'idées, concepts, expériences que on a de l'objet est en suite automatiquement associé au nom de l'objet en question.

La représentation mentale est notre description sensorielle de l'objet ; en général on peut pas être sur que ça corresponde à l'objet physique réel : c'est l'interprétation gnoséologique que notre cerveau a élaborée de la nature ontologique de l'objet, a partir de l'expérience vécue.

- La *représentation matérielle* est utilisée pour remplacer une chose inatteignable par un objet tangible, déjà présent en nature ou produit par nos mains ; on espère, avec cet approche, de pouvoir récupérer des informations sur la chose absente par la directe interaction avec l'objet remplaçant.

C'est important de remarquer que dans tous les deux cas, représentation mentale et représentation matérielle, c'est impossible d'atteindre une représentation absolue en parfait accord avec la chose représentée, il y aura toujours un écart entre notre modèle, mental ou matériel, et la réalité décrite. On peut alors constater une identité de comportement dans un champ limité mais jamais d'égalité absolue.

La substitution d'un objet par sa représentation comporte donc une *réduction* : les propriétés de notre représentation forment une petite partie de la totalité de celles que l'objet réel contient en soi ; mais la création d'une représentation provoque aussi une extension : la représentation que prend la place de la chose originale aura des propriétés en plus, différentes par rapport à l'objet substitué. Sont ces propriétés « en plus » que déterminent un stimulus artistique différent ; la création par l'homme d'une représentation des choses détermine un premier pas, volontaire ou non, vers la constitution d'un langage artistique, vers une interprétation, une description du monde extérieur ; le modèle construit (mental ou physique), limite le champ d'action et adresse les choix suivantes dans le langage artistique.

- La *représentation analogique / symbolique* ; on subdivise le phénomène de représentation entre signifiant (l'objet physique qui vient représenté), le signifié (la représentation mentale que on a de l'objet), et le référent (l'ensemble de tous les objets avec même caractéristiques).

Evolution de la musique occidentale

Analysant l'histoire de la musique occidentale, on peut s'apercevoir que son développement a été très fortement influencé par la naissance de la notation et de l'écriture musicale. Si on tient compte de cette remarque dans l'analyse du présent, on peut constater que le récent changement de la notation musicale, introduit par la révolution numérique, et les nouvelles formes d'écriture, peut conduire à une nouvelle révolution sonore et à une conception différente de la musique.

L'introduction de l'écriture dans l'histoire de la musique humaine a permis la transmission détaillée des pratiques gestuelles nécessaires dans le jeu instrumental, pour décrire les différents événements sonores produits, ou encore pour décrire la relation entre les deux ; de ce point de vue, l'écriture est née pour étendre la portée de la mémoire « subjective ».

La condition nécessaire et suffisante pour reproduire un événement sonore est la présence des documents musicaux relatifs, la conservation temporelle de l'instrument et du corps de l'instrumentiste qui l'a joué.

La diffusion de la pratique d'écriture dans une société, sa compréhension et l'acceptation des conventions par plusieurs individus aboutissent à un langage sonore social. La mémoire n'est plus alors « subjective » mais on peut déjà parler d'une « mémoire sociale » : les individus de la société ont des représentations très semblables du processus de création musical. Par l'écriture l'humanité essaie d'élaborer des moyens de mémoire objective, d'extériorisation de la mémoire, pour constituer une mémoire sociale, ou la technologie et la société sont indissociables¹².

Histoire

Les premières formes d'écriture musicale dans l'histoire humaine sont très anciennes : les chinois furent les premiers à avoir introduit un système d'écriture musicale. Les grecs connaissaient des formes pour transmettre la pratique musicale dès le IV^e siècle avant JC, même si c'était plutôt des formes phénoménologiques d'écriture (l'intérêt se concentrait surtout sur le placement des mains sur l'instrument, plutôt que sur la représentation de la musique).

Ensuite, les formes d'écriture ont été oubliées jusqu'au Moyen Âge (IX^{ème}, X^{ème} siècles), quand l'écriture musicale a commencé à être nouvellement utilisée pour pouvoir classer la musique et les chants religieux, pour reproduire les textes sacrés. L'écriture musicale est très proche de l'écriture du langage : tous deux sont une traduction sur papier des sons que l'on produit, ce sont les mêmes processus à la base.

¹² On renvoie au prochain paragraphe l'explication de la relation entre technologie et écriture

La boucle notation – théorie - représentation

L'écriture est une forme de franchissement du temps et de l'espace, c'est un moyen durable et transmissible, un système objectif et clair avec des règles explicites. Dans ce point de vue, la musique peut être comparée avec les mathématiques. Ecrire c'est suivre des règles, donc rationaliser le son : étymologiquement, étudier les rapports (*ratios*) logiques à travers les nombres. Mais rationaliser a comme effet une réduction de l'expressivité humaine : on essaie de décrire les phénomènes interprétables par des nombres on opère une discrétisation du réel, on rend plus facile la notation mais on introduit une perte dans les phénomènes représentables.

Une telle description discrète détermine la naissance d'une théorie : un moyen très fort pour économiser les moyens de notation, être pragmatique et réduire l'effort de la mémoire. De la théorie à la notation ; « noter » signifie donner une représentation sur un support permanent, augmenter la portée de l'information.

Comme toutes les représentations, l'écriture limite, mais au même temps introduit un nouveau modèle d'interprétation des phénomènes sensibles et du monde : c'est une description des propriétés du son ; l'écriture introduit intrinsèquement une vision artistique différente et un nouveau système conceptuel que soutient autant l'écriture, que la création artistique : la théorie.

On remarque que rationalisation et théorie sont indissociables par le moyen des règles. Notation et théorie sont en contact, évoluent l'une et l'autre en dépendance. Chaque système d'écriture, donc de rationalisation, possède des règles, donc une théorie ; chaque théorie est déterminée par un système conceptuel à la base, et produit des règles pour décrire les phénomènes sensibles ; il est donc automatique de dériver une description sous forme d'écriture à partir d'une théorie.

Un aspect important à souligner ici, c'est la liaison entre système formel de description et création artistique. Dans l'histoire humaine on peut observer différents exemples à regard : d'un simple signe sur papier, il est souvent possible traire inspiration et déduire des concepts importants et profonds (Newton et les intégrales, Dirac et ses diagrammes). Les nouvelles théories introduisent de nouveaux signes, mais aussi de nouvelles idées.

La technologie dans la notation

Dans ce paragraphe, on s'intéressera à la relation entre notation et technologie, en particulier on essaiera de répondre à la suivante question : où se trouve l'aspect technologique dans le processus d'écriture musicale ?

La notation nécessite, pour sa production, de la technologie ; en examinant attentivement le processus d'écriture on peut s'apercevoir que l'aspect technologique est présent dans :

- le support, qui sert à la conservation de l'information ; sa propriété principale est la durée, chaque support a une durée finie ; selon l'utilisation que l'on veut en faire, on doit choisir le plus approprié ;
- le marqueur : l'outil, l'instrument qui permet d'écrire, graver l'information dans le support ;
- la trace, plus ou moins durable selon le cas, c'est-à-dire plus ou moins d'information ;
- le rapport entre le support et le marqueur, représente le nombre des signes que l'on peut produire ou apercevoir par une technologie précise donnée ;
- les outils complémentaires à l'écriture, par exemple la règle, le compas, etc.
- l'imprimerie, pour produire un grand nombre d'exemplaires ;

Technologie et création musicale

Quel rôle joue la technologie de la notation dans la création musicale ? Quel rôle spécifique jouent les contraintes technologiques dans la composition artistique ? Quelles sont les limites expressives introduites ?

Il y a d'abord un lien très fort entre *la graphie* et la technologie du marquage :

- du point de vue *qualitatif*, une esthétique particulière peut conduire vers des formes différentes de représentation de la musique, et donc également de sa création;
- du point de vue *quantitatif*, la grande variété des formes différentes peut stimuler la créativité et la composition peut changer selon le caractère des formes en jeu (liaison images ↔ idées produites);
- une problématique instrumentale fondamentale est liée à l'aspect ergotique du geste graphique ; cela dépend de :

- la géométrie du support : bidimensionnelle, tridimensionnelle
- la staticité : la trace reste dans le temps. Il n'y a aucun moyen de représenter une évolution temporelle. En musique, l'écriture de gauche à droite introduit une notation monodimensionnelle pour représenter le temps. Donc, même avec un support apparemment 2D, il est possible de représenter des logiques causales comme l'évolution temporelle d'une partition musicale.

Grâce à l'utilisation d'un support, on acquiert la possibilité d'intervenir sur l'aspect temporel de la musique : on peut choisir directement la structuration, la construction de nos phrases musicales, changer leur ordre, les ranger ; quand on la joue, en absence du support, on ne peut pas revenir pour changer l'ordre des notes jouées

Résumons : grâce à la notation et à la théorie (donc l'écriture dans son ensemble), on a acquis un nouvel outil pour atteindre différentes représentations abstraites des phénomènes sensibles.

2.4 Composition / Interprétation

Dans les paragraphes précédents, nous avons introduits des concepts importants comme le sujet, l'objet (instrument musical) et l'écriture. Les relations entre eux déterminent l'acte de création artistique. Toutes les relations à l'intérieur de ce système sont en boucle : il y a échange réciproque dynamique entre les composantes. La boucle est à la base du processus artistique.

Les relations bouclées entre les composantes du processus de création.

Le sujet avec ces gestes manipule l'objet.

L'objet entre en contact avec le sujet par sa représentation mentale objective.

Le sujet utilise pour sa composition l'écriture, mais l'écriture elle-même influence le sujet à travers la représentation mentale (discrétisation des sons naturels sous forme numérique) que le sujet a de l'écriture. L'écriture est un phénomène matériel (technologique), c'est déjà un modèle, l'introduction est production d'un système de création.

Il reste à analyser les relations directes entre l'objet instrumental et la notation. Dans la conception musicale moderne, on sait bien qu'un ordinateur peut interagir avec un instrument selon une notation interne, sans nécessité de passer à travers un sujet humain : dans cet exemple la notation entre en relation directe avec l'objet.

Mais l'instrument a aussi une partie active vers la notation : par exemple, dans le cas des tablatures, la notation musicale s'intéresse aux relations entre sujet et instrument, plutôt qu'aux sons produits, la notation indique au sujet où placer les mains, comment interagir correctement avec l'instrument pour produire une mélodie. Dans cet exemple, on voit comment l'instrument peut influencer activement l'écriture musicale.

Le processus de création mentale abstrait

La création musicale d'un sujet peut être définie comme « mentale », parce qu'il a une représentation mentale de l'écriture et des objet extérieurs, et dans le cas spécifique musical de la notation et de l'instrument que le sujet veut utiliser. Par le moyen de sa représentation mentale, le sujet est capable d'imaginer comment interagir avec l'instrument pour produire les sons voulus, et

comment il pourrait noter sur un support les gestes instrumentaux qu'il vient de créer dans sa tête; à ce moment, la composition n'est plus seulement mentale mais aussi « abstraite », parce que elle est retirée de la réalité motrice, le sujet a réussi à composer sans produire aucune note audible.

Interprétation

L'interprétation est définie comme le processus de création en soi ; c'est tout de même un processus de création musicale parce que l'instrumentiste ajoute des couleurs, change le tempo, l'intensité sonore, modifie l'oeuvre d'art selon des directions que le compositeur n'aurait jamais pu imaginer ; l'introduction du facteur humain dans la production artistique ne peut pas être négligée.

Qu'est-ce qu'une oeuvre musicale ?

La question présente différents points de vue :

- cela pourrait être la partition musicale ; dans ce cas, l'instrumentiste est seulement un exécuteur, un « lecteur » de partition, le compositeur est le vrai artiste.
- ou cela pourrait être la reproduction musicale : dans certains cas, l'interprétation exprime un travail intérieur et interprétatif tellement important du sujet, que le résultat peut sembler être une nouvelle composition, que étend de manière inédite, le travail originel du compositeur.
- ou cela pourrait être l'instrumentiste qui exécute la vraie oeuvre d'art ; dans l'exécution, toutes les différentes parties du système musical sont mélangées : la composition, l'interprétation et la technique de l'instrumentiste, sans lesquelles les sons ne pourraient être produit.

3 Musique et nouvelles technologies

Avec ce chapitre, nous avons l'intention de répondre à une question fondamentale : « Qu'y a-t-il de nouveau dans les nouvelles technologies ? »

La technologie, comme nous l'avons déjà précisé au début de ce traité, n'est pas une nouveauté pour l'homme, n'est pas un produit de ce siècle dernier, ni de l'avant dernier : on peut définir le début de la technologie avec l'affranchissement de l'homme de sa posture à « quatre pattes ». Parfois on serait tenté de dire que l'introduction des ordinateurs constitue une nouveauté essentielle dans le processus d'évolution de l'humanité, mais en analysant leurs composantes, on peut comprendre que les principes qui interviennent à la base des ordinateurs ne sont pas des idées nouvelles : l'amplification, l'automatisme, les transducteurs, la notion de programme, la mécanisation du calcul, la numérisation. Examinons-les :

Amplification

Normalement, on se réfère à l'amplification électronique, introduite pour la première fois en 1906 avec la construction de la lampe triode¹³. Mais si on examine le principe qui détermine l'amplification, on s'aperçoit qu'il y a deux circuits énergétiques en relation : un dit primaire et un autre secondaire, comme nous l'avons déjà vu dans la partie précédente pour les systèmes à relais. L'amplification peut être linéaire ou peut introduire une saturation, si le gain de l'amplificateur est trop haut ; la sortie dans ce cas produira des 0 ou des 1, ON ou OFF, un petit changement du signal analogique en entrée détermine une différence dans le signal discrète en sortie.

