



Les descripteurs harmoniques : étude théorique et applications musicologiques

Manuel Gaulhiac

► To cite this version:

Manuel Gaulhiac. Les descripteurs harmoniques : étude théorique et applications musicologiques. Musique, musicologie et arts de la scène. Sorbonne Université, 2021. Français. NNT : 2021SORUL021 . tel-03360582

HAL Id: tel-03360582

<https://theses.hal.science/tel-03360582>

Submitted on 30 Sep 2021

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

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



SORBONNE UNIVERSITÉ

ÉCOLE DOCTORALE V

IReMus - LAM

T H È S E

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ SORBONNE UNIVERSITÉ

Discipline : Musique et musicologie

Présentée et soutenue par :

Manuel GAULHIAC

le : 22 juin 2021

Les descripteurs harmoniques
étude théorique et applications musicologiques

Sous la direction de :

M. Jean-Marc CHOUVEL – Professeur, Sorbonne Université

M. Christophe D'ALESSANDRO – Directeur de recherche, Sorbonne Université

Membres du jury :

M. Julio ESTRADA – Professeur, Facultad de Música, UNAM, Mexique

M. Xavier HASCHER – Professeur, Université de Strasbourg

M. Moreno ANDREATTA – Directeur de recherche, Université de Strasbourg

Mme Myriam DESAINTE-CATHERINE – Professeur, Bordeaux INP

Mme Marie TAHON – Maître de conférence, Le Mans Université

Mme Maria Edith ALONSO – Professeur, Universidad Internacional de la Rioja, Espagne

À ma grand-mère

Remerciements

Mes remerciements vont en premier lieu à mes directeurs de thèse, Jean-Marc Chouvel et Christophe d'Alessandro, à qui ce travail doit tant. Par la confiance qu'il m'a témoignée dès le début, Jean-Marc m'a permis de réaliser une transition heureuse des mathématiques fondamentales vers la musicologie. Son esprit critique et ouvert, sa curiosité et la profondeur de ses réflexions m'ont grandement inspiré, élargissant considérablement mon champ de vision. Son écoute, sa bienveillance et sa générosité sont précieuses. Les colloques chez lui dans la Corrèze profonde ont occasionné des rencontres passionnantes au milieu d'une effervescence artistique et intellectuelle d'une rare qualité. Je lui suis reconnaissant de m'avoir emmené au Mexique à deux reprises dans le cadre du projet de master « Muxic », et pour tous les moments partagés là-bas. Je remercie vivement Christophe pour m'avoir intégré à l'équipe du LAM. Ses conseils avisés sur la méthode scientifique m'ont appris à mieux aborder l'activité de recherche. Je salue en lui sa culture, sa gentillesse, son humour et les dégustations de thé japonais. Surtout, je lui suis redevable de m'avoir fait rencontrer ma femme, un beau jour d'avril 2018.

Je remercie les rapporteurs Xavier Hascher et Julio Estrada, ainsi que les autres membres du jury Moreno Andreatta, Myriam Desainte-Catherine, Marie Tahon et Edith Alonso.

Merci à Christophe Guillotel-Nothmann pour m'avoir suivi au tout début de ce travail, dans le cadre du stage de Master ATIAM. Boris Doval et Pierre Couprie, en tant que membres de mon comité de suivi, ont su m'aiguiller quand cela était nécessaire, je les en remercie.

Je souhaite remercier Julio Estrada pour les échanges riches sur les plans humain et scientifique, qui m'ont encouragé à utiliser les descripteurs harmoniques pour décrire acoustiquement les identités de la « carotte ». J'ai un souvenir radieux de son accueil chez lui, où nous sirotions du mezcal « Don

Julio » au milieu des colibris et des palmiers. Je lui suis reconnaissant d'avoir accepté d'être rapporteur de cette thèse. Un grand merci à mon ami César Birschner pour m'avoir fait découvrir la musique d'Almeida Prado, ainsi que pour cette collaboration fructueuse d'analyse des *Cartas Celestes I*. Je tiens à remercier tout particulièrement Yaroslav Stanishevskiy pour son amitié précieuse et pour le soutien et l'intérêt qu'il m'a témoignés. Je lui suis reconnaissant de m'avoir invité à présenter mes recherches au conservatoire de Moscou. Nos échanges ont été plus qu'utiles, et j'espère qu'ils aboutiront à de futures collaborations. Спасибо Слава, и Лена тоже!

Je souhaite remercier mes collègues du LAM et de l'IReMus, à la fois les membres de l'équipe et les doctorants. Parmi eux, une mention spéciale revient à Jean-Théo, dont le panache n'a d'égal que la faconde. Je garde des souvenirs mémorables de nos discussions sans fin aussi originales que passionnantes. Merci bien sûr à Grégoire, ainsi qu'à Yann pour les duos avec clarinette. La thèse n'aurait pas été la même sans mes collègues de bureau : Jeanne et sa bonne humeur, Franck, Cécile, Antoine, Toufik, Alexis.

La rédaction de la thèse, coïncidant avec la pandémie du Covid, fût marquée par de précieux voyages. Je salue en particulier l'Arménie, terre incroyablement riche d'histoire, de culture, de nature, qui n'en finit pas, hélas, de rebander ses plaies. Je suis infiniment reconnaissant envers le peuple arménien pour son accueil, sa chaleur, sa fierté, son mysticisme. Je remercie Nahapet, qui travaillait à la sueur de son front en me voyant écrire la thèse sur son balcon ombragé. Sa générosité est un cadeau. Avec le recul, j'aurais bien pris un peu de sa peine. Merci à Samuel, cet ours philosophe empli de douceur qui m'a fait bourlinguer et voir du pays. Merci à Ambartsum, grand mystique ayant une théorie sur tout. Nous avons, en compagnie de sa femme Myriam, partagé des moments inoubliables dans la maison des compositeurs de Dilijan. Merci à Satenik, Martin et leurs deux filles pour leur sourire et leur gentillesse. Merci à David pour m'avoir permis d'écrire humblement une partie de cette thèse là où Shostakovitch écrivait son quatuor n°14. Merci à Babadjanian pour nous avoir prêté des bûches. Je souhaite remercier tous les compagnons de randonnée : Zohrab, Anna, Mané, David. Enfin, j'ai une pensée pour tous les chiens du quartier, dont les jeux sous la fenêtre égayaient grandement la rédaction. Et j'oubliais Nerpaline, qui m'écoutait patiemment quand je doutais.

La période polonaise fut plus courte, mais non moins inspirante. Je remercie la forêt de Białowieża pour le cadre idéal d'immensité et de silence qu'elle m'a offert. Les bisons n'étaient pas les seuls à en profiter. Je suis redevable

à Janusz et sa famille de leur accueil et leur bienveillance.

De la période basque, je souhaite remercier Valérie et Hugo pour m'avoir accueilli si spontanément. Je remercie aussi la mer agitée, qui ouvrait l'horizon devant lequel j'ai mis le point final à ce travail.

Je tiens à exprimer toute ma gratitude à ma famille. À ma grand-mère en premier lieu, qui vient de quitter ce monde. Je garde d'elle le souvenir de son sourire, et lui dédie ce travail. Une pensée pour mon grand-père qui se retrouve seul. Je rends bien sûr un hommage ému à mon père, qui m'a donné le goût des mathématiques et de la musique. Ma chère maman reste un soutien constant, elle s'inquiète pour moi, m'encourage, et je lui suis reconnaissant de tout ce que je lui dois. Je remercie mes beaux-parents, Xiaocheng et Ke, pour avoir beaucoup prié pour moi. Leurs prières ont joué un grand rôle dans le bon déroulé de ce travail. Je suis satisfait de me rendre compte que ma thèse, malgré sa taille, est moins absconse que celle de mon cher frère Sylvain. Qu'il soit remercié de m'offrir cette flatteuse comparaison. Je ne parle là que de la thèse, bien entendu. À lui et à Yasya je souhaite tout le meilleur.

Enfin, je remercie celle qui m'a permis de trouver l'harmonie aussi dans le couple, celle qui ne cesse de m'inspirer, celle qui a fait grandir ma thèse, celle que j'aime énormément, Xiao.

Résumé

Cette thèse propose des modèles acoustiques adaptés à l’usage musicologique capables de rendre compte du phénomène harmonique, décrit comme l’ensemble des phénomènes sonores et des relations formelles impliquant l’interaction simultanée ou successive de plusieurs notes. Nous introduisons dans ce travail la notion de descripteur harmonique, comme type particulier de descripteur audio décrivant ce type d’interaction. Les modèles développés s’attachent aux aspects à la fois statique (accords isolés) et dynamique (enchaînements d’accords) de l’harmonie. Les descripteurs proposés — rugosité, harmonicité, concordance, concordance totale, tension, changement harmonique, concordance différentielle et rugosité différentielle — sont définis à travers une implémentation hybride, à l’interface entre les approches symbolique et signal.

Les descripteurs harmoniques fournissent des représentations que nous mettons à profit pour étudier les ensembles d’accords et de transitions. Nous proposons des cartographies des espaces harmoniques continus mettant en avant des structures liées à différentes propriétés spectrales. Le recours à des spectres de Shepard permet de généraliser les modèles aux classes d’accords. Nous appliquons cette généralisation à deux classifications : le potentiel combinatoire d’Estrada et le Tonnetz, en associant une « couleur harmonique » aux identités et aux transitions entre classes d’accords.

Les descripteurs harmoniques offrent des outils d’analyse à même de mettre en évidence des logiques acoustiques à l’œuvre dans la musique. Nous en tirons parti pour analyser un corpus de trois œuvres de la même époque qui font appel à des langages radicalement différents : *Fratres* d’Arvo Pärt (1977), la *Pièce pour orgue n°2* d’Alfred Schnittke (1980) et les *Cartas Celestes I* d’Almeida Prado (1974).

Mots-clés

Harmonie ; descripteurs harmoniques ; analyse musicale ; modèles de perception ; consonance et dissonance ; espaces d’accords et de transitions entre accords ; Tonnetz ; potentiel combinatoire des intervalles.

Abstract

This thesis proposes acoustic models designed for musicological use that describe harmonic phenomena, defined as the set of sound phenomena or the formal relationships between several simultaneous or successive notes. We introduce the notion of harmonic descriptors, which characterize the interactions between notes, and model both static (isolated chords) and dynamic (chord sequences) aspects of harmony. Proposed descriptors — roughness, harmonicity, concordance, total concordance, tension, harmonic change, differential concordance and differential roughness — are defined through a hybrid implementation at the intersection of symbolic and signal-based approaches.

Harmonic descriptors provide representations that are then used to study sets of chords and transitions. We present maps of continuous harmonic spaces, highlighting structures linked to different spectral properties. The use of Shepard spectra makes it possible to generalize the models to chord classes. This generalization is applied to two classifications : the combinatorial potential of Estrada and the Tonnetz, by associating a “harmonic color” to identities and transitions between chord classes.

Finally, we demonstrate the analytical potential of harmonic descriptors in highlighting the acoustic logics at work in a corpus of three works from the same period that use radically different languages : *Fratres* by Arvo Pärt (1977), the *Piece for organ n°2* by Alfred Schnittke (1980) and Almeida Prado’s *Cartas Celestes I* (1974).

Keywords

Harmony ; harmonic descriptors ; musical analysis ; perception models ; consonance and dissonance ; chord and transition spaces ; Tonnetz ; combinatorial potential of intervals.

Table des matières

1	Introduction	29
1.1	Motivations	29
1.1.1	Panorama des modèles perceptifs de l'harmonie	32
1.1.2	Usage en musicologie : approches statistiques	35
1.1.3	Suivi temporel par un modèle acoustique : enjeux et difficultés	39
1.1.4	Catégories acoustiques et catégories musicales	42
1.1.5	Périmètre et limitation de l'approche acoustique	51
1.2	Problématique	53
1.3	Méthodologie	54
1.4	Annonce du plan	55
1.5	Définitions	57
1.5.1	Définitions de base	57
1.5.2	Classes d'accords	60
I	État de l'art sur la consonance et la dissonance	71
2	Théories explicatives de la consonance et de la dissonance : les courants de pensée	81
2.1	Approches arithmétiques et géométriques	83
2.1.1	Courant pythagoricien	83
2.1.2	Approches géométriques	84
2.1.3	Ordre et raison au XVII ^e siècle : la théorie des coups	85
2.1.4	Euler, ou la tentation de l'abstraction arithmétique	88
2.1.5	Les successeurs	90
2.2	Approche physicaliste	92
2.2.1	Rameau et d'Alembert : le corps sonore résonant	92

2.2.2	Au-delà du corps sonore harmonique : lien entre spectre et échelle	98
2.3	Approches perceptives	106
2.3.1	La révolution perceptive	106
2.3.2	Étude psychologique de la sensation : la fusion tonale	113
2.3.3	Héritage d’Helmholtz et de Stumpf	117
3	Théories de la perception de l’harmonie : état de la recherche actuelle	121
3.1	Multitude des approches et des modèles	122
3.2	Volonté de synthèse	126
II	Les descripteurs harmoniques : introduction et implémentation	131
4	Des descripteurs audio aux descripteurs harmoniques	133
4.1	Les descripteurs audio	133
4.1.1	Différentes catégories de descripteurs audio	133
4.1.2	Usage en musicologie numérique	136
4.2	Des descripteurs audio aux descripteurs harmoniques	141
4.2.1	Les descripteurs audio traitant de l’harmonie	141
4.2.2	Descripteurs audio et modèles de perception	146
4.2.3	La notion de descripteur harmonique	148
4.3	Usage des descripteurs harmoniques	151
4.3.1	Origine des descripteurs	151
4.3.2	À l’interface entre le symbolique et le signal	152
4.3.3	Approches statiques et dynamiques	158
4.3.4	Prise en compte du contexte harmonique	159
5	Méthode et implémentation des descripteurs harmoniques	161
5.1	Vue générale	161
5.1.1	Approches signal et symbolique : avantages et usages	161
5.1.2	Les différentes configurations	163
5.1.3	Modélisation du contexte harmonique	164
5.2	Approche symbolique	169
5.2.1	Spectre en gaussiennes	170
5.2.2	Spectres de Shepard	174

5.3	Approche signal	181
5.3.1	Analyse spectrale	181
5.3.2	Synchronisation du spectre	188
5.3.3	Détection des partiels	193
6	Présentation des descripteurs harmoniques	197
6.1	Les descripteurs statiques	198
6.1.1	Rugosité	198
6.1.2	Concordance	208
6.1.3	Concordance totale et concordance d'ordre $m \geq 2$. . .	214
6.1.4	Harmonicité	219
6.1.5	Tension	222
6.2	Les descripteurs dynamiques	226
6.2.1	Changement harmonique	227
6.2.2	Concordance différentielle	230
6.2.3	Rugosité différentielle	235
7	Contraintes en entrée et normalisations	241
7.1	Contraintes sur les sons utilisés	241
7.1.1	Temps d'établissement de l'harmonie	242
7.1.2	Variation temporelle de l'amplitude	243
7.2	Comportement de descripteurs harmoniques vis-à-vis des nuances	249
7.2.1	Conditions d'indépendance vis-à-vis des nuances	251
7.2.2	Normalisation des descripteurs par l'énergie	252
7.2.3	Cas limite : les nuances par note	257
7.3	Influence du nombre de notes	258
7.3.1	Considérations perceptives et théoriques	258
7.3.2	Normalisation des descripteurs sur les unissons	263
7.3.3	Cas limite : unisson au sein d'un accord	269
III	Exploration de l'univers harmonique	271
8	Les espaces continus d'accords et de transitions	273
8.1	Structuration des dyades	274
8.1.1	Histoire des classements des intervalles	274
8.1.2	Courbes de descripteurs harmoniques	276
8.2	Structuration des triades	283

8.2.1	Classements et représentations des triades	283
8.2.2	Topographie de l'espace des triades	290
8.3	Aspect dynamique : ajout d'une troisième note à une dyade .	309
8.3.1	Représentation des espaces de transition	309
8.3.2	Cartes de gain des descripteurs statiques	311
8.3.3	Surfaces de descripteurs dynamiques	317
8.3.4	Descripteurs dynamiques triviaux et non triviaux : syn- thèse des apports	320
9	Classes et classifications d'accords	321
9.1	Calcul des descripteurs harmoniques sur les classes d'accords .	321
9.1.1	Affranchissement de la dépendance à la disposition . .	321
9.1.2	Spectres de Shepard	324
9.1.3	Descripteurs de classes normales, descripteurs de classes premières et descripteurs dynamiques entre ensembles de hauteurs.	326
9.1.4	Espaces continus de classes de dyades et de triades . .	327
9.1.5	Structuration des classes d'intervalles	334
9.1.6	Topographie des classes premières de triades	337
9.1.7	Topographie des classes normales de triades	340
9.2	Continuum et discontinuum des classes : application au po- tentiel combinatoire des intervalles	342
9.2.1	Continuum et discontinuum	342
9.2.2	Le cas des tetrachordes	345
9.2.3	La « carotte » ou potentiel combinatoire des intervalles	359
9.2.4	Coloration de la « carotte » par les descripteurs har- moniques	367
9.2.5	Conclusions et perspectives	382
9.3	Enchaînements harmoniques dans le Tonnetz	383
9.3.1	Le Tonnetz : présentation	383
9.3.2	Coloration du Tonnetz par les descripteurs harmoniques dynamiques	388
9.3.3	Description acoustique des résolutions cadentielles . . .	394
9.3.4	Conclusions et perspectives	401

IV Exemples musicaux	405
10 Arvo Pärt - <i>Fratres</i>	409
10.1 Rôle harmonique et perceptif de la voix tintinnabulante	410
10.2 Ajout progressif d'accords	413
11 Alfred Schnittke - <i>Pièce pour orgue op.149 n°2</i>	417
12 Almeida Prado - <i>Cartas Celestes I</i>	429
12.1 Analyse statique du catalogue d'accord	435
12.1.1 Un alphabet d'accords	435
12.1.2 Cartographies des accords	439
12.2 Analyse de <i>Scorpio</i>	448
12.2.1 Résidus cadentiels	450
12.2.2 Incidence de la compression rythmique sur la résonance	454
Conclusion	465
Résumé des travaux	465
Apports	469
Perspectives	471
Bibliographie	473
Annexes	497
Annexe 1 : Partition de <i>Fratres</i> d'Arvo Pärt	499
Annexe 2 : Partition de la <i>Pièce pour orgue n°2</i> d'Alfred Schnittke	511
Annexe 3 : Partition des <i>Cartas Celestes I</i> d'Almeida Prado	517

Table des figures

1.1	Prévalence des triades - approche diachronique	37
1.2	Courbes de rugosité sur une sonate de Scarlattti	40
1.3	Courbe de dissonance d'Helmholtz	45
1.4	Cercle des hauteurs	61
1.5	Transposition des ensembles de classes de hauteurs	62
1.6	Inversion des ensembles de classes de hauteurs	63
1.7	Contenu intervallique d'un ensemble de hauteurs.	64
1.8	Ensembles de hauteurs homométriques	67
1.9	Ensembles de hauteurs d'une même identité	68
1.10	Frise chronologique des CDC	77
2.1	Échelle adaptée au son de cloche	103
2.2	Spectre associé à l'échelle 10-tet, courbe et surface de dissonance	105
4.1	Table des descripteurs audio	136
4.2	Logiciels utilisés en analyse musicale	138
4.3	Suite d'accords	143
4.4	Analyse en chromas	145
4.5	Analyse avec <i>Chordino</i>	146
4.6	Schéma descripteurs harmoniques/descripteurs audio/modèle de perception	149
4.7	Modèle « Attaque Chute Entretien Extinction »	156
4.8	Schéma des approches statiques et dynamiques	158
4.9	Le contexte dans le calcul des descripteurs statiques et dyna- miques	160
5.1	Modélisation du contexte, tiré de la littérature	167
5.2	Modélisation du contexte : coefficients de pondération	168
5.3	Schéma de l'approche symbolique	169

5.4	Spectre en pics	173
5.5	Spectre en gaussienne	174
5.6	Spectre de Shepard, tiré de la littérature	175
5.7	Spectre de Shepard en pics	177
5.8	Spectre de Shepard en gaussiennes	178
5.9	Spectre de Shepard - énergie constante sur les hauteurs	179
5.10	Spectres de Shepard - rôle d'écart-type de l'enveloppe spectrale	180
5.11	Schéma de l'approche signal	181
5.12	Spectrogramme obtenu par TQC	183
5.13	Cadence à 4 voix	184
5.14	Position des fenêtres d'analyse dans le calcul de la TQC	184
5.15	Spectrogramme obtenu par TQC avec relocalisation des fenêtres	185
5.16	Séparation en partie harmonique et partie percussive + bruit	188
5.17	Synchronisation sur les verticalités - sans retard	190
5.18	Synchronisation sur les verticalités - avec retard	191
5.19	Synchronisation sur les verticalités - la cas du vibrato	192
5.20	Étalement des partiels	193
5.21	Détection des partiels	195
6.1	Courbe de rugosité de deux sons simples - tiré de la littérature	200
6.2	Courbe de rugosité des sons simples	202
6.3	Évolution de la rugosité, sur 3 accords	204
6.4	Accord diminué : interaction des partiels	205
6.5	Évolution de la rugosité - à partir du spectre global	206
6.6	Évolution de la rugosité - à partir des spectres non ponctuels	207
6.7	Coïncidence des harmoniques chez Helmholtz	209
6.8	Coïncidence des partiels sur les intervalles	210
6.9	Produit de gaussiennes	212
6.10	Évolution de la concordance	213
6.11	Évolution de la concordance totale - accords de 3 sons	216
6.12	Coïncidence simultanée des harmoniques sur l'accord mineur	217
6.13	Série des harmoniques inférieurs	218
6.14	Évolution de l'harmonicité - accords de 3 sons	221
6.15	Hauteurs fondamentales virtuelles - sur 3 accords	222
6.16	Courbe de tension	223
6.17	Évolution de la tension - accords de 3 sons	225
6.18	Utilisation de la TFD - <i>Suite Lyrique</i> de Berg	228
6.19	Évolution du changement harmonique - <i>Via Crucis</i> de Liszt	230

6.20	Schéma des successions d'accords	231
6.21	Évolution de la concordance différentielle - <i>Via Crucis</i> de Liszt	232
6.22	Coïncidence des partiels entre accords successifs	233
6.23	Évolution de la concordance différentielle - effet des fautes . .	234
6.24	Évolution de la rugosité différentielle - <i>Via Crucis</i> de Liszt . .	237
6.25	Contributions à la rugosité différentielle - <i>Via Crucis</i>	238
6.26	Dépendance de la rugosité différentielle d'un ton avec le registre	239
7.1	Temps d'établissement de l'harmonie	243
7.2	Modulation d'amplitude - effet du timbre	244
7.3	Modulation d'amplitude - effet des battements	245
7.4	Modulation d'amplitude - effet sur les descripteurs harmoniques	247
7.5	Absence de modulation en amplitude significative	248
7.6	Courbes d'isotonie	249
7.7	Indépendance des descripteurs statiques relativement aux nuances	254
7.8	Comportements des descripteurs dynamiques vis-à-vis des nuances.	256
7.9	Cas limite : les nuances par note	257
7.10	Indépendance des descripteurs statiques au nombre de notes sur les unissons	267
7.11	Évolution de la tension sur les unissons, jeux Tutti	268
7.12	Évolution des descripteurs dynamiques sur les transitions d'unis- sons.	269
7.13	Dépendance des descripteurs statiques aux doublures de notes	270
8.1	Évolution des système de classification des intervalles, IX^e - XVI^e s.	275
8.2	Courbes de rugosité	276
8.3	Courbes de rugosité sur l'octave - influence de K	277
8.4	Courbes de rugosité sur l'octave - influence de la hauteur de f_0	278
8.5	Courbe de concordance sur deux octaves	280
8.6	Courbe de concordance sur deux octaves - influence de σ . . .	281
8.7	Courbe d'harmonicité	282
8.8	Table pour déterminer les accords de trois sons consonants . .	285
8.9	Triades : tableau comparatif des expériences perceptives et des modèles.	286
8.10	Classement des triades, tiré de (Johnson-Laird <i>et al.</i> , 2012) . .	288
8.11	Espace des triades, tiré de (Mach, 1866)	289

8.12	Surface de dissonance : représentation plate et représentation 3D	290
8.13	Plan des triades - localisation accords majeurs et mineur	291
8.14	Surface de concordance - visualisation en 3D	292
8.15	Surface de concordance - visualisation plane	294
8.16	Concordance des accords majeurs dans différentes positions	295
8.17	Coïncidence des partiels de l'accord majeur	295
8.18	Surfaces de concordance - variation de K	296
8.19	Surface de rugosité des triades	297
8.20	Surfaces de rugosité - variation de la note inférieure	298
8.21	Surface de tension des triades	299
8.22	Surface de tension des triades avec subdivision tempérée	301
8.23	Surface d'harmonicité des triades	302
8.24	Surface d'harmonicité - position des accords	303
8.25	Surface de concordance d'ordre 3 des triades	304
8.26	Surface de concordance d'ordre 3 - position des accords	305
8.27	Triades au sein de la série harmonique inférieure	307
8.28	Espace de transitions : ajout d'une troisième note à une dyade	310
8.29	Carte de perte de rugosité	312
8.30	Carte de gain d'harmonicité	314
8.31	Carte de gain de concordance totale	316
8.32	Ajout d'une note : surfaces de concordance différentielle et changement harmonique	318
8.33	Ajout d'une note - surface de rugosité différentielle	319
9.1	Présentations réduites - visualisation sur l'espace des triades	328
9.2	Localisation des formes normales de triades	331
9.3	Localisation des formes premières des triades	333
9.4	Shepard vs. non Shepard : comparaison des courbes de concordance	334
9.5	Courbe de concordance de classe	336
9.6	Surface de concordance des classes premières de triades	338
9.7	Surface de rugosité des classes premières de triades	339
9.8	Surface de concordance d'ordre 3 des classes normales de triades	340
9.9	Surface d'harmonicité des classes normales de triades	341
9.10	Localisation des identités de triades avec subdivision de l'octave	343
9.11	Tetrachordes : localisation des formes identitaires	346
9.12	Tetrachordes : rugosité des classes identitaires	348

9.13	Tetrachordes : concordance des formes identitaires	349
9.14	Tetrachordes : localisation des formes premières	351
9.15	Tetrachordes : localisation des classes normales	352
9.16	Orbite identitaire dans l'espace des formes normales	353
9.17	Concordance sur une orbite identitaire	355
9.18	Tension sur une orbite identitaire	356
9.19	Orbite de classe normale	357
9.20	Potentiel combinatoire de l'échelle chromatique.	361
9.21	Orbite identitaire associée à la forme identitaire <i>do-réb-mib-sol</i>	362
9.22	Dénombrement des identités pour différentes subdivisions . . .	366
9.23	Interface du logiciel de coloration de la carotte par les descrip- teurs harmoniques statiques	368
9.24	Carotte colonnes 3 à 5 - harmonicité des formes identitaires . .	372
9.25	Carotte colonne 3 - tension des formes identitaires	373
9.26	Carotte - rugosité des formes identitaires	375
9.27	Carotte - rugosité des formes normales	376
9.28	Carotte - concordance des formes identitaires	377
9.29	Carotte - harmonicité des formes identitaires	378
9.30	Carotte - tension des formes identitaires	379
9.31	Carotte - concordance totale des formes identitaires	380
9.32	Carotte - coloration mixte par la rugosité, l'harmonicité et la concordance	381
9.33	Tonnetz	384
9.34	Tonnetz - 2 géométries	384
9.35	Transformations élémentaires de la théorie néo-riemannienne .	386
9.36	Tonnetz augmenté	387
9.37	Logiciel de coloration du Tonnetz - interface	389
9.38	Transitions à partir de <i>Do M</i> : coloration par les descripteurs dynamiques	392
9.39	Transitions à partir de <i>Do M</i> : coloration mixte	393
9.40	Résolution d'une cadence mineure - coloration mixte du Tonnetz	397
9.41	Résolution d'une cadence mineure - coloration mixte du Ton- netz - spectre à 1 partiel	398
9.42	Cadence majeure avec suspension de la tierce - coloration mixte du Tonnetz	400
10.1	<i>Fratres</i> : versions droite et rétrogradées du thème	411
10.2	<i>Fratres</i> - le rôle harmonique de la voix <i>tintinnabulante</i>	412

10.3	<i>Fratres</i> - influence de l'ajout de notes sur les trajectoires harmoniques	415
11.1	<i>Pièce pour orgue n°2</i> d'Alfred Schnittke — premières mesures.	418
11.2	<i>Pièce pour orgue n°2</i> de Schnittke — exploration des identités	419
11.3	<i>Pièce pour orgue n°2</i> de Schnittke - spectrogramme	421
11.4	<i>Pièce pour orgue n°2</i> de Schnittke - rugosité sur la carotte . .	422
11.5	<i>Pièce pour orgue n°2</i> de Schnittke - harmonicité sur la carotte	423
11.6	<i>Pièce pour orgue n°2</i> de Schnittke - tension sur la carotte . . .	424
11.7	<i>Pièce pour orgue n°2</i> de Schnittke - comparaison de 3 interprétations	427
12.1	<i>Cartas Celestes I</i> - accord ϵ	437
12.2	<i>Cartas Celestes I</i> - accords qui sont dans la même classe normale	438
12.3	<i>Cartas Celestes I</i> - Cartographie du catalogue d'accord dans l'espace (rugosité, harmonicité)	441
12.4	<i>Cartas Celestes I</i> - Hauteurs fondamentales virtuelles et hauteurs fondamentales virtuelles inverses	445
12.5	<i>Cartas Celestes I</i> - Cartographie du catalogue d'accord dans l'espace (harmonicité, harmonicité inverse)	446
12.6	<i>Cartas Celestes I</i> - Cartographie du catalogue d'accord dans l'espace (rugosité, harmonicité, harmonicité inverse)	447
12.7	<i>Cartas Celestes I</i> - premier cycle de <i>Scorpio</i>	449
12.8	<i>Cartas Celestes I</i> - aspect cadentiel de <i>Scorpio</i>	453
12.9	<i>Cartas Celestes I</i> - <i>Scorpio</i> : effet de la pédale sur les transitions	458
12.10	<i>Cartas Celestes I</i> - <i>Scorpio</i> : effet des répétitions sur les transitions	459
12.11	<i>Cartas Celestes I</i> - <i>Scorpio</i> : effet de la compression rythmique sur les transitions	461

Liste des tableaux

1.1	Résumé des CDC	78
4.1	Rapports de fréquences et déviations en cents des intervalles justes	144
7.1	Descripteurs statiques : effet de la multiplication du spectre par α	250
7.2	Descripteurs dynamiques : effet de la multiplication des spectres par α	250
7.3	Effet de la normalisation par l'énergie sur les descripteurs statiques	253
7.4	Effet de la normalisation par l'énergie sur les descripteurs dynamiques	255
7.5	Classement des classes normales par ADC extensive	262
7.6	Classement des classes normales par ADC intensive	263
9.1	Dénombrement des identités	344
9.2	Dispersion des couleurs sur les orbites identitaires	371

Introduction

Chapitre 1

Introduction

1.1 Motivations

Partons d'un exemple simple, qui va permettre d'introduire la problématique de cette thèse. Imaginons une expérience où un pianiste joue simultanément l'intervalle *do-si* sur la même octave dans le médium du piano, puis rejoue les mêmes notes avec en plus le *mi* de la même octave, pour former l'accord *do-mi-si*.

Si on demande à un auditeur de juger la sonorité des accords, il est probable qu'il réponde que l'intervalle *do-si* est désagréable à l'oreille, tandis que l'accord *do-mi-si* sonne mieux. S'il connaît les termes il dira que le premier accord est plus dissonant que le second, ou que le second est plus consonant que le premier.

Les modèles usuels de dissonance ([Plomp et Levelt, 1965](#); [Kameoka et Kuriyagawa, 1969a,b](#); [Hutchinson et Knopoff, 1979](#)) expliquent la dissonance par la rugosité, qui est la sensation désagréable due au battements des partiels situés dans la même bande critique. Or, toujours selon cette théorie, les battements, et donc la dissonance, ne peuvent qu'augmenter avec l'ajout d'une note. Les modèles de rugosité sont donc impuissants à expliquer le jugement précédent.

Parncutt, qui reprend une expérience de pensée similaire dans ([Parncutt, 2006](#)), fournit quatre hypothèses :

- **Explication 1** : Les battements de l'intervalle *do-si* sont partiellement masqués quand le *mi* est rajouté.

- **Explication 2** : La rugosité de l'accord *do-mi-si* est distribuée sur les bandes critiques de manière plus homogène que pour l'intervalle *do-si*.
- **Explication 3** : L'accord *do-mi-si* est un accord de septième majeure sans la quinte, on l'identifie en tant que tel et il est plus familier que l'intervalle *do-si*.
- **Explication 4** : L'accord à 3 notes aurait une propriété de fusion ou d'euphonie que n'a pas l'intervalle *do-si*. La fusion est la tendance à percevoir un amas sonore comme une seule note.

Les deux premières explications sont d'ordre psychoacoustique, la psychoacoustique étant la science qui se propose de relier les caractéristiques acoustiques d'un stimulus à sa perception ; la troisième d'ordre culturel, et la dernière d'ordre psychologique.

Parncutt nuance ses hypothèses en prévenant de l'insuffisance des modèles : aucun modèle de dissonance ne tient compte à la fois du masquage et de la distribution spectrale des contributions à la dissonance. Il ajoute que la notion d'euphonie, qui contribuerait positivement à la consonance de manière indépendante des autres facteurs, si tant est qu'elle existe, échappe pour l'instant à la compréhension.

Plusieurs difficultés peuvent venir nuancer la portée de cette expérience de pensée, citons-en quelques-unes :

1. **Multi-dimensionnalité de la perception** Quelle consigne a été donnée à l'auditeur ? Est-ce celle de juger les deux accords en matière d'agrément, de préférence, de douceur, d'absence de tension, d'harmonie, de pureté, de simplicité, voir directement de consonance ? Ces notions sont synonymes, mais non équivalentes ([Lahdelma et Eerola, 2020](#)). La perception a de multiples facettes, qu'il convient de clarifier.
2. **Diversité des auditeurs** Selon la troisième explication, l'accord *do-mi-si* apparaîtrait comme plus consonant du fait qu'il a une interprétation tonale plus évidente que la dyade *do-si* pour un auditeur familier du système tonal. Mais alors, comment réagirait l'indigène de la forêt amazonienne qui n'a jamais été confronté à la musique occidentale ? Ce point pose la question de l'universalité du jugement perceptif.

3. **Nature du jugement** Le jugement de l'auditeur peut a priori se situer autant au niveau de la sensation immédiate, entendue comme impression passive, qu'à celui de la perception, entendu comme interprétation active de la sensation mettant en jeu des procédés psychiques et cognitifs tels la mémoire ou l'imagination.
4. **Influence des facteurs acoustiques** Les deux premières explications se réfèrent aux interactions entre partiels. Or l'emploi d'un autre instrument que le piano modifie le spectre des notes, et donc les interactions entre partiels. De même, un changement de registre modifie les largeurs de bandes critiques. Les jugements de perception dépendent donc du timbre et du registre.
5. **Contexte harmonique** Les accords ont été joués dans un ordre donné. Autrement dit, le premier accord est resté dans la mémoire de l'auditeur lors du second accord. Le jugement sera-t-il le même si l'ordre des accords est inversé ? Pas nécessairement, car la mémoire et l'attente peuvent jouer un rôle majeur dans le processus de perception.
6. **Contexte musical** Les accords ont été isolés de tout contexte musical. Le jugement serait-il similaire si l'enchaînement était joué au sein d'une séquence d'accords ? Là encore, rien n'est moins sûr. On sait que le contexte musical et le langage influent grandement la perception. Cette remarque limite grandement la portée des conclusions de l'expérience quand celles-ci se voudraient musicales.

Nous voyons par cet exemple que la perception des accords soulève de nombreuses questions en partie irrésolues, et constitue un phénomène complexe qui touche aux deux aspects de l'opposition nature/culture. Les approches dites naturelles considèrent la perception comme innée, déterminée par les propriétés physiques des accords ou par le fonctionnement du système auditif. Les approches culturelles, quant à elles, défendent l'idée que la perception des accords est acquise, résultat de la familiarité, de la mémoire et de l'apprentissage.

Une question importante soulevée par les remarques précédentes consiste à savoir si la compréhension du phénomène qui sous-tend le jugement perceptif de l'expérience de pensée apporterait une meilleure compréhension de la présence de cet enchaînement dans une œuvre où il apparaîtrait. La question se formule autrement, en se généralisant : la compréhension des phénomènes

perceptifs est-elle utile à l'analyse musicale ?

La réponse est oui, c'est l'objet de cette introduction. Nous commençons par un panorama des modèles perceptifs de l'harmonie, avant de présenter des études statistiques sur des corpus musicaux, qui mettent en évidence l'apport de ces modèles à l'analyse des processus compositionnels. À partir de ces exemples, nous généralisons le propos à des modèles acoustiques, non uniquement perceptifs, et introduisons les enjeux sémiotiques du rapprochement des faits acoustiques avec les faits musicaux.

1.1.1 Panorama des modèles perceptifs de l'harmonie

Nous présentons ici les enjeux des recherches perceptives sur l'harmonie.

Expériences perceptives et modèles de perception Les premières expériences sur la perception des intervalles datent du début du XXe siècle. Elles traitent le problème de la consonance et de la dissonance ([Malmberg, 1918](#); [Guernsey, 1928](#)) ou celui de la fusion tonale ([Stumpf, 1883, 1890](#); [Brues, 1927](#); [Guthrie et Morrill, 1928](#)). Depuis lors, la perception des accords est un champ de recherche fécond, fruit de nombreux débats au cœur de l'actualité. Le problème épineux de l'origine de la consonance et de la dissonance, considéré comme « *l'une des questions les plus fondamentales de la perception musicale*¹ », agite particulièrement la communauté de recherche ces dernières années ([Harrison et Pearce, 2020](#); [Eerola et Lahdelma, 2020](#); [Lahdelma et al., 2020](#); [Smit et al., 2020](#); [Friedman et al., 2021a](#)).

Les études sur la perception de l'harmonie visent à déterminer les mécanismes influençant la perception des accords et des enchaînements d'accords. Elles aboutissent à des modèles prédictifs qui prennent en entrée tantôt une partition, tantôt un signal audio. Certaines approches modélisent directement le fonctionnement des organes de l'appareil auditif ([Helmholtz, 1863](#)) ou la réponse du cerveau au stimulus sonore ([Shapira Lots et Stone, 2008](#); [Bidelman et Krishnan, 2009](#)). Les approches psychoacoustiques court-circuitent quant à elle l'aspect biologique en tentant de relier directement les propriétés acoustiques au jugement perceptif. Les approches culturelles, enfin, considèrent la

1. "The origins of tonal consonance—the tendency to perceive some simultaneously sounded combinations of musical tones as more pleasant than others—is arguably among the most fundamental questions in music perception." Tiré de ([Smit et Milne, 2021](#)), traduction de l'auteur.

perception comme objet culturel, façonné par la familiarité et l'apprentissage, et non réductible aux propriétés acoustiques ou biologiques.

Sensation brute vs. interprétation cognitive Dans tout phénomène perceptif on distingue généralement deux degrés, le premier, de bas-niveau, étant le niveau physiologique ou sensoriel, qui convoque les organes de l'audition et donne lieu à la sensation immédiate et passive. Le second est le niveau cognitif, qui inclut la mémoire, la familiarité, la projection ([Stevens, 1957](#)), les jugements esthétiques. Il peut être interprété comme une réévaluation active de la sensation immédiate par les fonctions cognitives.

La majorité des modèles psychoacoustiques visent à décrire la sensation. Plusieurs raisons encouragent cette approche :

1. Il est parfois sous-entendu que la sensation est déterministe et universelle, ne dépendant que des propriétés acoustiques du signal en entrée, contrairement à la perception de plus haut niveau, qui dépend en grande partie de facteurs culturels, plus complexes à modéliser. Dès lors, la sensation se prête mieux aux tests perceptifs et à la modélisation.
2. Depuis la théorie des battements d'Helmholtz ([Helmholtz, 1863](#)), l'approche physiologique de la sensation a prouvé son efficacité avec des théories relativement simples.
3. L'idée que la compréhension des mécanismes fondamentaux de la sensation est une condition nécessaire avant de pouvoir aborder des réponses plus sophistiquées comme les émotions ou le jugement esthétique incite à se concentrer sur la sensation.

Or, la perception de l'harmonie en musique est un phénomène complexe qui inclut les aspects tant physiologiques que cognitifs ([Lahdelma, 2017](#); [Parncutt *et al.*, 2019](#); [Harrison et Pearce, 2020](#)). Notamment, l'aspect culturel joue un rôle central même dans la perception des accords isolés ([Tramo *et al.*, 2001](#); [Johnson-Laird *et al.*, 2012](#); [Bowling *et al.*, 2017](#)), ce qui rend illusoire l'idée de « sensation pure » qui ne serait qu'une réaction physiologique.

Contexte musical et contexte tonal Un grand nombre de recherches traite la perception des accords isolés de tout contexte musical. Cette position très actuelle (Lahdelma, 2017; Harrison et Pearce, 2020; Eerola et Lahdelma, 2020; Friedman *et al.*, 2021a) se justifie par la volonté de comprendre les mécanismes fondamentaux de la perception, et notamment l'influence de paramètres comme le timbre, le tempérament, le registre (Carcagno *et al.*, 2019).

Pourtant, la mémoire à court terme joue un rôle central dans la perception de l'harmonie (Rogers, 2010) lorsqu'il s'agit d'une séquence d'accords ou d'un extrait musical. La perception des accords comme entités isolées serait même illusoire, étant donnée l'importance des mécanismes de comparaison (Arthurs, 2015).

Certains modèles perceptifs de l'harmonie proposent une modélisation du contexte musical (Jensen et Hjortkjær, 2012; Johnson-Laird *et al.*, 2012; Cong, 2016; Chan *et al.*, 2019).

Ces deux types de perception de l'harmonie, la perception hors contexte musical, et la perception en contexte musical, sont source de nombreuses confusions terminologiques. Terhardt (Terhardt, 1984) résout cette confusion à propos de la dissonance et de la consonance en distinguant deux notions : la dissonance sensorielle (*sensory dissonance*), qui traite des accords isolés, et la dissonance musicale (*musical dissonance*), parfois appelée dissonance harmonique (*harmonic dissonance*), qui traite des accords en contexte musical. Terhardt considère que la dissonance sensorielle ne dépend que de facteurs naturels, l'influence de la culture n'étant présente qu'au niveau de la dissonance musicale¹. Cette position n'est plus tenable maintenant, la familiarité à un langage étant reconnue comme influençant la perception d'accords isolés, même situés en-dehors de ce langage. Rogers (Rogers, 2010) propose le terme de dissonance cognitive (*cognitive dissonance*), pour prendre en compte autant la mémoire à court terme (contexte musical) que la mémoire à long terme (familiarité, acculturation).

1. Pour Terhardt, la dissonance sensorielle est déterminée par plusieurs facteurs : rugosité, "sharpness", "loudness" et "toneness", tandis que la dissonance musicale est déterminée par la dissonance sensorielle, l'affinité tonale, la compatibilité et la "root relationship/fundamental-note relation". Se reporter à (Terhardt, 1984) pour plus de détails.

Aspects statiques et transitionnels L’aspect statique concerne ce qui se passe à un instant donné, et l’aspect transitionnel, ou dynamique, la transition entre deux entités se succédant temporellement.

Les premières théories physiologiques et les premières expériences perceptives se sont attachées à l’aspect statique de l’harmonie (Helmholtz, 1863; Stumpf, 1883, 1890; Faist, 1897; Malmberg, 1918; Guernsey, 1928).

Dès la fin du XIX^e siècle des expériences perceptives sont menées en faisant écouter aux auditeurs deux intervalles successifs (Meinong et Witasek, 1897; Pear, 1911). Le but est alors de comparer les degrés de consonance, et non de juger l’enchaînement. C’est la méthode dite de comparaison par paires.

Ce n’est que des années 80 que datent les premières expériences perceptives sur les enchaînements d’accords (Krumhansl et Kessler, 1982). Plusieurs théories psychoacoustiques pour décrire les transitions d’accords sont proposées, parmi lesquelles des modèles d’affinité (Terhardt, 1984; Parncutt, 1989; Thompson et Parncutt, 1997; Milne *et al.*, 2016) ou de tension de transition (Chan *et al.*, 2019). L’étude des transitions peut être, comme dans l’étude statique, réalisée hors contexte musical ou en contexte musical. Dans (Bharucha et Krumhansl, 1983), les auteurs incluent des éléments de contexte dans leur modélisation de la perception des enchaînements d’accords.

Malgré l’existence de modèles dynamique, l’approche statique est majoritaire aujourd’hui (Stolzenburg, 2015; Harrison et Pearce, 2020), et reste plus fouillée que l’approche dynamique.

1.1.2 Usage en musicologie : approches statistiques

Revenons maintenant à la question posée dans le prologue : la compréhension des phénomènes perceptifs est-elle utile à l’analyse musicale ? La réponse est affirmative. Nous le montrons dans cette partie en nous référant à des études qui mettent en rapport la fréquence d’apparition des intervalles ou des triades dans un corpus donné à des modèles perceptifs.

Distribution des intervalles sur le registre Plomp et Levelt, dans l’article *Tonal Consonance and Critical Bandwidth* (Plomp et Levelt, 1965), améliorent le modèle de rugosité d’Helmholtz en mettant en avant le rôle dans la dissonance de la largeur de bande critique, notion liée à l’intervalle maximal entre deux sons simples pour que ceux-ci interagissent, au sens de l’intensité sonore (Zwicker *et al.*, 1957). Plus précisément, ils montrent que l’intervalle

de dissonance maximale entre deux sons simples, le son grave étant fixé, correspond à 25% de la largeur de bande critique. Comme cette dernière dépend de la hauteur relative du son grave, la dissonance d'un intervalle dépend de son registre.

Dans ce même article, les auteurs tentent de montrer le rôle de la bande critique dans le choix des intervalles par les compositeurs, à travers l'étude statistiques des intervalles présents dans un corpus formé de la *Sonate pour orgue n°3* de Bach et du 3^e mouvement du *Quatuor op.51* en *mi*b majeur de Dvořák. En faisant varier le nombre de partiels, les auteurs étudient les fréquences d'occurrence des intervalles statiques (partiels compris) en fonction de leur registre. Ils montrent que l'augmentation du nombre de partiels entraîne une augmentation de la proportion des intervalles situés dans la bande critique. Notamment pour un nombre de partiels modélisant bien les sons musicaux (de 8 à 10 selon les auteurs), plus de 50% des intervalles sont situés dans la bande critique.

Plomp et Levelt ont montré, avec des moyens rudimentaires, que dans les œuvres considérées la distribution des intervalles statiques sur le registre est liée à la notion de bande critique. Ils interprètent ces résultats comme un moyen pour le compositeur de réaliser une intention musicale : celle de faire varier dans l'œuvre la consonance et la dissonance.

Consonance et clarté des voix chez Bach Dans l'article *Tonal Consonance versus Tonal Fusion in Polyphonic Sonorities*, ([Huron, 1991](#)), Huron corréle les fréquences d'occurrence des intervalles dans les *Inventions* à deux voix et les *Sinfonia* à trois voix de Bach avec des modèles de consonance et de fusion. La fusion est la tendance à percevoir un accord comme une seule note. Ce faisant, il teste deux hypothèses : la recherche de la part de Bach de sonorités consonantes, qui entraînerait une corrélation positive de la distribution des intervalles avec la consonance, et la recherche de clarté et d'indépendance des voix, qui entraînerait une corrélation négative avec la fusion.

Les résultats sont les suivants : la distribution des intervalles est bien corrélée à leur consonance, mis à part pour les intervalles les plus consonants, qui sont aussi ceux avec la plus grande fusion. Sur ces derniers, elle est négativement corrélée au degré de fusion. Ainsi, dans l'écriture de Bach, l'usage des intervalles les plus consonants et avec la plus grande fusion serait guidé par le critère de clarté des voix, tandis que l'usage des autres intervalles serait

guidé par la volonté de maintenir un certain degré de consonance.

Psychohistorique Plus récemment, Parncutt ([Parncutt *et al.*, 2019](#)) corréle différents modèles perceptifs avec la fréquence d’occurrence des triades dans un corpus vocal allant du XIII^e au XIX^e siècle. La démarche est à la fois psychologique et historique, d’où la qualification de « psychohistorique ». Il s’agit d’abord de comprendre si la fréquence d’apparition des triades, à une époque donnée, est guidée par un modèle perceptif ; puis, dans une dimension diachronique, de suivre l’évolution de ces modèles, en rapport avec l’évolution de la pratique musicale.

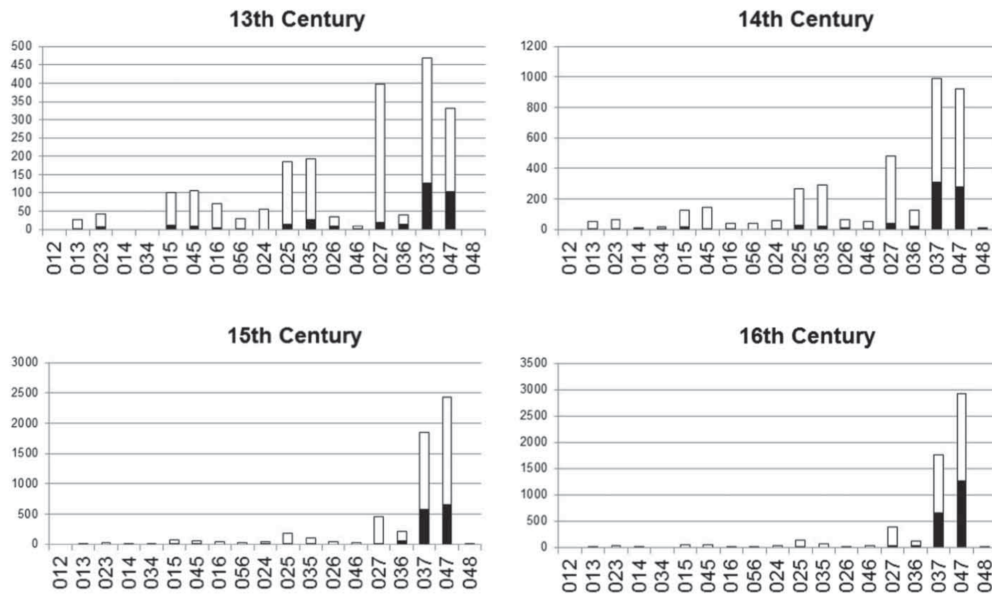


FIGURE 1.1 – Prévalence (nombre d’occurrences) des 19 triades et évolution de la distribution au fil des siècles. Tiré de ([Parncutt *et al.*, 2019](#), p.10). Les labels des accords se réfèrent aux numéros des notes dans l’échelle chromatique à partir de la note de base. L’accord majeur est 047, le mineur 037. Les barres noires correspondent aux accords non préparés, les barres blanches aux accords préparés.

La figure 1.1 montre la distribution statistique des triades selon les siècles. En corrélant ces distributions à des modèles de perception, Parncutt montre que jusqu’au XVI^e siècle la prévalence des accords est guidée essentiellement

par la perception de leurs propriétés acoustiques (rugosité et harmonicité entre autres). À partir du moment où le langage tonal se stabilise, la familiarité prend le relais. L'importance relative de plus en plus grande de l'accord majeur serait due à une sensibilité de plus en plus grande à l'harmonicité.

Discussion Les trois études présentées ici montrent comment des aspects de la perception peuvent guider des choix compositionnels. La démarche de Huron se distingue des deux autres en ce que ces aspects perceptifs sont interprétés en matière d'intention musicale.

Les études ont en commun une approche statistique qui éclaire le matériau de base du vocabulaire harmonique, mais cependant ne dit rien sur la syntaxe, sur l'organisation temporelle de ce matériau. Parncutt avoue lui-même : « *Sur le plan de la théorie musicale, notre approche est superficielle et même triviale ; elle se concentre sur l'image globale et ignore les détails*¹ ».

Se pose dès lors la question de la possibilité de l'influence de la perception non plus uniquement sur le lexique, mais sur la syntaxe d'un langage harmonique. Cette direction nécessite une approche plus fine que l'étude statistique par la fréquence d'occurrence des éléments de langage ; elle nécessite un suivi temporel. L'enjeu pour l'analyse musicale est de comprendre comment les modèles de perception structurent une œuvre, et de pouvoir dégager des logiques perceptives présentes dans une œuvre. Nous exposons dans la partie suivante les difficultés d'une telle démarche.

Les modèles perceptifs comme la rugosité, l'harmonicité ou la fusion sont des modèles spectro-temporels, formulés en termes acoustiques. La capacité de ces modèles d'éclairer la compréhension lexicale autant que syntaxique d'une musique, interprétée en matière d'intention compositionnelle ou de langage, peut provenir autant, a priori, de la formulation acoustique que de la motivation perceptive. En d'autres termes, on peut généraliser l'approche à des modèles acoustiques ne provenant pas de la perception, par exemple motivés par des considérations acoustiques ou mathématiques. La musique, en effet, ne se réduit pas à une sensation. Nous parlons dorénavant de modèle acoustique pour désigner un modèle qui se formule en termes acoustiques, indépendamment de sa motivation initiale.

1. "From a music-theoretical perspective, our approach has been quite superficial and perhaps even trivial, focusing on the big picture and ignoring the detail." Tiré de (Parncutt et al., 2019, p.17).

1.1.3 Suivi temporel par un modèle acoustique : enjeux et difficultés

Les modèles acoustiques peuvent être calculés sur un signal audio, et donc sur un enregistrement d'œuvre. Au moins deux desseins viennent motiver l'usage des modèles acoustiques dans le suivi temporel d'une œuvre :

1. **Approche structurelle** : il s'agit de mettre en évidence une logique acoustique qui sous-tend l'œuvre, et peut être interprétée en matière d'intention compositionnelle ou interprétative. Le modèle acoustique acquiert alors une dimension structurante.
2. **Approche comparative** : calculés sur des enregistrements, les modèles acoustiques, en donnant à visualiser des différences, fournissent des outils de comparaison et d'analyse des interprétations.

Quelques exemples tirés de la littérature vont nous permettre de mettre en avant les difficultés que peuvent poser les deux approches.

Danner ([Danner, 1985](#)), en calculant l'évolution de la rugosité dans *Canon for 3* de Carter, est l'un des premiers à utiliser un modèle psychoacoustique dans l'analyse temporelle d'une œuvre. Bien qu'il ne tire pas de conclusion significative de son analyse, il anticipe avec beaucoup de justesse :

*Parmi les applications musicologiques possibles de ces suggestions théoriques figure la possibilité d'utiliser les mesures acoustiques de dissonance pour caractériser des éléments du style musical.*¹

Depuis lors, les modèles de rugosité sont régulièrement employés pour décrire le suivi temporel d'une musique.

Lalitte se sert des courbes de rugosité pour analyser des œuvres de Varèse ([Lalitte, 2011](#)) et de Ligeti ([Lalitte, 2015](#)), Mamedov les études de Chopin ([Mamedov, 2018](#)). La rugosité constitue un aspect de la dissonance sensorielle, pour cette raison les courbes de rugosité sont souvent appelées courbe de dissonance.

Sethares, dans son livre *Tuning, timbre, spectrum, scale* ([Sethares, 1999](#)), présente différentes utilisations d'un modèle de rugosité. Il propose notam-

1. "Musicological applications beyond these theoretical suggestions include utilizing measures of acoustic dissonance in helping to define elements of musical style." Tiré de ([Danner, 1985](#)), traduction de l'auteur.

ment une analyse comparée de différentes interprétations au clavecin d'une sonate de Scarlatti (cf. fig 1.2).

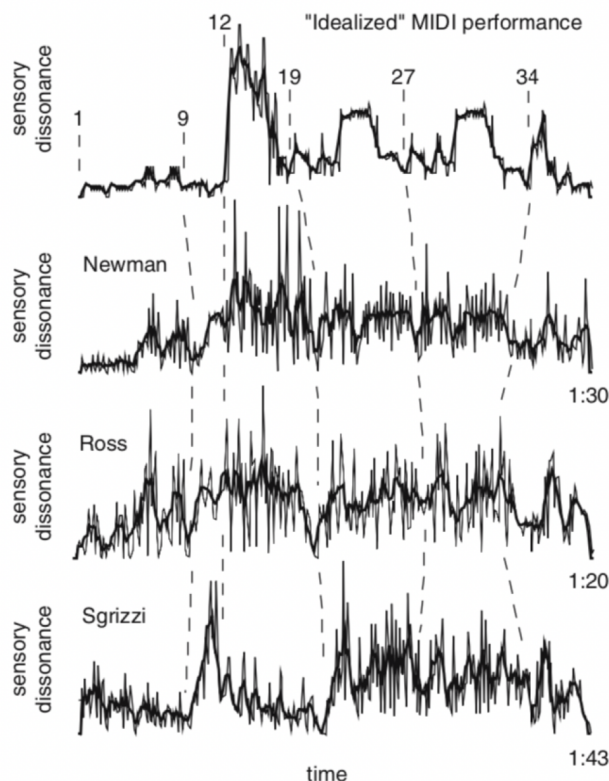


FIGURE 1.2 – Comparée d'interprétations de la *Sonate K.380* de Scarlatti sur des profils de dissonance sensorielle (rugosité). Tiré de (Sethares, 1999, p.223). La première courbe correspond à des sons MIDI, les trois autres à des enregistrements de clavecinistes. Ces courbes sont obtenues avec une fenêtre d'analyse constante de 0.2s, et le calcul de la dissonance - inspiré de (Plomp et Levelt, 1965; Hutchinson et Knopoff, 1979), est précédé d'un algorithme de détection des pics. En haut sont indiqués les numéros de mesure.

Sethares, conscient des difficultés d'interprétation des courbes, prévient :

*Les courbes de dissonance ne doivent pas être interprétées sans précautions*¹.

1. "Dissonance scores must not be viewed carelessly." Tiré de (Sethares, 1999, p.221),

Plusieurs facteurs rendent l'interprétation difficile, Sethares en cite trois :

1. Il n'y a pas de synchronisation des descripteurs avec les notes, le lien avec la partition est difficile à établir. Cela vient de la méthode de calcul. En effet, la rugosité étant calculée sur des fenêtres de 0.2 secondes, elle n'est pas synchronisée avec les notes et fait apparaître des variations de dissonance sensorielle au sein d'une même verticalité.
2. Du fait des fortes variations au sein d'une courbe, il est difficile de distinguer ce qui prend sens, qui est de l'ordre du signe et peut être interprété, de ce qui est de l'ordre de l'artefact dû à l'analyse.
3. Il y a une grande variabilité des courbes de dissonance d'une interprétation à l'autre, et les points communs sont difficiles à repérer. Dès lors se pose le problème de distinguer ce qui relève du choix d'interprétations de ce qui relève de la partition.

Les usages présentés de la rugosité montrent que les applications musico-logiques des modèles acoustiques dans l'analyse d'une œuvre, en particulier des modèles perceptifs, est limité. Cela pour plusieurs raisons :

- La rugosité ne constitue qu'un des aspects du phénomène de dissonance sensorielle ([Harrison et Pearce, 2020](#)). En cela, il n'est pas représentatif du phénomène et encore moins de la perception dans son ensemble. La prise de conscience récente du caractère composite de la consonance et de la dissonance, dues à plusieurs facteurs, n'a pas encore entraîné une approche multi-dimensionnelle chez les musicologues, qui ne prennent en considération généralement qu'un aspect de la perception.
- Les modèles généralement utilisés sont des modèles de dissonance sensorielle, qui caractérisent la perception hors contexte musical. Ils sont cependant calculés en contexte musical, sur une œuvre, sans prise en compte de facteurs tels que la mémoire à court terme ou l'attente.
- Les modèles utilisés s'attachent à l'aspect statique de l'harmonie. Les modèles décrivant les transitions comme l'affinité ou la tension de transition restent confinés dans le domaine des recherches perceptives, sans trouver d'application en analyse musicale. En cela, l'application des modèles acoustiques à l'analyse musicale contraste avec la théorie musicale tonale, qui contient de nombreuses notions dynamiques

traduction de l'auteur.

(conduit des voix, enchaînement cadentiel, modulation...).

Les raisons expliquant l’usage restreint des modèles acoustiques de l’harmonie en musicologie peuvent se résumer en deux points :

1. **Modèles inadaptés** : les modèles utilisés, qu’ils proviennent ou non du champ de la perception, sont parfois inadaptés à un usage musical. Il peut y avoir plusieurs raisons : ils sont parfois de trop bas niveau, ne tiennent pas compte du contexte musical, sont limités à l’approche statique, ou ne permettent à eux seuls d’inférer une logique à l’œuvre.
2. **Difficulté d’interprétation** : Les modèles acoustiques donnent lieu à des représentations, et la possibilité d’interprétation de ces représentations conditionne leur usage en analyse musicale. Or ces représentations sont souvent peu lisibles.

Pour être adapté à l’usage musical, un modèle doit fournir des représentations lisibles et pertinentes.

Tâchons dans la partie suivante de cerner les enjeux du problème de l’interprétation des représentations issues des modèles acoustiques à partir d’un signal audio, telles les courbes de rugosité de la figure 1.2.

1.1.4 Catégories acoustiques et catégories musicales

Il convient d’avoir conscience de la sensibilité de ces courbes comme celles de la figure 1.2 aux paramètres d’analyse et aux choix de représentation. Le modèle de Sethares prend en entrée un signal audio, le représente dans l’espace temps-fréquence par une transformée de Fourier à court terme (TFCT), applique un algorithme de détection de partiels avant de calculer la rugosité par des modèles issus de la littérature sur des fenêtres temporelles de 0,2 s. La courbe grise ainsi obtenue étant peu lisible, car très fluctuante, Sethares la lisse en la moyennant sur une fenêtre de deux secondes. Pour un signal et un modèle de rugosité donnés, il y a autant de courbes de rugosité différentes que de choix de paramètres dans l’analyse.

On en arrive à la question suivante : quels paramètres d’analyse font qu’une courbe est interprétable ?

Catégorisation acoustique Les bons paramètres sont ceux qui aboutissent à une représentation permettant d’identifier des régularités, des similarités et des différences. Cette condition apparaît comme nécessaire à

l'interprétabilité d'une courbe. La détection des régularités met en avant certains comportements (augmentation brusque de la rugosité, petites oscillations de la rugosité...), tandis que les similarités permettent de systématiser ces comportements, et les différences de les discriminer. Les régularités, les similarités et les différences sont impliquées dans le processus de catégorisation du signal. Même si le modèle acoustique considéré incorpore des aspects perceptifs, physiologiques, psychologiques ou cognitifs, les catégories qui découlent de cette catégorisation sont avant tout de nature acoustique, car calculées à partir du signal audio.

Recherche d'unités sémiotiques Une fois les unités repérées à partir de la représentation, il reste à interroger leur potentiel sémiotique, à savoir si elles font signe. Cette question est intimement liée à l'usage que l'on fait du modèle acoustique, et au type d'analyse attendu. En effet, pour reprendre Saussure ([Saussure, 1916](#)), la signification est une relation de correspondance entre un signifiant (en l'occurrence une catégorie acoustique déduite de la représentation) et un signifié. La nature de ce dernier peut varier selon le but de l'analyse. Il peut être de l'ordre de l'intention, compositionnelle ou interprétative, se référer à la perception de l'auditeur, ou encore aux catégories symboliques d'une partition. Une condition nécessaire dans l'identification des signes est que les signifiants potentiels, autrement dit les catégories acoustiques, soient suffisamment distincts. Rastier ([Rastier, 1999](#)) l'exprime en termes de catégories linguistiques :

*La signification n'est pas (ou pas seulement) constituée par la référence à des choses, ou par l'inférence entre concepts, mais aussi et d'abord par la différence entre des unités linguistiques*¹.

L'appréhension des significations passe par un aller-retour constant entre la représentation fournie par le modèle et l'objet étudié (le signal audio) ou sa représentation symbolique (la partition). Cet aller-retour fait partie intégrante de la démarche interprétative, en constitue la première étape. Rastier ([Rastier, 1994](#)) note à ce propos :

*Tout signifié résulte d'un parcours interprétatif : il n'est ni découvert, ni inventé, mais constitué dans une interaction entre le texte et l'interprète*².

1. Tiré de ([Rastier, 1999](#)).

2. Tiré de ([Rastier, 1994](#), p.329).

Les significations ne préexistent pas à l'interprète, dont le point de vue est nécessairement orienté :

*L'identification d'unités sémiotiques repose toujours sur une lecture située et circonstanciée. Elle est le fruit d'une grille de lecture choisie en amont et d'une sélection des structures signifiantes en aval*¹.

Le rôle du contexte dans la sémiose Les significations sont dépendantes de l'objet d'étude (signal ou partition), et elles n'ont de pertinence que dans le contexte qu'elles délimitent. En effet, un signe n'existe qu'au sein d'un système sémiotique, autrement dit un système de références, d'inférences et de différences, lié à l'objet d'étude.

En guise d'illustration, considérons les courbes de dissonance dessinées par Helmholtz dans sa *Théorie physiologique de la musique* (Helmholtz, 1863, p.247). Ces courbes ont une importance historique majeure, puisqu'elles inaugurent une nouvelle approche de la théorie de la musique². Helmholtz considère que la dissonance est due aux battements entre partiels proches. En quantifiant la dissonance entre deux partiels et en sommant les contributions des partiels qui interagissent, il représente le profil de la dissonance des intervalles en fonction du nombre d'harmoniques dans le spectre. La figure 1.3 montre les courbes de dissonance sur deux octaves. Sur une courbe donnée, les creux correspondent aux consonances.

Avec 3 harmoniques, le modèle d'Helmholtz ne permet pas de distinguer la tierce mineure de la tierce majeure, alors que le rajout des harmoniques 4 et 5 fait apparaître un minimum local de rugosité au niveau de la tierce majeure (de rapport 5/4).

On voit par cet exemple que l'objet d'étude influence, si ce n'est détermine, autant la catégorisation acoustique qui crée les signifiants que la mise en place des significations. De même, dans l'étude d'œuvres musicales, les éléments signifiants sont dépendants du corpus, et leur signification est conditionnée par le corpus.

1. Tiré de (Guillotel-Nothmann, 2020)

2. Nous aurons l'occasion d'y revenir dans ce travail.

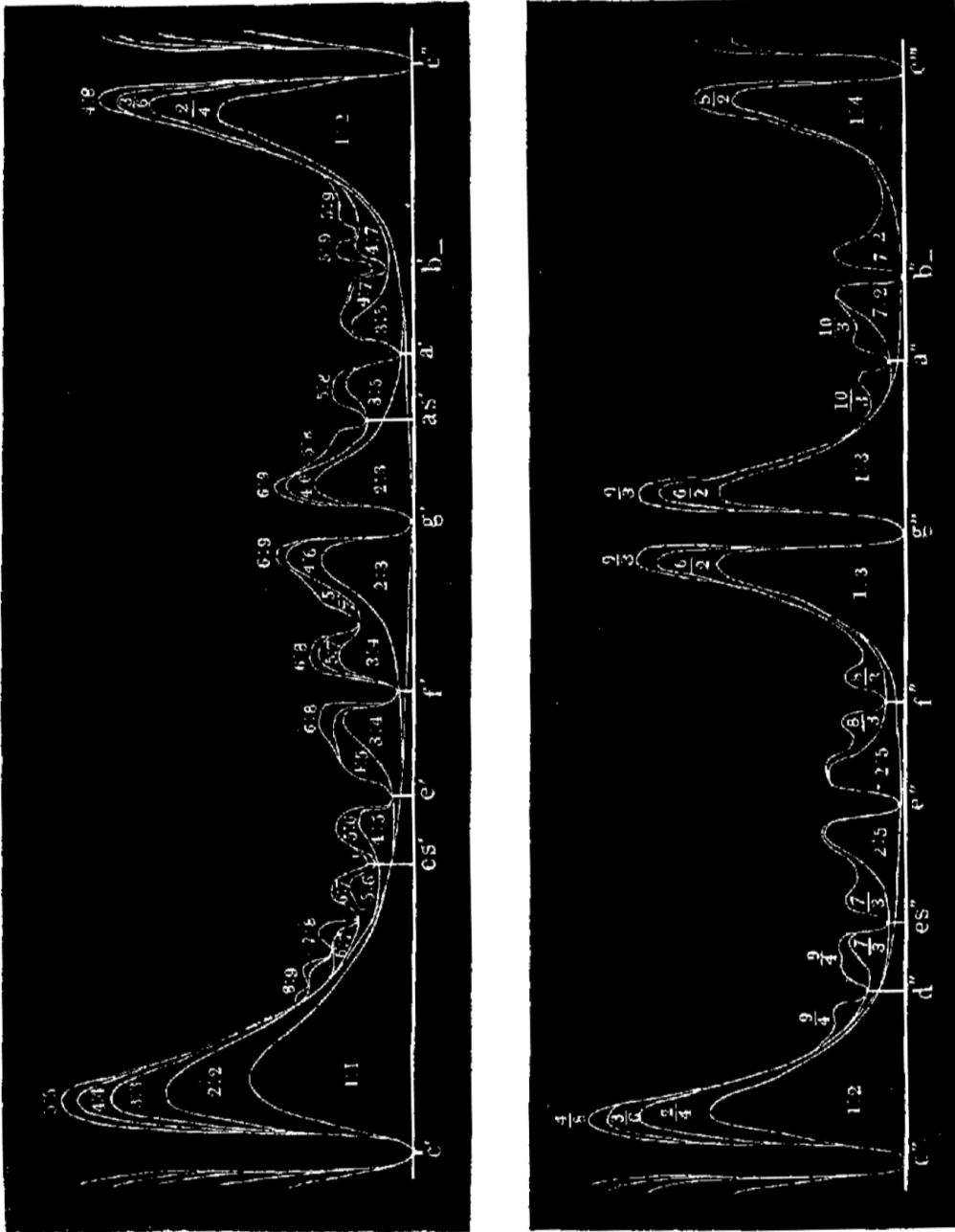


FIGURE 1.3 – Courbe historique d’Helmholtz de dissonance des intervalles, tiré de (Helmholtz, 1863, p.247). Helmholtz modélise ici la rugosité due aux battements entre partiels. L’axe des abscisses représente les intervalles, qui s’étendent sur deux octaves, et l’axe des ordonnées la dissonance (ou rugosité). Les courbes 1 à 5 sont calculées avec respectivement 1, 3, 5, 7 et 9 partiels.

Relation signifiante et parcours interprétatif : deux axes pour le musicologue Malgré leur dépendance à l’objet d’étude, les significations restent une propriété attachée aux signes, et non une propriété globale de l’objet d’étude (oeuvre, corpus, suite d’accords...). Il apparaît nécessaire à ce niveau de distinguer *signification* et *sens*, en reprenant les termes de Rastier (Rastier, 1994). La signification est une propriété assignée aux signes, résultant d’un processus de décontextualisation, tandis que le sens est une propriété du texte (l’objet d’étude), qui suppose une contextualisation maximale.

Comme le souligne Guillotel-Nothmann, la signification est un ensemble de « *valeurs essentiellement statiques et structurales qui relèvent de la forme et découlent de la position du signe au sein du système sémiotique* », et le sens un « *parcours interprétatif dynamique susceptible de transcender le texte [l’objet d’étude]* »¹. Cette distinction invite à distinguer à leur tour deux axes de recherche possibles pour le musicologue qui souhaite utiliser des modèles acoustiques pour analyser l’harmonie.

1. Le premier axe, centré sur les significations, consiste en une description formelle de l’organisation interne du matériau harmonique. Il s’agit ici d’adopter le point de vue acoustique fourni par les modèles pour décrire les structures des espaces harmoniques (espace d’accords ou de transitions entre accords). Les espaces harmoniques étudiés peuvent couvrir des degrés différents de généralité ou de spécificité. Ils peuvent être :
 - Les plus généraux possibles, sans ancrage historique ou géographique particulier (par exemple l’ensemble des accords à trois notes sur l’ambitus d’une octave, avec une distribution continue des fréquences).
 - Situés historiquement, géographiquement ou stylistiquement, sans pour autant être reliés à une oeuvre (notamment l’ensemble des accords ou classes d’accords au sein d’un système musical, par exemple le système tempéré avec une subdivision des octaves en 12 demi-tons).
 - Reliés à une oeuvre ou à un contexte musical particulier, et concerner l’ensemble des accords présents, ou l’ensemble des transitions. La dimension temporelle vient alors s’ajouter à la dimension spatiale.

1. Tiré de (Guillotel-Nothmann, 2020).

Dans cette démarche descriptive, la question de la représentation est cruciale, puisque c'est à travers elle que se mettent en place les unités sémiotiques. Le point de vue orienté de l'analyste se manifeste par les choix concernant l'espace de départ et la représentation des données acoustiques : type de représentation, type d'axes (linéaires, logarithmiques), attribution des couleurs... L'espace des accords à trois notes sur l'ambitus d'une octave peut être représenté par une surface continue où l'axe des abscisses figure l'intervalle inférieur et l'axe des ordonnées l'intervalle supérieur (Mach, 1866). L'ensemble des classes d'accords au sein d'un système tempéré peut être représenté par le potentiel combinatoire d'Estrada (Estrada, 1994). Les accords d'une œuvre, quant à eux, sont visualisables directement sur la partition, quand celle-ci existe. Les modèles acoustiques viennent structurer ces espaces, ce qui peut se faire concrètement par un code couleur, une mise en relief, des valeurs numériques ou des annotations.

Cet axe descriptif, au-delà de son aspect purement théorique, peut notamment servir dans une démarche de création musicale. En effet, la structuration acoustique du matériau harmonique est susceptible de guider le compositeur dans son exploration harmonique.

2. Le second axe de recherche concerne le sens, en tant que « *parcours interprétatif dynamique susceptible de transcender le texte.* » Il s'agit ici d'analyser une œuvre ou un extrait d'œuvre et de bâtir une interprétation dépendant du point de vue particulier de l'analyste. L'utilisation des modèles acoustiques de l'harmonie peut constituer un aspect parmi d'autres de l'analyse, qui sera complété par d'autres approches. Le choix du modèle est ici crucial, il doit servir la pertinence de l'analyse.

Catégories acoustiques vs catégories symboliques L'utilisation des modèles de perception de l'harmonie en musicologie pose le problème de l'articulation entre données qualitatives et données quantitatives. En les appliquant à la description des espaces d'accords ou à l'analyse musicale, le défi consiste à « *s'approcher, jusqu'à un certain degré, d'objets culturels à partir d'une interrogation quantitative des objets* »¹. Guillotel-Nothmann ajoute :

1. *Ibid.*

*Le signal audio ou vidéo est, au contraire, complètement vierge de ces catégories [les catégories symboliques] et ne comporte que des données quantitatives dont le séquençage s'effectue a priori indépendamment de considérations qualitatives. Il présente l'avantage d'une exploration d'éléments qui se seraient soustraits aux catégories symboliques, mais pose le défi considérable d'une catégorisation sémiotique effectuée a posteriori soit par le dispositif informatique, soit par l'analyste*¹.

Les catégories symboliques dont il est question sont celles présentes sur la partition, ou qui peuvent en être déduites : note, hauteur, accord, durée, rythme, tonalité, nuance... La catégorisation qui découle des modèles de perception, est, quant à elle, une catégorisation acoustique, quantitative du signal. Puisque les deux types de catégories coexistent, la question se pose du lien entre elles, et de leurs rôles respectifs dans la démarche du musicologue. Guillotel-Nothmann, après avoir pointé du doigt les disparités en matière de méthodologie et d'objectifs analytiques des approches symbolique et signal, souligne leur complémentarité :

*Il me paraît important d'insister sur la nature plus complémentaire qu'antagoniste des rapports entre données acoustiques (quantitatives) et représentations symboliques (qualitatives)*².

Sethares consacre son livre *Tuning, timbre, spectrum, scale* (Sethares, 1999) à montrer l'interdépendance des données acoustiques avec les catégories musicales. En considérant des sons et des échelles non traditionnels (sons inharmoniques, micro-tons), il met en évidence le lien fort qui unit le spectre à l'échelle des notes et aux accords utilisés.

Cette complémentarité et cette interdépendance des catégories physiques et symboliques incite à aller plus loin dans l'étude de ces rapports, à comprendre leur nature.

Le premier axe dégagé plus haut, la description formelle de l'organisation interne du matériau harmonique, pose frontalement la question du rapport entre les catégories acoustiques et les catégories symboliques. En effet, il s'agit de représenter des données acoustiques sur des espaces d'accords organisés

1. *Ibid.*

2. *Ibid.*

selon des critères relatifs à la structure intervallique des accords¹ ou des critères combinatoires².

Les catégories symboliques sont étagées selon plusieurs degrés d'abstraction, ce qui donne lieu à une grande diversité potentielle de rapports avec les catégories acoustiques. Prenons l'exemple de l'ensemble des triades dans le système tempéré à 12 demi-tons. La description la moins abstraite consiste à associer à chaque accord les hauteurs des trois notes. Si l'on se restreint à une octave, cela fait $\binom{12}{3} = 220$ accords. En regroupant les accords transposés l'un l'autre, on obtient $\binom{11}{2} = 55$ classes d'accords. Si on regroupe à partir de là les accords dont l'un est un renversement de l'autre (on dit qu'ils ont la même forme normale), on réduit le nombre à 19 classes d'accords. On peut monter encore en abstraction en regroupant les accords selon leur contenu intervallique, sans s'occuper de l'ordre des intervalles (on dit qu'ils ont la même forme identitaire), cela ne fait plus que 12 classes d'accords.

Le lien entre les catégories acoustiques et symboliques peut être questionné à chaque niveau d'abstraction des catégories symboliques, et si les modèles de perception peuvent être appliqués au plus bas niveau d'abstraction, ils ne peuvent pas forcément l'être aux plus hauts niveaux d'abstraction, du fait de la multiplicité des réalisations des classes d'accords.

Le premier type de relations possibles entre les catégories acoustiques et les catégories symboliques sont des relations de correspondance. Il s'agit alors de caractériser les accords, les classes d'accords ou les enchaînements harmoniques par un ou plusieurs modèles acoustiques. C'est la visée de tout classement des intervalles en fonction du degré de dissonance. Cette démarche est appliquée sur le vocabulaire harmonique spécifique à un langage (Cong, 2016), sur des accords choisis pour leurs propriétés (Kuusi, 2010), et de manière systématique sur l'ensemble des accords à trois sons (Danner, 1985) et l'ensemble des classes d'accords dans une échelle donnée.

Cong (Cong, 2016) cherche à décrire par des modèles perceptifs les types d'accords les plus utilisés dans la musique tonale. Ces résultats sont représentatifs de la démarche :

Les accords contenant un ou deux demi-tons ont la plus grande consonance sensorielle, les accords suspendus de quarte [du type do-fa-sol] ont la plus grande "ambiguïté" [ambiguïté dissonance],

1. Voir la représentation plane des triades page 291.

2. Voir la représentation des classes normales de triades page 331.

*les accords diminués ont la plus grande "obscurité" [gloom dissonance], les accords augmentés ont la plus grande tension [tension dissonance], et, de tous les types d'accords, les minimum des quatre types de dissonances sont atteints sur les accords majeurs*¹.

La recherche de caractéristiques acoustiques trouve son intérêt principalement lorsque le modèle n'est pas *ad hoc*, construit spécialement dans ce but, mais lorsque qu'il a des motivations extérieures. Le modèle de tension de Cook (Cook, 2009), par exemple, est construit pour reconnaître les triades dont les intervalles inférieur et supérieur sont égaux, c'est à dire, dans un contexte tonal, essentiellement les accords non résolus. Ce modèle se justifie essentiellement lorsqu'il est combiné à un modèle de dissonance pour modéliser l'instabilité.

La caractérisation des catégories symboliques par des données acoustiques est une voie possible dans l'exploration des liens entre catégories acoustiques et symboliques, mais ces liens ne se réduisent pas à cette caractérisation. Si les modèles de perception n'étaient capables que de retrouver les catégories symboliques présentes dans la partition, alors leur intérêt pour l'analyse musicale serait très limité, puisqu'ils n'apporteraient rien de plus que ce qui est déjà connu dans la partition.

