

Modélisation en musique

Table des matières

1. Différents types de modèles	2
2. Modèles mathématiques	2
a) Formalisation algébrique de la musique	3
i. Historique	3
ii. Quelques outils de formalisme	3
iii. Les structures de groupe	6
iv. Intérêt dans les sciences cognitives	10
b) Quelques exemples d'applications aux sciences cognitives	11
3. Modèle de simulation : LM.....	12
a) Intérêt	12
b) Principe	13
c) Fonctionnement	13
d) Exemple	14
e) Limites	15

1. Différents types de modèles

De nombreux modèles ont été développés dans l'étude de la cognition musicale. Parmi les modèles principaux, on trouve :

- Les modèles d'inspiration linguistique (utilisant des grammaires ou autres éléments dérivés du langage) ;
- Les modèles basés sur les schémas, des ensembles structurés de relations entre les attributs d'un phénomène ;
- Les modèles mathématiques, inspirés par les statistiques, la géométrie, la théorie des ensembles ou la théorie des groupes ;
- Les modèles de simulation, incorporés dans un programme ou mis en œuvre de toute autre façon.

Nous avons choisi ici de ne développer que 2 exemples de modèles : un mathématique et un computationnel.

2. Modèles mathématiques

En 1739, le mathématicien suisse Leonhard Euler publie un essai dans lequel il souhaite expliquer pourquoi la musique apporte du plaisir. Selon lui, l'élément clé est la perfection, qu'il recherche dans les rapports de nombres représentant les accords.

Depuis, l'étude des mathématiques dans la musique s'est développée et est devenue un champ disciplinaire à part entière. L'informatique a été un des catalyseurs de ce nouveau champ de recherche, donnant ainsi naissance à une discipline, la « musicologie computationnelle », ayant pour but d'analyser les œuvres musicales afin d'y déceler des structures mathématiques sous-jacentes.

Ainsi, cette partie a pour but de donner quelques exemples du formalisme de la musique et de son utilisation dans l'analyse des œuvres musicales.

a) Formalisation algébrique de la musique

i. Historique

L'utilisation des méthodes algébriques en musique met en œuvre 3 aspects souvent interdépendants : les aspects théoriques et analytiques, ainsi que ceux d'aide à la composition. Parmi les compositeurs et théoriciens emblématiques du mouvement théorique, on peut en retenir 3 : Milton Babbitt, Iannis Xenakis et Roumain Anatol Vieru. Ils ont découvert le caractère algébrique du tempérament égal, c'est-à-dire que dans la gamme, chaque note est séparée d'un-demi ton de sa voisine, soit une gamme à 12 demi-tons (*do, do#, ré, ré#, mi, fa, fa#...*), ce qui diffère de la gamme diatonique à 7 tons (*do, ré, mi, fa, sol, la, si*).

Plus précisément, ils ont mis en avant la notion mathématique de groupes comme concept unificateur. Cette théorie regroupe des concepts abstraits, mais ces structures seraient en réalité inhérentes à notre cerveau. Cette notion est née d'études sur les racines de polynômes, et va permettre d'aider à analyser la musique. Elle a été découverte par Joseph-Louis Lagrange, mais ne fut utilisée pour la 1^{ère} fois qu'un demi-siècle plus tard par M. Babbitt, qui a observé que le système dodécaphonique est « un groupe de permutations qui est façonné par la structure mathématique de ce modèle ». La musique dodécaphonique diffère de la musique tonale, où une des 7 notes de la gamme diatonique prédomine sur les autres et leur impose une hiérarchie. C'est en 1923 qu'Arnold Schoenberg va proposer la méthode de compositions avec 12 tons, d'où l'échelle dodécaphonique, dans laquelle aucune note n'est prédominante. Une composition dodécaphonique se fonde sur une séquence de 12 sons musicaux distincts, sans répétition, nommée série élémentaire. L'œuvre est une combinaison de cette série et d'autres séries dérivées par des symétries.