Mais ce concept n'est pas nouveau et n'est pas une découverte de l'électronique : déjà, en 300 avant J.C., l'orgue traditionnel se basait sur ce principe : le circuit primaire était basé sur la pression de l'aire et le deuxième circuit que déterminait le son était la pression du doigt de l'instrumentiste sur une touche.

Automatisme

Le principe qui est à la base de l'automatisme est le lien entre sortie et entrée, souvent appelé *feedback* : ce que l'on a à la sortie règle l'entrée de notre système par un mécanisme qui introduit un

¹³ Il est intéressant de remarquer que le premier nom de la lampe triode était « *audion* » ; elle était en fait d'abord conçue pour des utilisations audio.

asservissement des deux extrémités de notre système. Déjà Héron avec sa fontaine à vin avait exploité ce concept.

Les traducteurs

Ce sont des interfaces physiques : ils permettent de transformer un signal physique en entrée en un signal d'une autre nature en sortie. On pense souvent à des traducteurs électroniques, mais par exemple le baromètre domestique opère une transformation analogique entre la pression et un déplacement mécanique de l'aiguille sur une échelle angulaire étalonnée pour nous « montrer » la pression.

La numérique et le calcul

Cette une conception dérive du calcul, étymologiquement la pierre, le cailloux utilisés pour compter dans la préhistoire, et du digital, mot que dérive étymologiquement des doigts ; aussi pour compter. Pascal fut le premier à avoir introduit le calcul mécanique mais un autre exemple célèbre est aussi la machine arithmétique de Babbage (1840).

La notion de programme

Chaque logiciel est essentiellement un ensemble de commandes écrites sur un support. Une machine que lit le programme écrit, exécute les séquences d'actions prédéterminées, selon les caractéristiques, les paramètres et l'énergie stockée à l'intérieur du système. Le produit de ce processus d'interaction de la machine avec la lecture du support donne un résultat.

C'est le cas du théâtre roulant de Héron d' Alexandrie que au II^{ème} siècle avant J.C., exploitant la chute d'une pierre (stockage d'énergie) pour déclencher le mouvement de personnages et en même temps déplacer le théâtre selon des séquences écrites.

Cette rétrospective historique montre clairement qu'il n'y a rien de nouveau d'un point de vue conceptuel. La nouveauté réside simplement dans l'intégration de tous les principes au sein d'un objet unique, et à une échelle beaucoup plus petite.

	Le siècle passé	Aujourd'hui
Echelle spatiale	A la place d'un engrenage (échelle 1 m)	Les électrons (10^{-15})
Echelle temporelle (transmission des données)	Vitesse du son	Vitesse de la lumière
Echelle Energétique	Energie mécanique	Electrons électrique

Ce changement d'échelle comporte un changement radicale de notre conception de la réalité, tout devient signal, représentation des phénomènes sensibles, il se crée une discontinuité dans la substance, une coupure ontologique ; on a à faire à un univers des signaux. C'est à dire un monde à une échelle tellement différente, que nos organes ne peuvent pas avoir un accès direct aux phénomènes sensibles. La communication, la transformation de l'information par le moyen électrique nous oblige à communiquer entre nous via des signaux, à traiter des signaux, en un mot des signes. Il se pose ici une nouvelle question à laquelle on répondra dans la suite : l'ordinateur est il un nouveau instrument ?

Eléments historiques

On essayera en ce paragraphe de parcourir les principales étapes du développement technologique moderne, examinant le fondamentales découvertes scientifiques, pour poser les bases nécessaires à la compréhension de la composition musicale moderne. En substance comment est-il possible de commander à un ordinateur de calculer une composition sonore, suivant des contraintes précises.

Le premier système de production de son électronique fut inventé, en 1898, il s'appelle le Thelarmonium. Le principe de fonctionnement ressemble à celui d'une dynamo : un rotor produit un signal, courante ou tension, qui évolue sinusoïdalement dans le temps. Par des boutons c'était possible de sélectionner les harmoniques en fréquence et de faire varier leur intensité. Un réseau téléphonique était employé pour véhiculer les signaux vers une sortie sans nécessité pratique d'amplification.

C'était, on peut dire, une "centrale électrique à son" : elle pouvait produire toutes les fréquences désirées, pas seulement 50 Hz. Le Thelarmonium était polyphonique.

Dans l'année 1921 le russe Theremin, fabrique un instrument électronique monophonique constitué d'un seul oscillateur contrôlable par deux antennes : une pour la fréquence et l'autre pour l'amplification. L'interaction gestuelle de l'instrumentiste, à travers la variation capacitive avec l'antenne, détermine l'hauteur et le volume du son produit.

Les Ondes Martenot est un instrument électro-mécanique monophonique : le système est basé sur un oscillateur commandé ou par un clavier, ou par un ruban. Le contact du doigt de l'instrumentiste avec le ruban détermine la fréquence, comme si c'était un « clavier continu ». La diffusion du son est réalisée par une caisse de résonance contenant des cordes électro-aimantées que vibrent selon les notes choisies par le jeu de l'instrumentiste.

ENIAC est le nom du premier ordinateur créé par l'homme : 18000 tubes à vide, 6000 commutateurs (cette capacité de calcul est comparable avec celles des processeurs actuels de nos PC).

L'année 1948 signe une étape très importante pour la musique moderne : l'invention du transistor influence le développement de l'amplification et des techniques d'enregistrement, et détermine la naissance de la musique concrète.

Le progrès avance vite et après seulement quatre ans, on passe de la musique qui exploite des sons déjà présents en nature à des sons synthétiques, la recherche des sons artificiels mène à la naissance d'une nouvelle conception de musique : la musique électronique, qui apparaît en 1950.

Le progrès avance vite et après seulement quatre ans, on passe de la musique qui exploite des sons déjà présents en nature à des sons synthétiques, la recherche des sons artificiels mène à la naissance d'une nouvelle conception de musique : la musique électronique, qui apparaît en 1950.

Jusqu'à ce moment de l'histoire humaine, la production musicale était toujours de nature analogique : bandes magnétiques pour enregistrer, vinyles, cire. Vers la moitié de vingtième siècle les connaissances informatiques sont tellement avancées qu'il semble possible de les exploiter pour produire de la musique. 1956 est la date officielle pour la naissance de la musique assistée par l'ordinateur. Puisque la musique classique répond à des règles précises (éviter le quintes, les octaves parallèles, règle d'harmonie et contrepoint) ça devrait être possible de codifier ces règles en langage informatique et faire « calculer » une composition par un ordinateur simplement en évitant les fautes ; les choix étaient faits selon la méthode de Monte-Carlo. Le choix des accords et des mélodies était effectué purement de manière statistique, l'ordinateur était en quelque sorte un lanceur de dés automatique qui sélectionnait parmi les combinaisons qui ne présentaient pas d'erreurs pour composer. La symphonie produite aux Etats Unis avec l'ordinateur Illiac s'appelle Suite Illiac Quatuor Accords.

Bien évidemment la composition résulterait étrange comparé à celle de Bach, même si les règles employées par les deux « compositeurs » étaient les mêmes ; ce n'est pas la question de fixer des règles qui détermine l'expressivité. Les règles nous disent ce que le compositeur ne doit pas choisir mais jamais ce qui est bien. Le possible ne suffit pas pour créer le beau.

La logique, la raison, les expériences humaines du compositeur entrent dans les choix que lui fait parmi les notes possibles, on ne peut pas espérer de réussir à synthétiser une émotion. Un ordinateur est capable de produire le gros du phénomène mais c'est la partie plus subtile du phénomène de composition que ne peut pas être assimilé. C'est une idée anthropomorphe que de donner à un ordinateur la sensibilité ou l'intelligence d'un compositeur.

C'est pour cette raison que Xenakis n'accepte pas a priori tout ce que sorte de l'ordinateur, mais le classifie, le travaille avant être satisfait. C'est la naissance de la musique stochastique en 1960.

En 1957 Max Mathews, considéré comme le père de la musique électronique, crée le premier convertisseur numérique, c'est le début de la synthèse du son. Grâce aux techniques de synthèse commence aussi les premières études de psycho-acoustique.

Après la naissance de la musique électronique dans la moitié du siècle dernier, se créent deux courants différents qui se concentrent sur deux aspects complémentaires de la musique : une sur la synthèse des sons, l'autre sur la composition programmée.

4 La musique concrète

La découverte du transistor dans l'année 1948 permet d'introduire le concept très important d'amplification électrique, les premières notions d'échantillonnage, la théorie de Shannon et donc les outils nécessaires à l'enregistrement et élaboration sonore.

En 1948 naît, avec Pierre Schaeffer ce qu'il appellera « musique concrète », à partir d'expériences au sein de la radio, autour des techniques d'enregistrement qui permettent pour la première fois dans l'histoire de l'humanité de capturer le fugitif phénomène sonore d'une manière objective.

Notation vs enregistrement : la mémorisation d'un événement sonore

La notation musicale, comme on a vu, est une tentative humaine de lutter contre la volatilité d'un phénomène sonore ; mais l'écriture n'arrive pas à capturer la totalité d'un élément sonore. On réduit la son à un phénomène musicale, à une représentation mathématique ; même pour une œuvre musicale on décrit la hauteur des notes, le rythme, on ne pourrait pas décrire le timbre, la puissance employée. Dans le cas d'un événement sonore non musical le problème de mémorisation est encore majeur et la notation ne pourra pas nous aider pour le reconstruire.

L'enregistrement est le fruit de l'envie humaine de pouvoir conserver le son, réussir à rendre objectif, présent un phénomène sonore.

Mémoriser consiste à remplacer (représenter) une chose volatile par une autre qui existe tout le temps.

On pourrait distinguer en premier approche quatre catégories pour la chose remplaçante, elle peut être mentale (un modèle, une représentation) ou matérielle (un objet) et pour chacun de ces deux cas elle peut se présenter sous un état statique ou en évolution¹⁴. On peut réduire de moitié les catégories préférant une chose qui soit statique pour conserver nos idées, plutôt que quelque chose qui change et qui donc peut modifier l'information que l'on voulait stocker et garder identique à l'original.

Dans les deux catégories qui restent on choisirait un objet matériel sachant par expérience commune que la mémoire humaine n'est pas fiable comme un objet matériel stable pour le stockage d'information. On cherchera donc un objet physique en état statique.

Dans la peinture la toile sert comme instrument de production artistique et comme objet artistique en lui-même (propriété auto-graphique de l'art). La musique comme on a déjà vu est différente, on nécessite d'un objet physique au-delà de l'instrument pour la conserver (propriété allo-graphique de la musique) : l'enregistrement sur un support.

La mémorisation d'un phénomène physique qui évolue dans le temps, comme un événement sonore est un processus complexe qu'en général l'on peut diviser en trois étapes fondamentales :

1) une transduction : une transformation par un moyen analogique qui crée une correspondance entre un phénomène d'une nature (par exemple acoustique) en un phénomène d'une autre nature (par exemple magnétique).

2) une conservation : une transformation intermédiaire qui change l'état évanescence du phénomène physique en question en un état permanente, stable, consultable.

¹⁴ C'est évidente que tout peut être considéré comme "en évolution", mais on considère de se référer à l'échelle temporelle et spatiale humaine dans ce cas.

3) adressage : c'est un moyen de faire correspondre une évolution temporelle en disposition spatiale. Pour cela on nécessite d'un moteur : la relation plus simple entre temps et espace est la vitesse (s/t) ; un moteur à vitesse constante peut donc servir à transférer l'évolution temporelle du phénomène sur un support qui possède une dimension spatiale. C'est à ce niveau de création sonore qu'entre la technologie : un transducteur permet la transformation du temps en espace. Le gravage est pratique à la fin des phénomènes d'enregistrement : par ce moyen on transforme un état fugitif d'un système en état permanent stable, par le moyen de la déformation d'un support.

Un phénomène sensible stocké dans cette façon n'est pas forcément accessible par nos sens, donc pour pouvoir le consulter on nécessite un processus contraire : une transduction inverse.

À la fin de ce processus de mémorisation (commutation du temps en espace) on obtient un objet physique qui conserve la mémoire de l'objet sonore : le support ; on manipule l'objet physique produit par l'enregistrement : l'étirer ou comprimer temporellement, modifier son amplitude, inverser le déroulement du temps, le dupliquer¹⁵ ; on peut exécuter des opérations que jamais on aurait pu appliquer au phénomène physique.

L'objet sonore

L'idée de Schaeffer sort d'une expérience accidentelle : une rayure sur un disque, provoquait la répétition en boucle d'une partie sonore ; l'effet fut une totale abstraction du son de son contexte, le son était vidé du sens général et apparaît comme une entité indépendante, un phénomène isolé.

La boucle du disque sort de son contexte la séquence sonore, efface son sens, qui dérive des sons qui la précède et la suit, et donne une nouvelle dimension à l'écoute ; c'est la prise de conscience de la relativité de l'écoute, le phénomène sonore, chaque phénomène sonore prend une autre valeur si analysé séparément de son contexte¹⁶.

De cette expérience née une nouvelle approche à la musique, fondée sur l'enregistrement du son et sur les techniques de manipulation et traitement du support même de cet enregistrement (on traite pas le son, on change pas ses composantes fréquentielles, son amplitude mais on travaille sur le support même sur lequel le son est déjà présent).