Variabilité aux paramètres sonores Les catégories acoustiques comme les classes d'accords possèdent une multitude de réalisations sonores, dont l'ensemble forme ce que nous appelons leur "univers sonore". Les différentes réalisations diffèrent par les paramètres timbraux, les nuances, le registre, la disposition des notes, les doublures éventuelles, la durée... Les propriétés acoustiques des objets dépendent de ces paramètres. Ainsi, toute démarche visant à expliciter les liens entre catégories symboliques et acoustiques se doit de tenir compte de cette multiplicité, et de tester la rigidité des catégories acoustiques vis-à-vis de ces paramètres.

1. "The secondary chord structures (simultaneous chords containing one or two semitone intervals) have the highest level of sensory dissonance; the suspended 4th chords have a higher level of ambiguity dissonance; diminished chords have the highest level of gloom dissonance (followed by minor structures); the augmented chord structures contain the highest level of tension; and, of all chord structures, all the dissonance types are at their lowest level in the major chords.", tiré de (Cong, 2016, p.2), traduction de l'auteur.

De plus, la dépendance des modèles à ces paramètres peut servir à l'étude des tempéraments¹ ou des timbres, par exemple dans une démarche historiquement informée, et éclairer des questions liées aux choix d'orchestration.

Ces modèles fournissent également des outils de comparaison des dispositions d'accords, ce qui peut être utilisé pour confronter différentes réalisations d'un enchaînement harmonique en faisant varier la disposition, l'étalement des accords, le registre, les doublures. Les différenciations qui découlent des modèles peuvent alors être mises en relation avec les règles énoncées dans les traités, et peuvent éclairer sous un angle acoustique des choix d'écriture. Au-delà du lien entre catégories acoustiques et symboliques, il s'agit ici d'étudier le lien entre catégories acoustiques et catégories musicales.

1.1.5 Périmètre et limitation de l'approche acoustique

Il convient de nuancer la portée de l'approche acoustique dans l'analyse musicale en cernant son périmètre. L'approche acoustique décrit l'organisation acoustique du matériau musical. Dans le cas de l'harmonie, elle s'attache au vocabulaire harmonique de base : les intervalles, les accords, les transitions, les enchaînements harmoniques. L'univers harmonique peut être vu comme un éventail des possibilités, que les modèles acoustiques viennent structurer. Mis en rapport avec une œuvre ou des traits de langage, ils sont capables d'éclairer l'usage musical, mais en aucun cas de le justifier ou même de le légitimer. Autrement dit, ils ne dictent pas, mais suggèrent. Dans un langage donné, ils peuvent à la rigueur faire office de conditions, mais c'est alors l'usage qui détermine la condition, et non l'inverse. Même lorsque la démarche compositionnelle se fonde explicitement sur des données acoustiques, comme dans le spectralisme, c'est l'intention du compositeur qui guide leur emploi.

Les approches perceptives, acoustiques et combinatoires sont d'ailleurs à ce niveau-là sur le même plan : elles décrivent le matériau musical avec des critères perceptifs, acoustiques ou combinatoires, et ces descriptions peuvent être utiles à la compréhension du phénomène musical, mais elles ne le constituent pas, au sens d'une réductibilité. Le son n'est pas la musique, de même

1. Pour une application sur l'étude des tempéraments, voir <https://alpof.wordpress.com/2015/10/11/consonance-calculations-2-tunings/>. L'auteur, Popoff, calcule la dissonance de tous les accords majeurs et mineurs sur 96 tempéraments, et analyse ces données par une analyse en composantes principales (PCA), ce qui lui permet de regrouper les tempéraments selon leur ressemblance. Voir aussi (da Fontoura Costa, 2019).

que la proportion n'est pas la musique, et la perception sonore n'est pas la musique.

La beauté, le sens musical, le jugement esthétique sont d'un autre ordre. Helmholtz l'a bien compris lorsqu'il distingue bien-être des sens, objet de la science, au beau esthétique, ou « *principe de style* ». Le premier, pour Helmholtz, ne fournit au second que des moyens, pas des fins. Le philosophe Charles Lalo considère que seule la dimension sociale confère une valeur esthétique à la musique :

*Les lois sociologiques nous semblent dominer la nature psychologique des éléments de l'art musical. Seules elles peuvent leur conférer le caractère esthétique, dont, en dehors de la vie sociale, ils n'ont par leur nature individuelle et abstraite que la virtualité, la possibilité indéterminée*¹.

Francès, à la question de savoir comment l'analyse de la perception de la musique permet de comprendre le jugement esthétique, répond :

*On peut dire que ce jugement suppose la perception de toutes ces qualités, de toutes ces significations, mais qu'il ne s'identifie à aucune en tant qu'il est l'énonciation d'une valeur et non la constatation d'une donnée*².

Le jugement esthétique n'est pas la constatation d'une donnée perceptive, pas plus que d'une donnée acoustique ou mathématique. Pour Francès, l'irréductibilité de la musique à la description de ses éléments, de ses données, provient de l'action réciproque de l'ensemble sur les éléments, action qui change leur statut :

*Les éléments sont à la base du système, mais celui-ci, une fois expérimenté, devient chez l'auditeur une seconde nature et qualifie les éléments*³.

Le jugement esthétique proviendrait de cette interaction réciproque.

1. Tiré du livre *Esquisse d'une esthétique musicale scientifique* de Charles Lalo ([Lalo, 1908](#), p.323).

2. Tiré du livre *Perception de la musique* de Robert Francès ([Francès, 1984](#), p.383).

3. *Ibid.*, p.382.

1.2 Problématique

L'objectif de cette thèse est de proposer des modèles acoustiques adaptés à l'usage musicologique, à même de rendre compte du phénomène harmonique.

Détaillons les enjeux de la problématique, à la lumière de l'introduction qui précède.

Un modèle acoustique prend en entrée un signal audio, ou, à défaut, une partition associée à un modèle de spectre, et fournit des représentations. Ces dernières sont adaptées à l'usage musicologique lorsqu'elles sont à la fois lisibles et signifiantes.

Les modèles doivent pouvoir rendre compte du phénomène harmonique dans ses dimensions à la fois statique et dynamique, ainsi qu'être sensibles aux aspects comme le timbre, le nombre de notes, la disposition ou encore les nuances, qui apparaissent comme essentielles dans l'appréciation de la musique.

L'univers harmonique, considéré dans sa généralité comme l'ensemble des éléments constitutifs de l'harmonie (espaces d'accords, de classes d'accords, de transitions, de réalisations sonores), peut être vu comme l'ensemble du vocabulaire harmonique potentiel. Il apparaît important de dégager les implications et conséquences des modèles acoustiques sur le vocabulaire harmonique.

Appliqués à un langage donné, les modèles acoustiques permettent d'éclairer des éléments de langage. De même, dans une optique d'analyse musicale, ils mettent en avant des logiques acoustiques à l'œuvre capables d'éclairer des choix d'écriture.

1.3 Méthodologie

Nous introduisons dans ce travail la notion de descripteur harmonique, type particulier de descripteur audio qui traite des interactions spectrales entre les notes. Nous proposons deux implémentations des descripteurs harmoniques, une implémentation signal prenant en entrée un signal audio, et une implémentation symbolique prenant en entrée une partition et un modèle de spectre. Par certains aspects de l'implémentation, les deux approches se situent à l'interface entre le signal et le symbolique.

L'approche symbolique a l'avantage, de par sa maniabilité, d'offrir une vue globale sur les espaces harmoniques. De plus, pour des modèles de spectre bien choisis, elle permet d'abstraire le calcul des descripteurs harmoniques aux classes d'accords, et de trouver par-là des applications aux classifications d'accords. L'approche signal a quant à elle l'avantage d'être plus réaliste que la première, notamment en ce qui concerne le paramètre timbral.

Dans l'optique d'un usage musicologique, le calcul des descripteurs harmoniques est réalisé sur les verticalités, autrement dit à l'échelle temporelle des notes. Ce paradigme permet une lecture parallèle des représentations issue des descripteurs harmoniques avec la partition, mais implique des contraintes fortes sur les types de sons employés, et met de côté le suivi temporel de phénomènes comme le vibrato.

Pour satisfaire au mieux à ces contraintes tout en assurant une certaine réalité musicale, une partie des exemples traités utilisent des sons d'orgue tenus, issus d'une synthèse à partir de sons échantillonnés. Deux des analyses musicales abordées prennent en entrée des enregistrements.

1.4 Annonce du plan

La première partie de la thèse s’attache à replacer les descripteurs harmoniques dans une perspective historique. Le problème de l’origine de la consonance et de la dissonance est central dans toute l’histoire de la théorie musicale occidentale en ce qu’il questionne les fondements de l’harmonie. Le chapitre 2 dégage les courants de pensée traversant l’histoire de la consonance et de la dissonance. Ce faisant, il met en avant les enjeux épistémologiques des descripteurs harmoniques. Le chapitre 3 dresse une revue des recherches actuelles sur la perception de l’harmonie, et montre que cette même question de la consonance et de la dissonance reste au cœur des débats actuels, dominée maintenant par une volonté de synthèse.

La deuxième partie traite des descripteurs harmoniques. Le chapitre 4 introduit cette notion dans le contexte des descripteurs audio, tandis que le chapitre 5 détaille leur implémentation dans ce travail. Le chapitre 6 présente les descripteurs statiques et dynamiques utilisés par la suite. Dans un but pédagogique, des illustrations sont fournies, avec des accords et des enchaînements hors contexte musical. Le chapitre 7 expose les contraintes sur les sons et propose des normalisations visant à rendre les représentations cohérentes avec l’intuition musicale.

La troisième partie s’attache à mettre en avant la manière dont les descripteurs harmoniques structurent le matériau harmonique. Cette partie adopte l’approche signal. Le chapitre 8 traite des espaces continus d’accords et de transitions entre accords. Les descripteurs harmoniques fournissent de véritables cartographies de ces espaces, révélant leur structure acoustique. Dans le chapitre 9, les descripteurs harmoniques sont calculés sur les classes d’accords, via l’utilisation de spectres particuliers, les spectres de Shepard. Cela permet de superposer un point de vue acoustique aux considérations combinatoires à l’origine des classifications d’accords déjà existantes. Une description géométrique des espaces de classes d’accords est proposée, ce qui amène à tirer parti des descripteurs statiques pour caractériser acoustiquement les identités de la classification de Julio Estrada, connue sous le nom de « potentiel combinatoire des intervalles ». Les descripteurs dynamiques sont quant à eux employés pour caractériser les transitions dans le Tonnetz. Dans les deux applications, l’apport acoustique est matérialisé par une « couleur harmonique » venant colorier les représentations.

La dernière partie de la thèse est dédiée à l’apport des descripteurs harmoniques en analyse musicale. Le corpus d’étude est constitué de trois œuvres de

la même époque faisant appel à des langages radicalement différents : *Fratres* d'Arvo Pärt (1977), la *Pièce pour orgue n°2* d'Alfred Schnittke (1980) et les *Cartas Celestes I* d'Almeida Prado (1974). Les descripteurs harmoniques se montrent capables de révéler des logiques acoustiques à l'œuvre dans l'écriture, non accessibles par les méthodes classiques d'analyse.

1.5 Définitions

Avant d'exposer la problématique de cette thèse, il paraît nécessaire de définir les termes dont nous aurons un usage courant : note, intervalle, accord, harmonie.

1.5.1 Définitions de base

Note Une note peut être de nature symbolique ou sonore. Une partition, tout du moins dans l'acception classique, est un agencement spatial et temporel de notes, où la note désigne alors une abstraction caractérisée par une fréquence, une durée, éventuellement une nuance et ou un mode de jeu. Dans le protocole midi, une note a une hauteur, une vitesse (équivalent de la nuance), un début et une fin.

On désigne également par notes les différentes réalisations sonores de ces abstractions. Un musicien joue des notes, autant lorsqu'il interprète une partition que lorsqu'il n'a pas cet intermédiaire symbolique. Lorsqu'une note désigne un son, celui-ci possède nécessairement une fréquence fondamentale f_0 . On appelle son simple un son sinusoïdal, et son complexe un son non sinusoïdal. La transformée de Fourier décompose tout son en une somme de sons simples, appelés partiels. On dit que le spectre d'un son est harmonique lorsque les partiels sont des multiples de la fréquence fondamentale.

Dans la suite nous désignons par note à la fois l'objet abstrait, caractérisé par une fréquence fondamentale et une durée, qu'une réalisation sonore de cet objet abstrait.

Le délimitation des notes à partir d'un son est un problème complexe d'ordre à la fois acoustique et psychologique. La notion de fusion, introduit par Carl Stumpf ([Stumpf, 1883, 1890](#)), caractérise la tendance d'un son à être perçu comme une seule note. Sous certaines conditions - rapport harmonique des partiels, synchronicité des phases, similarité des profils temporels des partiels ([Schneider, 1996](#)), un son composé de plusieurs notes peut être perçu comme une seule note, de même qu'un son composé d'une note peut être perçu comme composé de plusieurs notes, notamment pour les sons d'orgue.

La hauteur d'une note peut être exprimée en *Hertz*, ou par son nom dans le système tempéré. Nous notons les notes en minuscule, et *do4* représente le *do* central du piano à 361 *Hz*.

Intervalle Un intervalle désigne tantôt une combinaison, simultanée ou successive, de deux notes, tantôt la distance qui sépare deux notes sur l'axe logarithmique des fréquences.

*Ce n'est qu'à l'intérieur d'une relation qu'une note devient un phénomène musical en soi, au lieu d'un simple fait acoustique.*¹

Accord Un accord est une combinaison simultanée de notes. La cardinalité d'un accord est le nombre de notes de cet accord. Nous utilisons la nomenclature suivante : un accord à 2 notes est appelé une dyade, de 3 notes une triade, de 4 notes un tetrachorde. Nous notons les accords avec des majuscules, de manière à les différencier des notes, notées en minuscule. À côté du nom de la note fondamentale de l'accord, *M* désigne un accord majeur, *m* un accord mineur. Par exemple *Do M* ou *Réb m*.

Verticalité Dans un contexte musical, partition ou signal audio, une verticalité est la combinaison de toutes les notes présentes délimitée temporellement par deux événements harmoniques successifs ; un événement harmonique désignant l'apparition ou la disparition d'une note. En présence d'une succession d'accords, les verticalités correspondent aux accords.

Harmonie Le terme harmonie peut se référer à plusieurs réalités musicales. Carl Dahlhaus, dans son ouvrage *La Tonalité harmonique* (Dahlhaus, 1967), distingue cinq acceptions du terme harmonie, qui sont essentiellement des distinctions historiques.

1. Au début du Moyen-Âge, l'harmonie se réfère à la succession de notes au sein d'une mélodie. On distingue le concept d'*harmonie*, qui se réfère à la mélodie, du concept de *concorde*², qui se réfère à un agrégat de notes.
2. À partir du XIIIe siècle, l'harmonie peut désigner également une combinaison simultanée de deux ou trois notes. Dans *De Harmonia* de Gafurius (1518) (Gafurius, 1518), le terme *harmonie* est réservée aux combinaisons simultanées de trois notes (triades), et le terme *consonance* aux combinaisons de deux notes (dyades).

1. Tiré de (Dahlhaus, 1967).

2. Nous utiliserons dans ce travail le terme « concordance » dans une autre acception.

3. À partir du XVe siècle, l'harmonie se met à désigner des progressions de dyades. Dahlhaus fait remonter à 1412 un tel usage, dans le *Traité de Contrepoint* de Prosdocimo de Beldemandi (de Beldemandi, 1412).
4. Au XVIe siècle, notamment chez Zarlino (Zarlino, 1558), l'harmonie a une acception très large, qui englobe de nombreux aspects de la musique polyphonique : l'enchaînement de plusieurs notes au sein d'une même voix, la combinaison simultanée de deux notes dans une dyade, la combinaison des dyades au sein d'une triade, ainsi que les relations rythmiques entre les différentes notes et les différentes voix. En ne distinguant pas les combinaisons successives des combinaisons simultanées des notes, Zarlino voit dans l'ordre simultané ou successif des mêmes notes la manifestation de la même harmonie.
5. Au XVIIIe siècle, l'harmonie désigne une succession d'accords au sein d'une progression tonale. D'Alembert écrit dans ses *Éléments de Musique Théorique et Pratique suivant les principes de M. Rameau* (d'Alembert, 1752) :

*On appelle accord le mélange de plusieurs sons qui se font entendre à la fois. L'harmonie est proprement une suite d'accords qui en se succédant flattent l'organe.*¹

La notion d'accord est donc intrinsèquement une notion tonale, car liée au système tonal. Les accords sont décrits comme un empilement de tierces, et c'est leur structure qui définit la progression tonale.

Dans la suite, nous désignons par harmonie, ou encore phénomène harmonique, l'ensemble des phénomènes sonores et des relations formelles impliquant l'interaction simultanée ou successive de plusieurs notes. Cette définition large englobe les différentes acceptions du terme identifiées par Dahlhaus. Nous mettons cependant de côté les relations rythmiques pour nous concentrer sur les interactions spectrales entre les notes. Avec cette définition, l'étude de l'harmonie peut se scinder en deux branches :

1. Étude statique : c'est l'étude des accords isolés et des interactions simultanées entre notes.
2. Étude dynamique : c'est l'étude des enchaînements des notes ou des accords.

1. Tiré de (d'Alembert, 1752).

Nous appelons espace harmonique tout espace de phénomènes harmoniques. On peut distinguer les espaces harmoniques abstraits, constitués d'entités formelles : espaces d'accords, de transitions, classifications d'accords, des espaces harmoniques sonores, par exemple l'univers sonore d'un enchaînement harmonique, obtenu en faisant varier le timbre.

1.5.2 Classes d'accords

Il est courant en musique de regrouper des notes ou des accords en classes d'équivalence. Introduisons ici le vocabulaire approprié. Les notions d'ensemble de hauteurs, de classe d'intervalles, de classes normale et première et de formes normale et première sont tirées de la *Set Theory* de Forte ([Forte, 1973](#)). La notion d'identité vient de la théorie d'Estrada ([Estrada, 1994](#)).

Hauteur On appelle « hauteur » l'ensemble des notes à distance d'une ou de plusieurs octaves l'une de l'autre. La considération des hauteurs présuppose l'équivalence des octaves, et permet de s'abstraire du registre. Dans la suite du travail, l'utilisation des hauteurs sous-entend l'emploi de sons harmoniques, assurant ainsi l'équivalence des octaves. L'échelle chromatique, fondée sur une subdivision de l'octave en douze demi-tons, contient 12 hauteurs, de *do* jusqu'à *si*.

Dans la *Set Theory*, la hauteur est appelée *pitch-class*, traduit en français par « classe de hauteur ». Cette terminologie, qui amène à considérer des « classes d'ensembles de classes de hauteurs », est trop lourde à notre avis, et nous préférons le terme « hauteur » pour désigner les *pitch-class*. Dans la littérature musicologique, la notion de hauteur est parfois nommée « hauteur nominale », étant donné que dans l'échelle chromatique les hauteurs se confondent avec les 12 noms de notes. Ce travail dépassant le cadre de l'échelle chromatique, le terme « hauteur » est jugé préférable.

L'ensemble des hauteurs est usuellement représenté par le cercle chromatique, comme sur la figure [1.4](#).

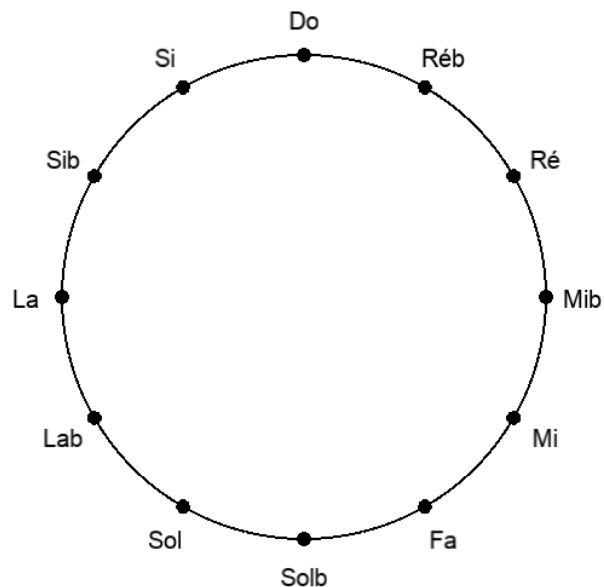


FIGURE 1.4 – Cercle des hauteurs. Les subdivisions correspondent à l’échelle chromatique.

Ensemble de hauteurs On appelle ensemble de hauteurs tout ensemble non ordonné de hauteurs. À un accord donné on peut associer un ensemble de hauteurs en quotientant par la relation d’octave, mais un ensemble de hauteurs peut être obtenu à partir d’accords différents. On peut ainsi voir un ensemble de hauteurs comme l’ensemble des accords qui lui sont associés. En guise d’exemple, les accords $do_4-mi_4-sol_4$ et $mi_3-do_4-sol_4-do_5$, bien qu’ayant des cardinalités différentes et étant chiffrés différemment, correspondent au même ensemble de hauteurs $\{do, mi, sol\}$. La cardinalité d’un ensemble de hauteurs est le nombre de hauteurs qu’il contient.

Par abus de langage, on parlera parfois d’accord pour désigner l’ensemble de hauteurs associé.

Relations de transposition et d'inversion

Transposition Deux ensembles de hauteurs sont transposés l'un de l'autre lorsque l'on passe d'une représentation à l'autre par rotation sur le cercle chromatique. C'est le cas des ensembles de hauteurs $\{do, mi, sol\}$ et $\{ré, fa\sharp, la\}$, illustrés ci-dessous.

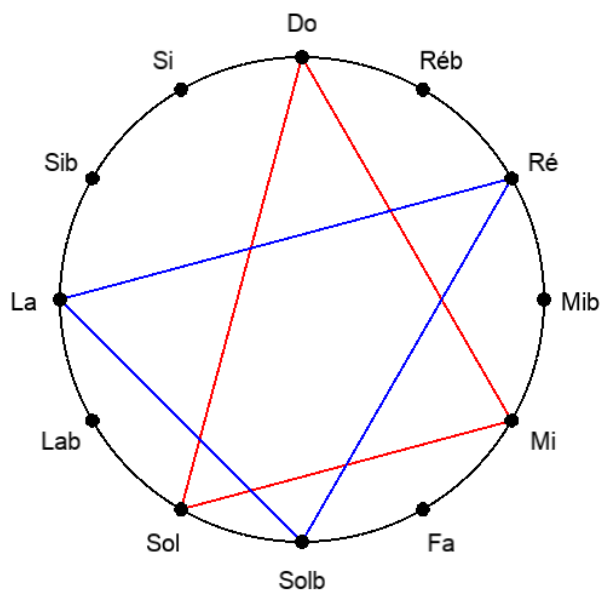


FIGURE 1.5 – Ensembles de hauteurs équivalents sous la relation de transposition.

Inversion Deux ensembles de hauteurs sont inverses l'un de l'autre lorsque l'on peut passer d'une représentation à l'autre sur le cercle chromatique par symétrie axiale.

La figure 1.6 montre deux ensembles inverses l'un l'autre : $\{do-mi-sol\}$ et $\{fa-lab-do\}$. La relation d'inversion rend le majeur mineur, et inversement.

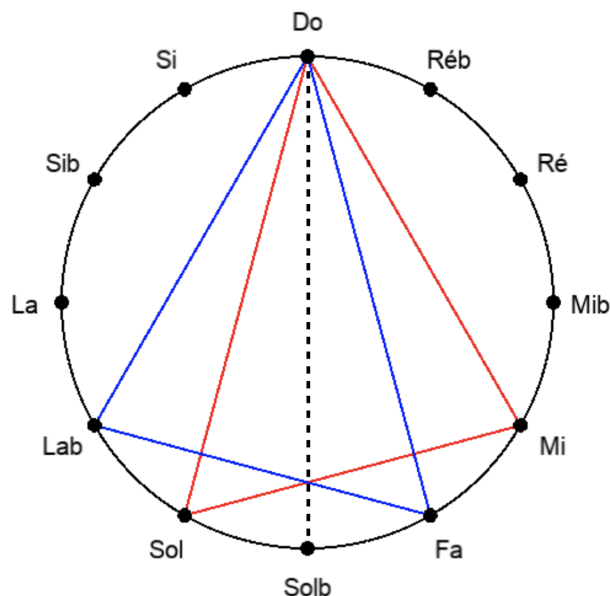


FIGURE 1.6 – Ensembles de hauteurs équivalents sous la relation d’inversion. En pointillés, l’axe de symétrie qui permet de passer de l’un à l’autre.

Classe d’intervalles Étant donné un intervalle inférieur à l’octave, on définit son complémentaire par l’intervalles à ajouter pour atteindre l’octave. De là, une classe d’intervalles est constituée de l’ensemble des intervalles égaux ou complémentaires modulo l’octave.

L’échelle chromatique contient 7 classes, numérotées dans la suite de 0 à 6 : la classe de l’unisson et de l’octave, celle de la seconde mineure et de la septième majeure, celle de la seconde majeure et de la septième mineure, celle de la tierce mineure et de la sixte majeure, celle de la tierce mineure et de la sixte majeure, celle de la quarte et de la quinte, et celle de la quinte diminuée.

On peut associer sans ambiguïté à un ensemble de deux hauteurs une classe d’intervalles.

Contenu intervallique Le contenu intervallique d’un ensemble de hauteurs est défini comme l’ensemble des classes d’intervalles contenues dans

l'ensemble de hauteurs. Si N désigne la cardinalité de l'ensemble de hauteurs, son contenu intervallique contient $\binom{N}{2}$ classes d'intervalles.

Avec l'échelle tempérée, le contenu intervallique s'écrit usuellement sous la forme d'un vecteur à 6 composantes, appelé vecteur intervallique, qui correspondent au nombre d'occurrence de chacune des classes d'intervalles, de la classe du demi-ton à celle de la quinte diminuée.

La figure 1.7 représente chaque classe d'intervalle de l'ensemble de hauteurs $\{do, mi, sol, sib\}$, de contenu intervallique $\langle 0,1,2,1,1,1 \rangle$.

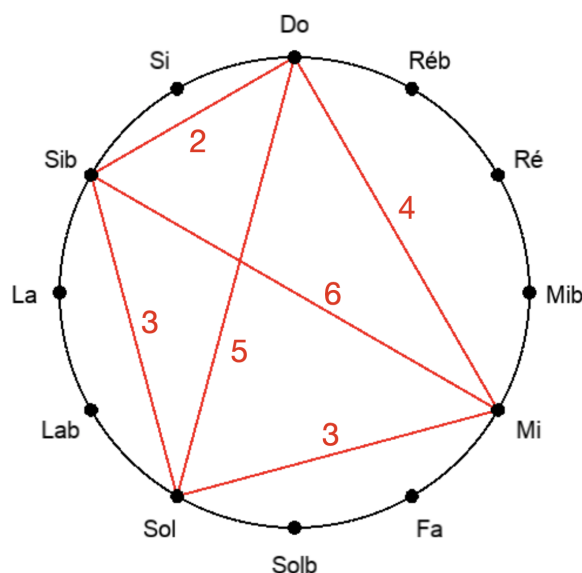


FIGURE 1.7 – Ensemble de hauteurs $\{di-mi-sol-sib\}$, de contenu intervallique $\langle 0,1,2,1,1,1 \rangle$

Deux ensembles de hauteurs ayant le même contenu intervallique sont dits « homométriques », ou encore reliés par la « relation Z^1 ». Le contenu intervallique est stable par transposition et inversion.

Présentation réduite (structure intervallique) Étant donné un ensemble ordonné (autrement dit une liste) de hauteurs, on peut lui associer sa

1. Ce terme vient de Forte (Forte, 1973).

présentation réduite, parfois appelée structure intervallique, qui est la liste des intervalles entre les notes consécutives de l'accord obtenu en fixant une note de la première hauteur, en ajoutant les notes unes à unes dans l'ordre des hauteurs vers l'aigu de manière à ce que les intervalles entre notes consécutives ne dépassent pas l'octave, et en ajoutant enfin la note de la première hauteur qui forme un intervalle pas plus grand que l'octave avec la dernière note. Nous préférons dans ce travail le terme « présentation réduite » à celui de « structure intervallique », étant donnée la confusion possible entre les termes « contenu intervallique » et « structure intervallique ».

En guise d'illustration, la présentation réduite du singleton $\{\text{ré}\}$ est [12], celle de $[\text{do}, \text{mi}, \text{sol}]$ est [4,3,5], et celle de $[\text{do-sol-mi}]$ est [7,9,8]. À partir de la présentation réduite d'un ensemble ordonné de hauteurs, on obtient la présentation réduite de ses renversements en faisant agir le groupe des permutations cycliques.

Forme normale et classe normale Deux ensembles de hauteurs appartiennent à la même classe normale lorsqu'ils sont transposés l'un de l'autre.

À un ensemble de hauteurs on peut associer sa forme normale. Celle-ci est obtenue parmi les présentations réduites de toutes les listes ordonnées de l'ensemble de départ, comme celle dont la dernière composante majore les autres composantes. Si plusieurs présentations réduites vérifient cette propriété, alors la forme normale est la plus petite dans l'ordre alphabétique des composantes.

Par exemple, la forme normale de l'ensemble $\{\text{do}, \text{fa}, \text{sib}\}$ est [2,5,5], celle de $\{\text{ré}, \text{fa}\sharp, \text{la}\}$ est [4,3,5].

La forme normale est invariante par transposition, et caractérise les classes normales. Les classes normales forment une partition de l'ensemble des ensembles de hauteurs.

Les classes normales caractérisent les accords au renversement et à la disposition près. Ainsi, les types d'accords comme les accords majeurs, les accords mineurs ou les accords diminués sont en fait des classes normales.

Forme première et classe première Deux ensembles de hauteurs appartiennent à la même classe première lorsqu'ils sont reliés par des relations de transposition et/ou d'inversion.

À un ensemble de hauteurs on peut associer sa forme première. Celle-ci est la présentation réduite la plus petite dans l'ordre alphabétique parmi la

forme normale de l'ensemble de départ ou la forme normale d'une de ses inversions.

L'ensemble de hauteurs $\{do, mi, sol\}$, de forme normale $[4,3,5]$, a comme inversion la forme normale $[3,4,5]$, qui est sa forme première. La forme première des accords majeurs est la forme normale des accords mineurs.

La forme première est invariante par transposition et inversion, et caractérise les classes premières. La partition de l'ensemble des ensembles de hauteurs en classes premières est moins fine que la partition en classes normales. Les accords majeurs et mineurs sont dans la même classe première.

Le contenu intervallique, stable par transposition et inversion, reste constant sur tous les ensembles de hauteurs d'une classe normale. Cependant, le contenu intervallique ne caractérise pas les classes premières. Sur la figure 1.8 sont ainsi exhibés deux ensembles de hauteurs ayant le même contenu intervallique, en l'occurrence $\langle 1,1,1,1,1,1 \rangle$, mais n'appartenant pas à la même classe première (on ne passe pas de l'un à l'autre en composant des inversions et transitions).

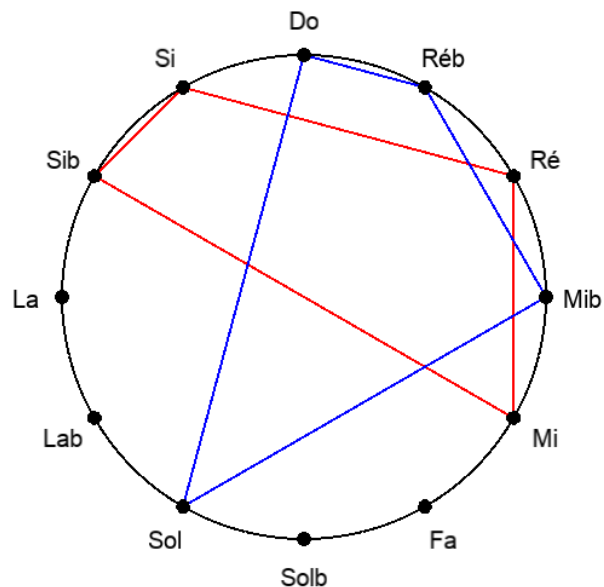


FIGURE 1.8 – Ensembles de hauteurs ayant le même contenu intervallique, mais pas la même forme première, ni la même forme identitaire.

Identité et forme identitaire La forme identitaire d'un ensemble de hauteurs est obtenue à partir de sa présentation réduite en permutant les intervalles de manière à ce que ceux-ci soient rangés par ordre croissant.

On appelle identité l'ensemble des ensembles de hauteurs qui ont la même forme identitaire. La partition de l'ensemble des ensembles de hauteurs en identités est moins fine que la partition en classes premières.

Prenons l'exemple d'un accord de septième de dominante, de présentation réduite $[3,3,2,4]$. Sa forme identitaire est $[2,3,3,4]$, et l'identité associée contient 3 classes normales : celle de l'accord de septième de dominante, celle de l'accord de septième mineure avec quinte diminuée, soit $[2,3,3,4]$, et celle de l'accord de septième mineure $[3,2,3,4]$. Les deux premières classes normales constituent une classe première, tandis que la troisième constitue une classe première distincte.

On désigne par « orbite identitaire » l'ensemble des classes normales d'une même identité

Deux ensembles de hauteurs d'une même identité n'ont pas nécessairement le même contenu intervallique. La figure 1.9 exhibe des ensembles de hauteurs appartenant à une même identité, mais n'ayant pas le même contenu intervallique. En effet, l'ensemble rouge a pour forme première $\langle 1,1,0,1,2,1 \rangle$, le bleu $\langle 1,1,1,1,1,1 \rangle$. De même, deux ensembles partageant le même contenu intervallique n'ont pas forcément la même forme identitaire, comme le montre la figure 1.8.

Ces considérations montrent que la partition des ensembles d'ensembles de hauteurs en classes homométriques (caractérisées par le contenu intervallique) et la partition en identités ne sont pas comparables, autrement dit aucune n'est plus fine que l'autre.

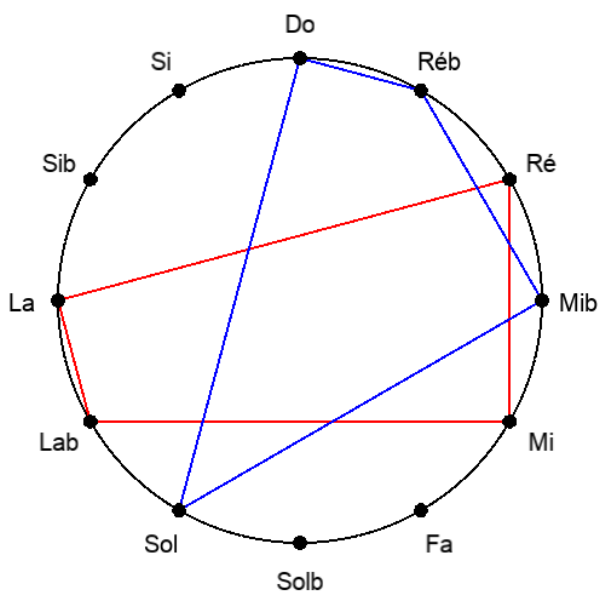


FIGURE 1.9 – Ensembles de hauteurs ayant la même forme identitaire, mais pas le même contenu intervallique.

Nous avons défini plusieurs partitions de l'ensemble des ensembles de hauteurs : la partition en classes normales, premières, homométriques, et celle en identités. Si elles sont désignées respectivement par $\mathcal{P}_{\text{normales}}$, $\mathcal{P}_{\text{premières}}$,

$\mathcal{P}_{\text{homom.}}$ et $\mathcal{P}_{\text{identités}}$, on a l'inégalité suivante, où $<$ désigne la relation de finesse entre partitions d'un même ensemble :

$$\begin{cases} \mathcal{P}_{\text{identités}} \\ \mathcal{P}_{\text{homom.}} \end{cases} < \mathcal{P}_{\text{premières}} < \mathcal{P}_{\text{normales}}.$$

Conventions de rédaction

Ce travail étant à cheval entre les sciences dures et les sciences humaines, nous faisons le choix d'opter pour des renvois bibliographiques dans le corps du texte sous la forme (Nom, année).

Première partie

État de l'art sur la consonance et la dissonance

Au cœur des préoccupations pratiques et théoriques sur l’harmonie dans la musique occidentale se trouvent les notions de consonance et de dissonance, qui traversent les époques, de l’Antiquité jusqu’à nos jours. Pourtant, loin d’unifier l’histoire de la musique et de la théorie musicale, elle évolue avec eux, devance ou précède l’usage, s’adapte à l’époque, quitte à se métamorphoser et se contredire.

De cette évolution résulte une multiplicité sémantique des notions, qui peut faire l’effet d’un flou sémantique. La consonance, pour ne citer que le *Oxford English Dictionary*, serait une « *combinaison plaisante de sons* », contre une « *combinaison dure et inharmonieuse de sons* » pour la seconde. D’autres approches donnent pour critère de la consonance non plus l’aspect agréable, mais la stabilité, la résolution, l’unité perceptive, la facilité d’identification, la familiarité, voir la dimension ennuyeuse et inintéressante. Certaines définitions sont fonctionnelles, attachées au langage tonal, tandis que d’autres ne le sont pas¹. La consonance et la dissonance désignent tantôt des propriétés, des qualités d’un agrégat de sons, tantôt les agrégats sonores qui ont ces propriétés. Lorsqu’elles se réfèrent à des qualités, ces dernières peuvent être de nature perceptive, de nature acoustique, ou des propriétés abstraites. Les notions et les théories les expliquant peuvent alors être difficilement identifiables.

Dans son livre de référence *A History of ‘Consonance’ and ‘Dissonance’* (Tenney, 1988), James Tenney insiste sur la nécessité de distinguer quatre approches possibles de la consonance et de la dissonance :

- L’approche sémantique. Il s’agit de comprendre à quels objets se réfèrent ces notions, et quels types de propriétés entrent dans leur définition. Tenney distingue cinq phases ou conceptions différentes de la consonance et de la dissonance (CDC - « *consonance-dissonance concept* »), chacune associée à une époque et à une pratique musicale.
- Les théories explicatives. Celles-ci sont d’ordre métaphysique, religieux ou scientifique. Elles recoupent en partie les différentes conceptions de la consonance et de la dissonance.
- L’attitude esthétique à leur égard. Tout un discours esthétique est généralement produit par les savants, les philosophes et les musiciens autour de la consonance et de la dissonance, qui traite du beau, du

1. Voir (Tenney, 1988, p.1-5)

- jugement de goût ou des relations aux autres arts.
- Leur usage pratique. Il s'agit de confronter les consonances et les dissonances aux œuvres pour comprendre leur rôle dans le discours musical.

On peut ajouter à cela un cinquième point, qui est l'expérience psychologique face à la consonance et la dissonance.

Multiplicité sémantique des notions de consonance et de dissonance

Les cinq conceptions de la consonance et de la dissonance identifiées par Tenney diffèrent autant par les objets auxquels elles se réfèrent que par les qualités qu'elles décrivent. Chacune d'elles est associée à une pratique musicale et à une époque, durant laquelle elle est dominante. Les voici :

- **CDC 1 : caractère mélodique** Pendant l'époque qui précède la polyphonie, les notions de consonance et de dissonance ont en grande partie un caractère mélodique. Elles traduisent une affinité entre deux notes jouées successivement. Chez les Grecs, les intervalles consonants sont ceux qui sont directement accordables par la méthode pythagoricienne de division de la corde : l'octave, la quinte, la quarte, l'octave + quinte, l'octave + quarte, la double octave, la double octave + quarte et la double octave + quinte¹. Ce classement des consonances est lié à la « *tetraktys* » pythagoricienne, c'est-à-dire qu'elles sont composées des intervalles dont les rapports ne font qu'intervenir les chiffres 1, 2, 3 et 4. Pratiquement le chant et le jeu en octaves, appelé « *magadizein* », les grecs reconnaissent aux consonances d'octaves simultanées un caractère de fusion.
- **CDC 2 : caractère sonore des dyades simultanées** Aux débuts de la polyphonie, à partir du IX^e siècle, une nouvelle notion de consonance apparaît, qui se rapporte à la sonorité de deux notes jouées non plus successivement, mais simultanément. Le critère de la consonance est alors perceptif, c'est la capacité de la dyade à sonner comme une seule note, préfigurant les travaux de Stumpf un millénaire plus tard². Dans la pratique musicale, les voix gagnent petit à petit en indépen-

1. L'octave + quarte, la double octave + quarte et la double octave + quinte sont refusés par les Pythagoriciens mais admis par Aristoxène.

2. Voir (Stumpf, 1883, 1890).

dance mélodique, et les tierces et les sixtes entrent dans la catégorie des consonances, malgré des résistances dues à la vision pythagoricienne. Le conflit entre l'intonation pythagoricienne et la justesse des tierces et des sixtes est symptomatique de la rencontre de deux systèmes différents. Au XIII^e siècle, des classements élaborés voient le jour avec des savants comme Jean de Garlande ([Garlande, 1250](#)), qui distingue 3 types de consonances et 3 types de dissonances : les parfaites, les intermédiaires et les imparfaites. Ces classements ne suivent pas l'ordre de simplicité des ratios, ce qui suggère que la perception joue un rôle primordial dans leur établissement.

- **CDC 3 : usage contrapuntique** Au tournant des XIII^e et XIV^e siècles, avec le développement du contrepoint, la consonance change de statut, et sa définition devient opérationnelle. C'est l'usage dans l'écriture contrapuntique qui détermine le classement des consonances et des dissonances, et non plus seulement les qualités sonores propres. Le contexte musical devient déterminant dans leur attribution. Ce changement va de pair avec une simplification des catégories, qui se réduisent à trois catégories opérationnelles, correspondant à des usages distincts : les consonances parfaites, les consonances imparfaites et les dissonances. À la fin du XV^e siècle on commence à prôner les intervalles justes par rapport aux intervalles Pythagoriciens, et Zarlino ([Zarlino, 1558](#)) légitime théoriquement l'appartenance des tierces et sixtes aux consonances, passant de la *tetraktys* au *senario*, qui ajoute à la liste 2, 3, 4 les chiffres 5, 6 et 8. Dans les écrits de l'époque, les qualités sonores font également partie les critères de consonance et de dissonance, mais sont plus subjectives qu'à l'époque du CDC-2 : on parle de la douceur, du caractère agréable, et la connotation consonance-plaisir et dissonance-déplaisir s'impose.
- **CDC 4 : consonance d'une note par rapport à la basse fondamentale** Au XVII^e s, les termes « consonant » et « dissonant » sont utilisés pour qualifier une note individuelle dans un agrégat, et non plus l'agrégat lui-même. Cette notion devient centrale dans les écrits de Rameau, dans un sens bien précis. Si une note a un rapport d'unisson, de tierce ou de quinte par rapport à la basse fondamentale, elle est considérée comme une note consonante, sinon elle devient une note dissonante. C'est donc ici le rapport à la basse fondamentale qui

est déterminant, et des ambiguïtés par rapport à cette basse induisent des ambiguïtés sur le statut de la note. Les ambiguïtés peuvent être résolues en considérant le contexte tonal et l'aspect fonctionnel. Une idée de mouvement est associée au CDC-4 : une note dissonante est tenue de se résoudre, contrairement à une note consonante. A ce titre, la consonance et la dissonance de manière purement fonctionnelle, sans se référer aux qualités sonores.

- **CDC 5 : Consonance perceptive** Avec Helmholtz ([Helmholtz, 1863](#)), une nouvelle conception de la consonance et de la dissonance apparaît, cette fois placée sous l'autorité scientifique de la biologie et de la psychophysique naissante. Celle-ci peut être vue comme une réponse à l'élargissement de la palette sonore de l'orchestre, à l'importance grandissante des dynamiques et des timbres et à une écriture qui privilégie les contrastes et les chromatismes. La dissonance devient une propriété perceptive associée à un son ou agrégat sonore, due aux battements rapides entre partiels, qui interagissent sur l'appareil auditif par déformation de la membrane basilaire. En tant que propriétés perceptives, la consonance et la dissonance dépendent des paramètres acoustiques que sont le timbre, l'intensité, le mode de jeu, et ne peuvent pas être déduits d'un simple raisonnement sur les hauteurs écrites. Une unique note, si son timbre est riche, peut d'ailleurs contenir un certain degré de dissonance. Malgré une méthode de quantification des battements, Helmholtz continue à adopter une catégorisation des consonances et des dissonances. Cette conception scientifique de la consonance et de la dissonance a eu de nombreux développements, et constitue de nos jours l'approche dominante lorsqu'il s'agit de trouver des corrélats acoustiques, physiologiques ou psychologiques de la consonance.

Le tableau [1.1](#) ainsi que la figure [1.10](#) résument les différentes conceptions de la consonance et de la dissonance.

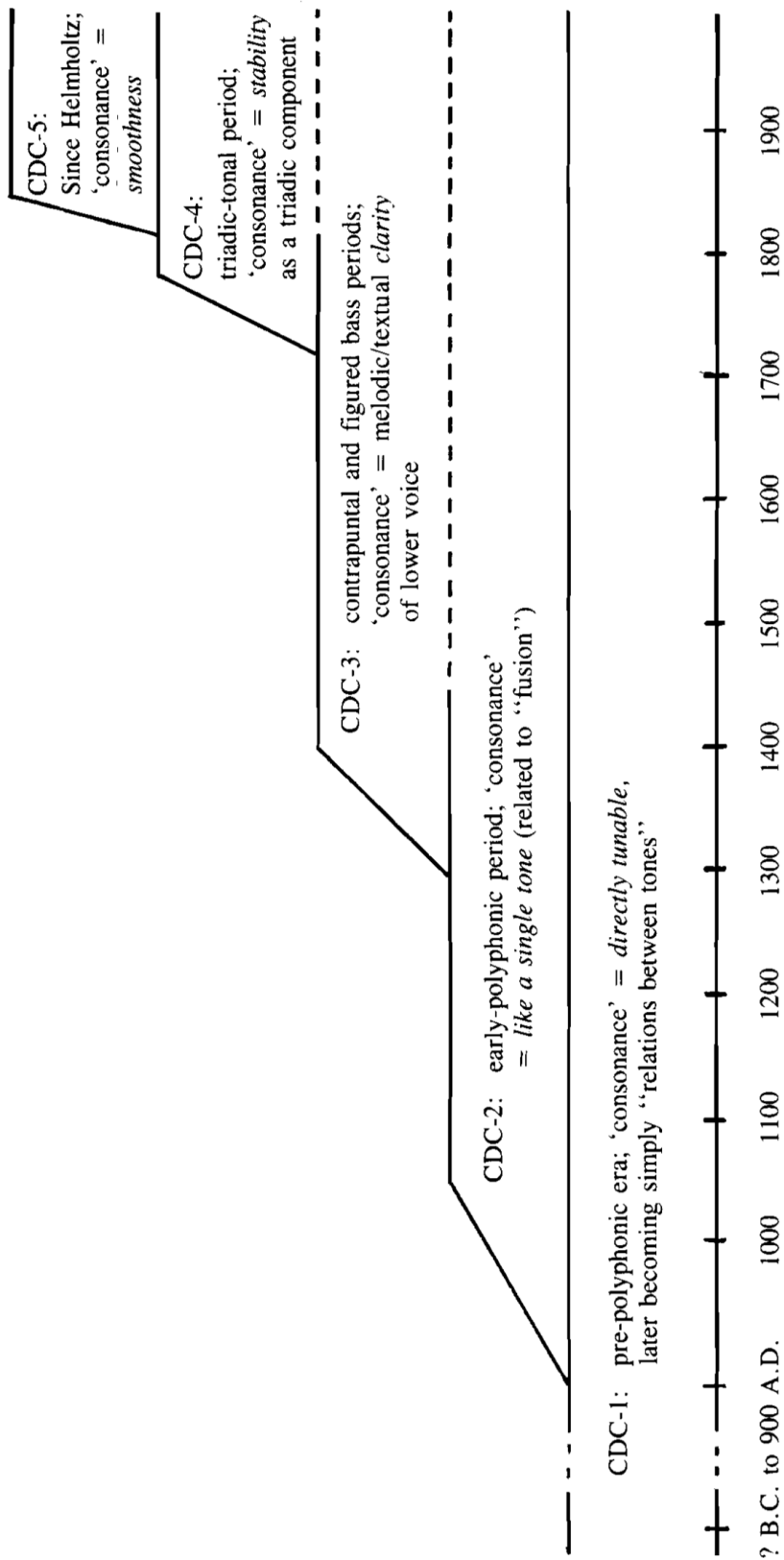


FIGURE 1.10 – Frise chronologique des différentes conceptions de la consonance et de la dissonance, tirée de (Tenney, 1988, p.115).

	CDC 1 : C/D mélodique	CDC 2 : C/D diphonique	CDC 3 : C/D contrapuntique	CDC 4 : C/D triadique	CDC 5 : C/D timbrale
Objet	intervalles mélodiques	dyades simultanées		notes indivi- duelles dans un accord	n'importe quel son
Qualité	relations entre notes	aspect sonore	aspect sonore + rôle contrapun- tique	tendance au mouvement	aspect sonore
Définition de la consonance	directement accordable	sonne comme une note	clarté mélodique et textuelle de la voix inférieure	stabilité (ne nécessite pas de résolution)	absence de rugosité

TABLE 1.1 – Résumé des différentes conceptions de la consonance et de la dissonance (CDC), telles que présentées dans (Tenney, 1988).

Comme le suggère la figure 1.10, le processus d'évolution des conceptions de consonance et dissonance est cumulatif, c'est-à-dire qu'une conception ancienne survie à l'émergence d'une nouvelle conception. Cet aspect cumulatif est cause de nos jours d'une grande confusion sémantique, le même mot étant employé dans des acceptions différentes¹. Cette confusion n'étant pas toujours consciente, Tenney propose d'employer des termes distincts pour désigner les différentes conceptions de la consonance et de la dissonance, respectivement consonance et dissonance *mélodiques*, *diaphoniques*, *contrapuntiques*, *triadiques* et *timbrales*.

Tenney insiste sur la nécessité de distinguer les différentes notions de consonance et de dissonance avant de se lancer dans la recherche de corrélats et de théories explicatives, que celles-ci soient d'ordre acoustique, biologique, psychologique ou cognitif. En effet, il émet l'hypothèse que les corrélats diffèrent selon les conceptions de la consonance et de la dissonance mises en

1. Les exemples d'une telle confusion sont nombreux, elle apparaît par exemple dans le traité de Garbouzov (Garbouzov, 1954, chap.2), devenu un classique en Russie. Chez l'auteur, les notions de consonance et de dissonance se réfèrent successivement aux sons simultanés, aux sons successifs, à une note ou à un agrégat sonore, et les critères sont tantôt perceptifs (sensation agréable / désagréable), tantôt fonctionnels. Les mêmes notions sont employées pour désigner les cinq conceptions de la consonance et de la dissonance dégagées par Tenney.

jeu.

Le chapitre 2 est consacré à une revue historique des théories explicatives de la consonance et de la dissonance. Différents paradigmes sont mis en avant, qui sont mis en rapport avec les conceptions de la consonance et de la dissonance (CDC) présentés précédemment. Le chapitre 3 fait le point sur l'état actuel des recherches sur la consonance et de la dissonance.

Chapitre 2

Théories explicatives de la consonance et de la dissonance : les courants de pensée

La conception de la consonance et de la dissonance, nous l'avons vu, a évolué au fil des époques et des pratiques musicales. Tout au long de leur histoire, ces notions deviennent l'objet de théories qui tentent de les expliquer. Ces théories jouent un rôle, à côté de la pratique musicale, dans l'élaboration des conceptions de consonance et dissonance, et la part respective des deux facteurs varie avec l'histoire. Le passage de la consonance et dissonance mélodiques (CDC 1) à la consonance et dissonance diphoniques (CDC 2) est une réponse à la polyphonie naissante, bien plus qu'à un changement de point de vue théorique. À l'opposé, l'émergence de la consonance et dissonance timbrale (CDC 5) est une conséquence directe de l'approche psychoacoustique d'Helmholtz, plus qu'une conséquence directe de l'élargissement de la palette dynamique et timbrale dans la pratique musicale. Aussi, l'évolution des théories explicatives de la consonance et de la dissonance est-elle en partie indépendante de l'évolution des conceptions de la consonance et de la dissonance.

L'histoire des théories explicatives de la consonance et de la dissonance est traversée par plusieurs courants de pensée. À l'instar des différentes conceptions de la consonance et de la dissonance, chacune d'elles est associée à une époque, durant laquelle elle constitue le paradigme dominant. Elles suivent cependant un processus cumulatif, cause de recoupements et d'affrontements. Plusieurs débats témoignent d'ailleurs d'une histoire des idées mouvementée :

entre l'école pythagoricienne et Aristoxène, Rameau et Rousseau, Helmholtz et Stumpf.

Cette partie, sans viser à l'exhaustivité, se propose d'exposer les différents courants théoriques en montrant comment ceux-ci sont toujours vivants, et en mettant en avant leur relation avec les conceptions de la consonance et de la dissonance.

2.1 Approches arithmétiques et géométriques

Dans l'Antiquité grecque, la musique fait partie, à côté de l'arithmétique, de la géométrie et de l'astronomie, du *quadrivium*, distinct du *trivium*, qui regroupe la grammaire, la dialectique et la rhétorique. A ce titre, la théorie de la musique, et particulièrement la théorie de la consonance, est le lieu de raisonnements abstraits, qui font intervenir des nombres, des rapports, et des formes. Si les consonances sont reconnues par l'oreille, et utilisées en pratique pour accorder les instrument de musique, l'explication de la consonance reste abstraite, se situant en-dehors de la perception ou de la réalité physique de la musique. Cette abstraction de la théorie règnera en maître pendant plus de deux mille ans, pour trouver son chant du cygne chez Euler, à l'époque où un Rameau ou un d'Alembert renverseront l'approche en considérant les sciences physiques.

2.1.1 Courant pythagoricien

L'expérience de la division de la corde vibrante sur le monocorde a amené les Pythagoriciens à associer un intervalle musical à un rapport entier de longueur de corde, et la théorie de la musique devient une théorie des proportions. Ces proportions, du moins jusqu'à la Renaissance, ne sont pas pensées comme des rapports de fréquence, mais correspondent, du moins le croit-on dans l'Antiquité, aux rapports entre les vitesses des sons, les sons aigus allant plus vite que les sons graves. Cette hypothèse se révélera bien sûr fausse, puisque l'air n'est pas un milieu dispersif pour le son.

N'admettant pour consonances que l'octave, la quinte et la quarte, de rapports respectifs $2/1$, $3/2$ et $4/3$, l'arithmétique développée par les grecs est fondée sur la *tetraktys* pythagoricienne, c'est-à-dire sur les entiers 2, 3 et 4. Elle combine leurs rapports selon trois sortes de moyennes, ou médiétés : la médiété arithmétique, la médiété harmonique et la médiété géométrique¹. Cette arithmétique n'est pas propre à la musique : elle est sous-tendue par une conception holistique dans laquelle la musique traduit l'harmonie du monde, harmonie régie par des rapports simples qui reflètent l'ordre cosmique, universel.

La Division du Canon d'Euclide, à l'origine controversée, constitue une source importante de la théorie musicale de l'Antiquité. L'auteur part d'un

1. Se référer au Timée de Platon.

postulat qui établit un rapport entre la qualité sensible de la consonance et une certaine propriété arithmétique. À partir de ce postulat, Euclide établit un système logique qui justifie les consonances, ainsi que la constitution interne des tétracordes en vigueur dans la musique grecque.

Un jalon important est Ptolémée, auteur des *Harmoniques* (Ptolémée, 1930). Ptolémée divise les intervalles en trois classes sous la forme d’une gradation, partant de l’identité (l’unisson). Il y a d’abord les homophones, avec l’octave et l’octave redoublée ; puis les « *symphonies* » avec la quarte, la quinte, la onzième et la douzième ; enfin les sons « *emmeleis* » — c’est-à-dire justes ou harmonieux — qui sont les sons propres à la mélodie, comme le ton.

La conception pythagoricienne traversera les siècles et sera en vigueur jusqu’à la Renaissance, malgré l’apparition de la polyphonie et du contrepoint. Le bouleversement de la pratique musicale n’agit qu’en surface sur le paradigme théorique. En passant de la *tetraktys* pythagoricienne au *senario*¹, Zarlino étend le système des grecs par des explications d’ordre non seulement arithmétique, mais aussi religieux et mystique, justifiant ainsi théoriquement l’appartenance des tierces et des sixtes à la catégorie des consonances dites imparfaites, en opposition aux consonances parfaites des grecs.

2.1.2 Approches géométriques

Alors que Zarlino reste fidèle à l’approche pythagoricienne, Kepler s’en détache par les moyens employés, troquant l’arithmétique contre la géométrie, les nombres entiers contre les formes. De la vision grecque, l’astronome garde la dimension holistique et universelle, comme le montre le titre de son ouvrage majeur, *Harmonices Mundi* (Kepler, 1619). Celui-ci est divisé en cinq volumes, les deux premiers consacrés à des questions de géométrie (polygones réguliers et congruence des figures), le troisième aux proportions harmoniques en musique, et les deux derniers à des questions de cosmologie (configurations harmoniques dans l’astrologie et harmonie dans le mouvement des planètes). Pour Kepler, les problématiques cosmologiques, théologiques et musicales sont étroitement liées. Le monde étant une création de Dieu,

1. Le *senario* comprend les intervalles dont les rapports sont composés des chiffres 2, 3, 4, 5, 6, 8, autrement dit les chiffres inférieurs à 8 faisant intervenir les facteurs premiers 2, 3 et 5, quand la *tetraktys* pythagoricienne ne faisait intervenir que les facteurs premiers 2 et 3. Le passage du *tetraktys* au *senario* consiste essentiellement en l’ajout du facteur premier 5.

il se voit doté d'attributs manifestant sa provenance divine. De là, la musique reflète l'harmonie divine de l'univers. Kepler voit l'ordre divin non plus dans les rapport arithmétiques, mais dans l'alignement des corps célestes, ce qui l'amène à des considérations géométriques sur les polygones réguliers, le rapprochant de la vision de Platon. Aussi Kepler fait-il appel aux polygones réguliers pour justifier le classement des consonances et l'absence du 7 dans le *senario* de Zarlino (chiffres 2, 3, 4, 5, 6, 8). La rigueur géométrique va de pair avec des preuves très générales, de nature philosophique. Le mysticisme du nombre cède la place à la mystique des formes, qui perd l'empirisme des divisions entières du monocorde.

Par son approche théologique, cosmologique, philosophique et géométrique, Kepler reste éloigné des problématique musicales. Par son mysticisme et son abstraction, il se montre résolument réactionnaire, vis-à-vis d'un Descartes qui publie au même moment son *Compendium Musicae* ([Descartes, 1618](#)).

2.1.3 Ordre et raison au XVII^e siècle : la théorie des coups

Dans la première moitié du XVII^e siècle, on assiste, avec des savants comme Descartes, Beeckman, Mersenne et Galilée, à un véritable changement de paradigme concernant la place de la musique, dominée par un désir de rationalisation. La nature, si elle est l'œuvre de Dieu, est avant tout rationnelle, et il s'agit de déchiffrer ses secrets, sans faire appel au mysticisme des nombres et des formes. C'est l'époque où Beeckman tente de donner une démonstration physique de la loi des cordes vibrantes. La musique ne vise plus à célébrer l'ordre divin, mais à flatter les sens et la raison ; de l'ordre cosmique, elle passe à l'ordre humain. « *Sa fin est de plaire et d'exciter en nous diverses passions* », écrit Descartes au début du *Compendium Musicae*, c'est un « *art destiné à produire des effets sur l'âme par l'intermédiaire des sens* », rajoute Mersenne dans son *Harmonie universelle* ([Mersenne, 1636](#)).

Les consonances et les dissonances sont des phénomènes humains, et n'ont de sens que pour ceux qui sont capables de les percevoir ou de les apprécier. Les consonances, pour Mersenne, n'existent que par le plaisir qu'elles fournissent. Ce plaisir est rattaché à la raison, et non au jugement particulier, ce qui garantie l'objectivité de leur classement. Pour Descartes, le « *le plaisir des sens consiste en une certaine proportion et correspondance de l'objet avec*

le sens¹ », et cette proportion devient le critère de la consonance. Les consonances sont précisément les intervalles qui font que les sens ne se fatiguent pas, tandis les dissonances paraissent confuses à l'oreille. Cette argumentation accorde un rôle important à la perception, sans lui conférer pour autant une autonomie, puisqu'elle est toujours subordonnée à la raison. Sens, raison et beauté sont étroitement liés : la beauté est rattachée à l'ordre, et la régularité produit le beau.

Cet ordre se révèle dans la théorie de la coïncidence des coups, qui se met en place au début du XVII^e siècle pour constituer l'approche dominante jusqu'aux travaux de Rameau. Les connaissances acoustiques à cette époque sont encore très sommaires, on ignore la vraie nature du son, et la théorie corpusculaire de Beeckman n'a que peu de succès par rapport à la vision de Mersenne, qui considère que le son provient de battements de l'air où chaque battement crée un son. Dans les deux théories, le son produit des coups sur la membrane du tympan, à une régularité correspondant à leur fréquence. Lors de l'audition de deux sons, le tympan est soumis à deux fréquences de coups, qui tantôt coïncident, lorsque les rapports de fréquences sont en proportions rationnelles, tantôt ne coïncident pas. La consonance provient de la simultanéité des coups, la dissonance de l'irrégularité des coups. Le rapport du nombre de coups qui coïncident sur le nombre total de coups sert alors à classer les consonances. La théorie de la coïncidence des coups fait donc appel à l'arithmétique, comme le faisait l'approche pythagoricienne, avec l'édifice cosmico-mystique en moins. Pourtant, dans le discours de Mersenne, la religion intervient souvent, à côté de la métaphysique, et l'unisson est comparé à Dieu.

Pour Leibniz, qui reprend à son compte la théorie de la coïncidence des coups au siècle suivant, cette arithmétique est inconsciente : « *la musique est une pratique occulte de l'arithmétique dans laquelle l'esprit ignore qu'il compte*² ».

Dans la théorie des coups, les consonances et les dissonances n'entretiennent plus une différence de nature, mais une différence de degré. En outre, l'hégémonie des nombres premiers 2, 3 et 5, qui date de Zarlino, est remise en cause, et Mersenne accepte, du moins en pratique, les intervalles

1. Tiré du *Compendium Musicae* (Descartes, 1618).

2. « *Musica est exercitium arithmeticae occultum nescientis se numerare animi* », tiré d'une lettre à Goldbach du 17 avril 1712. Traduction de Bailhache tirée de (Bailhache, 1992).

dont le rapport contient le chiffre 7 :

*Puisque le long exercice a coutume de rendre doux et facile ce qui semblait auparavant rude et fâcheux, je ne doute nullement que les intervalles dissonants [...], à savoir les raisons de 7 à 6 et de 8 à 7, qui divisent la quarte, ne puissent devenir agréables, si l'on s'accoutume à les ouïr et à les endurer, et que l'on en use comme il faut dans les récits et dans les concerts, afin d'émouvoir les passions, et pour plusieurs effets, dont la Musique ordinaire est privée.*¹

Il faut avoir conscience que dans le tempérament mésotonique en usage à l'époque de Mersenne, il est possible de faire sonner des intervalles (*sib-sol*♯) dont le rapport est proche de $7/4$, donnant une septième plus proche du 7e harmonique que la septième *do-sib*.

La ségrégation des consonances et des dissonances devient une affaire de sensibilité et de familiarité, plus qu'une question théorique, ce qui affaiblit la dépendance de la pratique musicale à la théorie. Leibniz se montre plus traditionnel que Mersenne sur cette question lorsqu'il affirme qu'« *en musique, nous ne comptons pas au-delà de cinq*² ». Il concède cependant : « *Si nous avions en partage un peu plus de finesse, nous pourrions aller jusqu'au nombre premier 7. [...] Mais il n'y aura guère de gens, qui iraient jusqu'aux nombres premiers [suivants] les plus proches, 11 et 13*³. »

Un point crucial de la théorie des coups consiste en l'assimilation de l'ordre à la beauté, et souvent à la moralité, qui semblait évidente aux savants de l'époque. Cette assimilation s'inscrit dans la séparation cartésienne âme - corps, qui induit un saut entre la matérialité des coups physiques sur le tympan et la nature spirituelle du son. D'ailleurs, l'apport physiologique à la théorie s'arrête aux coups sur le tympan, le reste relève de l'arithmétique, via cette assimilation fondamentale. Ce qu'il se passe après le tympan est inconnu à l'époque, et il faudra attendre plus de deux siècles avant qu'Helmholtz n'aille au-delà.

1. (Mersenne, 1636), avec ajustement de l'orthographe selon les normes actuelles par l'auteur.

2. Tiré d'une lettre à Goldbach du 17 avril 1712. Traduction de Bailhache tirée de (Bailhache, 1992).

3. *Ibid.*

2.1.4 Euler, ou la tentation de l'abstraction arithmétique

Le mathématicien Leonhard Euler constitue le dernier grand jalon de l'approche mathématique de la consonance. S'intéressant à la musique, il publie tout au long de sa vie plusieurs ouvrages sur le sujet, dont le premier, le *Tentamen novae theoriae musicae* (Euler, 1739), alors qu'il n'a que 24 ans. Malgré sa lecture de Rameau et d'Alembert, qu'il commente¹, Euler est résolument tourné vers le passé par rapport à son époque, autant par sa théorie, version améliorée de la théorie des coups, que par sa conception de la consonance et de la dissonance. Euler s'occupe en effet de la sonorité globale des accords, sans adopter l'approche fonctionnelle de la dissonance de Rameau.²

Euler réinvestit la théorie des coups pour la pousser plus loin sur le plan des calculs. Comme chez Descartes ou Mersenne, il s'agit de découvrir les vérités de la nature, révélées aux sens chez les personnes éclairées. Il écrit :

*Le jugement de l'oreille, auquel tout doit être rapporté, quelque bizarre qu'il paraisse souvent, n'est cependant rien moins qu'arbitraire ; il se règle toujours sur certains principes qui sont ceux de la véritable harmonie*³.

Le plaisir musical prend sa source dans la perception de cette véritable harmonie, qui résulte d'un ordre, lequel fonde le principe de goût. Cet ordre ce calcule, et c'est ce qu'essaye de montrer Euler en réduisant les principes de l'harmonie à des nombres. Pour se faire, il introduit la notion de suavité, mesure arithmétique de la consonance d'un accord.

Le degré de suavité d'un accord est obtenu à partir de la décomposition en facteurs premiers du PPCM (plus petit commun multiple) des nombres entiers intervenant dans les rapports des fréquences des sons.

Si le PPCM en question se décompose sous la forme $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_n^{\alpha_n}$, avec $(p_k)_{1 \leq k \leq n}$ des entiers premiers distincts et $\alpha_k \geq 1$ pour tout k , alors le degré

1. Par exemple dans son article *Du véritable caractère de la musique moderne* (Euler, 1766).

2. Pour reprendre les distinctions de Tenney (Voir page 74), il reste dans CDC 2 et CDC 3, quand le paradigme dominant est devenu la CDC 4.

3. Tiré de (Euler, 1766, p.174).

de suavité δ est défini de la manière suivante :

$$\delta = 1 + \sum_{k=1}^n (\alpha_k - 1) p_k.$$

Le degré de suavité détermine, selon Euler, le plaisir et l'agrément à l'écoute, les accords de faible degré étant plus consonants que les accords de degré important. Ainsi, une note seule est de degré 1, l'octave est de degré 2, la quinte (de rapport $3/2$) de degré 4. L'accord majeur (de rapport $6/5/4$) et l'accord mineur (de rapport $15/12/10$) ont le même PPCM, égal à 60, soit un degré de suavité égal à 9.

En mathématicien, Euler met tous les nombres premiers sur le même plan, sans privilégier 2, 3 et 5, ce qui l'amène à dire, raillant Leibniz : « *la musique a maintenant appris à compter jusqu'à 7*¹ ». Le raisonnement a l'avantage d'englober les accords de plus de deux notes, ce qui incite le mathématicien à dresser de grands tableaux de classification des accords de toutes les cardinalités. L'accord parfait majeur *do4-mi4-sol4* et l'accord à 12 notes *do1-do2-sol2-do3-mi3-sol3-mi4-sol4-si4-mi5-si5-si6*, en intonation juste, ont tous deux le même PPCM et donc le même degré de suavité², ce qui, selon Euler, assure qu'ils procurent aux sens le même agrément.

On atteint là la limite de l'approche abstraite d'Euler, dont les résultats, peu confrontés à l'expérience du sonore, semblent plus le fruit d'un esprit mathématique fécond que motivés par des considérations musicales³. Un siècle plus tard, les travaux d'Helmholtz permettront de réinterpréter les résultats d'Euler en matière de coïncidence des partiels harmoniques, et de dresser une analogie entre le degré de suavité et l'harmonicité. La faiblesse de la démarche d'Euler tient dans le fait qu'il considère une note comme une entité

1. Cité dans (Bailhache, 1992).

2. Ces accords ont pour rapports respectivement $6/5/4$, de PPCM égal à 60, et $60/30/20/15/12/10/6/5/4/3/2/1$, de même PPCM. Le second accord est dit complet, car on ne peut lui ajouter de note sans que son PPCM, en même temps que son degré de suavité, n'augmentent.

3. Ce qui ne l'empêche pas de tirer des conclusions d'ordre musical. Dans son article *Du véritable caractère de la musique moderne* (Euler, 1766), Euler défend la *musique moderne* (l'harmonie classique) en montrant que comme la *musique ancienne* (la polyphonie de la Renaissance) elle est « *conforme aux principes de la véritable harmonie* », et que par rapport à celle-ci elle étend les bornes des accords consonants, en incluant des consonances que « *les anciens n'ont pas eu la hardiesse ou l'adresse d'employer* ».

abstraite, dotée d'une fréquence, et non comme un objet physique, comme le fait Rameau à la même époque.

2.1.5 Les successeurs

Même si elles ne sont plus dominantes, les approches arithmétiques perdurent aux XIX^e et XX^e siècles. Présentons-en quelques unes.

Lipps Tandis que les travaux de Helmholtz et de Stumpf ont révolutionné la théorie de la consonance et de la dissonance par la rigueur et l'empirisme de leurs méthodes, le psychologue allemand Theodor Lipps, dans son ouvrage *Faits fondamentaux de la vie psychique* (Lipps, 1880), fait marche arrière en réinvestissant la coïncidence des coups.

Lipps considère que la consonance provient d'une coïncidence rythmique d'excitations inconscientes. Ces excitations, qu'il nomme « *micro-psychiques* », constituent un intermédiaire entre le stimulus physique et l'effet conscient. Il y a une correspondance entre le stimulus physique et les processus *micro-psychiques*, chaque vibration (autrement dit chaque période) créant une excitation *micro-psychique*. Une combinaison simultanée de sons crée une combinaison de séries d'expériences *micro-psychiques*, dont la coïncidence peut fusionner en une expérience consciente. Ce qui était acceptable au temps de Leibniz, faute de connaissances suffisantes sur la nature du son et le fonctionnement de l'oreille, est fortement critiqué par les contemporains de Lipps, Stumpf en première ligne, pour qui le recours à des processus inconscients et invérifiables pour expliquer des processus conscients constitue une capitulation de la part de la psychologie, qui cesse d'être scientifique quand elle cesse d'être empirique. Stumpf avance l'argument que deux sons, quand bien même leurs fréquences respectives entretiendraient un rapport simple, ne peuvent être en coïncidence rythmique s'ils sont bien déphasés. Moore (Moore, 1914) y voit d'ailleurs une « *rétrogradation d'Helmholtz vers l'approche vague et spéculative des prédécesseurs*¹ ».

Daniélou Grandement influencé par la musique indienne qu'il a fréquentée, Alain Daniélou développe dans son ouvrage *Sémantique musicale* (Daniélou, 1967) une théorie musicale qui troque le *tetraktys* pythagoricien (chiffres 2,

1. « *retrogression from Helmholtz back to the speculative vagueness of his predecessors* », tiré de (Moore, 1914, p.8), traduction de l'auteur.

3, 4) contre les nombres premiers 2, 3 et 5. Ces rapports ne se rapportent plus à un ordre universel, mais au fonctionnement du cerveau humain, qui « *classe immédiatement les facteurs 2, 3 et 5, et certains de leurs multiples ou produits même assez élevés, mais cesse de fonctionner lorsqu'on nourrit son mécanisme avec des nombres premiers supérieurs à 5* »¹. Selon Daniélou, la perception des rapports de hauteurs par le cerveau humain est basée sur la recherche inconsciente de rapports de fréquences uniquement dérivés des nombres 2, 3 et 5, et un intervalle qui ferait intervenir le facteur 7 ou tout autre nombre premier supérieur ne peut pas être reconnu par l'oreille. L'oreille humaine identifierait et classerait les hauteurs en s'appuyant sur ces systèmes binaires, ternaires et quaternaires.

Le classement des intervalles de Daniélou a un caractère avant tout mélodique, en raison probablement de l'influence de la musique indienne. On est là du côté de la CDC 1. Ardent défenseur de l'intonation juste, Daniélou propose d'ailleurs une échelle musicale qui divise l'octave en 53 intervalles inégaux², présents dans la musique indienne, dont les rapports font intervenir les nombres premiers 2, 3 et 5.

Une dimension morale est présente dans la théorie, puisque chaque intervalle provoque une réaction émotionnelle précise et universelle. La musique influence l'esprit par l'effet bénéfique ou nuisible de certains modes sur la formation du caractère.

Kayser Contrairement à Daniélou, qui se réfère explicitement à la tradition indienne plus qu'aux traditions occidentales, le musicologue Hans Kayser se place comme continuateur de Pythagore et de ses successeurs, en particulier Kepler. Dans son ouvrage *Le livre de l'Harmonique* (Kayser, 1950), Kayser développe l'idée de l'*Harmonique* comme science englobante qui étudie l'harmonie du cosmos, au sens de la musique des sphères. Il distingue cette Harmonique de l'étude des accords, qui en constitue une sous-branche. Son approche traite la musique dans le cadre de spéculations mystiques.

Si son influence n'est pas négligeable dans les cercles littéraires et ésotériques, elle n'a laissé que peu de traces dans la théorie musicale.

1. Tiré de (Daniélou, 1967).

2. La division de l'octave en 53 intervalles égaux est ancienne, elle approxime la division de l'octave en commas pythagoriciens. Daniélou, lui, approxime cette division égale avec des intervalles justes dont les rapports sont composés des nombres 2, 3 et 5. Pour plus de détails, se référer à (Daniélou, 1967).

2.2 Approche physicaliste

2.2.1 Rameau et d'Alembert : le corps sonore résonant

C'est Rameau qui viendra faire basculer la théorie de la musique, qui cesse d'être un jeu formel régi par des théories abstraites pour gagner une dimension empirique. Rameau publie en 1722 un *Traité de l'harmonie réduite à ses principes naturels* (Rameau, 1722), premier d'une série de plusieurs ouvrages, que le mathématicien, physicien et encyclopédiste d'Alembert reprend et simplifie dans ses *Éléments de musique suivant les principes de M. Rameau* (d'Alembert, 1752).

Dans le discours préliminaire de la seconde édition des *Éléments de musique*, d'Alembert précise sa démarche. Tout d'abord, il commence par rejeter les conceptions grecques et celles qui s'en inspirent, c'est-à-dire les conceptions géométriques, la coïncidence des coups, ainsi que la théorie de l'ordre et de la simplicité, arguant, tout empiriste qu'il est, que toutes ces théories ne font pas appel à l'expérience et sont construites sur des bases n'ayant pas la solidité d'observations concrètes. Il considère les constructions abstraites (Euler est sans doute visé, sans être cité explicitement) comme illusoires, du fait qu'elles ne sont pas « *nécessaires à l'Art* ». Aussi remet-il en question l'assimilation de l'ordre à la beauté : « *Mais pourquoi la concurrence des vibrations, c'est-à-dire leur direction dans le même sens, et la propriété de recommencer fréquemment ensemble, est-elle une si grande source de plaisir ? Sur quoi est fondée cette supposition gratuite*¹ ? »

Pour d'Alembert, la théorie musicale n'a pas à former un ensemble beau et harmonieux, mais doit servir la musique. Il fait d'ailleurs l'aveu de l'impossibilité d'aboutir à la rigueur mathématique des ouvrages d'Euler, estimant que le raisonnement par analogie et par déduction doit primer sur la rigueur de la démonstration. C'est là que se situe le tournant fondamental dans la démarche par rapport à un Descartes ou un Euler : pour eux la nature est divine, elle recèle un ordre caché qu'il convient de découvrir, tandis que pour d'Alembert la nature est avant tout physique, il s'agit par l'expérience de se l'approprier et de comprendre ses lois. La théorie musicale ne relève pas, pour d'Alembert, des mathématiques, mais des sciences de la nature.

La démarche physicaliste de Rameau doit beaucoup à Joseph Sauveur, père de l'acoustique musicale, qui le premier a mis en évidence l'existence

1. Tiré des *Éléments de musique*, (d'Alembert, 1752).

des différents modes de vibration d'une corde¹. La théorie se construit non plus sur un postulat ou une assimilation arbitraires, mais sur deux constats empiriques. Le premier est l'observation qu'un corps sonore qui résonne « *contient* » 5 harmoniques ; le second est la ressemblance des octaves, qui amène d'Alembert à les considérer comme équivalentes. Ces deux constats font office d'axiomes, à partir desquels est déduite la théorie, qui est un mélange de démonstrations logiques et de « *raisons de convenance* » au caractère plus arbitraire qui relève de l'analogie et de la déduction.

Les deux observations posées, l'accord majeur émerge naturellement, puisqu'à partir d'une note *do1*, la suite des cinq premiers harmoniques est la suivante : *do1*, *do2*, *sol2*, *do3*, *mi3*, et par l'équivalence des octaves, il reste *do*, *mi*, *sol*, soit l'accord de *Do* majeur, ce qui fait dire à d'Alembert : « *cet accord est l'ouvrage de la nature.* » Dans le contexte des Lumières, la nature est extérieure à l'homme, opposée à l'arbitraire des conventions humaines.

Cette « nature » devient alors la norme du goût, et en particulier de la consonance. Un postulat est sous-entendu dans la démarche de Rameau et de d'Alembert, celui de l'assimilation de la beauté et du plaisir subjectif à ce « *caractère naturel* ». L'assimilation de l'art à la nature n'est pas nouveau, puisque présente chez Aristote. Ce saut des phénomènes physiques aux éléments subjectifs sera déplacé par Helmholtz dans le système auditif.

Le corps sonore résonant Revenons ici sur les fondements physiques de la théorie de Rameau. N'ayant pas connaissance des travaux de Sauveur lors de la publication de son *Traité de l'harmonie* en 1722, la notion de sons harmoniques provient d'abord de l'expérience de résonance relatée par Descartes dans le *Compendium Musicae* :

*De deux termes que l'on suppose être en consonance le plus grave est de beaucoup le plus puissant et contient l'autre en quelque façon. On le voit sur les cordes d'un luth : si l'une d'elles est touchée, celles qui sont plus aiguës d'une octave et d'une quinte tremblent et résonnent spontanément*².

Le mot « *contient* » a ici un sens avant tout géométrique, qui concerne le rapport de longueurs de cordes, puisque Descartes postule une égalité de proportion entre le rapport des longueurs de cordes et le rapport de hauteurs

1. Se référer à son ouvrage *Principes d'acoustique et de musique, ou Système général des intervalles des sons* (Sauveur, 1701).

2. Tiré de (Descartes, 1618, p.64-66)

des sons. Descartes ne propose pas une interprétation proprement physique de cette résonance, pas plus que Rameau dans son *Traité*, qui écrit : « *Nous laisserons à la physique le soin de définir le son*¹. » Ce ne sont pas les fondements physiques, mais les conséquences musicales qui intéressent le musicien, lequel radicalise les résultats de Descartes en considérant que cette relation de contenance détermine le caractère musical des notes. Le son grave fonde le caractère consonant des sons aigus :

*De l'union de ces différents intervalles, il se forme différentes consonances, dont l'harmonie ne peut être parfaite, si ce premier son ne règne au-dessous d'eux, comme étant la base et le fondement, [...] donc ce premier son est encore le principe de ces consonances et de l'harmonie qu'elles forment*².