ii. Quelques outils de formalisme

Les transformations

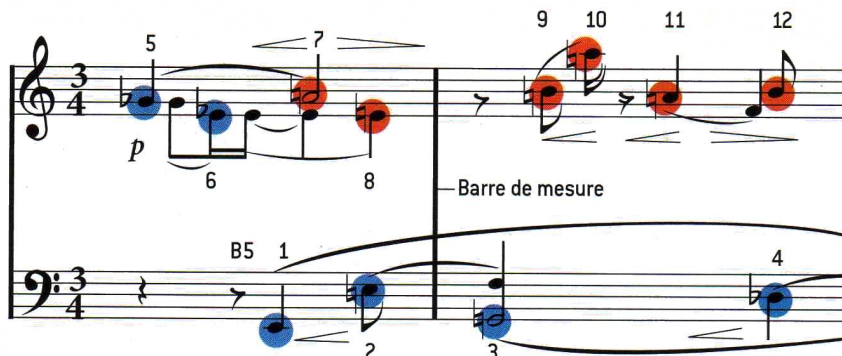
Chaque note peut être représentée par un nombre qui renvoie à une classe d'équivalence modulo 12 : chaque nombre représente lui-même, ainsi que ce nombre additionné d'un multiple de 12. Par exemple, 1 est équivalent à 49 ($=1+4*12$) ou -11 ($=1-12$). L'addition de 2 nombre devient une somme modulo 12 : par exemple, $(3 + 8)_{12} = 11, (5 + 9)_{12} = 2$.

Partant d'une série élémentaire P, on peut définir la notion de symétrie. On appelle ainsi série rétrograde R, la série P jouée à l'envers. On peut aussi définir la série renversée I, dans

laquelle les nombres sont les opposés de ceux de P, modulo 12. On peut alors appliquer ces deux opérations, pour obtenir la série renversée rétrograde RI de P. Enfin, la série transposée de P par k demi-tons est obtenue par l'addition modulo 12 de k à tous les nombres de la série P. De même, on obtient les transposées par k demi-tons d'une série rétrograde $T_k R$, d'une série renversée $T_k I$ et d'une série renversée rétrograde $T_k RI$.

La symétrie axiale généralisée

Cette notion renseigne sur la composition d'une polyphonie sérielle constituée de 2 hexacordes (séries de 6 notes) liées par une telle symétrie (cf. : Figure 1).



2. LA SYMÉTRIE AXIALE GÉNÉRALISÉE

permet d'analyser cet extrait d'une polyphonie sérielle (ci-dessus) à deux voix. Ces deux voix correspondent à deux hexacordes (l'un en bleu, l'autre en rouge) liés par une symétrie axiale généralisée, comme le montre la représentation circulaire (ci-contre).

On peut aussi vérifier que si on segmente l'extrait en deux parties correspondant aux barres de mesure, on obtient également deux hexacordes en rapport de symétrie axiale généralisée.

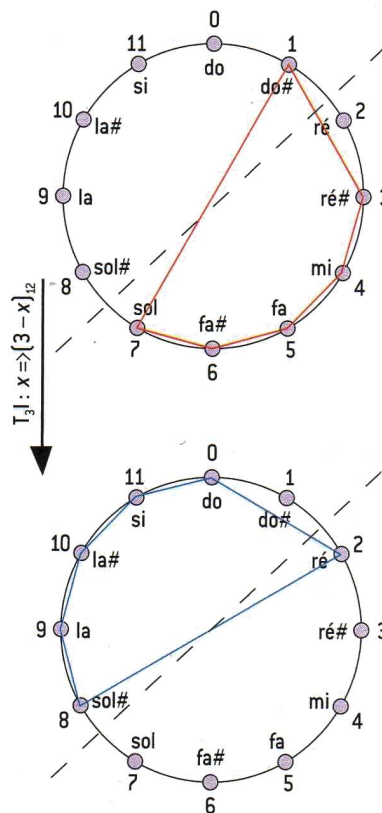
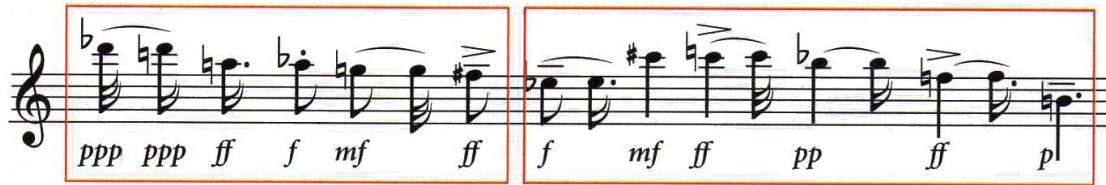


Figure 1 : Symétrie axiale généralisée utilisée dans le sérialisme

Cette notion de symétrie axiale est fondée sur le concept de groupe et est utilisée dans la musique sérielle, mais aussi dans la musique modale, comme par exemple celle du compositeur français Olivier Messiaen (cf. : Figure 2).



3. LA SYMÉTRIE AXIALE GÉNÉRALISÉE n'est pas utilisée uniquement dans la musique sérielle. Ainsi, le compositeur Olivier Messiaen en fait usage entre les deux hexacordes du mode utilisé dans la pièce *Mode des valeurs et d'intensités*, composée en 1950.