Pour décrire sa conception Schaeffer parle d'« objet sonore » pour indiquer l'événement sonore privé de son contexte ; pour pouvoir le percevoir il faut réussir à opérer une réduction matérielle et spirituelle d'un enregistrement sonore. L'objet créé, coupé sur la bande magnétique, vient décontextualisé et perd sa propriété d'exprimer sensations, et idées, . Le son peut plus être utilisé pour se renseigner sur autres choses : c'est le son lui-même. On atteint de cette façon une objectivité matérielle acquise par le phénomène sonore du fait de son enregistrement et par la possibilité de le reproduire à l'identique, indépendamment des actions et des positions du sujet ; l'objet sonore est une chose permanente, extérieure au sujet et indépendante de lui. Souvent pour décrire la musique concrète on emploie le terme « acousmatique ». Acousmate c'est un son dont on ne voit pas les causes, la musique concrète est donc acousmatique dans le sens que l'on crée une séparation entre le son et son origine, mais c'est encore plus, c'est même créer une distance entre l'événement sonore et les parties qui le précède et le suivent, c'est décontextualiser un moment sonore et le reproduire tout seul donc enlevant les causes mais aussi les effets, et donc le sens totale logique de n'importe quelle production sonore. Cette réduction est beaucoup plus qu'une simple élimination des causes est une idéalisation du phénomène, un phénomène d'abstraction : on enlève toutes nos opinions, pour simplement écouter le contenu objectif.

¹⁵ Les très communes opérations de copie/collage que on exécute tous les jours sur nos ordinateurs sont originaire des mêmes opérations que Schaeffer appliquait physiquement à ses bandes magnétiques.

¹⁶ La nouvelle vision des phénomènes sonores de Schaeffer ouvre une porte sur le monde de la psychologie de la perception sonore. Après la sortie de son livre : « Traité des objets musicaux » commencent les premières expériences modernes en psycho-acoustique.

Tout phénomène sonore (concret) est susceptible de devenir musical : c'est peut être la plus importante conséquence de cette nouvelle forme de penser la musique.

Musique abstraite et musique concrète

Le terme musique concrète a été employé par Schaeffer pour s'éloigner de la musique que lui définit comme abstraite, se référant à la tradition musicale habituelle : elle est abstraite dans le sens où elle est d'abord conçue par l'esprit, « entendue », imaginée (par l'oreille interne du compositeur) et pensée, et puis notée théoriquement, ce qui constitue une opération d'abstraction. L'écriture consiste en fait à inscrire la musique dans un espace à deux dimensions : les hauteurs et le temps. Puis, dans un deuxième temps, vient enfin réalisée dans une exécution instrumentale. La musique de Schaeffer est donc « concrète » parce que elle est « constituée à partir d'éléments préexistants empruntés à n'importe quel matériau sonore, qu'il soit bruit ou musique habituelle, puis composée expérimentalement par une construction directe ». La musique habituelle est « abstraite » car elle procède de l'« abstraction » de l'écriture vers sa concrétisation : un objet sonore est d'abord pensé est en suite matérialisé dans un son ; la musique « concrète » en revanche, escamote l'étape de l'écriture, elle prend un objet sonore concret et le rend abstrait par sa decontextualisation ; de cette façon l'on perd le concept initial.

Conséquences philosophiques sur la composition

Tout objet sonore peut devenir musical. L'idée que soutient cette affirmation c'est que des fois l'art est dans l'objet avant d'être dans la tête de l'artiste, chose que n'arrive avec les concepts de la musique abstraite. La question que se pose est alors : est ce que c'est encore de l'art de découvrir par hasard un phénomène sonore ? Personnellement je suis d'accord avec l'idée qu'il est juste de conserver n'importe quel objet sonore si il est dense digne de d'une importance artistique, le choix de garder un phénomène, le fait de s'apercevoir de la beauté d'un échantillon sonore est déjà une démarche artistique, le degré zéro comme on l'a déjà dit du processus de création artistique, choisir consciemment est déjà composer, être présents et perceptifs au moment du phénomène sonore est déjà un moment artistique.

5 La musique électronique et la synthèse numérique du son

À la différence de la musique concrète de Pierre Schaeffer dans laquelle l'oeuvre d'art est produite à partir de l'enregistrement d'un son déjà existant, dans la musique électronique on a aucun son à l'origine. Le mouvement artistique électronique est donc conceptuellement à l'opposé de la conception Schaefferienne de la création sonore.

5.1 Analyse sonore

L'enregistrement sonore donne comme on l'a déjà dit un rôle central au son lui-même (le statut d'« objet » dans la vision Schaefferienne), les techniques d'enregistrement nous permettent de transmettre un son dans l'espace et dans le temps, de surpasser les limites physiques naturelles des phénomènes sonores et de les rendre (presque) éternellement exploitables. Mais il y a une grande conséquence à cette disponibilité des sons sur un support : la possibilité de modifier le support et donc les sons enregistrés jusqu'à produire d'autres sons ou même des sons que dépassent les possibilités naturelles. C'est à cette étape que naît la notion du signal d'information et de transmission de signaux.

Le signal

Est un objet qui présente un double visage :

- il a un côté concret, matériel : c'est alors une courbe plus ou moins complexe gravée sur un support physique ; ça peut être une tension, ou un courant, en général c'est la représentation d'un phénomène difficilement palpable par le moyen de transduction. Mais pour que un

signal puisse être produit à partir d'un phénomène physique il faut que il soit possible le mesurer instant par instant pendant son évolution.

- Il a aussi une côté abstrait, conceptuel : la courbe représente la variation quantitative des variables sonores dans le temps, c'est une représentation durable.

Les deux concepts comme les pôles d'un micro-aimant, composant une bande magnétique, sont indissociables : la mesure des grandeurs physiques, matérielles du son nous plonge dans l'univers mathématique, conceptuel qui nous permet de parler du « signal ».

J'explique mieux ce point important ; le phénomène n'est pas analysable, ni mesurable, c'est son signal que nous pouvons traiter analyser que nous mesurons. En général c'est facile de confondre les concepts et de ne pas comprendre le sens profond de cette distinction fondamentale. Dans la vie de tous les jours on traite en général trois types d'objets : les phénomènes réel qui ne sont ni mesurables, ni donc traitable, les signaux, qui sont la mesure des phénomènes réels, et les représentations des signaux (courbes dessinées : fonctions, la partition musicale) qui nous permettent de les visualiser et analyser faciles.

Concepts d'analyse des signaux

Dans la musique concrète le processus d'analyse sonore comporte d'abord un approche matérielle : capturer et enregistrer les échantillons sonores sur un support (bande magnétique) Avec cette méthode on décompose le son et on prélève des échantillons, transformant un phénomène sonore en phénomène électrique. On peut ensuite couper la bande magnétique pour examiner les différents échantillons ou groupes d'échantillons. Les différentes composantes sonores du signal son enlevé de leur contexte pour permettre de les percevoir différemment : on arrive à considérer le tout comme composé de parties interdépendantes (approche mental) ; essayant de comprendre comment les différentes parties sont sémantiquement reliées entre elles dans la constitution de l'ensemble on arrive à comprendre la structuration sonore du phénomène enregistrée¹⁷.

Le processus d'analyse est donc d'abord matériel (coupure de la bande), mais c'est le début d'un processus d'analyse mental (séparation sémantique des parties sonores) ; cette approche est une pratique que on peut apprendre à effectuer systématiquement pour analyser les sons de tous les jours ; par la pratique matérielle et l'analyse patiente de plusieurs exemplaires d'enregistrement sonores on peut développer une capacité à l'écoute et à l'analyse qui nous permet à l'oreille de décomposer les événements sonores d'une façon qu' on n'aurait jamais pu atteindre sans le moyen de l'analyse matérielle. C'est une nouvelle écoute qui nous permet d'atteindre un niveau plus haut dans nos facultés de cognition sonore et d'analyse mentale, jusqu'à arriver à décomposer les sons dans la tête beaucoup plus finement qu'avec les appareils matériels d'un studio.

Analyse de Fourier

La théorie des signaux nous fournit différents outils mathématiques pour commencer une analyse détaillée des propriétés fondamentales des signaux en général ; mais il faut clarifier le domaines de validités de ces théories avant de s'aventurer à leur utilisation, sans être conscient des avantages et des inconvénients qu'une mauvaise application peut entraîner.

Pour caractériser univoquement un signal on aurait besoin de définir pour chaque instant sa valeur; on devrait pouvoir le représenter sur un graphique cartésienne sans devoir laisser des « trous » dans le dessin de sa courbe (on pourrait définir ça comme degré 0 du processus de mesure d'un signal). Evidement tout ça est théorique : dans la vie courante on peut ne pas penser à réaliser des mesures infiniment détaillées, même si la technologie nous aide tous les jours en trouvant un moyen plus précis pour analyser les choses. Soit les limites technologiques (temps de latence des processeurs ou des appareils mécaniques) et soit les limites économiques (nombre des données par unité de temps)

¹⁷ Bien évidemment il n'y a pas une façon unique de décomposer le tout en parties, l'analyse c'est un processus quand même subjectif et donc guidé par la logique de chacun, un travail personnelle selon la sensibilité individuelle.

nous imposent une limite et nous obligent à prendre des instants finis, petits selon le besoin, pour la mesure et le stockage des signaux.

Analyse de Fourier d'un signal périodique

La théorie de Fourier est basée sur le concept de courbe sinusoïdale : une courbe périodique caractérisée par quatre variables : T (période), f (fréquence), φ (phase), A (amplitude)¹⁸.

Le théorème de Fourier affirme que toute fonction périodique peut être représentée comme une somme des fonctions sinusoïdales (série de Fourier). C'est avec cette affirmation que commence déjà le processus d'analyse parce que on peut penser à superposer différentes courbes (dites harmoniques) chacune caractérisées que par trois variables : A , f , φ .

Cette approche nous mène à construire ce que l'on appelle le *spectre du signal* : un graphique qui représente entièrement le signal analysé par deux variables : amplitude (en ordonnée) et fréquence (en abscisses) de ses composantes harmoniques¹⁹.

Dans le cas des signaux périodiques il prend la forme d'un peigne : des raies positionnées correspondant aux différentes fréquences des harmoniques et de hauteur correspondant à l'amplitude de chaque harmonique. On pense alors pouvoir avoir une liaison biunivoque entre signal et son spectre, et donc avoir le droit de représenter, et même de remplacer les signaux naturels par leur analyse. Est-ce que c'est juste de faire cela, dans quels limites ?

Si on analyse n'importe quel signal de la nature, on peut dire mathématiquement que son support est borné, que donc sa durée dans le temps est finie. Comment peut-on décrire quelque chose de fini par la superposition de fonctions infinies comme les fonctions sinusoïdales ? Chaque courbe sinusoïdale périodique se répète sans fin, c'est une abstraction mathématique idéale, mais non applicable à la réalité dans laquelle chaque phénomène est fini.

On essaiera dans la suite de donner une réponse à ces questions et de définir un domaine de validité pour la théorie de Fourier ou au moins de décider les avantages et les désavantages que porte une telle approche.

Pour représenter des signaux finis on peut penser appliquer la théorie de Fourier sans changer ses concepts ou la mathématique qui est à sa base, simplement en changeant les fonctions employées : à la place des sinusoïdes infinies on peut prendre des pseudo-sinusoïdes : des fonctions périodiques nulles en dehors de leur support de définition ; par leur superposition l'on crée alors des signaux réels (c'est cela la base de la synthèse additive dont on parlera dans la suite). Tout signal réel peut être décomposé en un certain nombre des signaux selon une base des fonctions réelles (finies).

Analyse de Fourier d'un signal quelconque (non périodique)

On peut considérer un signal quelconque variable continûment au cours du temps comme une fonction périodique de période infinie ; la transformée de Fourier (en forme intégral : \int) est alors la généralisation de la série de Fourier (en somme de fonctions : Σ) dans le cas des fonctions non périodiques. Le spectre d'une telle fonction est alors continu, et n'est pas discret (peigne) comme dans le cas périodique, mais est une fonction contenant toutes les fréquences. La transformée de Fourier, comme la série de Fourier, entraîne des problèmes conceptuels profonds :

- on a à traiter des quantités infinies : les raies discrètes dans le cas discret se rétrécissent jusqu'à créer une fonction continue dans le cas continu (on a infinies fréquences entre les intervalles discrètes d'avant), on a une densité infinie des harmoniques ;
De plus les fréquences des harmoniques composants les signaux peuvent varier entre 0 jusqu'à une valeur infinie en théorie, le spectre peut être mathématiquement infini.

¹⁸ On sait bien que période et fréquence sont l'un l'opposé physique de l'autre, donc en général ça suffisent trois variables pour définir une courbe sinusoïdale.

¹⁹ La phase est moins importante à cet niveau d'analyse ; la plupart de l'information nécessaires sur un signal sonore sont véhiculés par l'amplitude et fréquence de composantes.

Heureusement en nature les signaux n'ont jamais un spectre infini²⁰, ce que nous permet d'être sûrs de trouver une fréquence maximale F_m dans n'importe quel spectre et donc de pouvoir appliquer le théorème d'échantillonnage. Mais la densité infinie des fréquences composantes reste encore un grand problème : par des signaux avec un spectre infiniment dense on peut pas penser de prendre toutes les fréquences, les limites techniques nous imposent de choisir, donc d'enlever des composantes dans l'analyse.