Ce tournant marque l'avènement de la consonance de notes³ (CDC 4) : la consonance du son aigu est déterminée par le son grave. Rameau critique par là l'interprétation de Descartes de l'expérience de résonance, et la dépasse. Cette subordination de propriétés des sons à un son source, à un centre harmonique, est centrale chez Rameau, puisqu'elle va le mener vers la notion de basse fondamentale.

La lecture par Rameau des travaux de Sauveur lui fournit un cadre qui l'incite à aller plus loin dans la justification des éléments de la théorie par des phénomènes physiques, comme le montre les recours plus nombreux à des considérations sur la résonance dans le *Nouveau système de musique théorique* (Rameau, 1726) et surtout la *Génération harmonique* (Rameau, 1737). Sauveur développe la notion d'harmonique à partir de l'expérience de résonance des cordes par sympathie, en observant que la vibration de la corde contient les vibrations de ses divisions physiques. Rameau retient les cinq premiers harmoniques pour ériger, via l'équivalence des octaves, l'accord majeur en objet naturel contenu dans la fondamentale. C'est la résonance du corps sonore.

*Tout corps sonore, pris en particulier, est toujours censé porter avec lui la même harmonie qu'il fait résonner, il en est le générateur*⁴.

1. Tiré de (Rameau, 1722, liv.I, chap.I).

2. *Ibid.*

3. Voir page 74.

4. Tiré de la *Démonstration du Principe de l'harmonie* (Rameau, 1750, p.8).

Rousseau, grand adversaire de Rameau, critique l'arbitraire de la restriction aux cinq premiers harmoniques :

Le corps sonore ne donne pas seulement, outre le son principal, les sons qui composent avec lui l'accord parfait, mais une infinité d'autres sons [...]. Pourquoi les premiers sont-ils consonants, et pourquoi les autres ne le sont-ils pas, puisqu'ils sont tous également donnés par la nature¹ ?

La justification de Rameau fait appel à la perception, il répond que l'oreille ne perçoit bien que les sons harmoniques qui forment un intervalle de douzième (3^e harmonique) et de dix-septième (5^e harmonique) avec le son fondamental. D'Alembert, quant à lui, s'en remet au statut expérimental de la théorie musicale, qui ne peut avoir la rigueur d'un raisonnement mathématique.

À la limite de la naturalité : le cas de l'accord mineur À partir de la *Génération harmonique* ([Rameau, 1737](#)) et dans les textes suivants, Rameau évoque une justification physique de l'accord mineur à partir d'une seconde expérience de résonance. Cette expérience est la suivante, décrite par d'Alembert :

Si on accorde avec le même corps sonore deux autres corps dont l'un fait à la douzième au-dessous du corps sonore, et l'autre à sa dix-septième majeure en-dessous, ces deux derniers corps frémiront dès qu'on fera raisonner le premier, mais ils ne frémiront point dans leur totalité ; en frémissant ils se diviseront par une espèce d'ondulation, l'un en trois, l'autre en cinq parties égales ; en sorte qu'il y aura pendant le frémissement des point qui resteront en repos².

Autrement dit, la corde *do*5 fait vibrer les cordes *la*3 et *lab*2, dont les fréquences correspondent aux divisions par 3 et par 5 de la fréquence de *do*5, mais ces cordes ne vibrent pas à leur fréquence fondamentale plus grave, elles « *rendent l'unisson* » *do*5. En termes modernes, on dit que la corde aiguë fait résonner leurs modes 3 et 5.

L'accord mineur est de cette manière indirectement rapporté, via l'équivalence des octaves, à la résonance du corps sonore. Pour rendre le raisonnement cohérent avec l'explication du caractère naturel de l'accord majeur,

1. Tiré du *Dictionnaire de musique* de Rousseau ([Rousseau, 1768](#)), article « Harmonie ».

2. Tiré des *Éléments de musique*, ([d'Alembert, 1752](#), chap.I, §23).

Rameau fait appel à la pirouette des harmoniques inférieurs, obtenus à partir d'un son fondamental en prenant les diviseurs de sa fréquence. Ces harmoniques inférieurs se fondent sur un principe de symétrie, qui, selon Rameau, est prescrit par la nature. D'Alembert se montre plus empirique, s'en tenant à l'expérience. Quoi qu'il en soit, la génération de l'accord mineur par le son supérieur fait à son tour apparaître un problème théorique : le son générateur (*do* dans l'exemple précédemment cité) n'est plus le fondamental de l'accord mineur (*fa*).

Des difficultés épistémologiques apparaissent cependant dans l'explication du caractère naturel de l'accord mineur, dont sont conscients Rameau et d'Alembert, et qui seront la source des critiques virulentes de Rousseau. En effet, comme le phénomène physique de résonance est différent, l'accord mineur ne peut avoir le même statut que l'accord majeur. L'accord mineur « *n'est que suggéré, et non véritablement produit, par la résonance du corps sonore*¹ ». D'Alembert précise :

*Comme les sons fa, la bémol et leurs octaves ne font que frémir quand ut résonne, et ne résonnent pas avec lui comme les octaves de mi et de sol ; il s'ensuit de là que le mode ou genre mineur est donné par la nature moins immédiatement et moins directement que le mode majeur*².

Ou encore : « *La nature ne fait que nous indiquer le mode mineur par le frémissement des cordes fa et la bémol*³. »

L'accord mineur, s'il est donné par la nature, ne l'est pas comme phénomène acoustique, mais par l'intermédiaire de signes. Il est virtuellement présent dans les vibrations des cordes graves, qui fournissent par-là des indications à la perception.

*La nature n'offre rien d'inutile, et nous voyons le plus souvent qu'elle se contente de donner à l'art de simples indications, qui le mettent sur les voies*⁴.

À la même époque que Rameau, un théoricien comme Serre, fasciné par la symétrie en musique, pousse plus loin, dans ses *Essais sur les principes de l'harmonie* (Serre, 1753), les implications musicales de cette symétrie entre

1. Tiré de (Charrak, 2001, p.124).

2. Tiré des *Éléments de musique* (d'Alembert, 1752, chap.II, §31)

3. *Ibid.*, (d'Alembert, 1752, chap.IX, §83)

4. Tiré de la *Démonstration du Principe de l'harmonie* (Rameau, 1750, p.66).

le majeur et le mineur, ouvrant la voie au dualisme harmonique du siècle suivant¹.

Implications sur la consonance et la dissonance La subordination des consonances aux sons graves desquels elles sont issues mène Rameau au principe de basse fondamentale d'un accord parfait. Cette fondamentale fait office de centre harmonique, de principe premier fournissant à l'accord sa qualité d'accord. Rameau écrit :

*On ne peut pas déterminer les propriétés d'un intervalle à moins d'avoir au préalable déterminé les propriétés du son fondamental et de l'accord complet qui l'accompagne*².

On a une inversion des relations entre intervalles et accords. Jusqu'ici, et en particulier dans la conception contrapuntique de la consonance et de la dissonance, les propriétés d'un accord dépendent de ses intervalles constituants ; tandis qu'avec Rameau les propriétés des notes et des intervalles constituants sont déterminées par l'accord, lui-même déterminé par le son fondamental. L'accord est une entité première, qui « *s'applique au matériau sonore pour l'organiser*³ ». Rousseau écrit à ce propos :

*On voit que le système harmonique n'est point composé de parties qui se réunissent pour former un tout, mais qu'au contraire c'est de la division du tout ou de l'unité intégrale que se tirent les parties ; que l'accord ne se forme point des sons, mais qu'il les donne*⁴.

Les accords étant des « *unités données immédiatement*⁵ », l'appartenance ou non des notes à ces unités définit leur caractère de consonance ou de dissonance.

Cette nouvelle conception marque un retour à une complète dichotomie des consonances et des dissonances, comme dans la conception mélodique

1. L'école dualiste en Allemagne se fonde sur une opposition symétrique entre les triades majeures et mineures, ainsi qu'entre les modes majeurs et mineurs. Elle regroupe des théoriciens comme Hauptmann, Ettingen et surtout Riemann. Sekimoto ([Sekimoto, 2007](#)) montre que la théorie dualiste est déjà présente dans ses grands traits chez Serre.

2. Tiré du *Traité* ([Rameau, 1722](#)).

3. Tiré de ([Charrak, 2001](#), p.34).

4. Tiré du *Dictionnaire de musique* ([Rousseau, 1768](#)), article « *Système* ».

5. Tiré de ([Dahlhaus, 1967](#)).

(CDC 1), alors que les CDC 2 et CDC 3 impliquaient toutes deux un classement étagé des consonances et des dissonances en plusieurs catégories¹.

Si l'aspect mélodique n'entre pas en compte² dans la détermination du caractère consonant ou dissonant, il subit son influence. En effet, dans l'écriture musicale alors en vigueur, les dissonances ont l'obligation de se résoudre. Elles impliquent un mouvement, contrairement aux consonances, considérées comme stables. On assiste là-aussi à une inversion des causes et des effets par rapport à la conception contrapuntique. En effet, dans la CDC 3, la dissonance était en partie le résultat des mouvements mélodiques, tandis qu'avec Rameau elle devient une de ses causes³.

2.2.2 Au-delà du corps sonore harmonique : lien entre spectre et échelle

Limites de la notion ramiste de naturalité La naturalité selon Rameau se situe dans un corps sonore particulier, le corps sonore harmonique, qui correspond à la corde souple avec extrémités fixes. C'est justement sur cette corde que Descartes et Sauveur ont expérimenté la résonance.

Or, on peut objecter que tous les corps sonores ne sont pas harmoniques, et surtout que tous les corps vibrants employés dans la musique ne sont pas harmoniques. Si Rameau était parti d'une expérience de résonance par sympathie des tambours, il n'aurait pas pu conclure qu'un son fondamental contient son octave, sa douzième et sa dix-septième, et donc à la naturalité de l'accord majeur. Aurait-il conclu à la naturalité du nouvel accord formé des premiers modes de vibration du tambour⁴? Probablement pas, car la théorie ramiste n'est pas rigoureusement fondée sur l'exigence du principe naturel, mais fait appel à des « *raisons de convenance* », pour reprendre le terme de d'Alembert, qui visent à respecter les données perceptives ainsi que les données de la musique tonale de l'époque. Sont révélateurs de ces *raisons de convenance* la restriction aux cinq premiers harmoniques ainsi que la justification de l'accord mineur, toutes deux vivement critiquées par

1. Voir pages 74 et 75.

2. À nuancer toutefois, car le contexte harmonique peut jouer un rôle. Notamment, Tenney (Tenney, 1988, p.82) montre le rôle important de celui-ci dans la détermination de la basse fondamentale et donc de l'accord lorsqu'il y a ambiguïté.

3. Voir (Tenney, 1988, p.78).

4. Pour savoir lequel, se rapporter au livre de référence *Acoustique de instruments de musique* de Chaigne et Kergomard (Chaigne et Kergomard, 2008).

Rousseau.

En outre, le recours à un *principe naturel* implique pour Rameau l'unicité de ce principe, soit une certaine unicité de la nature. Sa définition du son par la propriété de faire résonner les harmoniques est sur ce point révélatrice :

Toute cause qui produit sur mon oreille une impression une et simple, me fait entendre du bruit ; toute cause qui produit sur mon oreille une impression composée de plusieurs autres, me fait entendre du son. J'appelai le son primitif, ou générateur, son fondamental, ses concomitants sons harmoniques, et j'eus trois choses très distinguées dans la nature, indépendantes de mon organe, et très sensiblement différentes pour lui ; du bruit, des sons fondamentaux, et des sons harmoniques¹.

Puisque la musique s'occupe des sons, elle ne peut s'occuper que des sons harmoniques, ce qui exclut, du moins comme fondement théorique, l'usage de sons inharmoniques. Helmholtz, qui par ailleurs loue la volonté de Rameau et d'Alembert de fonder la théorie musicale sur une base scientifique et non métaphysique, critique la restriction de Rameau aux sons harmoniques, et par-là l'étroitesse de sa conception de la naturalité :

En outre, en frappant des verges, des cloches, des membranes, en soufflant dans des cavités, Rameau aurait pu obtenir un grand nombre d'accords dissonants, tout autre que ceux qu'il trouvait avec les cordes et les autres instruments de musique. Il aurait fallu, pourtant, les considérer².

Du reste, les connaissances physiques de l'époque ne sont pas suffisantes pour appuyer théoriquement la décomposition des sons inharmoniques en sons simples. Il faudra attendre les travaux de Fourier sur la propagation de la chaleur (Fourier, 1822), la théorie physiologique de la consonance d'Helmholtz (Helmholtz, 1863) et les expériences perceptives de Plomp et Levelt (Plomp et Levelt, 1965) avant de voir apparaître les premiers travaux sur la consonance et la dissonance des sons inharmoniques (Pierce, 1966; Geary, 1980). Ces approches modernes, qui relèvent de la conception timbrale de la consonance et de la dissonance (CDC 5), bousculent le règne des rapports simples en montrant que la consonance et la dissonance dépendent du spectre, et ne se réfèrent plus à des intervalles ou accords symboliques sur

1. Tiré de la *Démonstration du principe de l'harmonie* (Rameau, 1750, p.12-13)

2. Tiré de la *Théorie physiologique de la musique* (Helmholtz, 1863, p.300)

la partition. Cette dépendance crée une interdépendance très forte entre le spectre, donc le timbre, et l'échelle musicale, autrement dit entre des objets de nature physique et musicale.

Les courbes de dissonance : vers un lien entre spectre et échelle De nos jours, l'approche dominante de la consonance et de la dissonance est celle d'Helmholtz¹. Le scientifique relie la dissonance à la rugosité, qui est la sensation désagréable due aux battements rapides entre partiels. L'explication de l'harmonie ne se situe plus directement dans le corps sonore résonnant, mais dans la sensation et dans le fonctionnement des organes auditifs.

L'apport majeur d'Helmholtz, qui permet de relier le spectre aux intervalles consonants, consiste en la représentation des courbes de dissonance². Sur l'échelle continue de deux octaves, qui figure l'intervalle entre deux sons, des maxima et minima locaux apparaissent, qui deviennent d'autant plus nombreux que le nombre d'harmoniques est important. Chaque « bosse » est due à l'interaction entre deux partiels : les battements qu'ils créent lorsqu'ils sont suffisamment proches contribue à la dissonance, tandis que leur coïncidence crée un minimum local de dissonance. Les courbes font ainsi apparaître plusieurs intervalles remarquables qui maximisent la consonance et correspondent aux intervalles entre les partiels du son utilisé. Étant donné qu'Helmholtz utilise des sons harmoniques, ces minima tombent sur les intervalles justes, qui correspondent aux rapports simples. Ces intervalles constituent une échelle qui en un certain sens maximise la consonance étant donnée une distribution harmonique des partiels. Ériger cette échelle en échelle musicale pose de nombreux problèmes théoriques et pratiques, et l'histoire de tempérament peut être interprété comme l'histoire d'un compromis entre la justesse des intervalles et la pratique musicale.

Imaginons maintenant que Helmholtz soit allé un cran plus haut jusqu'au 11^e harmonique dans la représentation des courbes de dissonance. À côté du minimum local de dissonance correspondant au rapport $7/5$, à seulement 17 cents en-dessous de la quinte diminuée dans la gamme de tempérament égal, apparaîtrait le minimum local de rapport $11/8$, à 49 cents en-dessous de la quinte diminuée, soit quasiment sur le quart-de-ton entre la quarte et la quinte diminuée.

1. L'approche d'Helmholtz sera présentée dans la partie [2.3](#).

2. Présentées page [45](#).

Sons inharmoniques et échelles de notes Si les sons qui servent à calculer les courbes de dissonance sont des sons inharmoniques, les minima des courbes de dissonance ne tombent plus sur les rapports simples. Les intervalles considérés comme consonants comme l’octave peuvent devenir dissonants, au sens de la rugosité, tandis que des intervalles considérés comme dissonants peuvent à leur tour devenir consonants. Avec de tels sons, l’échelle chromatique sonnera comme désaccordée, car les intervalles la constituant ne minimiseront plus la dissonance.

Dans son livre *Tuning, timbre, spectrum, scale* (Sethares, 1999), Sethares montre, en s’appuyant sur des extraits sonores, à quel point la dépendance entre spectre et échelle musicale est profonde.

Un chapitre du livre est consacré au gamelan javanais. Le gamelan est un orchestre traditionnel constitué d’instruments à percussion, parmi lesquels des instruments formés de barres métalliques, nommés *saron*, et des instruments formés de résonateurs métalliques en forme de pots retournés, nommés *bonang*. Les *saron* et *bonang* produisent des sons inharmoniques. Les instruments sont accordés selon deux échelles : une échelle pentatonique nommée *slendro*, ou une échelle heptatonique nommée *pelog*. Les deux échelles coexistent au sein d’un même gamelan, et ont généralement une note en commun. Les échelles ne sont généralement pas octaviantes, et aucune des échelles ne recoupe l’échelle chromatique. En étudiant les instruments et les accords de deux gamelans, Sethares montre que le spectre des instruments est relié aux intervalles des échelles utilisées. Plus précisément, les intervalles du *slendro* sont fortement corrélés avec les minima de dissonance d’un son du *bonang* avec un son harmonique, tandis que les intervalles du *pelog* sont corrélés avec les minima de dissonance d’un son du *saron* avec un son harmonique. Ainsi, les échelles sont adaptées aux instruments tout en étant compatibles avec l’ajout d’un instrument harmonique, courant dans la musique de gamelan de Java. Ces résultats montrent la puissance des courbes de dissonance et mettent en évidence une certaine universalité de la notion de consonance timbrale (CDC 5), universalité par ailleurs remise en question par d’autres études interculturelles (McDermott *et al.*, 2016).

Sethares fournit une méthode générale pour jouer de la musique à partir de sons inharmoniques. Voici les différentes étapes :

1. Analyse spectrale ;

2. Simplification du spectre en un spectre en pics ;
3. Calcul de la courbe de dissonance ;
4. Choix de l'échelle, constituée d'intervalles compris entre deux minima de la dissonance ;
5. Construction d'un instrument qui joue avec ce spectre et sur cette échelle (éventuellement à l'aide d'un *sampler*) ;
6. Composition et performance.

Ces étapes impliquent plusieurs conditions, notamment d'avoir des sons relativement stables d'un point de vue spectral, pour pouvoir associer un spectre, ou encore d'avoir un son suffisamment riche pour que la courbe de dissonance ait plusieurs minimums locaux, afin de pouvoir choisir une échelle intéressante du point de vue musical. De plus, le son ne doit pas être trop bruyé, sans quoi les partiels seront trop nombreux ou indistinguables. La figure 2.1 montre la procédure préalable à la composition de la pièce *Tingshaw* pour cloche. Sethares part d'un son de cloche, en fait une analyse spectrale, simplifie le spectre en un spectre en amplitude, dessine la courbe de dissonance sur deux octaves, et choisit une échelle possible parmi les échelles dont les intervalles sont compris entre deux minima de dissonance. Le choix de l'échelle en question, à huit notes, est motivé par l'intuition et l'intention musicale. Sethares réalise un *sampling* du son de cloche initial et fait correspondre les huit notes de l'échelle avec huit touches d'un synthétiseur. L'échelle retenue n'est pas égale (les intervalles entre les notes successives ne sont pas égaux) et non octaviante (l'octave est légèrement dilatée, avec un rapport 2,02 à la place de 2).

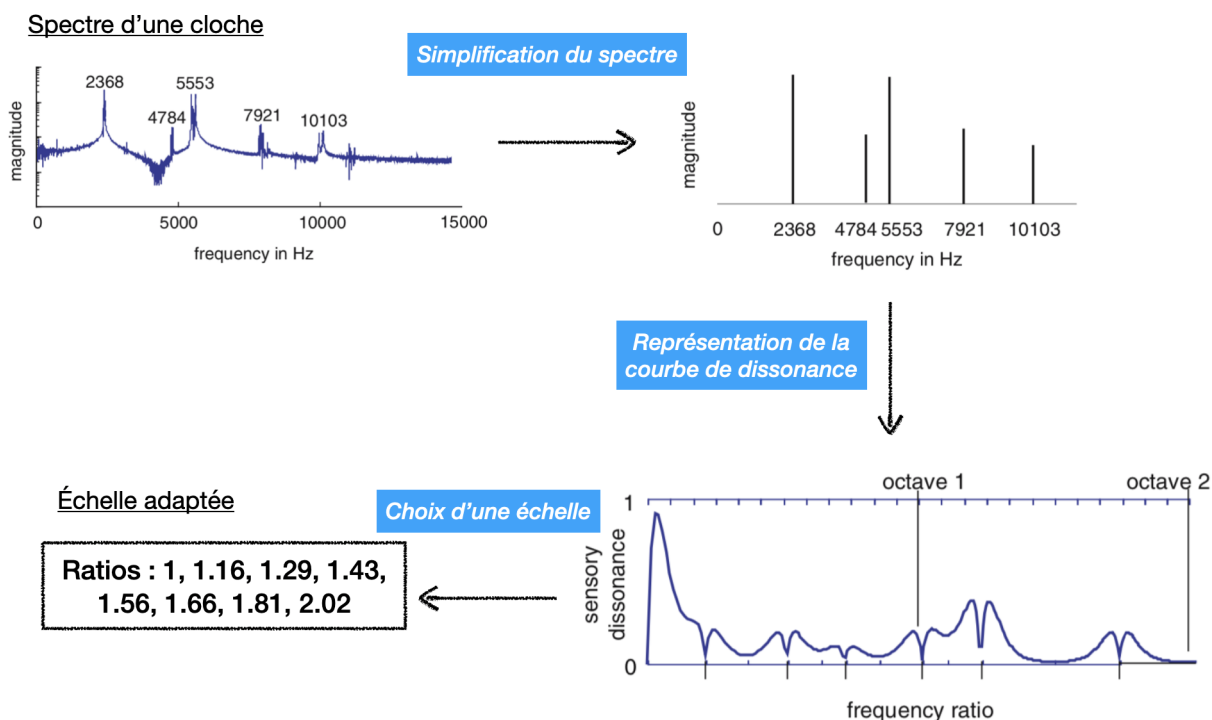


FIGURE 2.1 – Procédure proposée par Sethares (Sethares, 1999, chap.7) pour associer une échelle de hauteurs à un instrument inharmonique, ici une cloche. Sethares part du spectre de la cloche, le simplifie en spectre à pics, trace la courbe de dissonance sur deux octaves, et choisit certains intervalles parmi les intervalles formés par les minimums de rugosité pour construire une échelle. Celle-ci est adaptée à la cloche, en ce sens qu'elle optimise en un certain sens la consonance.

Dans la suite du livre, Sethares propose une méthode pour associer un spectre à une échelle donnée, de manière à maximiser la consonance des intervalles de l'échelle. L'idée est la suivante : de même que n'importe quel intervalle est une consonance pour un spectre adapté, n'importe quelle échelle est accordée (c'est-à-dire constituée d'intervalles justes) pour un bon spectre. À moins de fixer des conditions fortes sur les spectres et les échelles, la correspondance entre spectre et échelle n'est pas unique.

Du Rameau tempéré et inharmonique Pour Sethares, le lien entre l'objet physique qu'est le son et la pratique musicale va plus loin que la correspondance entre spectre et échelle. Il montre qu'il est possible d'associer à un spectre et à une échelle des accords et des progressions entre accords.

Sethares part de l'échelle qui subdivise l'octave en dix intervalles égaux (échelle appelée 10-tet), très éloignée de l'échelle chromatique puisqu'elle ne partage avec cette dernière que l'intervalle de quarte augmentée. Les notes de cette échelle, qu'il nomme E, F, G, H, I, J, A, B, C et D, sont séparées d'un rapport $r = 2^{1/10}$, contre $r = 2^{1/12}$ pour l'échelle chromatique. Sethares associe un spectre adapté à l'échelle, c'est-à-dire tel que les minima de la courbe de dissonance coïncident avec des intervalles de l'échelle. La distribution du spectre est donnée figure 2.2. Le spectre est octaviant, et la comparaison avec des spectres harmoniques fournit des analogues de la quinte et de la tierce majeure : ce sont les intervalles formés respectivement de 6 et de 3 intervalles élémentaires. L'accord formé de la fondamentale, la tierce majeure et la quinte est appelée *accord neutre*. Rameau dirait que l'accord neutre est *contenu* dans un corps sonore résonnant émettant ce spectre, ce qui lui confère une naturalité ramiste analogue au caractère naturel de l'accord majeur vis-à-vis du corps sonore résonnant harmoniquement. Faute de corps sonore résonnant avec ce spectre présent dans la nature, le caractère naturel change de statut épistémologique pour se situer dans les sons ayant ce spectre, éventuellement produits ad hoc par des moyens numériques. La justification devient alors physicaliste, plus que naturaliste.

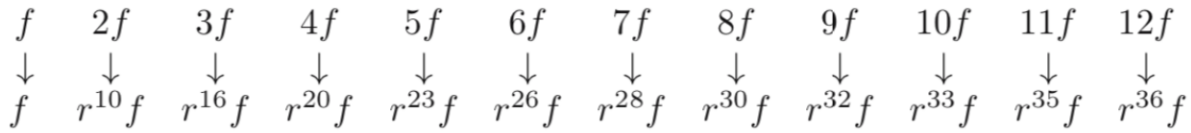
Étant donné que la tierce majeure divise la quinte en deux intervalles égaux, la justification ramiste de l'accord mineur¹ donnerait pour analogue à l'accord mineur ce même accord neutre, d'où son nom.

Un autre moyen d'aboutir à l'accord neutre est de dessiner à partir du spectre la surface de dissonance, analogue à la courbe de dissonance pour les triades. Celle-ci fait apparaître plusieurs minima locaux, parmi lesquels l'accord neutre et ses renversements (région A sur la figure 2.2).

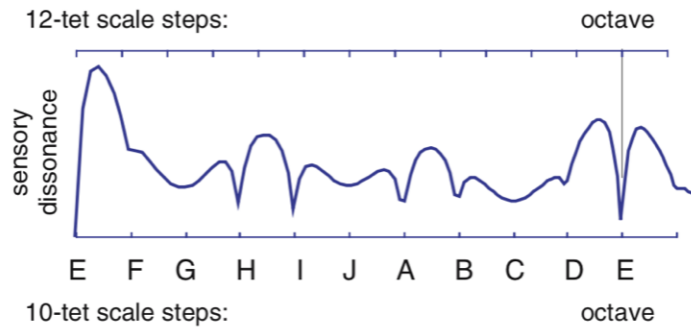
À partir de cet accord neutre, Sethares construit un équivalent au cycle des tierces par un raisonnement de type néo-riemannien, puis, par analogie formelle avec les comportements tonaux, des enchaînements cadentiels I-IV-V, un accord de septième de dominante, et plusieurs gammes dans cette échelle, entre lesquelles il est possible de moduler.

1. Voir page 95.

Spectre



Courbe de dissonance



Surface de dissonance

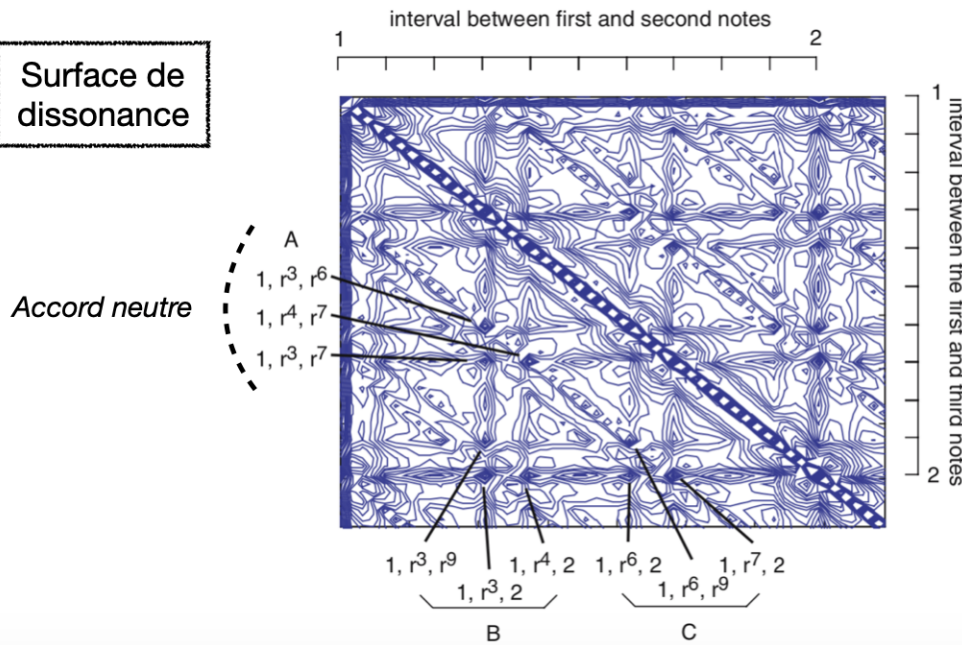


FIGURE 2.2 – Échelle subdivisant l'octave en 10 intervalles égaux (10-tet). Association d'un spectre adapté, puis à partir de ce spectre représentation des courbe et surface de dissonance. Figure tirées de (Sethares, 1999, chap.14). Pour la construction du spectre, $r = 2^{1/10}$.

2.3 Approches perceptives

Depuis Pythagore, la perception joue un rôle important dans la détermination des intervalles consonants et dissonants. Certains auteurs lui accordent une place d'honneur dans la théorie de la consonance et de la dissonance. C'est le cas d'Aristoxène de Tarente, élève d'Aristote, auteur d'un des plus anciens ouvrages de musique connus, les *Éléments harmoniques* ([Aristoxène, 1870](#)). Aristoixène dénonce les abstractions pythagoriciennes, qu'il considère comme des « *spéculations étrangères au sujet et directement contraires aux phénomènes*¹ ». Les phénomènes dont parle Aristoixène sont précisément les phénomènes auditifs. Dans la théorie des coups², les phénomènes auditifs sont assimilés aux coups produits par les sons sur le tympan, et la perception des intervalles est ramenée à la régularité de ces coups. Il faudra attendre Helmholtz au XIX^e siècle pour voir apparaître une théorie de la consonance et de la dissonance fondée rigoureusement sur les phénomènes auditifs.

2.3.1 La révolution perceptive

L'ouvrage d'Helmholtz *Théorie Physiologique de la musique fondée sur l'étude des sensations auditives*, paru en 1863, marque une véritable révolution épistémologique dans la théorie de la consonance et la théorie de la musique en général, en l'ouvrant aux champs de la physiologie et de l'étude des sensations auditives. Plus empirique que ses prédécesseurs, il donne un véritable fondement scientifique à la théorie de l'harmonie, en usant de principes expérimentaux de physique et de psychologie. Il est en outre l'un des premiers à appliquer la théorie de Fourier à l'analyse des sons.

Helmholtz s'intéresse à la structure anatomique de l'oreille, et cherche les causes de la consonance et de la dissonance non plus directement dans le phénomène sonore, mais dans son effet sur les organes sensoriels. On assiste ici à un véritable glissement de paradigme, qui fait passer des principes naturels extérieurs à l'homme à la nature humaine.

Ce glissement de paradigme a été initié un siècle plus tôt par Hume, père fondateur de l'esthétique expérimentale, pour qui le beau ne pré-existe pas à l'homme mais se situe dans une adéquation de l'objet avec celui qui l'expérimente : « *La beauté n'est pas une qualité qui se trouve dans les choses*

1. Tiré de ([Aristoxène, 1870](#))

2. Voir partie [2.1.3](#).

elles-mêmes, elle n'existe que dans l'esprit qui les contemple¹. » Au même moment, Pierre Estève, pour qui « *le principe des sons est mécanique*² », amorce ce glissement dans la théorie de la musique, en définissant l'agrément que procurent les consonances par l'action physique des sons sur l'organe de l'ouïe.

L'oreille comme analyseur spectral Étant donné que « *dans la musique, les sensations auditives sont précisément ce qui forme la matière de l'art*³ », Helmholtz se propose d'élaborer une théorie des sensations auditives, au même titre qu'il élaborera une théorie des sensations visuelles dans son *Optique physiologique* (Helmholtz, 1867). Cette théorie tombe à la fois dans le domaine de l'acoustique physique, qui est la science du son en tant que mouvement vibratoire de l'air, et celui de l'acoustique physiologique, qui décrit le fonctionnement des organes auditifs jusqu'aux terminaisons des voies nerveuses, qui permettent de passer du son à la sensation.

À l'aide de résonateurs, de membranes, de sirènes, et par le recours à la sensation auditive, Helmholtz montre expérimentalement la loi d'Ohm, qui s'énonce ainsi :

L'oreille décompose tout mouvement périodique [non pendulaire⁴] de l'air en une série de vibrations pendulaires qui correspondent chacune à la sensation d'un son simple⁵.

L'oreille se comporte comme un analyseur de fréquences qui rend effective et concrète la transformation de Fourier d'un son complexe en une somme de sons simples. Reprenant à son compte le phénomène de résonance, Helmholtz considère que les fibres qui tapissent la membrane basilaire de l'oreille interne sont de résonateurs, au même titre que les cordes du piano. Par cette représentation il réinvestit l'idée, répandue à son époque sous des formes vagues, de clavier auditif⁶.

1. Tiré de *The Standards of Taste* (Hume, 1757), traduction par Philippe Folliot, disponible à l'adresse http://philotextes.info/spip/IMG/pdf/regle_du_gout.pdf.

2. Pierre Estève, *Nouvelle découverte du principe de l'harmonie avec un examen de ce que M. Rameau a publié sous le titre de Démonstration de ce principe* (Estève, 1752, p.50)

3. Tiré de (Helmholtz, 1863, p.4).

4. Une vibration pendulaire est une vibration sinusoïdale

5. Tiré de (Helmholtz, 1863, p.78)

6. L'idée de clavier auditif est présente dès le XVII^e siècle. Duverney, dans son *Traité de l'organe de l'ouïe* (Duverney, 1683), place le clavier dans la spirale du limaçon de l'oreille interne.

La vibration de l'air fait vibrer le tympan, qui lui-même transmet ses vibrations au liquide de la cochlée par l'intermédiaire des osselets. Ce liquide à son tour excite les fibres de la membrane basilaires, distribuées sur la membrane selon leur fréquence, et chaque fibre est excitée sur un voisinage de sa fréquence propre. Ainsi, deux fréquences différentes mettront en vibration différentes fibres.

L'idée fondamentale, qui reste vraie malgré les erreurs d'Helmholtz sur le fonctionnement de la cochlée, est celle d'une topographie de la membrane selon les fréquences : les régions vibrantes de la membrane sont distribuées selon les fréquences d'excitation. Lors de l'audition d'un son complexe, chaque partiel fait vibrer la région correspondante.

La sensation de rugosité Helmholtz observe que lorsque deux sons simples sont suffisamment proches pour ne pas être clairement distinguables, et suffisamment éloignés pour ne pas être confondus en un seul son, la sensation produite à l'oreille est désagréable, du fait des battements. Cela arrive lorsque les sons sont séparés d'environ 30 Hz. À partir de là, Helmholtz relie la sensation de dureté d'un intervalle de deux sons complexes aux battements produits par les harmoniques, et il assimile cette dureté, ou rugosité, à la dissonance. La consonance correspond alors à l'absence de battements. Une telle explication de la dissonance par les battements n'est pas nouvelle, on la trouve déjà au siècle précédent chez Sauveur¹ ou chez l'anglais Robert Smith². Mais avec la décomposition de Fourier d'un son complexe, Helmholtz connaît l'origine physique des battements et est capable de quantifier le phénomène.

De plus, Helmholtz appuie cette explication par une hypothèse physiologique. En effet, deux partiels proches font vibrer deux régions proches de la membrane, qui partagent un même voisinage, lequel vibre selon l'addition des deux vibrations. Cette addition reproduit sur le voisinage les battements entre les sons, dont la fréquence correspond à la différence des fréquences.

1. « *Les battements ne plaisent pas à l'oreille, à cause de l'inégalité du Son, et l'on peut croire avec beaucoup d'apparence, que ce qui rend les Octaves si agréables, c'est qu'on n'y entend jamais de battements. En suivant cette idée, on trouve que les accords dont on ne peut entendre les battements, sont justement ceux que les Musiciens traitent de Consonances, et que ceux dont les battements se font sentir, sont les Dissonances, et que quand un accord est Dissonance dans une certaine octave et Consonance dans une autre, c'est qu'il bat dans l'une et qu'il ne bat pas dans l'autre.* », Joseph Sauveur, tiré de l'article *Sur la détermination d'un son fixe* (Sauveur, 1700, p.143).

2. Robert Smith, *Harmonics, or the philosophy of musical sounds* (Smith, 1759).

Ainsi, cette région ne vibre plus seulement à sa fréquence excitatrice, mais à la fréquence des battements. L'intermittence des battements perturbent les terminaisons nerveuses reliées.

*Les battements produisent sur les nerfs de l'audition une impression dure et désagréable, parce que toute excitation intermittente de nos appareils nerveux agit plus vivement qu'une excitation régulière et durable*¹.

Helmholtz en vient à envisager la consonance et la dissonance selon l'aspect continu ou discontinu de la sensation :

*La consonance est une sensation continue, et la dissonance une impression intermittente. Deux sons consonants résonnent régulièrement l'un à côté de l'autre, sans provoquer l'un dans l'autre des perturbations réciproques ; deux sons dissonants se divisent en une série de secousses isolées*².

La *sensation continue* est synonyme de d'agrément, et l'*impression intermittente* de désagrément. On n'est pas loin ici de la théorie des coups en vigueur au XVII^e siècle³. Aussi Helmholtz y apporte-t-il une justification physiologique absente et inaccessible au XVII^e siècle.

Helmholtz généralise aux sons complexes⁴ la dissonance des sons simples en sommant les contributions à la dissonance des couples de partiels, avec une pondération en fonction des amplitudes des partiels. Il est ainsi en mesure de dresser les courbes de dissonance des intervalles en faisant varier le nombre de partiels⁵.

En complément des battements des harmoniques, Helmholtz introduit les battements des *sons résultants*, notamment pour expliquer la dissonance entre des sons simples sans harmoniques. Ces sons, décrits un siècle plus tôt⁶,

1. Tiré de (Helmholtz, 1863, p.290).

2. *Ibid.*, p.291

3. Voir partie 2.1.3.

4. Un son est dit complexe lorsqu'il n'est pas simple, c'est-à-dire sinusoïdal. Un son complexe périodique peut être décomposé, via l'analyse de Fourier, en une somme de sons simples.

5. Voir figure 1.3 page 45.

6. Il semble que les sons résultants aient été découverts presque simultanément en Italie par Tartini (Tartini, 1754), en France par Romieu (1751) et en Allemagne par Sorge (1745-1747).

proviennent des défauts de linéarité sous certaines conditions de production et la propagation du son. On les entend « *lorsque deux sons musicaux de différente hauteur sont émis simultanément avec force, et se prolongent avec régularité*¹ ». Helmholtz montre expérimentalement leur réalité physique.

Les sons résultants nécessitent la production simultanée dans un même espace d'au moins deux sons. Helmholtz distingue les *sons différentiels* des *sons additionnels*, les premiers correspondant à une vibration à une fréquence égale à la différence des fréquences des sons primaires, les seconds à leur somme. En présence de sons complexes, les sons additionnels et différentiels peuvent se créer entre les sons fondamentaux ou les partiels.

Les sons résultants peuvent battre entre eux, ou avec les harmoniques, participant ainsi de la rugosité. Dans la théorie de la rugosité, Helmholtz traite les sons résultants et les sons harmoniques sur le même plan.

Avec Helmholtz, la consonance et la dissonance sont de l'ordre de la sensation. Elles ne se réfèrent plus à une dimension dynamique comme c'était le cas de la CDC 3 et de la CDC 4 via les notions de résolution, de mouvement obligé, de note étrangère, pas plus qu'à l'idée de stabilité et d'instabilité. Elle se réfère uniquement à une qualité perceptive. À ce titre, la nouvelle conception de la consonance et de la dissonance se rapproche de la CDC 2. Elle en diffère néanmoins par sa dépendance forte aux caractéristiques acoustiques des sons : timbre, registre, nuances ; dépendance absente dans les formes précédentes de la consonance et de la dissonance. De plus, la consonance et la dissonance étant maintenant quantifiables, elles ne sont plus seulement la propriété d'agrégats sonores, mais peuvent se référer à des notes isolées, si sont prises en compte les interactions entre partiels proches d'un même spectre :

*Pour certains timbres où les harmoniques supérieures sont très développés, il peut se produire, même dans un seul son, des dissonances dont la dureté est appréciable à l'oreille*².

Cet aspect de la consonance est propre à la conception timbrale de la CDC 5.

Coïncidence des harmoniques À côté de la définition de la consonance par l'absence de battements, une seconde définition est présente chez Helm-

1. Tiré de (Helmholtz, 1863, p.191)

2. Tiré de (Helmholtz, 1863, p.240).

hertz, où la consonance est caractérisée positivement par la coïncidence de partiels entre les notes :

*Nous avons caractérisé les consonances par la coïncidence de deux harmoniques des deux sons fondamentaux*¹.

Là-encore cette approche n'est pas nouvelle, mais date du siècle dernier. On la doit à Pierre Estève, qui la juge meilleure que la théorie des battements².

La théorie des battements donne une définition positive de la dissonance, la théorie des harmoniques communes une définition positive de la consonance. Malgré le fait qu'Helmholtz traite plus de la première que de la seconde, les deux approches coexistent dans son ouvrage, sans qu'il n'en vienne à aucun moment à les confronter réellement, ce qui invite à considérer la consonance et la dissonance autrement que comme deux pôles opposés d'un même continuum.

L'approche de la consonance par la coïncidence des harmoniques est moins fertile pour Helmholtz que celle par les battements, car il ne la généralise pas aux triades ou aux accords de plus de deux sons.

Statut épistémologique de la théorie Helmholtz est le premier, dans l'histoire de la théorie de l'harmonie, à dépasser la « *barrière du tympan* », pour aller voir dans la « *boîte noire* »³ située au-delà.

Au siècle précédent, l'oreille est vue essentiellement comme un organe qui reproduit dans le corps les mouvements de la nature. « *La configuration harmonique produirait ainsi son double dans l'organe sensoriel* », note Chartrak (Chartrak, 2001). Rameau reprend cette hypothèse mécaniste dans la *Génération harmonique* :

Ce qu'on a dit des corps sonores doit s'entendre également des fibres qui tapissent le fond de la conque de l'oreille; ces fibres

1. *Ibid.*

2. « Je ne dis point que les battements soient tout à fait étrangers au sentiment de l'Harmonie, ils se trouvent dans les accords, mais n'en sont pas le principe premier. [...] La différence essentielle et fondamentale des accords, n'est ni les combinaisons qu'avait donné M. Rameau, ni une sympathie, ni un jugement de l'âme, ni un sentiment réfléchi, ni la variété des battements. Cette différence est établie par l'action mécanique des harmoniques conservés dans les consonances et détruits dans les dissonances », tiré de la *Nouvelle découverte du principe de l'harmonie* (Estève, 1752, p.53).

3. Expressions de Bailhache tirées de (Bailhache *et al.*, 2011, p.166)

*sont autant de corps sonore auxquels l'air transmet ces vibrations, et d'où le sentiment des sons et de l'harmonie est porté jusqu'à l'âme*¹.

Un corps sonore harmonique dans la nature fait vibrer son analogue dans l'oreille. Ainsi, la nature à laquelle se réfèrent Rameau et d'Alembert est une nature extérieure à l'homme. Avec Helmholtz, les principes fondant l'harmonie *a priori* s'effondrent, le système des sons est organisé par les lois de la nature humaine, en l'occurrence par l'appareil auditif qui transforme les sons en sensations.

Ce transfert de naturalité change en profondeur les prétentions d'une approche se voulant naturelle. En effet, le beau perdant son statut de principe immuable et universel présent dans la nature, il devient historique. Une théorie de la musique ne peut donc être qu'une dogmatique, au sens où l'emploie Dahlhaus, c'est-à-dire une explication structurée de principes situés historiquement et géographiquement.

Aussi Helmholtz ne prétend-il pas fonder scientifiquement le système musical :

*Le système des gammes, des modes et de leur enchaînement harmonique, ne repose pas sur des lois naturelles invariables, mais il est, au contraire, la conséquence de principes esthétiques qui ont varié avec le développement progressif de l'humanité, et qui varieront encore*².

Le titre original de l'ouvrage d'Helmholtz, improprement traduit en français par « *Théorie physiologique de la musique...* » mais plus proche de « *Théorie physiologique des sensations auditives...* », est révélateur de la démarche du scientifique. Il s'agit dans cet ouvrage de donner des explications acoustico-physiologiques des sensations, qui constituent les « *éléments de la construction de notre système musical* ».

Helmholtz distingue le bien être des sens, réductible au fonctionnement de l'oreille, du beau esthétique, ou « principe de style », lesquels constituent selon lui la finalité de la musique, mais ne découlent pas d'un principe naturel. En effet, ce « principe naturel », qui pour Rameau donnait à la fois les moyens en même temps que le but (imitation de la nature), ne fournit pour Helmholtz que les moyens, entendus au sens de conditions de possibilités.

1. Tiré de la *Génération harmonique*. (Rameau, 1737, Prop.XII, p.7)

2. Tiré de (Helmholtz, 1863, p.306).

Dahlhaus note à ce propos :

La « nature des choses » n'est pas un principe duquel on pourrait déduire un quelconque « système naturel » de la théorie musicale, mais un ensemble de possibilités qui ont été ou non réalisées selon un choix fluctuant, dépendant des « principes de style » qui changent au cours du temps¹.

Pour Helmholtz, le classement des consonances et les courbes de dissonance sont fondés dans la nature des choses, ils constituent un matériel de base pour la construction musicale, mais faute de la déterminer ils ne peuvent que la motiver :

La science peut cependant s'efforcer de déterminer, les mobiles psychologiques ou techniques, qui ont agi dans cette recherche de l'artiste. Les mobiles psychologiques tombent dans le domaine de l'esthétique scientifique, les mobiles techniques, dans celui de la science².

2.3.2 Étude psychologique de la sensation : la fusion tonale

Helmholtz, nous l'avons vu, entend fonder l'étude des sensations sur l'acoustique physique et l'acoustique physiologique. La loi de Fourier appliquée aux sons devient non seulement une réalité physique, mais également une réalité physiologique, puisque chaque vibration simple d'un son complexe fait vibrer des zones spécifiques de l'oreille interne, reliées aux terminaisons nerveuses. Ce qui se passe dans le cerveau reste alors un mystère. Autrement dit, la boîte noire qui séparait le tympan à la sensation est repoussée aux extrémités des terminaisons nerveuses, mais elle demeure une boîte noire. Le philosophe Charles Lalo écrit, à propos de l'hypothèse physiologique d'Helmholtz :

La théorie est tout au moins inutile : à quoi sert de doubler la vibration extérieure, ou celle du tympan, d'une deuxième vibration des extrémités nerveuses ? Comprendrons-nous mieux pour cela qu'elles soient perçues ? Pourquoi ne pas lui ajouter une troisième

1. Dahlhaus, *Hermann von Helmholtz et le caractère scientifique de la théorie de la musique*, traduction de Céline Vautrin tirée de (Bailhache *et al.*, 2011, p.203).

2. Tiré de (Helmholtz, 1863, p.306).

*vibration des centres, ou du trajet nerveux ? Où cessera enfin la vibration, pour que la conscience commence*¹ ?

Le réductionnisme de la sensation à ses aspects mécaniques est vivement critiqué par les contemporains d'Helmholtz², qui lui opposent la primauté de la sensation. Lalo décrit l'inanité de l'approche physiologique :

*Nous ignorons l'essentiel des faits physiologiques. Mais les connaissons-nous jusque dans leur détail chimique et dans leur mécanisme moléculaire, cela ne nous permettrait que deux choses : unifier les faits de plus haut, et expliquer leurs exceptions et leurs limites, que nous constatons sans les expliquer. Mais cette prétendue « réduction » du psychique au physique n'ajouterait rien à notre connaissance du fait psychologique de la consonance comme tel : il est en lui-même et tout entier connaissable*³.

Le fait psychologique de la consonance appelle donc une théorie psychologique des sensations. C'est Carl Stumpf qui la propose.

La fusion tonale Stumpf est un psychologue et philosophe allemand dont les idées, notamment les travaux sur l'articulation entre le tout et les parties dans la perception, ont eu une grande influence sur la psychologie de la forme (théorie de la Gestalt), ainsi que sur la phénoménologie de Husserl, qui fut son élève.

Pour Stumpf, la tendance qu'ont deux sons simultanés à être perçus comme un seul son est à l'origine du phénomène de consonance, c'est la *fusion*, ou « *Verschmelzung* ». Loin d'être neuve, cette idée se trouve déjà chez Euclide⁴, et la notion de *sonner comme un* (« *quasi sonus unus*⁵ ») joue un rôle essentiel dans les écrits théoriques de l'Ars Antiqua⁶.

1. Charles Lalo, *Esquisse d'une esthétique musicale scientifique*, (Lalo, 1908, p.105). Dans son ouvrage, Lalo entreprend d'unifier la science et l'esthétique. Ce faisant, il livre une remarquable synthèse critique des théories de la consonance et de la dissonance.

2. Parmi lesquels le psychologue Carl Stumpf (Stumpf, 1883, 1890) ou le physicien Ernst Mach.

3. *Ibid.*, p.122

4. « *La consonance est l'association de deux sons, l'un aigu, l'autre grave. La dissonance, au contraire, est l'inaptitude de deux sons à s'associer.* », Euclide, cité et traduit par Helmholtz (Helmholtz, 1863, p.291).

5. Tiré du *Speculum musicae* de Jacques de Liège (Liège, 1325, Livre VII, chap.IV). Jacques de Liège définit la concorde par cette propriété de *sonner comme un*.

6. « *Il est possible que par le terme de "concorde parfaite", les théoriciens du XIII^e*

Stumpf définit ainsi la fusion, dans son ouvrage *Tonpsychologie* (Stumpf, 1883, 1890) :

*Ce rapport de deux contenus, spécialement de deux contenus sensoriels, en vertu duquel ils ne forment pas une simple somme, mais un tout. Il résulte de ce rapport que plus ses degrés s'élèvent, plus l'impression totale se rapproche, toutes choses égales d'ailleurs, de celle d'une seule et unique sensation, et que l'analyse en devient de plus en plus difficile. Ces conséquences peuvent servir à définir la fusion*¹.

La fusion est une expérience consciente qui se rattache à l'analyse des sensations. Stumpf la voit comme un phénomène primaire, indivisible de la conscience, une loi fondamentale de la perception qui est la cause de la consonance. En cela il s'oppose à la vision réductionniste et atomiste d'Helmholtz. Ce faisant, il définit la consonance positivement, et la dissonance négativement, comme absence de fusion, quand la théorie des battements approchait la consonance de manière apophatique. Une définition positive de la consonance est toutefois présente chez Helmholtz, quoique minoritaire, c'est la coïncidence des harmoniques. Celle-ci est liée au phénomène psychologique de ressemblance, que Stumpf ne voit que comme une conséquence de la fusion.

La démarche de Stumpf, différente de celle du physicien Helmholtz, tient de la psychologie expérimentale. Il fait écouter des intervalles à un grand nombre de sujets, tous non musiciens, leur demandant de juger s'ils entendent un ou deux sons. Le degré de fusion d'un intervalle est fonction du pourcentage de sujets ayant perçu un son lorsqu'il y en avait deux. L'octave vient en premier, puis la quinte, la quarte, les tierces et les sixtes.

Stumpf se défend de subjectivité de la notion de fusion en la considérant comme une propriété des agrégats sonores, qui n'est pas établie par l'auditeur, mais reconnue par lui. Stumpf distingue deux types d'écoutes : l'écoute

siècle aient eu en vue cette propriété des dyades simultanées à sonner comme un seul son, et qu'ils distinguaient différents degrés de consonance et de dissonance des dyades par cette condition de 'singularité'. » Tenney, tiré de (Tenney, 1988, p.28), traduction de l'auteur. "I suggest that what 13th-century theorists may have actually meant by 'perfect concord' was a condition in which a simultaneous dyad sounded like a single tone, and that they distinguished varying degrees of consonance (and dissonance) according to the extent to which a given dyad satisfied this condition of 'singularity'."

1. Carl Stumpf, *Tonpsychologie* (Stumpf, 1883, 1890, II, p.128), traduit par Lalo dans (Lalo, 1908, p.124).

intégrale, qui perçoit l'unité d'un tout, plus ou moins grande selon son degré de fusion, et l'écoute plurielle, qui perçoit la multiplicité des parties. Ces deux modes de perception ne sont pas exclusifs, mais coexistent. Tandis que l'écoute intégrale est accessible à tous, la distinction des notes dans un accord requiert un entraînement et est favorisée par l'attention. Celle-ci ne modifie pas le degré de fusion, mais permet de permuer les écoutes. Les notes d'un accord restent distinguables si l'oreille est suffisamment exercée, et c'est le pouvoir de l'attention de permuer les écoutes.

Stumpf énonce une série de *lois de la fusion*, qui sont établies « *par de méthodes purement psychologiques, d'après l'observation des faits, sans aucun emprunt à une hypothèse quelconque*¹ » :

- Le degré de fusion dépend des rapports des nombres de vibrations sonores.
- Le degré de fusion est indépendant de la hauteur absolue des sons considérés.
- Le degré de fusion est indépendant de l'intensité, aussi bien de l'intensité absolue que de l'intensité relative des deux sons.
- La fusion de deux sons donnés n'est pas influencée par l'addition d'un troisième son.
- Le degré de fusion de deux sons n'est pas modifié par une altération minime du rapport de leurs vibrations.
- La fusion reste la même quand on entend chacun des deux sons par une oreille différente.
- La fusion reste identique pour deux sons imaginés et pour les deux mêmes sons perçus objectivement.
- La fusion d'un intervalle n'est pas modifiée par l'addition d'une ou plusieurs octaves à l'intervalle simple.
- La durée des sons ne change rien à leur degré de fusion.

Par l'indépendance du degré de fusion à la hauteur absolue et aux nuances relatives, Stumpf se place aux antipodes des conclusions d'Helmholtz, de même qu'avec l'absence d'influence d'un troisième son. La quatrième loi sur l'écoute dichotique² répond à l'une des critiques principales adressées à Helm-

1. Tiré de (Lalo, 1908, p.132)

2. On appelle *écoute dichotique* la configuration d'écoute où des sons différents sont

holtz par ses contemporains : présentées aux oreilles différentes, deux sons ne peuvent créer des battements dans l'oreille interne, d'où il résulte, si les battements sont la cause unique de la dissonance, que celle-ci ne peut pas être perçue en écoute dichotique ; ce qui n'est pas vérifié en pratique.

Tandis que le timbre joue un rôle primordial dans la théorie de la rugosité d'Helmholtz, il perd de son importance dans la théorie de la fusion. En effet, Stumpf considère le timbre et le degré de fusion comme indépendants. Toutefois, bien qu'il considère que la fusion est un phénomène primaire de la conscience, Stumpf admet que des paramètres acoustiques et physiologiques peuvent favoriser l'écoute intégrale d'un agrégat de sons complexes. Parmi eux figurent l'harmonicité des partiels des sons complexes, la justesse de l'intonation des intervalles, l'identité spatiale de la source, la synchronicité des partiels, propriété de l'enveloppe spectrale, et des facteurs psychologiques tels l'attention ou l'entraînement. Toutefois, Stumpf refuse de considérer ces facteurs comme des causes de la fusion. La fusion est justement pour lui un fait psychologique, en tant que tel dernier et irréductible. De ce fait, il rejette toute explication psychologique, comme celles par la ressemblance, le rythme inconscient¹ ou l'habitude. Ainsi, la ressemblance de deux sons qui fusionnent est pour Stumpf une conséquence de leur fusion, et non l'inverse.

La non-dépendance de la fusion au timbre exclue la conception timbrale de la consonance (CDC 5). La fusion peut-être vue comme une reformulation moderne de la consonance diphonique (CDC 2).

2.3.3 Héritage d'Helmholtz et de Stumpf

L'héritage d'Helmholtz et de Stumpf est considérable et se perpétue dans la recherche moderne. En outre, le domaine de la perception de l'harmonie a connu un essor fulgurant à partir des années 1960, en incluant les branches des sciences cognitives et de la neurobiologie.

Le débat entre Stumpf et Helmholtz est virulent au tournant du siècle, alimenté, en plus de Stumpf, par des psychologues de renom, parmi lesquels Wundt ([Wundt, 1902](#)), un des fondateurs de la psychologie expérimentale, Krueger ([Krueger, 1903](#)), Meyer ([Meyer, 1901](#)), ou encore Faist ([Faist, 1897](#)), chacun apportant sa propre contribution à l'édifice théorique. Les synthèses

entendus aux deux oreilles, en opposition à l'*écoute diotique*.

1. Le rythme inconscient est cher à Lipps (voir page [90](#)).

de Malmberg ([Malmberg, 1918](#)), Moore ([Moore, 1914](#)), ou, dans une visée différente, Lalo ([Lalo, 1908](#)) constituent de bons indicateurs des idées débattues. Dans la théorie d’Helmholtz, on critique le recours à l’hypothèse physiologique, qui se révèle d’ailleurs fausse. Chez Stumpf, on critique le protocole expérimental, notamment le recours à des sujets non-musiciens. L’irréductibilité psychologique du phénomène de fusion est sérieusement remise en cause. Certains théoriciens comme Wundt ([Wundt, 1902](#)) proposent des approches multi-dimensionnelles qui concilient la rugosité d’Helmholtz avec la fusion de Stumpf. De nombreuses améliorations sont réalisées, notamment d’ordre méthodologique. C’est d’ailleurs l’empirisme de l’approche qui constitue la part la plus importante de l’héritage. Après Helmholtz et Stumpf, l’expérimentation physique et surtout psychologique (avec un grand nombre de sujet et le recours aux statistiques) devient la norme de tout apport dans le champ de la perception de la consonance et de la dissonance, plus généralement de la perception de l’harmonie.

Les arguments psychologiques issus d’études perceptives prennent le dessus sur les arguments physiologiques. Ainsi, l’apport majeur de Plomp et Levelt à la rugosité ([Plomp et Levelt, 1965](#)) se situe dans la notion de bande critique et dans la courbe élémentaire de dissonance qui moyenne les résultats expérimentaux obtenus sur un ensemble de sujets. Les arguments physiologiques ne sont plus invoqués. C’est le règne de la psychoacoustique. En cela, Stumpf gagne sur Helmholtz. Helmholtz, par contre, a le dessus sur Stumpf par la théorie de la rugosité. Celle-ci a connu des développements considérables dans la deuxième moitié du XX^e siècle ([Plomp et Levelt, 1965](#); [Kameoka et Kuriyagawa, 1969a,b](#); [Hutchinson et Knopoff, 1979](#)) et reste de nos jours invoquée comme cause principale, mais pas unique, de la dissonance. La théorie des sons résultants a été poussée plus loin par Krueger ([Krueger, 1903](#)), qui considère à la fois leurs battements et leurs coïncidence, pour ensuite tomber en désuétude, apparaissant comme un phénomène secondaire sur le plan perceptif¹.

La popularité de la théorie de la rugosité est due en grande partie à sa prédictibilité. En se ramenant aux interactions élémentaires entre partiels, elle fournit un algorithme pour calculer la rugosité de tout accord à partir de données spectrales ou sonores. À l’inverse, les lois de la fusion formu-

1. Citons tout de même Benade ([Benade, 1990](#)), qui reprend les mesures de battements entre sons différentiels et sons additionnels en lien avec la sensation de consonance et de fusion. À titre anecdotique, les sons résultants sont réinvestis dans une étude récente ([Chan et al., 2019](#)).

lées par Stumpf ne suffisent pas à rendre cette notion opératoire. En effet, elle constitue pour le psychologue un fait dernier et indivisible de la perception, irréductible à une description de faits acoustiques ou physiologiques. Malgré des tentatives relativement récentes de description psychoacoustique (DeWitt et Crowder, 1987; Schneider, 1996), de description cognitive (McAdams, 1982), ainsi que des modèles mathématiques de fusion fondés sur des mécanismes neuronaux de détection de périodicité (Ebeling, 2008), la fusion n'est plus que rarement invoquée comme cause de la consonance. L'harmonicité, qui est un des facteurs acoustiques majeurs influençant la fusion, est par contre souvent associée à la théorie de la rugosité dans l'explication du phénomène de consonance et de dissonance¹.

La théorie de la rugosité d'Helmholtz a fait basculer la consonance et la dissonance du côté de la conception timbrale (CDC 5), qui reste aujourd'hui largement majoritaire. Cette dépendance au timbre fournit notamment des moyens théoriques pour aborder les musiques micro-tonales ou étudier la musique produite par des instruments inharmoniques².

1. Trois études récentes, de Parncutt (Parncutt *et al.*, 2019), Harrison (Harrison et Pearce, 2020) et Friedmann (Friedman *et al.*, 2021a) s'accordent à ranger l'harmonicité parmi les trois causes principales de la consonance et de la dissonance, à côté de la rugosité et de la familiarité.

2. Le lien entre spectre et échelle musicale est présenté dans la partie 2.2.2.

Chapitre 3

Théories de la perception de l'harmonie : état de la recherche actuelle

En 1908, Lalo, dans l'introduction de son *Esquisse d'une esthétique musicale scientifique*, écrit :

Pour l'antiquité grecque et sans doute aussi pour les civilisations orientales antérieures, la science musicale fut essentiellement et presque uniquement arithmétique. C'est encore dans ce sens qu'elle entre dans le quadrivium du Moyen-Age.

Au XVII^e siècle seulement, avec Descartes et Mersenne, elle change de caractère ; et la génération suivante prend conscience, avec Sauveur, Rameau ou le P. André, que la science musicale est double : physique et mathématique.

Les savants croient alors pouvoir négliger le premier point de vue : l'élévation récente de la psycho-physiologie au rang des sciences positives a semblé permettre à Helmholtz de ne plus s'appuyer que sur elle.

Plus récemment, les contemporains dégagent et isolent la psychologie comme si elle se suffisait à elle-même ; pour Stumpf, Lipps ou H. Riemann, la théorie musicale est essentiellement psychologique et non physiologique.

Enfin quelques penseurs clairvoyants commencent, de nos jours seulement, à entrevoir un nouvel ordre de réalités par-delà tous

*les autres : la réalité sociale. Tard venue et plus réfléchie, on peut espérer que la sociologie ne poursuivra pas plus longtemps l'illusion permanente de ses devancières, qui est de se croire chacune à soi toute seule l'unique réalité scientifique*¹.

Lalo, tout en dénonçant l'erreur et la prétention d'une approche scientifique de la musique qui se restreindrait à une discipline en excluant toutes les autres, donne des directions à suivre pour les recherches futures. Il prône les approches intégrant une multiplicité de points de vue, d'« *ordres de réalité* », qui nécessitent en amont une synthèse historique².

3.1 Multitude des approches et des modèles

Que deviennent les propos de Lalo, un siècle plus tard, sur la question précise de la perception de l'harmonie ? Aux ordres de réalités déjà cités s'ajoute la réalité cognitive, qui pare les questions perceptives d'une forte dimension individuelle. Sous l'angle cognitif, les jugements de consonance et de dissonance sont l'effet de la familiarité et de l'apprentissage (Lundin, 1947; Cazden, 1980). De là à rejeter les explications physiologiques, il n'y a qu'un pas, que certains auteurs n'hésitent pas à franchir (McDermott *et al.*, 2016), aussitôt rappelés à l'ordre par d'autres (Bowling *et al.*, 2017). Toutefois, conscients de la stérilité d'une vision univoque, nombreux sont ceux qui défendent la complémentarité des différents « *ordres de réalité* », notamment l'inséparabilité des approches sensorielles et culturelles (Tramo *et al.*, 2001; Johnson-Laird *et al.*, 2012; Bowling *et al.*, 2017; Parncutt *et al.*, 2019; Harrison et Pearce, 2020; Eerola et Lahdelma, 2020; Friedman *et al.*, 2021a).

Les recherches actuelles sur la perception des accords traitent la perception dans une optique plus large que celle de la seule consonance et dissonance, qui reste néanmoins la question la plus débattue. On est loin de la démarche d'Helmholtz ou de Stumpf, qui élaborent une théorie de la sensation à partir d'un unique aspect de la perception : la rugosité pour le premier, la fusion pour le second. D'autres dimensions perceptives des accords isolés sont

1. Tiré de l'*Esquisse d'une esthétique musicale scientifique* (Lalo, 1908, p.9).

2. Lalo écrit « *Une esthétique qui veut se fonder sans tenir compte des lois de l'évolution est comme une histoire naturelle qui ignorerait l'anatomie comparée.* » (Lalo, 1908, p.7). Lalo cherche à faire de l'esthétique une science, mais on peut élargir son propos aux approches scientifiques de la musique.

explorées, parmi lesquelles la valence, la tension ou l'énergie (Lahdelma et Eerola, 2016), l'ambiguïté, l'obscurité (Cong, 2016). Une littérature conséquente traite des émotions transmises par les accords (Oelmann et Laeng, 2009; Cousineau *et al.*, 2012; Lahdelma, 2017), de même que l'attente créée par les accords (Huron, 2006). Une distinction est faite entre la consonance, associée à l'agrément, et la préférence, qui ne sont pas synonymes (Lahdelma et Eerola, 2020). En effet, des études ont montré qu'en isolation les accords légèrement dissonants sont préférés aux accords consonants, la courbe de préférence en fonction de la dissonance suivant une loi en forme de « U » inversé (Parncutt, 1989; Lahdelma et Eerola, 2016).

La perception des accords isolés, de l'ordre de la sensation, n'est pas la perception musicale ¹, et de nombreuses études ont montré l'importance de la prise en compte du contexte musical (Rogers, 2010; Arthurs, 2015; Johnson-Laird *et al.*, 2012). Les modèles de perception en contexte prennent tantôt appui sur les éléments d'un langage donné, qui délimite par là leur champs d'application (Johnson-Laird *et al.*, 2012), tantôt intègrent la mémoire à court terme (Rogers, 2010; Jensen et Hjortkjær, 2012).

La transition entre accords, isolés ou en contexte musical, constitue encore une autre dimension de la perception, traduite par l'affinité (Milne *et al.*, 2016) ou la tension de transition (Chan *et al.*, 2019).

Face aux nombreux aspects de la perception de l'harmonie, plusieurs modèles théoriques sont développés pour en rendre compte. La plupart sont des modèles psychoacoustiques, c'est-à-dire issus d'une confrontation des propriétés acoustiques avec les résultats d'expérience perceptives.

Une grande partie des modèles ont un fondement neurobiologique, ils modélisent alors les mécanismes neuronaux qui codent des aspects précis de la perception. Ces modèles poussent l'approche physiologique plus loin que ne pouvait le faire Helmholtz ², et se situent dans sa continuité. Ainsi, on sait que la dissonance et la consonance ont une influence directe sur les potentiels d'action des nerfs auditifs (Tramo *et al.*, 2001; Bidelman et Heinz, 2011), et sont encodés dans le tronc cérébral (Bidelman et Krishnan, 2009; Bones

1. Lalo écrit à ce propos : « *De par sa passivité, la sensation brute, enregistrement pur et simple de l'impression psycho-physiologique, est impuissante à donner leur sens musical aux sons. Ils ne l'acquiescent que lorsqu'ils sont interprétés comme une partie d'un ensemble, réel ou imaginé. C'est cette synthèse active que nous entendons par le mot de perception.* », tiré de (Lalo, 1908, p.163).

2. Et par là confirment l'exaspération de Lalo. Se reporter à la citation page 113.

et al., 2014). De plus, la consonance et la dissonance n'activent pas les mêmes régions du cerveau (Fishman *et al.*, 2001; Minati *et al.*, 2009).

Les modèles dits "pitch-based" (Terhardt, 1984; Parncutt, 1989; Hofmann-Engl, 2008; Cong, 2016) diffèrent des modèles physiologiques ou acoustiques en partant des fréquences fondamentales perçues par l'oreille, et non directement d'un spectre. Cette approche permet de contourner l'aspect physiologique.

Citons quelques uns des principaux modèles psychoacoustiques de la littérature contemporaine :

- Rugosité et battements (Plomp et Levelt, 1965; Kameoka et Kuriyagawa, 1969a,b; Hutchinson et Knopoff, 1979; Pressnitzer et McAdams, 1999; Sethares, 1999; Vassilakis, 2001)
- Harmonicité : approche spectrale (McDermott *et al.*, 2010; Cousineau *et al.*, 2012; Milne *et al.*, 2016; Bowling *et al.*, 2018; Carcagno *et al.*, 2019; Friedman *et al.*, 2021a)
- Périodicité temporelle et autocorrelation (Patterson, 1986; Ebeling, 2008; Shapira Lots et Stone, 2008; Stolzenburg, 2015; Trulla *et al.*, 2018; Chan *et al.*, 2019).
- Entropie, approche probabiliste et théorie de l'information (Temperley *et al.*, 2007; Cox, 2010; Strohmaier, 2014; Smit *et al.*, 2019)
- Tension, instabilité (Pressnitzer *et al.*, 2000; Cook, 2009; Hjortkjær, 2011; Farbood, 2012; Balsach, 2016; Herremans et Chew, 2016; Chan *et al.*, 2019)

À ces modèles traitant la perception comme innée s'ajoutent ceux qui la considèrent comme acquise, fruit de l'acculturation. La familiarité influence la mémoire à long terme, et la consonance ainsi que la dissonance seraient déterminées par la confrontation d'un stimulus à cette mémoire, via le processus de reconnaissance. C'est le cas de l'incongruence cognitive (McLachlan *et al.*, 2013). D'autres approches tiennent compte de l'effet de la familiarité à un langage donné (Johnson-Laird *et al.*, 2012) ou à une échelle donnée (Smit *et al.*, 2019), et des théories hybrides modélisent l'effet de la mémoire à court terme sur les modèles psychoacoustiques (Jensen et Hjortkjær, 2012).

Face à la multiplicité des aspects perceptifs de l'harmonie, à la multiplicité des modèles et des théories, et à la confusion qui en résulte, une réelle volonté de synthèse se fait jour ces dernières années, rendant le sujet de la perception

de l'harmonie très actuel.

3.2 Volonté de synthèse

Un grand nombre de publications récentes sur la perception de l'harmonie sont caractérisées par un état de l'art extrêmement fourni. Que les états de l'art portent sur les dimensions perceptives et les émotions véhiculées par les accords isolés (Lahdelma, 2017), sur les différents aspects du débat nature vs. culture (Rogers, 2010; Arthurs, 2015), sur l'influence du contexte musical dans les jugements perceptifs (Arthurs et Timmers, 2016), sur les approches biologiques de la consonance (Bowling *et al.*, 2017), sur les mécanismes neurologiques à l'œuvre (Carcagno *et al.*, 2019), sur les modèles de périodicité (Stolzenburg, 2015), d'entropie (Strohman, 2014), de tension (Hjortkjær, 2011), et surtout sur les théories de la consonance et de la dissonance (Parncutt *et al.*, 2019; Harrison et Pearce, 2020; Friedman *et al.*, 2021a), cette tendance est révélatrice d'une volonté non cachée d'apporter de l'ordre dans la profusion des modèles et des discours.

Cette volonté se traduit par une remise en cause, une rectification ou une amélioration constantes des modèles (Mashinter, 2006; Eerola et Lahdelma, 2020; Friedman *et al.*, 2021a), sources de vifs débats. Ainsi, la remise en question par McDermott de l'existence de causes naturelles à la consonance (McDermott *et al.*, 2016), via des expériences perceptives sur un peuple amazonien, donne lieu à une publication de Bowling (Bowling *et al.*, 2017) critiquant la méthode et les conclusions du premier, et dressant en même temps un inventaire des approches biologiques. Plus récemment, l'article de Friedman (Friedman *et al.*, 2021a) sur les corrélats de la consonance et de la dissonance dans une échelle micro-tonale entraîne plusieurs réactions sous forme d'articles-commentaires dans le même journal (Harrison, 2021; Smit et Milne, 2021; Bowling, 2021; Friedman *et al.*, 2021b). En outre, c'est le seul article du journal à avoir été commenté, ce qui montre l'actualité brûlante de la question de la dissonance et de la consonance.

La volonté de synthèse se traduit également par une approche comparative des modèles, en particulier sur cette même question de la consonance et de la dissonance (Stolzenburg, 2015; Parncutt *et al.*, 2019; Harrison et Pearce, 2020; Eerola et Lahdelma, 2020). Parmi tous les modèles, plusieurs sont redondants et tous ne sont pas pertinents. Il s'agit, dans ces études, de corréler un corpus de modèles avec des expériences perceptives pour trouver des indicateurs prédisant la consonance et la dissonance. La publication de Harrison (Harrison et Pearce, 2020) est impressionnante d'exhaustivité :

après un état de l’art très fourni, l’auteur compare une vingtaine de modèles avec les résultats de quatre études perceptives et un corpus musical de plus de 100 000 œuvres, pour conclure que la concordance et la dissonance statiques dépendent essentiellement de la rugosité, de l’harmonicité et de la familiarité. D’autres études arrivent au même consensus (Parncutt *et al.*, 2019; Smit *et al.*, 2020; Eerola et Lahdelma, 2020).

On aboutit de ce fait à une modélisation multi-dimensionnelle de la consonance et de la dissonance, vues comme des phénomènes complexes, composites (Smit et Milne, 2021), faisant intervenir des mécanismes de natures différentes. En prônant l’indissociabilité des aspects naturels et culturels, cette approche met à mal certains fantasmes, comme l’existence d’une perception « pure » des accords isolés, innée et indépendante de l’exposition ou de l’apprentissage. Eerola (Eerola et Lahdelma, 2020), après avoir corrélié un grand nombre de modèles et procédé à une analyse par composantes principales, conclut que les facteurs culturels expliqueraient 46% des variations des jugements de consonance et de dissonance des accords isolés, contre 19% pour la rugosité et l’harmonicité.

Cette approche multi-dimensionnelle concerne aussi la diversité des aspects de la perception. Des études corrélient les modèles disponibles à plusieurs dimensions perceptives, afin non seulement de comprendre les mécanismes fondamentaux, mais aussi d’éclaircir les concepts guidant les jugements perceptifs et les liens qu’ils entretiennent entre eux (Smit *et al.*, 2019; Lahdelma et Eerola, 2020). Lahdelma (Lahdelma et Eerola, 2020) tire d’un corpus de 60 expériences perceptives sept concepts associés à la consonance et à la dissonance : la consonance, la douceur, la pureté, la dimension harmonieuse, la tension, l’agrément et la préférence. Il montre que les corrélations entre ces concepts dépendent fortement de la familiarité aux stimuli ainsi que de la pratique musicale, suggérant par-là l’utilisation de définitions différentes de la consonance et de la dissonance selon le type d’expérience. De plus, la familiarité et la pratique musicale agissent différemment sur les concepts : ils influencent beaucoup l’agrément, mais très peu la tension.

La double notion de consonance et de dissonance est au cœur des recherches actuelles sur la perception de l’harmonie. Et pour cause, l’origine de la consonance est considérée comme « *l’une des questions les plus fondamentales de la perception musicale*¹ ». Les recherches vont toujours plus loin dans

1. Tiré de (Smit et Milne, 2021), voir note 1 page 32.

la dissection du phénomène, pour en comprendre les mécanismes et dresser son « *anatomie*¹ » En outre, une volonté se dégage de résoudre de manière définitive la question, comme si suffisamment d'encre avait déjà coulé. Suite à un récent travail de synthèse, des corpus entiers de modèles perceptifs² sont mis en ligne pour accélérer les recherches, ainsi que des bases de données audio d'accords³ et des corpus de résultats perceptifs⁴. Les chercheurs restent toutefois modestes, conscients de la grande complexité du phénomène : « *Il est encore tôt pour dresser l'anatomie définitive de la consonance et de la dissonance*⁵. »

Approches comparatives de la consonance et de la dissonance De même qu'« *une esthétique qui veut se fonder sans tenir compte des lois de l'évolution est comme une histoire naturelle qui ignorerait l'anatomie comparée*⁶ », la compréhension du phénomène de consonance et de dissonance doit tenir compte de toutes les formes qu'il peut prendre, à travers l'histoire, les cultures, les âges, les espèces animales, les handicaps.

En corrélant différents modèles perceptifs avec la fréquence d'occurrence des triades dans un corpus vocal allant du XIII^e au XIX^e siècle, Parncutt (Parncutt *et al.*, 2019), dans une démarche qu'il qualifie de « *psychohistorique* », retrace l'histoire des modèles qui prédisent à une époque donnée la prévalence des accords, et par-là la conception de la consonance et de la dissonance. Selon l'auteur, jusqu'au XVI^e siècle ces modèles se réfèrent essen-

1. Expression tirée du titre "The Anatomy of Consonance/Dissonance" de l'article (Eerola et Lahdelma, 2020).

2. Citons ici trois corpus de modèles d'analyse audio, incluant des modèles perceptifs :
 — La librairie « incon », pack open-source développé sur R (Harrison et Pearce, 2020), disponible à l'adresse <https://github.com/pmcharrison/incon>.
 — La librairie « Essentia » (Bogdanov *et al.*, 2013), disponible à l'adresse <https://essentia.upf.edu>.
 — « MIRtoolbox » (Lartillot *et al.*, 2008), disponible à l'adresse <https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox>.

3. Citons la base d'accords « *Durham Chord Dataset* » (Eerola et Lahdelma, 2020) comprenant plus de 4755 accords de 2 à 6 notes, générés avec des sons de piano, disponible à l'adresse <https://github.com/tuomaseerola/DCD>.

4. Les résultats des quatre expériences perceptives utilisés par Harrison (Harrison et Pearce, 2020) sont disponibles à l'adresse <https://github.com/pmcharrison/inconData>.

5. "It is early to decisively conclude the exact anatomy of Consonance/Dissonance." Tiré de (Eerola et Lahdelma, 2020, p.42)

6. Citation due à Lalo, voir note 2 page 122.

tiellement à la sonorité des accords (rugosité et harmonicité), et ce n'est qu'à la Renaissance que la familiarité devient un facteur important. L'approche de Parncutt peut être vue comme le pendant explicatif du travail sémantique de Tenney¹ (Tenney, 1988).

De nombreux travaux visent à limiter l'influence de la familiarité, de l'apprentissage et des facteurs culturels dans les jugements perceptifs, de manière à isoler les aspects acoustiques et biologiques. C'est le cas des études menées sur des peuples peu familiers avec la musique tonale (McDermott *et al.*, 2016), sur des enfants (Masataka, 2006; Plantinga et Trehub, 2014) ou sur des bébés animaux (Sugimoto *et al.*, 2010; Chiandetti et Vallortigara, 2011) ou des animaux plus âgés (Toro et Crespo-Bojorque, 2017). Elles posent la question du caractère inné ou acquis de la préférence pour la consonance. Une autre manière de limiter les effets de la familiarité consiste à utiliser des échelles micro-tonales non-usuelles (Smit *et al.*, 2019, 2020; Friedman *et al.*, 2021a) et des spectres inharmoniques (Milne *et al.*, 2016).

Des accidents ou des handicaps peuvent avoir un effet sur les mécanismes de la perception de la consonance et de la dissonance. Afin de mieux cerner les rôles respectifs de la rugosité et de l'harmonicité, des tests ont été effectués sur des sujets souffrant d'amusie congénitale, maladie neurogénétique caractérisée par une déficience de la perception des hauteurs (Cousineau *et al.*, 2012; Gosselin *et al.*, 2015; Marin *et al.*, 2015). Les préférences de ces sujets sont corrélées à la rugosité, mais non à l'harmonicité, contrairement aux sujets sains.

Enfin, les notions de consonance et de dissonance sont confrontées à des traditions musicales diverses, par exemple le jazz (McGowan, 2008), des musiques privilégiant les intervalles dissonants (Vassilakis, 2005; Ambrazevičius, 2017) ou des musiques extra-européennes utilisant des instruments inharmoniques² (Sethares, 1999).