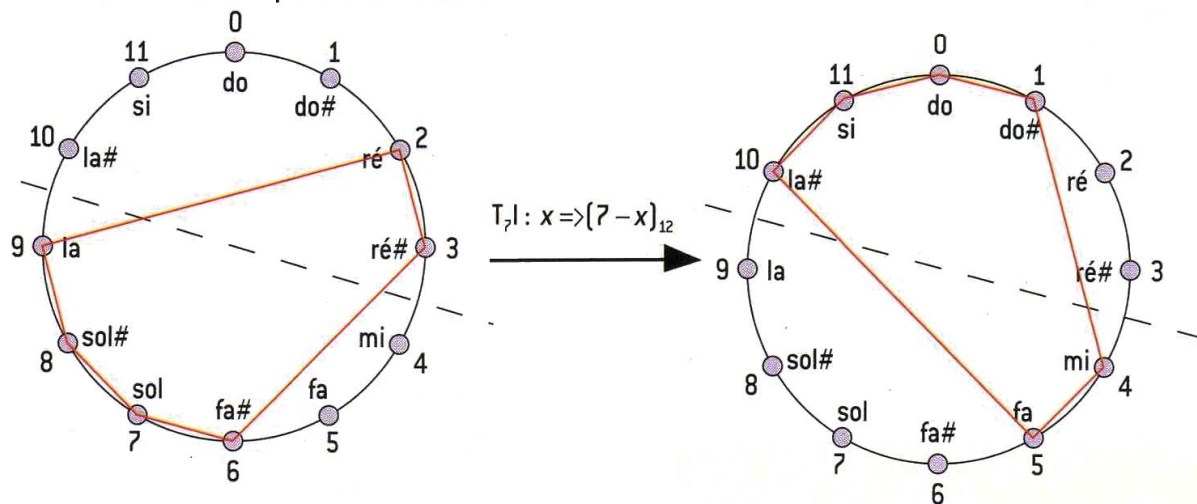


Figure 2 : Symétrie axiale généralisée utilisée dans la musique modale

En musique, le mode d'une gamme est constitué des mêmes notes que la gamme dont il est issu, mais a une sonorité qui lui est propre, caractérisée par une tonique est par les intervalles entre cette tonique et les autres notes. Par exemple, à partir de la gamme de *do* majeur, dont la tonique est *do* (*do*, *ré*, *mi*, *fa*, *sol*, *la* et *si*), avec pour structure intervallique en demi-tons (2, 2, 1, 2, 2, 2, 1), on peut déplacer l'axe tonal sur la deuxième note afin d'obtenir un nouveau mode (en *ré*) ayant une nouvelle structure intervallique (2, 1, 2, 2, 2, 1, 2) en conservant les mêmes notes (*ré*, *mi*, *fa*, *sol*, *la*, *si* et *do*). Dans ce type de composition, on privilégie des notes ou des intervalles au détriment des autres.

iii. Les structures de groupe

Le groupe $\mathbb{Z}/12\mathbb{Z}$

L'ensemble des entiers modulo 12 donne un premier exemple musical de structure de groupe, celui noté $\mathbb{Z}/12\mathbb{Z}$. Il peut être interprété musicalement de plusieurs façons :

- Comme le groupe des intervalles musicaux (avec l'addition modulo 12) ;
- Comme un groupe de transformations, celui engendré par les transpositions. On peut représenter ce groupe par un cercle, correspondant à une octave et divisé en 12 parties.

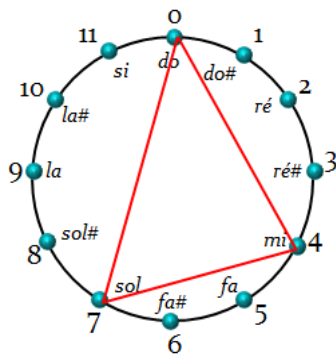


Figure 3 : Exemple sur la gamme tempérée

Sur l'exemple de la figure ci-contre est représentée la gamme tempérée, c'est-à-dire dont toutes les notes sont séparées d'un demi ton. Elle est représentée par un cercle où chaque note est numérotée (le *do* correspondant au 0 dans la tradition américaine, et au 1 pour les Européens). On a ici une structure intervallique, qui correspond aux intervalles séparant les sommets du polygone à n côtés (figure en rouge) correspondant à un ensemble de notes.

Par exemple, la structure intervallique de l'accord majeur (constitué d'une fondamentale, d'une tierce majeur et d'une quinte, tels *do*, *mi*, *sol*) est égale à (4, 3, 5). En effet, 4 demi-tons séparent *do* et *mi*, 3 séparent *mi* et *sol* et 5 séparent *sol* et le *do* supérieur.

Ainsi, l'hypothèse sous-jacente de cette représentation circulaire est qu'elle permet de formaliser tout accord musical.