Techniques d'échantillonnage

Créer un signal numérique à partir d'un signal réel de base est une des problématiques technologiques les plus importantes des dernières années. A la base de ce phénomène est l'impossibilité technologique de mesurer l'amplitude d'un signal originelle à chaque instant. On est donc amené à choisir un nombre fini des points dans un espace temporel correspondant lesquels seront les mesures du signal. Pour simplifier le problème on décide d'abord de prendre des intervalles réguliers et après de définir une fréquence (inverse de la période minimum) à laquelle effectuer les mesures; ce processus appelé technique d'échantillonnage se base sur une double discrétisation : une substitution du signal original par un nombre des mesures d'amplitude prises à des intervalles égaux et une quantification des valeurs en amplitude (que peuvent aussi par nature prendre infinis valeurs) sur différents niveaux, selon le nombre de bits à disposition dans l'échelle d'amplitude. On a donc une double discrétisation : sur la fréquence et sur l'amplitude, ce que détermine que le signal échantillonné puisse « ressembler » (avec un certain degré de fidélité dépendant de la largeur des intervalles choisis) au phénomène sonore mais ce n'est jamais pareil. La théorie développée par Claude Shannon (1916-2001) nous donne une règle pratique pour choisir la fréquence d'échantillonnage suffisante à stocker les principales informations d'un phénomène physique. La fréquence d'échantillonnage F_c doit être au moins le double d'une fréquence maximale F_m du spectre relatif au phénomène naturel.

En pratique, technologiquement, la conversion Analogique-Numérique est réalisé par l'unité CNA (convertisseur numérique-analogique) construite dans l'année 1956 par John Pierce et Max Matthews. Le fonctionnement de ce appareil a plusieurs étapes : d'abord on stocke une série d'échantillons suivant les discrétisations décrites auparavant; on effectue une interpolation des données par exemple avec un filtre passe-bas qui élimine les « sauts » entre les différentes amplitudes. Cela comporte deux approximations : sur les amplitudes, à cause de l'interpolation on introduit un écart par rapport au signal original, et sur la phase, le temps d'élaboration des données crée un décalage minimum d'une période (cela crée un problème importante sur les systèmes bouclés, systèmes en temps réel, et donc sur tous les systèmes à retour d'effort).

Mais l'introduction d'un tel système de conversion n'a pas seulement permit la réalisation de l'analyse sonore (et l'enregistrement numérique) mais aussi le processus inverse : la synthèse. On peut commencer créant des signaux sinusoïdales différents à les additionner, et les convertir en signaux analogiques et les reproduire sur des enceintes : produire un phénomène sonore à partir d'un calcul numérique, c'est la naissance de la synthèse sonore.

²⁰ On peut voir mathématiquement ce problème écrivant des équations différentielles par différents phénomènes physiques naturels, les linéariser par le moyen de la transforme de Laplace et analyser l'équation que on trouve. En général on obtient un fraction de deux polynômes, par l'étude des pôles et zéros est possible déduire des importantes caractéristiques du phénomène. Les phénomènes naturels ont toujours le polynôme au dénominateur d'un degré majeur de celui du numérateur, ce qui implique que la solution de l'équation différentielle est toujours trouvable a priori en nature.

5.2 La synthèse numérique du son

5.2.1 Les premiers concepts de synthèse

La synthèse sonore numérique commence dans l'année 1957 quand Max Matthews réalise le logiciel « MusicV » pour exploiter les idées de composition des signaux sinusoïdales et réussir à créer des sons complexes et des créations musicales totalement numériques. L'intérêt des travaux de Max Matthews est la création sonore, la composition musicale à partir d'aucun son. C'est le concept inverse de l'idée artistique de Schaeffer, qui utilisait des sons trouvés dans la nature ; Matthews part de rien pour créer²¹.

Par rapport à la musique traditionnelle la musique concrète abandonne le geste instrumental, il n'y a pu d'interaction matérielle directe entre son et homme, l'action humaine est produite sur le support mais pas sur le son lui-même. Dans la musique de synthèse on trouve même pu du son au début, il a aucun phénomène vibratoire, pas du geste du retour tactile, on perd la matérialité, le contact réel avec l'extérieur. Est-ce que on peut vraiment parler d'« instrument électronique », par exemple d'un ordinateur ?

Les problèmes technologiques liés à la composition des sons numériques ne sont pas simples à résoudre.

Le premier obstacle c'est le temps mort introduit par le calcul : la présence d'un déphasage empêche de penser de créer du son en temps réel²² comme avec un instrument traditionnel, c'est la naissance du concept de *temps diffère*. Par exemple avec Music V la fabrication de 17 seconds du son dans les années 60 pouvait prendre 1 jour et ½. On a récemment atteint le but de produire et contrôler la musique en temps réel (depuis le 1983), et aujourd'hui différents systèmes nous permettent plus ou moins de créer des compositions sans déphasage.

On peut classer les systèmes dits à *temps réel* en trois catégories selon les temps de calcul que emploient :

- 1) Systèmes à temps réel fort quand Latence $\text{Entrée/Sortie} < T_{\text{échantillonnage}}$; on sait que le temps de latence 0 est impossible.
- 2) Systèmes à retour d'effort (temps réel moins fort) : à l'entrée on a des signaux gestuelles (commande instrumentale) échantillonnées à 1-2 kHz²³ ; le temps de calcul pour produire une sortie est plus grande que dans le cas précédent.
- 3) Systèmes à temps réel simple : on a rien à l'entrée et on se préoccupe de rien stocker, on produit simplement des éléments sonores en temps réel

La recherche temps réel a été un des principaux intérêts des années 70. Plusieurs systèmes ont été inventés ;

Max Matthews – Groove (fin des années 60)

²¹ Le même Schaeffer avec un parallélisme explique très bien le rapport entre musique concrète et synthétique comme la relation entre peinture et sculpture : dans la musique synthétique comme dans la peinture l'artiste doit partir d'une toile vide, blanche et dessiner créer son monde artistique, dans la musique concrète et dans la sculpture l'artiste enlève, travaille ce que il a déjà, à partir de la matière informe crée l'objet artistique.

²² On parle en général du « temps réel » si l'émission sonore d'un échantillon a lieu avant le calcul du prochain échantillon, si, donc, on a que le temps de l'écriture, modification et sortie des données est mineur du période d'échantillonnage.

²³ On sait bien que la fréquence de gestes humains ne peut pas atteindre fréquence majeures que 50 Hz donc, en général, pour transmettre à la machine les informations sur nos mouvements gestuels suffirait un échantillonnage à 150 Hz idéalement, selon le théorème de Shannon. Le problème est le retour d'effort des vibrations de l'instrument : un archet de violon par exemple engendre des vibrations que arrivent à des kHz ; si on veut rendre réelle la perception d'un instrumentiste avec le retour d'effort on est obligés (test à l'ACROE) à prendre des hautes fréquences d'échantillonnage que rend les calculs plus longs.

G. Di Giugno - K-trix (processeur numérique spécialisé) c'est un système séquentiel, mais avec un temps de réponse très rapide => simulation temps réel. IRCAM

Sitter (GRM)

Fin années 80 avec machines parallèles => on simule un processeur spécialisé sur une machine monoprocesseur

Années 90 parallélisme => avec un seul processeur on simulait jusqu'à 1024 processeurs.

Aujourd'hui redimensionnement : maximum 7 -8 processeurs simulés.

Le deuxième obstacle que limite fortement la synthèse numérique est le problème d'économie/généralité : on sait que l'espace perceptif de l'oreille humaine est compris entre 20 et 20.000 Hz ce que il veut dire selon le théorème de Shannon, échantillonner le son au moins à 40.000 Hz. La choix officiellement accepté est de échantillonner à 44.100 Hz (qualité CD)^{24, 25} ; cela veut dire que si on veut créer des son synthétiques on doit être capables de produire 44.100 échantillons sonores par chaque second de musique désirée. On est donc mené à trouver des règles de génération que à partir d'un nombre petit de données arrivent à créer un grand nombre d'échantillons. Pour faire cela on écrit des algorithmes, mais le problème général est : comment les définir ? Comment les régler et comment les évaluer à la fin ? Ces sont des questions cruciales qui ont influencé l'histoire de la musique des dernières dizaines d'années ; les chercheurs se sont concentrés produisant idées d'intérêt remarquable, mais aussi algorithmes que dans le nome de l'économie mènent a des création sonores prive de valeur esthétique.

C'est aujourd'hui évident que seulement notre système cognitif peut juger la qualité sonores des résultats produits par un algorithme et que ni l'économie (donc les propriétés des généralité) ni la facilité de mise en oeuvre des algorithmes sont des paramètres suffisants pour employer un système de création musicale numérique.

Music V

C'est Max Matthews qui en premier résoud le problème de la synthèse en proposant un outil (Music V) contenant dans sa conception les concepts fondamentaux de plusieurs approches de synthèse qui seront expliquées dans les chapitres suivants.

L'objet fondamentale de Music V est l'« OSC » un module symbolisant un oscillateur électronique ; il possède deux entrées : fréquence et amplitude, et une sortie : un signal déterminé par la table d'onde d'amplitude et de fréquence fixées par les entrées.

En assemblant différents modules, on peut créer des sons complexes : les additionner, les mélanger et utiliser la sortie d'un oscillateur comme entrée d'un autre.

Cet caractéristique fais de Music V le premier système de blocs fonctionnels modulaires et rend très simple la syntaxe de création musicale.

²⁴ Le débat pour la meilleure fréquence pour l'échantillonnage est encore ouvert même si la plupart des gens reconnaissent la valeur optimale à 44.100 Hz, appliquant en gros le théorème de Shannon sur la fréquence maximale théoriquement audible par l'oreille humaine : 20.000 Hz. Mais il y a quand même des divergences par exemple le compositeur Iannis Xenakis estime être important échantillonner à 300.000 Hz parce que des phénomènes de bâtiment être fréquences à 299.000 et 300.000 Hz peuvent entraîner des effets autour des 1000 Hz (donc audibles) importantes pour une pièce musicale, que seraient évidemment éliminées par les conventions d'échantillonnage officielles.

²⁵ Au delà de l'échantillonnage en fréquence les standard CD on déterminé aussi un standard de quantification sur l'amplitude : 24 bit. Cela veut dire que par 1 second de son on pourrait représenter $7,6 \times 10^{80}$: 2^{24} signaux différents pour chaque instant d'échantillonnage.

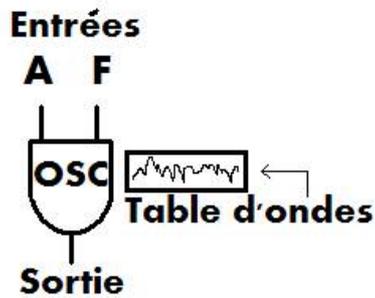


Figure 1 - L'oscillateur Music V avec les entrées (Amplitude et Fréquence), la sortie et la table d'onde

Un autre concept existe en Music V : l'ensemble des opérations que Schaeffer pouvait appliquer sur une bande magnétique est totalement transféré de la réalité à la machine en traitant les « tables d'ondes » qui sont en général un ensemble de 512 échantillons ; les tables d'ondes peuvent être copiées et collées, coupées exactement comment il était possible sur bande magnétique, mais avec une précision supérieure. La fréquence des tables d'ondes est calculable par la formule $F_t = F_e / N$. Si N est 512, le nombre d'échantillons par table, et F_e est la fréquence d'échantillonnage : 44.100 Hz, avec un calcul rapide on peut trouver que la fréquence de la table d'onde est de 86 Hz. On stocke alors sur 86 Hz à la place de 44.100, ce qui nous permet d'économiser en calcul. Une table peut être employée par chaque oscillateur et peut subir des modifications en amplitude (enveloppe) selon la forme du signal à l'entrée d'amplitude de l'oscillateur.

Pour diminuer la fréquence du son à la sortie, il suffit de prendre un petit nombre d'échantillons et d'en ajouter pour l'augmenter virtuellement la taille de la table d'onde. Si l'on veut des valeurs de fréquence non multiple de la fréquence d'échantillonnage on est alors obligé de prendre des valeurs intermédiaires entre les échantillons à l'aide d'une interpolation.

Dans Music V, comme nous l'avons déjà vu, on peut composer les blocs fonctionnels de manière à contrôler l'entrée d'un bloque par la sortie d'un autre bloque. C'est le cas dans l'exemple de la modulation de fréquence découverte par John Chowning (et appelée ensuite *Synthèse FM*²⁶) ; la même configuration était exploitée pour produire des vibratos (à fréquence de 5 – 6 Hz) mais quand Chowning a utilisé des fréquences radio qui dépassaient les fréquences audios, il découvrait que les spectres produits étaient très riches en fréquences et donc très intéressants en terme d'économie de composition.

Avec Music V, nous pouvons aussi construire des circuits additionneurs et donc créer des sons à partir de plusieurs oscillateurs : en composant différentes fréquences nous sommes prêts à produire de sons plus riches.

De plus dans le logiciel nous avons à disposition une unité appelée RAN que est un générateur aléatoire des sons (bruits), un concepts encore jamais utilisé dans l'histoire de la musique et hérité directement du langage mathématique.

Music V est doué aussi d'une « partition » : un fichier dans lequel on peut stocker les informations sur les différents instruments que l'on utilise (liste qui se termine par « END »), décrivant les configuration des oscillateurs et leurs caractéristiques. Une deuxième partie est dédiée à la

²⁶ Exploitée à niveau commercial par Yamaha dans le synthétiseur DX-7. C'est intéressant de remarquer ici que Music V fut inspirée par des principes du monde matériel : les tables d'ondes à la bande magnétique, l'oscillateur à des composants électroniques, et comme aussi le monde technologique se soit en suite inspiré à MusicV pour la construction d'objets réels: par exemple Yamaha avec son DX-7 qui utilise la synthèse FM ou le synthétiseur MOOG qui utilise un knob de control appelé « Voltage Control », inspiré à l'oscillateur Music V pour sa construction.

description des notes : le moment de l'exécution (« ACTION TIME ») et les paramètres de chacune.