De la sensation à la valeur affective Les recherches précédemment citées visent à comprendre les mécanismes qui sous-tendent la sensation de consonance et de dissonance. Or cette compréhension, aussi complète et prédictive qu'elle puisse l'être, n'explique pas la valeur affective associée à cette sensation, qui peut être exprimée pour la consonance en termes d'agrément,

1. Voir page 73.

2. Voir partie 2.2.2.

de préférence, voir de beauté¹, et pour la dissonance en termes antonymes. La phrase de Lalo « *Où cessera enfin la vibration, pour que la conscience commence*² ? » est toujours d’actualité. Helmholtz formule une hypothèse, qui n’est plus tenable telle quelle aujourd’hui : celle que les intermittences créées par les battements sur les terminaisons des nerfs auditifs causent une fatigue des nerfs, et sont donc nuisibles à l’organisme³.

Plusieurs hypothèses évolutives sont avancées (Harrison, 2021), qui restent spéculatives, et qui se réfèrent aux explications naturelles :

- Similarité vocale : la préférence pour les sons consonants proviendraient de leur similarité acoustique (en matière d’harmonicité) avec les vocalisations humaines (Bowling et Purves, 2015; Bowling *et al.*, 2018; Friedman *et al.*, 2021a). Ainsi, l’homme aurait évolué pour reconnaître ce type de sons, porteurs d’informations sur les congénères.
- Analyse de scènes auditives : les sons harmoniques et avec peu de battements nécessiteraient moins de traitement cognitif que les sons rugueux et inharmoniques (Huron, 2001).
- Facteurs sociaux : l’homme aurait évolué en préférant les sons à haute harmonicité et à faible interférence car cette préférence encouragerait la création musicale collaborative, qui aurait à son tour des fonctions adaptatives favorisant le lien social (Savage *et al.*, 2020).

À côté de ces explications évolutives figure l’explication culturelle, qui consiste à dire que la consonance et la dissonance sont seulement conditionnées par la familiarité et l’exposition. De ce fait, l’homme n’aurait pas évolué pour apprécier les sons à forte harmonicité et à faible interférence, l’agrément qu’il en tire viendrait uniquement d’habitudes et de conventions culturelles (McDermott *et al.*, 2016).

1. Attention toutefois à ne pas confondre ces termes, qui ne sont pas strictement équivalents entre eux. Voir page 123 pour une distinction entre préférence et agrément.

2. Voir la citation page 113.

3. Voir page 109. Helmholtz ne parle pas explicitement de nuisance, mais celle-ci est sous-entendue dans l’hypothèse physiologique.

Deuxième partie

Les descripteurs harmoniques : introduction et implémentation

Chapitre 4

Des descripteurs audio aux descripteurs harmoniques

4.1 Les descripteurs audio

L'appellation « descripteurs audio » regroupe un ensemble de méthodes qui mettent en évidence des aspects spécifiques du signal audio, en proposant des représentations numériques ou visuelles. Ils sont largement utilisés dans le domaine de la recherche d'information musicale (M.I.R), où de nombreux articles passent en revue les descripteurs et testent leur efficacité : (McKinney et Breebaart, 2003; McKay *et al.*, 2005; Mitrović *et al.*, 2010; Pachet et Roy, 2009; Peeters, 2004). Ils sont également utilisés dans d'autres domaines, liés ou non à la musique : analyse musicale, classification musicale, reconnaissance de style, acoustique des instruments de musique, analyse de la parole...

De nombreux logiciels et bibliothèques permettent l'extraction de descripteurs audio à partir du signal. Citons *MIR Toolbox* et *Timbre ToolBox* (Peeters *et al.*, 2011), écrits en Matlab, *Ircamdescriptor* développé par l'IRCAM, les *plugins vamp* utilisés sur Sonic Visualiser et Audacity, les bibliothèques *Librosa* et *TorchAudio* pour python, *Zsa descriptors* pour du temps réel sur MaxMSP, *Essentia*, *Marsyas*, *Yaafe*.

4.1.1 Différentes catégories de descripteurs audio

Les descripteurs audio, nous l'avons vu, regroupent une multitude de méthodes, qui se différencient par leur nature, leur niveau d'abstraction,

leur complexité, l'échelle temporelle de ce qu'ils décrivent, le type de représentations auxquels ils aboutissent, leur lien avec la perception et leur domaine d'application. Les descripteurs audio peuvent être séparés en plusieurs grandes catégories.

Descripteurs spectraux/descripteurs d'enveloppe La première étape dans le calcul des descripteurs à partir du signal consiste à obtenir une représentation numérique du signal qui soit exploitable. On distingue essentiellement deux types de représentations en entrée, les représentations de la forme d'onde (ou de l'enveloppe temporelle), et les représentations spectro-temporelles. Ces dernières mettent en jeu des transformées comme la transformée de Fourier à court terme (TFCT, ou STFT — *short term Fourier transform* en anglais), ou les transformées à facteur Q constant (TQC, ou CQT — *constant-Q transform* en anglais). Ces représentations sont dépendantes des paramètres de l'analyse tels la taille et la forme de la fenêtre glissante, le recouvrement des fenêtres, la précision spectrale pour la TQC...

Dans (Peeters *et al.*, 2011), les auteurs incluent un autre type de représentation en entrée : les représentations issues de modèles auditifs, modélisant la perception humaine. Celles-ci tiennent compte de la notion de bande critique, ensemble de filtres à bande passante dont la largeur de bande dépend de la position sur l'axe des fréquences.

Certains descripteurs peuvent être calculés à partir de différentes représentations en entrée, comme le barycentre spectral (qui peut être extrait d'un spectrogramme obtenu par TFCT ou d'une détection des partiels harmoniques), tandis que d'autres ne peuvent être extraits que d'un seul type de représentation (comme les coefficients *attack*, *decay* et *release*, calculés sur l'enveloppe temporelle).

Voici une liste, non exhaustive, des descripteurs issus de la représentation de la forme d'onde, et de ceux issus des représentations spectrales. Comme le développement des descripteurs est relativement récent, leur dénomination est généralement en anglais. Pour plus de commodité, nous laissons les noms des descripteurs en anglais.

- Descripteurs issus de la représentation de l'enveloppe temporelle : *RMS-energy envelope*, descripteurs issus du modèle *attack-decay-release*, *temporal centroid*, *frequency and amplitude of energy modulation*.
- Descripteurs issus des représentations spectrales : *spectral centroid*,

spectral spread, spectral skewness, harmonicity, harmonic spectral deviation, fundamental frequency, chroma, cepstre.

Descripteurs locaux / globaux Chaque descripteur est caractérisé par une échelle temporelle, liée à l’aspect du signal qu’il décrit. Il peut être plus ou moins localisé en temps. Des descripteurs sont calculés sur chaque fenêtre de l’analyse (typiquement de l’ordre de quelques dizaines de millisecondes), par exemple le barycentre spectral ou l’étalement spectral, tandis que d’autres sont calculés sur une échelle de quelques secondes, voire sur le signal entier. C’est le cas des descripteurs qui estiment le tempo ou la tonalité.

Contrairement aux descripteurs globaux qui sont généralement présentés tels quels, les descripteurs locaux peuvent nécessiter après extraction un traitement statistique (corrélation, écart-type, clustering) pour donner une information globale sur le signal (matrices d’auto-similarité). L’information locale peut-être exploitée telle quelle pour des analyses dynamiques.

Descripteurs de bas niveau/moyen niveau/haut niveau Nous reprenons ici la distinction opérée par Serra ([Serra *et al.*, 2013](#), p.12).

- **Descripteurs de bas niveau** : ils sont calculés par des formules mathématiques simples. Cette catégorie contient à la fois des descripteurs spectraux (barycentre spectral, étalement spectral) mais aussi des descripteurs partant de la forme d’onde (énergie, intensité). Certains descripteurs globaux sont de bas niveau, comme le *temporal centroid*.
- **Descripteurs de moyen niveau** : ils sont calculés par des algorithmes plus complexes, souvent paramétrables. C’est le cas des algorithmes de détection des *onsets* ou de la pulsation. Ce sont souvent des estimateurs, pour lesquels le choix des paramètres joue un rôle déterminant. Ces descripteurs n’ont pas la robustesse des descripteurs de bas niveau.
- **Descripteurs de haut niveau** : ils permettent de caractériser des données perceptives ou sémantiques, comme la tonalité, le rythme, le genre ou l’humeur. Ils font appels à des modèles complexes, qui peuvent inclure des modèles statistiques (chaînes de Markov cachées,

modèles bayésiens), du *clustering* ou des modèles d'apprentissage.

La figure 4.1, issue de (Peeters *et al.*, 2011), liste les descripteurs audio implémentés dans *Timbre Toolbox*, en les classant suivant leur échelle temporelle et le type de représentation en entrée.

	Audio descriptor	Units	Abbreviation	Input representation
Global descriptors	Attack	s	Att	Temporal Energy Envelope
	Decay	s	Dec	
	Release	s	Rel	
	Log-Attack Time	log(s)	LAT	
	Attack Slope	a/s	AttSlope	
	Decrease Slope	log(a)/s	DecSlope	
	Temporal Centroid	s	TempCent	
	Effective Duration	s	EffDur	
	Frequency of Energy Modulation	Hz	FreqMod	
	Amplitude of Energy Modulation	a	AmpMod	
Time-varying descriptors	Autocorrelation (12 coefficients)	-	AutoCorr	Audio Signal
	Zero Crossing Rate	s ⁻¹	ZcrRate	
	RMS-Energy Envelope	a	RMSEnv	Temporal Energy Envelope
	Spectral Centroid	F	SpecCent	STFTmagnitude (STFTmag) STFTpower (STFTpow) ERBfft (ERBfft) ERBgammatone (ERBgam) Harmonic
	Spectral Spread	F	SpecSpread	
	Spectral Skewness	-	SpecSkew	
	Spectral Kurtosis	-	SpecKurt	
	Spectral Slope	F ⁻¹	SpecSlope	
	Spectral Decrease	-	SpecDecr	
	Spectral Rolloff	F	SpecRollOff	
	Spectro-temporal variation	-	SpecVar	
	Frame Energy	I	FrameErg	
	Spectral Flatness	-	SpecFlat	STFTmag, STFTpow, ERBfft, ERBgam
	Spectral Crest	-	SpecCrest	
	Harmonic Energy	a ²	HarmErg	Harmonic
	Noise Energy	a ²	NoiseErg	
	Noisiness	-	Noisiness	
	Fundamental Frequency	Hz	F0	
	Inharmonicity	-	InHarm	
	Tristimulus (3 coefficients)	-	TriStim	
	Harmonic Spectral Deviation	a	HarmDev	
	Odd to even harmonic ratio	-	OddEveRatio	

FIGURE 4.1 – Table des descripteurs audio, issue de (Peeters *et al.*, 2011). La dernière colonne représente les représentations en entrée. ERB est l'acronyme de Equivalent Rectangular Bandwidth, représentation issue d'un modèle auditif

4.1.2 Usage en musicologie numérique

L’usage des descripteurs audio en musicologie a connu un essor depuis les années 2000, s’inscrivant dans le développement de la musicologie numérique ou musicologie computationnelle. La musicologie numérique est un champ pluridisciplinaire incluant l’utilisation de la technologie et des outils audio-numériques. Elle est intimement liée aux musiques électroacoustiques et aux corpus numériques, qui obligent les chercheurs et les musiciens à se tourner vers de nouveaux outils et de nouvelles approches incluant l’acoustique, le traitement du signal et l’informatique. Ces nouvelles approches ne sont pas réductibles à un transfert des méthodes traditionnelles, mais constituent une réelle rupture épistémologique en musicologie (Couprie, 2020). Notamment, les outils numériques ouvrent de nouveaux champs de recherche (Cook, 2005, p.5), intimement liés à la création et à l’interprétation.

Les descripteurs audio dans l’analyse musicale Dans le contexte de la musicologie numérique, l’analyse musicale se scinde en deux sous-disciplines, qui diffèrent par la nature de leur objet. D’un côté, un pan de l’analyse traite les données symboliques, encodées dans des formats comme le MIDI, MusicXML ou MEI. L’automatisation des analyses par les logiciels dédiés permet le traitement de vastes corpus. De l’autre, l’analyse prend en entrée des données audio, et implique l’utilisation de descripteurs audio. La figure 4.2, tirée de (Couprie, 2020), liste les principaux logiciels utilisés en analyse musicale, avec à gauche ceux partant de la notation symbolique, à droite du signal. La multiplicité des logiciels proposant des représentations à partir du signal audio est révélatrice de l’importance des descripteurs audio dans l’analyse musicale.

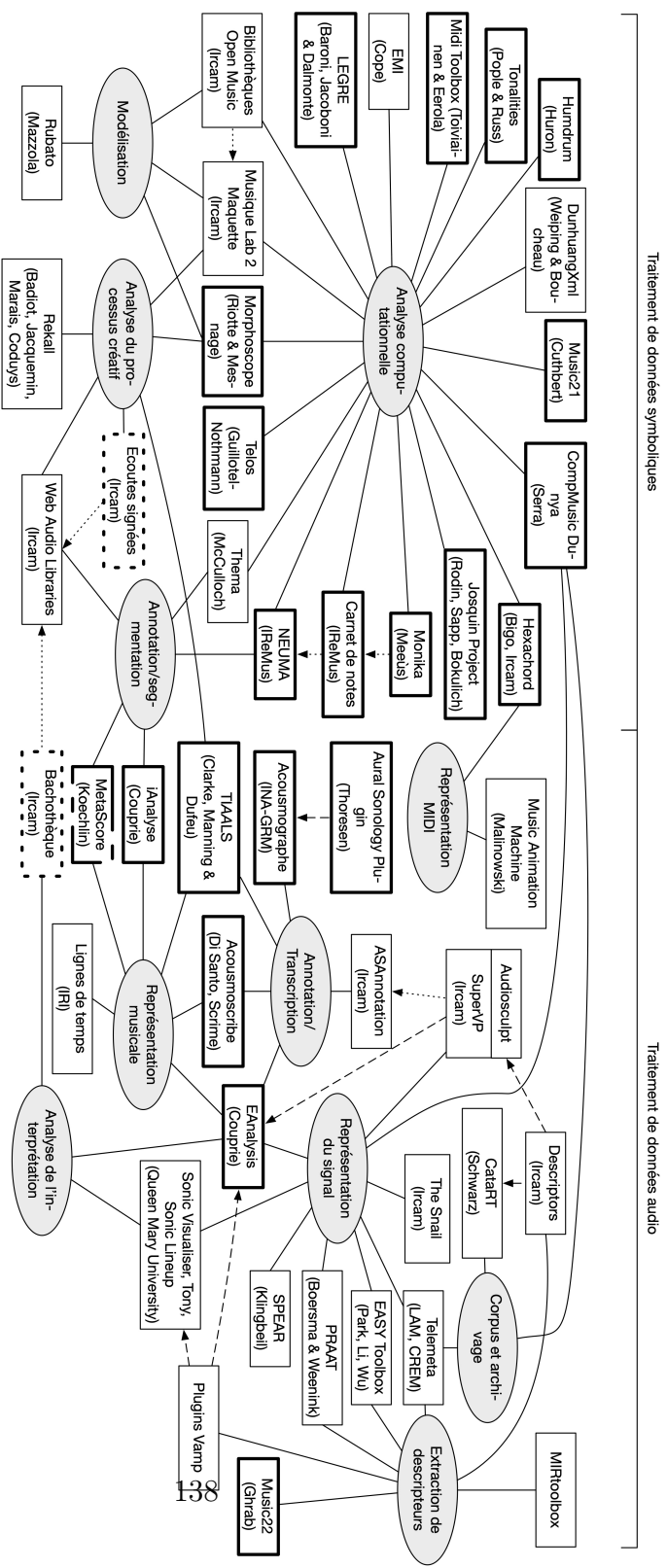


FIGURE 4.2 – Les principaux logiciels utilisés en analyse musicale, extrait de (Coupré, 2020). À gauche, les logiciels traitant des données symboliques et à droite ceux partant du signal. Les logiciels développés spécifiquement pour la musicologie apparaissent entourés en gras, les autres proviennent de l'acoustique, de l'analyse et du traitement de la parole ou de la création musicale.

Couprie ([Couprie, 2020](#)) distingue quatre étapes dans toute démarche d'analyse musicale à partir de fichiers audio :

1. Extraction d'informations du fichier audio. Cette étape passe par la représentation de la forme d'onde ou d'un spectrogramme, et va généralement jusqu'à l'extraction de descripteurs audio.
2. Calcul de coefficients à partir des descripteurs : méthodes statistiques comme l'autocorrélation des valeurs prises par un descripteur, leur écart-type, leur dispersion. Cette étape optionnelle est utilisée principalement avec des descripteurs de bas niveau.
3. Visualisation des informations extraites. Cette visualisation directe peut être difficilement exploitable par le musicologue, soit que ce dernier n'ait pas les clés pour la lire, soit que les catégories mises en évidence soient éloignées des catégories musicales. Ainsi, une autre représentation, plus exploitable, peut-être nécessaire.
4. L'interprétation des résultats, l'analyse et la création de représentations musicales. Cette étape est purement musicologique. Elle peut prendre la forme d'annotations d'une partition, de schémas, de diagrammes statistiques...

La distinction entre descripteur audio et visualisation issue de ce descripteur n'est pas claire, car les descripteurs peuvent fournir automatiquement des représentations visuelles (par exemple sous forme de matrices d'auto-similarité).

Couprie incite à prendre conscience de la rupture épistémologique entre l'objet d'étude, le signal musical, et les représentations de ce signal servant de point de départ à l'analyse. Les représentations sont de nature acoustique, et non à proprement parler musicales. Elles sont parfois éloignées des intuitions perceptives.

L'audio est très difficile à étudier sans un support visuel, c'est particulièrement le cas avec la création contemporaine ou avec des enregistrements de longue durée. Il faut donc passer par des représentations, toujours imparfaites, qui seront à la base du travail d'analyse. Or, la lecture même du fichier audionumérique et ses représentations ne sont que des interprétations, des médiations technologiques pour lesquelles il est impossible de vérifier la fidélité [...]. De plus, les représentations ne sont généralement

*pas musicales : la forme d'onde n'est qu'une vague représentation de l'intensité sonore, le sonagramme n'est qu'une estimation approximative du contenu spectral ou le descripteur de barycentre spectral n'est qu'une représentation de la « brillance », particulièrement difficile à traduire musicalement. Avec l'étude de fichiers audionumériques, le musicologue navigue donc dans des données très approximatives pour lesquelles il va devoir construire les outils et les méthodes permettant de révéler leurs qualités musicales.*¹

1. ([Couprie, 2020](#))

4.2 Des descripteurs audio aux descripteurs harmoniques

4.2.1 Les descripteurs audio traitant de l'harmonie

Dans l'introduction de ce travail nous avons défini l'harmonie comme l'ensemble des phénomènes sonores et des relations formelles impliquant l'interaction simultanée ou successive de plusieurs notes. L'existence de descripteurs audio extrayant des informations de nature harmonique est loin d'être évidente, car elle se heurte à deux difficultés, la première liée à l'opposition entre discret et continu, la seconde à une confusion possible des systèmes de discrétisation.

Opposition continu/discret Les descripteurs audio prennent en entrée une seule piste audio. Dès lors, ils ne discriminent pas a priori le continuum sonore en unités différenciées d'un point de vue temporel (notes successives) ou spatial (notes simultanées). Cela implique que les descripteurs ne séparent pas a priori les propriétés relevant des notes isolées des propriétés relevant spécifiquement de l'interaction entre ces notes.

Confusion possible des systèmes de discrétisation La représentation sur laquelle sont calculés les descripteurs est discrétisée temporellement et spectralement. Tout d'abord, la piste audio, numérisée, est discrétisée temporellement avec un pas temporel égal à $1/f_r$, où f_r est la fréquence d'échantillonnage. Ensuite, la représentation spectro-temporelle implique une discrétisation spectrale (qui dépend du type de transformation ainsi que de ses paramètres) et crée une seconde discrétisation temporelle¹. Cette discrétisation dans l'espace temps-fréquence, liée à l'analyse du signal, est artificielle en ce qu'elle n'est de nature ni perceptive ni musicale. Comme l'écrit Guillo-
tel : « *Le signal audio [...] ne comporte que des données quantitatives dont le séquençage s'effectue a priori indépendamment de considérations qualitatives.* »²

Dans le cas d'une musique notée, la segmentation temporelle en verticalités n'a rien à voir avec la segmentation temporelle de l'analyse automatique.

1. La discrétisation temporelle dans l'analyse dépend notamment de la taille des fenêtres d'analyse et du taux de recouvrement.

2. (Guillotel-Nothmann, 2020), voir page 48 pour la citation entière.

La première est variable au sein d’une œuvre, une verticalité pouvant durer de quelques centièmes de secondes à quelques secondes, tandis que la seconde est constante, avec une fenêtre typiquement de l’ordre du centième de seconde. On peut dire la même chose de la segmentation fréquentielle. Dans le cas d’une musique à échelle, les hauteurs sont fixées, et ne peuvent prendre qu’un nombre fini de valeurs.

Or, avec une TFCT, la segmentation de l’axe fréquentiel ne correspond aucunement à une échelle musicale. Avec des transformations plus fines il est possible de faire correspondre les deux, notamment avec celles qui permettent un contrôle du facteur de qualité Q .

Cette coexistence de deux systèmes de discrétisation, l’un devant rendre compte de l’autre, peut a priori être source de confusion, et l’on peut légitimement se demander si les descripteurs audio sont à même de saisir des propriétés spécifiquement harmoniques. Il convient cependant de garder à l’esprit que les pas de quantification fréquentielle de l’analyse sont généralement d’un ordre de grandeur plus petit que les intervalles présents dans les échelles musicales.

Ces difficultés tiennent à la différence de nature entre la partition, qui manipule des objets symboliques, et le signal physique. L’harmonie, telle qu’elle a été définie, fait intervenir à la fois l’acoustique et le symbolique, puisqu’il s’agit de considérer les interactions entre les différentes notes.

Si elles méritent d’être posées, ces difficultés ne constituent pas pour autant un frein qui rendrait impossible l’existence de descripteurs audio à même de toucher le phénomène harmonique. Citons trois raisons qui vont dans ce sens.

1. **Passage du physique au symbolique.** Un pan très important dans le domaine du MIR est l’analyse de contenu audio (ACA)¹, qui consiste à extraire du signal audio des concepts musicaux, c’est-à-dire essentiellement des concepts liés à la musique écrite dans la tradition européenne : note, mélodie, accord, tempo, rythme, structure... Dès lors que l’on sait extraire du signal les notes fondamentales et les contributions spectrales de chaque note, on est capable d’isoler les notes qui interagissent. Nous reviendrons sur les méthodes issues du ACA qui concernent l’harmonie.

1. Voir (Serra *et al.*, 2013, p.25) pour une présentation générale des enjeux de l’ACA

2. **Le passage par le symbolique n'est pas forcément nécessaire.**

Il est possible de décrire les effets d'une interaction sans pour autant savoir isoler les unités qui interagissent. Lorsque deux sons sinusoïdaux en phase et de fréquences proches f_0 et $f_0 + \delta$ sont joués simultanément, se produisent des battements de fréquence δ . Ces battements sont détectables dans la forme d'onde sans avoir à isoler les deux partiels.

3. **L'interaction peut se situer à une autre échelle.** Les phénomènes issus de l'interaction de plusieurs notes ne se situent pas forcément aux mêmes échelles temporelles et spectrales que ces notes (battements rapides).

De plus, il est possible de contourner la difficulté de la différenciation des sons issus de différentes notes si l'on a des descripteurs audio prenant en entrée non pas uniquement une piste audio, mais aussi les pistes audio de chaque voix, dans l'idée qu'une voix ne contient pas plus d'une note simultanément.

Examinons les descripteurs de la littérature se référant explicitement à une musique polyphonique. La représentation des chromas, la reconnaissance d'accords ou encore la transcription de notes en font partie.

Pour illustrer les descripteurs présentés nous prenons l'exemple suivant de cinq accords de trois notes :

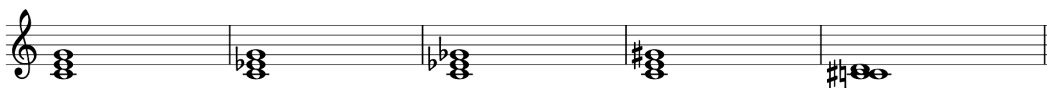


FIGURE 4.3 – Suite d'accords utilisés comme exemples dans la suite : accord majeur, accord mineur, accord diminué, accord augmenté, et agrégat de deux demi-tons

Pour générer l'audio nous utilisons le logiciel Musescore associé à librairie de sons d'orgue échantillonnés *Orgue de salon*¹. L'exemple sonore est généré avec le jeu *Bourdon + Cheminée + Flûte*₄.

1. Disponible à l'adresse <http://sonimusicae.free.fr/orguedesalon-en.html>

Nous utilisons un plugin Muscore nous permettant de régler l'intonation de chaque note¹. Nous utilisons pour cet exemple des intervalles justes par rapport à *do*. Ci-dessous sont résumés les rapports de fréquence utilisés.

	<i>do</i> - <i>do</i> ♯	<i>do</i> - <i>ré</i>	<i>do</i> - <i>mi</i> b	<i>do</i> - <i>mi</i>	<i>do</i> - <i>sol</i> b	<i>do</i> - <i>sol</i>	<i>do</i> - <i>sol</i> ♯
Rapports des fréquences	$\frac{5}{3} \times \frac{5}{4} \times \frac{1}{2}$	$\left(\frac{3}{2}\right)^2 \times \frac{1}{2}$	$\frac{6}{5}$	$\frac{5}{4}$	$\frac{7}{5}$	$\frac{3}{2}$	$\left(\frac{5}{4}\right)^2$
Déviations (en cents)	-30	4	-16	-14	-18	-2	-28

TABLE 4.1 – Rapports de fréquences et déviations en cents par rapport aux intervalles également tempérés des intervalles utilisés dans l'exemple.

Chromas Nous reprenons ici le modèle de Matthias Mauch dans (Mauch et Dixon, 2010), utilisé pour la transcription de notes et la reconnaissance d'accords. Ces algorithmes sont disponibles en un *plug-in Vamp*² pour Sonic Visualiser ou Audacity.

Les chromas représentent les contributions spectrales des fréquences fondamentales dans chaque demi-ton de la gamme chromatique, à l'octave près.

La figure 4.4 montre la représentation en chromas sur la suite d'accords (considérée). On a bien le résultat qui était attendu, mis à part pour le dernier accord, où le *do*♯ n'est pas bien repéré.

1. Voir <https://github.com/billhails/MuseScore-plugins/tree/master/src/Tuning>

2. Disponible à l'adresse <http://isophonics.org/nls-chroma>

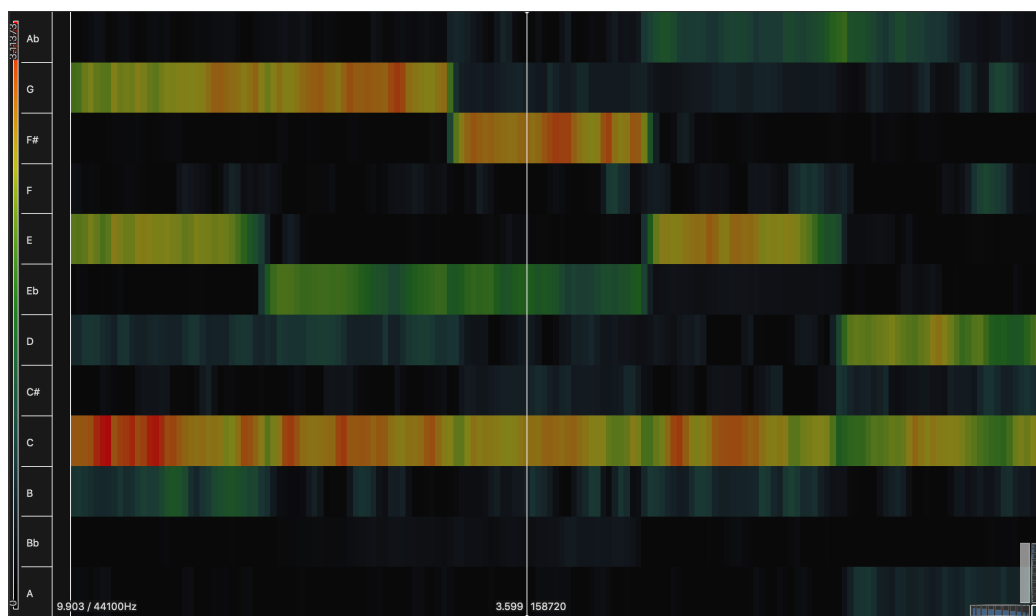


FIGURE 4.4 – Analyse en chromas de la suite d’accords (fig. 4.3) sur Sonic Visualiser, avec l’algorithme « Chromagram » de Matthias Mauch. Les paramètres de l’analyse sont les paramètres par défaut.

Transcription de notes La représentation précédente en chromas estime la hauteur des notes fondamentales, mais ne renseigne pas sur l’octave de ces notes. D’autres algorithmes (Klapuri, 2003; Yeh *et al.*, 2009) estiment les fréquences fondamentales des notes présentes simultanément dans l’audio, sans hypothèse sur le nombre de notes. Cette tâche, appelée *multipitchtracking* est plus hardue que la détection d’une unique fréquence fondamentale, car elle implique d’estimer le nombre de notes et de séparer les contributions spectrales via la considération de l’harmonicité, de l’enveloppe spectrale et de la synchronisation de l’évolution de l’enveloppe temporelle des partiels pour chaque source potentielle.

Changement harmonique Il détecte le changement spectral d’un accord à l’autre. Certains algorithmes sont uniquement des algorithmes de traitement du signal (Mauch, 2010), tandis que d’autres incluent des considérations géométriques et formelles (Harte *et al.*, 2006).

Reconnaissance d'accords Les algorithmes de reconnaissance d'accords (Harte, 2010; Mauch, 2010; Wu *et al.*, 2019) estiment les types d'accords. Ils passent généralement par le calcul des chromas. Ils peuvent être associés à un indice de confiance de l'estimation.

La figure 4.4 montre l'analyse de la suite d'accords (fig. 4.3) sur Sonic Visualizer avec les descripteurs audio de *Chordino* : estimation du changement harmonique, estimation du type d'accord, et indice de confiance dans l'estimation. L'estimation des accords est correcte, sauf pour le dernier, qui n'est pas un accord classé. L'algorithme donne un accord de trois sons qui contient deux des trois hauteurs réellement présentes.

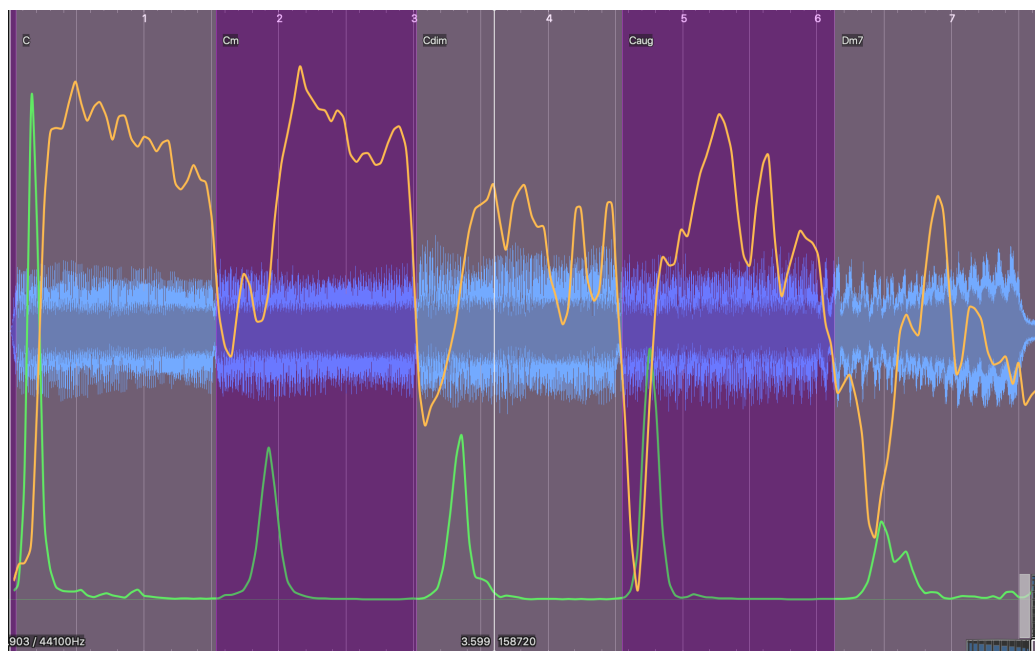


FIGURE 4.5 – Analyse de la suite d'accords (fig 4.3) sur Sonic Visualiser, avec le plugin *Chordino* de Matthias Mauch. En vert, une estimation du changement harmonique. Bandes verticales : estimation des types d'accords. Orange : indice de confiance dans l'estimation du type d'accords. Les paramètres des analyses sont les paramètres par défaut.

4.2.2 Descripteurs audio et modèles de perception

Dans l'introduction de ce travail nous avons discuté des modèles de perception de l'harmonie et de leur usage en musicologie, relativement restreint. Il convient de préciser le lien entre modèles de perception et descripteurs audio.

Une partie des modèles de perception sont implémentés à partir du signal, et peuvent ainsi être considérés comme des descripteurs audio. Notamment, le domaine des MIR fait un usage important des modèles perceptifs. On peut citer les MFCC (Mel-frequency cepstral coefficients), coefficients cepstraux mettant en jeu l'échelle Mel, échelle de hauteur des sons basée sur l'audition humaine (Stevens *et al.*, 1937).

La différence entre les modèles de perception et les descripteurs audio tient essentiellement au domaine d'application : les premiers sont utilisés en perception, les seconds principalement en traitement du signal et dans le domaine du MIR. Les descripteurs audio regroupent un ensemble de méthodes variées, qui diffèrent grandement par leurs objectifs, leurs hypothèses, les représentations auxquelles ils aboutissent et leur niveau de sophistication. Les modèles perceptifs sont plus unifiés quant à leur objectif initial : il s'agit de construire des modèles qui rendent compte de la perception. Le rôle des expériences perceptives est double : elles permettent à la fois de bâtir les modèles et de vérifier leur validité. Les modèles de perception, comme les descripteurs audio, peuvent être hautement sophistiqués - modèles de *virtual pitch* (Terhardt, 1984), modèles de *pitch commonality* (Thompson et Parncutt, 1997).

On peut noter une différence dans la manière d'utiliser les deux types de modèles, due essentiellement à la divergence des objectifs. Les descripteurs audio sont souvent regroupés entre eux pour former des vecteurs harmoniques, qui peuvent servir à des tâches comme la segmentation automatique (Tahon *et al.*, 2019). Les modèles de perception, quant à eux, sont souvent utilisés isolément, comme c'est le cas des courbes de dissonance utilisées pour analyser le signal musical¹. L'approche multi-dimensionnelle, nous l'avons vu, est toutefois de plus en plus souvent adoptée dans la littérature², avec idée qu'un seul modèle ne suffit pas pour modéliser correctement la perception. Cook (Cook, 2009), pour s'approcher de la perception des triades, combine des modèles de dissonance, de tension et de modalité (appartenance majeur-

1. voir *Introduction* page 39

2. Voir partie 3.2.

mineur), tandis que Cong (Cong, 2016) modélise la perception de l’harmonie en combinant quatre grandeurs indépendantes : la dissonance sensorielle, l’ambiguïté, l’obscurité et la tension³. Harrison (Harrison et Pearce, 2020) inclut la familiarité dans un modèle de perception de la consonance simultanée, en la combinant avec deux modèles acoustiques : la rugosité et l’harmonicité/périodicité. Dans les exemples cités, il ne s’agit pas uniquement de modéliser la perception, mais de comprendre les mécanismes à l’œuvre. En outre, la combinaison des modèles est fine en ce qu’elle découle d’une réflexion approfondie à partir des expériences perceptives, et que les auteurs définissent les rôles relatifs des modèles utilisés.

Une autre différence entre descripteurs audio et modèles de perception est dans l’objet de départ. Les premiers, comme leur nom l’indique, prennent en entrée des fichiers audio, tandis que les seconds prennent soit un fichier audio, soit une partition¹. Cette distinction n’est en fait pas fondamentale, étant donné que la plupart des approches symboliques prennent en entrée des paramètres de spectre pour pouvoir associer à chaque note une série de partiels. Dans ce cas, l’implémentation peut être adaptée pour prendre en entrée directement le signal audio.

Les descripteurs audio et les modèles de perception de l’harmonie sont des notions qui se recoupent en partie, et leurs différences, bien qu’elles existent, sont avant tout superficielles, car liées à l’usage. Certains modèles appartiennent d’ailleurs aux deux catégories, comme la rugosité et l’inharmonicité - implémentées dans la MIRtoolbox, ou encore les *mel-frequency cepstrum coefficients*.

4.2.3 La notion de descripteur harmonique

Nous introduisons ici la notion de descripteur harmonique, comme type particulier de modèle acoustique décrivant des phénomènes harmoniques, c’est-à-dire l’ensemble des phénomènes sonores et des relations formelles impliquant l’interaction simultanée ou successive de plusieurs notes.

Cette notion recoupe en partie celle de descripteur audio, en partie celle de modèle de perception de l’harmonie, mais n’est réductible ni à l’une ni à

3. voir citation page 50.

1. Se référer à l’article "*Simultaneous consonance in music perception and composition*" (Harrison et Pearce, 2020) pour une classification des modèles de perception de l’harmonie selon le type d’objets en entrée.

l'autre, pas plus qu'à leur intersection. Le schéma 4.6 représente d'un point de vue ensembliste les rapports entre les trois notions.

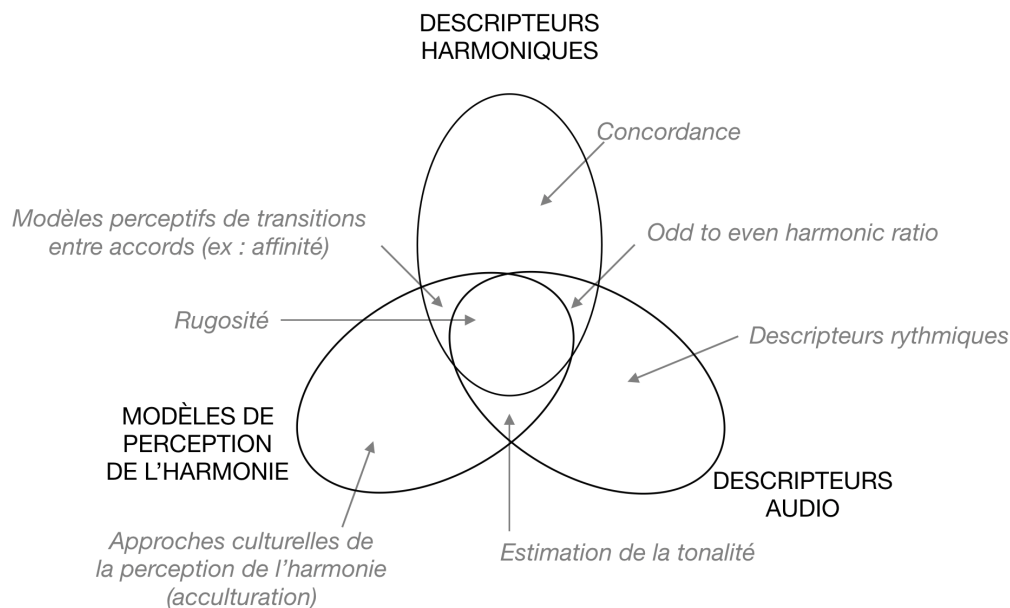


FIGURE 4.6 – Schéma des relations d'appartenance entre les descripteurs harmoniques, les descripteurs audio et les modèles de perception de l'harmonie. Les modèles notés en gris constituent des exemples de chaque région.

La rugosité appartient aux trois catégories, de même que l'harmonicité. Parmi les modèles de perception de l'harmonie, les approches culturelles qui tiennent compte de l'apprentissage ou de l'exposition à un système musical ne sont pas de nature uniquement acoustique, ils ne rentrent donc pas dans la catégorie des descripteurs harmoniques, pas plus que dans celle des descripteurs audio puisqu'ils ne peuvent pas en général être calculés à partir du signal seul, mais nécessitent de prendre en compte des paramètres sur l'auditeur. Les descripteurs audio donnant une estimation de la tonalité sont des modèles de perception de haut niveau, qui incorporent un modèle de connaissance, la théorie tonale. De fait, ils ne sont pas de nature purement acoustique, ce qui incite à les mettre en-dehors de la catégorie des descripteurs harmoniques. Les modèles perceptifs de transition entre accords sont bien des descripteurs harmoniques, qui concernent l'aspect dynamique de l'harmonie, mais ils ne sont pas considérés comme des descripteurs audio, étant donné qu'ils sont la

plupart du temps calculés à partir de la partition ([Terhardt, 1984](#)). Le rapport entre harmoniques impairs et harmoniques pairs (*odd to even harmonic ratio*) est un descripteur audio utilisé essentiellement pour décrire le timbre, il constitue donc un modèle de perception du timbre, et non à proprement parler un modèle de perception de l'harmonie. On peut cependant l'inclure dans la catégorie des descripteurs harmoniques. Les descripteurs rythmiques sont des descripteurs audio qui ne concernent pas l'harmonie au sens où nous l'avons définie. Enfin, la concordance, qui quantifie les fréquences communes aux spectres de plusieurs notes dans un accord, est issue du formalisme vectoriel des signaux¹. Le terme « concordance » a été introduit par Chouvel ([Chouvel, 1998](#)), qui s'en sert comme outil pour guider le compositeur dans la musique micro-tonale. Même si son lien avec la perception n'est pas exclu, celui-ci n'a jamais été approfondi ; elle n'est donc pas un modèle perceptif à proprement parler. Étant donné que son calcul nécessite les spectres des différentes notes de l'accord, elle ne peut être calculée sur le signal brut, et n'est donc pas considérée au sens strict comme un descripteur audio.

1. La concordance sera définie dans la partie [6.1.2](#).

4.3 Usage des descripteurs harmoniques

Nous l'avons vu, ni les modèles de perception, ni les descripteurs audio ne peuvent suffire pour répondre à la problématique de ce travail. Étant donné que la notion de phénomène harmonique regroupe à la fois des phénomènes sonores et des relations formelles entre éléments symboliques, les modèles acoustiques, pour rendre compte du phénomène harmonique, doivent intégrer les deux aspects. Or, les descripteurs audio, pour pouvoir incorporer les relations formelles, nécessitent de passer par des estimateurs (estimateur des notes, séparation de sources...), et nous privilégions dans ce travail les approches déterministes. Les modèles de perception, s'ils peuvent traiter à la fois les aspects symboliques et sonores, montrent quant à eux des limites dans l'usage fait en musicologie.¹

La notion de descripteur harmonique a été introduite dans l'optique de rendre compte acoustiquement du phénomène harmonique. Elle est donc adaptée pour répondre à la problématique initiale, et nous développons désormais cette notion.

Les descripteurs harmoniques utilisés dans la suite du travail, ainsi que leur implémentation, seront détaillés dans les chapitres 5 et 6. Nous exposons dans cette partie différents points caractéristiques des descripteurs harmoniques.

4.3.1 Origine des descripteurs

Les modèles de perception de l'harmonie constituent une source importante de descripteurs harmoniques. Parmi eux s'en trouvent qui modélisent directement le fonctionnement d'un ou des organes du système auditif, comme le modèle de rugosité d'Helmholtz ([Helmholtz, 1863](#), chap.8), dans lequel les battements entre partiels proches provoquent des déformations de la membrane basilaire.

D'autres modèles de perception trouvent leur origine dans les travaux des psychologues, et donnent des explications d'ordre psychologique, comme la notion de fusion introduite par Carl Stumpf, qui se réfère à la tendance de deux sons simultanés à être perçus comme un seul son. Cette notion présente de nombreuses analogies avec les principes de la théorie de la Gestalt,

1. Voir partie [1.1.3](#).

sur laquelle il a eu une grande influence¹. La fusion est intimement liée à l’harmonicité, un de ses corrélats acoustiques.

Enfin, des modèles de perception sont construits pour reconnaître certaines catégories d’accords, et sont directement inspirés des structures de ces accords. C’est le cas de la tension et de la modalité de Cook (Cook, 2009), qui caractérise, dans une triade, l’écart entre l’intervalle supérieur et l’intervalle inférieur. La tension met en valeur les accords diminués et augmentés, tandis que la modalité distingue les accords majeurs des accords mineurs.

Parmi les descripteurs utilisés dans ce travail, plusieurs adaptent à l’aspect dynamique des modèles de perception statique. C’est le cas de la concordance différentielle ou de la rugosité différentielle.

Une partie des descripteurs harmoniques utilisés n’a pas une origine perceptive. C’est le cas de la concordance, naturellement issue du formalisme vectoriel des signaux, car s’identifiant dans ce cadre avec le produit scalaire canonique entre signaux, calculé dans le domaine temporel ou dans le domaine fréquentiel. Bonnet (Bonnet, 1968) l’interprète comme une énergie d’interaction. Dans ce travail nous développons plusieurs descripteurs dérivés de la concordance.

4.3.2 À l’interface entre le symbolique et le signal

Dans notre approche, l’harmonie, nous l’avons vu, désigne à la fois des relations entre entités formelles et des phénomènes sonores. Les descripteurs harmoniques, puisque qu’ils décrivent l’harmonie, doivent être capables de traiter ces deux types d’objets. Ils doivent notamment :

- Être sensibles aux aspects sonores comme le timbre et les nuances, et rendre compte de l’influence de ces paramètres sur l’harmonie.
- Prendre en entrée autant un signal audio que des objets abstraits comme une partition ou des classes d’accords.

Ces conditions sont a priori difficilement conciliables, étant donné qu’un objet abstrait comme une classe d’accords peut donner lieu à une grande diversité de réalisations sonores. Celles-ci se distinguent par leur timbre, leur

1. Voir *Carl Stumpf and the training of scientists in Berlin* (Ash, 1998, chap.2) consacré à l’influence de Stumpf sur la théorie de la Gestalt

nuance, leur durée, mais également le nombre de notes (il peut y avoir des doublures), le registre, la disposition des notes.

Nous surmontons le problème en adoptant une approche hybride, à l'interface entre le signal et le symbolique, qui inclut les deux aspects sans jamais se réduire à l'un ni à l'autre.

Traces sonores dans l'approche symbolique Les descripteurs harmoniques utilisés dans la suite sont des descripteurs spectraux, c'est-à-dire calculés sur des représentations spectrales.

Lorsque l'objet d'entrée est un signal audio, ces représentations sont obtenues par analyse spectrale¹.

Lorsque l'objet en entrée est une partition, l'analyse spectrale n'est pas envisageable. Nous contournons la difficulté en associant un spectre à chaque note via un modèle de spectre². Cette association a cela d'arbitraire qu'elle ne traduit pas une réalité sonore, mais seulement un choix de paramètres d'un modèle de spectre. Elle doit être vue comme une modélisation d'une réalisation sonore parmi d'autres de la partition ou de l'objet symbolique, mais ne prétend en aucun cas être représentative de l'ensemble des réalisations sonores. L'intérêt de passer par un modèle de spectre réside dans sa maniabilité, et la variation de paramètres permet d'isoler l'effet de certains facteurs.

Lorsque l'objet symbolique en entrée est une classe d'accords, son univers sonore dépend, en plus du timbre, du renversement, du registre, et plus généralement de la disposition des notes. Une manière de contourner la dépendance à la disposition consiste à utiliser un modèle de spectre sous forme de spectre de Shepard généralisé³, dans lequel chaque partiel est démultiplié sur tous les intervalles de duplication de fréquence. Dans le cas de sons harmoniques, qui sera le parti pris de ce travail, cet intervalle devient l'octave. Ce type de modèle permet de se départir de la dépendance des descripteurs harmoniques à la disposition, et ainsi de les calculer, via le choix des paramètres de spectre, sur les classes d'accords.

1. Nous utilisons une transformée à facteur Q constant (TQC), se référer à la section 5.3.

2. Voir partie 5.2.1.

3. Voir partie 5.2.2.

Traces de la partition dans l’approche signal L’approche symbolique à partir d’une partition n’est pas totalement symbolique puisqu’elle contient une trace sonore sous forme de spectres modélisés. De la même manière, l’approche à partir du signal audio n’est pas une approche signal au sens strict du terme, puisqu’elle inclut des traces de la partition. Ces traces sont à la fois temporelles et spectrales.

Traces temporelles Pour que les descripteurs harmoniques soient adaptés à l’usage musicologique, il faut tout d’abord que les représentations soient lisibles. À partir de la figure 1.2, nous avons énuméré plusieurs raisons pouvant rendre la lisibilité des représentations temporelles difficiles, en particulier lorsqu’elles sont obtenues à partir du signal. Parmi elles figure la difficulté d’identifier l’emplacement des notes et de se repérer par rapport à la partition, difficulté due aux différences d’échelles temporelles entre les fenêtres d’analyse et les verticalités¹ de la partition. Pour le musicologue, qui a l’habitude de manipuler des notes, ce repérage est pourtant crucial.

Afin de régler ce problème, nous calculons les descripteurs harmoniques non pas sur des fenêtres d’analyse constantes, mais sur chaque verticalité. La méthode, détaillée dans la partie 5.3, consiste à moyenner les spectres entre deux verticalités. Elle implique de connaître les emplacements temporels des verticalités. Ceux-ci peuvent être estimés automatiquement ([Hainsworth *et al.*, 2003](#); [Parascandolo *et al.*, 2016](#)), être repérés manuellement ou être pré-enregistrés.

Traces spectrales L’harmonie implique l’interaction simultanée ou successive de plusieurs notes. Lorsque les interactions sont successives, la connaissance des verticalités permet d’isoler les termes de cette interaction, mais cela n’est pas le cas lorsqu’elles sont simultanées. Dans ce dernier cas, il faut être capable d’isoler chaque note, c’est-à-dire connaître à la fois la fréquence fondamentale et le spectre de chacune d’elles.

Certains descripteurs utilisés, comme la concordance, nécessitent en effet d’isoler les termes de cette interaction simultanée. Là encore, les notes peuvent être estimées automatiquement² via des algorithmes de séparation de sources, ou bien être pré-rentées. Dans la suite du travail, pour calculer ces descripteurs par l’approche signal, nous exigeons en entrée non seulement

1. Une verticalité est l’intervalle temporel séparant deux notes successives.

2. Voir p.145

le signal global, mais également le signal des voix séparées, où chaque voix contient au maximum une note à un instant donné. Cette hypothèse forte restreint nécessairement le corpus d'application.

Continuation et nouveautés dans l'approche hybride Les approches hybrides entre partition et signal, loin d'être une nouveauté, sont déjà présentes dans les premiers modèles psychoacoustiques de la dissonance. Ceux-ci prennent en entrée une partition, et les calculs sont réalisés avec des spectres modélisés, le plus souvent avec la série des partiels associée à leurs amplitudes respectives. C'est le cas du modèle de dissonance d'Helmholtz ([Helmholtz, 1863](#)), dans lequel la sensation de dissonance est expliquée par les battements entre partiels proches. Ainsi, les courbes historiques de la figure 1.3 sont obtenues en faisant varier le nombre de partiels dans le modèle de spectre. Malgré le développement du traitement du signal dans les années 1960 et 1970, le calcul de modèles psychoacoustiques à partir de la partition via des modèles de spectre en pics reste très fréquente ([Parncutt, 1989](#)).

L'approche consistant à partir du signal et à combiner une analyse temps-fréquence à des informations d'ordre symbolique est une contribution nouvelle des descripteurs harmoniques implémentés dans ce travail. Si les traces temporelles de la partition, c'est-à-dire l'emplacement des verticalités, peuvent être inférées à partir du signal avec des algorithmes performants sur certains types de sons (avec des attaques marquées) et certains répertoires, la séparation des spectres des notes est plus ardue, et des descripteurs comme la concordance totale, très sensibles aux variations des spectres des notes, supportent mal des spectres faussés.

Contraintes sur le répertoire Cette hybridation dans l'approche entre données symboliques et données sonores, ou modélisation de données sonores, impose des contraintes sur le répertoire et sur la méthode d'analyse.

Contraintes temporelles Le fait que les descripteurs harmoniques, dans ce travail, soient calculés sur les verticalités impose d'avoir une musique séparable en verticalités, c'est-à-dire une musique de notes, que ces notes soient écrites ou non sur une partition. Cela met de côté toute une partie de la musique qui manipule plus spécifiquement des textures ou des continuums sonores, notamment dans la musique électroacoustique.

De plus, associer un spectre à une note n'est acceptable musicalement que du moment où le spectre de cette note présente une faible variation temporelle. Cette contrainte a deux aspects, puisqu'un changement du spectre d'une note peut être dû à une variation de l'enveloppe temporelle, se traduisant par une variation de l'amplitude des partiels, simultanée ou non, ou à un déplacement des partiels, comme c'est le cas dans un *vibrato* ou un *glissando*. Le son de gong, dont l'amplitude relative des partiels évolue temporellement, ne convient à ce type d'analyse.

La première contrainte met de côté les sons présentant une trop forte décroissance spectrale, comme les sons percussifs ou les sons de piano, de même que les sons dont l'amplitude est modulée, comme les *tremolo*. Aussi, dans le modèle classique d'enveloppe temporelle « Attaque Chute Entretien Extinction », seule la phase d'entretien doit-elle être prise en compte dans la construction du spectre synchronisé.

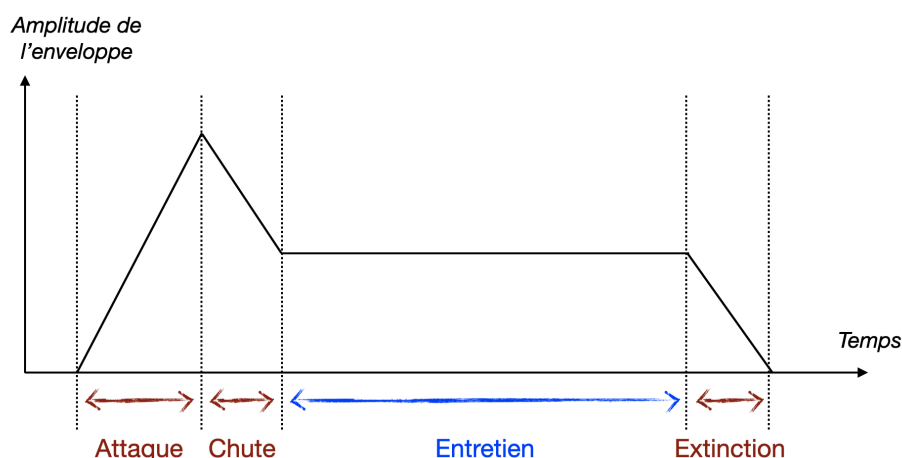


FIGURE 4.7 – Modèle « Attaque Chute Entretien Extinction ». Seule la phase constante d'entretien doit être considérée dans la construction du spectre synchronisé pour le calcul des descripteurs harmoniques.

L'usage des descripteurs harmoniques sur tout le cycle « Attaque Chute Entretien Extinction » ou sur des sons du type « Attaque Extinction » comme les sons de piano n'est pas pour autant exclu, mais l'interprétation des résultats devra alors être faite avec beaucoup de précautions, puisque les spectres synchronisés sur les notes seront moins représentatifs de la réalité sonore que pour des sons stables.

Les effets de *tremolo* et de *vibrato* peuvent être liés à une nécessité de saillance contextuelle, c'est-à-dire une nécessité de ressortir dans la perception pour attirer l'attention. Dans le cas de la voix, le soliste use plus volontairement de ces techniques que les chanteurs d'un chœur, étant donné qu'un chœur vise essentiellement l'homogénéité¹. Pour éviter les problèmes que peuvent poser ces effets dans la synchronisation des spectres sur les verticalités, nous privilégions des sons relativement stables et des répertoires dont l'esthétique relève de l'homogénéité plus que de la saillance.

Dans la suite de la thèse, nous nous focalisons particulièrement sur les sons d'orgue. Ceux-ci, en plus de couvrir un large panel de timbres, par l'intermédiaire des jeux et de leurs combinaisons, ont l'avantage d'être relativement stables temporellement. En effet, ils ne présentent qu'une faible attaque, et l'entretien de la soufflerie assure la stabilité de l'enveloppe temporelle.

Contraintes spectrales La nécessité d'avoir les spectres de chaque note d'un accord pose de fortes contraintes dans l'approche signal : en plus du signal global, certains descripteurs exigent d'avoir les voix séparées. Dans le cas d'un enregistrement, cela nécessite d'enregistrer les voix l'une après l'autre s'il s'agit d'un instrument polyphonique comme l'orgue, ou bien, si les instruments sont monodiques (orchestre à cordes sans utilisation des doubles cordes par exemple), d'enregistrer chaque instrument avec des micros de proximité. Des bases de données d'enregistrements d'œuvres musicales avec voix séparées sont disponibles sur internet².

Dans une partie de la thèse nous utilisons des sons d'orgue échantillonnés, provenant de la base *Orgue de salon*³. Cette banque de sons reproduit note par note et registre par registre un orgue de salon construit en 1988 comprenant 250 tuyaux, 5 registres et deux claviers. Des traitements numériques ont été réalisés par les auteurs pour éliminer les bruits mécaniques, notamment la soufflerie. Les sons ont été enregistrés dans une acoustique sèche avec peu de réverbération.

1. Voir (Ternström, 2003).

2. Pour l'orchestre, voir par exemple (Miron *et al.*, 2016), disponible à l'adresse <https://www.upf.edu/web/mtg/phenicx-anechoic>, base de données qui contient les enregistrements anéchoïques synchronisés de chaque instrument de l'orchestre dans des extraits d'œuvres de Mozart, Beethoven, Mahler et Bruckner.

3. Disponible à l'adresse <http://sonimusicae.free.fr/orguedesalon-en.html>

4.3.3 Approches statiques et dynamiques

Les descripteurs harmoniques peuvent se scinder en deux catégories, les descripteurs statiques, décrivant l'aspect statique de l'harmonie, et les descripteurs dynamiques, décrivant les transitions entre accords ou entre verticalités. La figure 4.8 différencie, sur un enchaînement de trois accords, les accords eux-même des deux transitions.

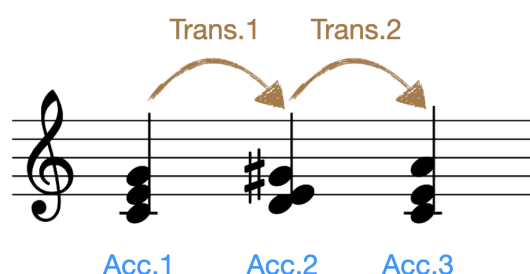


FIGURE 4.8 – Schéma des approches statiques et dynamiques, indiquant trois accords et les deux transitions correspondantes.

Sur cet exemple, un descripteur statique prend trois valeurs, une sur chaque accord, tandis qu'un descripteur dynamique prend deux valeurs, sur les deux transitions. On peut s'intéresser à l'évolution du descripteur statique en considérant la différence des valeurs de proche en proche. L'évolution d'un descripteur statique est donc un descripteur dynamique, qui prend des valeurs négatives s'il y a une perte, et positives dans le cas d'un gain.

Pour tout descripteur, statique comme dynamique, on peut considérer ses valeurs sur toutes les verticalités ou toutes les transitions, mais aussi son évolution, qui est obtenue en prenant la différence des valeurs de proche en proche. Ainsi, l'évolution d'un descripteur statique décrit la transition entre deux verticalités, tandis que l'évolution d'un descripteur dynamique décrit l'enchaînement entre deux transitions. Nous appelons « descripteurs dynamiques triviaux » les descripteurs dynamiques correspondant à l'évolution d'un descripteur statique.

Nous qualifions de « symétrie temporelle » la propriété que peuvent avoir les descripteurs dynamiques d'être invariables par inversion de l'ordre des accords sur les transitions,.

4.3.4 Prise en compte du contexte harmonique

Deux raisons nous incitent à développer des descripteurs harmoniques qui tiennent compte du contexte musical, c'est-à-dire des accords qui précèdent.

La première raison est purement acoustique et perceptive. Dans des salles avec beaucoup de réverbération, un son continue à être entendu après que l'instrumentiste a arrêté de le jouer, et lors d'une transition entre accords, le premier accord est encore perçu lorsque le second se fait entendre, ce qui crée un chevauchement perceptif des deux accords. Plus la salle est réverbérante, plus ce chevauchement est important. Ce phénomène est un effet du contexte.

La seconde raison est cognitive. On sait l'importance du contexte dans la perception des accords¹, que l'on soit en contexte musical ou hors contexte musical, et notamment l'importance de la mémoire à court terme, indépendamment de l'acoustique ou du mode d'écoute. Contrairement à la mémoire à long terme qui, en convoquant l'apprentissage et la familiarité, fait basculer immédiatement du côté des approches dites *culturelles*, la mémoire à court terme peut être modélisée par des considérations acoustiques ([Jensen et Hjortkjær, 2012](#)).

Les raisons acoustiques et cognitives nous poussent à proposer un modèle paramétrable qui intègre à un instant donné les accords précédemment joués, de manière à pouvoir modéliser autant les acoustiques sèches que les acoustiques réverbérantes. Le modèle consiste à associer à une verticalité donnée une moyenne spectrale pondérée de cette verticalité avec les $n - 1$ verticalités précédentes, où n est un paramètre du modèle². On obtient ainsi le spectre du contexte harmonique, qui vient remplacer dans le calcul des descripteurs harmoniques le spectre de la verticalité donnée.

Du point de vue cognitif, cette façon de procéder est très rudimentaire, et mériterait d'être confrontée à des expériences perceptives. Elle a néanmoins pour elle l'avantage de la simplicité. Conscients de ses limites, nous l'appliquons dans la suite du travail seulement au niveau abstrait des transitions entre classes d'accords³.

La modélisation du contexte peut servir à caractériser les aspects à la fois statiques et dynamiques de l'harmonie. Les descripteurs de contexte sta-

1. Voir page [123](#).

2. Voir section [5.1.3](#) pour les détails de l'implémentation.

3. Dans la partie [9.3](#).

tiques, sur une verticalité donnée, sont calculés à partir du spectre de contexte de cette verticalité, tandis que les descripteurs de contexte dynamiques décrivent la transition d'un contexte harmonique vers un nouvel accord. Ces relations sont résumées sur le schéma ci-dessous.

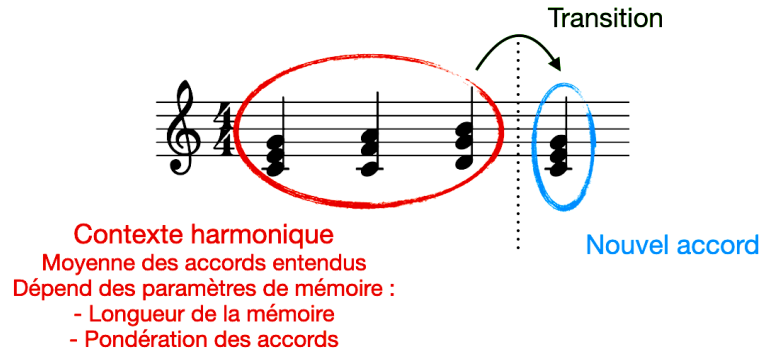


FIGURE 4.9 – Mode opératoire des descripteurs de contexte statiques et dynamiques, sur un exemple d’enchaînement d’accords.

Sur le schéma, la longueur de la mémoire est de trois accords. Ainsi, le contexte de l’accord 3 est obtenu par une moyenne spectrale pondérée des trois premiers accords, entourés en rouge, et c’est à partir de ce contexte que sont calculés les descripteurs de contexte statiques sur cet accord. Sur la transition entre les accords 3 et 4, les descripteurs de contexte sont calculés entre le contexte de l’accord 3 et l’accord 4.

Cette modélisation généralise l’approche hors contexte, puisqu’avec une mémoire de longueur 1 on retrouve les accords isolés et l’approche hors contexte.

Dans la suite du travail, seuls les descripteurs de contexte dynamiques seront étudiés.

Concernant l’aspect transitionnel, il peut être tentant d’étudier la transition entre deux contextes, plutôt que la transition entre un contexte et un accord. Elle ne sera cependant pas traitée dans le cadre de ce travail.

Chapitre 5

Méthode et implémentation des descripteurs harmoniques

5.1 Vue générale

5.1.1 Approches signal et symbolique : avantages et usages

Afin de pouvoir traiter l'harmonie à la fois en tant que phénomène sonore et ensemble de relations formelles, nous suivons deux approches. La première, dite symbolique, calcule les descripteurs harmoniques à partir d'une partition, tandis que la seconde, dite signal, part du signal audio. Nous avons vu ² que l'approche symbolique nécessite en entrée, en plus de la partition, un modèle de spectre, tandis que l'approche signal nécessite, pour certains descripteurs, la connaissance des *onsets* ainsi que le signal des voix séparées, en plus du signal global.

Les descripteurs harmoniques utilisés dans la suite du travail sont des descripteurs spectraux, c'est-à-dire calculés à partir d'une représentation spectro-temporelle du signal. Aussi, l'approche signal, qui part de l'objet sonore, est-elle plus directe que l'approche symbolique, qui court-circuite le son à travers un modèle de spectre. À ce titre, l'approche symbolique peut sembler artificielle, il n'en demeure pas moins qu'elle présente des avantages sur l'approche signal, à la fois en matière d'interprétation et d'usage.

2. Voir partie [4.3.2](#)

Approche symbolique Le principal avantage de l’approche symbolique réside dans la facilité et le faible temps de calcul, particulièrement utiles lorsqu’il s’agit de traiter des gros corpus musicaux ou des catalogues d’accords. Nous utilisons cette approche dans la partie III, pour étudier la structuration par les descripteurs harmoniques des espaces harmoniques abstraits comme les classifications d’accords ou les représentations continues des dyades et triades.

De plus, les éventuelles difficultés, dans l’approche signal, liées à la synchronisation des spectres sur les verticalités¹, ne se posent pas dans l’approche symbolique. En outre, l’usage de spectres de même forme assure une homogénéité aux valeurs des descripteurs et facilite leur interprétation.

Enfin, l’utilisation de modèles de spectres plutôt que de spectres réels présente un avantage théorique. En effet, le jeu des paramètres du modèle permet de se rendre compte directement de l’influence de facteurs timbraux comme le nombre de partiels ou leurs amplitudes respectives. D’autre part, le choix des spectres permet de construire des spectres *ad hoc* qui vérifient des propriétés particulières. C’est le cas des spectres de type Shepard², dans lequel les partiels sont reproduits sur toutes les octaves, avec une décroissance des amplitudes aux extrémités du spectre. Ils ont la propriété d’être associés aux hauteurs, et non directement aux fréquences fondamentales. Cela est particulièrement utile lorsqu’on souhaite se départir de la dépendance des descripteurs harmoniques au registre, et notamment la dépendance à la disposition des notes au sein d’un accord³.

Dans l’approche symbolique, les descripteurs harmoniques sont calculés directement sur les spectres modélisés. Une autre possibilité serait de réaliser une synthèse sonore de chaque note à partir de son spectre modélisé, par exemple une synthèse additive, et de calculer les descripteurs à partir de ces sons via une analyse spectrale. Cette étape supplémentaire alourdirait cependant la méthode, et si elle préserverait les avantages théoriques, elle perdrait la facilité et la rapidité de calcul. Nous ne passons donc pas par une synthèse sonore à partir du spectre modélisé.

1. Voir pages 155 et 243.

2. Voir section 5.2.

3. Les spectres de Shepard sont utilisés dans la partie III.

Approche signal Dans l’approche signal, les spectres sont calculés directement à partir de l’audio. L’avantage principal par rapport à l’approche symbolique est méthodologique, puisque l’objet d’étude est clairement identifié : c’est le signal audio. Les traces symboliques, à savoir la liste des *onsets* et l’exigence des pistes séparées, n’altère en rien la nature sonore de l’objet d’étude. Dans l’approche symbolique par contre, les traces sonores, c’est-à-dire le modèle de spectre, remettent en cause sa nature symbolique, car il n’est plus formé de la seule partition, mais de la partition associée à des paramètres spectraux d’un modèle de spectre.

L’approche signal a l’avantage d’être plus proche de la réalité sonore que l’approche symbolique. De ce fait, elle est plus adaptée pour rendre compte du timbre. Le choix de calculer les descripteurs sur les verticalités met cependant de côté le suivi temporel de phénomènes comme le vibrato ou le tremolo¹. La contrepartie de la relative fidélité au phénomène sonore, par rapport à l’approche symbolique, est la difficulté d’interprétation, qui peut apparaître du fait à la fois de la complexité du phénomène et des nombreux paramètres qui influencent les valeurs des descripteurs harmoniques, ainsi que des artefacts d’analyse, lesquelles sont absents de l’approche symbolique.

5.1.2 Les différentes configurations

Nous distinguons trois configurations différentes des données en entrée pour le calcul des descripteurs harmoniques. Ces configurations dépendent à la fois des descripteurs, des données de départ et des choix de calcul.

Configuration enregistrement C’est lorsque les descripteurs sont calculés sur le signal global, sans prise en compte du signal des voix séparées. Dans l’approche signal, c’est la configuration en vigueur lorsqu’on a à disposition seulement un enregistrement global, et non les voix séparées.

Certains descripteurs, comme l’harmonicité, sont calculés dans cette configuration, même en présence du spectrogramme de toutes les voix. D’autres descripteurs, comme la concordance ou la concordance totale, ne peuvent pas être calculés dans cette configuration, car nécessitent les voix séparées.

Configuration mélodie-accompagnement Dans cette configuration, une voix est discriminée par rapport aux autres voix, et les descripteurs harmo-

1. Voir page [155](#)

niques décrivent alors l'interaction de cette voix particulière avec les autres voix. Musicalement, une telle discrimination ne peut avoir de sens que lorsque cette voix possède un statut particulier, ce qui est le cas lorsqu'on est en présence d'une mélodie et de son accompagnement. Cela nécessite de pouvoir distinguer mélodie et accompagnement, ce qui n'est pertinent que pour un corpus musical précis.

Dans l'approche signal, cette configuration nécessite d'avoir les pistes séparées de la mélodie et de l'accompagnement. L'accompagnement peut alors contenir plusieurs notes simultanément, il n'y a pas la contrainte d'une note maximum par piste.

Cette démarche consistant à dégager les propriétés d'une note relativement à l'ensemble de la verticalité à laquelle elle appartient est située historiquement. Elle est initiée avec Rameau, qui considère la consonance d'une note dans un agrégat sonore relativement à la basse fondamentale, laquelle est déduite de l'agrégat¹. Dans le cas de la consonance et la dissonance, James Tenney s'y réfère sous la dénomination de CDC 4².

Certains descripteurs comme la concordance d'ordre 3 ou la concordance totale nécessitent d'avoir toutes les voix séparées, et ne peuvent pas être calculés dans cette configuration.

Configuration voix séparées Cette configuration nécessite les spectres de chaque note, donc les spectrogrammes de chaque voix, où une voix contient au maximum une note. Dans l'approche signal, cela implique d'avoir le signal des voix séparées, en plus du signal global.

Certains descripteurs ne peuvent être calculés que dans cette configuration, par exemple la concordance totale.

Dans la suite du travail, la configuration « mélodie-accompagnement » ne sera pas traitée. Nous nous concentrerons sur les configurations « enregistrement » et « voix séparées ».

5.1.3 Modélisation du contexte harmonique

La prise en compte du contexte harmonique dans l'implémentation des descripteurs harmoniques est motivée, nous l'avons vu, à la fois par des rai-

1. Voir le *Traité de l'harmonie réduite à ses principes naturels* (Rameau, 1722).

2. Voir (Tenney, 1988), et page 75 pour la présentation des différents CDC.

sons d'ordre acoustique et d'ordre cognitif¹.

Les premières traduisent un chevauchement temporel, il s'agit de tenir compte du phénomène de réverbération de la salle, à cause duquel un son se fait entendre encore après que sa production a cessé. Les acoustiques réverbérantes comme les églises sont recherchées pour certains répertoires (orgue, musique chorale), mais sont considérées comme peu adaptées pour d'autres. La réverbération influence grandement la perception de la musique, et agit directement, à la fois qualitativement et quantitativement, sur les interactions entre les notes, autrement dit sur le phénomène harmonique. Il paraît ainsi essentiel de proposer une modélisation de la réverbération dans l'implémentation des descripteurs harmoniques.

La seconde raison est cognitive, il s'agit de modéliser la mémoire à court terme, qui joue un rôle prépondérant dans la perception de l'harmonie².

État de l'art sur la modélisation de la mémoire à court terme dans la perception de l'harmonie La mémoire à court terme agit sur la perception de la musique par intégration temporelle à un instant donné des éléments précédemment entendus. Son effet est bien connu en ce qui concerne l'intensité sonore (Munson, 1947; Algorn et Marks, 1990). Pour ce qui est de la rétention de la hauteur des sons isolés, elle est de l'ordre de 30 secondes (Winkler *et al.*, 2002). Rogers (Rogers, 2010) a montré que la mémoire des intervalles est plus longue que la mémoire des notes isolées, allant jusqu'à 48 secondes.

La mémoire à court terme est cruciale dans la perception de l'harmonie, mais son fonctionnement dans ce cadre reste encore peu étudié. Jensen souligne ce manque :

*Les mesures de dissonance sensorielle ne prennent pas en compte l'intégration temporelle ou la mémoire. Pourtant, la perception de la dissonance est probablement affectée par le contexte temporel, au même titre que les autres propriétés psychophysiques du son*³.

Les modèles de mémoire à court terme sont de deux types : ceux qui

1. Voir partie 4.3.4.

2. Voir page 123.

3. "Measures of sensory dissonance do not take temporal integration or memory into account. The perception of dissonance is, nonetheless, likely to be affected by the local temporal context as other psychophysical sound properties are.", tiré de (Jensen et Hjortkjær, 2012), traduction de l'auteur.

considèrent que le contenu de la mémoire est limité par le nombre d'éléments déjà présents, d'autres par le facteur temporel. Miller ([Miller, 1956](#)) considère qu'on peut retenir jusqu'à 7 ± 2 éléments, tandis que Cowan ([Cowan, 2001](#)) restreint ce chiffre à 4 éléments. Malgré les résultats de Winkler et Rogers, les modèles temporels s'accordent sur une durée de la mémoire à court terme de 3 à 7 secondes ([Radvansky, 2015](#); [Snyder et Snyder, 2000](#)), avec une décroissance temporelle logarithmique du coefficient d'activation d'un évènement ([Anderson et Lebiere, 2014](#)).

Certains modèles tiennent compte des deux facteurs ([Barrouillet *et al.*, 2004](#); [Jensen et Hjortkjær, 2012](#)). Parmi eux, le modèle de Jensen et Hjortkjær est développé directement pour prendre en compte l'influence du contexte harmonique dans la dissonance perceptive.

Dans ce modèle, la dissonance à un temps t est calculée en faisant la somme de la dissonance statique — d'après les modèles de ([Plomp et Levelt, 1965](#); [Sethares, 1999](#)) — et de la dissonance entre les verticalités présentes dans la mémoire et la verticalité courante au temps t . Le coefficient d'activation d'une verticalité dans la mémoire s'exprime sous la forme :

$$A = 2 - 0.5 \ln(t + 1) - 0.5 \ln(N_c + 1), \quad (5.1)$$

où t est le temps écoulé en secondes, et N_c le nombre d'évènements présents dans la mémoire au temps considéré. Dans la mémoire sont présentes seulement les verticalités passées qui ont un coefficient d'activation positif. Une verticalité passée occupe d'autant plus de place dans la mémoire qu'elle est récente et qu'il y a eu peu d'évènements depuis.

La figure [5.1](#) illustre le profil des coefficients d'activation pour chaque élément d'une séquence isochronique d'évènements séparés de 0.5s. La taille de la mémoire se stabilise ici à 10 évènements.

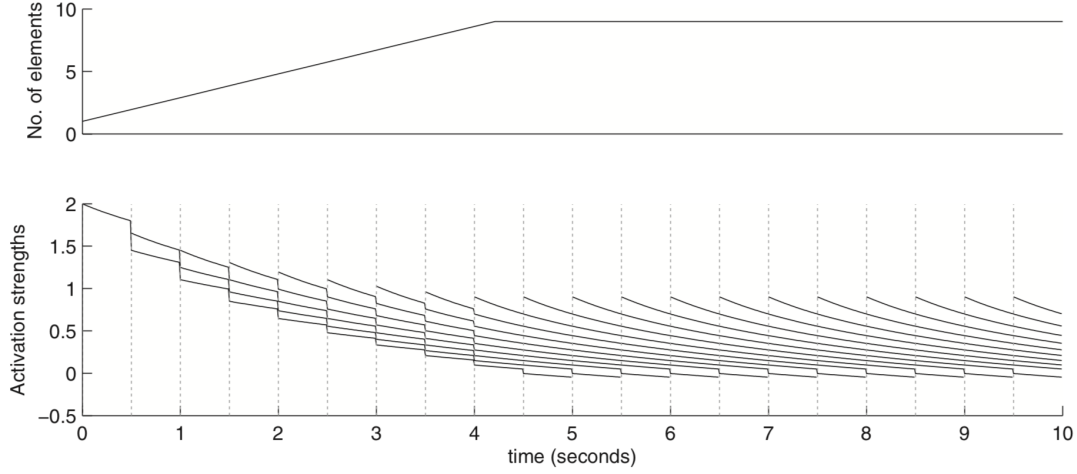


FIGURE 5.1 – Profil des coefficients d’activation pour chaque élément d’une séquence isochronique d’évènements, selon l’équation 5.1. Tiré de ([Jensen et Hjortkjær, 2012](#))

Implémentation Nous proposons une modélisation du contexte harmonique sous forme d’une mémoire à court terme limitée par le nombre de verticalités uniquement. Cette restriction est due au contexte particulier d’application de ce modèle dans ce travail¹, limité aux transitions entre classes d’accords, sans prise en compte de paramètres rythmiques ou métriques.

Dans une suite de verticalités (œuvre, extrait d’œuvre, suite d’accords...), on associe à chacune son spectre de contexte, lequel est calculé par une moyenne pondérée des spectres des verticalités précédentes. Cette moyenne est réalisée sur chaque plage de fréquence du spectre, et elle dépend de deux paramètres : le nombre maximal de verticalités présentes dans la mémoire (taille de la mémoire), ainsi que le coefficient de décroissance pour la pondération, qui accorde plus d’importance aux verticalités récentes.

Le spectre de contexte $\mathcal{S}_t^{ctx}$ de la verticalité t s’exprime sous la forme :

$$\mathcal{S}_t^{ctx} = \frac{\mathcal{S}_t + \frac{\mathcal{S}_{t-1}}{2^d} + \dots + \frac{\mathcal{S}_{t-l+1}}{l^d}}{1 + 1/2^d + \dots + 1/l^d}, \quad (5.2)$$

1. Voir partie 9.3.

où $\mathcal{S}_t$ désigne le spectre de la verticalité t , d le coefficient de décroissance, et l la longueur de la mémoire ou le nombre de verticalités à partir du début, ci celui-ci est inférieur à la longueur de la mémoire.

La figure 5.2 représente le profil des coefficients de pondération des verticalités dans une séquence d'évènements, avec $l = 4$ et $d = 1$.