Le groupe de Klein

M. Babbitt va proposer d'exprimer des transformations de la structure du groupe cyclique $\mathbb{Z}/12\mathbb{Z}$. Effectivement, il remarque que l'on peut interpréter les 4 formes d'une série dodécaphonique (la série elle-même, la série rétrograde R, inversée I et renversée rétrograde RI) comme les 4 transformations suivantes :

$$P: (a, b) \Rightarrow (a, b)$$

$$I: (a, b) \Rightarrow (a, (12 - b)_{12})$$

$$R: (a, b) \Rightarrow (11 - a, b)$$

$$IR: (a, b) \Rightarrow (11 - a, (12 - b)_{12})$$

Ici, la série dodécaphonique P est représentée par une suite de couple (a, b) , a indiquant la position de la note dans la série, et b , la note (par rapport à l'origine 0).

Ces 4 groupes de transformations d'une série dodécaphonique et, plus généralement, d'un profil mélodique, constituent les éléments d'une structure algébrique nommée groupe de Klein de 4 éléments. Il tient compte de toutes les transformations de la musique, alors que le groupe cyclique $\mathbb{Z}/12\mathbb{Z}$ ne rend compte que des transpositions.

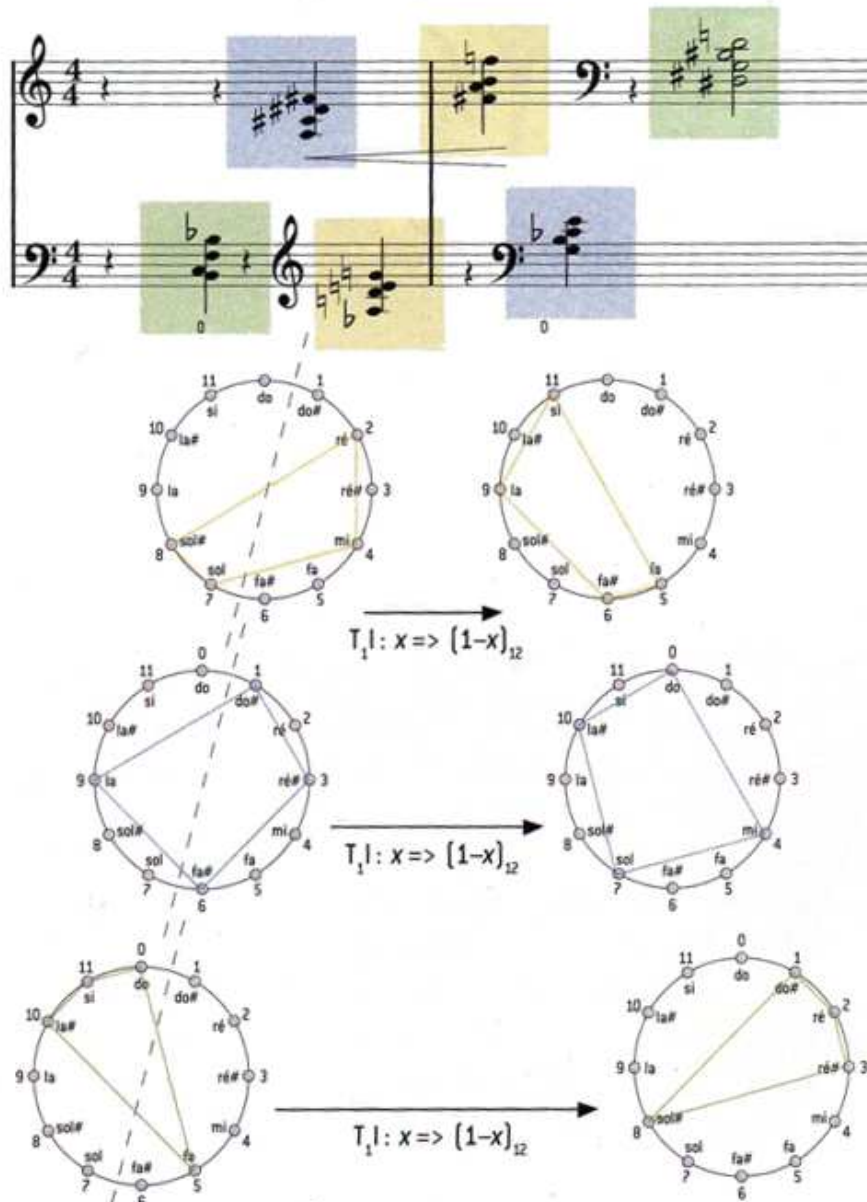
Le groupe diédral

Il s'agit de l'ensemble de toutes les compositions (au sens mathématique) des transpositions et des inversions. En d'autres termes, c'est le groupe des symétries axiales généralisées. Le nom diédral signifie « à 2 faces », et indique que d'un point de vue géométrique, ce groupe renvoie au groupe des symétries d'un polygone régulier de n côtés dans le plan.