Une note peut être composée de plusieurs sons : plusieurs instruments additionnés pour produire un timbre riche. Il peut aussi arriver que plusieurs instruments dépendent d'une même note (fonction random). Tous les concepts développés dans Music V (qui est maintenant diffusé dans une version actuelle sous le nom de « *C Sound* ») viennent du monde réel.

Grâce à tous les outils de Music V, on semble avoir trouvé un langage universel pour décrire un signal sonore et pour pouvoir reproduire parfaitement chaque son de la réalité par des simples modules de synthèse.

Dans le 1969 Jean Claude Rissent réalise un catalogue des timbres en forme écrite et fixe pour MusicV dans lequel on stocke les informations relatives à chaque timbre d'instrument réel synthétisé dans le langage MusicV. Par cette approche, il s'aperçoit qu'on ne peut pas penser de reproduire le timbre d'un instrument simplement à partir de ses composantes spectrales (comme on avait toujours pensé dans le passé), mais que le phénomène sonore d'un instrument est beaucoup plus complexe, c'est la variation dans le temps de ses caractéristiques spectrales qui rend compte de son timbre ; par exemple une trompette est bien décrite par une fonction spectrale qui augmente le poids relative des hautes fréquences quand le trompettiste joue plus fort. Trouver le timbre d'un instrument requiert une longue écoute attentive pour essayer d'identifier les fondamentales caractéristiques qui déterminent le son de l'instrument, c'est une analyse des petites vibrations, des bruits aléatoires, une analyse fine et importante, sans laquelle nous ne pourrions pas reproduire le même phénomène ; nous devons nous concentrer sur l'évolution des composantes dans le temps.

5.2.2 Le midi et le sampler

Dans les années 1983, d'une série d'études portant sur la reproduction sonore des instruments de musiques, naît le MIDI (Musical Instrument Digital Interface) qui est diffusé avec le synthétiseur Dx-7 Yamaha. C'est dans cette période que croît un grand intérêt pour l'informatique musicale et que l'on assiste à une vulgarisation de la musique électronique.

Le MIDI a de nombreux avantages : tous les instruments MIDI ont une grande compatibilité entre eux, ce nouveau langage permet la communication entre différents modules instrumentaux, des échanges d'informations, la synchronisation, etc. ; malgré qu'il soit utile et rapide c'est malheureusement un codage très simplifié et très réducteur dans les possibilités de composition sonore.

Par ailleurs, la grande magie qu'apporte le langage MIDI, est le fait de pouvoir composer la musique d'un orchestre entier sans nécessiter de connaître les différents instruments, mais il suffit de n'en connaître qu'un : par un piano on peut réussir à jouer une trompette ; grâce à un ordinateur on pourrait concevoir le son d'un ensemble complexe autant que l'on veut de différents instruments, sans savoir les jouer physiquement.

Le MIDI est un concept économique très lié aux concepts musicaux classiques : hauteur, fréquence, intensité sonore ; mais il n'arrive pas à rendre compte des subtiles interactions instrumentales qui rendent un son *vrai* ou intéressant.

C'est la période de la musique « simple » : composer avec des banques de sons déjà préparées et de créer des logiciels qui le gèrent n'a jamais été aussi facile qu'avant. Mais l'on perd le contact matériel avec l'instrument l'essence de la musicalité, le jeu est transformé en engendrement de signaux musicaux.

Les techniques du sampling commencent à se diffuser (années 80). Les machines de calcul avaient deux grands problèmes qui empêchaient un travail continu : temps de calcul très grands et une mémoire disponible très petite par rapport au nombre des données à produire. Les développements de la technologie ont résolu rapidement le deuxième problème, dans les années 80 les mémoires deviennent beaucoup moins chères et sont donc disponibles sur le marché à des prix accessibles ; les musiciens qui avaient beaucoup de sons mais pas assez de mémoire dans laquelle les stocker ont maintenant la possibilité de mémoriser tous leurs sons. Mais il faut se rappeler que le sampling n'est

pas une technique de synthèse parce que l'on part déjà d'un son de base, provenant de la nature ou synthétisé mais jamais avec le sampling on crée des son a partir de rien²⁷.

5.3 Les différents types de synthèse

Les années 1965-1975 représentent une période de recherche intense de toute la communauté musicale mondiale pour la découverte des nouvelles méthodes de synthèse (algorithmes d'engendrement), après cette période, l'intérêt s'est éparpillé ailleurs et aujourd'hui on se concentre sur d'autres problématiques du son.

Les différents types de synthèse développées en cette période ne sont pas indépendants : MusicV peut être défini comme l'environnement commun pour la synthèse additive, soustractive, la distorsion non linéaire et la modulation de fréquence ; mais on a besoin d'autres concepts de base pour les autres types de synthèse : des nouveaux formalismes.

Un chapitre sera dédié à la synthèse des sons par modélisation physique qui se distingue de toutes les autres par une approche lointaine du traitement des signaux.

Comme nous le verrons par la suite, la communauté musicale n'a pas encore réussi à créer une méthode valable dans chaque situation et optimale pour produire toutes types de signaux, mais on a plusieurs approches qui nous permettent de bien recouvrir toutes les typologies des domaines sonores. Il est alors à l'utilisateur, par son expérience de décider la méthode qu'il préfère appliquer dans chaque situation particulière.

5.4.1 La synthèse additive

L'analyse de Fourier représente un signal comme une addition de sinusoides. Le processus inverse de l'analyse de Fourier est la synthèse additive : on peut composer un son en additionnant des composantes périodiques.

Comme nous l'avons déjà vu, il y a un problème de base : les sons naturels sont finis, un phénomène sonore peut être représenté par un signal bornée, comment est-t-il possible de transférer la théorie de Fourier, qui traite des signaux infinis, dans un monde plus « réel » ? Comment utiliser à la place de fonctions périodiques infinies des pseudo-sinusoides ? Dans Music V, grâce à l'utilisation des oscillateurs on peut construire des signaux périodiques avec une décroissance en intensité (enveloppe d'amplitude). Nous arrivons comme ça a créé une sinusoides à temps finie. C'est par le même mécanisme que l'on pourrait réaliser un oscillateur électronique réel : on fournis de l'énergie au système (électronique ou mécanique) pour contrôler l'intensité d'oscillation (relais amplificateur).

Avec la superposition de sinusoides décroissantes, nous arrivons a produire des signaux complexes.

Un autre problème se pose : est ce qu'il est possible de changer la phase du signal ? Si le signal commence par un zéro ou par un max dans sa fonction, la perception de l'instrument simulé peut changer fortement ; dans la synthèse additive c'est un problème, on trouve des limites fixées par la théorie mathématique de Fourier.

Les critères d'évaluation des algorithmes

Comme nous l'avons déjà vu, il n'y a pas aujourd'hui une méthode de synthèse privilégié, il est donc important de définir des critères d'évaluation pour chaque méthode afin de nous permettre de choisir en fonction de la situation particulière que l'on traite.

Il y a plusieurs critères qui permettent de juger une méthode de synthèse et donc d'analyser ses problèmes et ses faiblesses.

²⁷²⁵ En utilisant des son échantillonnées c'est encore notre oreille que doit nous guider dans l'évaluation. Des fois les son échantillonnées sont très différents de la réalité : par exemple dans le cas du piano, les études de l'IRCAM ont montré comme le son produit par une touche fait vibrer par sympathie toutes les autres cordes produisant un son complexe que l'oreille reconnait comme le timbre du piano (par confrontation avec le stockage d'expériences mnésiques de notre passée). Un piano avec des sons échantillonnées n'aura pas cette résonance de fond et le son même si correctement construit semblera artificiel, plat.

Complétude

La *complétude* nous informe si un processus de synthèse est capable de reproduire un grand nombre des sons présents en nature, ou si sa portée d'action est limitée à quelque cas particulier. C'est un paramètre important de jugement parce qu'il rend compte si la méthode est versatile et souple dans les différents domaines sonores.

Ergonomie

Le principe d'*ergonomie* lie la méthode ou le système choisi avec le résultat produit. Un instrument ou un système ergonomique est un système simple à traiter de telle manière que l'utilisateur sait déjà ce que il va produire pour chaque opération que applique sur l'instrument.

Générativité

Si l'on analyse la synthèse additive d'un point de vue économique on s'aperçoit que pour produire un son on nécessite de 44.100 échantillons par seconde pour chaque IN (donc deux) de chaque Oscillateur MusicV pour obtenir un seul OUT.

On définit un autre paramètre de *générativité* comme le rapport Information sortie / Information entrée ; avec un calcul rapide on voit que dans le cas de la synthèse additive, nous avons une *générativité* = $\frac{1}{2}$.

Pour avoir un système efficace on veut atteindre au moins un niveau de *générativité* égale ou supérieure que 1.

Sons réel ou son artificiels ?

Tous les signaux ne sont pas tous intéressants à produire (choix esthétique), pour pouvoir juger sur une base commune de complétude, on choisit de réussir développer des sons réels, naturels : les sons des instruments de musique. La synthèse des instruments sert évidemment de test parce que jamais nous ne pourrions espérer reproduire un son synthétique qui soit parfaitement identique à celui de notre instrument. L'oreille humaine interne, nous le traiterons dans la section de psycho-acoustique, fonctionne dans une première approche comme un analyseur de spectre, mais c'est un analyseur intelligent que s'interroge sur les causes du bruit, sur sa provenance (on arrive avec seulement deux oreilles à détecter d'où vient un son, à décider s'il est loin ou proche), sur sa nature ; il est difficile de tromper l'oreille humaine et même si la technologie atteindra des niveaux très hautes, il y aura quand même quelque chose qui nous communiquera que le son n'est pas réel : un bruit, une artificialité que l'on ne peut éliminer. Même inconsciemment notre oreille est capable de détecter les caractéristiques les plus fines du son et de s'apercevoir s'il est réel ou pas (test de psycho-acoustique)²⁸.

Instruments et modes de vibration

En analysant les caractéristiques physiques des instruments musicaux, on peut s'apercevoir qu'ils sont constitués de matière viscoélastique stable : les composants utilisés dans la construction des instruments sont faits à partir de matériau élastique linéaire ; "linéaire" signifie que la déformation de la structure matérielle est proportionnelle à la force employée pour le jouer. Par conséquence du formalisme mathématique des systèmes élastiques linéaires on peut voir un instrument comme une superposition des modes vibratoires : la fréquence de vibration que l'on mesure est alors liée au rapport entre inertie et élasticité, c'est une propriété stable du système physique étudié.

Les modes de vibrations caractérisent chaque instruments et influencent son émission sonore. Un instrument peut donc être analysé comme une superposition de modes (en analysant sa

²⁸ Un des problèmes technologiques fondamentales et peut être pas résoluble est celui de la diffusion. Même si un son est synthétisé dans une façon parfaite doit être émis par des enceintes : émettre des son par une membrane ne peut pas rendre le phénomène d'un son émis par du bois par exemple, il y a une limite liée à la matérialité des objets en jeu, même avec les plus compliquées effets de spatialisation on ne peut pas rendre compte de la nature différente de la matière.

Mais le problème par l'instant est encore avant : la numérisation, l'oreille humaine ne se laisse pas trahir même par le meilleur son de synthèse.

conformation géométrique) avec certaines propriétés modales (en analysant d'un point de vue plus physique). À l'aide d'études physiques, nous pouvons détecter la fréquence de résonance, l'amplitude et la phase de chaque mode et donc arriver à reproduire les caractéristiques essentielles de chaque instrument musicale.

La synthèse des sons d'instruments réels est très importante :

1) Comme critère de complétude : il est impossible de vérifier la complétude d'une méthode sur tous les sons réels, il faut donc se concentrer sur un petit ensemble de sons que l'on connaît bien : les sons des instruments musicaux. Pour les simuler il faut en analyser leurs propriétés spectrales au cours du temps : la variation dynamique des fréquences et des amplitudes des modes vibratoires caractéristiques.

2) On peut synthétiser mais aussi analyser les sons réels par ce moyen : comprendre quels sont les composantes essentielles pour un son de trompette ou de violon, peut nous aider aussi à définir comment fonctionne notre oreille, la sensibilité qu'elle arrive à atteindre dans l'analyse sonore, ce que veut vraiment dire « entendre ».

Analyseur de spectre

Le premier analyseur de spectre inventé dans l'histoire humaine est l'analyseur de spectre de Helmholtz : il est constitué d'un ensemble de lames de différentes longueurs et épaisseurs qui entrent en vibration selon des fréquences caractéristiques de vibration f_i et avec une amplitude A_i .

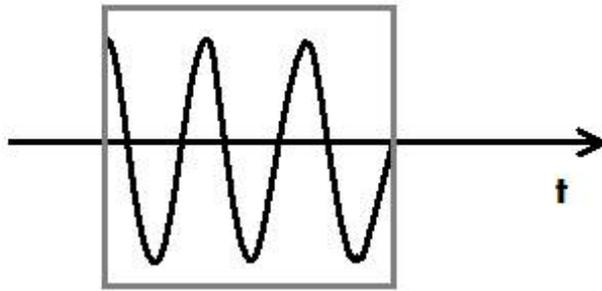
Produisant un phénomène sonore, un certain nombre de lames entrent en vibration par sympathie : leurs fréquences caractéristiques de vibration sont les composantes fréquentielles du phénomène sonore original : grâce à leurs fréquences f_i et aux amplitudes A_i de vibration on peut reconstruire le spectre du signal original.