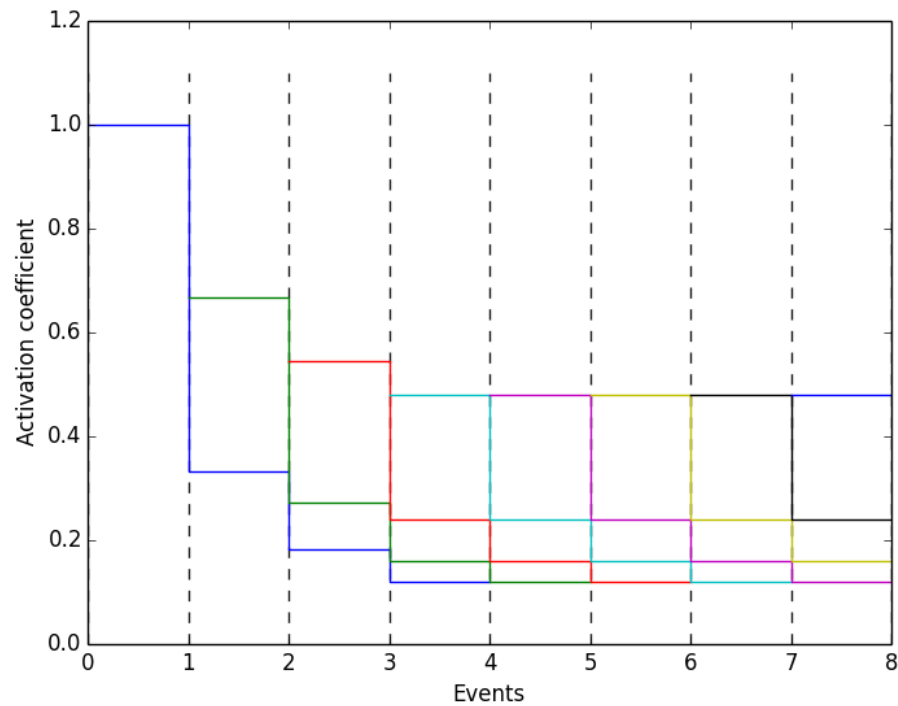


FIGURE 5.2 – Profil des coefficients de pondération des verticalités pour le calcul du spectre de contexte, dans une séquence d'évènements, d'après l'équation (5.2). Les paramètres sont les suivants - longueur de mémoire : 4, coefficient de décroissance : 1. La somme des coefficients sur chaque verticalité égale 1.

5.2 Approche symbolique

Dans l'approche symbolique, l'objet de départ est une partition, et le signal audio est court-circuité via un modèle de spectre qui associant un spectre à chaque note de la partition. Le schéma général est le suivant :

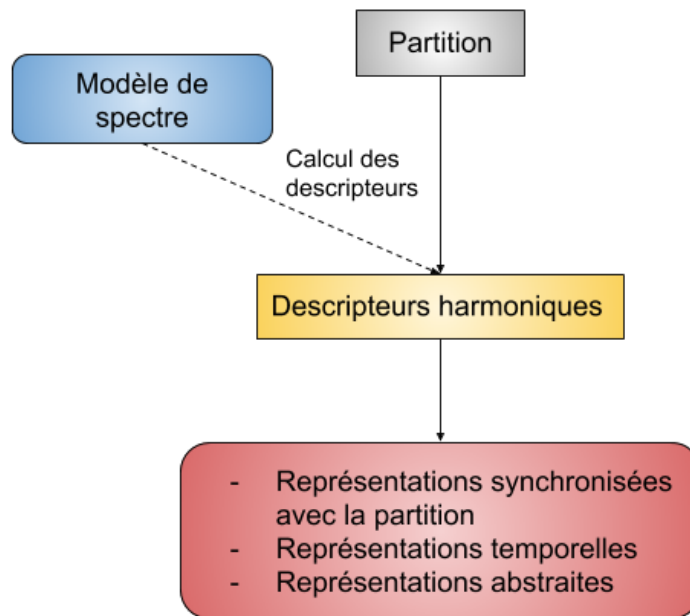


FIGURE 5.3 – Schéma de l'approche symbolique

Nous proposons deux modèles de spectre différents : la modélisation d'un spectre réel où chaque partiel est représenté par une gaussienne, et un spectre de type Shepard, utilisé pour se départir de la dépendance à la disposition. Dans les deux cas, le spectre modélisé est un spectre en amplitude. Ce choix est motivé par les descripteurs harmoniques considérés, qui ne tiennent pas compte des effets de phase. De plus, étant donné que les descripteurs harmoniques pris en compte décrivent l'interaction entre partiels, nous modélisons la composante harmonique du son, sans tenir compte des composantes percussives ou bruitées.

5.2.1 Spectre en gaussiennes

Partiels ponctuels vs. continus Nous cherchons à modéliser les sons stables spectralement, donc décomposables par transformée de Fourier en une somme de sons simples, c'est-à-dire de sons sinusoïdaux. Le modèle de spectre le plus simple que l'on puisse imaginer pour de tels sons est le modèle de spectre en pics, où chaque partiel est représenté par un pic sur l'axe des fréquences, et les hauteurs des pics déterminent l'amplitude des partiels.

De nombreux modèles de perception partent de telles représentations de spectres en pics ([Parncutt, 1989](#); [Cook, 2009](#)).

En pratique, cependant, la durée des verticalités n'est pas infinie, mais varie entre quelques dizaines de millisecondes et quelques secondes, voire plus, ce qui rend la représentation en pics irréaliste du point de vue physique. En effet, sans hypothèse de périodicité sur les sons utilisés, le calcul des coefficients de Fourier au sens des distributions de Dirac nécessite une intégration sur une fenêtre temporelle infinie. En outre, sur une analyse spectrale, la longueur finie des fenêtres implique une discrétisation des fréquences en bandes de fréquences, dans laquelle un partiel occupe au mieux une bande de fréquence.

Une amélioration des spectres en pics consiste à donner une largeur spectrale aux partiels. Cette étape peut être justifiée d'au moins trois manières, par :

- les propriétés de l'analyse spectrale : des artefacts de calcul peuvent contribuer à la largeur spectrale des partiels.
- des propriétés inhérentes au son : évolution temporelle du son au sein d'une fenêtre d'analyse (vibrato, glissando).
- des propriétés de la perception, de nature physiologique ou psychologique. Dans le premier cas, la largeur des partiels vient des propriétés mécaniques des vibrations sur la membrane basilaire de l'oreille¹. Dans le second cas, elle traduit l'aspect probabiliste de la perception des hauteurs, qui résulte en particulier de la perception de la même hauteur pour deux partiels proches². Dans cette approche le spectre n'est pas une représentation purement acoustique, mais acquiert une dimension perceptive.

1. Comme dans l'hypothèse physiologique d'Helmholtz ([Helmholtz, 1863](#)).

2. Voir ([Milne et al., 2011](#))

Modèle en gaussiennes Nous proposons un modèle de spectre où chaque partiel est représenté par une gaussienne centrée sur sa fréquence, en nous inspirant du modèle de Milne (Milne *et al.*, 2011, 2016). Helmholtz, dans son hypothèse physiologique, assimile chaque fibre de la membrane basilaire à un résonateur suivant une loi d’influence gaussienne. Dans notre modèle, un spectre est donc donné par la liste de ses partiels, où chaque partiel est caractérisé par sa fréquence, son amplitude (hauteur de la gaussienne), et sa largeur spectrale (écart-type de la gaussienne).

Un tel modèle permet de modéliser une grande variété de situations sonores, et notamment les sons inharmoniques. Dans la suite du travail, nous nous restreignons cependant aux sons harmoniques, pour lesquels nous faisons les hypothèses suivantes, qui visent à simplifier l’implémentation, et peuvent être mises de côté si nécessaire :

- Tous les partiels ont la même largeur sur l’axe logarithmique des fréquences. Nous appelons σ l’écart-type commun à tous les partiels, où l’axe des fréquence est en $\log_2(f)$. Dans l’interprétation de σ comme amplitude de vibrato, cette simplification est équivalente à l’hypothèse que cette amplitude ne varie pas avec le registre. Celle-ci paraît raisonnable, comme le suggère la figure 5.19 qui montre un spectre synchronisé obtenu à partir de sons échantillonnés de cordes frottées avec vibrato, dans laquelle l’amplitude en fréquence des partiels est globalement homogène.
- L’amplitude des partiels décroît de manière homogène avec le numéro du partiel. La décroissance est caractérisée par le facteur de décroissance $decr$, où l’amplitude du k^e partiel est $\frac{1}{k^{decr}}$.

Même si la prise en compte de partiels étalés est justifiée par les considérations physiques et techniques citées plus haut ainsi que par la nature même de certains descripteurs utilisés dans la suite, le choix des gaussiennes est motivé avant tout par la perception et son aspect probabiliste. Chaque gaussienne peut alors être vue comme une loi normale de distribution de la perception autour d’une fréquence. L’égalité de l’écart-type des distributions au sein d’un spectre peut être discutée¹, elle permet tout du moins de séparer

1. Cette question est liée au seuil différentiel de perception des hauteurs, ou *just-noticeable difference* en anglais - voir (Moore, 1973). Il s’agit du plus petit écart de fréquence en dessous duquel deux fréquences sont perçues comme égales.

dans le spectre modélisé $\mathcal{S}_{gauss}$ la partie déterministe $\mathcal{S}_{pics}$ (spectre en pics) de la partie probabiliste $\mathcal{N}(\sigma)$ (gaussienne) :

$$\mathcal{S}_{gauss} = \mathcal{S}_{pics} * \mathcal{N}(\sigma), \quad (5.3)$$

où $*$ désigne le produit de convolution.

Ce modèle de spectre en gaussiennes est déjà présent chez Milne ([Milne et al., 2011, 2016](#)), qui interprète les paramètres σ et $decr$ comme des paramètres perceptifs.

Certains descripteurs présentés dans la suite nécessitent des partiels non ponctuels, ils sont alors calculés directement à partir des spectres en gaussiennes. D'autres, par contre, ne prennent en entrée que la série des partiels avec leur fréquence et leur amplitude, ils sont alors calculés à partir d'un spectre en pic tel celui de la figure [5.4](#).

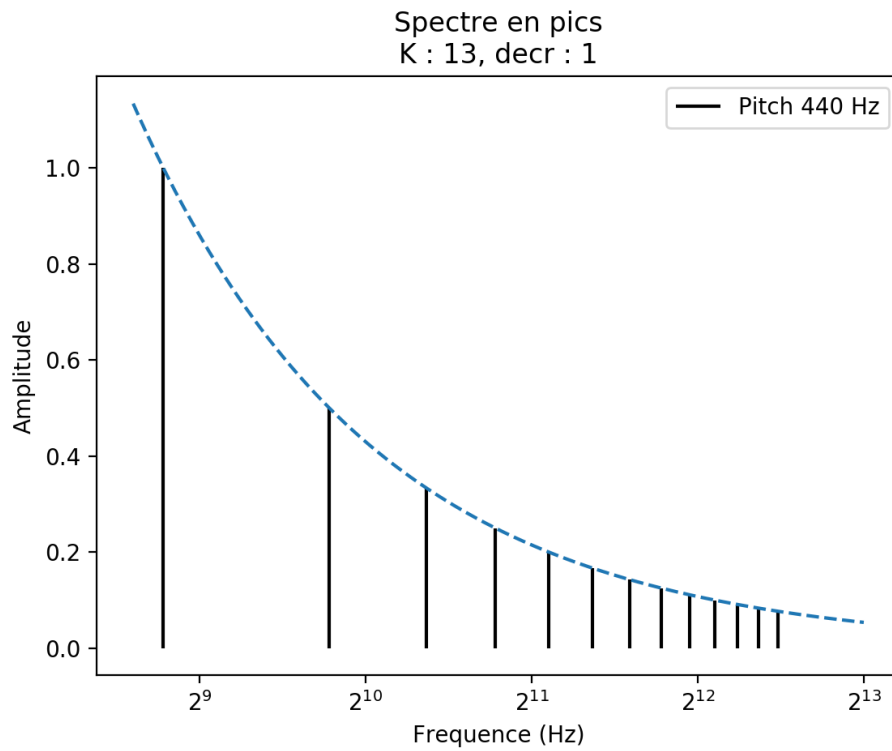


FIGURE 5.4 – Spectre en pics d’un son à 440 Hz, avec 13 partiels et une décroissance des partiels en $1/k$.

La figure 5.5 montre un exemple de spectre avec une largeur spectrale des partiels non nulle, en l’occurrence $\sigma = 0.02$.

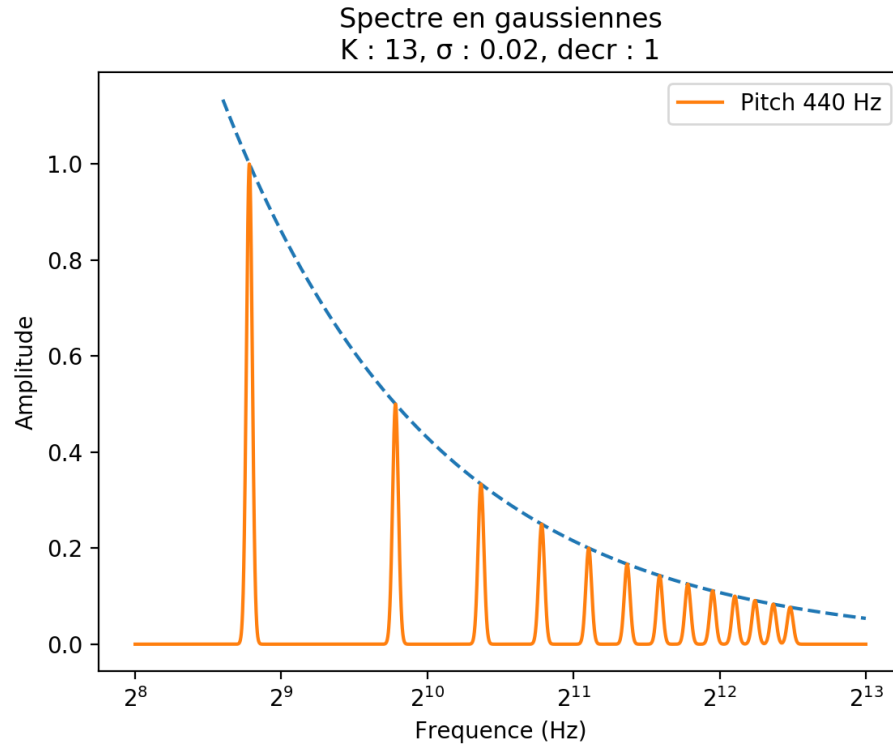


FIGURE 5.5 – Spectre en gaussiennes d’un son à 440 Hz, avec 13 partiels, une décroissance des partiels en $1/k$ et une largeur des partiels $\sigma = 0.02$.

5.2.2 Spectres de Shepard

Les spectres de la partie précédente sont associés à une fréquence fondamentale, qui est la fréquence du premier partiel. Or, nous avons besoin, dans certains cas, de pouvoir associer un spectre à une hauteur fondamentale, et non plus à une fréquence fondamentale. C’est notamment le cas lorsque nous voulons calculer les descripteurs sur des classifications d’accords ou des types d’accords, où les fréquences sont données modulo l’octave. L’emploi des hauteurs implique l’utilisation de spectres harmoniques¹. Il s’agit alors de se départir des considérations liées au registre, au renversement, aux doublures et plus largement à la disposition.

Les sons de Shepard (*Shepard tones*) sont construits en démultipliant

1. Voir page 60.

un partiel sur toutes les octaves du spectre, avec une enveloppe gaussienne pour assurer la décroissance des amplitudes aux extrémités du spectre. Ils répondent au problème précédent, car sont associés précisément aux hauteurs.

Usage des spectres de Shepard Les sons de Shepard ont été introduits par Shepard ([Shepard, 1964](#)), comme sons dont les partiels sont espacés d’une octave et s’étendent sur toute l’étendue du spectre, soit 10 octaves¹. Ces sons remettent en question la linéarité de la perception de la dimension grave-aigu pour mettre en avant la sensibilité à la hauteur² et privilégier une représentation circulaire.

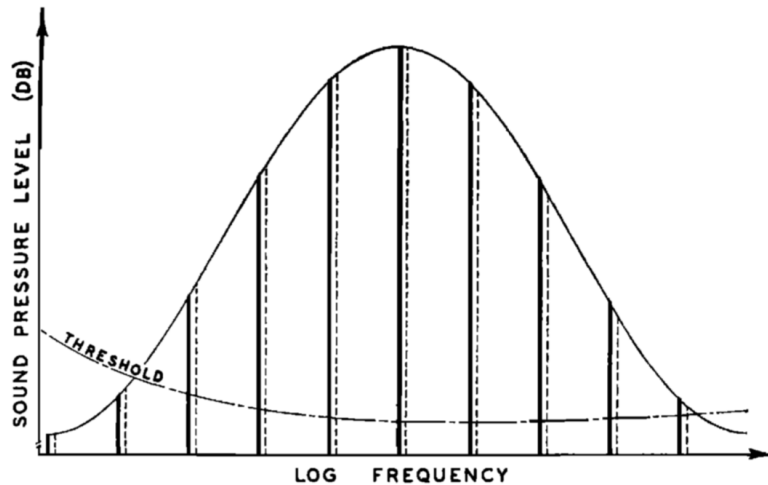


FIGURE 5.6 – Spectre de Shepard dans sa version originelle, tire de ([Shepard, 1964](#)). Les 10 composants spectraux sont séparés d’une octave.

Quelques années plus tard, Jean-Claude Risset utilise ces sons pour créer des paradoxes de hauteur avec des sons qui n’en finissent pas de monter, en *glissando* infini ([Risset, 1969](#)). Il donne une application musicale de ces paradoxes dans des pièces comme *Mutations* (1969).

Parncutt ([Parncutt, 1993](#)) introduit des accords à partir de sons de Shepard, et étudie des aspects perceptifs de l’harmonie formée de ces sons.

1. Voir figure 5.6.

2. Au sens défini page 60.

Deutsch ([Deutsch et al., 2008](#)) utilise des sons de Shepard généralisés à des séries harmoniques, avec plusieurs hauteurs de partiels.

Récemment, les recherches convoquant les sons de Shepard se concentrent sur un aspect particulier de leur perception : l’ambiguïté liée à l’enchaînement de deux sons de Shepard séparés par la classe d’intervalle du triton, qui est tantôt perçu comme ascendant ou comme descendant¹. Certaines études étudient l’influence du timbre ([Russo et Thompson, 2005](#)), d’autres la sensibilité au contexte ([Repp et Thompson, 2010](#); [Siedenburg, 2018](#)) sur ce problème.

Implémentation Nous utilisons des spectres de Shepard généralisés à des séries harmoniques de partiels, avec des modèles en gaussiennes comme dans la partie précédente².

Pour passer des amplitudes des partiels du spectre harmonique aux amplitudes relatives des partiels du spectre de Shepard, nous commençons par associer un coefficient à chaque hauteur présente dans la série des partiels en sommant les amplitudes des partiels correspondants.

Prenons l’exemple du spectre modélisé d’un *la* 3, avec 5 partiels et un facteur de décroissance égal à 1. Les 5 partiels *la* 3, *la* 4, *mi* 4, *la* 5 et *do*♯ 6 ont respectivement pour amplitudes $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ et $\frac{1}{5}$. Le spectre de Shepard de *la* est composé des hauteurs *la*, *mi* et *do*♯, affublés respectivement des coefficients $1 + \frac{1}{2} + \frac{1}{4} = \frac{7}{4}$, $\frac{1}{3}$ et $\frac{1}{5}$.

Les spectres utilisés s’étendent sur 16 octaves, avec une enveloppe spectrale sous forme de gaussienne centrée sur *do* 4, et un écart-type Σ ajustable, compté en octaves. Sauf mention contraire, Σ est fixé à deux octaves. Les partiels sont démultipliés sur toutes les octaves, avec une amplitude correspondant au coefficient relatif de leur hauteur multipliée par l’enveloppe.

Pour certains descripteurs ne nécessitant que la liste des partiels avec leur fréquence et leur amplitude, nous utilisons les spectres de Shepard en pics, tel celui de la figure [5.7](#).

1. Ambiguïté déjà relevée dans ([Shepard, 1964](#)).

2. Voir figure [5.5](#).

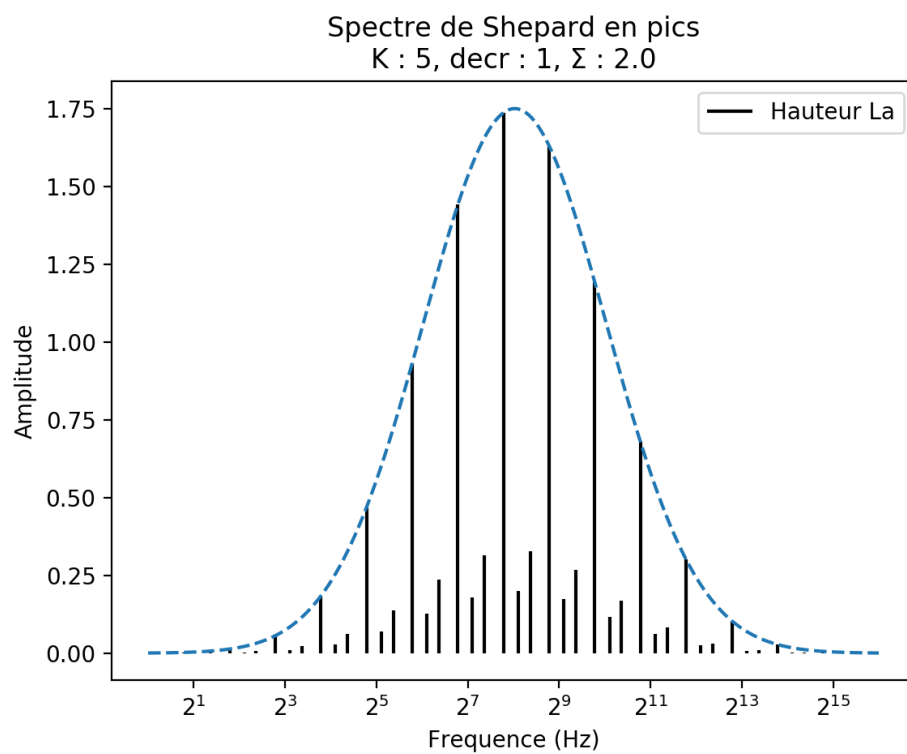


FIGURE 5.7 – Spectre de Shepard en pics d’un *la* (diapason à 440 Hz), avec 5 partiels, une décroissance des partiels en $1/k$ et $\Sigma = 2$ octaves.

La figure 5.8 montre le spectre de Shepard d’un *la*, avec un diapason à 440 Hz, 5 partiels, un facteur de décroissance égal à 1 et une largeur spectrale des partiels $\sigma = 0.03$.

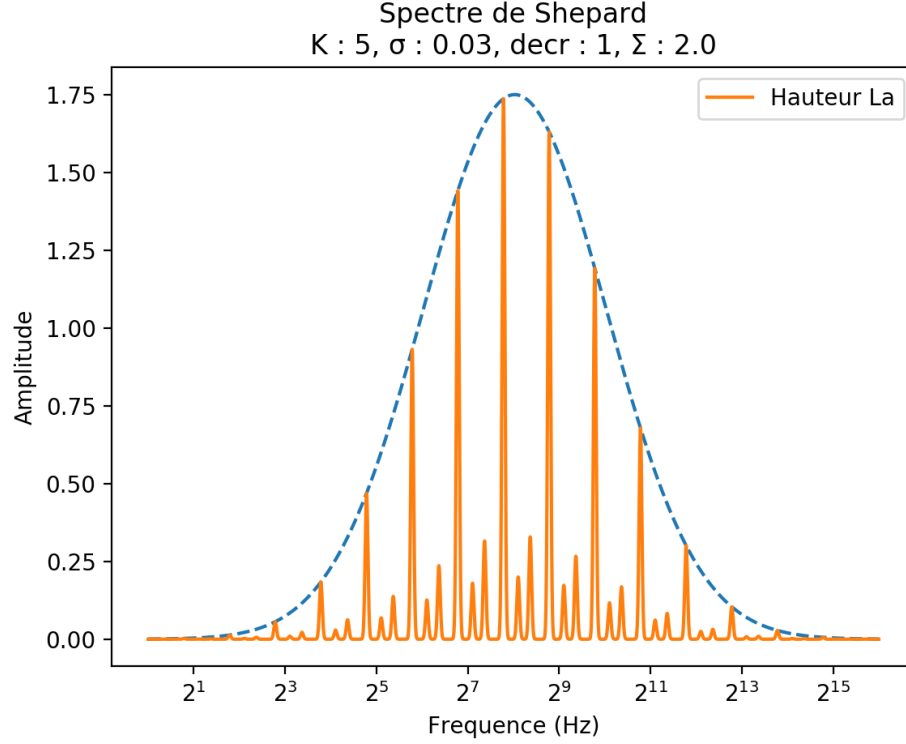


FIGURE 5.8 – Spectre de Shepard en gaussiennes d’un *la* (diapason à 440 Hz), avec 5 partiels, une décroissance des partiels en $1/k$, une largeur des partiels $\sigma = 0.03$ et $\Sigma = 2$ octaves.

L’enveloppe spectrale en gaussienne a pour but d’homogénéiser l’énergie du spectre sur toutes les hauteurs, celle-ci étant définie comme le carré de la norme $\mathcal{L}_2$ du spectre. L’énergie $\mathcal{E}$ d’un spectre en amplitude $\mathcal{S}$ s’exprime de la manière suivante :

$$\mathcal{E}(\mathcal{S}) = \sum_f \mathcal{S}(f)^2, \quad (5.4)$$

où f parcourt tout l’axe logarithmique des fréquences.

Cette condition d’homogénéité est importante pour le calcul des descripteurs harmoniques, car des variations d’énergie provoqueraient des effets indésirables, tels la dépendance de la concordance des classes d’intervalles à la hauteur de départ. Cette condition d’homogénéité est vérifiée pour les spectres en gaussiennes qui ne sont pas de type Shepard, pour lesquels la

forme du spectre, et donc aussi son énergie, sont constants sur l'axe des fréquences.

La figure 5.9 montre la variation de l'énergie sur toutes les hauteurs. L'énergie est bien constante.

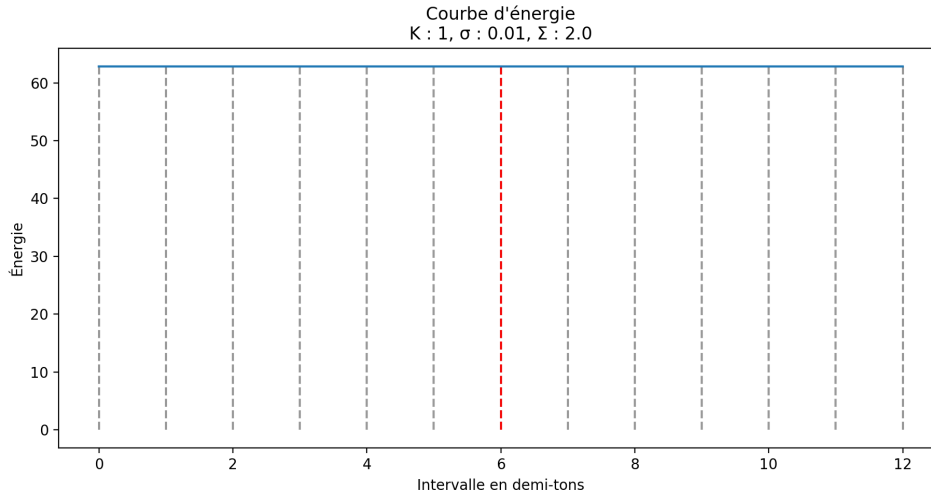


FIGURE 5.9 – Variation de l'énergie des spectres de Shepard à 1 partiel sur tous les hauteurs. Paramètres spectraux : $\sigma = 0.03$ et $\Sigma = 2$ octaves.

Pour se convaincre du rôle de l'enveloppe en gaussienne, notamment de son écart-type, représentons sur la figure 5.10 le coefficient de variation (rapport écart-type sur moyenne) de l'énergie sur toutes les hauteurs en fonction de l'écart-type de l'enveloppe spectrale.

On distingue deux comportements : le coefficient de variation commence à décroître jusqu'à un écart-type situé autour de $\Sigma = 1.6$, les variations de l'énergie sont alors dues aux inhomogénéités sous le dôme de la gaussienne, puis croissent, car les contributions en dehors du dôme de la gaussienne, qui jouaient un rôle dans l'homogénéisation des valeurs de l'énergie, s'amenuisent. Le choix de $\Sigma = 2$ est motivé par la position du minimum de variation relative de l'énergie, dont il est l'arrondi entier. De plus, avec cette valeur, l'intervalle spectral correspondant à une enveloppe située à plus de 6% du sommet de la gaussienne s'étend sur 11 octaves, qui est l'ambitus maximum d'audition humaine (de 20 à 20000 Hz).

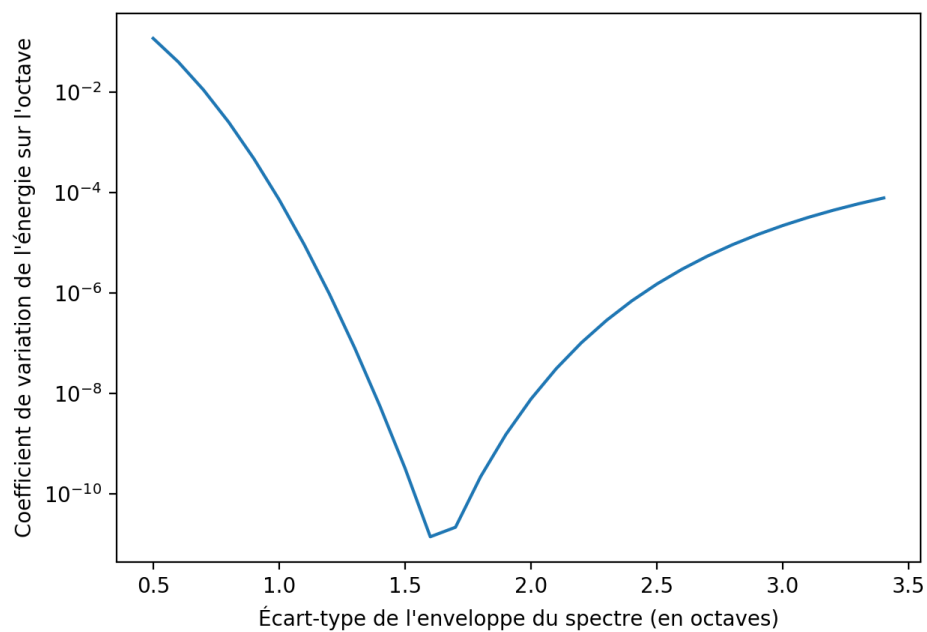


FIGURE 5.10 – Évolution du coefficient de variation de l'énergie d'un spectre de Shepard à 1 partiel sur tous les hauteur en fonction de l'écart-type de l'enveloppe spectrale. On a pris $\sigma = 0.01$.

5.3 Approche signal

La figure 5.11 montre le schéma général de l'approche signal. Les descripteurs calculés dans la configuration « enregistrement »¹ ne nécessitent que le signal global en entrée, tandis que la configuration « mélodie-accompagnement » nécessite les pistes séparées de la mélodie et de l'accompagnement, et la configuration « voix séparées » les pistes séparées de toutes les voix, où chaque voix contient au plus une note.

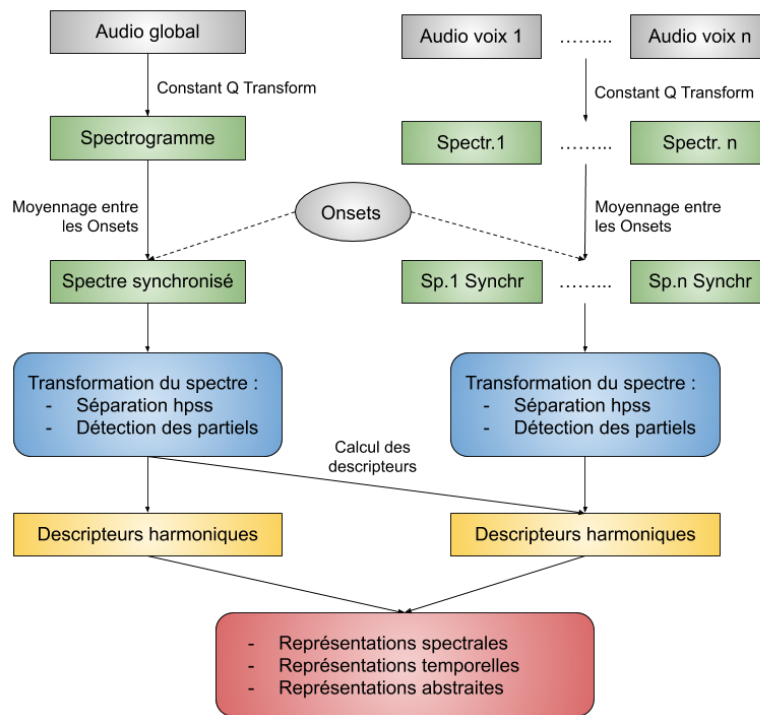


FIGURE 5.11 – Schéma de l'approche signal

5.3.1 Analyse spectrale

Les descripteurs harmoniques utilisés dans la suite étant des descripteurs spectraux, c'est-à-dire calculés sur des spectrogrammes, la première étape dans le traitement du signal est de représenter le signal dans le plan temps-fréquence. Nous utilisons une transformée à facteur Q constant.

1. Voir page 163.

Transformée à facteur Q constant La transformée à facteur Q constant (TQC) a l'avantage, par rapport à la transformée de Fourier à court terme (TFCT), de préserver la même précision spectrale sur l'axe logarithmique des fréquences, sans dégradation dans les graves. Cela la rend particulièrement adaptée à l'usage musical, puisque dans les échelles tempérées l'espacement des notes est constant sur tout l'ambitus en logarithme. En outre, le choix de la précision spectrale peut-être adapté à l'échelle musicale utilisée. Nous reprenons ici l'algorithme décrit dans (Schörkhuber et Klapuri, 2010). Le facteur de qualité Q exprime le rapport d'une fréquence f_k sur la précision spectrale (plage de largeur Δf_k) autour de cette fréquence :

$$Q(f_k) = \frac{f_k}{\Delta f_k}. \quad (5.5)$$

Avec Q constant sur tout le spectre, on peut le réécrire sous la forme :

$$Q = \frac{1}{2^{1/B} - 1}. \quad (5.6)$$

$B = 12$ correspond à une précision au demi-ton, $B = 24$ à une précision spectrale au quart de ton.

Comme par ailleurs

$$Q = \frac{N_k f_k}{\Delta \omega f_s}, \quad (5.7)$$

où f_s désigne la fréquence d'échantillonnage, N_k la longueur de la fenêtre d'analyse pour la fréquence f_k , et $\Delta \omega$ la largeur de bande à -3dB du lobe principal de la fenêtre d'analyse, on a :

$$N_k = \frac{\Delta \omega f_s}{f_k(2^{1/B} - 1)}. \quad (5.8)$$

Ainsi, $N_k(f_k)$ est inversement proportionnelle à la fréquence f_k . La contrepartie d'une précision fréquentielle constante est une précision temporelle variable, qui se dégrade dans les fréquences graves.

Dans la suite, nous utilisons une précision fréquentielle au 16^e de ton, ce qui correspond à $B = 8 * 12$. Cette précision est nécessaire pour des descripteurs harmoniques comme la rugosité, qui sont très sensibles aux petits écarts de fréquence.

Les sons utilisés sont échantillonnés à la fréquence $f_s = 22050Hz$, et la fenêtre utilisée est une fenêtre de Hanning, avec $\Delta \omega = 1.5$

Voici les longueurs de fenêtres en secondes pour certaines fréquences :

Fréquence d'analyse	Longueur de la fenêtre
do 6 (1046 Hz)	0.20 s
do 4 (261 Hz)	0.79 s
do 3 (130 Hz)	1.60 s

Le pas temporel entre les fenêtres d'analyse est choisi constant, ne dépendant pas de la fréquence. Nous le fixons égal à 512 échantillons, soit $23ms$.

Nous utilisons l'implémentation de la TQC disponible sur la librairie *Librosa*¹. La figure 5.12 représente la TQC d'une réalisation de la cadence à 4 voix de la figure 5.13 avec les jeux *Bourdon + Flûte 4*².

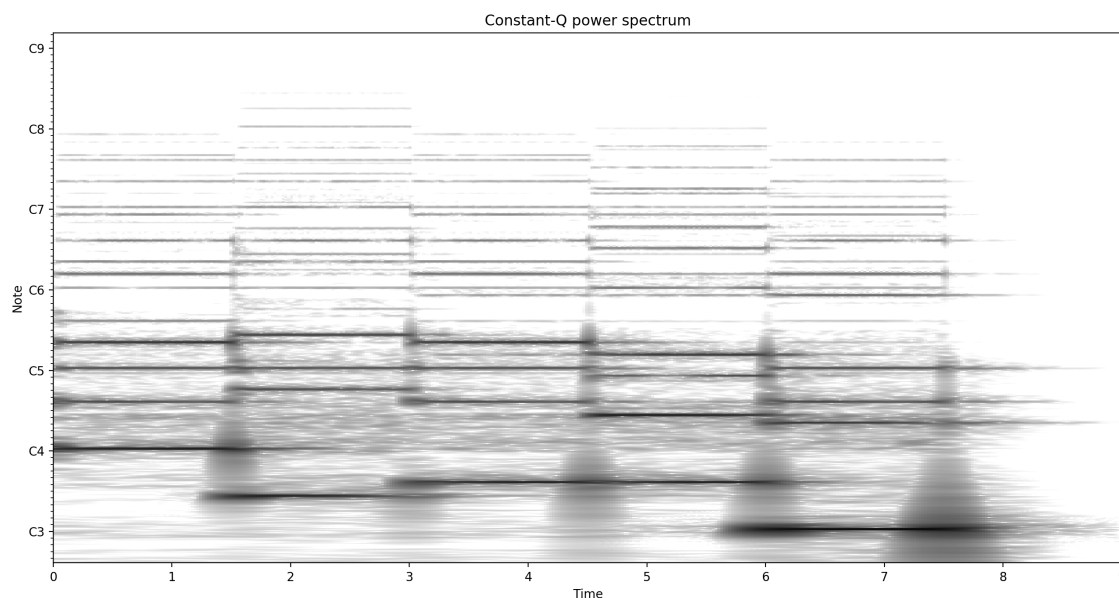


FIGURE 5.12 – Transformation à facteur Q constant d'une cadence à 4 voix à l'orgue, avec une précision spectrale au 16^e de ton. Jeux *Bourdon + Flûte 4*. Ambitus de l'analyse : *sol 2 - do 9*. Les fenêtres d'analyse sont centrés sur les temps d'analyse.

1. Version 7.2, disponible à l'adresse <https://librosa.org>.
2. Avec la librairie *Orgue de Salon*, voir page 4.2.1



FIGURE 5.13 – Cadence à 4 voix

Sur la figure 5.12, la dégradation de la précision temporelle dans le grave est bien visible, elle résulte de l'élargissement des fenêtres d'analyse. Le fait que les fenêtres d'analyse soient centrées sur les temps d'analyse¹ donne l'impression que les fréquences graves commencent avant les sons aigus, car elles sont détectées avant qu'elles ne soient jouées. Pour remédier à ce problème, nous modifions l'implémentation de manière à ce que les fréquences détectées aient toutes été déjà jouées. Cela revient à décaler les fenêtres d'analyse vers la gauche de manière à ce qu'elles se ferment sur le temps d'analyse, comme indiqué sur la figure 5.14.

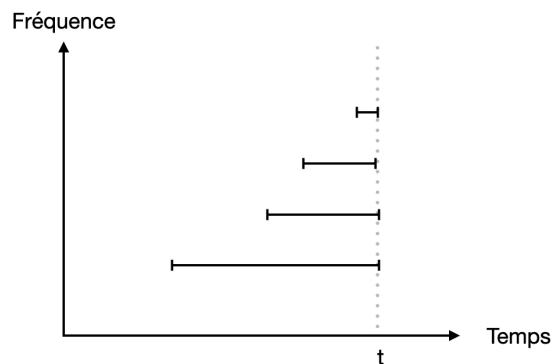


FIGURE 5.14 – Position des fenêtres d'analyse en fonction de la fréquence par rapport aux temps d'analyse dans le calcul de la TQC.

1. Implémentation par défaut de la TQC dans *Librosa*.

Nous obtenons de cette manière le profil de TQC de la figure 5.15. Cette fois les partiels graves ne sont pas détectés avant leur apparition. En contrepartie, le débordement vers la droite des fréquences graves après extinction du son est plus important. Ce débordement ne correspond pas à une réalité sonore, mais est dû à un artefact de calcul. Une solution sera proposée dans la suite¹ pour le minimiser.

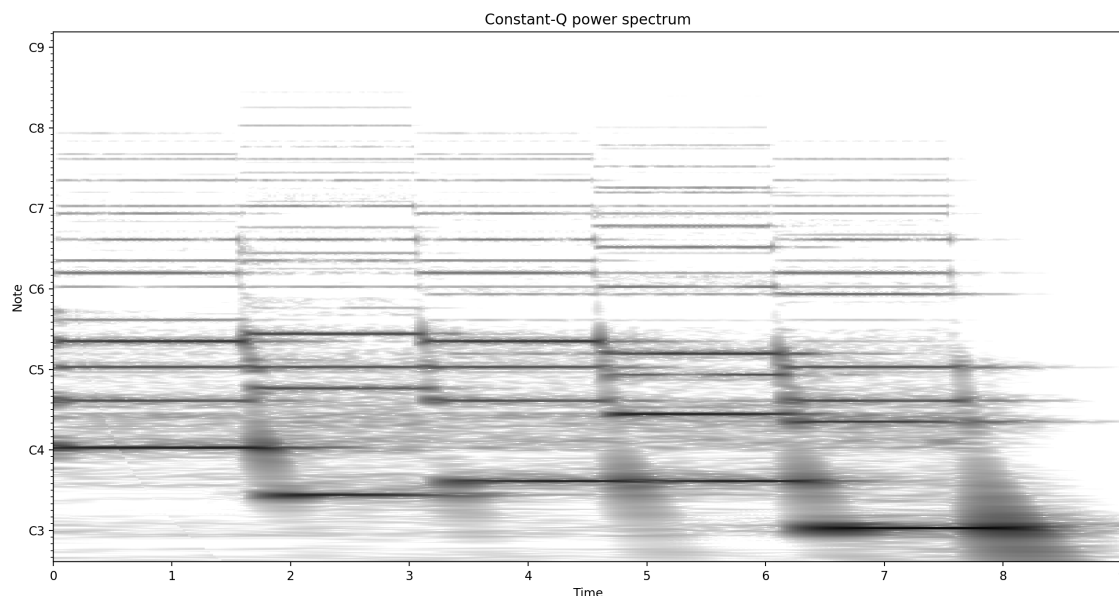


FIGURE 5.15 – Transformée à facteur Q constant d’une cadence à 4 voix à l’orgue, avec une précision spectrale au 16^e de ton. Jeux *Bourdon + Flûte 4*. Ambitus de l’analyse : *sol 2 - do 9*. Les fenêtres d’analyse se ferment sur les temps d’analyse.

Décomposition partie harmonique/partie percussive Le modèle *sines + transients + noise* (STN)² décompose le signal audio en une partie périodique, une partie percussive et une partie bruitée. Dans le cas des descripteurs harmoniques, nous nous intéressons aux interactions entre partiels, lesquels sont contenus essentiellement dans la partie périodique.

Les algorithmes de séparation spectrale des composantes harmoniques et

1. Voir page 191.

2. Voir (Levine et Smith, 1998).

percussives sont largement utilisés dans le cadre des MIR, notamment pour la transcription automatique et la détection des accords ou des tonalités¹.

Nous utilisons l'algorithme de Fitzgerald (Fitzgerald, 2010) avec l'amélioration de Drieger (Driedger *et al.*, 2014), implémentés dans Librosa. Cet algorithme utilise des filtres médians sur les parties horizontales et verticales du spectre en amplitude pour détecter les composantes harmoniques et percussives, plus une composante résiduelle qui est la différence entre le spectrogramme et la somme des composantes harmoniques et percussives. Il tire partie du fait que les composantes harmoniques sont relativement constantes sur une bande de fréquence (ligne du spectrogramme), tandis que les composantes percussives sont caractérisées par des pics (colonne du spectrogramme) sur une fenêtre d'analyse donnée. Le principal avantage de ce modèle est la présence d'un coefficient paramétrable qui définit un degré d'harmonicité.

La méthode est la suivante : à partir du spectrogramme en amplitude X obtenu par TQC, on commence par appliquer un filtre médian horizontal pour obtenir un spectre intermédiaire Y_h et un filtre médian vertical pour obtenir un spectre intermédiaire Y_p :

$$\begin{aligned} Y_h(t, k) &= \text{median}(X(t-l, k), \dots, X(t+l, k)) \\ Y_p(t, k) &= \text{median}(X(t, k-l), \dots, X(t, k+l)) \end{aligned} \quad (5.9)$$

pour tout couple (t, k) de coordonnées temporelles et spectrales, où l détermine la longueur du filtre, fixé sauf mention contraire à $l = 15$.

De là, le masque binaire $\mathcal{M}_h$ est défini de la manière suivante :

$$\mathcal{M}_h = \frac{Y_h(t, k)}{Y_p(t, k)} > \beta \quad (5.10)$$

où $\beta \geq 1$ est un coefficient déterminant le degré d'harmonicité.

La composante harmonique X_h du spectre est alors obtenue en multipliant coordonnée par coordonnée le spectre par le masque :

$$X_h(t, k) = \mathcal{M}_h(t, k)X(t, k) \quad (5.11)$$

La différence entre le spectrogramme initial et sa composante harmonique correspond à la somme des composantes percussive et résiduelle (bruit). Il est

1. Voir (Müller, 2015).

possible d'isoler la composante percussive de manière analogue à la composante harmonique, en introduisant un masque percussif $\mathcal{M}_p$, mais nous n'en aurons pas usage dans la suite.

C'est le spectre de la composante harmonique que nous utilisons pour le calcul des descripteurs harmoniques. Sauf mention contraire, nous fixons $\beta = 20$ dans le calcul de la composante harmonique.

La figure 5.16 représente la séparation en partie harmonique et partie percussive + bruit du spectrogramme de la figure 5.15.

On observe toujours dans les fréquences graves un débordement sur la verticalité suivante, dû à la grande largeur des fenêtres d'analyse sur ces fréquences.

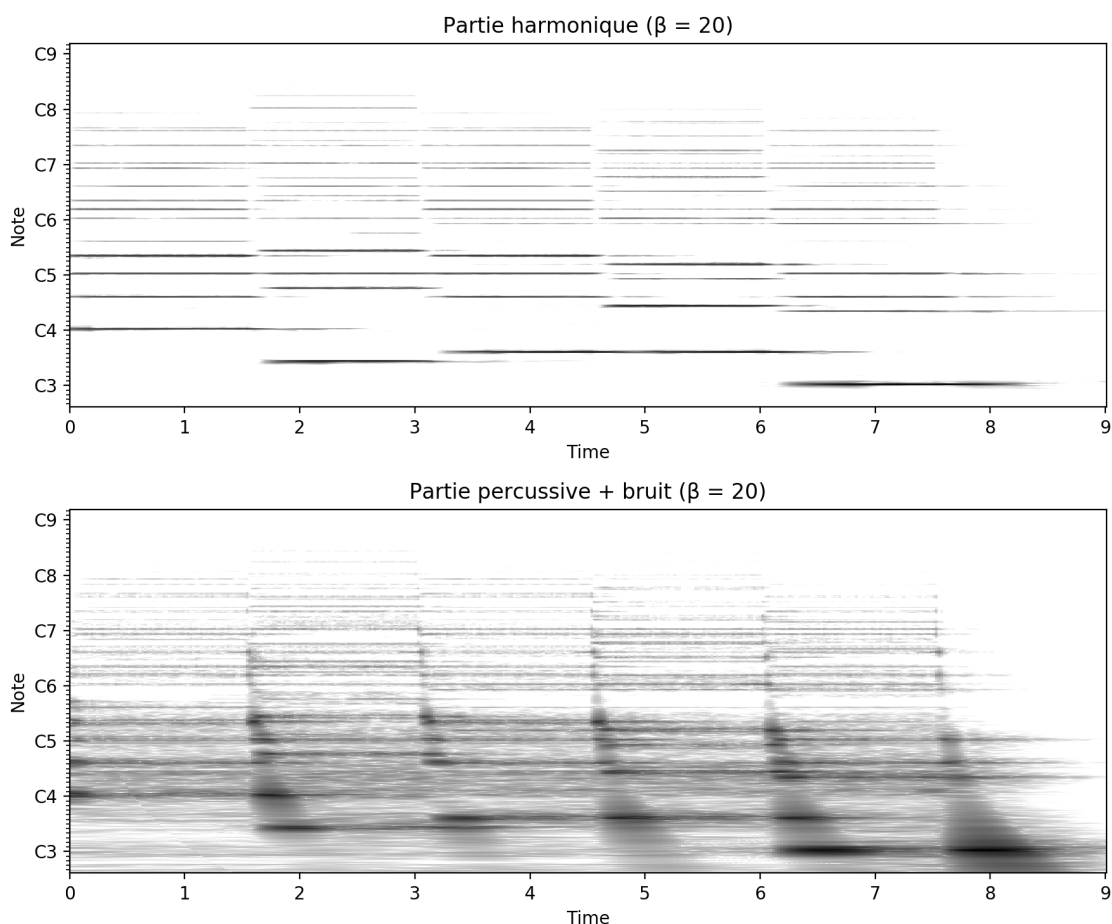


FIGURE 5.16 – Séparation du spectre de la figure 5.15 en partie harmonique (en haut) et partie percussive + composante résiduelle (en bas), avec un coefficient d’harmonicité $\beta = 20$.

5.3.2 Synchronisation du spectre

Une partie des difficultés d’application des descripteurs audio dans le cadre de l’analyse musicale tient, nous l’avons vu, à l’absence de synchronisation des représentations avec les notes. Sur la figure 1.2 page 40, la dissonance est calculée sur des fenêtres d’analyse de 200 ms, et la fluctuation de la courbe rend difficile sinon impossible la correspondance avec les notes de la partition.

Pour résoudre ce problème, nous proposons de calculer les descripteurs harmoniques sur chaque verticalité, une verticalité étant délimitée temporellement par l'apparition ou la disparition d'une note. Comme les descripteurs partent des spectrogrammes, il est nécessaire de les homogénéiser sur la durée de chaque verticalité. Cela implique de connaître tous les *onsets* (début des notes) et *offsets* (fin des notes), ce qui peut se faire de trois manières différentes :

- Informations préalablement rentrées ;
- Détection automatique ;
- Détection manuelle, à partir d'une enveloppe spectrale ou d'un spectrogramme.

Une grande partie des exemples donnés dans la suite sont synthétisés à partir de sons d'orgue échantillonnés, par l'intégration d'un *soundfound* dans l'éditeur de partition Musecore. Dans ce cas, les séparations entre verticalités, ou *onsets* et *offsets* pour reprendre la terminologie du MIDI, sont extraits de la synthèse et pré-enregistrés pour la synchronisation des spectres. Pour les enregistrements, on utilise une annotation manuelle des *onsets* à partir d'un spectrogramme¹, mais une détection automatique est possible.

Plusieurs méthodes de détection automatique des *onsets* existent, parmi lesquelles de méthodes fondées sur les variation énergétiques (Dixon, 2001), sur les variations de phase (Duxbury *et al.*, 2003) ou sur les variations harmoniques (Harte *et al.*, 2006).

Médiane des spectres Une fois les *onsets* détectés, les verticalités correspondent à l'intervalle entre deux *onsets* successifs. Il reste alors à homogénéiser le spectre sur chaque verticalité. Nous travaillons pour cela sur chaque plage de fréquence de l'analyse (ligne du spectrogramme) en prenant la médiane des valeurs situées entre les *onsets* délimitant la verticalité. En l'assignant à la verticalité et en répétant l'opération sur chaque plage de fréquence, on obtient un spectre médian sur chaque verticalité.

Le choix de la médiane plutôt que de la moyenne est justifié par la volonté de minimiser l'importance des valeurs extrêmes sur chaque bande de fréquence.

La figure 5.17 représente le spectre de la figure 5.16 (composante harmo-

1. Cette annotation est réalisée dans *Sonic Visualiser*, et les instants sont ensuite exportés sous forme de fichier texte.

nique) synchronisé sur les verticalités, les *onsets* étant extraits directement de la synthèse sonore.

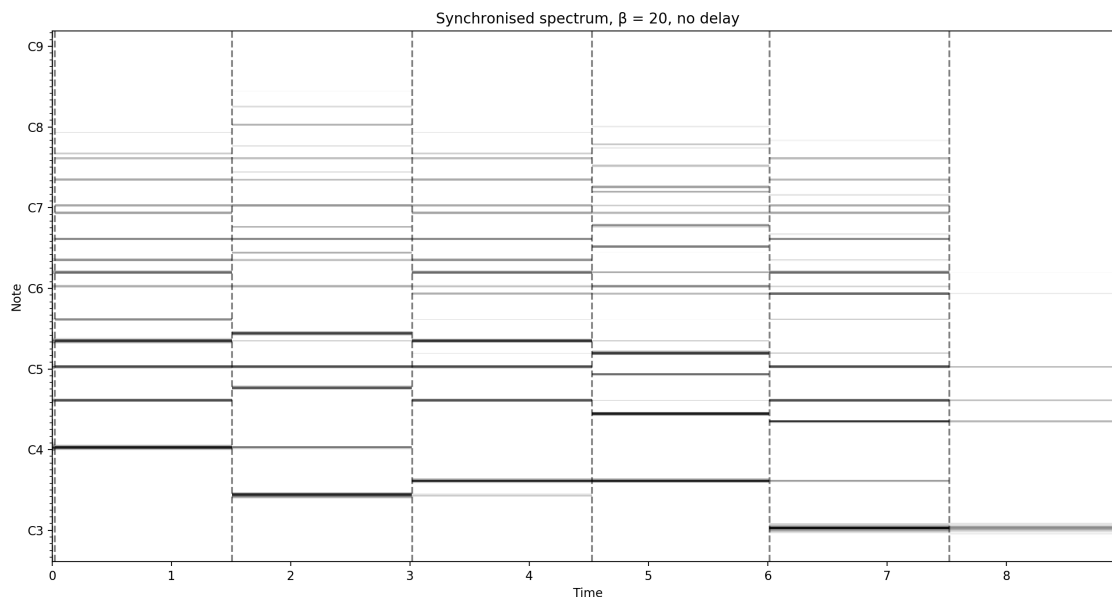


FIGURE 5.17 – Synchronisation sur les verticalités du spectre de la figure 5.16 par prise de la médiane sur chaque plage de fréquence. On observe toujours dans les fréquences graves un débordement sur la verticalité suivante.

Cette manière de procéder a des désavantages qui sont facilement visibles sur la figure précédente, où certaines composantes spectrales d’une verticalité débordent sur la verticalité suivante, alors qu’ils ne le devraient pas. C’est le cas du *do* 4 de la verticalité 1, qui débord sur la verticalité 2, du *fa* 3, ou encore de toutes les composantes de l’avant-dernière verticalité dépassant sur la dernière verticalité, qui correspond au silence entre l’extinction du son et la fin de la piste. Cet effet de débordement est dû à l’élargissement des fenêtres d’analyse dans le grave¹, il ne correspond pas à une réalité sonore.

Pour corriger ce problème, nous ajoutons pour chaque plage de fréquence f_k un retard temporel $\tau(f_k)$ au début de l’intervalle servant à calculer la médiane. Nous prenons ce retard égal à la longueur de la fenêtre d’analyse,

1. Voir page 183 et figure 5.14.

soit

$$\tau(f_k) = \frac{\Delta\omega}{f_k(2^{1/B} - 1)}, \quad (5.12)$$

avec B la précision spectrale, égale sauf mention contraire à 12×8 , et $\Delta\omega = 1.5$. Le tableau page 183 montre la longueur de la fenêtre d'analyse sur quelques fréquences.

La médiane sur la verticalité délimitée par t_{start} et t_{end} pour la plage de fréquences f_k est calculée sur l'intervalle $[t_{start} + \tau(f_k), t_{end}]$. Dans le cas où $t_{end} - t_{start} < \tau(f_k) + \delta$, avec δ marge temporelle égale à 0.1 s, la médiane est calculée sur l'intervalle $[t_{end} - \delta, t_{end}]$, avec un retard plus petit que $\tau(f_k)$.

La figure 5.18 montre le spectre de la figure 5.16 synchronisé par cette méthode. Les débordements observés sur la figure 5.17 sont pour la plupart supprimés.

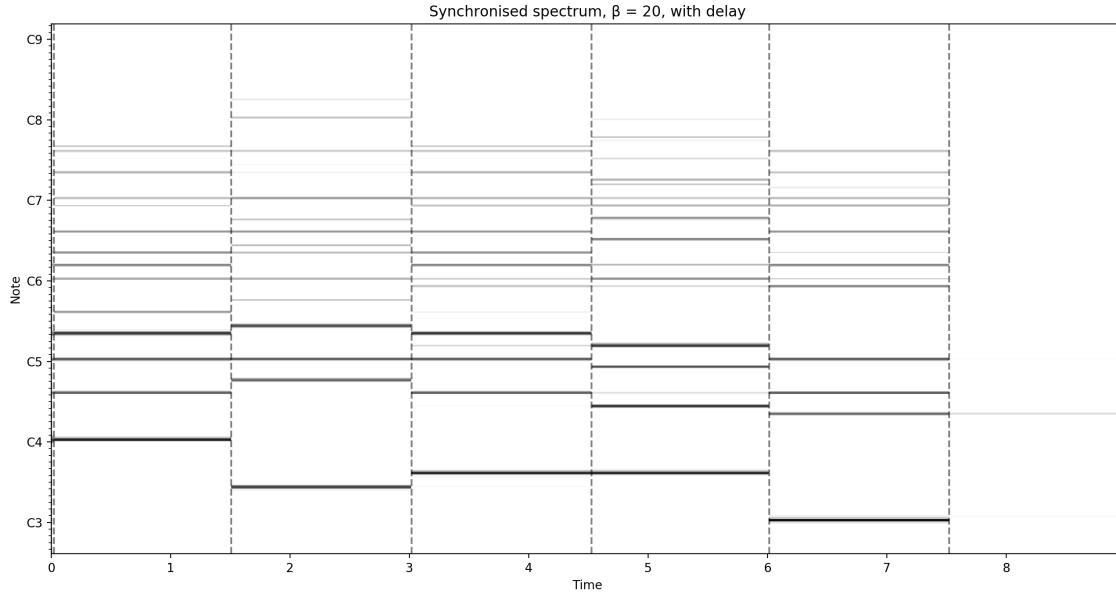


FIGURE 5.18 – Synchronisation sur les verticalités du spectre de la figure 5.16. La médiane est calculée sur chaque *onset* avec un retard de $\tau(f_k)$ dépendant de la plage de fréquence, ce qui diminue le débordement des fréquences sur la verticalité suivante.

L'avantage de l'ajout d'un retard dans le calcul des spectres synchronisés au début de chaque verticalité est double : en plus de réduire drastiquement le

débordement des fréquences graves sur les verticalités, il permet dans certains cas d'éviter pour certaines fréquences la prise en compte des phases d'attaque et de chute ¹. Il minimise donc la contribution de ces phases de variation de l'enveloppe pour favoriser la partie stable, l'entretien.

Limites de la méthode et contraintes sur les sons utilisés La construction d'un spectre médian sur chaque verticalité sous-entend une certaine stabilité temporelle du contenu spectral. Une variation spectrale des sons du type vibrato aura pour effet d'épaissir chaque composante spectrale, comme c'est le cas sur la figure 5.19 où des sons échantillonnés de cordes frottées avec vibrato ont été utilisés.

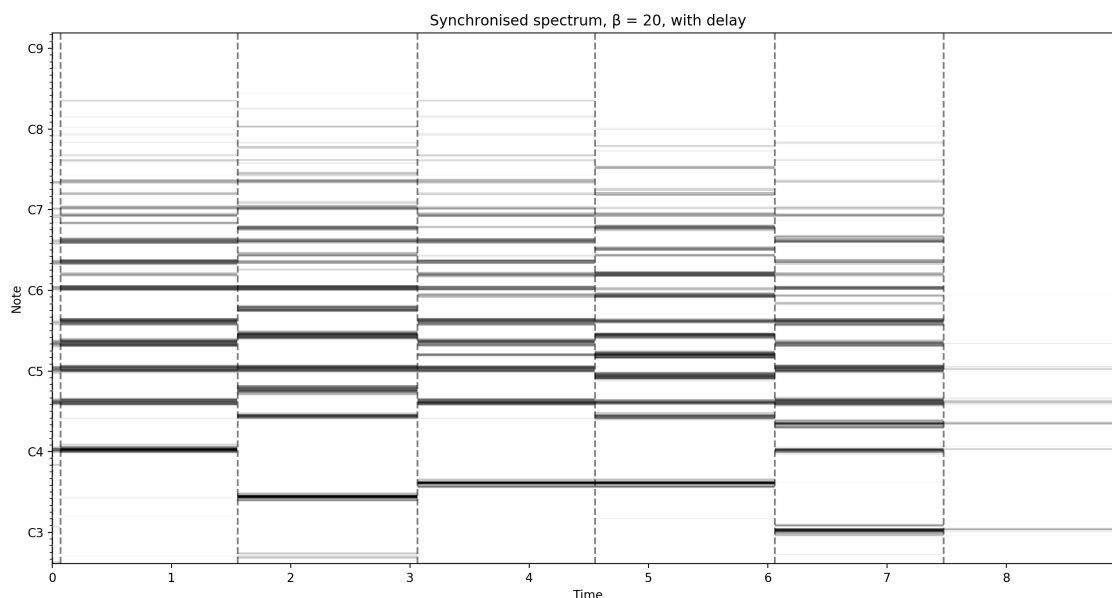


FIGURE 5.19 – Spectre synchronisé à partir d'une synthèse avec des cordes frottées de la cadence figure 5.13. Le vibrato crée des partiels plus épais qu'avec des sons d'orgue (figure 5.18), du fait de la prise de la médiane sur chaque verticalité. La largeur des partiels est de l'ordre du demi-ton.

Avec des sons présentant une variation spectrale comme le vibrato, la synchronisation sur les verticalités fait perdre l'information temporelle sur ces variations fines, d'où une forte imprécision spectrale. Cette dernière n'est

1. Voir page 156.

pas nécessairement problématique puisqu'elle peut correspondre à une réalité perceptive, et est directement liée au paramètre σ de l'approche temporelle¹. Aussi, comme nous le verrons bientôt, cette imprécision spectrale joue-t-elle un rôle crucial dans le calcul de la concordance et des descripteurs dérivés².

Cependant, la prise en compte du vibrato nécessite que la période du vibrato soit d'un ordre grandeur inférieur à la durée des verticalités, sous peine d'incohérences pouvant mettre à mal l'interprétation. Dans la suite du travail, afin d'éviter les écueils, nous privilégions les sons stables du point de vue spectral. C'est l'une des raisons pour lesquelles nous utilisons des sons d'orgue.

5.3.3 Détection des partiels

Malgré l'utilisation de sons relativement stables, les partiels ont invariablement une certaine largeur spectrale, qui peut s'étaler sur plusieurs bandes de fréquences, surtout lorsque la précision spectrale est importante, comme le montre la figure 5.20.

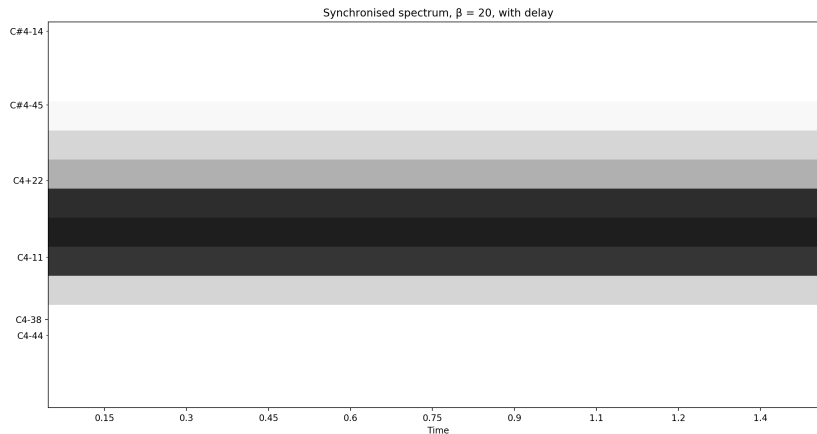


FIGURE 5.20 – Zoom sur le premier partiel de la première verticalité de la figure 5.18, centré sur *do* 4. L'analyse étant en 16^e de tons, le partiel s'étale sur près d'un demi-ton.

Cet étalement des partiels, s'il est crucial dans le calcul de certains des-

-
1. Voir page 170 pour les interprétations de σ .
 2. Voir figure 6.9 page 212.

cripteurs harmoniques comme la concordance, a une influence négligeable sur d'autres descripteurs comme la rugosité ou la tension, à condition de rester suffisamment inférieur à l'écartement minimal des partiels — condition vérifiée sur les sons d'orgue utilisés¹. Ainsi, il est possible de calculer ces derniers sur des spectres en pics, où toute l'énergie d'un partiel a été concentrée en une seule plage de fréquence. Cette manipulation a l'avantage d'accélérer les calculs et de rendre les spectres plus lisibles. Puisqu'elle est réalisée sur les spectres en amplitude, elle ne conserve en revanche pas l'énergie, calculée à partir du spectre étalé.

Dans la suite, on considère qu'une composante spectrale constitue un pic lorsque sa valeur majore un voisinage spectral fixé. Pour construire le spectre en pics X_{pics} à partir du spectre synchronisé X , on fait glisser sur chaque verticalité une fenêtre spectrale de largeur 2δ , et lorsque la composante spectrale centrale majore les autres composantes spectrales sur la fenêtre, on lui assigne la somme des composantes spectrales sur la fenêtre, sinon on lui assigne la valeur 0. Deux partiels suffisamment proches ne sont ainsi pas distingués.

Pour la bande de fréquence k et la verticalité n , on a :

$$X_{pics}(k, n) = \begin{cases} \sum_{l=k-\delta}^{k+\delta} X(l, n) & \text{si } \forall l \in [k - \delta, k + \delta], X(k, n) \geq X(l, n) \\ 0 & \text{sinon,} \end{cases} \quad (5.13)$$

où δ est un entier déterminant l'espacement minimal possible entre deux partiels. Dans la suite nous fixons, sauf mention contraire, $\delta = 4$, ce qui correspond, avec une analyse au 16^e de ton, à un espacement minimal au quart-de-ton.

La figure 5.21 présente le spectre en pics obtenu à partir du spectre de la figure 5.18.

1. Voir figure 6.6

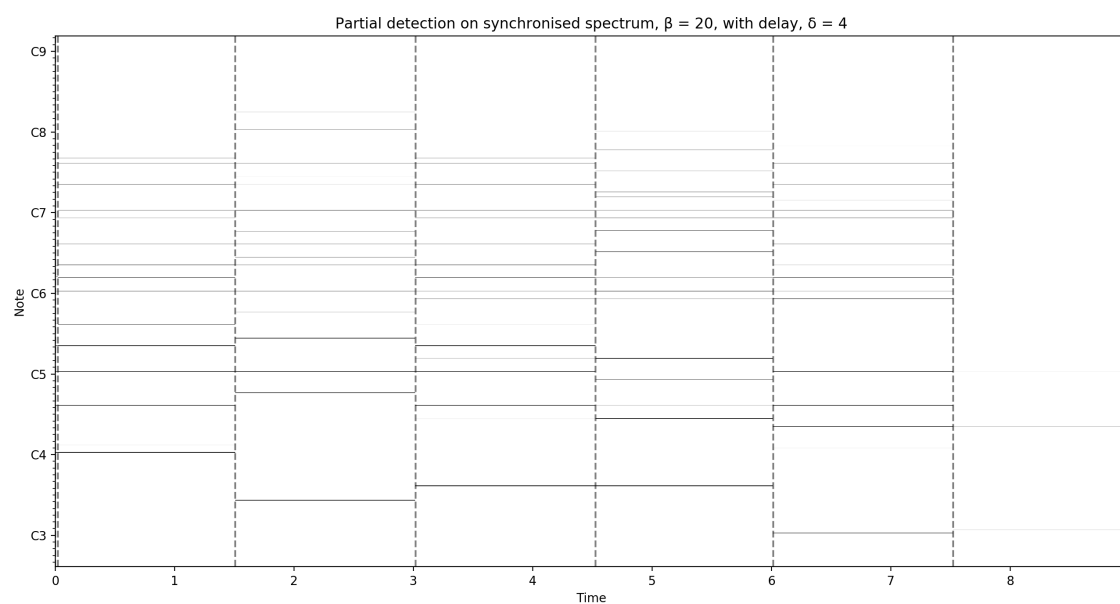


FIGURE 5.21 – Détection des partiels à partir du spectre de la figure 5.18, avec $\delta = 4$, soit un espacement minimal des partiels d'un quart-de-ton.

Chapitre 6

Présentation des descripteurs harmoniques

Nous présentons dans ce chapitre les descripteurs harmoniques utilisés dans la thèse, en détaillant leur implémentation. Le choix des descripteurs est motivé par leur potentiel musicologique ainsi que par leur pertinence acoustique ou perceptive. Tout d’abord, c’est une volonté délibérée de nous restreindre à des descripteurs de bas niveau qui impliquent des calculs simples. Nous nous limitons d’ailleurs à des descripteurs spectraux, calculables à partir d’un spectrogramme en amplitude, sans prendre en compte des descripteurs de forme d’onde ou de phase.

La concordance, qui consiste en un produit des spectres, est motivée par son interprétation physique comme énergie d’interaction. La rugosité et l’harmonicité sont quant à eux choisis pour leur rôle crucial dans la perception de la consonance et de la dissonance², et la concordance différentielle comme l’un des rares descripteurs harmoniques dynamiques de la littérature³. La tension est choisie pour sa simplicité, sa corrélation avec la perception, et comme exemple de « *psychophysique à 3 notes* »⁴. La concordance d’ordre m fixé est choisie dans l’optique de mettre en avant des propriétés acoustiques propres aux accords d’au moins m notes, et la concordance totale dans celle de comparer des accords de toutes les cardinalités. Le changement harmonique est choisi pour son usage en traitement du signal, et la rugosité différentielle

2. Voir partie 3.2.

3. La concordance différentielle correspond à la notion d’affinité chez Milne (Milne *et al.*, 2016), voir partie 6.2.2.

4. Voir (Cook, 2009).

comme extrapolation dynamique d'un modèle de perception statique.

Par ce choix, nous ne visons aucunement l'exhaustivité, qui pourrait être justifiée dans le cadre d'une comparaison des modèles, ce qui n'est pas l'objet de ce travail.

6.1 Les descripteurs statiques

6.1.1 Rugosité

La rugosité d'un son est la sensation désagréable due au battements entre partiels proches. Quand deux sons simples (sinusoïdaux) sont joués simultanément, ils interagissent pour former des battements de fréquence égale à la différence des fréquences des sons simples. Lorsque les fréquences sont suffisamment proches, les battements sont clairement audibles, mais lorsque qu'elles s'éloignent, ils deviennent trop rapides pour rester audibles et créent une sensation désagréable, qu'on appelle rugosité. En s'éloignant encore, les fréquences sont perçues comme distinctes, et la sensation de rugosité disparaît.

Histoire de la rugosité La notion de battements est formalisée par Sauveur ([Sauveur, 1700](#)), qui voit déjà en eux la cause de la dissonance¹. Helmholtz ([Helmholtz, 1863](#)) les étudie expérimentalement, et les quantifie à l'aune de la décomposition de Fourier, ce qui l'amène à représenter les courbes de rugosité², qui viennent révolutionner la théorie de la consonance et de la dissonance³. La rugosité, cette « *impression dure et désagréable* »⁴ à l'oreille, est due aux battements entre harmoniques et entre sons résultants. Helmholtz trouve que la rugosité entre deux partiels est maximale pour des battements de 33 Hz, indépendamment du registre, et généralise ce résultat aux sons complexes en sommant les contributions de tous les couples de partiels qui interagissent.

Un siècle plus tard, à partir d'expériences perceptives, Plomp et Levelt ([Plomp et Levelt, 1965](#)) relie la rugosité à la notion de bande critique et

1. Voir citation de la note 1 page [108](#).

2. Voir les courbes de rugosité d'Helmholtz page [45](#)

3. Voir partie [2.3](#).

4. Tiré de ([Helmholtz, 1863](#), p.290)

proposent une courbe de dissonance des sons simples où le maximum de dissonance correspond à une différence des fréquences égale à un quart de la bande critique¹. La bande critique est une notion introduite quelques années auparavant ([Zwicker et al., 1957](#)), construite sur des résultats perceptifs concernant les effets de masquage, d'intensité et de sensibilité à la différence de phase entre deux sons sinusoïdaux. Elle donne un ordre de grandeur de l'écart fréquentiel en-dessous duquel deux sons simples sont assez proches pour interagir.

Les résultats de Plomp et Levelt marquent le début d'une longue série de modèles raffinant le celui d'Helmholtz. Les sons résultants ne sont plus d'actualité, seuls les battements entre partiels sont considérés. Parmi les plus importants, citons celui de Kameoka et Kuriyagawa ([Kameoka et Kuriyagawa, 1969a,b](#)), qui tient compte de la dépendance aux amplitudes respectives des composantes spectrales, et celui de Hutchinson et Knopoff ([Hutchinson et Knopoff, 1979](#)), qui se démarque en prenant en compte les interactions internes à un son, avec un raffinement du calcul des bandes critiques.

À côté de ces modèles partant de la décomposition spectrale de l'énergie, largement étudiés, d'autres modèles de rugosité s'occupent de l'aspect temporel, notamment des fluctuations d'amplitude ([Terhardt, 1974](#)) ou des différences de phase entre les composantes spectrales ([Pressnitzer et McAdams, 1999](#)). Chez Terhardt, la rugosité est déterminée par les fluctuations de l'enveloppe du signal à travers un filtre auditif. Certains modèles combinent les approches spectrales et temporelles ([Vassilakis, 2001](#)).

1. Voir figure [6.1](#)

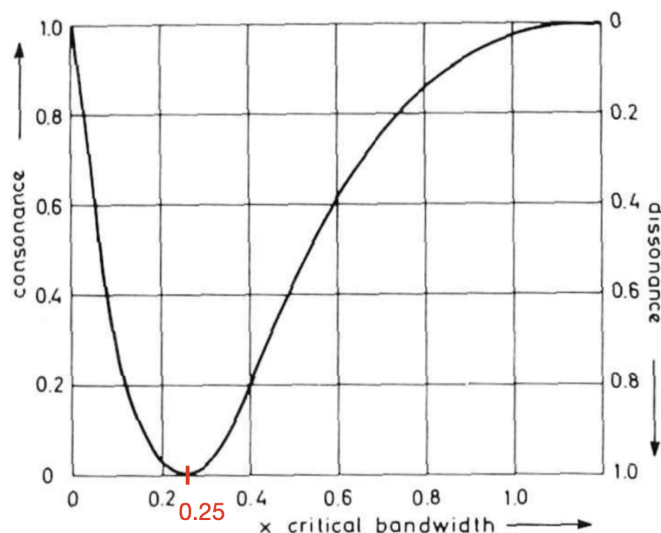


FIGURE 6.1 – Courbe de rugosité de deux sons simples, tirée de (Plomp et Levelt, 1965). Le maximum de dissonance, indiqué ici en rouge par l’auteur, correspond à une différence des fréquences égale à un quart de la bande critique.

Rugosité vs. dissonance sensorielle La rugosité est l’objet d’une confusion terminologique, puisqu’elle est souvent assimilée, à tort, à la dissonance sensorielle, dont elle n’est qu’un aspect parmi d’autres. Pour Terhardt (Terhardt, 1984), la dissonance sensorielle inclut, outre la rugosité, trois autres facteurs : *sharpness*, *loudness* et *toneness* (tendance d’un son à être perçu comme une note plutôt que comme un bruit), tandis que pour Harrison (Harrison et Pearce, 2020) et Parncutt (Parncutt et al., 2019) elle inclut, en plus de la rugosité, l’harmonicité et la familiarité. Cette assimilation des concepts, due en grande partie à Helmholtz (Helmholtz, 1863), s’est perpétuée tout au long du siècle et demi de recherches sur la rugosité (Plomp et Levelt, 1965; Kameoka et Kuriyagawa, 1969a,b; Hutchinson et Knopoff, 1979), le statut ambigu des notions de consonance et dissonance¹ n’aidant pas à éclaircir ce qui est devenu, au fur et à mesure des avancées, une confusion terminologique, laquelle est toujours présente dans la littérature (Mashinter, 2006), quoique pointée du doigt (Parncutt, 2006). Pour éviter toute confusion, nous

1. Arthurs (Arthurs et Timmers, 2016) emploie l’expression « *fluidité de la consonance et de la dissonance* ».

employons ici le terme de « rugosité » sans se référer à la notion de « dissonance ».

Implémentation Nous avons vu que la notion de rugosité, d'ordre perceptif, donne lieu à des interprétations différentes et à de nombreux modèles acoustiques. Le propos n'est pas ici de choisir le modèle le mieux corrélé avec les résultats perceptifs, mais d'avoir un modèle suffisamment représentatif pour pouvoir étudier les applications musicales de la rugosité ainsi que la manière dont elle structure les espaces harmoniques.

Nous utilisons un modèle largement inspiré de Sethares ([Sethares, 1999](#)). La courbe de rugosité de la figure 6.1 (renversée) est approchée par une différence d'exponentielles décroissantes de la forme

$$r(x) = e^{-b_1 x} - e^{-b_2 x}, \quad (6.1)$$

avec b_1 et b_2 paramètres contrôlant l'emplacement du maximum et la décroissance après le maximum. Les valeurs de b_1 et b_2 sont obtenues par Sethares par minimisation de l'erreur avec les données expérimentales de plusieurs expériences¹ : $b_1 = 3.5$ et $b_2 = 5.75$.

Étant donnée la dépendance de la bande critique à la hauteur, x ne s'exprime pas comme la simple différence des fréquences, mais comme la différence pondérée par un terme exprimant la longueur de la bande critique, qui dépend de la fréquence basse. La rugosité de deux sons simples de fréquences f_1 et f_2 et d'amplitude 1, avec $f_1 \leq f_2$, s'exprime de la manière suivante :

$$\begin{aligned} \mathcal{R}(f_1, f_2) &= r((f_2 - f_1)s(f_1)) \\ &= e^{-b_1(f_2 - f_1)s(f_1)} - e^{-b_2(f_2 - f_1)s(f_1)}, \end{aligned} \quad (6.2)$$

avec

$$s(f_1) = \frac{x^*}{s_1 f_1 + s_2}, \quad (6.3)$$

où $x^* = 0.24$, repérable sur la figure 6.1, correspond à la part de la bande critique au maximum de dissonance, et $s_1 = 0.021$ et $s_2 = 19$ décrivent la variation de la bande critique en fonction de la fréquence inférieure. Ils sont

1. Celles décrites dans ([Plomp et Levelt, 1965](#); [Kameoka et Kuriyagawa, 1969a,b](#)).

obtenus par Sethares par minimisation de l'erreur avec les données expérimentales.

La figure 6.2 montre l'évolution de la courbe de rugosité de deux sons simples en fonction de la hauteur du son grave. L'intervalle de rugosité maximale décroît avec la fréquence du son supérieur. Pour le *do* 4 médium, ce maximum se situe entre le demi-ton et le ton, tandis qu'à l'octave d'en-dessous il se situe au niveau de la tierce mineure, et pour l'octave d'au-dessus au niveau du demi-ton.

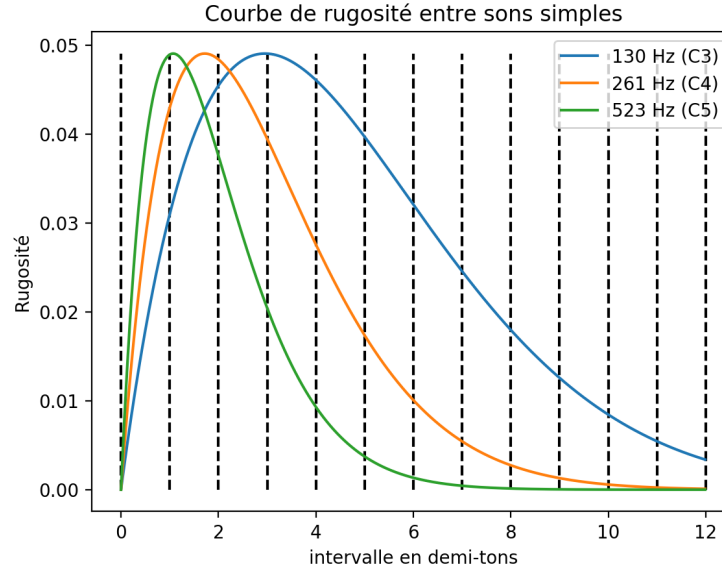


FIGURE 6.2 – Courbes de rugosité entre deux sons simples, pour trois valeurs de fréquence basse. L'intervalle de rugosité maximale diminue avec la fréquence.