Ces symétries sont de 2 types : rotations et réflexions (ou miroirs par rapport à un axe). Musicalement, les rotations correspondent aux transpositions, et les réflexions sont des inversions par rapport soit à une note choisie comme pôle, soit à une note « imaginaire » qui se trouve entre 2 notes séparées d'un demi-ton, quand l'axe de symétrie ne passe pas par une note du cercle chromatique.

Ce type de groupe peut aussi être utilisé dans les musiques atonales.

Un exemple de l'application de ce type de groupe concerne la *Pièce pour piano op.33a* de Schoenberg, 1929 (cf. : Figure 4). Les accords obtenus suite à la segmentation de l'œuvre se disposent de façon symétrique dans la partition.



4. LES ACCORDS DE LA PIÈCE POUR PIANO OP 33A de Schoenberg sont disposés de façon symétrique dans la partition. Il s'agit d'un exemple du paradigme du groupe diédral, car la symétrie axiale généralisée relie chaque couple d'accords.

Figure 4 : Premier exemple d'application du groupe diédral pour l'analyse musicale

Plus généralement, les parties résultant de la segmentation de l'œuvre peuvent avoir des intersections : on parle alors de segmentation par imbrication. Un exemple de cette démarche est l'analyse du théoricien David Lewin du *Klavierstück III* de Stockhausen (cf. : Figure 5).