En réalité on ne peut pas parler du « spectre » pour deux raisons : on opère une discrétisation, même si l'on prend un grand nombre de lames notre sensibilité en fréquence sera toujours finie et donc introduit une erreur, mais deuxièmement, comme nous l'avons déjà remarqué plusieurs fois, dans le formalisme de Fourier, les sinusoides qui déterminent les spectres sont infinies, dans la vie réelle elles décroissent. En plus, les spectres changent au cours du temps selon différentes variables : un spectre est en général dynamique dans le temps; on peut alors penser que l'on réalise une analyse de Fourier valide sur un intervalle de temps fini : un spectre temporel ! Une analyse de Fourier à court terme par le moyen d'une fenêtre temporelle, à travers laquelle l'amplitude et la fréquence de nos composantes sonores ne varient pas trop. On peut penser alors à les remplacer par un morceau d'une fonction infinie.

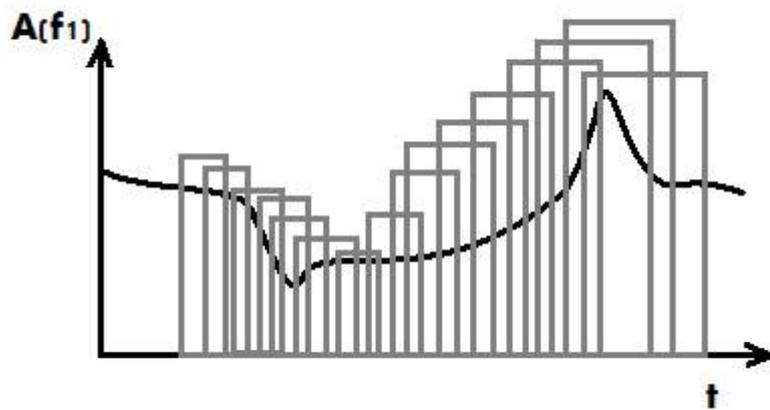
$$S^*(t) = S(t) \cdot p_d(t)$$

On peut choisir la forme de porte que l'on trouve la plus performante pour chaque cas.

Pour chaque composante on obtiendra un signal filtré par la fonction porte :

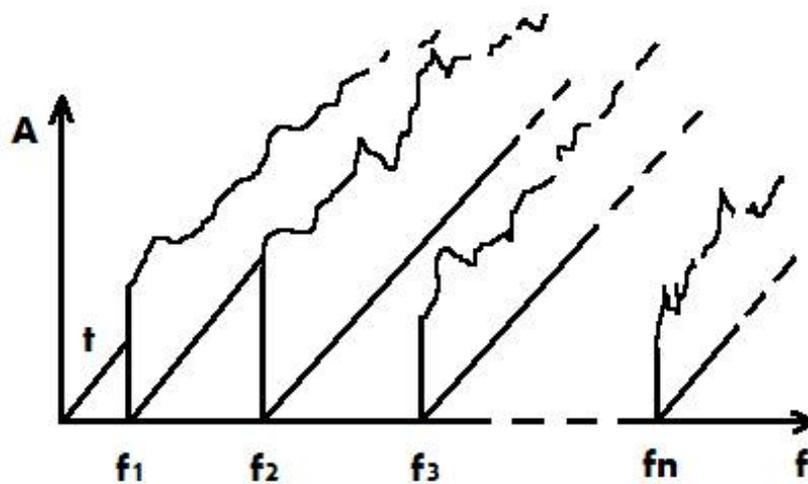


Mais l'amplitude de chaque composante varie dans le temps :



La fenêtre glisse au cours du temps en mesurant pour différents instants l'amplitude A de chaque composante f_i du signal.

On peut représenter sur un graphique tridimensionnel l'évolution du spectre dans le temps :

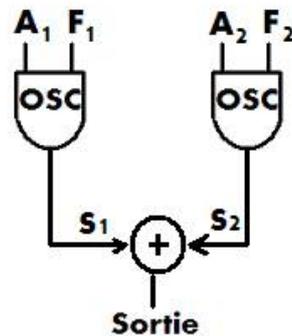


Ou chaque composante a une variation d'amplitude au cours du temps.

En général, les fréquences elles-même peuvent changer au cours du temps. Le système devient complexe : deux variables dépendantes du temps : fréquence et amplitude, la représentation d'un spectre semble être difficile dans ce conditions. Un système semble résoudre très bien le problème : le suiveur hétérodyne ; il divise le spectre des fréquences à analyser (par exemple de 20 Hz a 20 KHz) en intervalles (plus petits et nombreux pour avoir une fidélité supérieure) et cherche par chacun le maximum d'amplitude au cours du temps en stockant au même temps la valeur de fréquence relative : on a de cette manière les composantes et leur variation dans le temps. A partir d'un signal on produit des oscillateurs de fréquence et d'amplitude variables au cours du temps.

La synthèse additive

La synthèse additive produit le phénomène inverse : on imagine de « lames » électroniques vibrante selon des valeur fréquentielles et d'amplitudes choisies et par leur addition dynamique on crée un phénomène sonore en sortie.



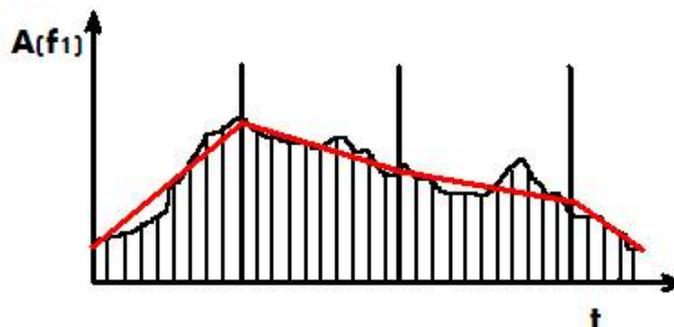
La synthèse additive est alors inspiré par l'analyse de Fourier mais il faut la transférer au niveau réel, dynamique dans le temps. On peut même penser par la synthèse additive à varier les caractéristiques des vibrations modales très rapidement, ou de mettre des paramètres qui sortent de la réalité instrumentale, de façon à créer des exemples d'instruments semi réels qui ne suivent pas l'analyse modale naturelle.

Expérience de J.C.Risset

Dans ces expériences on essaye d'évaluer l'importance perceptive de chaque caractéristique sonore introduite par MusicV dans la synthèse des sons d'instruments musicaux réels: on produit des sons de synthèse puis on enlève progressivement des composantes, on teste les sons obtenus avec un ensemble de sujets qui doivent dire s'ils perçoivent les sons encore comme réels ou non.

L'instrument choisi est le violon, dans le cas d'une générativité proche à 1, les sujets sont incapables de détecter une différence entre son de synthèse et sons réels.

On essaye d'augmenter la générativité et donc de réduire le nombre d'échantillons et donc la qualité de notre son:



A la place de prendre 44.100 échantillons par second on crée 5 ou 6 intervalles que l'on interpole linéairement : on introduit une approximation dans la courbe de réponse fréquentielle ; les expériences montrent que même dans ce cas les sujets ne s'aperçoivent d'aucune différence entre sons réels et sons de synthèse.

Le pas ultime est de rendre fixe un des paramètres: on choisit de rendre la fréquence indépendante du temps, l'amplitude continue à être une fonction du temps. La générativité augmente donc d'un facteur deux parce que on a la moitié d'entrées dans l'oscillateur Music V.

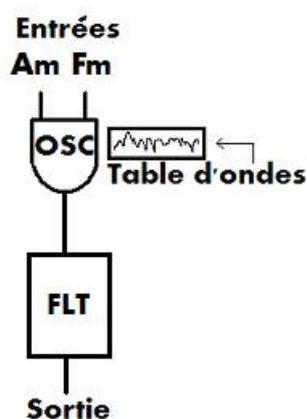
L'expérience faillit pour le violon (le vibrato ne peut pas être représenté) ; la même expérience sur un piano (qui ne permet pas aux joueur de changer la fréquence de ses touches pendant l'exécution) montre aucune différence : les sujets reconnaissent un son réel du piano ;on peut conclure que dans la synthèse d'instruments à « fréquences discrètes » on peut économiser des données considérant que des fréquences ont des valeur fixées.

Le dernière partie de l'expérience, un peu violente, consiste à éliminer totalement les parties transitoires du son : attaque et coda. Les tests montrent que les sujets s'apercevant bien de la différence entre le son synthétique et le son réel : les phénomènes transitoires sont essentielles pour tous les instruments et on doit les prendre fortement en compte pour une synthèse aussi fine que les sons réel²⁹.

Synthèse soustractive

On peut développer les concepts de synthèse soustractive encore à partir du formalisme MusicV l'idée est simple: partant d'un son riche on peut enlever des composantes et obtenir idéalement n'importe de quel son. Dans l'oscillateur Music V, il suffit d'insérer une table d'onde avec un spectre riche (par exemple un bruit blanc) et avec un filtre numérique soustraire des composantes.

Il est possible d'obtenir des sons très intéressants si à la place d'un oscillateur produisant un son périodique, on insère un générateur de signaux quasi aléatoires : un algorithme qui produise du bruit.



²⁹ Un autre expériment a été tenté pour voir les limites entre l'avant dernière partie et la dernière : on a défini pour chaque courbe un centre de gravité qui règle tous les paramètres de la courbe. Même cette expérience faillit : les tests montrent une reconnaissance facile des sujets entre sons artificiels et réels.

Les filtres

Un filtre est un composant logique (à l'origine électronique) qui permet de traiter un signal à l'entrée, le modifier par une fonction dite *de transfert* et d'en sortir un signal final. On peut classer différents types de filtres selon leur action sur le signal d'entrée :

- *filtres causaux* : il ne se passe rien à la sortie tant qu'il n'y a pas d'entrée,
- *filtres linéaires* : filtres qui conservent les propriétés d'addition et de multiplication des signaux.
- *filtres passe bas, passe haut, et passe bande* qui sélectionnent les fréquences de sortie d'un signal.

Pour caractériser le comportement d'un filtre on utilise des signaux impulsifs, que l'on peut décrire mathématiquement comme des deltas de Dirac et observer les signaux en sortie. Le rapport entre signal de sortie et signal d'entrée nous donne la réponse impulsionnelle du filtre : $h(t)$.

La Transformée de Fourier est un moyen très utile pour passer du domaine temporel au domaine fréquentiel. La transformée de Fourier de la réponse impulsionnelle d'un filtre nous donne la fonction de transfert $H(z)$. Cette fonction rend compte, à la sortie du filtre, du changement en amplitude du signal initial $Y(z) = X(z) \cdot H(z)$, de la réponse en fréquence et du déphasage introduit par le filtre.

La fonction de transfert d'un filtre peut être écrite comme un rapport de deux polynômes qui introduisent des pôles et des zéros dans la fonction du filtre. En analysant pôles et zéros on détermine l'intervalle de validité du filtre et son comportement.

La synthèse soustractive est très utile pour synthétiser les instruments réels caractérisés par un système de filtrage : le violon, par exemple, ainsi que tous les instruments à cordes rentrent dans cette famille. La corde a une grande émission de fréquences qui sont filtrées par la caisse de résonance. La voix fonctionne de la même façon : les cordes vocales produisent un spectre très riche. C'est la gorge qui filtre la plupart des harmoniques produites pour déterminer le timbre final de la voix.

La synthèse soustractive ne présente pas une grande complétude, même si théoriquement on pourrait penser partir d'un bruit blanc et par filtrage pouvoir obtenir tous les sons. En réalité on ne peut pas produire tous les filtres possibles à notre besoin.

La généralité ainsi que l'ergonomie ne sont pas mauvaises : les paramètres sont simples.

La Synthèse par Distorsion non linéaire (1979), CNRS – LMA de Marseille.

On parle en général de distorsion quand un amplificateur produit une saturation. Le signal en entrée dépasse le seuil et la sortie est coupée. L'effet, considéré en général comme désagréable, produit un enrichissement fréquentiel (des ondes carrées sont riches en fréquences !).

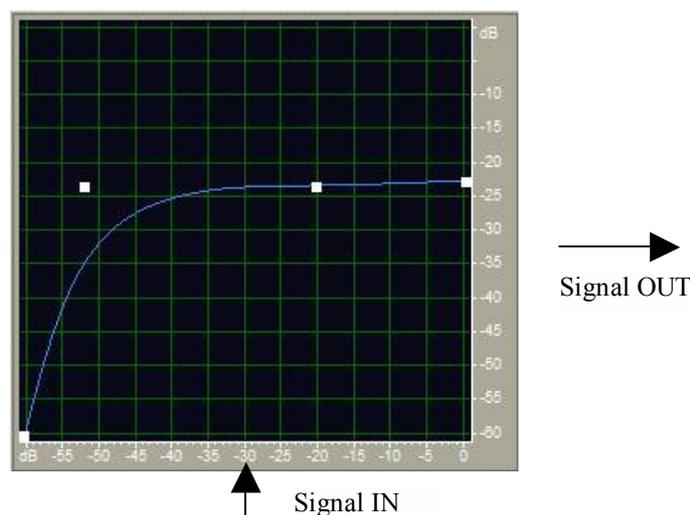


Figure 2 – Exemple de courbe typique de distorsion : l'amplificateur à lampe

Les courbes de distorsion possèdent une espèce de linéarité : on utilise deux courbes d'entrée In1 et In2 et on les distord par la même caractéristique de distorsion produisant Out_{DIST 1} et Out_{DIST 2}. Si on appelle In_{TOT} la somme de In1 et In2 on peut vérifier que Out_{DIST TOT} sortie de In_{TOT} est la somme de Out_{DIST 1} et Out_{DIST 2} :

$$D(a+b)=D(a)+D(b)$$

Les chercheurs de Marseille ont utilisé pour leurs fonctions de distorsion des séries des polynômes de Chebychev. Ces polynômes ont des caractéristiques particulières :

$$T_k(\cos(\alpha))= \cos (k \alpha) , \text{ ou } T_k \text{ est le polynôme de Chebychev d'ordre } k$$

On remarque dans cette formule qu'une fonction périodique est transformée par le polynôme en une fonction périodique de période rétrécie de k fois. On a donc une multiplication de toutes les fréquences du signal initial par le constant k.