Dans notre implémentation, si les fréquences f_1 et f_2 ont pour amplitudes respectives α_1 et α_2 , $\mathcal{R}(f_1, f_2)$ contribue à la rugosité totale avec un coefficient égal au produit des amplitudes, autrement dit :

$$\mathcal{R}(\alpha_1, \alpha_2, f_1, f_2) = \alpha_1 \alpha_2 \mathcal{R}(f_1, f_2) \quad (6.4)$$

Des coefficients plus fins que le produit des amplitudes ont été proposés, notamment par Sethares ([Sethares, 1999](#)), qui prend le minimum des intensités en *sonies*, ou encore Vassilakis ([Vassilakis, 2001](#)), qui tient compte des fluctuations d'amplitude.

Pour les sons complexes, les contributions des couples de partiels sont sommées.

Le calcul de la rugosité sur les verticalités se fait à partir du spectre synchronisé étalé ou en pics des pistes séparées. Si N désigne le nombre de notes, et $X_{k,t}^{(i)}$ la valeur du spectre en amplitude de la piste i sur la bande de fréquence k centrée en f_k à la verticalité t , alors la rugosité $\mathcal{R}(t)$ sur la verticalité t s'exprime de la manière suivante :

$$\mathcal{R}(t) = \sum_{i < j \leq N} \sum_{k,l} \mathcal{R}(X_{k,t}^{(i)}, X_{l,t}^{(j)}, f_k, f_l) \quad (6.5)$$

Nous illustrons la rugosité avec un exemple d'enchaînement de trois accords en position serrée : *Do m*, *Do M* et *Do dim*, synthétisés avec les jeux (*Bourdon + Flûte 4*). Le choix de prendre des accords du langage tonal est uniquement didactique, car ce sont sur les accords familiers que l'intuition des descripteurs harmoniques peut se forger le plus facilement, mais le champ des descripteurs harmoniques dépasse le langage tonal.

Le profil de rugosité des trois accords est représenté sur la figure 6.3, en même temps que la répartition spectrale des contributions à la rugosité. Cette dernière est construite de la manière suivante : la rugosité entre deux partiels est répartie équitablement sur les deux fréquences des partiels. Elle permet de localiser les partiels qui par leur interaction apportent la plus grande contribution à la rugosité.

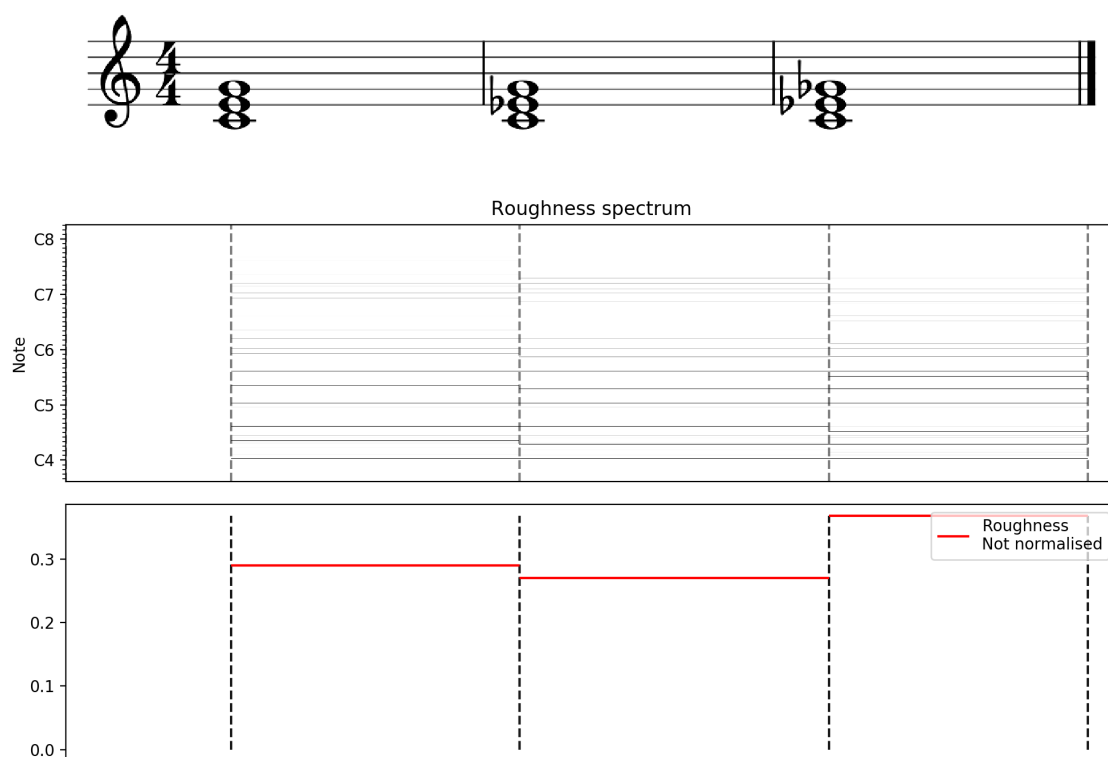


FIGURE 6.3 – Évolution de la rugosité calculée, sur 3 accords synthétisés avec les jeux *Bourdon + Flûte4* : *Do M*, *Do m* et *Do dim*. En haut, la répartition spectrale des contributions à la rugosité, en bas, les valeurs de rugosité. Les spectres en pics des pistes séparées ont été utilisés.

L'accord diminué a la rugosité la plus importante. La répartition spectrale permet de localiser la principale contribution, qui découle de l'interaction entre *solb* 5 et *sol#* 5, c'est à dire entre le 3^e harmonique de *do* 4 et le 2^e harmonique de *solb* 4, comme montré sur la figure 6.4.



FIGURE 6.4 – Accord diminué : interaction des partiels. En gris : premiers harmoniques. En rouge : les harmoniques dont l’interaction constitue la plus grande contribution à la rugosité.

Lorsque le spectre des pistes séparées n’est pas disponible¹, on peut tout de même calculer la rugosité à partir du spectre synchronisé en pics du signal global. Le modèle reprend alors l’idée de (Hutchinson et Knopoff, 1979) de prendre en compte les interactions internes à une note. Le calcul est le suivant :

$$\mathcal{R}(t) = \sum_{k < l} \mathcal{R}(X_{k,t}, X_{l,t}, f_k, f_l) \quad (6.6)$$

où $X_{k,t}$ est la valeur du spectre en pics sur la bande de fréquence k centrée en f_k et la verticalité t . Sur le même exemple que précédemment, on obtient le profil suivant :

1. Par exemple dans la configuration enregistrement (voir page 163). Dans la configuration mélodie-accompagnement, on calcule la dissonance entre la mélodie et l’accompagnement.

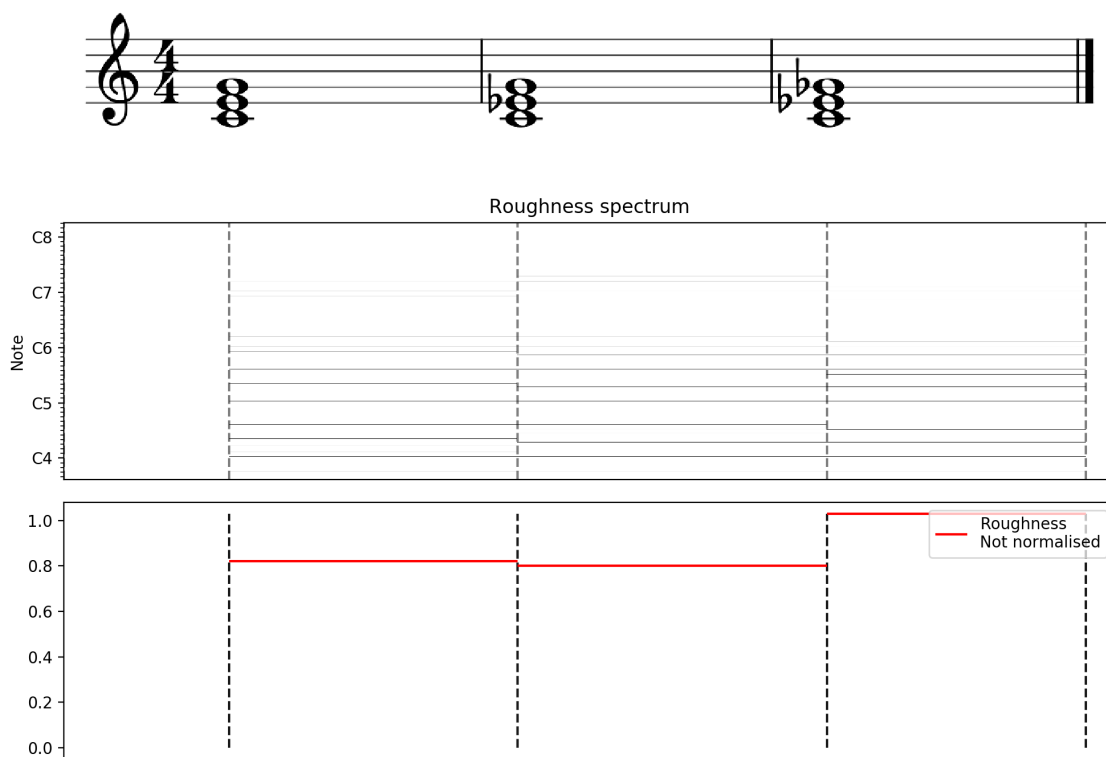


FIGURE 6.5 – Évolution de la rugosité calculée à partir du spectre global, sur 3 accords synthétisés avec les jeux *Bourdon + Flûte4* : *Do M*, *Do m* et *Do dim*. Les spectres en pics ont été utilisés.

Malgré quelques interactions non prises en compte, le profil de rugosité est similaire au cas où elle est calculée à partir des spectres en pics des pistes séparées.

Les modèles de rugosité cités sont construits pour des spectres discrets, c'est-à-dire des spectres en pics. Il est possible également de calculer la rugosité sans détection préalable des pics, en partant directement du spectre synchronisé, car le modèle est généralisable sans effort aux spectres continus, en suivant les formules (6.5) et (6.6).

En effet, si f_1 et $f_1 + \delta$ sont deux fréquences proches d'un même partiel dans un spectre en amplitude X , avec $X(f_1) > X(f_1 + \delta)$, et f_2 la fréquence

d'un autre partiel, alors le calcul de la rugosité sur le spectre étalé donne :

$$\begin{aligned}
 \mathcal{R}_{cont} &= X(f_1)X(f_2)\mathcal{R}(f_1, f_2) + X(f_1 + \delta)X(f_2)\mathcal{R}(f_1 + \delta, f_2) \\
 &\approx (X_1 + X(f_1 + \delta))X(f_2)\mathcal{R}(f_1, f_2) \\
 &= \mathcal{R}_{pics}
 \end{aligned}
 \tag{6.7}$$

sous l'hypothèse $\delta \ll f_1$.

La figure 6.6 montre que les deux approches aboutissent au même résultat.

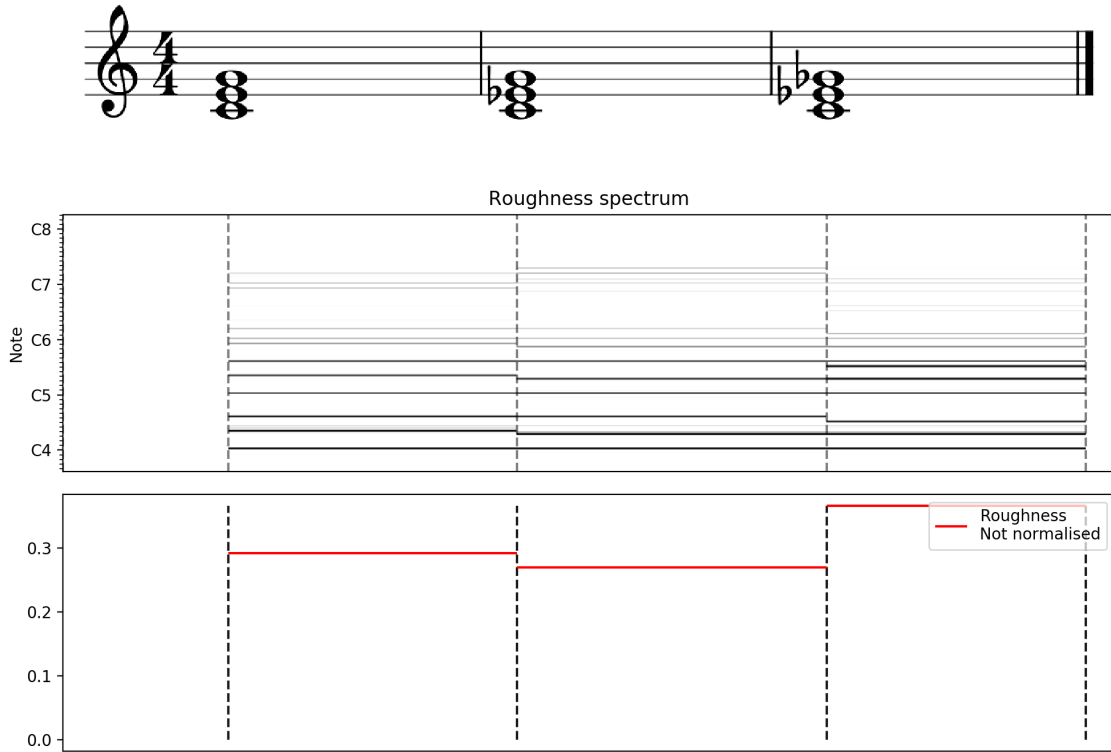


FIGURE 6.6 – Évolution de la rugosité calculée à partir des spectres non ponctuels des pistes séparées, sur 3 accords synthétisés avec les jeux *Bourdon + Flûte4* : *Do M*, *Do m* et *Do dim*.

6.1.2 Concordance

Origine et histoire de la notion La concordance est introduite ici comme mesure de la coïncidence des partiels de deux sons. Nous avons vu que la première référence à la coïncidence des harmoniques est postérieure d'un demi-siècle à leur découverte par Sauveur. Elle est due à Estève¹, qui caractérise de cette manière les consonances, en avance d'un siècle sur Helmholtz.

Helmholtz, à son tour, définit la concordance d'une dyade par la coïncidence des partiels des deux sons², bien que cette approche soit relayée au second plan dans la *Théorie physiologique de la musique* devant la théorie de la rugosité cause de la dissonance. La figure 6.7 montre le classement par Helmholtz des consonances des intervalles justes selon le rang des premiers harmoniques qui coïncident. Le rang de ces premiers harmoniques est en relation directe avec le rapport des fréquences. Ainsi, pour la tierce majeure de rapport $5/4$, le cinquième harmonique de la note inférieure coïncide avec le quatrième harmonique de la note supérieure. De part cette relation, une théorie de la consonance fondée sur le rang de la première coïncidence des partiels fait retomber, dans le cas de sons harmoniques, dans le règne des rapports simples qui domine de Pythagore jusqu'à Euler, en passant par la théorie des coups³. À ce titre, la définition positive d'Helmholtz de la consonance par la première coïncidence des partiels est réactionnaire, tout du moins dans ses conclusions, par rapport à sa définition positive de la dissonance par les battements. De plus, il ne la généralise pas à des accords de trois notes ou plus, alors qu'il généralise à ces accords le calcul de la dissonance par les battements.

Une manière de rendre la coïncidence des partiels plus moderne et plus féconde est de la considérer non pas sous l'angle des rapports de fréquences, mais sous celui de la ressemblance. Citons Lalo :

*De façon générale, sont consonants les sons dont l'un est directement ou indirectement une partie de l'autre. Si la ressemblance est une identité partielle, la consonance est la ressemblance de deux sons*⁴.

1. Nouvelle découverte du principe de l'harmonie avec un examen de ce que M. Rameau a publié sous le titre de *Démonstration de ce principe*, (Estève, 1752). Voir note 2 page 111.

2. Voir page 110.

3. Voir partie 2.1.

4. Tiré de *Esquisse d'une esthétique musicale scientifique* (Lalo, 1908, p.98)

Octave Douzième Quinte Quarte Sixte maj. Tierce maj. Tierce min.
 1 : 2 1 : 3 2 : 3 3 : 4 3 : 5 4 : 5 5 : 6 5

SONS ÉLÉMENTAIRES devant COÏNCIDER.	1	2	3	4	5
6	2 octaves et la quinte.	Douzième. 1 : 3	Octave. 1 : 2	Quarte. 2 : 3	Tierce mineure. 5 : 6
5	2 octaves et la tierce.	Dixième ma- jeure. 2 : 5	Sixte majeure. 3 : 5	Tierce majeure. 4 : 5	
4	Double oc- tave. 1 : 4	Octave. 1 : 2	Quarte. 3 : 4		
3	Douzième. 1 : 3	Quinte. 2 : 3			
2	Octave. 1 : 2				

FIGURE 6.7 – Coïncidence des premiers harmoniques chez Helmholtz. En haut : intervalles (en blanches) et séries harmoniques (en noires) jusqu'à la coïncidence des harmoniques, en intonation juste. En bas : tableau des intervalles classés selon le rang des partiels harmoniques qui coïncident. Helmholtz justifie sa restriction aux six premiers harmoniques en expliquant que le septième harmonique est pratiquement inaudible sur des instruments comme le piano. Tiré de ([Helmholtz, 1863](#), p.233, 235)

La fusion de Stumpf¹ a quelque chose à voir avec l'identité et la ressemblance, même si le psychologue refuse de considérer la ressemblance de deux sons comme cause de leur fusion². Une autre manière d'approcher la ressemblance de deux sons est de quantifier ce qu'ils partagent. On peut le faire de façon rudimentaire en comptant les partiels communs, comme sur la figure 6.8 où les harmoniques communes aux notes des intervalles sont indiqués en bleu. La ressemblance décrite est une ressemblance spectrale.

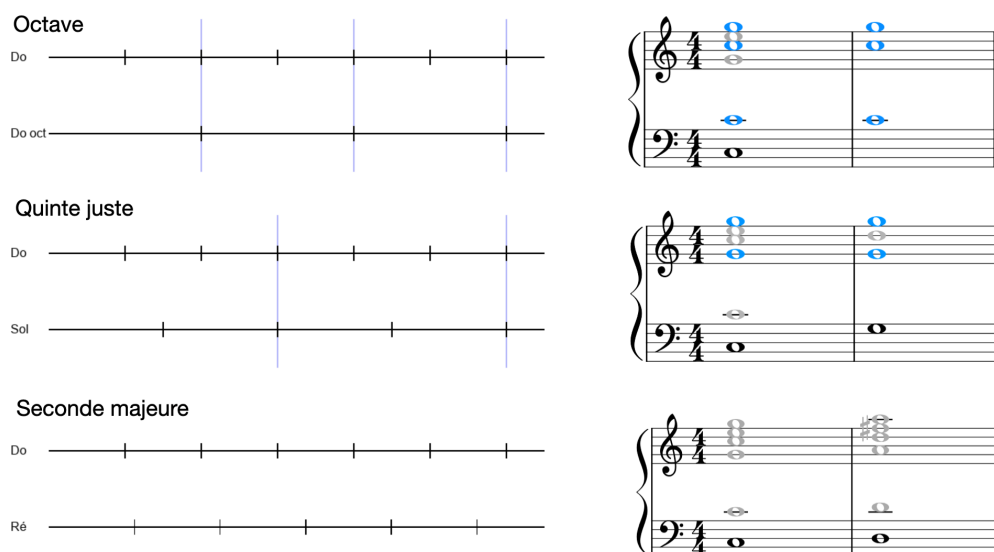


FIGURE 6.8 – Coïncidence des partiels sur les intervalles. Avec des spectres harmoniques, la coïncidence des partiels va de pair avec la simplicité des rapports des fréquences fondamentales. En bleu, les harmoniques communs. Rapports des intervalles justes : octave de rapport $2/1$, quinte juste $3/2$, seconde majeure $9/8$.

Une approche plus physique et plus moderne pour quantifier la ressemblance spectrale consiste à faire le produit des spectres. Nous introduisons la concordance harmonique de deux sons comme le produit de leurs spectres en amplitude. La concordance harmonique quantifie les composantes spectrales communes aux deux spectres.

1. Se référer à (Stumpf, 1883, 1890).
2. Voir page 117.

Dans le formalisme vectoriel des signaux, la concordance harmonique s'identifie avec le produit scalaire canonique dans le domaine temporel, ou le produit hilbertien dans le domaine fréquentiel, dans le cas particulier de spectres réels. Bonnet (Bonnet, 1968, p.4) l'interprète comme une énergie d'interaction entre deux sons.

En effet, en toute généralité, pour deux spectres complexes X_1 et X_2 , on a la relation :

$$\|X^{(1)} + X^{(2)}\|^2 = \|X^{(1)}\|^2 + \|X^{(2)}\|^2 + 2\Re\langle X^{(1)}, X^{(2)}\rangle, \quad (6.8)$$

où $\Re$ désigne la partie réelle.

Les spectres considérés ici sont des spectres en amplitude, il ne devient plus nécessaire de prendre la partie réelle.

La terme « concordance harmonique » a été introduit par Chouvel (Chouvel, 1998), qui l'utilise comme un guide à la composition, essentiellement dans les espaces micro-tonaux.

La concordance harmonique est la version statique de la notion de *pitch similarity* ou d'affinité telle que présente chez Milne (Milne et al., 2016). L'affinité mesure les similarités entres les spectres de notes ou d'accords successifs, et constitue un modèle perceptif de leur bon enchaînement¹. Milne reprend en simplifiant le modèle d'affinité de Terhardt et Parncutt (Terhardt, 1984; Parncutt, 1989), qui prend en compte les hauteurs virtuelles.

Implémentation La concordance est calculée à partir des spectres en amplitude des pistes séparées. La concordance $\mathcal{C}(X^{(1)}, X^{(2)})$ de deux notes de spectres $X^{(1)}$ et $X^{(2)}$ s'exprime comme le produit scalaire des spectres :

$$\begin{aligned} \mathcal{C}(X^{(1)}, X^{(2)}) &= \langle X^{(1)}, X^{(2)} \rangle \\ &= \sum_k X_k^{(1)} X_k^{(2)}. \end{aligned} \quad (6.9)$$

Dans l'approche symbolique, où chaque partiel est modélisé sur le spectre en amplitude par une gaussienne, la concordance de deux partiels peut être

1. "Tonal affinity is the perceived goodness of fit of successive tones", tiré de (Milne et al., 2016).

visualisée comme l'aire de la gaussienne produit¹. De cette manière, la localisation spectrale des contributions à la concordance garde la forme d'un spectre d'énergie en gaussiennes, dont la norme $\mathcal{L}_1$ correspond à la concordance.

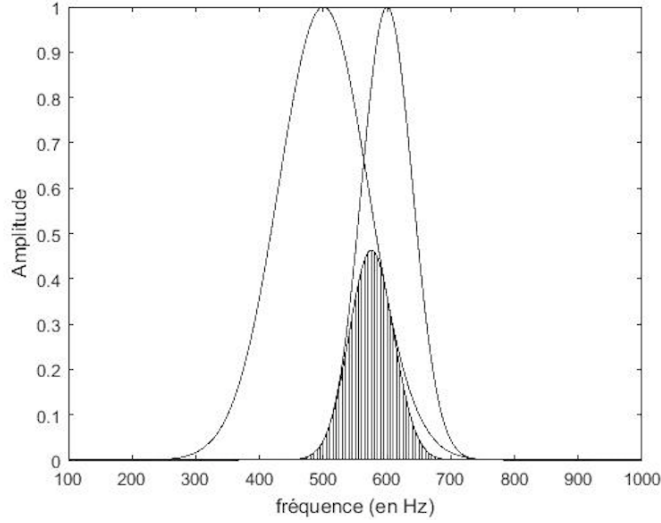


FIGURE 6.9 – Produit de gaussiennes, modélisant la concordance de deux partiels dans l'approche signal, où chaque partiel a la forme d'une gaussienne. L'aire hachurée correspond à la valeur de la concordance.

De là, si N désigne le nombre de notes, la concordance $\mathcal{C}(t)$ de la verticalité t correspond à la somme des concordances de tous les couples de notes :

$$\begin{aligned}\mathcal{C}(t) &= \sum_{i < j \leq N} \mathcal{C}(X_t^{(i)}, X_t^{(j)}) \\ &= \sum_{i < j \leq N} \sum_k X_{k,t}^{(i)} X_{k,t}^{(j)}.\end{aligned}\tag{6.10}$$

Dans l'approche symbolique, comme tous les spectres sont réels, la concordance peut s'exprimer au niveau temporel comme la différence entre l'énergie

1. Le produit de deux gaussiennes de moyenne x_0 et x_1 et d'écart type σ_0 et σ_1 est une gaussienne de moyenne $x = \frac{\sigma_1^2 x_0 + \sigma_0^2 x_1}{\sigma_0^2 + \sigma_1^2}$ et d'écart type $\sigma = \frac{\sigma_0 \sigma_1}{\sqrt{\sigma_0^2 + \sigma_1^2}}$.

globale et la somme des énergies des notes isolées, grâce à la relation (6.8) :

$$\mathcal{E}_t = \mathcal{E}_t^{(1)} + \dots + \mathcal{E}_t^{(N)} + 2 \sum_{i < j \leq N} \mathcal{C}(X_t^{(i)}, X_t^{(j)}),$$

soit

$$\mathcal{C}(t) = \frac{\mathcal{E}_t - \sum_{i=1}^N \mathcal{E}_t^{(i)}}{2}, \quad (6.11)$$

où $\mathcal{E}_t$ désigne l'énergie globale au temps t .

La figure 6.10 montre l'évolution de la concordance sur le même extrait de trois accords synthétisés avec des sons d'orgue.

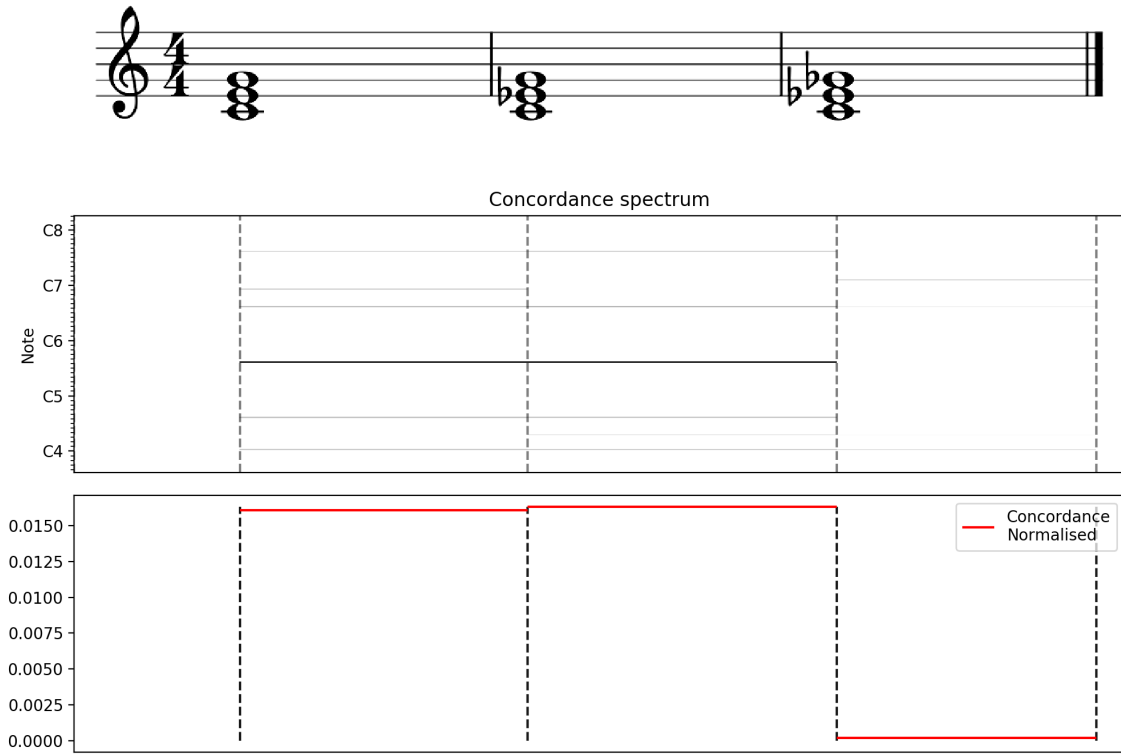


FIGURE 6.10 – Évolution de la concordance sur 3 accords synthétisés avec les jeux *Bourdon + Flûte4* : *Do M*, *Do m* et *Do dim*. En haut, la répartition spectrale des contributions à la concordance, en bas, les valeurs de concordance.

L'accord diminué a une concordance négligeable par rapport aux accords majeurs et mineurs, dont les concordances sont quasiment égales. La répartition spectrale des contributions montre que la contribution la plus importante est due à la quinte juste *do-sol*, puisqu'elle est située sur *sol5*, qui coïncide avec le 3^e harmonique de *do* 4 et le 2^e harmonique de *sol* 4¹. La tierce joue un rôle moindre, en tout cas avec les jeux *Bourdon + Flûte4*.

Contrairement à la rugosité, il n'est pas possible de calculer la concordance à partir du spectre global, elle nécessite d'avoir le spectre de chaque note². De plus, il est crucial de partir des spectres continus (étalés), car avec des spectres en pics seule la coïncidence exacte de partiels apporterait une contribution à la concordance. Avec des spectres continus, deux partiels apportent par contre une contribution du moment qu'ils se chevauchent, sans que les centres spectraux ne soient nécessairement confondus. À ce titre, l'épaisseur des partiels, qui correspond au paramètre σ dans l'approche signal, joue un rôle crucial dans la concordance.

6.1.3 Concordance totale et concordance d'ordre $m \geq 2$

Concordance totale La concordance d'un accord mesure les composantes spectrales communes à deux notes de l'accord. Dès lors, il est naturel de vouloir mesurer les composantes spectrales communes à toutes les notes de l'accord simultanément. C'est le rôle de la concordance totale.

La concordance totale peut être vue comme une des généralisations possibles aux accords de cardinalité supérieure de la théorie de la coïncidence des harmoniques d'Helmholtz³, à côté de la concordance. La concordance totale est en un certain sens plus contraignante que la concordance, étant donné qu'elle exige la coïncidence simultanée, et non plus deux-à-deux, des partiels issus de chaque note.

Nous définissons la concordance totale $\mathcal{C}_{Tot}$ de la verticalité t comme le produit N-linéaire des spectres synchronisés des N notes $X_t^{(i)}$ de cette verti-

1. Voir figure 6.8

2. La configuration mélodie-accompagnement est par contre possible, la concordance est alors concordance entre la mélodie et l'accompagnement.

3. Voir page 2.3.1.

calité :

$$\begin{aligned}\mathcal{C}_{Tot}(t) &= \langle X_t^{(1)}, \dots, X_t^{(N)} \rangle \\ &= \sum_k \left(\prod_{i=1}^N X_{k,t}^{(i)} \right).\end{aligned}\tag{6.12}$$

Dans le cas $N = 2$, on retrouve la concordance. Dans le chapitre 7, des ajustements seront réalisés afin de pouvoir comparer les concordances totales d'accords de cardinalités différentes sur des grandeurs homogènes.

La figure 6.11 montre l'évolution de la concordance totale sur le même extrait de trois accords synthétisés avec des sons d'orgue.

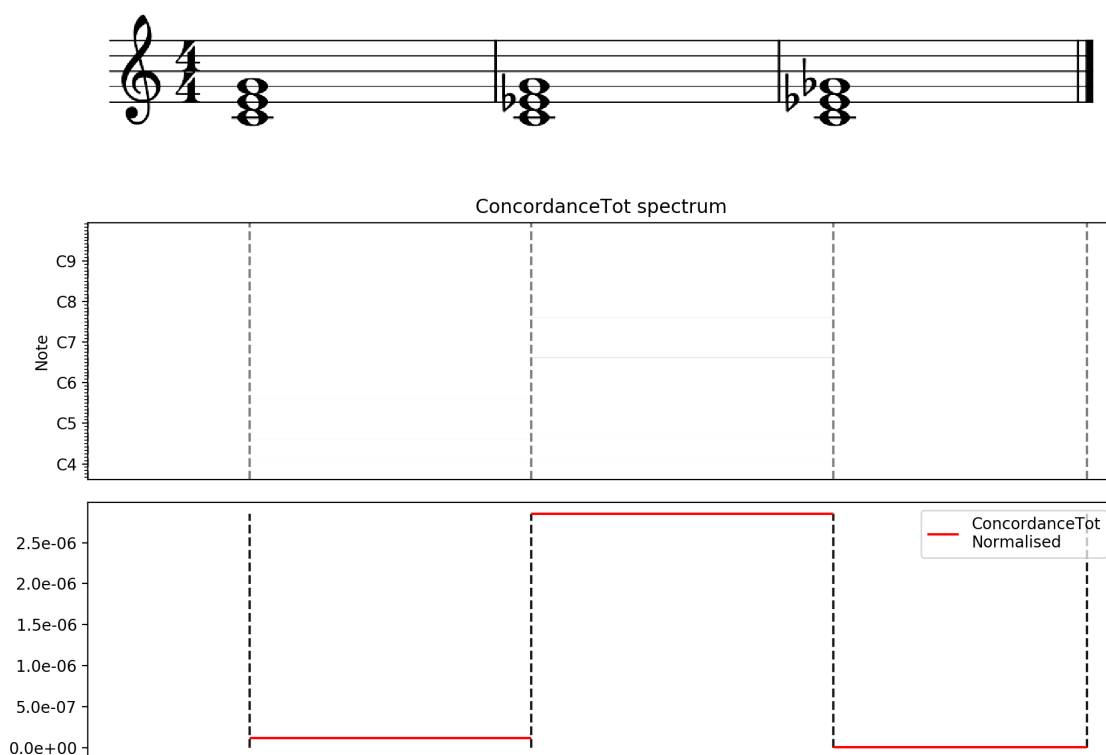


FIGURE 6.11 – Évolution de la concordance totale sur 3 accords de 3 notes synthétisés avec les jeux *Bourdon + Flûte4* : *Do M*, *Do m* et *Do dim*. En haut, la répartition spectrale des contributions à la concordance totale, en bas, les valeurs de concordance totale.

L'accord majeur et l'accord diminué ont une concordance totale négligeable par rapport à celle de l'accord mineur, dont les contributions sont situées sur deux bandes de fréquence : *sol* 6 et *sol* 7. *Sol* 6 coïncide avec l'harmonique 6 de *do* 4, l'harmonique 5 de *mib* 4 et l'harmonique 4 de *sol* 4, comme le montre la figure 6.12.

Cette coïncidence peut être interprétée par le prisme de la série des harmoniques inférieures employée par Rameau¹, laquelle est obtenue, à partir d'une fréquence de départ f , en considérant tous les sous-multiples $f, \frac{f}{2}, \frac{f}{3}, \frac{f}{4}, \dots$

1. Voir page 95.

En effet, une fréquence f fait partie de la série harmonique d'une fréquence fondamentale f_0 si et seulement si il existe un entier k tel que $f = kf_0$, autrement dit $f_0 = \frac{f}{k}$, c'est-à-dire si et seulement si f_0 est un harmonique inférieur de f . De là, si les spectres sont harmoniques, trois notes de fréquences f_1 , f_2 , f_3 partagent une composante spectrale si et seulement si il existe une fréquence f telle que f_1 , f_2 et f_3 appartiennent à la série des harmoniques inférieurs de f .

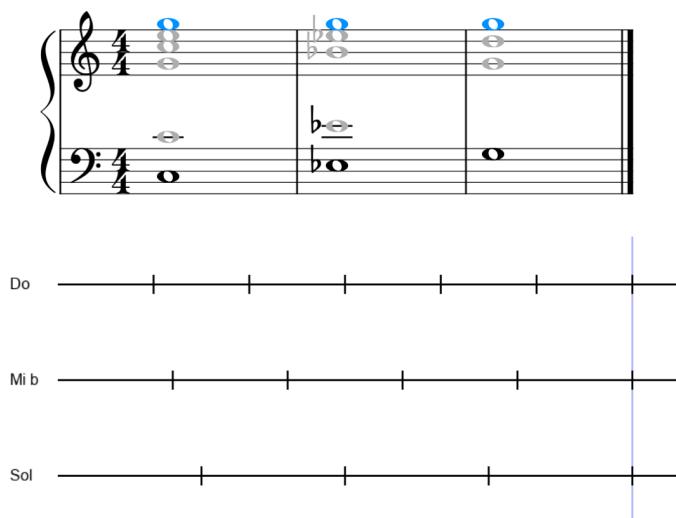


FIGURE 6.12 – Coïncidence simultanée des harmoniques sur l'accord mineur en position fondamentale. En bleu, les harmoniques communs.

De là, en exhibant la série des harmoniques inférieurs d'une fréquence (figure 6.13), on voit que la première triade à apparaître, modulo l'octave, est un accord mineur, d'où la concordance totale élevée de l'accord mineur sur la figure 6.11.



FIGURE 6.13 – Série des harmoniques inférieurs, à partir de *sol* 6. la première triade à apparaître, modulo l’octave, est un accord de *Do* mineur.

Dans le cas de spectres inharmoniques, comme celui des cloches, on peut, par des considérations de symétrie, exhiber la série des partiels inférieurs, qui ne seront cette fois pas harmoniques. Les triades ayant une forte concordance totale seront alors les triades présentes dans la série des partiels inférieurs.

Concordance d’ordre $m \geq 2$ La concordance coïncide avec la concordance totale sur les intervalles, et sur les accords à plus que deux notes elle est définie par la somme des concordances de chaque intervalle présent dans l’accord¹.

On peut procéder de manière similaire avec les triades ou les accords de cardinalité N supérieure en définissant la concordance d’ordre m , avec $m \geq 2$ fixé, comme étant la somme des concordances totales des $\binom{N}{m}$ sous-accords de cardinal m .

La concordance $\mathcal{C}_m$ d’ordre m de la verticalité t s’exprime de la manière suivante :

$$\begin{aligned} \mathcal{C}_m(t) &= \sum_{i_1 < i_2 < \dots < i_m \leq N} \langle X_t^{(i_1)}, \dots, X_t^{(i_m)} \rangle \\ &= \sum_{i_1 < \dots < i_m \leq N} \left(\sum_k \left(\prod_{\alpha=1}^m X_{k,t}^{(i_\alpha)} \right) \right), \end{aligned} \quad (6.13)$$

où N désigne le nombre de notes, avec $N \geq m$. Avec $m = 2$, on retrouve la concordance telle que définie précédemment.

Comme pour la concordance, la concordance totale et la concordance d’ordre m nécessitent d’avoir les spectres de chaque note, et donc d’être dans la configuration « pistes séparées ». Elles sont calculées à partir des spectres non ponctuels.

1. Voir équation (6.10)

6.1.4 Harmonicité

L’harmonicité mesure la ressemblance d’un spectre à un spectre harmonique, c’est-à-dire un spectre dont les partiels sont tous multiples d’une fréquence fondamentale.

Rôle dans la perception On sait que l’harmonicité joue un rôle fondamental dans la perception des hauteurs (Terhardt *et al.*, 1982; Plack et Oxenham, 2005), ainsi que dans la capacité de l’écoute à isoler les notes et à déterminer le nombre de notes (Darwin, 2005). À ce titre, elle constitue l’un des facteurs déterminant de la fusion tonale (Stumpf, 1883, 1890; DeWitt et Crowder, 1987; Schneider, 1996), qui est la tendance à percevoir deux notes comme une unité constituant une seule note. Avec la rugosité et la familiarité, elle est l’une des trois causes principales à la consonance (Parncutt *et al.*, 2019; Harrison et Pearce, 2020). Il a d’ailleurs été montré que la dimension plaisante des accords consonants est due à l’harmonicité plus qu’à l’absence de partiels (McDermott *et al.*, 2010).

Implémentation Les modèles d’harmonicité sont intimement liés aux modèles de détection de la hauteur (Terhardt, 1974). Ils suivent tous la même idée, bien qu’aucun consensus n’ait été trouvé dans l’implémentation.

L’idée des modèles d’harmonicité est de fixer un spectre harmonique initial, et de le faire glisser sur l’axe des fréquences pour repérer la concordance¹ maximale avec ce spectre. Il s’agit de trouver la corrélation maximale. Cela permet d’obtenir une hauteur fondamentale virtuelle ainsi que la ressemblance du spectre avec un spectre patron issu de cette hauteur fondamentale virtuelle.

Avec les spectres en gaussiennes de l’approche symbolique, le spectre patron est choisi comme ayant $K + 10$ partiels, K étant le nombre de partiels des spectres des sons, avec la même largeur spectrale des partiels σ et sans décroissance. Avec les spectres de Shepard, la corrélation est faite sur une octave.

Dans l’approche signal, le spectre patron est construit avec κ partiels harmoniques qui sont des fenêtres rectangulaires d’amplitude 1 et de largeur σ_{harm} constantes. Dans la suite, sauf mention contraire, on prend $\kappa = 15$ et $\sigma_{harm} = 3$ seizièmes de ton. En notant $X_{patr(k)}$ le spectre patron commen-

1. Au sens où on l’a définie dans la partie 6.1.2

cant sur la plage de fréquence k , la hauteur fondamentale virtuelle $\mathcal{V}(t)$ et l'harmonicit  $\mathcal{H}(t)$ de la verticalit  t sont d finies par :

$$\mathcal{V}(t) = \operatorname{argmax}_k \langle X_{\text{patr}(k)}, X(t) \rangle \quad (6.14)$$

et

$$\mathcal{H}(t) = \max_k \langle X_{\text{patr}(k)}, X(t) \rangle. \quad (6.15)$$

La figure 6.14 montre l' volution de l'harmonicit  sur le m me extrait de trois accords synth tis s avec des sons d'orgue. En haut sont repr sent s les profils des potentielles hauteurs fondamentales virtuelles, dont le maximum est atteint en la hauteur fondamentale virtuelle.

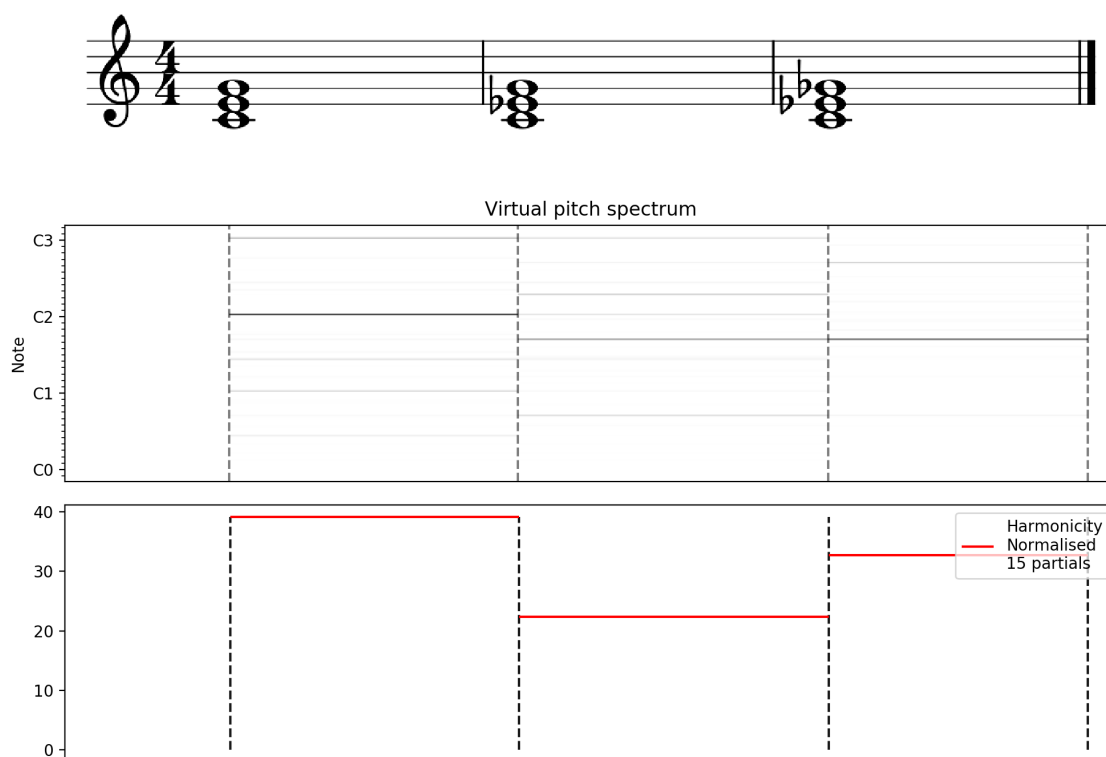


FIGURE 6.14 – Évolution de l’harmonicité sur 3 accords synthétisés avec les jeux *Bourdon + Flûte4* : *Do M*, *Do m* et *Do dim*. En haut, le profil des potentielles hauteurs fondamentales virtuelles, dont le maximum est atteint sur la hauteur fondamentale virtuelle, et en bas l’exponentielle des valeurs d’harmonicité. L’exponentielle sert à mieux séparer les valeurs pour la représentation. Les hauteurs fondamentales virtuelles sont les suivantes - *Do M* : *do 2*, *Do m* : *lab 1* et *Do dim* : *lab 1*.

Les hauteurs fondamentales virtuelles obtenues sont les suivantes - *Do M* : *do 2*, *Do m* : *lab 1* et *Do dim* : *lab 1*. La figure 6.15, qui représente l’emplacement des accords par rapport au spectre de leur hauteur fondamentale virtuelle, permet d’interpréter les différentes valeurs d’harmonicité. En effet, les notes des accords majeur et diminué sont toutes incluses dans le spectre harmonique de leur hauteur fondamentale virtuelle, alors dans le cas de l’accord mineur, seules la note fondamentale et la tonique le sont, d’où

une plus grande harmonicité des accords majeur et diminué. De plus, les notes de l'accord majeur occupent les rangs 4, 5 et 6 parmi les harmoniques de sa hauteur fondamentale virtuelle, tandis que celles de l'accord diminué occupent les rangs plus éloignés 5, 6, 7, d'où une harmonicité plus grande de l'accord majeur sur l'accord diminué.

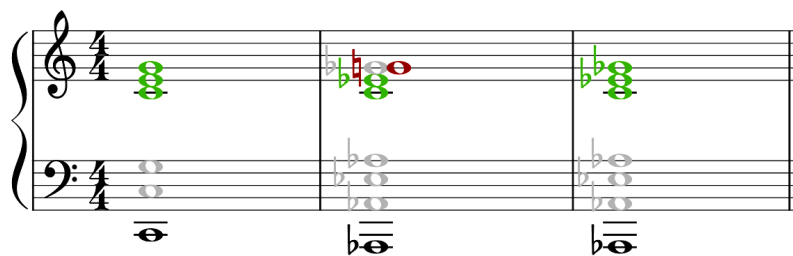


FIGURE 6.15 – Hauteurs fondamentales virtuelles sur les 3 accords précédents obtenues par l'approche signal. Les notes des accords majeur et diminué sont toutes incluses dans le spectre harmonique de leur hauteur fondamentale virtuelle, mais dans le cas de l'accord mineur, seules la note fondamentale et la tonique le sont.

6.1.5 Tension

La tension est une dimension essentielle de la perception de la musique et des émotions en musique, fortement liée à la projection et l'attente (Huron, 2006; Collins *et al.*, 2014), au suspens et à l'intensité émotive. Notion largement intuitive, elle est définie tantôt en termes statiques (Cook, 2009), tantôt en termes dynamiques (Farbood et Price, 2017)¹. Elle est cruciale dans la musique tonale puisqu'elle participe du schéma fondamental de tension-résolution.

Difficile à formaliser et à quantifier, le concept de tension est complexe et multidimensionnel, décrit comme émergent et holistique, résultant de l'interaction combinée de nombreux facteurs musicaux, tels l'intensité, le contour mélodique, les aspects dynamiques et rythmiques, les aspects structuraux, le

1. « Une augmentation de la tension revient à une augmentation de l'excitation ou de l'intensité, alors qu'une diminution de la tension est associée à un sentiment de relaxation et de résolution. » (*"a rise in tension has been equated with an increase in excitement or intensity while a decrease in tension has been described as a feeling of relaxation or resolution"*), tiré de (Farbood et Price, 2017), traduction de l'auteur.

timbre (Nielsen, 1987; Granot et Eitan, 2011; Farbood, 2012; Collins *et al.*, 2014; Gingras *et al.*, 2016; Farbood et Price, 2017), ainsi que les aspects harmoniques (Lerdahl et Krumhansl, 2007; Cook, 2009; Herremans et Chew, 2016). Les paramètres psychoacoustiques de bas niveau sont impliqués, notamment la rugosité (Pressnitzer et McAdams, 1999) et l’inharmonicité (Farbood et Price, 2017).

Implémentation Nous implémentons ici le modèle de tension de Cook (Cook, 2009), qui traite l’aspect statique du phénomène. Cook part de l’expérience musicale, et de l’observation que les triades vérifiant la propriété d’égalité entre l’intervalle supérieur et l’intervalle inférieur sont perçues, en isolation, comme tendues. C’est le cas de l’accord diminué en position fondamentale, formé de deux tierces mineures, de l’accord augmenté, ou de l’accord formé de deux demi-tons. À partir de là, Cook construit la courbe fondamentale de tension des triades de sons simples de la figure 6.16 en fonction de la différence entre les deux intervalles de la triade, de manière à mettre en avant les triades qui vérifient la propriété précédente. La tension de sons complexes est obtenue en sommant les contributions de tous les triplets de composantes spectrales.

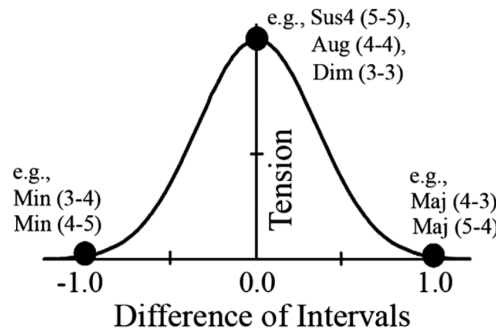


FIGURE 6.16 – Courbe de tension d’une triade de sons simples, tiré de (Cook, 2009). En abscisses, la différence en demi-tons entre l’intervalle supérieur et l’intervalle inférieur.

Par l’introduction de la tension, Cook cherche à modéliser la perception des triades, et il la combine avec un modèle de rugosité pour obtenir l’instabilité, qu’il teste sur des expériences perceptives. Le modèle de tension

n'est donc pas directement un modèle perceptif, mais avant tout un modèle théorique.

Le modèle de tension de Cook est à mettre en relation avec la notion d'*unevenness* (Tymoczko, 2010; Parncutt *et al.*, 2019), qui mesure l'irrégularité de l'espacement des notes sur le cercle chromatique¹. Parncutt a montré que cette notion est corrélée à la dissonance.

L'intérêt du modèle de tension dans ce travail tient à la fois à son aspect théorique, et au fait qu'il est pensé pour les triades, et non pour les dyades. Concernant le premier point, la courbe de tension de la figure 6.16 n'est pas une courbe expérimentale comme l'est la courbe de rugosité de deux sons simples de (Plomp et Levelt, 1965)², mais une courbe théorique qui sert, dans le cas de triades de sons simples, à repérer les triades satisfaisant une propriété particulière, laquelle est motivée par des considérations musicales. Construit *ad hoc* sur les triades de sons simples, ce modèle gagne en intérêt à être généralisé aux sons complexes ainsi qu'aux accords de plus de trois notes. De plus, à l'instar de la concordance d'ordre 3³, la tension traite des interactions entre 3 partiels simultanément, et non pas deux comme la rugosité ou la concordance.

Ce point pose la question de l'apparition de dimensions perceptives ou analytiques nouvelles avec l'ajout d'une note. Cook parle à ce sujet de « *psychophysique à 3 sons* »⁴.

La courbe de la figure 6.16 est paramétrisée de la manière suivante :

$$f(x) = e^{-\left(\frac{x}{\delta}\right)^2}, \quad (6.16)$$

avec $\delta = 0.6$.

De là, la tension $\mathcal{T}(t)$ de la verticalité t est calculée sur le spectre en pics

1. Parncutt propose deux modèles distincts. Étant donné un accord, son *unevenness* est définie soit comme l'écart entre le plus grand intervalle et le plus petit intervalle sur le cercle chromatique, soit comme l'écart-type des intervalles sur ce cercle. Comme le modèle de tension de Cook, ces modèles peuvent être généralisés aux interactions entre partiels.

2. Voir figure 6.1.

3. Voir page 218.

4. Voir (Cook, 2009).

global X en sommant toutes les contributions :

$$\mathcal{T}(t) = \sum_{i < j < k} X_{i,t} X_{j,t} X_{k,t} \times f\left(\frac{(k-j) - (j-i)}{\beta}\right), \quad (6.17)$$

où β est le nombre de subdivisions du demi-ton, ici $\beta = 8$.

La figure 6.17 représente l'évolution de la tension sur le même extrait de trois accords synthétisés avec des sons d'orgue. Comme attendu, l'accord diminué a une tension plus grande que les accords majeurs et mineurs.

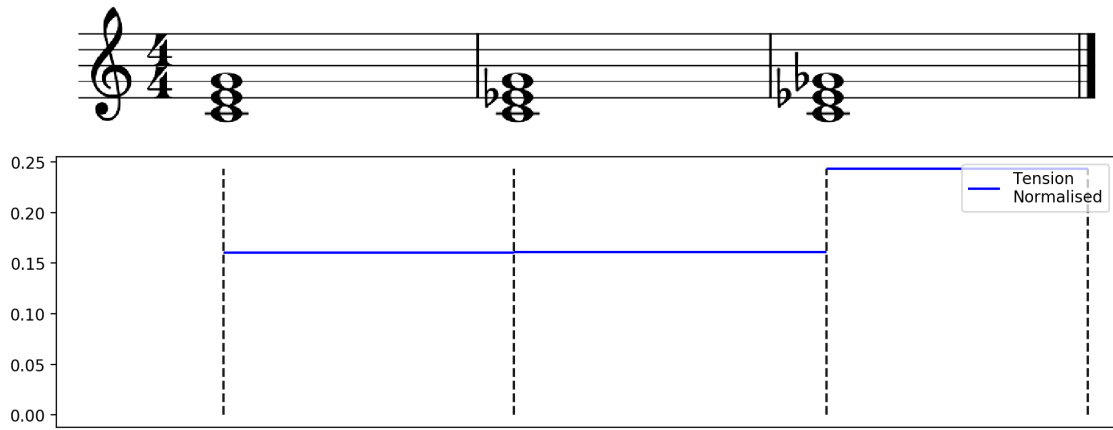


FIGURE 6.17 – Évolution de la tension sur 3 accords synthétisés avec les jeux *Bourdon + Flûte4* : *Do M*, *Do m* et *Do dim*.

6.2 Les descripteurs dynamiques

Les descripteurs dynamiques décrivent la transition entre deux verticalités. Ils peuvent être appliqués aux transitions d'accords autant qu'à la transition entre deux notes d'une mélodie. À ce titre, ils traitent l'aspect horizontal de l'harmonie.

La littérature offre relativement peu de descripteurs harmoniques dynamiques par rapport aux descripteurs harmoniques statiques. À l'époque où la notion de consonance et de dissonance est dynamique, en l'occurrence mélodique¹ (CDC 1), les « affinités » entre notes successives dépendent de la simplicité du rapport des fréquences. Avec l'apparition de la polyphonie, la consonance et la dissonance deviennent des notions essentiellement statiques (CDC 2). Le physicalisme de Rameau fait émerger du corps sonore l'accord majeur comme objet statique ; et les notions d'enchaînement et de résolution s'y ramènent. La théorie de la rugosité d'Helmholtz traite des accords isolés, de même que son approche par coïncidence des harmoniques. La fusion de Stumpf concerne elle aussi des sons entendus simultanément.

Une manière, a priori artificielle, de décrire acoustiquement une transition, consiste à partir d'un descripteur statique et à calculer sa valeur sur la somme des deux accords ou signaux qui constituent la transition, la rendant en quelque sorte figée. Euler le fait sans scrupules, en généralisant le calcul de la suavité à des transitions, des enchaînements harmoniques et même des œuvres entières². Helmholtz, de même, construit une notion d'affinité mélodique qui renvoie à la consonance statique :

Le lien musical entre deux notes consécutives peut être déterminé :

1. *Par l'affinité des sons. Celle-ci est directe, lorsque les deux sons consécutifs forment un intervalle consonant. [...]*
2. *Les notes peuvent se relier musicalement, par le peu de différence de leur hauteur³.*

Le « *peu de différence de leur hauteur* » peut renvoyer, même si Helmholtz ne s'y réfère pas explicitement, à la rugosité et à la dissonance ; on n'est pas loin de : « les notes peuvent se relier musicalement, lorsque les deux sons

1. Avant l'apparition de la polyphonie, les notions de consonance et de dissonance traduisent une notion d'affinité entre deux notes jouées successivement. Voir page 74.

2. Voir page 88.

3. Tiré de (Helmholtz, 1863).

consécutifs forment un intervalle dissonant. » Stumpf, lui aussi, développe une notion de fusion des notes successives en se ramenant au cas simultané. Il évoque pour cela la mémoire, qui peut justifier l’approche statique sur une transition.

Nous apprécions parfaitement la consonance de deux sons qui se suivent. [...] Or des sons successifs ne fusionnent évidemment pas à proprement parler. [...] La persistance du souvenir, qui produit un moment une fusion réelle des sons successifs en la représentant dans la mémoire simultanément, suffit à expliquer ce fait¹.

Helmholtz fait aussi appel à la mémoire pour justifier l’affinité des sons successifs. Ce recours à la mémoire pourrait difficilement justifier une rugosité des transitions, du fait de l’hypothèse physiologique.

Les descripteurs dynamiques présentés dans la suite sont calculés à partir du spectre global, ponctuel ou non, ils ne nécessitent donc pas les pistes séparées, dans le cas de plusieurs notes. De plus, ils sont temporellement symétriques², c’est à dire invariables par changement de l’ordre des accords. Cette symétrie temporelle sera brisée par les descripteurs de contexte, qui tiennent compte, pour caractériser une transition, de ce qui a été entendu auparavant³.

6.2.1 Changement harmonique

Le changement harmonique quantifie les différences spectrales entre les verticalités d’un enchaînement. Il est obtenu en faisant la somme sur toutes les bandes de fréquences des valeurs absolues des différences des composantes spectrales.

Usage Le changement harmonique est largement utilisé dans les techniques de détection d’*onsets* (Dixon, 2001; Hainsworth *et al.*, 2003; Harte *et al.*, 2006), où son calcul constitue une étape préliminaire. Il mesure alors les différences spectrales entre les fenêtres d’analyse, et assure un suivi fin de l’évolution spectrale.

1. Tiré de (Lalo, 1908, p.135).

2. Voir partie 4.3.3

3. Les descripteurs dynamiques de contexte seront utilisé dans la partie 9.3.

Ce suivi est capable de mettre en évidence des dynamiques à une échelle temporelle inférieure à la note, ce qui représente un grand intérêt musicologique. Chouvel ([Chouvel et al., 2007](#)) a introduit la transformée de Fourier différentielle (TFD), qui est une représentation spectro-temporelle des variations spectrales. L'absence de valeur absolue lui permet de distinguer sur chaque bande de fréquence les nouveautés des pertes spectrales, qu'il interprète en termes d'énergie entrante et d'énergie sortante, par analogie avec les systèmes physiques. La figure 6.18 montre la résultante (somme sur les composantes spectrales) de la transformée de Fourier différentielle sur une synthèse avec des sons de violon de la série de la *Suite Lyrique* d'Alban Berg.

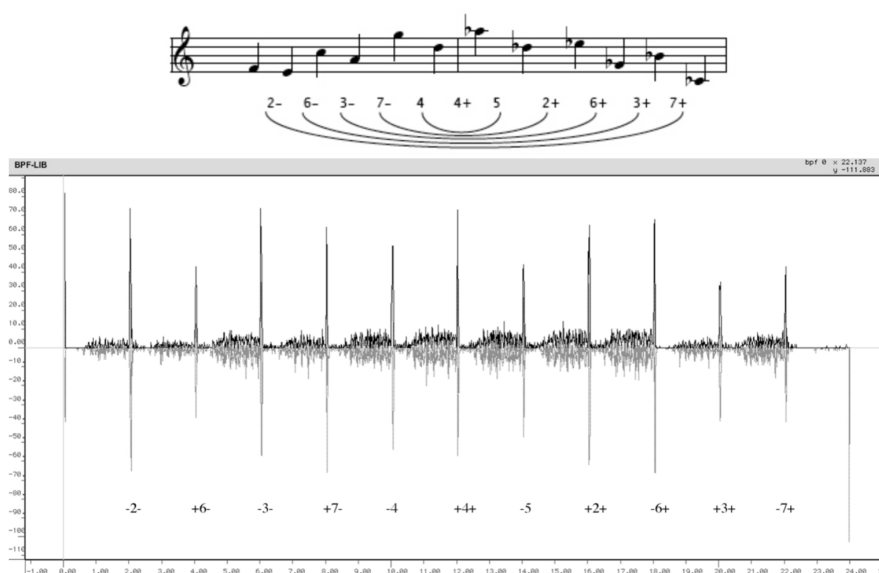


FIGURE 6.18 – Exemple d'utilisation de la résultante de la Transformée de Fourier Différentielle, sur une synthèse avec des sons de violon de la série de la *Suite Lyrique* d'Alban Berg - tiré de ([Chouvel et al., 2007](#)). Les valeurs positives (en noir) correspondent aux gains d'énergie, et les valeurs négatives (en gris) aux pertes d'énergie. Les discontinuités sont dues aux changements de notes, et les oscillations sur une note au vibrato.

Le changement harmonique diffère de la résultante de la transformée de Fourier différentielle en ce que toutes les contributions sont comptées positivement.

Implémentation Le spectre différentiel en amplitude $X_{t \rightarrow t+1}^{diff}$ de la transition entre les verticalités t et $t + 1$ est défini de la manière suivante :

$$X_{k,t \rightarrow t+1}^{diff} = |X_{k,t+1} - X_{k,t}| \quad (6.18)$$

Le changement harmonique $\mathcal{D}(t, t + 1)$ est alors obtenu en prenant la norme 1 du spectre différentiel :

$$\mathcal{D}(t, t + 1) = \|X_{t \rightarrow t+1}^{diff}\|_1 = \sum_k |X_{k,t+1} - X_{k,t}| \quad (6.19)$$

Du fait de la valeur absolue, on ne distingue pas la partie entrante de la partie sortante.

Sur la transformée de Fourier différentielle de la figure 6.18, le changement harmonique correspond au double de l'amplitude des discontinuités entre les notes.

La figure 6.19 montre l'évolution du changement harmonique sur un extrait du *Via Crucis* de Liszt pour orgue, tiré des mesures 12 et 13 du *Vexilla Regis* introductif, et synthétisé avec les jeux *Bourdon + Flûte 4*.

Les transitions vers les accords 3 et 5 admettent chacune une note commune, à cause de laquelle les différences spectrales sont moins importantes que pour les transitions vers les accords 2 et 4, qui n'admettent pas de note commune. La considération de la présence ou non de notes communes n'est bien entendu pas suffisante, étant donné que les recoupements spectraux peuvent avoir lieu entre les partiels d'ordre supérieur.

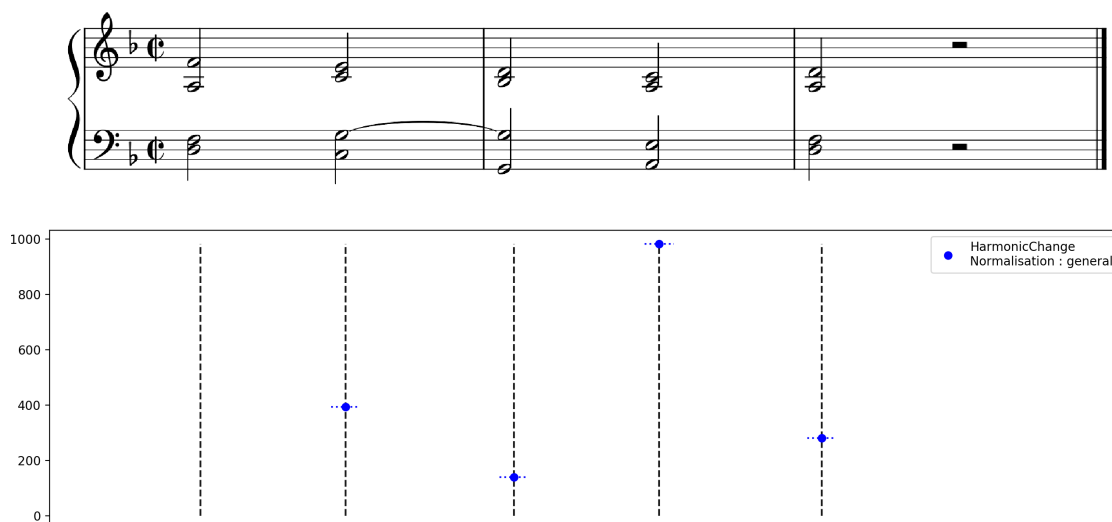


FIGURE 6.19 – Évolution du changement harmonique sur un extrait du *Via Crucis* de Liszt, synthétisé avec les jeux *Bourdon + Flûte4*. En bas, l'exponentielle du changement harmonique, de manière à mieux séparer les valeurs sur la représentation. Le graphique est synchronisé temporellement avec la partition, la valeur sous l'accord t correspondant à la transition vers lui.

6.2.2 Concordance différentielle

Si le changement harmonique mesure les différences spectrales entre deux verticalités successives, la concordance différentielle mesure au contraire leur similarité spectrale. Elle quantifie la partie constante du son. Pour Lalo, cette partie constante a une fonction de lien entre les deux accords :

Un harmonique commun à deux accords et retenu dans la mémoire peut servir de pivot ou de pont entre eux¹.

La concordance différentielle est obtenue en prenant la concordance entre les deux spectres successifs.

Usage La concordance différentielle est d'abord utilisée comme modèle perceptif de l'affinité, c'est-à-dire du degré de bon enchaînement de deux accords.

1. Tiré de (Lalo, 1908, p.100)

Les premiers modèles prennent en compte les similarités entre les hauteurs virtuelles (Terhardt *et al.*, 1982; Parncutt, 1989), et ce n'est que récemment qu'ils ont été simplifiés en considérant seulement les composantes spectrales réelles (Milne *et al.*, 2016). La notion d'affinité de Milne est identique, dans son calcul, à la celle de concordance différentielle.

Le terme de « concordance différentielle » est introduit par Chouvel (Chouvel, 1998), en même temps que la concordance statique. Ici, cette notion n'est pas considérée ni utilisée comme modèle perceptif, mais comme modèle physique et théorique qui sert à l'analyse musicale. Chouvel l'interprète comme l'énergie transférée d'un état à l'autre du système. La figure 6.20 montre les rapports entre concordance statique et concordance différentielle.

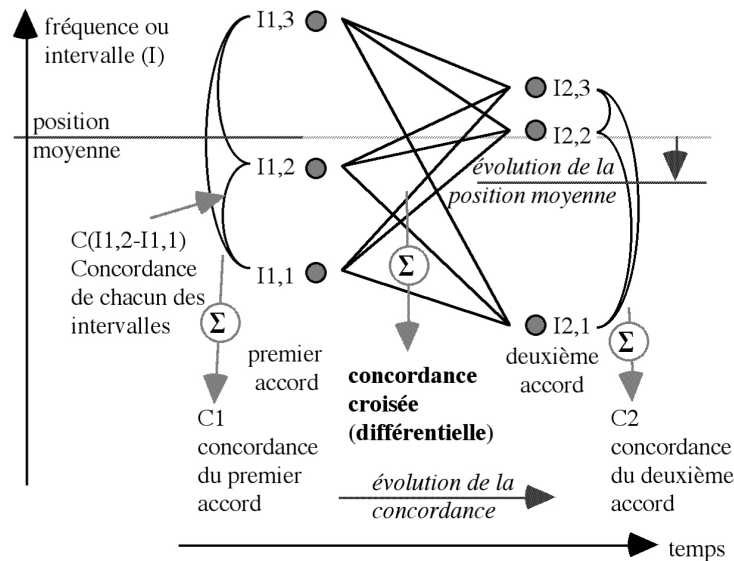


FIGURE 6.20 – représentation schématique des rapports mis en jeu dans la mise en relation séquentielle de deux accords, tiré de (Chouvel, 1998)

Chouvel calcule la concordance différentielle sur tous les couples de notes entre le premier et le second accord. Cette approche équivaut à calculer la concordance entre les deux spectres globaux, aux phases près cependant, qui peuvent avoir l'effet de diminuer les amplitudes sur le spectre global.

La concordance différentielle n'est pas un descripteur trivial, au sens où elle ne peut pas être déduite des concordances statiques des deux accords.

Implémentation La concordance différentielle $\mathcal{C}_{diff}(t, t+1)$ de la transition entre les verticalités t et $t+1$ est obtenue en prenant le produit scalaire des spectres X_t et X_{t+1} des deux verticalités :

$$\begin{aligned}\mathcal{C}_{diff}(t, t+1) &= \langle X_t, X_{t+1} \rangle \\ &= \sum_k X_{k,t} X_{k,t+1}\end{aligned}\tag{6.20}$$

Application La figure 6.21 montre l'évolution de la concordance différentielle sur le même extrait du *Via Crucis*.

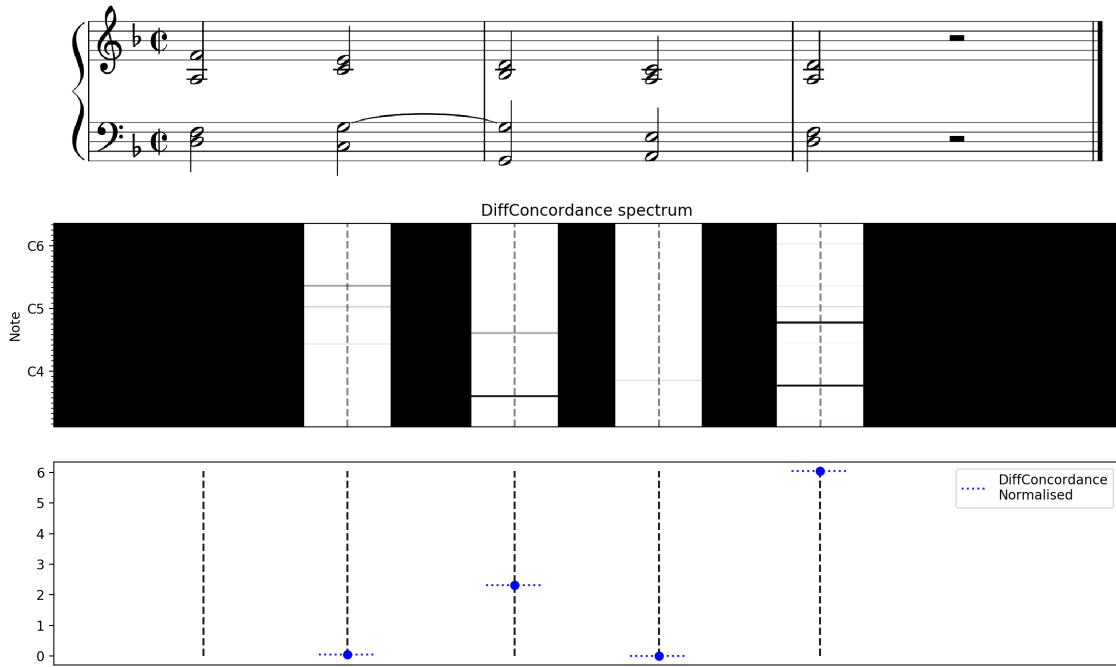


FIGURE 6.21 – Évolution de la concordance différentielle sur un extrait du *Via Crucis* de Liszt, avec des sons d’orgue synthétisés. En haut, la répartition spectrale des contributions à la concordance différentielle, en bas, les valeurs de concordance différentielle.

La répartition spectrale des contributions, reproduite sur la partition de la figure 6.22, permet de déterminer les notes du texte qui apportent la plus grande contribution.



FIGURE 6.22 – Coïncidence des partiels entre accords successifs, extrait du *Via Crucis* de Liszt. Les partiels représentés en bleu sont ceux visibles sur la répartition spectrale des contributions à la concordance différentielle de la figure 6.21.

La valeur élevée du dernier enchaînement est essentiellement due à la contribution de *la* 4, lequel provient de deux notes de l'avant-dernier accord : la basse *la* 2 et *la* 3, et deux notes du dernier accord : la basse *ré* 3 et *la* 2. De la même manière, le *sol* 4 du deuxième enchaînement provient de la basse et de la quinte du deuxième accord, et des deux *sol* de l'accord suivant. Dans les deux cas, la basse a un mouvement de quarte ou de quinte¹. Il est intéressant de noter que la réalisation de l'accord qui est en rapport de quinte ascendante avec l'autre (*Sol* m dans le deuxième enchaînement, *La* m dans le dernier), par le doublement de la basse, maximise la concordance différentielle. De plus, dans aucun des quatre accords de ces deux enchaînements les tierces ne contribuent de manière significative à la concordance différentielle. Par contre, dans le premier enchaînement, les tierces jouent un rôle primordial, puisque le *fa* 3 contribue au *do* 5, et le *mi* 4 au *mi* 5. D'ailleurs, la doublure sur le premier accord de la tierce à la main gauche, qui se retrouve à une quarte ascendante de la basse de l'accord suivant, interagit avec elle sur l'harmonique *do* 5.

Les considérations précédentes suggèrent que certains choix d'écriture vont dans le sens d'une maximisation de la concordance différentielle. Pour s'en convaincre, reprenons l'exemple précédent en trahissant Liszt par des « fautes » d'écriture, et observons leur effet sur la concordance différentielle sur la figure 6.23. Ces « fautes » sont simplement des modifications apportées sur les doublures, qui ne modifient pas les types d'accords.

1. Lesquels appartiennent à la même classe d'intervalle.

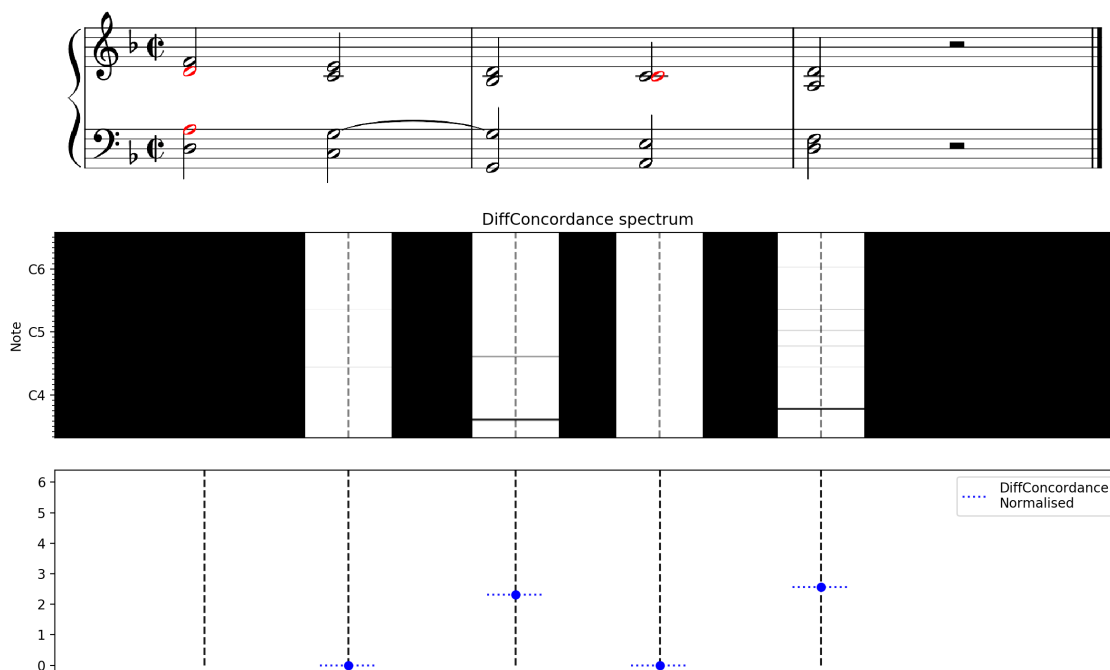


FIGURE 6.23 – Évolution du changement harmonique sur un extrait modifié du *Via Crucis* de Liszt. Les notes modifiées sont en rouge sur la partition. Ces modifications ont pour conséquence de faire diminuer la concordance différentielle de tous les enchaînements affectés.

Les modifications ont été apportées sur la disposition de deux accords : sur le premier accord, la tierce n'est plus doublée par la main gauche, créant ainsi des quinte et octave parallèles, tandis que sur le quatrième accord ce n'est plus la note fondamentale qui est doublée, mais la tierce. L'effet est radical : toutes les valeurs de concordance différentielle diminuent sur les enchaînements touchés par les modifications. Sur le dernier enchaînement elle diminue de moitié, du fait que la contribution du *la* 4 n'est due qu'à une seule note du quatrième accord, contre deux dans la disposition correcte. Dans le premier enchaînement, l'absence de la doublure de la tierce à la main gauche annule la contribution de *do* 5, ce qui divise la valeur de la concordance différentielle par 25¹. La concordance différentielle du troisième enchaînement

1. Cette diminution n'est pas visible sur la partie basse à cause de l'échelle en ordonnée, elle est en revanche visible sur la répartition spectrale des contributions à la concordance

diminue quant à lui de moitié. Le classement des enchaînements selon leur concordance différentielle n'est par contre pas modifié.

Ce résultat est important en ce qu'il est révélateur de la manière dont un descripteur harmonique, de nature acoustique, peut contribuer à justifier un usage musical. Dans l'exemple précédent, la disposition des notes telle que l'a conçue Liszt maximise la concordance différentielle, si l'on considère fixes la basse et le soprano. Cette disposition, avant d'optimiser un descripteur, est d'abord l'effet de l'usage musical et des choix esthétiques du compositeur. La corrélation avec la concordance différentielle montre le rôle que joue cette dernière dans la perception de ce que sont les bons enchaînements, question qui englobe de nombreux aspects de la perception, parmi lesquels figurent l'équilibre ou la conduite des voix. L'introduction de « fautes » d'harmonie ¹, notamment les quinte et octave parallèles sur le premier accord, apporte une diminution drastique de la concordance différentielle, qui suggère qu'elle peut être utilisée pour éclairer ces règles, lesquelles sont bien souvent énoncées sans justification théorique. Gardons cependant à l'esprit que les choix esthétiques et stylistiques dépassent de loin les considérations purement acoustiques, qui ne sauraient faire office de norme. En effet, quant bien même ces dernières contribueraient à éclairer un usage musical, celui-ci serait toujours situé à la fois historiquement et esthétiquement. Disons que la concordance différentielle, comme les autres descripteurs, est à même d'apporter un éclairage acoustique sur des choix qui peuvent être motivés par d'autres considérations.

6.2.3 Rugosité différentielle

La rugosité différentielle détecte, dans une transition, les partiels successifs qui sont suffisamment proches pour créer de la rugosité s'ils étaient entendus simultanément. Une caractérisation de l'harmonie par Dahlhaus vient éclairer sa position par rapport aux descripteurs dynamiques précédents :

*L'harmonie signifie la concordance de ce qui est distinct ou opposé*².