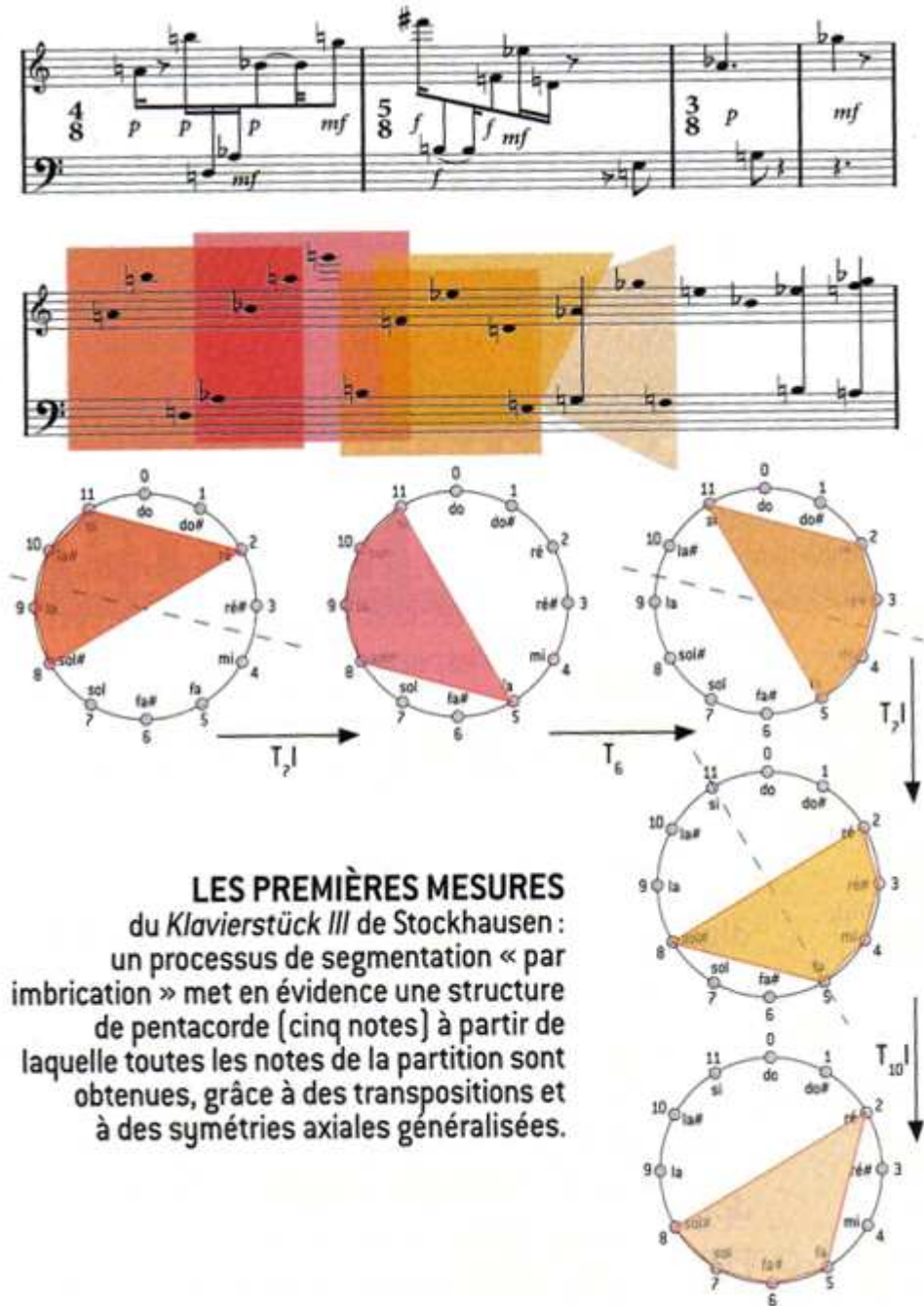


Figure 5 : Deuxième exemple d'application du groupe diédral

Dans cette analyse, 2 stratégies sont utilisées. Selon la 1^{ère}, les transformations sont organisées dans un ordre qui reflète le déroulement temporel de la pièce. Cette vision « chronologique » de l'organisation des transformations est une progression transformationnelle. Ici, le processus de segmentation par imbrication met en évidence une structure de pentacorde (série de 5 notes), où l'on passe de l'un à l'autre (les 2 ayant des notes en commun) grâce à une

symétrie axiale généralisée : tous les pentacordes sont reliés par des transpositions et des inversions.

La 2nde stratégie consiste à considérer les transformations comme une structuration possible d'un espace abstrait, un réseau transformationnel, des formes du pentacorde dans lequel on analyse le déroulement de la pièce. Dans le réseau transformationnel de l'exemple Figure 5, tous les pentacordes sont liés par des relations de transpositions et d'inverses (d'où les T et I). Ici, l'organisation des formes du pentacorde dans un réseau n'a aucun lien direct avec leur apparition chronologique.

Le groupe affine d'ordre 48

C'est l'ensemble des fonctions f qui transforment un élément x de $\mathbb{Z}/12\mathbb{Z}$ en $(ax + b)_{12}$, où a est premier avec 12 : c'est un facteur multiplicatif appartenant à l'ensemble $U = \{1, 5, 7, 11\}$. Quant à b , il appartient à $\mathbb{Z}/12\mathbb{Z}$. Dans ce cadre, une transformation affine consiste en une transposition quand $a=1$ et à une inversion quand $a=11$. Le théoricien Robert Morris a montré que ce type de transformation est compatible avec des techniques utilisées dans le Jazz.

iv. Intérêt dans les sciences cognitives

Dans un réseau, on peut retrouver les mêmes configurations de correspondances entre les pentacordes de la pièce dans des régions différentes : ce sont des isographies, à partir desquelles il pourrait y avoir un lien étroit entre les réseaux transformationnels et la perception musicale. L'écoute de la pièce deviendrait ainsi un des parcours à l'intérieur de ce réseau, avec la possibilité de repérer les isographies. Cependant, cette hypothèse n'a pas suffisamment fait l'objet de tests en psychologie expérimentale.

A partir de cet exemple, il apparaît que l'analyse transformationnelle implique la « construction » d'un réseau, mais aussi l'« utilisation » de cette architecture formelle pour dégager des critères de pertinence pour la réception de l'œuvre et pour son interprétation. Ainsi, l'intérêt de construire un réseau transformationnel réside dans la possibilité de l'utiliser pour « structurer » l'écoute, mais aussi pour établir des critères formels utiles à son interprétation. En effet, la construction d'un tel réseau permet de mettre en évidence une certaine logique musicale dans la pièce analysée.

L'approche transformationnelle représenterait donc un tournant en théorie et analyse musicale, et elle révélerait aussi une position particulière dans les rapports entre mathématiques, musique et cognition.

b) Quelques exemples d'applications aux sciences cognitives

Pour les études sur la cognition musicale, les modèles mathématiques ont été appliqués aux ensembles de hauteurs dans la musique tonale, à la représentation interne de séquences de hauteurs dans la musique tonale, aux facteurs tonals et harmoniques, aux aspects des échelles et aux relations psychologiques entre les timbres. Ces modèles se sont principalement attachés à analyser les relations structurales dans ces matériaux musicaux.

Un des spécialistes de la théorie des ensembles en musique est Hasty (1981, 1986), à l'origine de l'intégration d'une théorie de la segmentation en musique post-tonale dans une théorie du traitement de l'information, et ce dans le but de développer un ensemble de contraintes psychologiques quant à la façon dont les sons d'un morceau peuvent être regroupés.