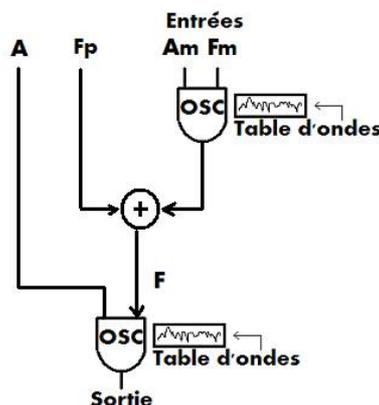
En utilisant une table de distorsion décrite par la somme de N polynômes de Chebychev on obtiendra un spectre très riche constitué des multiplications des fréquences du spectre initial par chaque degré des polynômes utilisés.

L'avantage d'une telle approche est évident. Une grande générativité car à partir d'un spectre simple on obtient de nombreuses multiplications qui aboutissent à un spectre très complexe et riche. C'est donc très économique, le problème vient de l'ergonomie : le calcul des effets attendus pour une table de distorsion est compliqué.

L'autre problème est la complétude : on ne peut pas faire de la modulation d'amplitude à l'entrée en espérant conserver la même forme d'amplitude en sortie. Tout est changé car le phénomène n'est pas linéaire. Il est donc difficile de prévoir ce que va faire ce genre de filtre. On ne peut donc pas imaginer produire n'importe quel son avec cette approche, mais seulement exploiter sa grande richesse.

La synthèse FM, John Chowning (CCRMA Stanford)

Pour appliquer cette méthode nous avons de 4 paramètres de contrôle et d'une table d'onde. Dans cette approche de synthèse on exploite le signal produit à partir de la modulation d'une onde dite porteuse par un signal modulant. Le signal est utilisé comme fréquence d'entrée d'un oscillateur avec une valeur fixe d'amplitude. La table d'ondes subit alors une modulation fréquentielle périodique qui peut, pour des petites valeurs, produire un vibrato comme dans le jeu du violon. Mais on peut aussi créer des signaux beaucoup plus extrêmes pour des grandes valeurs de variation fréquentielle³⁰.



³⁰ Dans la modulation de fréquence Music V on peut utiliser des valeurs quelconques de fréquence, par exemple 200 Hz comme dans la modulation FM. Mêmes les fréquences négatives qui n'ont pas une signification physique sont traitées en introduisant un déphasage π .

La synthèse FM est caractérisée par une grande efficacité. On produit des spectres très variés avec très peu de variation de paramètres. Le fait d'avoir seulement quatre paramètres et de pouvoir balayer le spectre entier des fréquences rend le réglage des paramètres très délicat. D'un changement minimal peut modifier énormément le spectre final. Le problème est donc ergonomique. En effet, la théorie mathématique qui est à la base de la synthèse FM dépend des fonctions de Bessel, leur calcul demandant du temps et n'étant pas facile. Il est donc évident que sans l'outil mathématique il est impossible de prévoir à l'avance ce que le système peut produire.

Il est intéressant de remarquer ici que le synthétiseur DX-7 Yamaha fut le premier à mettre en commercialement en oeuvre le concept de synthèse FM. La résolution du problème ergonomique a été la création de patches pré-programmés que l'utilisateur pouvait choisir. Cette simple approche conduit à une perte du potentiel de la méthode. En effet, prendre des patches pré-crés ne permet pas d'exploiter toute la multitude de timbres différents que cet approche peut offrir. On considère aussi qu'en théorie on peut fournir à l'onde porteuse n'importe quelle fréquence de modulation et pas forcément une note appartenant au spectre audible. En définitive on sait que le phénomène commercialisé n'est pas le même qu'au départ, comme cela arrive souvent.

Synthèse par Fonctions d'Ondes Formantiques (FOF), (1980) Xavier Rodet (IRCAM)

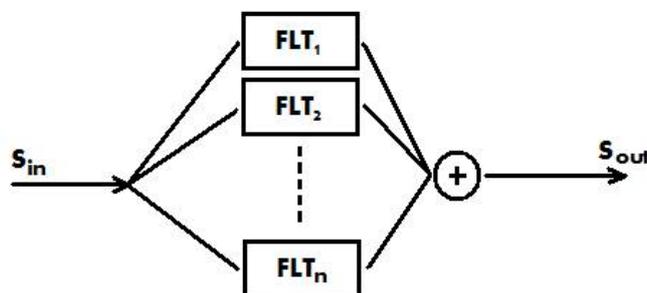
Le but de cette approche est de traiter tous les instruments comme des voix humaines, et étendre ainsi à tous les instruments le formalisme que l'on pourrait utiliser pour la synthèse vocale.

L'émission sonore vocale humaine peut être représenté par une source, dont le spectre est très riche en fréquences (comme c'est le cas de nos cordes vocales), ainsi que par un système de filtrage (opéré par la gorge, les lèvres, la bouche) qui sélectionne les fréquences nécessaires.

Les études sur les courbes liées aux signaux vocaux ont déterminé des formes presque constantes sur tous les sujets normaux dans la prononciation des lettres. En particulier les voyelles qui présentent des allures typiques (signature formantique).

Le méthode est caractérisée par une table d'impulsions à l'entrée du circuit logique et par plusieurs filtres passe bande en parallèle, chacun avec une réponse impulsionnelle caractéristique simple, qui reproduit une sinusoïde amortie (comme souvent dans les signaux vocaux relatifs aux voyelles) de fréquence et d'amplitude définie. Chaque filtre sert donc à reproduire une partie simple (un pic) d'un signal vocal complexe. L'addition de tous les sorties des filtres détermine la forme finale du signal : la superposition des pics typiques de la voix permet de synthétiser le signal d'une voix humaine.

Cette approche fonctionne bien pour les instruments similaires à la voix, comme par exemple le violon ou les instruments à corde, mais ce n'est pas optimal pour les instruments très différents, par exemple les percussions. De plus même l'attaque d'un archet de violon n'a pas un bon rendu, l'attaque est toujours lente et douce comme celui d'une voix.



6 Éléments de psycho acoustique

Dans le chapitre précédent on a vu que pour synthétiser les sons d'un instrument musical réel on ne doit pas négliger l'évolution spectrale dans le temps. L'évolution temporelle des amplitudes et des

fréquences est essentielle pour la reconnaissance sonore de l'oreille humaine³¹. On a vu aussi comment d'un même environnement très puissant comme Music V on peut rendre compte des variations temporelles de fréquence et d'amplitude par le moyen d'un Oscillateur. Quelles caractéristiques doit donc avoir un modèle de synthèse pour représenter des sons réels ? Et quelles sont les caractéristiques sonores des signaux synthétisés qui sont fondamentales dans notre processus perceptif ? Enfin comment peut on juger si un son de synthèse semble réel ou pas ?

Un modèle de synthèse doit avoir des spécifications structurelles (qualitatives), comment connecter les différentes parties (par exemple les différents oscillateurs Music V), et des spécifications paramétriques (quantitatives). A partir d'un phénomène sonore d'origine, par une analyse profonde des caractéristiques sonores et leur évolution temporelle, on peut reconstruire un signal (définition des connexions et des paramètres du modèle). La synthèse est le phénomène inverse de l'analyse : le signal obtenu est transformé en phénomène d'arrivée et comparé avec le signal d'origine. L'oreille experte du musicien est le meilleur juge dans cette phase. En général la confrontation entre les deux phénomènes : naturel et synthétique montre une partie commune mais aussi des caractéristiques différentes.

Il est normal à ce moment de création, une fois acquise la possibilité de créer des son, de se poser les questions suivantes : Quelles relations s'établissent entre le phénomène d'origine et le résultat (le phénomène synthétisé) ?

Comment peut on réaliser cette création rigoureusement ?

Quels effets un changement des paramètres de synthèse a-t-il sur le phénomène à la sortie ?

Pour répondre à ces questions, on utilise une stratégie d'exploration des paramètres pour déterminer leurs influences sur l'apprentissage. Notre système sera composé par un son de synthèse et par un sujet qui écoute les sons produits.

On se retrouve avec deux types de paramètres dans notre système : objectif (de l'objet lui même) et subjectifs/perceptifs (qui s'intéresse au phénomène perçu). La psycho-acoustique est la science qui étudie le lien entre ces deux systèmes.

Au début de cette science, l'analyse de Fourier semblait être l'outil idéal pour pouvoir traiter les relations entre phénomènes sonores et leurs perceptions. La psycho-acoustique était conçue comme lien entre phénomène perçu et paramètres de l'analyse de Fourier (fréquence, amplitude, phase).

Mais on s'est aperçu récemment que la théorie de Fourier n'est pas suffisante pour décrire de nombreux résultats expérimentaux.

Par exemple la notion de hauteur et de fréquence n'est pas la même : la première est une image mentale précise et stable du son (un LA ou un SI), la deuxième est une caractéristique physique du son. Les deux sont liées par une fonction logarithmique :

$$H = \text{Log}_2 F$$

Cela veut dire que l'oreille humaine perçoit les changements fréquentiels des sons du monde extérieur comme une fonction Log_2 .

Stevens a conduit des expériences pour vérifier cette loi, et a trouvé que la fonction n'est pas exactement logarithmique mais plus complexe. De plus elle change selon les individus, pour chaque age, selon le contexte... Ces résultats ont des grandes conséquences car elles nous montrent que chacun peut percevoir un même phénomène sonore de différentes façons. On arrive à concevoir une relativité perceptive sonore.

La perception dépend aussi de l'intensité sonore. Stevens montre qu'il y a des courbes isofréquentielles (qui représentent le niveau de perception des différentes fréquences selon la

³¹ On sait que la phase n'est pas perçue par l'oreille humaine. Tous les sons passant à travers les objets ou dans l'air, selon les différentes densités aériennes, subissent des déphasages incontrôlables. Si notre oreilles étaient sensibles à ce type de déphasage, l'information acquise serait très compliquée et discerner les différentes parties d'un même son parmi un ensemble de bruits incohérents serait un travail très dur. L'oreille humaine ne tient pas compte de la phase, tous les processus de perception sonore sont ainsi simplifiés.

variation de l'intensité sonore) et des courbes d'isotonie (qui représentent le même niveau de sonie pour les différentes fréquences sonores). Les différentes fréquences d'un son ne sont pas perçues au même niveau d'intensité par notre oreille : les fréquences moyennes sont perçues plus facilement que les autres, ce qui est justifié par le développement de la communication. La naissance du langage qui est véhiculée en grande partie par les fréquences moyennes peut avoir influencé notre perception sonore humaine qui s'est concentrée sur les fréquences indispensables pour communiquer. On voit donc que l'analyse des signaux ne nous informe pas sur « comment les sons sont perçus ».

Un test fondamental pour montrer l'insuffisance de la Théorie de Fourier a été réalisé par Schaeffer. Un son de piano reproduit à l'inverse n'est plus perçu comme un piano par une oreille humaine, pourtant le spectre sonore doit être pareil. Cela montre que le spectre est insuffisant pour percevoir et reconnaître des sons, mais, comme on l'a déjà dit, les phénomènes transitoires comme l'attaque, la coda et les évolutions du son au cours du temps sont indispensables.

J.C. Risset a exécuté d'intéressantes expériences pour montrer comment notre concept de hauteur est incomplet. Dans ses *glissandi paradoxaux*, un son semble descendre infiniment en hauteur mais on s'aperçoit paradoxalement qu'il ne bouge pas, après plusieurs minutes le son sur la bande est toujours le même. Risset montre comment on devrait considérer deux types de hauteur. Une tonale, par exemple toutes les notes Do à différentes octaves du piano, cyclique sur toute la tessiture de l'instrument. Et une absolue, liée aux concepts spectraux, par exemple un Do et un Ré. En créant deux mouvements contraires de hauteur tonale et de hauteur absolue selon une fonction d'enveloppe, on a un paramètre qui descend vers les basses fréquences et un autre qui monte vers les hautes, cela donne une émission sonore résultante qui ne bouge pas. Ces expériences montrent comment les concepts de hauteur employés d'habitude sont en réalité très délicats et nécessitent une réflexion profonde.

Un autre exemple qui décrit bien l'insuffisance d'une approche purement spectrale dans la synthèse sonore est celui de la synthèse d'un son de cloche. En écoutant attentivement le son d'une cloche la première remarque que l'on pourrait faire est qu'il existe une mise en avant des moyennes fréquences et que les basses fréquences durent très longtemps après que les autres ont déjà disparues. Dans le processus de synthèse on peut donc mettre une courbe d'enveloppe pour mettre en avant les moyennes fréquences à l'instant de percussion et une autre qui rend plus longue la coda des basses fréquences. Mais le phénomène le plus caractérisant dans le cas d'un son de cloche est le battement entre les différentes fréquences : la fabrication de la cloche n'est jamais parfaite et laisse forcément des impuretés matérielles à son intérieur, ce qui produit des doublements des modes vibratoires. Pour simuler cette propriété physique, on pourrait introduire des doubles des oscillateurs du début et les décaler pour produire une interférence. Le modèle fonctionne bien, l'approche classique de Fourier ne nous aurait pas aidé dans notre analyse et c'est l'évolution temporelle des phénomènes sonores qui rend réel les sons synthétisés ou pas.