Cette caractérisation paraît particulièrement pertinente dans le débat entre

différentielle

1. Telles qu'elles sont dictées par de nombreux traités d'harmonie, par exemple chez Bitsch (Bitsch, 1957).

2. Tiré de *La tonalité harmonique : étude des origines* (Dahlhaus, 1967).

Stumpf et Helmholtz¹, la fusion étant une concordance psychologique, pris dans le sens d'indistinction dans la sensation de ce qui est distinct physiquement, tandis que la rugosité d'Helmholtz provient d'une opposition. À partir de la courbe de rugosité fondamentale page 200, on peut distinguer trois modes d'interaction entre deux partiels, selon leur distance :

1. identité : les partiels sont confondus ;
2. opposition : les partiels créent des battements ;
3. distinction : les partiels sont suffisamment éloignés pour ne pas créer de battements.

Le changement harmonique, en détectant les différences entre deux spectres successifs, se situe au niveau de la distinction. La concordance différentielle, qui quantifie les points communs, est du côté de l'identité, tandis que la rugosité différentielle traite de l'opposition.

D'un point de vue perceptif, cependant, les rugosités statique et différentielle sont foncièrement différentes. En effet, lorsque les partiels s'enchaînent, ils ne créent pas les battements qui sont à l'origine de la sensation de rugosité dans le cas statique². L'intervalle de demi-ton *si* 4 - *do* 5, joué simultanément avec des sons simples, crée la sensation désagréable de rugosité, alors que joué successivement il donne un mouvement mélodique conjoint qui n'a a priori rien de désagréable pour l'oreille. Au contraire, il semble plutôt que l'écriture musicale privilégie les mouvements conjoints dans cette tessiture. Rappelons-nous en outre des deux modalités, selon Helmholtz, qui régissent le lien musical entre deux notes consécutives : l'« *affinité des sons* », et « *le peu de différence de leur hauteur* »³. La première est modélisée par la concordance différentielle, tandis que la seconde l'est par la rugosité différentielle.

La rugosité différentielle ne doit donc pas être associée à la valence négative comme l'est la rugosité statique.

Malgré la popularité des modèles de rugosité statique, c'est la première fois, du moins à la connaissance de l'auteur, qu'est introduit un modèle de rugosité dynamique ou différentielle.

Implémentation La rugosité différentielle $\mathcal{R}_{diff}(t, t + 1)$ de la transition entre les verticalités t et $t + 1$ est obtenue en prenant la rugosité entre les

1. Voir partie 2.3.
2. Du moins si l'on admet que les sons se succèdent strictement sans se chevaucher (pas de réverbération ou d'écho).
3. Voir la note 6.2 page 226.

spectres X_t et X_{t+1} des deux verticalités :

$$\mathcal{R}_{diff}(t, t+1) = \sum_{k,l} \mathcal{R}(X_{k,t}, X_{l,t+1}, f_k, f_l) \quad (6.21)$$

où la fonction de rugosité $\mathcal{R}$ est définie aux équations (6.2), (6.3) et (6.4).

Les spectres X_t et X_{t+1} peuvent être les spectres continus ou les spectres en pics, les deux approches étant équivalentes.

Application La figure 6.24 montre l'évolution de la rugosité différentielle sur le même extrait du *Via Crucis*.

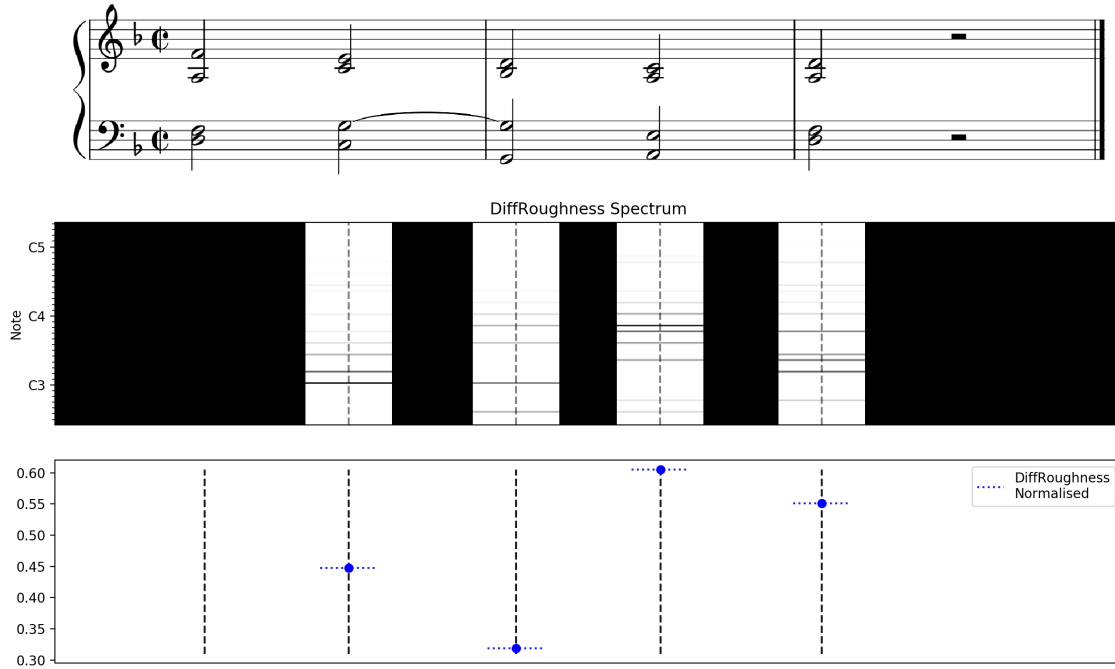


FIGURE 6.24 – Évolution de la rugosité différentielle sur un extrait du *Via Crucis* de Liszt, avec des sons d'orgue synthétisés. En haut, la répartition spectrale des contributions à la rugosité différentielle, en bas, les valeurs de rugosité différentielle.

La répartition spectrale des contributions à la rugosité différentielle nous montre que les principales contributions proviennent des relations entre les

notes fondamentales, sans faire intervenir les partiels d'ordre supérieur. La figure 6.25 représente ces contributions sur la partition.

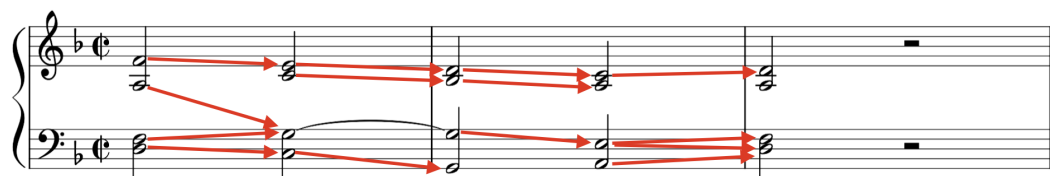


FIGURE 6.25 – Principales contributions à la rugosité différentielle, sur un extrait du *Via Crucis* de Liszt. Elles viennent majoritairement des interactions entre les fréquences fondamentales.

À la basse, les mouvements de quarte et de seconde majeure contribuent fortement à la rugosité différentielle, tandis qu'au soprano et à l'alto elle est due aux mouvements conjoints. En effet, la rugosité différentielle de deux notes successives est équivalente à la rugosité statique de l'intervalle formé par les mêmes deux notes entendues simultanément. La dépendance de la courbe de rugosité des sons simples à la fréquence de la note inférieure¹ a pour conséquence de rendre le maximum de rugosité dépendant du registre. Avec *do* 4 comme fréquence basse, l'intervalle de rugosité maximale entre deux sons simples est légèrement inférieur au ton, tandis qu'à l'octave inférieure il est situé sur la tierce mineure, et à l'octave supérieure sur le demi-ton.

La figure 6.26 montre la dépendance de la rugosité différentielle d'un ton sur la tessiture de 3 octaves. Le jeu utilisé, *Bourdon*, est pauvre harmoniquement, car ne contient que les partiels 1 et 3, ce qui a pour conséquence de rendre la dépendance au registre relativement proche de celle qui serait obtenue avec des sons simples. Cette courbe a un profil en *V* inversé, où le maximum est atteint au milieu de l'octave 3². Ce maximum augmente quand l'intervalle diminue.

1. Voir figure 6.2 page 202.

2. Ce qui est cohérent avec la figure 6.2

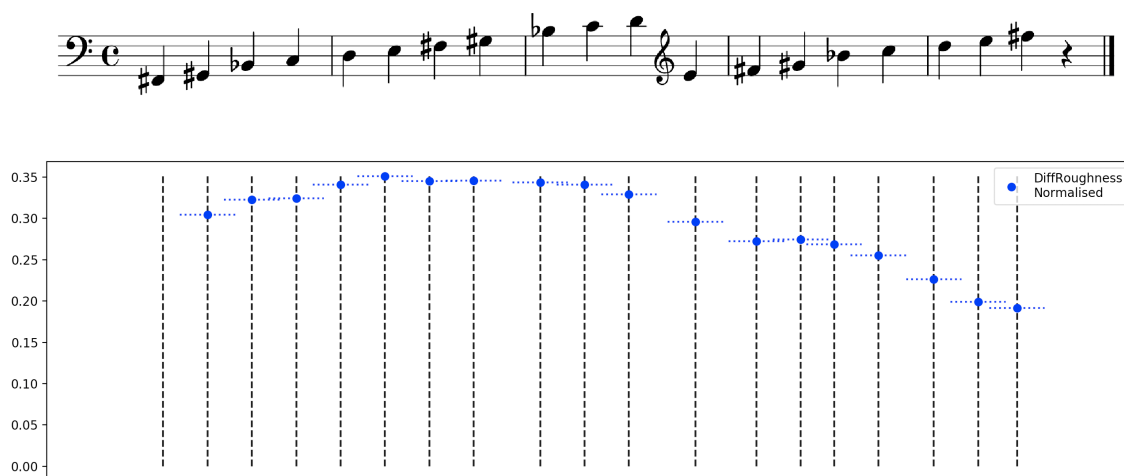


FIGURE 6.26 – Dépendance de la rugosité différentielle d’un ton sur la tessiture de 3 octaves allant de $fa\sharp 2$ à $fa\sharp 3$, avec le jeu de *Bourdon* contenant les harmoniques 1 et 3.

Du fait de cette dépendance, de manière générale, la présence de mouvements conjoints autour de l’octave 4 et de sauts de tierces autour de *do* 3 avec des sons complexes a pour conséquence de favoriser la contribution à la rugosité différentielle des fréquences fondamentales. C’est ce qui semble se passer dans l’exemple précédent extrait du *Crucis*.

Dans le cas de deux notes successives *do* 3 - *réb* 3 formant un demi-ton, l’apport à la rugosité différentielle de l’interaction des fréquences fondamentales devrait par contre être moindre vis-à-vis de l’apport de l’interaction entre les deuxièmes harmoniques *do* 4 - *réb* 4, qui se situent à l’octave au-dessus, comme nous le montre la figure 6.2. Cependant, la décroissance de l’amplitude des partiels peut venir contrebalancer la contribution de l’interaction des deuxièmes harmoniques.

Les remarques précédentes invitent à formuler trois hypothèses relatives au registre et au timbre de la cadence du *Via Crucis* étudiée :

Hypothèse 1 *La rugosité différentielle est due essentiellement aux interactions entre fréquences fondamentales successives, dont certaines forment un intervalle qui atteint le maximum de sa rugosité différentielle. Cette pro-*

priété reste vraie sur un voisinage autour du registre de départ¹, lequel ne dépend pas du timbre, dans la limite de sons harmoniques. Ce voisinage sépare deux zones où le comportement de la rugosité différentielle diffère.

Hypothèse 2 *Si on s'éloigne de ce voisinage dans l'aigu, alors les interactions entre fréquences fondamentales successives continuent de constituer la plus forte contribution à la rugosité différentielle, mais globalement les valeurs du descripteur diminuent. Cette diminution est indépendante du timbre, dans la limite de sons harmoniques.*

Hypothèse 3 *Dans le grave de ce voisinage, les interactions successives entre partiels d'ordre supérieur prennent une part importante à rugosité différentielle. Le profil de rugosité différentielle et son évolution dans cette région dépendent fortement du timbre.*

Ces hypothèses nécessiteraient de faire varier le registre et le timbre, ce qui ne sera pas abordé dans ce travail. Leur validation montrerait un résultat important, à savoir que pour un certain type d'écriture (qui resterait à délimiter), le timbre joue un rôle sur ce descripteur uniquement lorsque des fréquences suffisamment graves entrent en jeu, justifiant alors une approche spectrale. Dans les registres médium et aigu, les contributions sont dues aux fréquences fondamentales, justifiant une approche symbolique à partir de la partition².

1. Il s'agit ici de voisinage en matière de transposition dans l'aigu et dans le grave.

2. Ou approche *pitch-based*, comme dans (Cong, 2016).

Chapitre 7

Contraintes en entrée et normalisations

7.1 Contraintes sur les sons utilisés

Pour pouvoir appliquer les descripteurs harmoniques à des extraits musicaux, il est impératif de s'assurer qu'ils donnent des résultats cohérents avec les objets en entrée. En effet, les choix réalisés dans l'approche signal, en particulier celui de la transformée², ainsi que la réduction du spectre à un spectre synchronisé³ sur les verticalités, ne sont pas sans poser d'éventuelles difficultés.

Celles-ci concernent avant tout l'assignation du calcul des descripteurs harmoniques aux verticalités. Les verticalités sont caractérisées par des hauteurs de notes et des spectres. Elles sont situées en-dehors de toute notion de temps et de durée. Autrement dit, dans un extrait musical, deux accords qui ne diffèrent que par leur emplacement temporel et leur durée sont considérées comme des verticalités équivalentes, et on attend des descripteurs harmoniques qu'ils soient constants sur les deux verticalités. Or, deux effets peuvent remettre en cause cette propriété : une longueur trop courte des verticalités, et une variation temporelle des sons utilisés.

2. Transformée à facteur Q constant avec une précision spectrale au 16^e de ton

3. Voir l'implémentation dans la partie [5.3](#)

7.1.1 Temps d'établissement de l'harmonie

Dans l'approche signal, la transformation spectrale à facteur Q constant convoque des fenêtres de longueur différente selon la plage de fréquence, la taille augmentant dans les fréquences graves, selon l'équation (5.8). Avec une précision fréquentielle au 16^e de ton, la taille est vite de l'ordre de grandeur de la seconde : pour *do 4* elle est de 0.79 s, et pour *do 3* de 1.6 s¹, ce qui est relativement long pour une analyse spectrale. Cela peut mener à des inconsistances sur le spectrogramme synchronisé sur les verticalités. En effet, si la durée d'une verticalité est plus courte que la longueur de la fenêtre d'analyse sur sa composante spectrale la plus grave, la marge $\tau(f_k)$ ² ne peut être appliquée sur cette fréquence, et le calcul de la médiane déborde sur les verticalités précédentes. Cela peut avoir plusieurs conséquences sur le spectre synchronisé de cette verticalité :

- Débordement des composantes spectrales de la ou des verticalité(s) précédente(s).
- Mauvaise détection de la fréquence grave, voire absence de détection.

La figure 7.1 montre un exemple où la durée des notes est trop courte par rapport à leur fréquence, ce qui se traduit par des artefacts à la fois sur le spectre et sur le profil d'énergie spectrale, cette dernière étant définie comme le carré de la norme $\mathcal{L}_2$ du spectre en amplitude. Les doubles croches durent 25 ms, alors que la fenêtre d'analyse de *do3* dure 1.6 s, contre 1 s pour *sol4*.

Sur le spectre synchronisé, cela se voit par la détection partielle du *do3*, son débordement sur les deux verticalités suivantes, la non-détection de son 2^e harmonique, le débordement du *do4*, la non détection du *mi3* et son débordement. Le profil d'énergie n'est pas plus cohérent, puisque l'énergie de la double croche *do3* est inférieure à l'énergie de la ronde *do3*, alors qu'on attend une égalité des énergies.

Une bonne manière de traiter cet exemple serait de diminuer la résolution fréquentielle de l'analyse, mais avoir des fenêtres temporelles ne dépassant pas la longueur des notes au dessus de *do 3* oblige à descendre à une précision du demi-ton, ce qui est trop imprécis pour le calcul des descripteurs harmoniques.

1. Voir tableau page 5.3.1.

2. Voir l'équation (5.12) page 190.

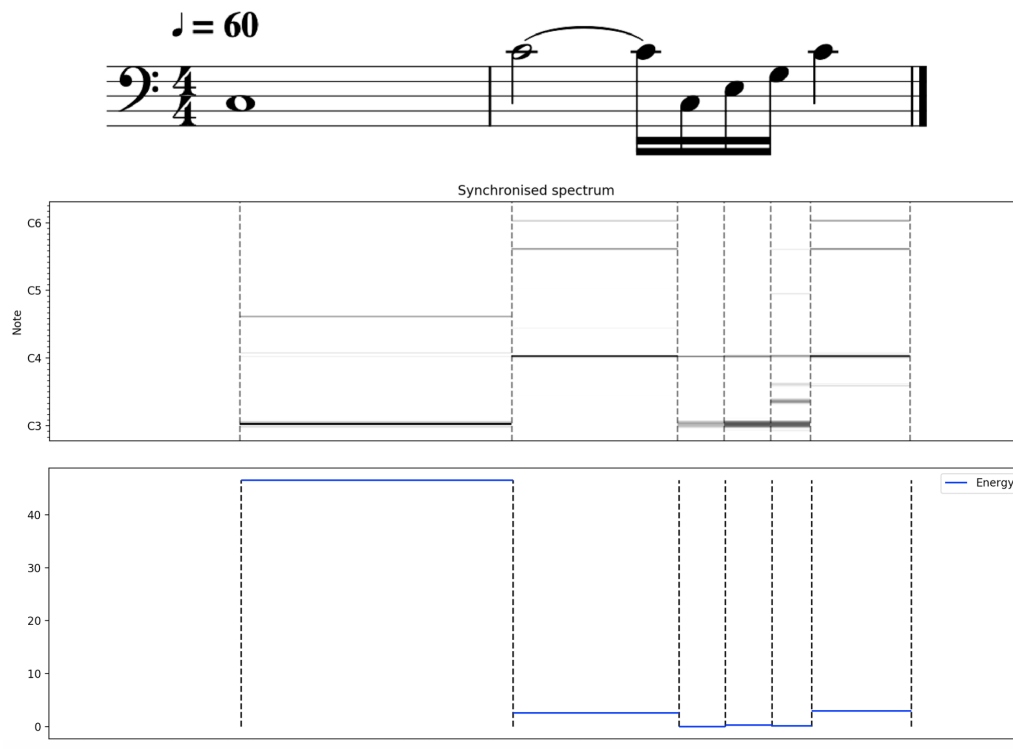


FIGURE 7.1 – Spectre synchronisé et évolution de l'énergie spectrale, avec le jeu *Bourdon*. Les doubles croches sont trop courtes par rapport aux largeurs de fenêtres d'analyse, et l'harmonie n'a pas le temps de s'établir.

7.1.2 Variation temporelle de l'amplitude

Un autre facteur problématique pour le calcul des descripteurs harmoniques peut être la variation d'amplitude au sein d'une verticalité, comme nous l'avons déjà évoqué page 155. Cette variation a plusieurs origines, elle peut être :

- une caractéristique de l'enveloppe temporelle des sons, notamment dans les phases d'attaque et d'extinction¹ ;
- une caractéristique du timbre employé (trémolo d'une chanteuse lyrique) ;
- une instabilité du contenu spectral (sur le piano, extinction plus rapide des partiels aigus que graves) ;

1. Voir figure 4.7.

— l’effet de battements entre fréquences proches.

Les modulations d’amplitude comme le trémolo d’un chanteur peuvent être liées à une nécessité de saillance contextuelle, c’est-à-dire à une nécessité de ressortir perceptivement pour guider l’attention. Si leur période temporelle est du même ordre de grandeur que la durée des verticalité, ces modulations risquent de fausser l’analyse. Pour éviter cet écueil, il vaut mieux privilégier un répertoire recherchant l’homogénéité plutôt que la saillance. C’est le cas dans beaucoup d’œuvres pour chœur, dont l’écriture et la stabilité des voix favorisent une texture homogène.

La figure 7.2 montre l’enveloppe spectrale de la note *do*2 avec deux jeux différents. À gauche les jeux *Bourdon + Flûte*4 donnent un son stable, tandis que les jeux *Bourdon + Cheminée + Flûte*4 donnent un son dont l’amplitude oscille avec une période de l’ordre de la seconde. Cette modulation d’amplitude peut être un artefact de l’échantillonnage, mais constitue aussi une réalité sonore sur de nombreux instruments.

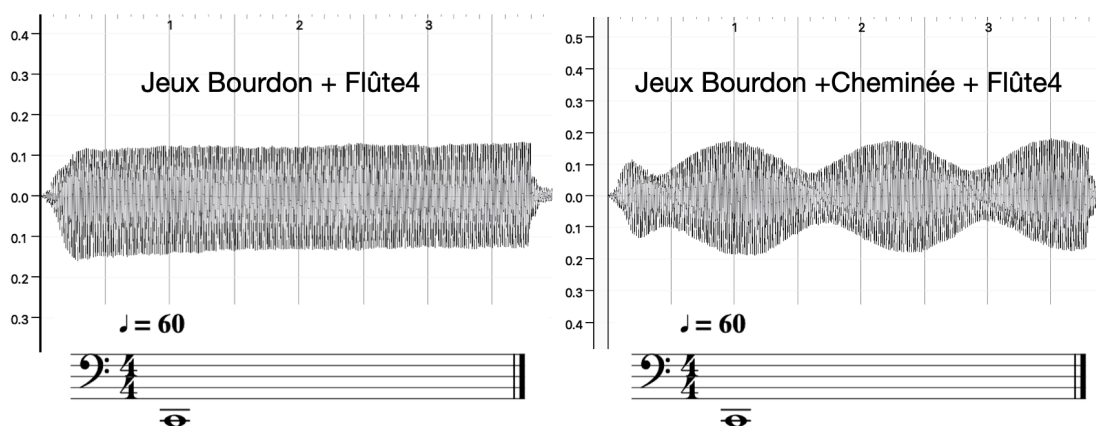


FIGURE 7.2 – Enveloppe temporelle de la note *do* 2. À gauche, les jeux *Bourdon + Flûte* : pas de modulation de l’amplitude. À droite : jeux *Bourdon + Cheminée + Flûte* : modulation de l’amplitude, avec une période égale à 1.3 s.

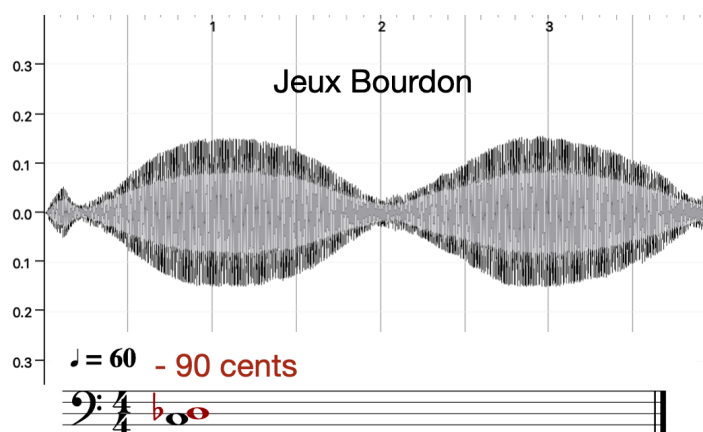


FIGURE 7.3 – Enveloppe temporelle d’un intervalle de 10 cents ($1/10^e$ de demi-ton) sur *do* 3, avec le jeu de *Bourdon*. La modulation d’amplitude est due aux battements.

La figure 7.3 montre quant à elle une variation d’amplitude de l’ordre de la seconde due aux battements d’un micro-intervalle de 10 cents.

Ce phénomène de modulation d’amplitude, quelque soit son origine, peut poser des problèmes pour le calcul des descripteurs harmoniques, lorsque son échelle est du même ordre que la durée des verticalités. La figure 7.4 montre l’évolution de trois descripteurs harmoniques statiques sur des subdivisions de 2 secondes d’un intervalle tenu de 10 cents. Les variations des descripteurs ne sont pas liées au temps d’établissement de l’harmonie, étant donné que la durée des verticalités est supérieure à la longueur de la fenêtre d’analyse de *do* 3 ; elles sont dues aux battements. En effet, les battements¹ créent des variations d’amplitude qui causent une variation significative des descripteurs. Une normalisation par l’énergie² ne permet pas de compenser ce problème.

Ces incohérences sur les descripteurs harmoniques arrivent lorsque les variations de l’amplitude sont de l’ordre de grandeur de la durée des verticalités. En effet, lors de la synchronisation des spectres sur les verticalités, la valeur de la médiane sur chaque composante spectrale dépend alors forte-

1. Voir figure 7.3

2. Cette normalisation sera présentée dans la partie 7.2, elle vise à se départir de la dépendance des descripteurs aux nuances.

ment de la phase de la variation de l’amplitude, qui varie d’une verticalité à l’autre. Si les battements sont plus rapides, la durée des verticalités permet d’homogénéiser les valeurs du spectre synchronisé, indépendamment de la phase de modulation d’amplitude de départ. La figure 7.5 montre l’évolution des descripteurs dans les mêmes conditions que sur la figure 7.4, mais avec un intervalle de demi-ton. La période des variations d’amplitude étant de 120 ms (8 Hz), soit vingt fois inférieure à la durée des verticalités, les battements n’affectent plus les valeurs des descripteurs harmoniques, qui restent constantes. Ils les affecteraient si les verticalités étaient de l’ordre de 100 ms, or dans ce cas elles deviendraient inférieures à la durée de la fenêtre d’analyse de *do* 3 et l’harmonie n’aurait pas le temps de s’établir, ce qui créerait les problèmes de détection présentés dans la partie 7.1.1.

Cet exemple montre qu’à moins de travailler sur de la musique microtonale, les variations d’amplitude dues aux battements seront rarement un problème dans le calcul des descripteurs harmoniques, tandis que les variations d’amplitude liées au timbre et au type de son utilisé sont plus problématiques. Dans la suite, nous veillerons systématiquement à utiliser les timbres qui minimisent cette variation, et quand elle sera tout de même présente, à ce que la durée des verticalités soient d’un ordre de grandeur plus grand.

Revenons un instant à la musique chorale. Bien que les voix d’un chœur présentent moins de fluctuations, tant en amplitude qu’en fréquence, que les voix opératiques, celle-ci sont invariablement présentes. De plus, comme les chanteurs ne peuvent chanter exactement à la même hauteur fondamentale, des phénomènes de battements similaires à ceux de la figure 7.4 apparaissent là-aussi invariablement. Ternström (Ternström, 2003) a montré que dans un chœur ces interférences s’homogénéisent pour créer l’« *effet chorus* ». Il en résulte que dans l’optique d’appliquer les descripteurs harmoniques à un répertoire choral, il conviendrait non pas d’isoler chaque chanteur, mais de considérer chaque pupitre isolément, c’est-à-dire que le spectre d’une note devrait être calculé à partir d’un enregistrement de tout le pupitre qui chante cette note. Il en est de même des cordes de l’orchestre, dans le cas d’une musique orchestrale.

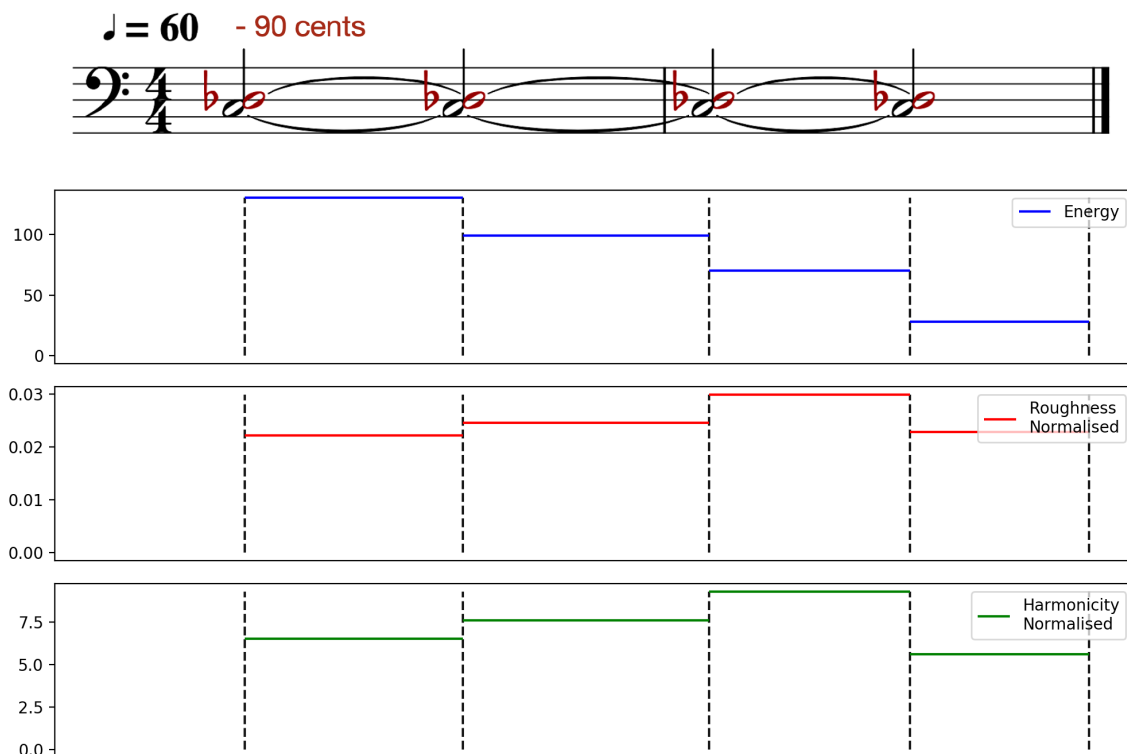


FIGURE 7.4 – Effet sur les descripteurs harmoniques de la modulation d’amplitude due aux battements, sur un micro-intervalle de 10 cents, avec les jeux *Bourdon + Flûte4*. En haut, l’évolution de l’énergie, au milieu, l’évolution de la rugosité et en bas l’évolution de l’harmonicité. Les variations sont causées par la modulation d’amplitude, elle-même due aux battements.

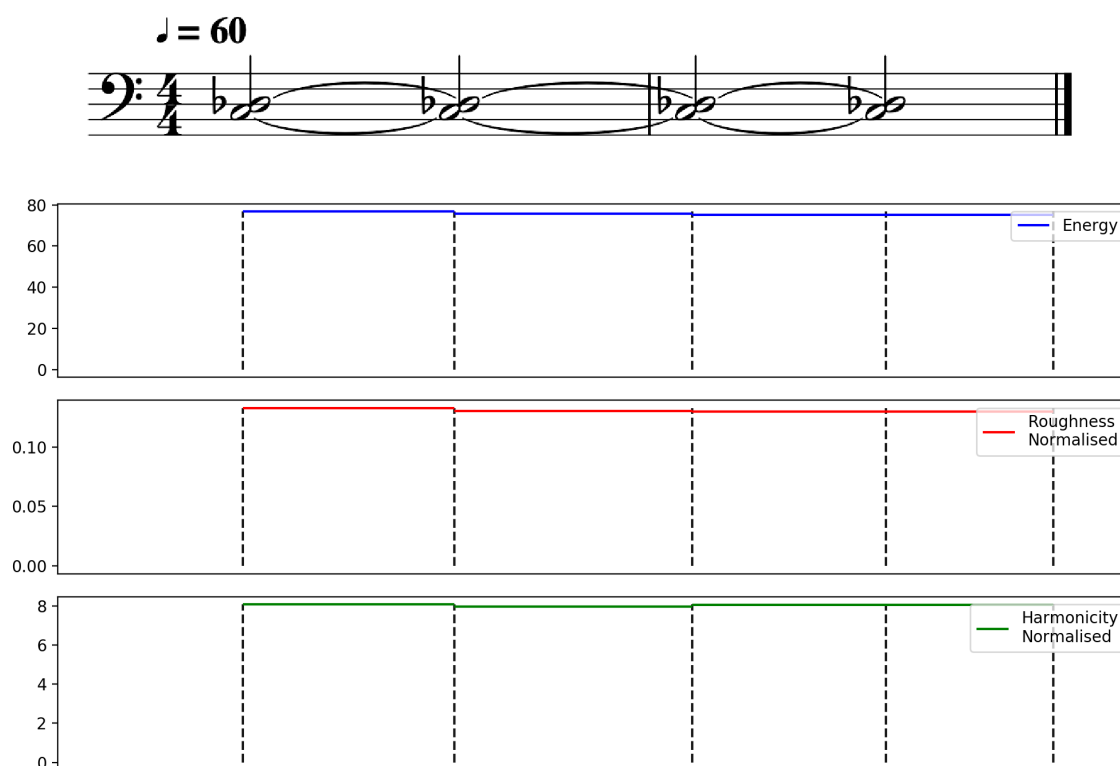


FIGURE 7.5 – Intervalle de demi-ton, avec les jeux *Bourdon + Flûte4*. En haut, l'évolution de l'énergie, au milieu, l'évolution de la rugosité et en bas l'évolution de l'harmonicité. Les battements sont d'un ordre de grandeur inférieur par rapport à la durée des verticalités, et ne perturbent pas les valeurs des descripteurs.

7.2 Comportement de descripteurs harmoniques vis-à-vis des nuances

Les descripteurs harmoniques, implémentés dans la partie 6, dépendent non seulement de la répartition spectrale des sons analysés, mais également de leur amplitude globale, liée à la pression sonore. Celle-ci détermine l'intensité sonore perçue, qui dépend également de la fréquence dans le cas de sons simples. La figure 7.6 représente les courbes d'isotonie des sons simples, c'est-à-dire les lieux d'intensité perçue égale, en fonction du niveau de pression acoustique, en dB SPL, et de la fréquence.

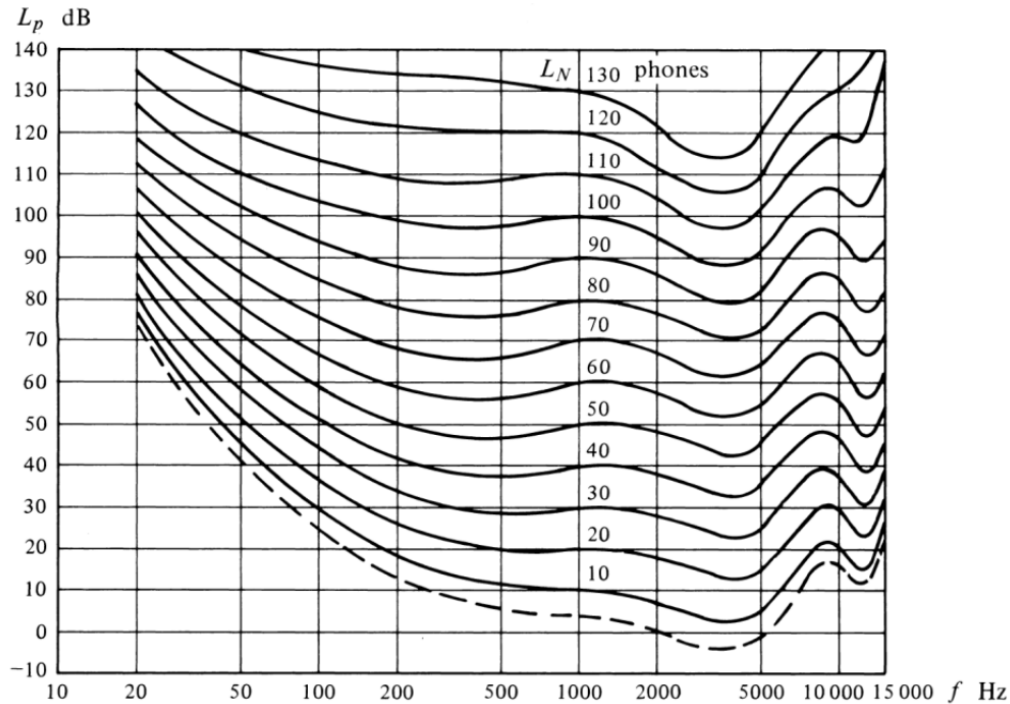


FIGURE 7.6 – Courbes d'isotonie des sons simples. Chaque courbe représente les lieux d'intensité perçue égale, en fonction du niveau de pression acoustique, en dB SPL, et de la fréquence. Tiré de (Rossi, 2007, p.127).

Dans notre approche, quand plusieurs partiels interagissent, les valeurs

2. https://fr.wikipedia.org/wiki/Courbe_isosonique

des descripteurs, que ceux-ci soient statiques ou dynamiques¹, sont proportionnelles à l'amplitude de chacun des partiels². Le tableau 7.1 montre l'effet de la multiplication du spectre en amplitude par α sur les valeurs des descripteurs harmoniques statiques, et le tableau 7.2 concerne les descripteurs dynamiques.

	Multiplication du spectre par α
Harmonicité	$\times \alpha$
Rugosité Concordance	$\times \alpha^2$
Concordance d'ordre 3 Tension	$\times \alpha^3$
Concordance totale	$\times \alpha^N$

TABLE 7.1 – Effet de la multiplication du spectre en amplitude par un facteur α sur les valeurs des descripteurs harmoniques statiques. N désigne le nombre de notes.

	Multiplication des deux spectres par α	Multiplication d'un des spectres par α
Changement harm.	$\times \alpha$	Indéterminé
Concordance diff. Rugosité diff.	$\times \alpha^2$	$\times \alpha$

TABLE 7.2 – Effet de la multiplication par un facteur α d'un spectre ou des deux spectres d'une transition sur les valeurs des descripteurs harmoniques dynamiques.

1. Mis à part le changement harmonique
2. Dans d'autres modèles (Sethares, 1999; Vassilakis, 2001), elles sont directement proportionnelles à l'intensité sonore en *sonies* de chaque partiels.

Dès lors, une variation des descripteurs peut être due à une variation qualitative des interactions, notamment lorsque qu'il y a une variation des accords ou des transitions, ou à une variation quantitative due à un changement de l'intensité sonore, autrement dit un changement de nuance.

7.2.1 Conditions d'indépendance vis-à-vis des nuances

On connaît la dépendance aux nuances de facteurs perceptifs comme la rugosité¹. En effet, l'amplitude des battements entre deux sons est d'autant plus grande que l'intensité des sons est importante. De plus, le tableau 7.1 montre que tous les descripteurs harmoniques statiques considérés, tels qu'ils ont été définis précédemment, augmentent avec l'intensité sonore du signal. Ainsi, étant donnés deux accords quelconques, on peut rendre à souhait l'un plus dissonant que l'autre et inversement, en jouant uniquement sur les nuances. Cela n'est pas très satisfaisant dans l'optique d'une étude des espaces d'accords, et implique en outre que la comparaison de deux accords, une fois le timbre fixé, n'est possible qu'à condition de fixer les nuances, ce qui peut paraître artificiel.

La question de la pertinence et de la légitimité de la dépendance aux nuances dans une approche essentiellement analytique et théorique se pose, et reste ouverte. Dans ce travail, nous faisons le choix de nous départir en partie de cette dépendance, afin de privilégier l'aspect qualitatif de l'harmonie. Pour ce faire, nous posons des conditions qui amènent à des normalisations énergétiques des descripteurs, fournissant une approche plus satisfaisante du point de vue théorique. On peut aussi faire le choix d'enlever ces normalisations, notamment pour l'étude de l'interprétation.

Condition d'indépendance pour les descripteurs statiques *Les valeurs statiques des descripteurs harmoniques ne doivent pas dépendre des nuances.*

Cette condition impose que pour tout descripteur harmonique statique $\mathcal{D}_{stat}$, tout spectre ou ensemble de spectres² X et tout facteur $\alpha > 0$, on a :

$$\mathcal{D}_{stat}(\alpha X) = \mathcal{D}_{stat}(X) \quad (7.1)$$

1. Voir (Kameoka et Kuriyagawa, 1969b).

2. Lorsque le calcul fait intervenir les spectres des différentes voix.

où αX indique la multiplication du ou des spectre(s) par le facteur α .

Autrement dit, deux accords sont équivalents du point de vue des descripteurs harmoniques statiques du moment que les rapports d'amplitude entre les partiels sont égaux.

Condition d'indépendance pour les descripteurs dynamiques *Les valeurs dynamiques des descripteurs harmoniques peuvent dépendre des nuances relatives des deux accords d'une transition, mais pas de leurs nuances globales.*

Autrement dit, deux transitions constituées des mêmes notes sont considérées équivalentes du point de vue des descripteurs harmoniques dynamiques lorsque le rapport des énergies des premiers accords et le rapport des énergies des seconds accords sont égaux. Cela s'exprime de la manière suivante, pour tout descripteur harmonique dynamique $\mathcal{D}_{dyn}$, tout couple de spectres X_1 et X_2 et tout facteur $\alpha > 0$:

$$\mathcal{D}_{dyn}(\alpha X_1 \rightarrow \alpha X_2) = \mathcal{D}_{dyn}(X_1 \rightarrow X_2). \quad (7.2)$$

Par contre $\mathcal{D}_{dyn}(\alpha X_1 \rightarrow X_2)$ n'est pas forcément égal à $\mathcal{D}_{dyn}(X_1 \rightarrow X_2)$.

Cette condition de dépendance aux nuances relatives des transitions est motivée à la fois par les études perceptives, qui montrent que les jugements d'affinité ou de tension de transition y sont sensibles ([Farbood, 2012](#); [Granot et Eitan, 2011](#); [Parncutt, 1989](#)), et par la pratique musicale dans laquelle elles sont cruciales.

7.2.2 Normalisation des descripteurs par l'énergie

Les descripteurs statiques La condition d'indépendance des descripteurs statiques aux nuances nécessite de normaliser les descripteurs par l'énergie, laquelle est définie sur la verticalité t comme le carré de la norme 2 du spectre :

$$\mathcal{E}(t) = \|X_t\|_2^2 = \sum_k X_{k,t}^2. \quad (7.3)$$

Cette normalisation vise à rendre les descripteurs sans dimension, en même temps qu'à vérifier la condition d'indépendance aux nuances. Elle dépend

des puissances du facteur α dans le tableau 7.1. Le tableau 7.3 indique l'effet de cette normalisation sur chacun des descripteurs statiques. Le choix de la norme 3 du spectre pour la concordance d'ordre 3 et de la norme N plutôt que $\mathcal{E}(t)^{N/2}$ pour la concordance totale sera justifié dans la partie 7.3.2.

	Effet de la normalisation énergétique - division par :
Harmonicité	$\sqrt{\mathcal{E}(t)}$
Rugosité Concordance	$\mathcal{E}(t)$
Concordance d'ordre 3 Tension	$\mathcal{E}(t)^{3/2}$ ou $\ X_t\ _3^3$
Concordance totale	$\mathcal{E}(t)^{N/2}$ ou $\ X_t\ _N^N$

TABLE 7.3 – Effet de la normalisation par l'énergie sur les descripteurs statiques. $\mathcal{E}(t)$ désigne l'énergie de la verticalité t , et N le nombre de notes.

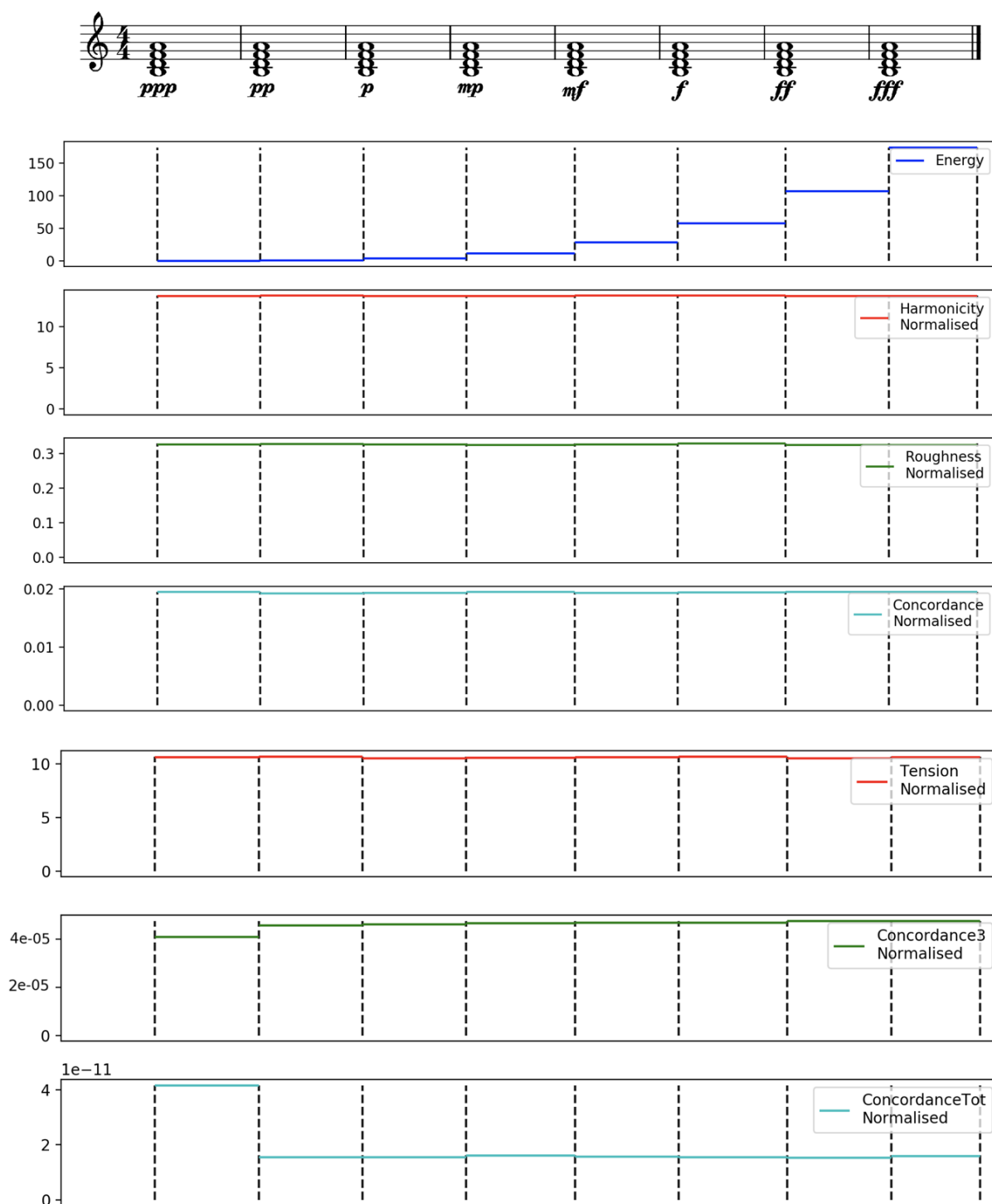


FIGURE 7.7 – Indépendance des descripteurs statiques relativement aux nuances. Synthétisé avec les jeux *Tutti*. Dans l'ordre : énergie, harmonicité, rugosité, concordance, tension, concordance d'ordre 3 et concordance totale.

La figure 7.7 montre les profils de descripteurs sur différentes présentations d'un même accord où l'on a fait varier les nuances en huit étapes, de *ppp* à *fff*. On observe bien une égalité des valeurs, mis à part pour la concordance d'ordre 3 et la concordance totale sur l'accord *ppp*. Cela est probablement dû à des artefacts du logiciel Muscore utilisé pour faire la synthèse sonore. Cependant, ces différences ne sont pas significatives par rapport aux variations des descripteurs concernés sur des accords différents¹.

Les descripteurs dynamiques De la même manière que pour les descripteurs statiques, la normalisation énergétique sur les transitions permet de rendre les descripteurs dynamiques sans dimension et de vérifier la condition d'indépendance aux nuances globales. Elle dépend des puissances du facteur α dans le tableau 7.2. Le tableau 7.4 indique l'effet de cette normalisation sur chacun des descripteurs statiques.

	Effet de la normalisation énergétique - division par :
Changement harm.	$(\mathcal{E}(t)\mathcal{E}(t+1))^{1/4}$
Rugosité diff.	$\sqrt{\mathcal{E}(t)\mathcal{E}(t+1)}$
Concordance diff.	

TABLE 7.4 – Effet de la normalisation par l'énergie sur les descripteurs dynamiques. $\mathcal{E}(t)$ et $\mathcal{E}(t+1)$ désignent l'énergie des verticalité t et $t+1$.

La figure 7.8 montre les profils de descripteurs dynamiques sur l'alternance de deux mêmes accords où l'on fait varier les nuances. L'égalité des descripteurs sur les transitions 2 et 3 vient du fait qu'ils sont temporellement symétriques, autrement dit l'ordre des accords dans une transition n'importe pas. L'égalité des première et dernière transitions traduit la condition d'indépendance aux nuances globales. La rugosité et la concordance différentielles sont indépendantes des nuances relatives des deux accords d'une transition, d'où leur constance sur toutes les transitions.

1. Voir figure 6.11.

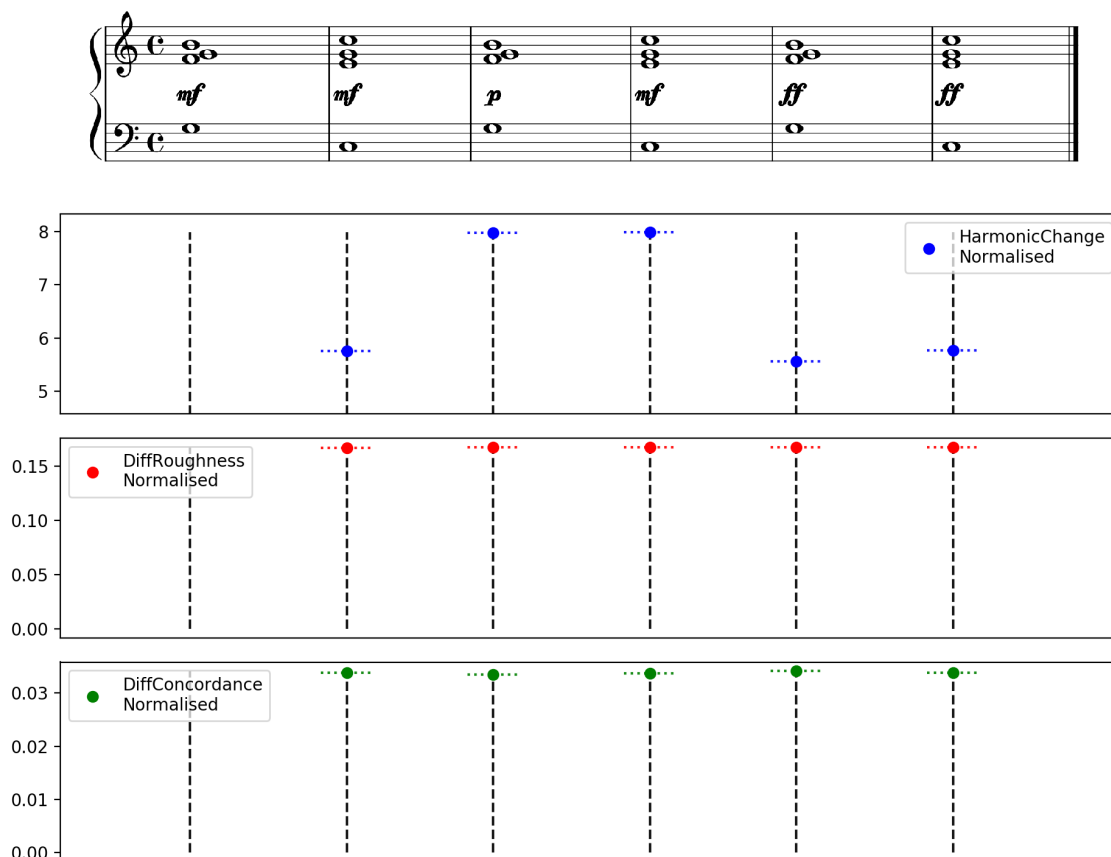


FIGURE 7.8 – Comportements des descripteurs dynamiques vis-à-vis des nuances. Une alternance des deux mêmes accords est présentée, avec des nuances différentes. Jeux utilisés : *Bourdon + Flûte4*.

Du fait de la condition d'indépendance des descripteurs statiques aux nuances, les descripteurs dynamiques triviaux¹ sont indépendants des nuances relatives et des nuances globales.

1. Autrement dit qui correspondent à l'évolution d'un descripteur statique, voir page 158.

7.2.3 Cas limite : les nuances par note

Les conditions précédentes d'indépendance des descripteurs harmoniques aux nuances ne concernent pas les nuances par note dans les accords. La figure 7.9 montre la sensibilité des descripteurs à la répartition des nuances sur les différentes notes d'un accord.

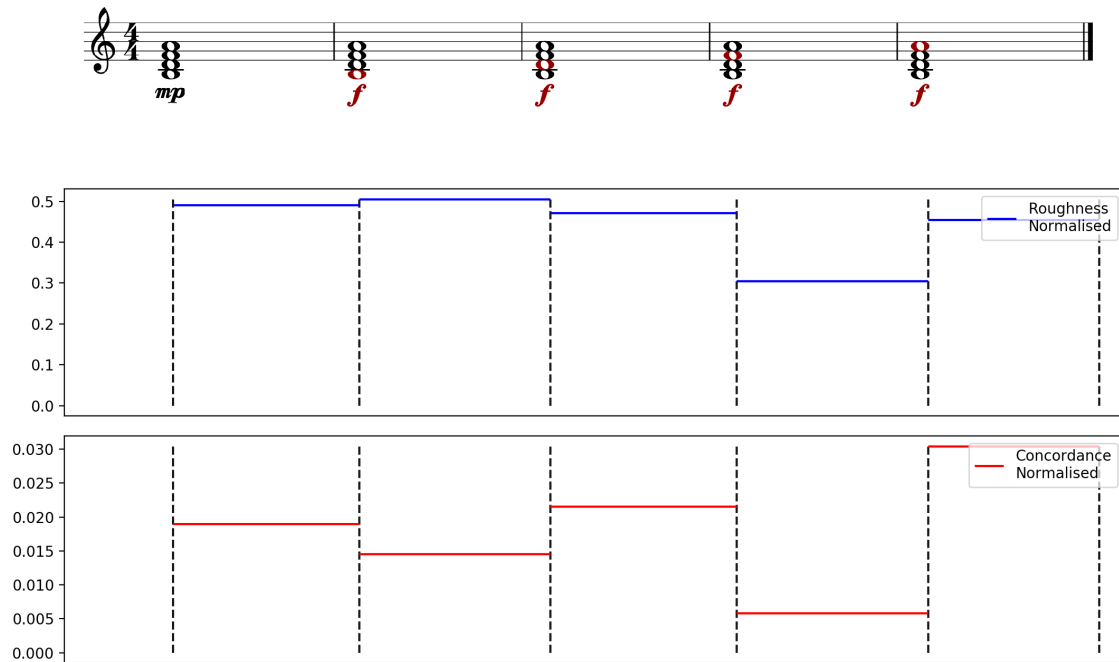


FIGURE 7.9 – Dépendance des descripteurs aux nuances des notes individuelles. Les notes en rouge sont jouées à la nuance *f*, les autres *mp*. Jeux utilisé : *Tutti*.

Cette sensibilité aux nuances par notes est une bonne nouvelle. Elle rejoint la sensibilité des descripteurs au timbre de chaque voix, lorsque des instruments différents jouent ensemble. Le lieu de cette rencontre est l'orchestration, puisque le travail de l'orchestrateur consiste à bien distribuer les timbres et les nuances sur l'ensemble des instruments et des possibilités sonores qui s'offrent à lui.

7.3 Influence du nombre de notes

7.3.1 Considérations perceptives et théoriques

Influence du nombre de notes sur la perception Revenons un moment à l'exemple initial donné dans le prologue de ce travail¹, qui concerne la perception d'un intervalle *do-si* auquel on rajoute une troisième note, le *mi* intermédiaire. L'ajout du *mi* fait augmenter la rugosité, ce qui devrait le rendre plus dissonant si l'on considère la rugosité comme principale cause de la dissonance. Cet argument est le plus souvent invoqué contre la théorie des battements d'Helmholtz². Cependant, dans l'état actuel de la recherche, plutôt que de mettre en défaut l'importance de la rugosité dans la perception des accords, il met en lumière la nécessité d'une approche complémentaire qui inclut, en plus de la rugosité, d'autres notions comme celles d'harmonicité ou de familiarité (Harrison et Pearce, 2020), dès lors qu'il s'agit de modéliser au plus près la perception.

Toutefois, la rugosité seule constitue un bon modèle de perception des intervalles, mais est insuffisante dès qu'elle est appliquée aux triades ou aux accords de plus de 3 notes (Cook, 2009; Johnson-Laird *et al.*, 2012). Johnson-Laird note :

*Un des problèmes majeurs de la rugosité est son inaptitude à modéliser la dissonance relative des triades*³.

Cet écart met en lumière des différences non pas uniquement quantitatives, mais également qualitatives dans la perception des intervalles et des triades, et plus généralement entre les intervalles et les accords de cardinalité supérieure. On sait par exemple que la présentation préalable d'un intervalle n'influence pas le jugement perceptif d'un accord cible qui le suit, alors que celui-ci est grandement influencé par la présentation préalable d'un accord de 4 notes (Lahdelma *et al.*, 2020), cela indépendamment de la rugosité des accords. Un accord de 3 ou 4 notes étant plus complexe acoustiquement qu'un intervalle, il aurait une empreinte cognitive plus forte que celui-ci. Il

1. Voir page 29.

2. Voir (Carcagno *et al.*, 2019) où sont exposés les autres arguments en défaveur de la théorie d'Helmholtz, parmi lesquels figurent la persistance de la sensation de dissonance en écoute dichotique ou la persistance de la sensation de consonance de la quinte dans le registre grave, quant bien même les battements sont importants.

3. "A critical problem with roughness is that it fails to predict the relative dissonance of the common triads", tiré de (Johnson-Laird *et al.*, 2012), traduction de l'auteur.

semblerait en outre que la familiarité et l'apprentissage jouent un rôle plus important dans la perception des triades que des intervalles ([Parncutt, 2006](#)), étant donné l'usage considérable des triades dans la musique tonale.

Cook, cependant, réfute la tendance à réduire la différence qualitative dans la perception des intervalles et des triades à une prépondérance du facteur culturel dans les derniers, même si celle-ci peut jouer un rôle. Il défend l'idée que de nouvelles dimensions perceptives propres aux triades apparaissent, qui se rapportent à des propriétés structurelles concernant les interactions de triplets de partiels. C'est le cas notamment de la tension¹, de la stabilité et de la modalité (caractère majeur ou mineur).

*Il est indéniable qu'il existe des propriétés structurelles des accords de 3 notes qui contribuent à la sensation de stabilité, et qui sont d'ordre acoustique, plus que culturel*².

L'échec des théories psychoacoustiques classiques à modéliser de manière satisfaisante la perception des accords de 3 notes serait due, selon lui, à l'importance excessive accordée aux interactions deux-à-deux entre partiels, à l'exclusion des interactions multiples impliquant simultanément plus de partiels³. Huron ([Huron, 1994](#)), conscient des limites de l'approche réductionniste⁴ dans la perception des accords à plus de deux notes, note d'ailleurs :

*Dans le cas des accords à trois ou quatre notes, il est important d'avoir à l'esprit que les mesures de consonances ne reflètent pas la consonance [au sens perceptif, N.D.T] d'un ensemble de notes qui sonnent simultanément (tel que la consonance d'une "triade majeure"). Elle reflète plutôt la consonance collective de toutes les paires possibles de notes tirées des notes initiales.*⁵

1. Implémentée dans ce travail, voir partie [6.1.5](#)

2. "There are clearly structural features of 3-tone chords that contribute to their overall stability and that have less to do with culture than with acoustics.", tiré de ([Cook, 2009](#)), traduction de l'auteur.

3. "This overemphasis on intervals to the complete exclusion of 3-tone psychophysics remains a problem in more modern attempts at explaining harmony perception .", tiré de ([Cook, 2009](#)).

4. On appelle approche réductionniste l'approche consistant à se ramener aux intervalles.

5. "In the case of these three- and four-note sets, it is important to recognize that the consonance measures do not reflect the consonance of the complete set of concurrently sounding tones (such as the consonance of 'a major triad'). Rather, the consonance values reflect the collective consonance of all possible pairs of tones drawn from the given pitch set.", tiré de ([Huron, 1994](#)), traduction de l'auteur.

Comparaison des accords de cardinalité différente Étant données les différences quantitatives et qualitatives dans la perception des accords de cardinalité différente, le problème se pose de tenir compte de cette cardinalité dans les modèles perceptifs. La plupart des expériences perceptives évitent le problème en fixant le nombre de notes, tendance qui s’est renversée récemment, avec la comparaison d’accords de 2 et 4 notes (Lahdelma *et al.*, 2020), 2 à 4 notes (Lahdelma et Eerola, 2020), 2 à 6 notes (Lahdelma et Eerola, 2016), 3 et 4 notes (Johnson-Laird *et al.*, 2012), 3 et 5 notes (Kuusi, 2010). Mis à part les deux études les plus récentes¹, ces travaux se concentrent sur différentes dimensions perceptives des accords sans que la cardinalité ne soit prise en compte comme paramètre et que son influence sur la perception ne soit explicitée. Plus précisément, ils affinent essentiellement la compréhension des corrélations entre les dimensions perceptives, sans proposer des modèles tenant compte directement de la cardinalité.

Modèles perceptifs tenant compte de la cardinalité Les modèles perceptifs tenant compte de la cardinalité des accords ne peuvent être de nature purement acoustique, étant donné que la note, et donc la cardinalité, sont des notions symboliques. Si ces modèles restent l’exception dans la littérature, on peut en citer trois :

- « *Gradus suavitatis* » Euler (Euler, 1739) considère que l’agrément d’un accord dépend du PPCM (plus petit commun multiple) des rapports de fréquences entre les notes, via une formule de la suavité qui fait intervenir sa décomposition en facteurs premiers². Cette formule a l’avantage de mettre sur les mêmes plans les accords avec des cardinalités différentes. Malheureusement pour Euler, son approche donne lieu à ce qui est considéré du point de vue de la perception comme des aberrations³.
- « **Consonance et dissonance tonales** » : **intégration de principes tonaux** Pour palier aux insuffisances des modèles purement

1. (Lahdelma et Eerola, 2020), qui montre que la rugosité, l’impureté, l’inharmonicité et la tension perçues augmentent globalement avec le nombre de notes, sans pour autant entraîner une baisse de la préférence, et (Lahdelma *et al.*, 2020), qui montre que les accords à 4 notes ont un effet d’amorçage, contrairement aux intervalles.

2. La suavité est présentée dans la partie 2.1.4.

3. Voir la note 2 page 89.

psychoacoustiques de perception des accords de 3 et 4 notes, Johnson-Laird ([Johnson-Laird et al., 2012](#)) intègre à un modèle de rugosité des considérations tonales, qui permettent de classer les accords selon leur appartenance ou non à des structures de la musique tonale (gamme majeure ou mineure, superposition de tierces ou de quintes). Ces principes d'organisation sont différents pour les triades ou les tetrachordes.

- « *Agregate Dyadic Consonance* » (ADC) Cette notion vient de Huron ([Huron, 1994](#)), qui associe une valeur de consonance aux objets abstraits que sont les classes d'accords (classes normales¹). Elle est calculée en associant d'abord une valeur de consonance à chaque classe d'intervalle, puis en sommant les valeurs de consonance sur toutes les classes d'intervalles constitutives de la classe d'accords. La consonance des classes d'intervalles est obtenue en moyennant les valeurs de consonance sur les deux réalisations de la classe d'intervalle inférieures à l'octave. En tant que telle, cette grandeur est extensive, et ne peut, du fait de la sommation, qu'augmenter avec l'ajout d'une note à une classe d'accords. Cependant, Kuusi ([Kuusi, 2010](#)) la rend intensive en moyennant la consonance sur toutes les classes d'intervalles qui entrent en jeu.

La démarche de Johnson-Laird, qui intègre dans une approche psychoacoustique des principes issus de la musique tonale, est représentative de la prise de conscience récente de la complémentarité des approches psychoacoustiques et culturelles. Les principes d'organisation tonale pris en compte semblent cependant relativement arbitraires et introduits de manière *ad hoc* pour expliquer les résultats expérimentaux, sans que la généralisation aux accords de plus de 4 notes ne soit assurée. L'approche consistant à moyennner les valeurs de la consonance sur tous les intervalles entrant dans le calcul, si elle est incomplète car n'inclut pas de considérations culturelles, est une manière plus satisfaisante théoriquement d'adapter le modèle de rugosité à la cardinalité des accords. Étonnamment, à la connaissance de l'auteur, l'article de Kuusi ([Kuusi, 2010](#)) est le seul exemple où un descripteur harmonique est obtenu par moyennage sur toutes les contributions des paires de notes.

Pourtant, cette démarche est très générale, et peut être appliquée aux accords réels, n'étant pas spécifiquement liée aux abstractions d'accords. En

1. Voir page [1.5.2](#)

prenant le rapport d'une grandeur extensive par une quantité, on la rend intensive. On passe ainsi de la consonance globale à la consonance par interaction spectrale en la divisant par le nombre d'interactions, de la même manière qu'en physique on passe de la masse m à la masse volumique μ en divisant par le volume. Perceptivement, ce moyennage est légitime du moment que la cardinalité représente une réalité perceptive, c'est-à-dire du moment qu'on entend le nombre de notes, question liée à celle de fusion tonale. En outre, ce moyennage constitue une réponse théorique, faute d'une justification rigoureusement perceptive, à l'exemple mentionné dans le prologue. En effet, dans le registre médium et avec des timbres raisonnables, la septième majeure est l'intervalle présentant la plus grande rugosité, après la seconde mineure¹. En ajoutant la tierce mi , deux nouveaux intervalles apparaissent, qui ont une rugosité plus faible que la septième majeure : la tierce majeure $do-mi$ et la quinte juste $mi-si$. Ainsi la rugosité moyenne baisse.

Les tableaux 7.5 et 7.6 montrent le classement des classes d'accords selon leur valeur d'ADC extensive, d'après (Huron, 1994), puis selon leur valeur d'ADC intensive, avec une subdivision de l'octave en 12 demi-tons. La consonance de classe extensive met en avant des échelles de notes, notamment la gamme pentatonique commune et la gamme majeure, tandis que la consonance de classe intensive met en avant les accords à cardinalité plus petite : la classe de la quarte et de la quinte, ou les accords mineurs et majeurs (qui ont le même contenu intervallique, donc la même valeur d'ADC).

1. Voir figure 1.3

Rang	Classement par ADC extensive	Cardinalité
1	Japanese Ritsu mode (CDEGAB)	6
2	Gamme pentatonique commune (CDEGA)	5
3	Gamme majeure	7
4	6-20* : mode demi-ton tierce mineure	6
5	Accord mineur avec septième mineure	4

TABLE 7.5 – Classement des 5 premières classes normales selon leur valeur d’ADC extensive, du plus consonant au moins consonant. Calculs tiré de (Huron, 1994).

Rang	Classement par ADC intensive	Cardinalité
1	Quarte / quinte	2
2	Triades majeures et mineures	3
3	3-9* : Accord de quintes superposées (CFG)	3
4	Tierce mineure / sixte majeure	2
5	Accord mineur avec septième mineure	4

TABLE 7.6 – Classement des 5 premières classes normales selon leur valeur d’ADC intensive, du plus consonant au moins consonant.

7.3.2 Normalisation des descripteurs sur les unissons

Les descripteurs statiques Une partie des descripteurs statiques présentés dans la partie 6.1 nécessitent le spectre des voix séparées, c’est le cas de la concordance, la concordance d’ordre k et la concordance totale. La tension et la concordance d’ordre 3, en décrivant les interactions entre 3 partiels, relèvent de ce que Cook appelle la « *psychophysique à 3 notes* »¹. La concordance d’ordre $k > 3$ et la concordance totale décrivent les interactions

1. Voir (Cook, 2009).

simultanées de plus de 3 partiels, et sont à ce titre particulièrement adaptées pour décrire des spécificités des accords de forte cardinalité. Mis à part ceux nécessitant les voix séparées, qui ne peuvent être calculés qu'à partir d'une certaine cardinalité, les descripteurs statiques peuvent être calculés sur des accords de toutes les cardinalités.

Dès lors, la question se pose d'imposer ou non des conditions sur les descripteurs harmoniques concernant le passage d'une cardinalité à l'autre. De même que nous avons posé des conditions d'indépendance aux nuances ², nous imposons une condition d'indépendance au nombre de notes, qui concerne les unissons.

Condition d'indépendance à la cardinalité sur les unissons *Les valeurs statiques des descripteurs harmoniques sur les unissons ne doivent pas dépendre du nombre de notes présentes, lorsque les notes ont le même spectre.*

La condition porte sur les unissons stricts, c'est-à-dire les verticalités avec une seule hauteur fondamentale, et non sur les doublures des accords.

Étant donné un unisson de N notes, l'ajout d'une $(N + 1)$ -ième note de même hauteur a deux conséquences :

1. Augmentation de la nuance.
2. Augmentation du nombre d'interactions prises en compte dans le calcul, pour les descripteurs calculés à partir des pistes séparées.

Le premier point, la nuance, a été déjà traité dans la partie 7.2. Ainsi, la condition d'indépendance à la cardinalité des unissons est déjà assurée par la normalisation par l'énergie pour les descripteurs harmoniques qui ne sont pas concernés par le second point, c'est-à-dire l'harmonicité, la tension, et la rugosité lorsque celle-ci est calculée à partir du signal global.

Quant au second point, il est dû au fait que les descripteurs concernés quantifient les interactions entre les notes, et non pas directement entre les partiels. Ainsi, dans un accord à N notes, le calcul de la concordance nécessite $\binom{N}{2}$ calculs, contre $\binom{N}{3}$ pour la concordance d'ordre 3. Dans le cas d'un unisson à N notes, ces calculs apportent la même contribution, à savoir le produit scalaire du spectre d'une note avec elle-même, ou son produit trinéaire pour la concordance d'ordre 3. Le fait que cette contribution soit

2. Voir partie 7.2.1.

comptée de multiples fois incite à diviser par le nombre de calculs, afin de se départir de la cardinalité. Cela ne saurait suffire cependant, en raison du terme de normalisation par l'énergie qui dépend lui aussi du nombre de notes. En effet, l'énergie de l'unisson à N notes est égale à N^2 fois l'énergie d'une seule note, sous la condition qu'elles soient toutes en phase, ce qui n'est pas forcément vérifié en pratique, mais que l'on supposera vérifié par commodité, et qui n'aura pas une grande influence sur les calculs, comme le montre la figure 7.10. Il reste alors à multiplier le tout par la bonne puissance de N .

Voici les corrections des équations de la partie 6.1 par les termes de normalisation par l'énergie et par le nombre de notes.

Concordance Le terme de droite de l'équation (6.10) est multiplié par le facteur $\frac{N^2}{\binom{N}{2}\mathcal{E}(t)}$, ce qui donne :

$$\begin{aligned}\mathcal{C}(t) &= \frac{N^2}{\binom{N}{2}\mathcal{E}(t)} \times \sum_{i < j \leq N} \sum_k X_{k,t}^{(i)} X_{k,t}^{(j)} \\ &= \frac{2N}{(N-1)\mathcal{E}(t)} \sum_{i < j \leq N} \sum_k X_{k,t}^{(i)} X_{k,t}^{(j)},\end{aligned}\tag{7.4}$$

où N désigne le nombre de notes, $\mathcal{E}(t)$ l'énergie de la verticalité t , et $X_t^{(i)}$ le spectre en amplitude de la i -ème note la verticalité t .

Sur les unissons, l'expression se simplifie et la concordance est égale à 1.

Rugosité De la même manière que la concordance, l'expression de la rugosité (équation (6.5)), lorsqu'elle est calculée à partir des pistes séparées, devient :

$$\begin{aligned}\mathcal{R}(t) &= \frac{N^2}{\binom{N}{2}\mathcal{E}(t)} \times \sum_{i < j \leq N} \sum_{k,l} \mathcal{R}(X_{k,t}^{(i)}, X_{l,t}^{(j)}, f_k, f_l) \\ &= \frac{2N}{(N-1)\mathcal{E}(t)} \sum_{i < j \leq N} \sum_{k,l} \mathcal{R}(X_{k,t}^{(i)}, X_{l,t}^{(j)}, f_k, f_l).\end{aligned}\tag{7.5}$$

Contrairement à la concordance, la rugosité n'égale pas 1 sur les unissons, et sa valeur diffère selon la hauteur et le timbre de l'unisson.

Concordance d'ordre 3 et concordance d'ordre supérieur Pour la concordance d'ordre m , le terme qui vient normaliser les nuances est exprimé selon la norme m du spectre, plutôt que selon la norme 2 qui apparaît dans l'expression de l'énergie. Ce choix vise à rendre les valeurs de concordance d'ordre m constantes sur tous les unissons et égales à 1.

La concordance d'ordre 3 est multipliée par $\frac{N^3}{\binom{N}{3}\|X_t\|_3^3}$, et la concordance d'ordre m par $\frac{N^m}{\binom{N}{m}\|X_t\|_m^m}$.

L'équation (6.13) devient ainsi :

$$\mathcal{C}_m(t) = \frac{N^m}{\binom{N}{m}\|X_t\|_m^m} \times \sum_{i_1 < i_2 < \dots < i_m \leq N} \langle X_t^{(i_1)}, \dots, X_t^{(i_m)} \rangle, \quad (7.6)$$

où $\|X_t\|_m^m = \underbrace{\langle X_t, \dots, X_t \rangle}_{m \text{ fois}}$

Concordance totale La concordance totale d'une verticalité à N notes s'exprime avec la norme N , l'équation (6.12) devenant :

$$\mathcal{C}_{Tot}(t) = \frac{N^N}{\|X_t\|_N^N} \langle X_t^{(1)}, \dots, X_t^{(N)} \rangle. \quad (7.7)$$

La concordance totale coïncide bien avec la concordance ayant pour ordre la cardinalité, et est égale à 1 sur les unissons. La formule (7.7) invite à prolonger en 1 la concordance totale sur les verticalités à une note, considérées comme des unissons à une note.

La figure 7.10 montre l'évolution des descripteurs statiques sur les unissons, avec les normalisations relatives aux conditions sur les nuances et au nombre de notes sur les unissons.

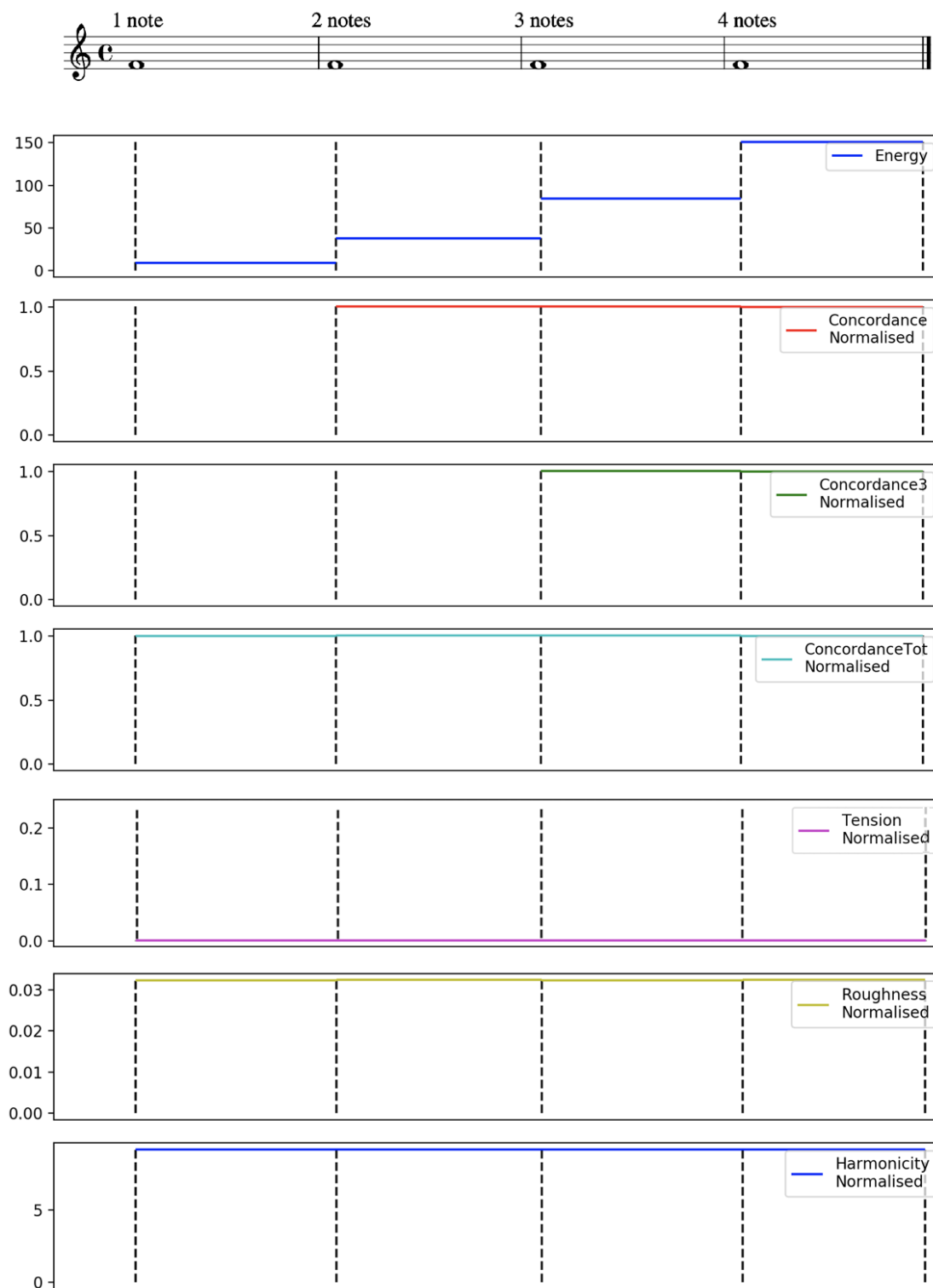


FIGURE 7.10 – Indépendance des descripteurs statiques au nombre de notes sur les unissons. Unissons de 1, 2, 3 et 4 notes, synthétisés avec les jeux *Bourdon + Flûte4*. Dans l'ordre : énergie, concordance, concordance d'ordre 3, concordance totale, tension, rugosité et harmonicité

Comme prévu, la concordance, la concordance d'ordre 3 et la concordance totale sont constantes et égales à 1 sur les unissons. La rugosité et l'harmonicité sont constantes. Les valeurs de tension sont quant à elle négligeables par rapport aux valeurs calculées sur les triades de la figure 6.17, dont nous avons repris ici l'échelle. Avec un spectre de note plus riche, la tension devient plus significative, tout en restant constante, comme le montre la figure 7.11, calculée avec les jeux *Tutti*.

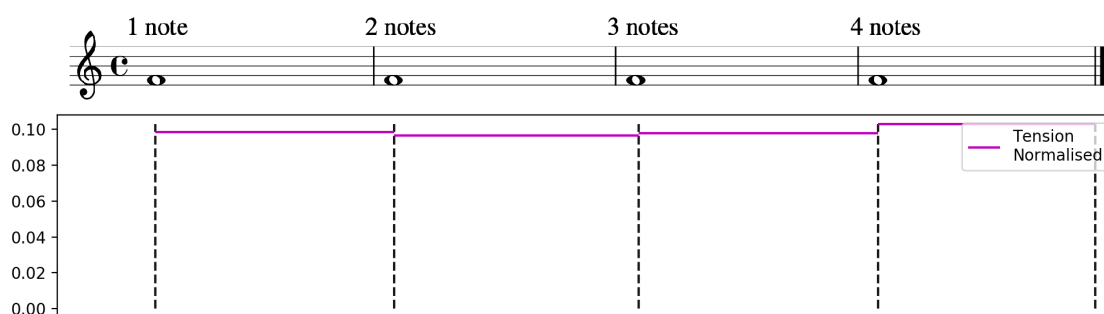


FIGURE 7.11 – Évolution de la tension sur les unissons, avec les jeux *Tutti*. La tension est plus stable qu’avec les jeux *Bourdon + Flûte4*.