Dans un autre domaine, Deutsch et Feroe (1981) ont proposé un modèle de l'organisation des structures pour la musique tonale. Ce modèle utilise une notion très développée de la hiérarchie pour les structures sonores, basée sur un certain nombre de primitives, appelées « alphabets », comme les triades, les échelles diatoniques et les échelles chromatiques. Cette notion est liée à l'idée de chunking (division en groupe), utilisée pour surmonter la complexité du stimulus en interaction avec les ressources limitées de la mémoire à court terme (Miller, 1956).

Une autre façon d'aborder le problème des modèles mathématiques de l'écoute musicale se retrouve dans un groupe d'étude sur la structure des échelles et sur les relations tonales et de timbres (Krumhansl, 1979 ; Wessel, 1979 ; Balzano, 1980 ; Shepard, 1982 ; Bharucha et Krumhansl, 1983). Ces modèles utilisent diverses techniques de base, dont l'échelle multidimensionnelle, mais tous se servent de représentations mathématiques ou géométriques des aspects structuraux de ces éléments musicaux.

Ainsi, les divers modèles proposés tendent à mettre en avant les interactions existant entre mathématiques et musique. Mais on peut dégager d'autres relations avec différentes disciplines, et notamment l'informatique. Les diverses théories proposées dans ce champ d'étude (en lien avec d'autres disciplines) permettent de modéliser diverses caractéristiques dans la musique, et notamment selon deux axes principaux :

- Modéliser les performances de l'Homme dans la musique, plus particulièrement l'apprentissage d'une mélodie par un individu (cf. : modèle LM présenté ci-après), ou encore la production musicale par un sujet pour l'adapter à une machine (exemple de la bouche artificielle, permettant la production de notes et donc l'analyse des phases dynamiques du jeu) ;
- Modéliser les caractéristiques plutôt propres aux instruments musicaux (ou au son), comme par l'extraction de leurs caractéristiques physiques, afin de mieux comprendre leur fonctionnement (comment ils peuvent moduler les ondes sonores créées, selon leur manipulation) et ce afin de créer des instruments virtuels aux sonorités réalistes. Il existe aussi les techniques de spatialisation du son, consistant à créer l'illusion que des sons proviennent de diverses directions de l'espace, et permettant ainsi d'organiser des scènes sonores en trois dimensions.

3. Modèle de simulation : LM

a) Intérêt

Un modèle computationnel intéressant est le système LM, créé par Richard D. Ashley et permettant de simuler un auditeur apprenant une mélodie. En effet, ce système est capable d'apprendre des mélodies en produisant des comportements similaires à ceux des musiciens qui apprennent des mélodies machinalement et les répètent en chantant pendant l'apprentissage. L'intérêt est qu'il permet de prendre en compte divers paramètres influençant les performances d'apprentissage du sujet, et pourrait ainsi mettre en évidence les différences entre musiciens et non musiciens dans l'analyse d'une mélodie, dans le sens où il peut rendre compte des erreurs des auditeurs. Le système essaie aussi d'utiliser certaines contraintes (relations entre matériaux musicaux, les problèmes de temps, les limites de la mémoire), ce qui permettra d'étudier plus en profondeur la façon dont les spécialisations détaillées de

l'architecture cognitive influencent les processus de la cognition musicale. De plus, il permet de mettre en avant la nature interactive de ces contraintes et des processus musicaux.

Ce programme a été conçu afin de rendre compte des relations entre les aspects structuraux dans des mélodies, des ressources attentionnelles accordées et du chunking dans la mémoire de travail.

b) Principe

Le programme LM se sert de 3 systèmes de production pour réaliser sa tâche :

- L'un analysant les données d'entrée en segments musicaux ;
- Un deuxième permet de mettre en évidence certains problèmes ;
- Un dernier est responsable des reproductions (la version chantée de LM).

LM est aussi composé d'une mémoire de travail regroupant un ensemble de mélodies construites à partir des schémas stockés dans une mémoire à long terme. Ces schémas sont organisés selon une hiérarchie et indexés selon le contour, le contenu des triades ou des échelles et le positionnement métrique. Ces caractéristiques sont posées pour donner des jugements du type « similaire » ou « apparenté » entre 2 schémas. Si LM ne parvient pas à trouver un schéma très spécifique, alors il se déplace vers le niveau qui correspond le mieux au percept dans sa mémoire de travail.

c) Fonctionnement

Lors de la première audition de la mélodie, LM analyse les données d'entrée perçues (représentées comme une liste d'évènements) et les divise en trames temporelles ou « trajectoires », grâce à un système de production qui utilise des règles inspirées des règles préférentielles de regroupement de Lerdahl et Jackendoff (règles tirées des lois de la gestaltthéorie, qui décrivent, dans le cas présent, une analyse primitive de notre environnement sonore). Il essaie ensuite soit d'activer un schéma existant, qui correspond au matériau, soit d'en construire un nouveau. Si le morceau n'est pas terminé, la trame temporelle est stockée comme groupe (chunk) ; sinon, elle est enregistrée comme une liste d'évènements intériorisés. Les trajectoires opèrent comme un réseau de renforcements et d'inhibitions mutuelles : les matériaux similaires se renforcent les uns les autres, tandis que les matériaux différents interfèrent, ce qui provoque un oubli partiel.