L'approche de la psychologie cognitive appliquée au monde des phénomènes sonores nous permet de comprendre que l'oreille humaine n'est pas un simple analyseur de spectre, comme on le pensait avant les expériences de psycho acoustique.

L'oreille humaine est un organe que l'on a développé pour percevoir l'environnement sonore qui nous entoure. C'est un moyen complexe qui nous permet de comprendre causes, sources, épisodes et c'est beaucoup plus complexe qu'un simple analyseur de flux :

- Elle nous fournit des informations essentielles à notre survie. On n'a pas qu'une représentation Bottom-Up, acquisition des informations sonores de l'extérieur vers le cerveau, mais aussi un phénomène Top-Down : le cerveau analyse et réorganise les informations reçues. La perception nous permet aussi d'identifier dans le signal reçu les caractéristiques qui nous informent sur l'origine du son, sa nature, le sens qu'il a pour nous (danger, félicité). On crée une interaction avec le monde réel extérieur. Par l'aide des perceptions sonores on peut comprendre les règles physiques qui dominent les mondes

extérieurs et on peut développer des stratégies pour réussir à déterminer et à fournir une description causale des phénomènes sonores.

- Pour des raisons d'économie du codage cognitif, quand le système auditif trouve le nombre minimum de sources pour identifier la scène et comprendre l'environnement dans lequel il se trouve, il arrête de se concentrer et peut analyser d'autres choses.

Des tests très simples peuvent nous démontrer les affirmations que l'on vient de faire :

1) Le système cognitif est influencé par le contexte (des cylindres de dimensions égales dans un contexte de perspective apparaissent différents), au niveau rationnel on ne pourrait pas voir la différence entre un objet grand et loin, et un objet petit et proche. C'est le contexte qui nous guide dans la perception. La même chose arrive dans le monde des phénomènes auditifs. Un son fort mais loin peut ressembler à un son proche et faible, mais nos oreilles sont très habiles à distinguer les deux cas. Le contexte est d'une importance fondamentale et est bien analysé par notre cerveau avant que l'on arrive à produire une représentation mentale des objets extérieurs.

2) D'autres exemples nous montrent comment des choses fausses ou pas importantes sont effacées automatiquement par notre système cognitif.

Steve Mc Adams a concentré ses études sur les phénomènes sonores de ségrégation et fusion³², c'est-à-dire quel est la limite de séparation temporelle pour que deux phénomènes sonores identiques soient perçus comme distincts par notre système cognitif.

Ses études ont démontré d'abord que même avec une seule oreille on arrive à distinguer des sons produits par différentes sources et on arrive à individualiser les sources séparées spatialement (un phénomène surprenant si on considère qu'une seule oreille véhicule un seul signal).

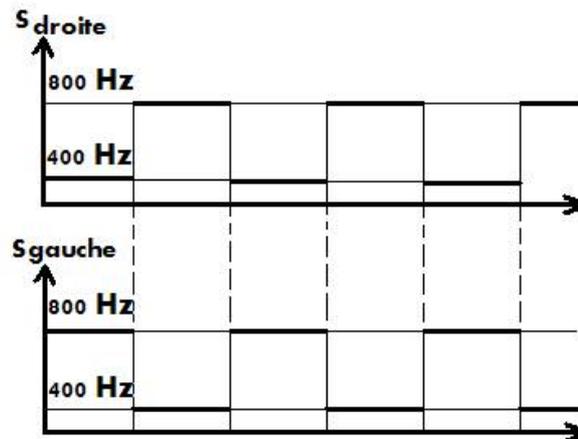
Un exemple qui clarifie les concepts de ségrégation et fusion est le cas de la cloche que l'on a décrit précédemment. Si l'on agrandit le retard temporel entre deux modes vibratoires de la cloche, on atteindra une valeur au delà de laquelle le son est perçu comme provenant de deux cloches séparées. En effet notre expérience sonore n'a jamais perçu une même cloche avec un retard tellement grand.

Si par contre on diminue le retard jusqu'à zéro le son ne semble plus être celui d'une cloche. Ces deux limites doivent être prises en compte dans la synthèse de sons de cloche, le bon paramètre de retard doit être entre les deux limites, c'est pour cela qu'il est fondamental de tenir compte des règles psycho-acoustiques dans la synthèse sonore.

Les expériences de fusion expliquent comment l'oreille humaine se refuse à croire que deux objets puissent produire le même phénomène sonore synchrone au cours du temps. La synchronisation entre deux objets perceptifs est pour notre cerveau la première preuve de l'existence d'un seul objet. Les deux objets sont perçus comme deux parties du même objet.

Un dernier exemple très intéressant dans lequel les concepts expliqués précédemment jouent un rôle fondamental dans la perception est appelé Paradoxe Dianna Deutsch (1980). Un sujet doté d'un casque écoute deux sons différents mais synchrones dans l'oreille droite et dans celle gauche selon le schéma suivant:

³² La fusion était un phénomène déjà employé par Debussy et Berlioz dans leurs compositions orchestrales : plusieurs instruments étaient superposés jouant la même mélodie pour produire l'image d'un seul objet complexe.



Quand l'oreille droite entend l'octave basse l'oreille gauche entend l'octave haute et vice-versa. Au niveau perceptif le sujet entend à gauche un son continu à 400 Hz et à droite un son continu à 800 Hz. Le résultat s'explique par des considérations vu précédemment : le cerveau se refuse de considérer qu'il existe deux objets parfaitement synchrones qui puissent passer d'une oreille à l'autre à une vitesse tellement élevée qui produit le phénomène d'alternance, et préfère percevoir la réalité comme constituée de deux objets immobiles et distincts, l'un placé à droite et l'autre à gauche.

Conclusions

Même si l'oreille humaine semble travailler comme un analyseur de spectre son comportement est beaucoup plus complexe. Notre image mentale fonctionne aussi selon notre représentation mnésique. La perception sonore est caractérisée en général par deux flux : un qui vient de l'extérieur (signal sonore) et l'autre de l'intérieur, une image mentale (qui dépend de nos expériences passées).

Le processus cognitif n'est pas donc une simple analyse des signaux sonores éternels mais c'est un profond bagage humain que l'on a accumulé pendant des siècles, c'est une interaction avec le monde réel³³. On s'aperçoit donc que l'analyse des signaux risque d'être insuffisante pour décrire nos perceptions et que le contexte ne doit pas être négligé pour une analyse plus fiable.

De plus il faut considérer, comme on l'a déjà dit au début de ce traité, que notre perception est multisensorielle. On devrait analyser l'ensemble des signaux dans tous les domaines sensibles humains et tenir compte de leur interaction dans l'étude des processus de cognition. La relation action-perception est un autre niveau de complexité qui s'ajoute. On a vu que chaque perception est à sa base active et que chaque action est perceptive. On doit, de plus, tenir compte du fait que chaque signal est unidirectionnel, mais que chaque sujet est un individu ayant une interaction avec le monde extérieur (pluridirectionnel).

On comprend que pour synthétiser des phénomènes sonores réels en conservant leur « degré de réalité », il ne suffit pas de reproduire simplement des signaux. Ce sont les causes que l'on recherche, la source de chaque signal que l'on écoute. Avec cette approche on peut produire des sons cohérents, que notre oreille pourra reconnaître comme réels. C'est la matérialité que détermine le caractère « réel » des simulations, la simple reconstruction des phénomènes apparents sera toujours perçue comme artificielle, sans une approche matérielle à la base. C'est pour ces raisons que l'on doit se poser d'importantes questions épistémologiques sur les modèles que l'on veut utiliser avant de commencer la synthèse d'un phénomène sensible.

³³ Par exemple on sait que dans la détection des contours d'une figure géométrique le cerveau du sujet en phase cognitive modifie les couleurs à proximité du contour pour percevoir mieux la division entre différents objets.

7 Trois niveaux de modélisation

Qu'est-ce que un modèle? Pourquoi crée t-on des modèles ?

Un modèle sert à donner une représentation d'une partie de la réalité en utilisant des moyens de calcul mathématique. On peut appliquer des actions, sur l'objet artificiel créé, que l'on ne pourrait jamais faire sur l'objet original.

Si on analyse les types de modèles de synthèse décrits dans les pages précédentes on peut regrouper tous les modèles de synthèse sonore en trois groupes logiques selon ce qu'ils cherchent à stocker :

- Modèles phénoménologiques
- Modèles fonctionnels
- Modèles structurels

La musique concrète se base sur la représentation de la réalité par ses phénomènes. L'enregistrement analogique est un moyen de stocker les phénomènes sonores naturels, c'est donc une modélisation phénoménologique. La bande magnétique est l'objet qui représente la réalité sur lequel on peut appliquer nos actions; mais les possibilités de ce modèle sont très limitées. On reste en surface, on ne peut pas descendre en profondeur dans le système sonore (la bande magnétique) et changer les sons qui y sont enregistrés. On peut les découper (pas infiniment non plus), changer leur ordre, mais pas changer leur nature, leurs paramètres physiques.

Avec Music V et tous les systèmes de synthèse sonore que l'on a traités, on ne modélise pas des phénomènes sonores, on les manipule. On crée des fonctions que l'on peut mathématiquement déformer, en changeant leurs paramètres. Les fonctions sont la représentation des signaux relatifs aux sons que l'on entend dans la réalité : on modélise des fonctions. Cela nous permet de manipuler les sons à un niveau très profond mais par cette approche on perd une caractéristique importante que l'on avait encore dans les modèles phénoménologiques : la bidirectionnalité. Dans les modèles fonctionnels on n'a plus d'isomorphie entre la structure de fonctions et les composantes de l'instrument. Le fonctionnement de notre modèle ne représente plus la nature. Cela veut dire que si l'on change un paramètre dans notre modèle cela ne correspond pas à un changement physique précis de l'instrument auquel on se réfère, il n'y a plus de relation directe entre paramètres de synthèse et réalité sonore de l'instrument. On peut donc facilement perdre le sens de notre modèle si les paramètres ne sont pas les bons.

Les modèles source-filtre (par exemple la synthèse FOF) conservent un fonctionnement similaire à celui de la nature, mais s'intéressent plus au phénomène qu'à la cause qui l'a produit. On décrit le son mais pas encore l'objet physique qui le cause.

Les modèles structurels s'intéressent d'abord à donner une représentation structurelle de l'objet qui produit le son et pas du son que l'on veut. L'étude se concentre sur les composantes instrumentales et sur la caractérisation de leurs propriétés pour arriver, à la fin, à la caractérisation de l'ensemble (qui est plus que la simple addition des composants).

8 La synthèse par modèle physique

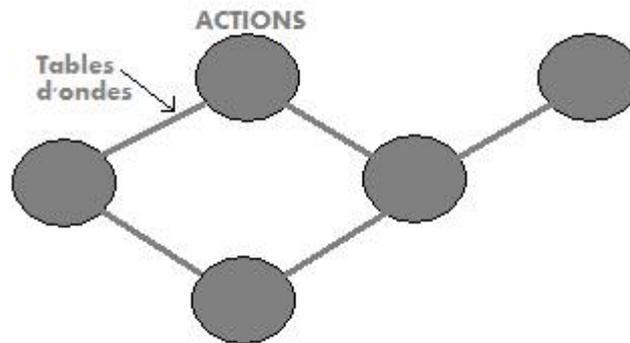
Quand on parle de modèle physique il faut d'abord faire attention : le modèle est physique et non pas la chose représentée. La nature de l'approche est physique (description avec des ressort, ou par les lois de la dynamique newtonienne), mais l'objet peut être un être humain ou un objet irréel d'un univers différents du notre (pas forcément des objets physiques naturels).

Waveguides (1983) by Karplus, Stray and Smith

Le concept à la base de cette approche récente de synthèse sonore est encore une fois la générativité : produire le plus grand nombre d'échantillons avec le minimum de données. Si on imagine pouvoir bouger la table d'onde de Music V, pendant que l'oscillateur qui l'utilise crée des signaux de sortie, on peut produire beaucoup de signaux différents qui changent en temps réel selon le désir de l'utilisateur. Cette approche permet de produire des algorithmes par récurrence et donc des sons périodiques qui représentent bien le son d'une corde percutée (récurrence harmonique). Ce fut le

premier exemple, dans l'histoire humaine, d'algorithme pouvant produire des sons variables en temps réel, un algorithme vif.

Smith implémentait le formalisme de l'idée de Karplus et Stray avec un système de filtrage numérique : un réseaux de lignes à retard avec des modules de traitements :



En analysant le système on peut commencer à voir un objet réel composé de différentes parties « vivantes ».

Synthèse Modale

Une autre approche tente de modéliser la constitution physique des phénomènes sonores : celle de la synthèse modale. Comme on l'a déjà vu, chaque composante d'un instrument a des propriétés modales, qui sont représentées par des modes vibratoires. Dans la synthèse modale (Modalys, IRCAM) l'utilisateur peut choisir la forme et les propriétés vibrantes de son objet.

Le problème de cette approche est le grain de décomposition de l'instrument : on a à faire à des parties trop grandes c'est pourquoi, finalement, on retombe facilement dans le formalisme fonctionnel. Le système est un hybride entre modèle structurel et fonctionnel.

Genesis – Cordis Anima, ACROE

Sujet du TP et de stage.