Une fois ces trajectoires construites et divisées en chunks, le rôle de LM est de découvrir l'endroit où la représentation interne de la mélodie n'est pas correcte, et d'effectuer les changements nécessaires pour l'améliorer.

d) Exemple



Figure 7 : Mélodie présentée pour l'apprentissage

Figure 6 : Protocole de LM pour un apprentissage mélodique

Après la première « audition » (*Playback 1*) de la mélodie d'origine (cf. : Figure 7), LM est capable de reproduire correctement la dernière phrase, et ce pour 2 raisons :

1. Un effet de récence vu qu'aucune autre musique n'interfère avec le matériau final ;
2. Un renforcement dû au parallélisme avec la 2^{ème} phrase (3^{ème} mesure) ;

Selon la « répétition structurale » qui a lieu entre les phrases lors de l'apprentissage, la phrase 4 devrait être retenue plus facilement, suivie de la phrase 1, puis la 2 et enfin la 3.

À la seconde « audition », LM a plus analysé la 2^{ème} phrase, qui été erronée. Il produit alors une approximation assez proche de la véritable phrase, mais « l'attention » qu'il y prête provoque d'autres erreurs, comme la modification de la dernière phrase. Ceci explique la cadence défailante que l'on retrouve dans la mesure 4 et l'absence des petits détails qui différencient le début de la phrase 3 de la 1^{ère} phrase.

À la 3^{ème} « audition », le système se rapproche le plus possible de la mélodie de départ. Ayant beaucoup analysé la phrase 1 afin de vérifier sa justesse, il arrive à en corriger la hauteur du début et la cadence. La 3^{ème} phrase a elle aussi été corrigée.

Cet exemple montre que les erreurs réalisées par LM sont similaires à celles que pourrait faire un apprenti sujet. En effet, la même expérience réalisée chez des sujets donne les résultats suivants (cf. Figure 8) :

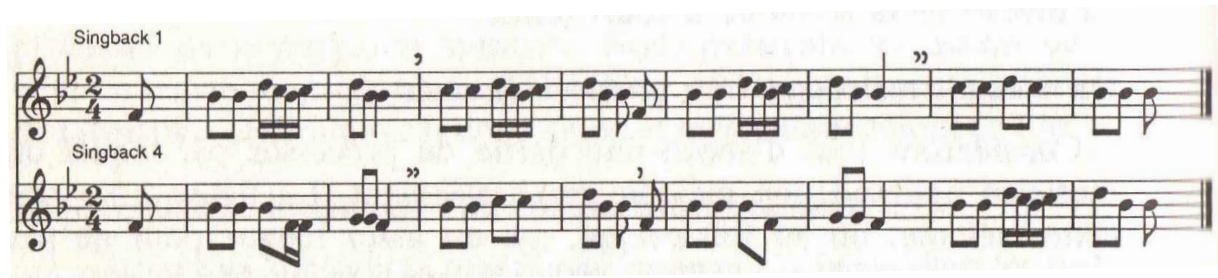


Figure 8 : Protocole humain pour un apprentissage de la mélodie

The image shows three staves of musical notation in 2/4 time, key of B-flat major. The top staff is labeled 'Playback 1', the middle 'Playback 2', and the bottom 'Playback 3'. Above the staves, the phrases are labeled 'Phrase 1' through 'Phrase 8'. The notation is identical for all three staves, representing LM performance data.

Protocole de LM pour un apprentissage mélodique

On constate que les erreurs faites par les sujets et par LM sont proches, ce qui prouverait que le programme simule des mécanismes similaires à ceux mis en jeu lors de l'apprentissage d'une mélodie par un humain.

De plus, les fautes s'expliquent par un nombre relativement restreint de facteurs, comme la puissance des associations entre les phrases dans la mémoire à court terme et les limites de l'attention.

e) Limites

LM n'est pas conçu pour imiter des expériences de l'apprentissage d'un individu. Il apprend plutôt d'une manière en apparence similaire à celle des humains, dans le sens où il produit les classes d'erreurs et les séquences d'apprentissage appropriées.