Descripteurs dynamiques Les descripteurs dynamiques considérés ne sont pas calculés à partir des voix séparées, comme c’est le cas de certains descripteurs statiques. De ce fait, lorsque dans une transition l’une des verticalités est un unisson, l’ajout d’une note supplémentaire à cet unisson n’a pour effet que d’augmenter la nuance de la verticalité. Le cas des nuances ayant déjà été traité dans la partie 7.2, on sait que la concordance et la rugosité différentielles restent inchangées, contrairement au changement harmonique. Nous n’imposons pas pour les descripteurs dynamiques de condition supplémentaire relative au nombre de notes.

La figure 7.12 illustre l’indépendance de la concordance et de la rugosité différentielles à la cardinalité sur les transitions d’unissons.

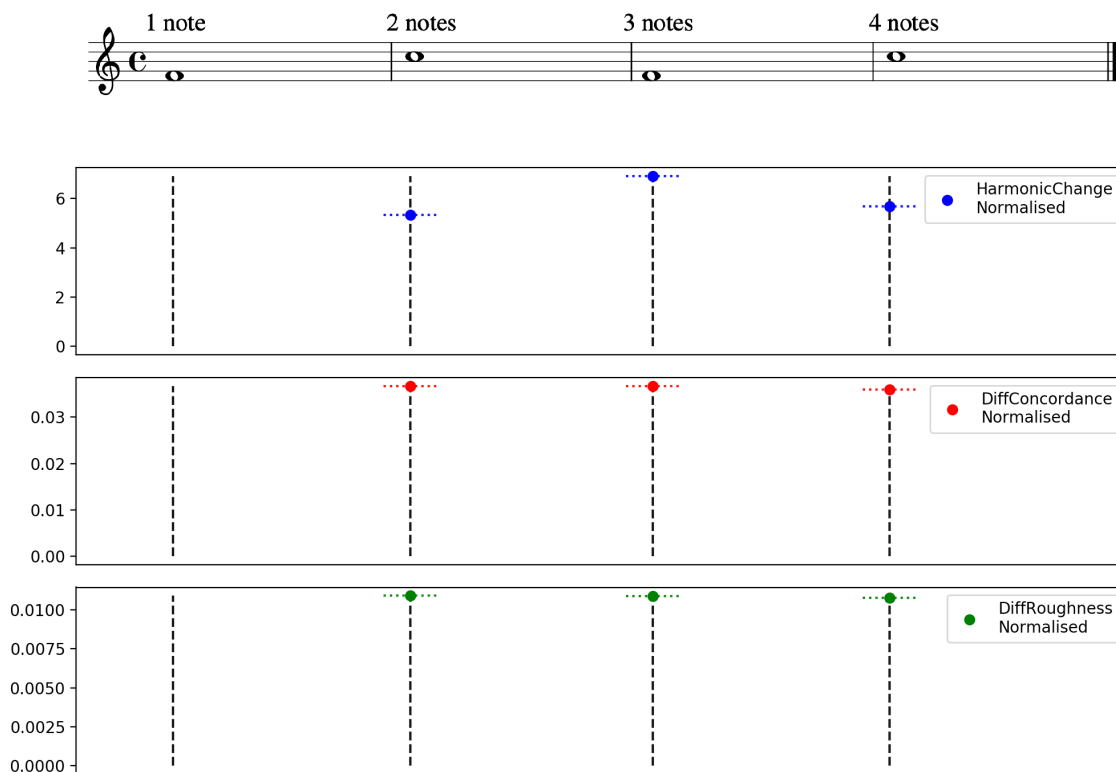


FIGURE 7.12 – Évolution des descripteurs dynamiques sur les transitions d’unissons. Synthétisé avec le spectre *Tutti*. La concordance et la rugosité différentielles restent indépendantes du nombre de notes.

7.3.3 Cas limite : unisson au sein d’un accord

La condition d’indépendance des descripteurs statiques à la cardinalité sur les unissons ne concerne que le cas précis des unissons stricts, c’est-à-dire les verticalités avec une seule hauteur fondamentale, et non les unissons au sein d’autres notes, comme l’illustre la figure 7.13, où chacune des notes d’un accord majeur est doublée successivement.

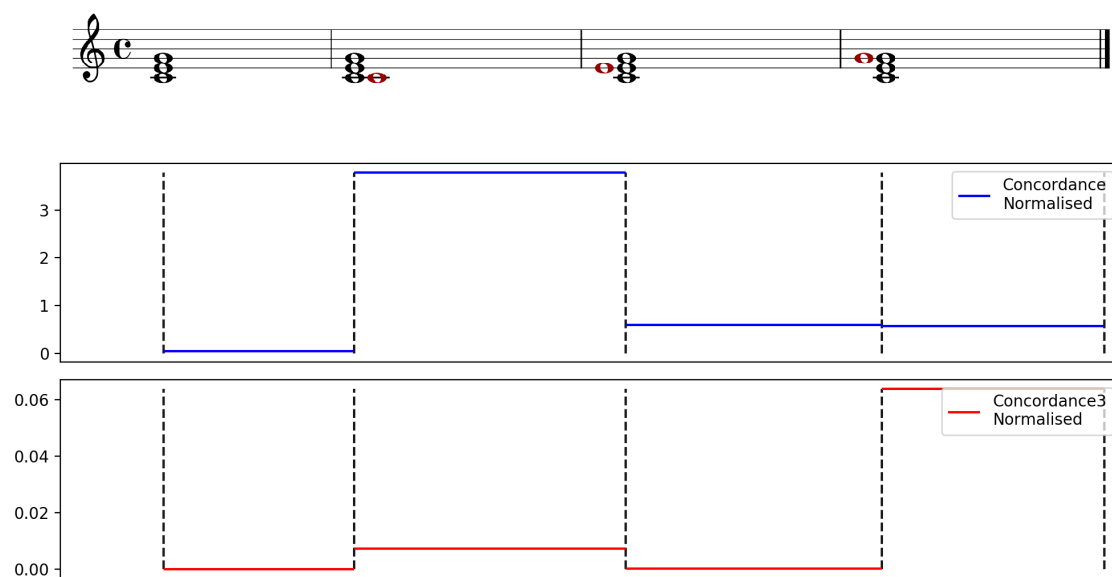


FIGURE 7.13 – Dépendance des descripteurs statiques aux doublures de notes.

Troisième partie

Exploration de l'univers
harmonique

Chapitre 8

Les espaces continus d'accords et de transitions

Les descripteurs harmoniques différentient et hiérarchisent les accords et les transitions. Ils révèlent ainsi les structures acoustiques des espaces harmoniques. Plusieurs types d'espaces d'accords et de transitions peuvent être considérés :

- Ensemble des accords ou transitions présents dans une œuvre musicale ;
- Ensemble des accords ou transitions dans une échelle fixée (par exemple dans l'échelle chromatique) ;
- Ensemble des accords ou transitions liés à un système musical (par exemple les accords classés du système tonal) ;
- Ensemble des accords ou transitions sans restriction en matière de système musical ni d'échelle.

Dans ce chapitre, nous traitons les espaces d'accords sans restriction en matière de système musical ni d'échelle, ce qui nécessite de considérer l'axe fréquentiel comme continu. Nous proposons des visualisations de ces espaces et de leur structuration par les descripteurs harmoniques.

8.1 Structuration des dyades

8.1.1 Histoire des classements des intervalles

Approche catégorielle Historiquement, les intervalles étaient classés par les théoriciens de la musique de manière catégorielle selon leur degré de consonance et de dissonance, lesquelles ont pris des significations différentes selon l'époque¹.

Avant l'avènement de la polyphonie, le classement des intervalles dépend de la simplicité des ratios des fréquences, et l'on distingue deux catégories : les consonances et les dissonances. Avec l'ars antiqua, en même temps que la polyphonie se développe, les classements se complexifient, et des théoriciens comme Jean de Garlande ([Garlande, 1250](#)) distinguent six catégories : les consonances parfaites, les consonances intermédiaires, les consonances imparfaites, les dissonances imparfaites, les dissonances intermédiaires et les dissonances parfaites. Le développement du contrepoint simplifie le classement pour le réduire à trois catégories : les consonances parfaites, les consonances imparfaites et les dissonances. Des théoriciens comme Prosdocius de Beldemandis ([de Beldemandi, 1412](#)) sont représentatifs de cette simplification.

La figure 8.1, tirée de ([Tenney, 1988](#), p.109), synthétise l'évolution des systèmes de catégorisation des intervalles dans la théorie musicale du *IX^e* jusqu'au *XVI^e* s, avec un grand nombre de jalons qui sont autant de traités. On observe une évolution à la fois qualitative (qui concerne les catégories utilisées) et quantitative (sur le contenu de ces catégories). Aussi, d'une classification à l'autre, le sens attribué à chaque catégorie peut-il varier, ainsi que le tempérament duquel elles sont issues. Tenney fait d'ailleurs l'hypothèse que la diversité des classements des intervalles intermédiaires tient en partie à la variabilité des tempéraments.

1. Voir ([Tenney, 1988](#)), dans lequel l'auteur distingue cinq conceptions différentes de la consonance et de la dissonance.

	1,8	5	4	M3	m3	M6	m6	M2	m7	M7	m2	T
1. <i>Musica</i> (and <i>Scholia</i>) <i>enchiriadis</i> (anonymous, 9th-10th c.)	C	C	C	(D)	(D)	(D)	(D)	(D)	(D)	(D)	(D)	D
2. <i>De Harmonica institutione</i> (ca. 900), Hucbald	C	C	C	(D)	(D)	(D)	(D)	(D)	(D)	(D)	(D)	(D)
3. <i>Micrologus</i> (1026-28), Guido d'Arezzo	C	C	C	(D)	(D)	(D)	(D)	(D)	(D)	(D)	(D)	(D)
4. <i>Ad organum faciendum</i> (anonymous, 11th-12th c.)	C	C	C	(D)	(D)	(D)	(D)	(D)	(D)	(D)	(D)	(D)
5. <i>Item de organo</i> (anonymous, 12th c.)	C	C	C	(C)	(C)	(D)	(D)	(D)	(D)	(D)	(D)	(D)
6. <i>Montpellier organum treatise</i> (anonymous, 12th c.)	C	C	C	(C)	(C)	(C)	(C)	(D)	(D)	(D)	(D)	(D)
7. <i>De musica libellus</i> (ca. 1220), Anon. VII (CS I)	pC	mC	mC	iC	iC	(D)	(D)	(D)	(D)	(D)	(D)	(D)
8. <i>De mensurabili musice</i> (ca. 1250) John of Garland	pC	mC	mC	iC	iC	iD	mD	mD	iD	pD	pD	pD
9. <i>Ars cantus mensurabilis</i> (ca. 1260), Franco of Cologne	pC	mC	mC	iC	iC	iD	pD	iD	iD	pD	pD	pD
10. <i>De mensuris et discantu</i> (ca. 1275), Anon. IV (CS I)	pC	mC	mC	iC	iC	D	(D)	(D)	(D)	(D)	(D)	(D)
11. <i>Tractatus de consonantiis musicalibus</i> (late 13th c.), Anon. I (CS I)	pC	mC	mC	iC	iC	iD	pD	iD	iD	pD	pD	pD
12. <i>Tractatus de discantu</i> (late 13th c.), Anon. II (CS I)	pC	mC	mC	iC	iC	iC	D	D	(D)	(D)	D	D
13. <i>De speculatione musicae</i> (ca. 1300), Walter Odington	pC	pC	ⁱ _p C	iC	iC	iC	(D)	(D)	(D)	(D)	(D)	(D)
14. <i>Speculum musicae</i> (ca. 1330), Jacobus of Liège	pC	pC	ⁱ _p C	ⁱ _m C	ⁱ _m C	iC	^m _i D	^m _i C	iC	^m _i D	^m _p D	^m _i p
15. <i>Tractatus de cantu perfecto et imperfecto</i> (14th c.), Henrici de Zelandia (CS III)	pC	mC	D	iC	iC	iC	D	D	D	D	D	D
16. <i>Quatuor principalia musicae</i> (1351), pseudo-Tunsted (CS IV)	pC	pC	iD	iC	iC	iC	iD	iD	pD	pD	pD	pD
17. <i>Ars contrapuncti</i> (late 14th c.), "secundum" Johannes de Muris (CS III)	pC	pC	(D)	iC	iC	iC	(D)	(D)	(D)	(D)	(D)	(D)
18. <i>Ars discantus</i> (late 14th c.), "secundum" Johannes de Muris (CS III)	pC	pC	(D)	iC	iC	iC	iC	(D)	(D)	(D)	(D)	(D)
19. <i>Tractatus de discantu</i> (late 14th c.), Anon. XIII (CS III)	pC	pC	D	iC	iC	iC	iC	D	D	D	D	D
20. <i>Ars contrapunctus</i> (late 14th c.), "secundum" Philippe de Vitry (CS III)	pC	pC	D	iC	iC	iC	iC	D	D	D	D	D
21. <i>Tractatus de contrapuncto</i> (1412), Prosdocius de Beldemandis	pC	pC	(i)D	iC	iC	iC	iC	D	D	D	D	D
22. <i>Compendium cantus figurati</i> (15th c.), Anon. XII (CS III)	pC	pC	(iC?)	iC	iC	iC	iC	D	D	D	D	(D?)
23. <i>Regilae supra contrapunctum</i> , Johannes Hothby (d. 1487)	pC	pC	(D)	iC	iC	iC	iC	(D)	(D)	(D)	(D)	(D)
24. <i>Liber de arte contrapuncti</i> (1477), Johannes Tinctoris	pC	pC	pC/D	iC	iC	iC	iC	D	D	D	D	D
25. <i>De praeceptis artis musicae...</i> (1480-90), Guilielmus Monachus	pC	pC	(D)	iC	iC	iC	iC	D	D	D	D	D
26. <i>Practica musicae</i> (1496), Franchinus Gafurius	pC	pC	(i)D	iC	iC	iC	iC	D	D	D	D	D
27. <i>Tetrachordum musices</i> (1511), Johannes Cochlaeus	pC	pC	D	iC	iC	iC	iC	D	D	D	D	D
28. <i>Isagoge in musicen</i> (1516), Henry Glarean	pC	pC	(D)	iC	iC	iC	iC	D	D	D	D	D
29. <i>Toscanello in musica</i> (1523), Pietro Aaron	pC	pC	C/D	iC	iC	iC	iC	D	D	D	D	D
30. <i>Le istitutioni harmoniche</i> (1558), Gioseffe Zarlino	pC	pC	pC/D	iC	iC	iC	iC	D	D	D	D	D

FIGURE 8.1 – Évolution des système de classification des intervalles, IX^e - XVI^e s, extrait de (Tenney, 1988, p.109). M = majeur, m = mineur, T = triton, C = consonance, D = dissonance, p = parfaite, m = intermédiaire, i = imparfaite

Approche continue Ce n'est qu'avec Helmholtz et l'avènement de l'approche psychoacoustique qu'est envisagée une organisation du continuum des intervalles, via une caractérisation acoustique de ceux-ci. Les courbes de rugosité d'Helmholtz¹ en constituent le premier exemple. Le continuum des fréquences y est figuré sur un axe, au-dessus duquel sont représentées les valeurs de rugosité des intervalles pour des sons contenant 1 à 9 partiels, avec des bosses et des vallées, les premières correspondant aux maximums locaux de rugosité, et les seconds aux minimums locaux, comme la quarte ou la quinte.

L'approche continue s'est vite imposée, puisque trois ans après la publication de la *Théorie physiologique de la musique*, Ernst Mach ([Mach, 1866](#)) écrit un texte de vulgarisation de la théorie d'Helmholtz où il reprend sa courbe de rugosité, cette fois sur deux octaves, en la faisant correspondre à un clavier de piano, les touches étant centrées sur leur fréquence. Cette représentation a l'avantage de mettre en évidence la correspondance entre les approches discrètes, qui prévalaient jusqu'alors lorsqu'il s'agissait de classer ou de caractériser les intervalles, et l'approche continue, qui va de pair avec les modèles acoustiques et perceptifs.

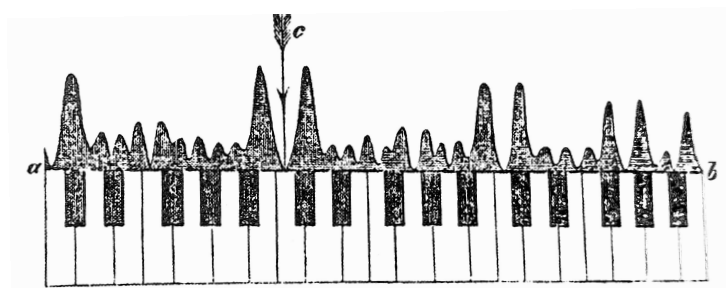


FIGURE 8.2 – Courbe de rugosité des intervalles, sur l'ambitus de deux octaves, tiré de ([Mach, 1866](#)) et reproduit dans ([Bailhache et al., 2011](#), p.139).

8.1.2 Courbes de descripteurs harmoniques

Dans la continuité du modèle de dissonance d'Helmholtz, une partie des descripteurs harmoniques statiques présentés dans le chapitre 6 peuvent être calculés sur le continuum des intervalles. C'est le cas de la rugosité, de la

1. Voir figure [1.3](#) page 45

concordance et de l'harmonicit . La tension et la concordance d'ordre 3 impliquent au minimum trois notes.

L'approche symbolique¹ est particuli rement adapt e au calcul des descripteurs sur des  chelles continues de fr quence, du fait qu'elle permet de contourner l'audio. Nous adoptons cette approche ici, avec le mod le de spectre en pics d crit dans la partie 5.2.1.

Courbes de rugosit  La figure 8.3 pr sente les courbes de rugosit  des intervalles sur l'ambitus d'une octave, avec *do* 5 comme note inf rieure, et respectivement 1, 3, 5, 7 et 9 partiels. Elle est   mettre en parall le des courbes de dissonance d'Helmholtz page 45, calcul es sous les m mes conditions.

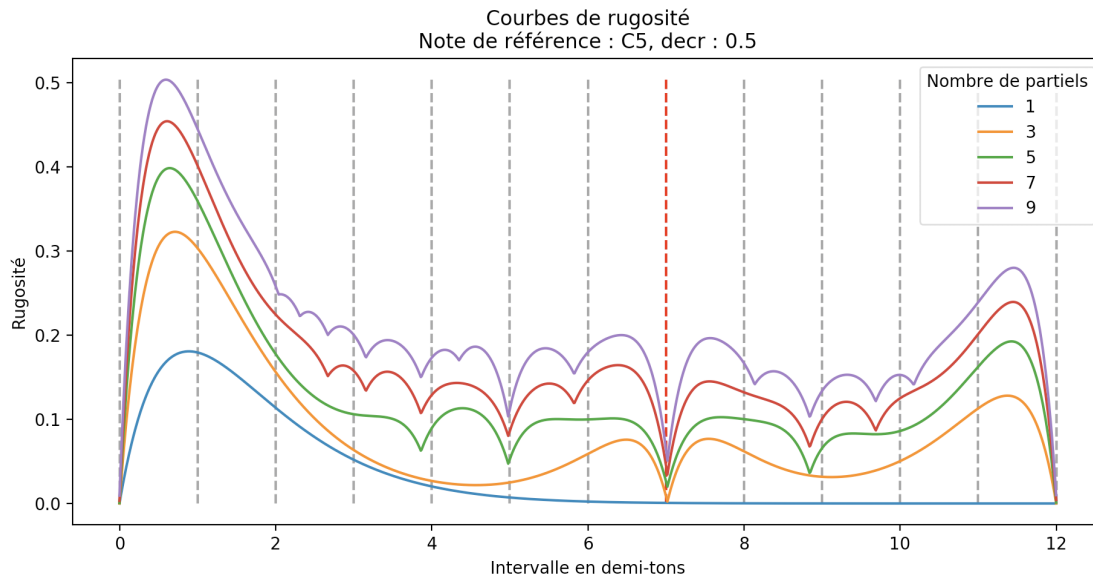


FIGURE 8.3 – Courbes de rugosit  des intervalles couvrant l'octave de *do* 5   *do* 6, avec respectivement 1, 3, 5, 7 et 9 partiels. En rouge est indiqu  l'intervalle de quinte. Param tres de spectre : $\sigma = 0.01$ et $decr = 0.5$.

La courbe   un partiel a d j   t  introduite². Sur les autres courbes, les interactions entre partiels d'ordre sup rieur cr ent des d pressions d'abord au niveau de l'octave, puis de la quinte, de la quarte, et ainsi de suite. Les

1. Voir partie 5.2

2. Voir figure 6.1 et 6.2.

intervalles des dépressions sont les intervalles justes, c'est-à-dire les intervalles correspondant aux ratios simples de fréquences. On peut visualiser la différence fréquentielle avec les intervalles tempérés, représentés par les traits verticaux. Notamment, la tierce majeure juste (de rapport $5/4$) est plus grave que la tierce majeure tempérée, alors que la tierce mineure juste (de rapport $6/5$) est plus aiguë que son homologue tempéré.

Les courbes de rugosité dépendent de la hauteur de la note inférieure¹, comme le montre la figure 8.4, où elle a été choisie parmi *do* 3, *do* 4 et *do* 5. La variation de la hauteur de la note ne modifie pas les positions des minimums locaux de rugosité, étant donné que ceux-ci correspondent aux ratios simples de fréquences, mais modifient en revanche la position des maximums locaux.

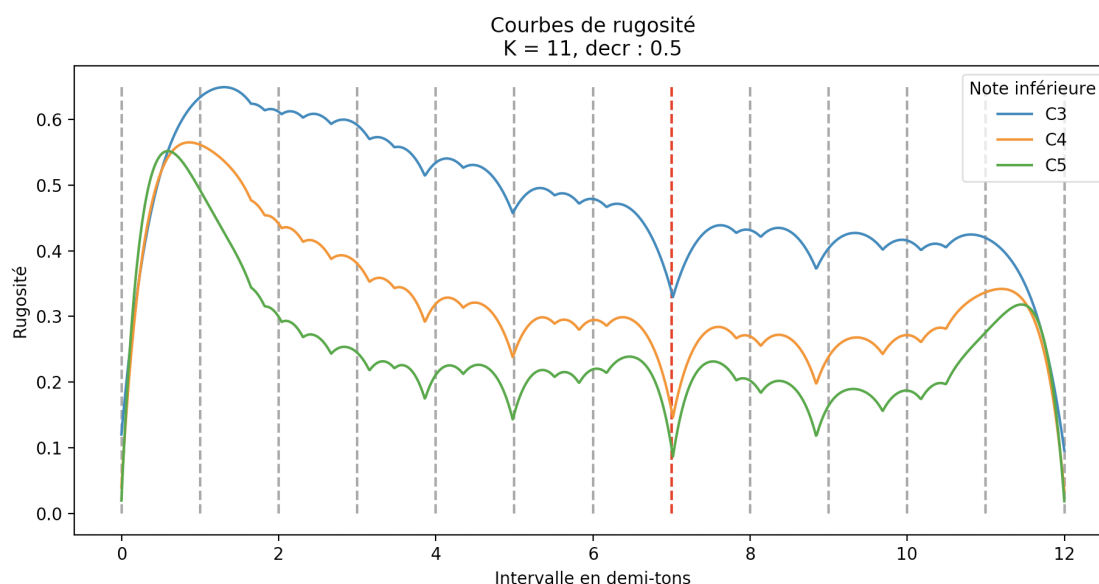


FIGURE 8.4 – Courbes de rugosité des intervalles sur une octave, avec pour note inférieure respectivement *do* 3, *do* 4 et *do* 5. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.01$ et $decr = 0.5$.

Les courbes de rugosité, prises comme courbes de dissonance, ont ceci de remarquable qu'elles peuvent être interprétées métaphoriquement comme des courbes d'énergie potentielle. En mécanique classique, un corps soumis à des forces conservatives (par exemple la pesanteur) a une énergie potentielle

1. Voir équation (6.3) et figure 6.2.

dépendant de sa position. Les minimums locaux de l'énergie potentielle caractérisent les équilibres stables (pendule au repos), les maximums locaux les équilibres instables (pendule à 180°). La stabilité est liée au mouvement : un pendule au repos ne bougera pas, alors qu'un pendule à 180° , sans vitesse, pour peu qu'on le perturbe légèrement, tombera d'un côté ou de l'autre. Aussi, un corps situé sur un maximum local d'énergie potentielle aura-t-il tendance à aller vers un minimum local d'énergie potentielle. De même, la dissonance, dès la CDC 3, implique une notion de mouvement, qui est alors le résultat de la conduite des voix. Avec Rameau, une dissonance est sommée de se résoudre¹, contrairement à une consonance, et c'est précisément sa qualité de dissonance qui lui impose le mouvement. L'association consonance-stabilité et dissonance-instabilité se renforce encore au siècle suivant².

Courbes de concordance De même qu'avec la rugosité, on peut tracer des courbes de concordance.

La figure 8.5 montre la courbe de concordance des intervalles sur deux octaves, avec un spectre comprenant 11 partiels. Les rapports simples de fréquences sont cette fois caractérisés par les maximums locaux. L'intervalle d'octave + quinte est plus concordant que la quinte. En effet, les harmoniques 3 et 1, qui se superposent dans le cas de l'octave + quinte, de rapport 3/1 apportent une contribution plus importante que les harmoniques 3 et 2 de la quinte, de rapport 3/2.

1. Voir page 98.

2. « Au milieu du XIX^e siècle, l'association de la consonance et de la dissonance avec le "phénomène de mouvement" a atteint une telle autonomie dans l'esprit de certains théoriciens, qu'elle leur semblait la seule piste valable pour définir ces notions. » Tiré de (Tenney, 1988, p.80) "By the middle of the 19th century, this association with 'consonance' and 'dissonance' with 'phenomena of motion' had attained such an autonomy in the mind of some theorists that it could seem to them the only valid basis for the definition of these terms." Traduction de l'auteur.

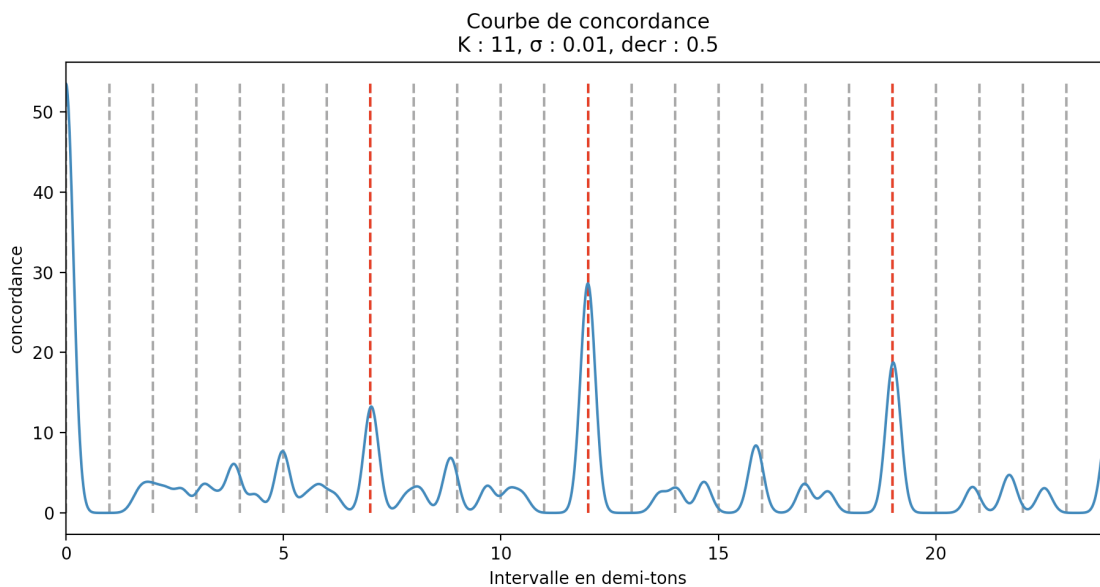


FIGURE 8.5 – Courbe de concordance sur deux octaves, avec les paramètres de spectre $K = 11$, $\sigma = 0.01$ et *decr* = 0.5. En rouge, la quinte, l’octave, et l’intervalle d’octave + quinte.

Le paramètre spectral σ , lié à la largeur des gaussiennes des partiels, influence directement la largeur des bosses, et donc la séparabilité des différents pics, comme le montre la figure 8.6, où σ a été divisé du tiers. Sur cette figure, tous les pics de concordance sont distinguables. Le paramètre σ a été interprété précédemment comme traduisant l’aspect probabiliste de la perception des hauteurs¹. La dépendance de la courbe de concordance des intervalles à σ invite à reformuler cette interprétation, où σ traduirait cette fois l’aspect probabiliste dans la perception d’un certain caractère de la sonorité des accords, en agissant directement sur la précision de la perception et la capacité à différencier les sonorités des accords.

1. Voir page 170.

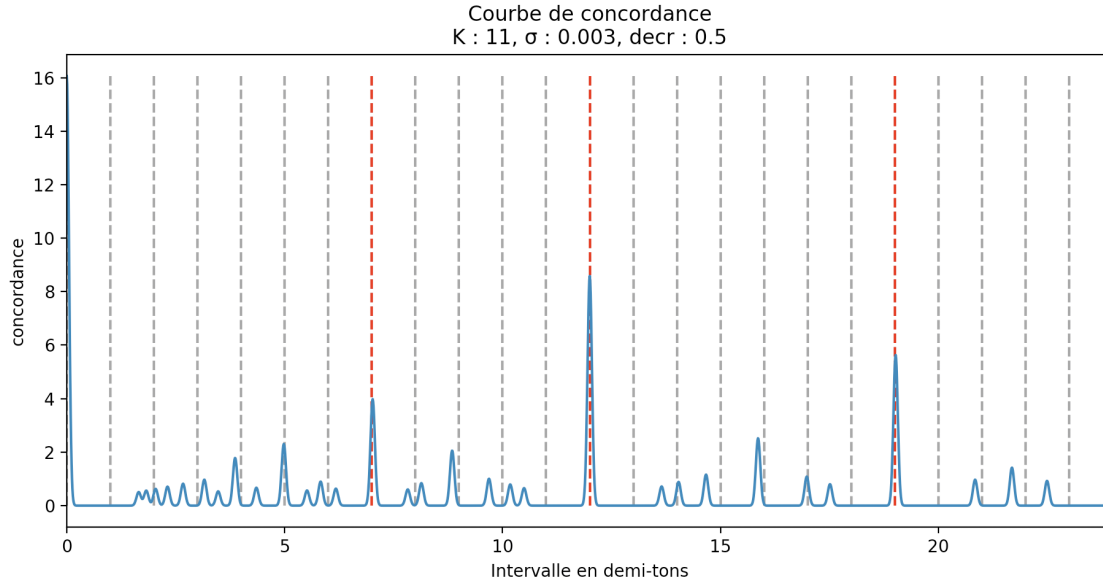


FIGURE 8.6 – Courbe de concordance sur deux octaves, avec les paramètres de spectre $K = 11$, $\sigma = 0.003$ et *decr* = 0.5. En rouge, la quinte, l’octave, et l’intervalle d’octave + quinte.

Courbe d’harmonicit  La figure 8.7 repr sente l’harmonicit  des intervalles inf rieurs   l’octave, avec 11 partiels. La courbe obtenue est tr s proche, au redimensionnement pr s,   la courbe de concordance¹, les maximums locaux correspondant aux rapports simples de fr quences. Comme pour la concordance, la s parabilit  des pics est directement li e au param tre σ .

1. Voir figure 8.5.

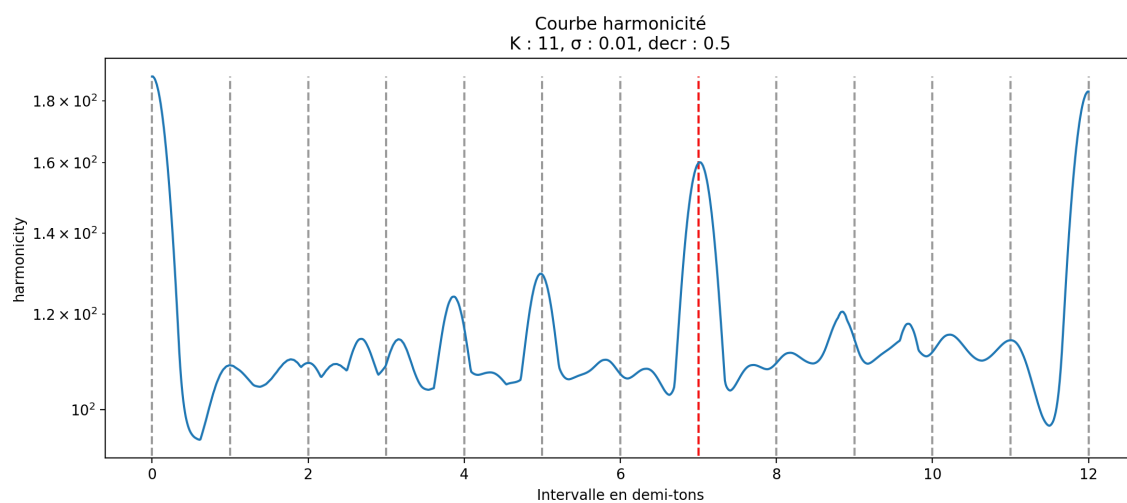


FIGURE 8.7 – Courbe d'harmonicit  sur une octave, avec les param tres de spectre $K = 11$, $\sigma = 0.01$ et *decr* = 0.5. L' chelle est logarithmique en ordonn es. La quinte est indiqu e en rouge.

8.2 Structuration des triades

8.2.1 Classements et représentations des triades

Dyades et triades : différences de statut Les dyades et les triades, si on les regroupe dans ce travail sous l'appellation d'accords, ont historiquement un statut très différent, à la fois dans la pratique musicale et a fortiori dans la théorie musicale.

Tout d'abord, pendant l'époque qui précède la polyphonie, la notion de consonance et de dissonance a un caractère avant tout mélodique, elle traduit l'affinité entre deux notes jouées successivement¹. Les termes de *concordia*, *discordia* ou *harmonia* désignent alors les intervalles en tant que succession de deux notes. Les premières polyphonies occidentales, avec l'organum et le déchant, sont d'abord à deux voix avant d'être à trois voix, et les théoriciens commencent à caractériser et classer les intervalles simultanés avant les triades.

Outre ces considérations historiques s'ajoutent des considérations combinatoires et perceptives qui distinguent fondamentalement les dyades des triades. Tout d'abord, dans une échelle à 12 notes, les triades potentiellement jouables sur une octave avec la note inférieure fixe sont au nombre de 55, contre 12 pour les intervalles, ce qui constitue une différence quantitative non négligeable.

De plus, en ce qui concerne la perception, les triades ont une empreinte cognitive plus forte que les dyades, et il semblerait que la familiarité et l'apprentissage jouent un rôle plus important dans la perception des premières que des secondes (Parncutt, 2006). Cook (Cook, 2009) considère en outre que des dimensions perceptives nouvelles apparaissent lorsqu'on passe de 2 à 3 notes.

Classements acoustiques et perceptifs des triades Pour les raisons à la fois historiques, perceptives et combinatoires citées précédemment, les triades n'ont pas été l'objet, contrairement aux intervalles, d'un classement ni d'une catégorisation systématique à travers l'histoire.

À ces raisons s'ajoute l'approche réductionniste qui consiste à considérer l'intervalle comme la brique de base de l'harmonie, dont l'étude, à la fois perceptive et théorique, suffirait pour comprendre les accords de cardinalité

1. Voir (Dahlhaus, 1967) et (Tenney, 1988).

plus élevée. Cette approche reste majoritaire dans la littérature perceptive depuis Helmholtz, malgré quelques propositions, pour reprendre l'expression de Cook, d'une « *psychophysique à 3 notes* » ¹.

Nous présentons ici différentes approche pour classifier les triades.

Helmholtz Dans sa *Théorie physiologique de la musique*, Helmholtz, après l'étude des battements et des sons résultants des intervalles, écrit à propos des accords :

Pour qu'un accord puisse être consonant, il est clair que les sons qui s'y trouvent doivent être deux à deux consonants ; car si deux d'entre eux seulement forment une dissonance et donnent des battements, l'harmonie de l'accord est détruite. ²

Cette condition est nécessaire et suffisante pour Helmholtz, qui réduit la consonance des triades à la consonance simultanée de ses trois intervalles constitutifs. L'étude des battements des intervalles l'a amené précédemment à distinguer 7 intervalles consonants, qui sont, dans l'ordre de consonance :

1. La quinte juste, de rapport $\frac{3}{2}$
2. La quarte juste, de rapport $\frac{4}{3}$
3. La sixte majeure, de rapport $\frac{5}{3}$
4. La tierce majeure, de rapport $\frac{5}{4}$
5. La tierce mineure, de rapport $\frac{6}{5}$
6. La sixte mineure, de rapport $\frac{8}{5}$
7. La septième mineure, de rapport $\frac{7}{4}$

À partir de ces intervalles consonants, Helmholtz se sert du tableau de la figure 8.8 pour déterminer les triades consonantes. Dans ce tableau, Helmholtz synthétise les accords formés d'intervalles consonants par rapport à la note inférieure, condition nécessaire à la consonance de la triade. Le troisième intervalle de la triade est indiqué à l'intersection des deux premiers intervalles.

1. Voir (Cook, 2009)

2. Tiré de (Helmholtz, 1863, p.272)

<i>ut</i> = 1	<i>sol</i> $\frac{3}{2}$	<i>fa</i> $\frac{4}{3}$	<i>la</i> $\frac{5}{3}$	<i>mi</i> $\frac{5}{4}$	<i>mi♭</i> $\frac{6}{5}$	<i>la♭</i> $\frac{8}{5}$
<i>sol</i> $\frac{3}{2}$						
<i>fa</i> $\frac{4}{3}$	Seconde majeure. $\frac{9}{8}$					
<i>la</i> $\frac{5}{3}$	Seconde majeure. $\frac{10}{9}$	<i>Tierce majeure.</i> $\frac{5}{4}$				
<i>mi</i> $\frac{5}{4}$	<i>Tierce mineure.</i> $\frac{6}{5}$	Seconde mineure. $\frac{16}{15}$	<i>Quarte.</i> $\frac{4}{3}$			
<i>mi♭</i> $\frac{6}{5}$	<i>Tierce majeure.</i> $\frac{5}{4}$	Seconde majeure. $\frac{10}{9}$	Quarte augmentée. $\frac{25}{18}$	Seconde mineure. $\frac{25}{24}$		
<i>la♭</i> $\frac{8}{5}$	Seconde mineure. $\frac{16}{15}$	<i>Tierce mineure.</i> $\frac{6}{5}$	Seconde mineure. $\frac{25}{24}$	Quarte diminuée. $\frac{32}{25}$	<i>Quarte.</i> $\frac{4}{3}$	
<i>si♭</i> $\frac{7}{4}$	<i>Tierce diminuée.</i> $\frac{7}{6}$	Fausse quarte. $\frac{21}{16}$	Seconde mineure. $\frac{21}{20}$	Quinte diminuée. $\frac{7}{5}$	Fausse quinte. $\frac{35}{24}$	Seconde majeure. $\frac{35}{32}$

FIGURE 8.8 – Table des triades, tirée de (Helmholtz, 1863, p.273). Ce tableau sert à déterminer les consonances parmi les triades. Les accords sont formés de *do*, d’une deuxième note dans la colonne de gauche, et d’une troisième note dans la ligne du haut. À l’intersection est indiqué l’intervalle entre les deuxième et troisième notes. Les intervalles consonants sont notés en italique.

Helmholtz en déduit les triades consonnantes au sein de l’octave, qui sont au nombre de 6, à savoir les accords majeur et mineur dans leur position fondamentale, leur premier et leur second renversement. Toutes les autres triades sont considérées comme dissonantes.

Expériences perceptives et modèles acoustiques Rompant avec des critères purement réductionnistes, les classements des triades après Helmholtz se veulent de nature perceptive, en s’occupant de leur sonorité globale.

Dans le tableau de la figure 8.9, Cook (Cook, 2009) synthétise le classement des triades communes dans la musique tonale avec leurs renversements, selon des expériences perceptives d’abord, puis d’après des modèles psychoacoustiques ensuite. Le propos de l’auteur est de montrer l’insuffisance des modèles de rugosité, performants sur les intervalles, à traduire les aspects de la sonorité globale des triades, ainsi que de montrer la supériorité à ce niveau-là de son modèle psychoacoustique, qui combine la rugosité et la tension. Les disparités du classement selon la théorie d’Helmholtz avec le propre classement d’Helmholtz¹ vient du fait que Cook, ici, ne tient pas compte des condition imposées par Helmholtz sur la consonance des triades.

Chord Class	Interval Structure	Empirical Sonority			Theoretical Sonority Based on Interval Consonance					Present Model
		Roberts (1986)	Eberlein (1994)	CFK (2007)	Helmholtz (1877/1954)	P&L (1965)	K&K (1969)	Parncutt (1989)	Sethares (1999)	
Major	4-3	1	1	1	3	4	1	1	4	1
	3-5	3	3	4	<u>9</u>	<u>11</u>	<u>11</u>	6	<u>8</u>	5
	5-4	2	7	2	1	2	6	3	2	4
Minor	3-4	4	2	3	3	4	1	4	4	2
	4-5	6	4	5	1	2	6	6	2	3
	5-3	5	6	6	<u>9</u>	<u>11</u>	<u>11</u>	<u>10</u>	<u>8</u>	6
Dim	3-3	7	9	7	13	13	6	9	12	12
	3-6	8	5	12	11	8	9	5	10	7
	6-3	9	10	10	11	8	9	8	10	10
Aug	4-4	10	11	13	<u>5</u>	10	13	<u>2</u>	12	13
Sus4	5-2	—	8	9	<u>5</u>	<u>6</u>	<u>1</u>	—	<u>6</u>	8
	2-5	—	11	8	<u>5</u>	<u>6</u>	<u>1</u>	—	<u>6</u>	11
	5-5	—	11	11	<u>8</u>	<u>1</u>	<u>1</u>	—	<u>1</u>	9

Note: CFK: Cook, Fujisawa, & Konaka, 2007; C&F: Cook & Fujisawa, 2006; K&K: Kameoka & Kuriyagawa, 1969; P&L: Plomp & Levelt, 1965.

FIGURE 8.9 – Tableau comparatif des classements des triades communes à travers trois lots de données empiriques et six modèles perceptifs. Le critère global donné par Cook est la *consonance totale*, ou *sonorité globale*. Les données empiriques sont constituées de deux expériences perceptives sur les triades (Roberts, 1986; Cook, 2009) et d’une étude des proportions relatives des triades dans un corpus de musique classique allant de Bach à Mendelssohn (Eberlein, 1994). Les modèles théoriques comptent cinq modèles de rugosité (Helmholtz, 1863; Plomp et Levelt, 1965; Kameoka et Kuriyagawa, 1969b; Parncutt, 1989; Sethares, 1999), et un modèle incluant la rugosité et la tension statique (Cook, 2009). Tiré de (Cook, 2009).

1. Voir page précédente.

Plus récemment, Johnson-Laird ([Johnson-Laird et al., 2012](#)) a proposé un modèle de dissonance des triades qui inclut des aspects tonaux aux considérations purement psychoacoustiques. En d’autres termes, il avance un modèle de dissonance musicale, et non plus uniquement de dissonance sensorielle¹. Les principes tonaux pris en compte sont au nombre de deux, et peuvent être déduits au niveau de la partition, car ne dépendent que des relations entre les fréquences fondamentales. Les voici, par ordre d’importance :

1. Les triades dont les notes sont incluses dans une gamme majeure sont plus consonantes que celles dont les notes sont incluses dans une gamme mineure, elles-mêmes plus consonantes que les accords dont les notes n’appartiennent ni à une gamme majeure, ni à une gamme mineure.
2. Les triades sont plus consonantes si elles sont formées (modulo l’octave) par une superposition de tierces ou de quintes.

Ces principes permettent de classifier les triades en six groupes d’accords en fonction de leur dissonance tonale, du plus dissonant au plus consonant. La théorie psychoacoustique vient discriminer les accords au sein de chaque groupe, créant ainsi un classement interne.

La figure [8.10](#) montre l’ensemble des 55 triades (à disposition et transposition près), classées par groupe d’accords selon les deux principes tonaux évoqués plus haut, du plus consonant au plus dissonant. Les groupes sont séparés par des doubles barres.

1. Voir page [34](#) pour la distinction entre *sensory dissonance* et *musical dissonance*.



FIGURE 8.10 – Classement des triades, tiré de (Johnson-Laird *et al.*, 2012). Les doubles barres divisent les accords en six groupes, en fonction de leur dissonance tonale déterminée par les principes tonaux sus-mentionnés. Sont également indiquées les valeurs calculées de rugosité, ainsi que l'évaluation perceptive des triades.

Représentation de l’espaces des triades Dans son texte de vulgarisation de Helmholtz, Mach ([Mach, 1866](#)) reprend la table des accords de la figure 8.8 en la faisant correspondre aux touches du clavier, avec un code de signes pour indiquer les intervalles et accords concordants.

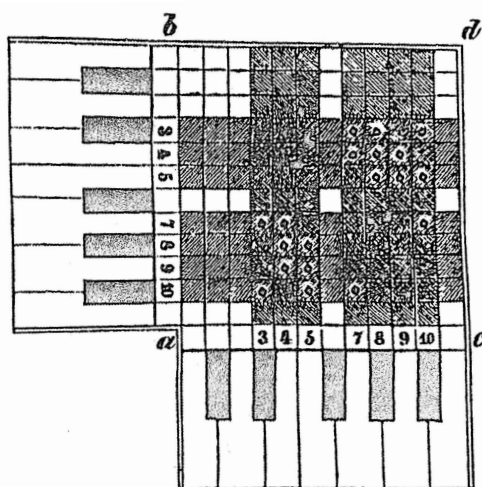


FIGURE 8.11 – Espace des triades, tiré de ([Mach, 1866](#)) et reproduit dans ([Bailhache et al., 2011](#), p.141), avec correspondances avec les touches de piano. Les lignes et colonnes grisées correspondent aux intervalles consonants, et les points aux accords potentiellement consonants, dont seulement 6 le sont réellement en tempérament juste (Voir figure 8.8).

De cette représentation à une représentation continue des triades, il n’y a qu’un pas, qui consiste à réduire les subdivisions des notes en passant à la limite. Comparons avec la surface de dissonance de Sethares ([Sethares, 1999](#), p.128 et 129), sur la figure 8.12. À gauche sont représentées les lignes de niveau de la rugosité des triades, et à droite la même surface en trois dimensions. Comme pour la figure 8.11, les deux intervalles lisibles indiqués sur les axes sont comptés par rapport à la note la plus basse.

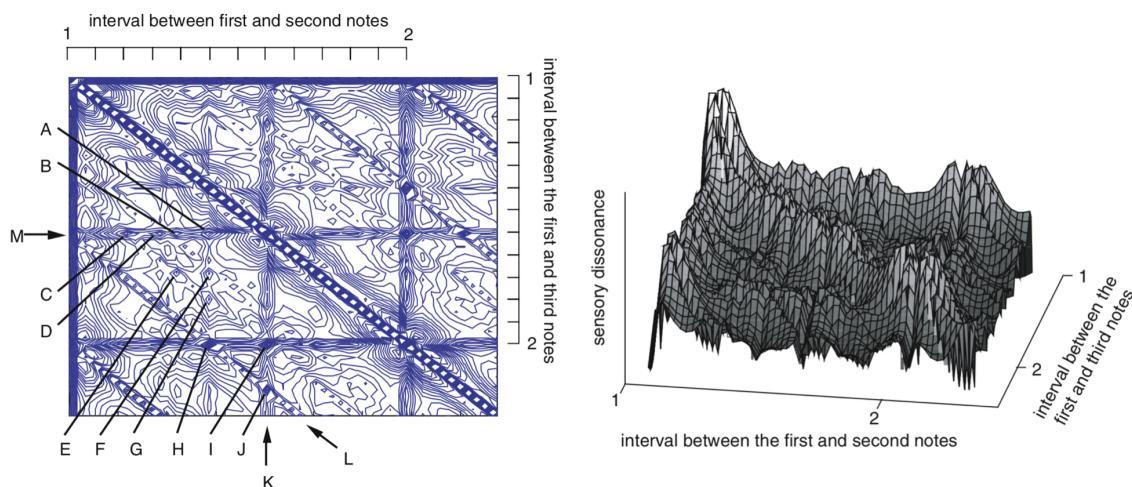


FIGURE 8.12 – Représentation plate et représentation en 3 dimensions d’une même surface de dissonance sensorielle avec 6 harmoniques, tiré de (Sethares, 1999, p.128 et 129).

Via cette représentation, l’ensemble des triades est assimilable à un quart de plan. La rugosité vient donner du relief à ce plan, relief pouvant être visualisé par des lignes de niveau, un code couleur ou une représentation en trois dimensions.

Des représentations similaires sont présentes dans (Cook, 2009), où l’auteur construit des courbes de dissonance, de tension et de modalité (aspect majeur-mineur).

8.2.2 Topographie de l’espace des triades

Plan des triades Pour représenter les surfaces de descripteurs harmoniques sur les espaces de triades nous utilisons un espace plan légèrement différent de celui utilisé par Mach et Sethares. L’axe des abscisses représente toujours l’intervalle entre la première et la deuxième note, mais l’axe des ordonnées, dirigé vers le haut, se réfère cette fois à l’intervalle entre la deuxième et la troisième note, et non plus entre la première et la troisième. Cela permet notamment d’éviter les redondances du fait de la symétrie diagonale de la première représentation.

La figure 8.13 montre le plan des triades, avec comme note inférieure *do* 4. La diagonale noire en pointillés figure le lieu des triades s’étendant sur une

octave. Les accords majeurs et mineurs de la gamme tempérée sont indiqués dans leur position fondamentale et leurs renversements, en position serrée ainsi que dans une position plus espacée qui dépasse l'octave. Les accords majeurs sont indiqués en noir, les accords mineurs en bleu.

On remarque que les six accords de chaque disposition (serrée ou espacée) forment un hexagone. L'échange des intervalles supérieur et inférieur, réalisé par la symétrie diagonale dont l'axe passe par l'origine (trait continu gris sur la figure), a pour effet de transformer un accord majeur en un accord mineur et inversement un accord mineur en un accord majeur.

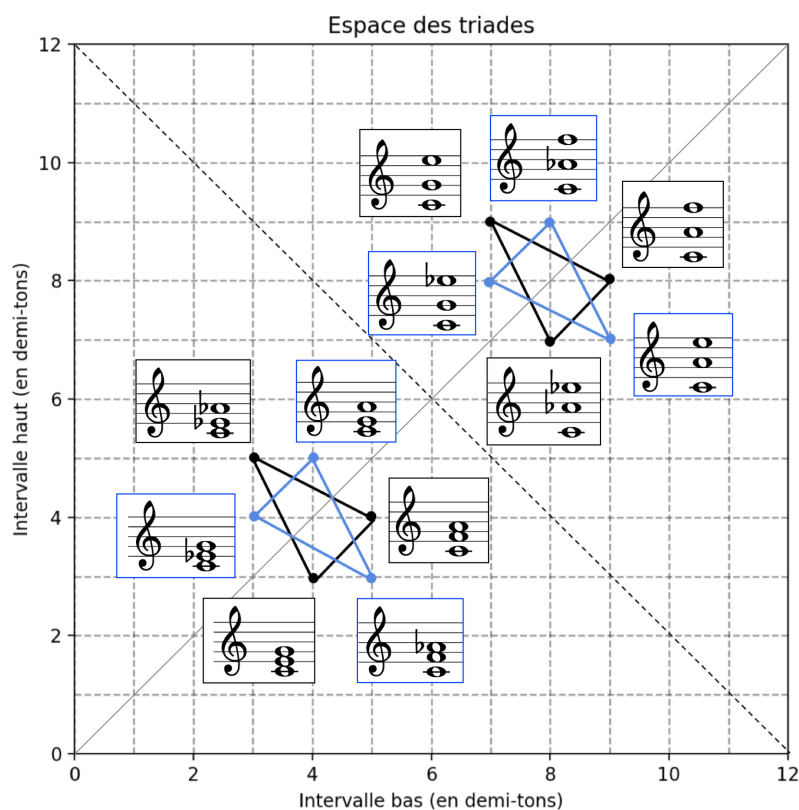


FIGURE 8.13 – Plan des triades - localisation des accords majeurs (en noir) et mineurs (en bleu), en position resserrée et dans une position plus espacée. En gris, l'axe de symétrie majeur-mineur. La diagonale noire en pointillés figure le lieu des triades s'étendant sur une octave.

Le plan des triades est l'analogue pour les dyades de la droite des inter-

valles. Les descripteurs harmoniques sont calculés sur cet espace, et peuvent être représentés via une déformation (mise en relief) ou une coloration de l'espace.

Surfaces de concordance La figure 8.14 montre la surface de concordance des triades, calculée avec 11 partiels.

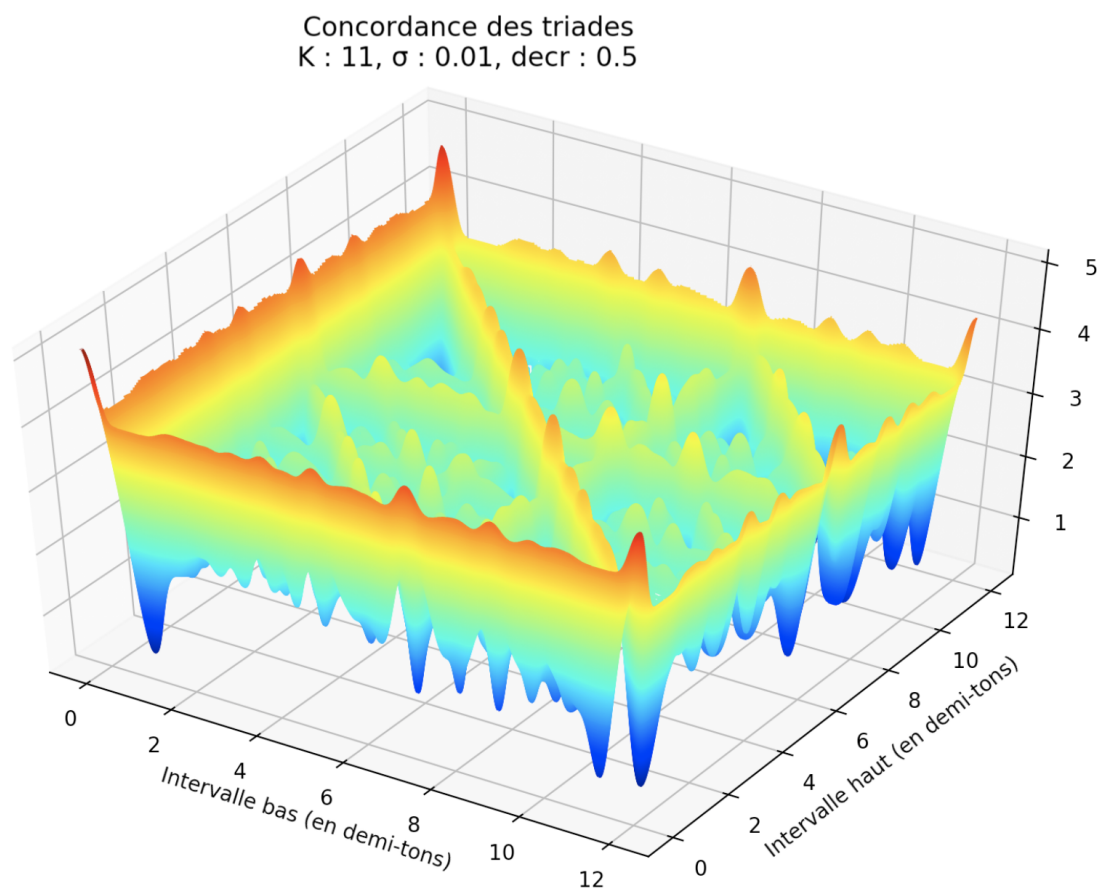


FIGURE 8.14 – Surface de concordance, visualisation en 3 dimensions, calculée avec 11 partiels. Les autres paramètres de spectre sont les suivants : $\sigma = 0.01$ (épaisseur des partiels) et *decr* = 0.5 (le k^e partiel a une amplitude égale à $1/\sqrt{k}$). Pour une meilleure visualisation, on a représenté ici le logarithme de la concordance - $\log(1 + \text{concordance})$. Les valeurs de concordance sont à la fois représentées sur l'axe z et figurées par un code couleur.

On préférera en général les représentations planes, comme celle de la figure 8.15. En noir sont représentés les accords majeurs et en rouge les accords mineurs (voir figure 8.13), cette fois avec les intervalle naturels.

La distinction entre intervalles tempérés et intonation juste est ici fondamentale. Les maximaux locaux des descripteurs comme la concordance correspondent à des accords formés avec des intervalles justes, et non des intervalles tempérés. Notamment, la tierce majeure juste (de rapport $5/4$) est située 14 cents en-dessous de la tierce majeure tempérée, tandis que la tierce mineure (de rapport $6/5$) est située 16 cents au-dessus de son homologue tempéré¹. De là vient la distorsion des positions de triades en intonation juste par rapport aux coordonnées tempérées.

1. Voir figures 8.5, 8.6 et 8.7

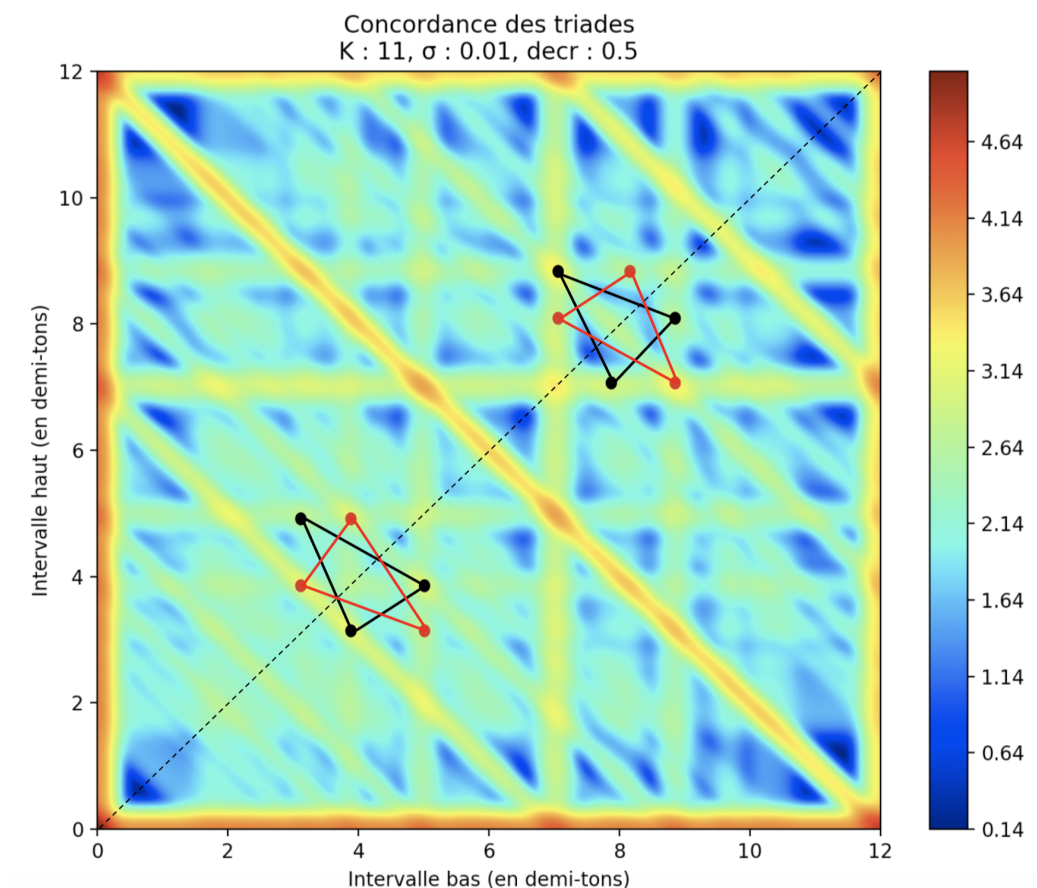


FIGURE 8.15 – Surface de concordance, visualisation plane, calculée avec 11 partiels, $\sigma = 0.01$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$). Le logarithme de la concordance et ici représenté - $\log(1 + concordance)$. En noir - les accords majeurs, en rouge - les accords mineurs, en intonation juste. En pointillé, l'axe de symétrie dû à la dépendance de la concordance uniquement au contenu intervallique.

La figure fait clairement apparaître une symétrie diagonale dont l'axe passe par l'origine, ici figuré par des pointillés, qui traduit l'invariabilité de la concordance des triades par échange des intervalles inférieur et supérieur. Celle-ci s'explique par le fait que la concordance d'un accord ne dépend que de son contenu intervallique, lequel demeure inchangé par cette symétrie.

Les accords majeurs et mineurs en intonation juste indiqués sur la figure tombent sur des maximum locaux de concordance. L'importance relative de

ces maximaux locaux est donnée sur la figure 8.16, où les accords sont en intonation juste. La concordance a été renormalisée.

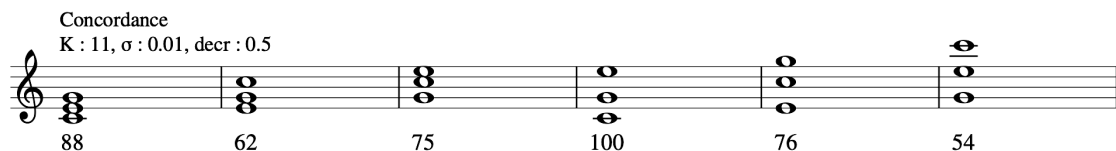


FIGURE 8.16 – Concordance des accords majeurs dans différentes positions en intonation juste, avec les paramètres de spectre $K = 11$ (nombre de partiels), $\sigma = 0.01$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

La position de l'accord majeur qui maximise la concordance est l'état fondamental en position écartée, ce qui est aussi visible sur la surface de la figure 8.15. Cela est dû à la forte coïncidence des partiels deux à deux, comme le montre la figure 8.17.

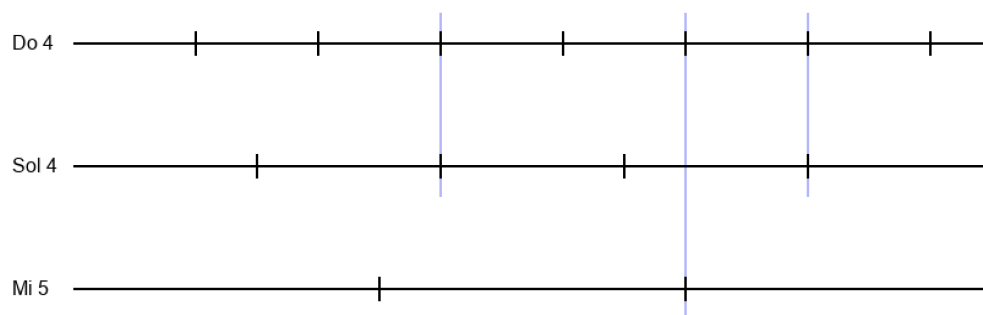


FIGURE 8.17 – Coïncidence des partiels de l'accord majeur dans son état fondamental en position écartée.

L'allure de la surface de concordance dépend fortement du nombre de partiels. Le nombre de crêtes augmente avec le nombre de partiels. La figure 8.18 montre l'ajout progressif de crêtes lorsque l'on passe de 3 à 5 puis à 7 partiels.

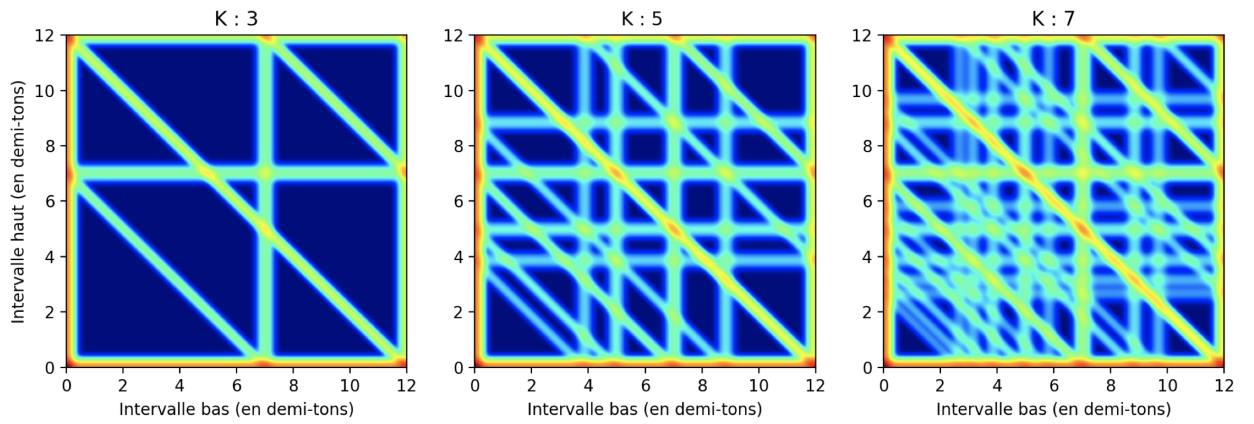


FIGURE 8.18 – Surfaces de concordance - variation du nombre K de partiels, avec $K = 3, 5$ et 7 . Les autres paramètres de spectre sont $\sigma = 0.01$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Surfaces de rugosité La figure 8.19 montre la surface de rugosité des triades, calculée avec 11 partiels.

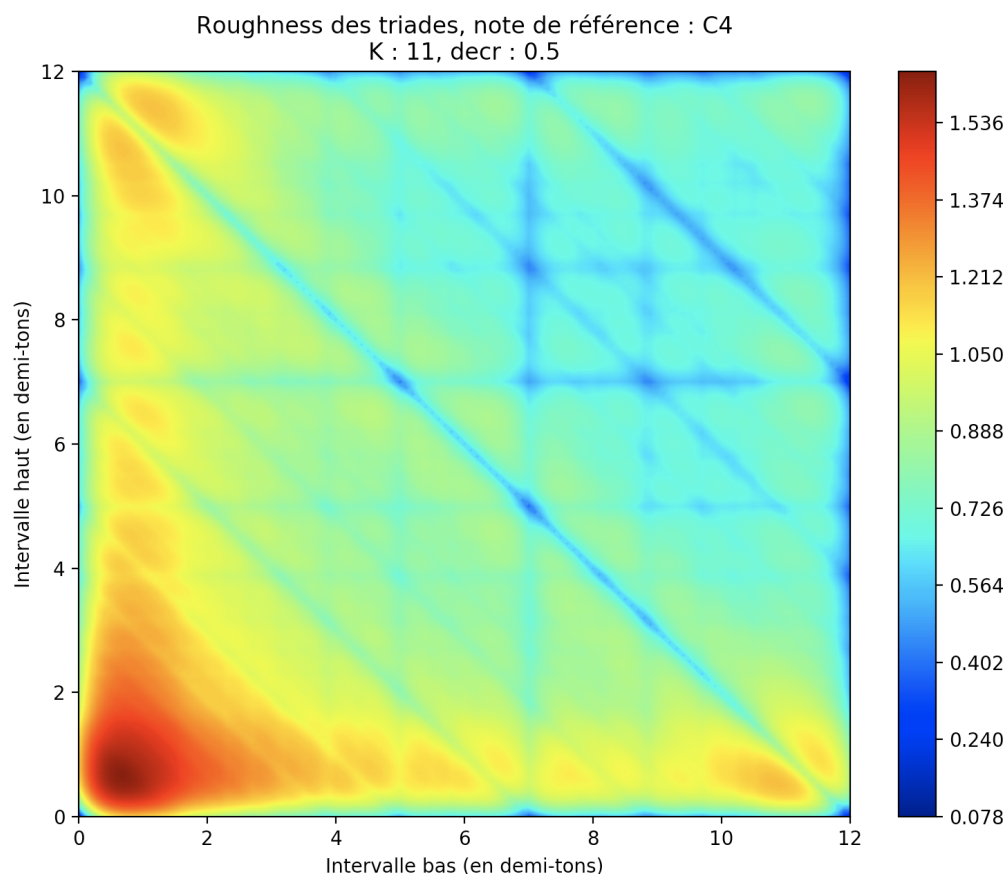


FIGURE 8.19 – Surface de rugosité des triades, avec comme note de référence *do* 4, et les paramètres de spectre suivants : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Comme pour la surface de concordance, la surface de rugosité admet une symétrie diagonale d'axe passant par l'origine, car la rugosité d'un accord ne dépend que de son contenu intervallique. La dépendance de la rugosité à la hauteur de la note inférieure est illustrée par la figure 8.20. Comme dans les courbes de rugosité¹, la position des minimums locaux ne dépend pas de

1. Voir figure 8.4.

la hauteur de la note inférieure, en revanche celle-ci agit sur la position des maximums locaux et sur l'allure générale de la courbe. Les accords majeurs et mineurs, en intonation juste, tombent sur des minimums locaux. La bosse principale est large avec *do* 3 pour note inférieure, et elle devient de plus en plus piquée, tout en se rapprochant de l'origine, pour les accords plus aigus.

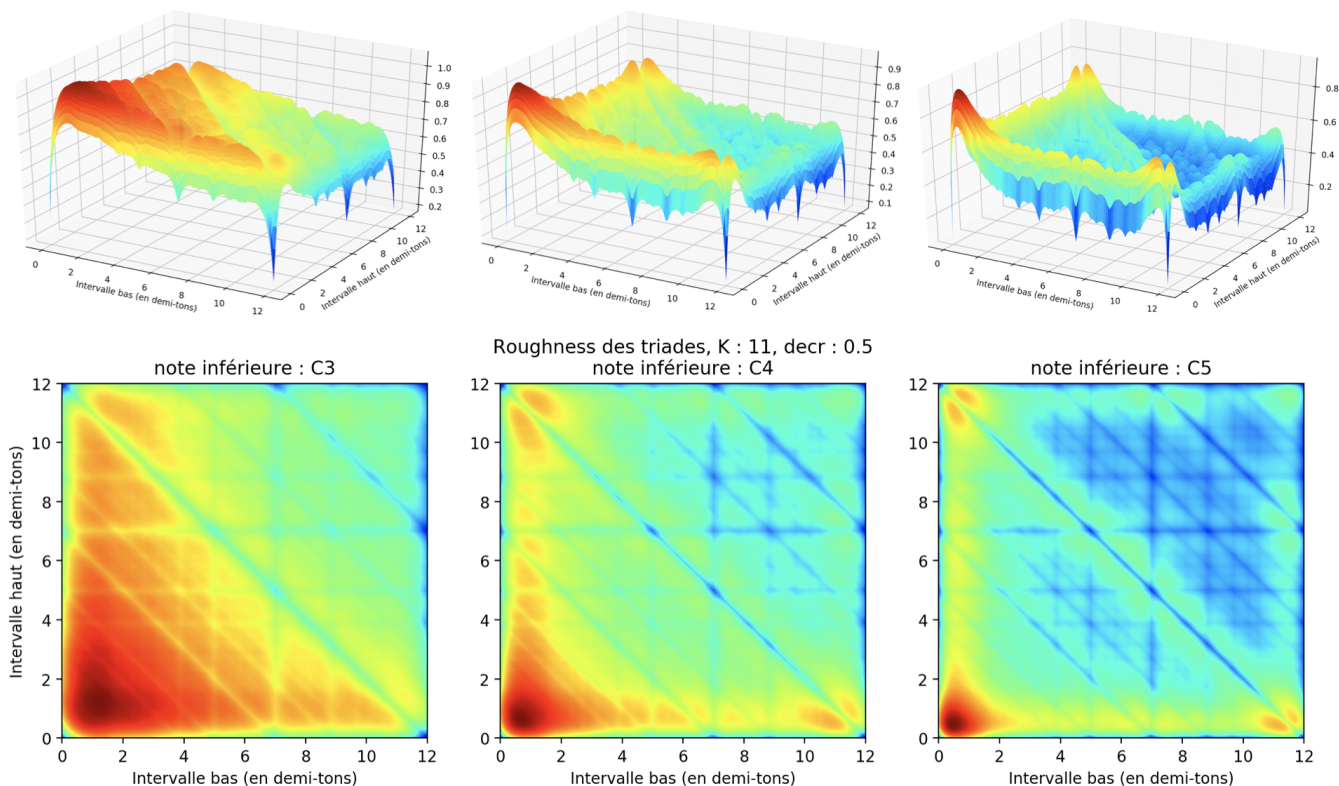


FIGURE 8.20 – Surfaces de rugosité - variation de la note inférieure, avec $f_0 = do\ 3, do\ 4$ et $do\ 5$. Les paramètres de spectre sont $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Surfaces de tension La figure 8.21 montre la surface de tension des triades, calculée avec 11 partiels.

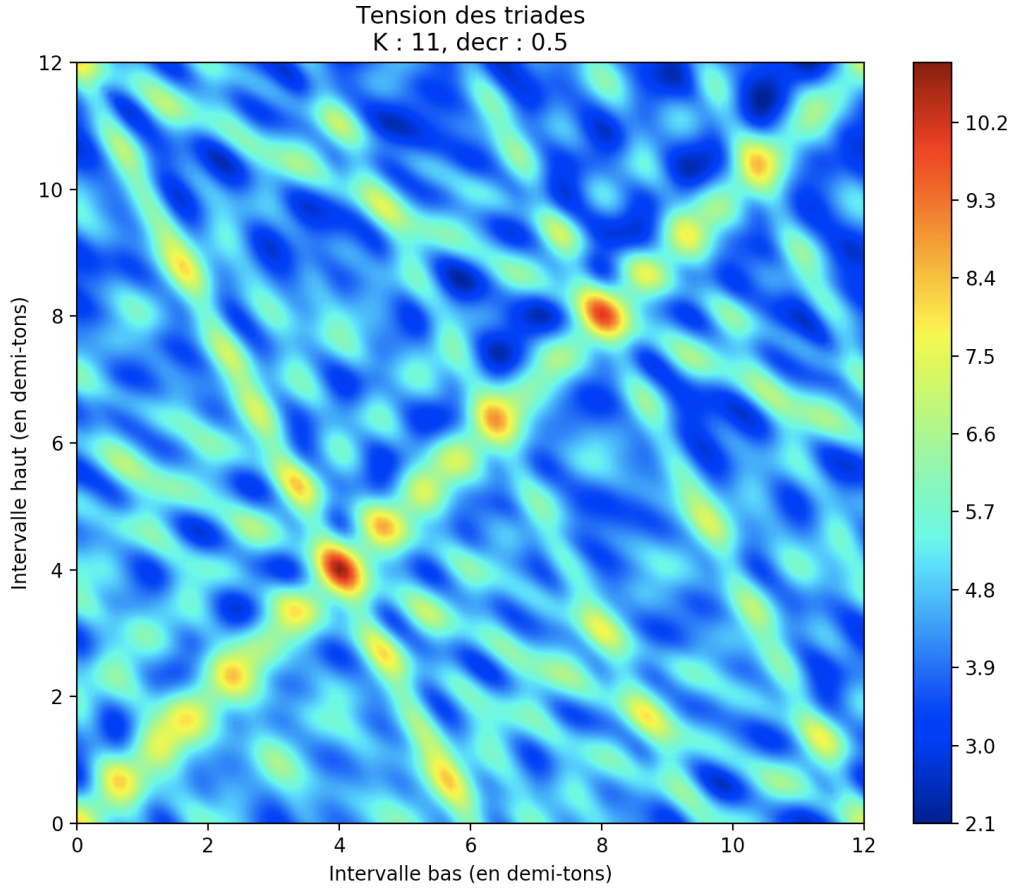


FIGURE 8.21 – Surface de tension des triades, avec $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

L'axe diagonal passant par l'origine n'est cette fois plus un axe de symétrie, étant donné que la tension d'un accord ne peut pas être déterminée par son contenu intervallique. En outre, cet axe, lieu de l'égalité des intervalles inférieur et supérieur, est un lieu de haute tension. Il n'est pas le seul, du fait de la présence des partiels. La tension maximale est atteinte en l'accord formée de deux tierces majeures tempérées, et ce dès que le nombre de partiels est supérieur ou égal à 2. Pour en comprendre la raison il suffit de représenter

l'accord en question avec ses premiers harmoniques, comme sur la partition ci-dessous, ou les harmoniques 2 sont indiqués en rouge.



Les harmoniques 1 et 2 forment une succession de tierces majeures égales, d'où une valeur de tension élevée. Les intervalles de tierce mis en valeur sont ainsi les intervalles tempérés. Cette propriété est générale : la tension fait ressortir les intervalles tempérés, c'est-à-dire les intervalles multiples d'une subdivision de l'octave. En cela la tension se différencie de la concordance et de l'harmonicité, qui font ressortir les intervalles justes. La figure 8.22 montre la surface d'e tension obtenue avec 3 partiels, et une subdivision en 6^e de tons est indiquée en pointillés. Tous les maximums locaux (il y en a plus d'une trentaine) coïncident avec des subdivisions en 6^e de tons.

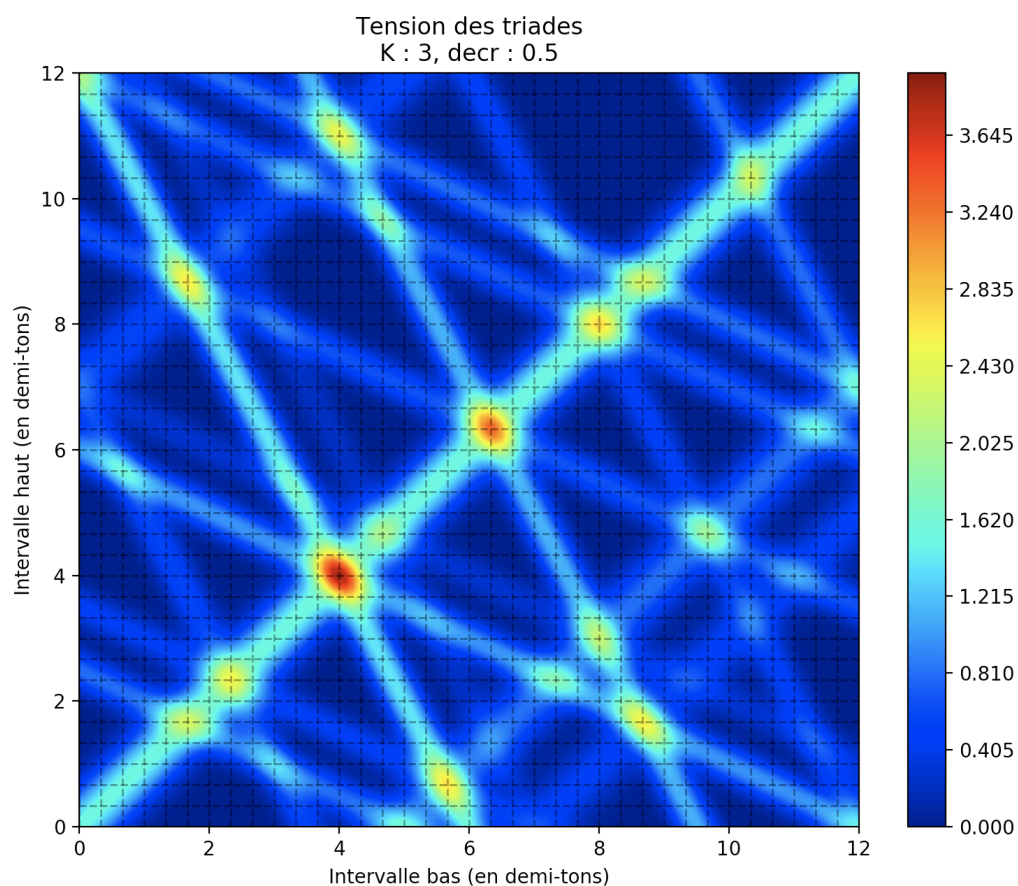


FIGURE 8.22 – Surface de tension des triades, avec $K = 3$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$). Les traits en pointillés représentent la subdivision tempérée en 6^e de tons. Les maximums locaux de la tension coïncident avec des subdivisions.

Harmonicité La figure 8.23 montre la surface d'harmonicit  des triades, calcul e avec 11 partiels.

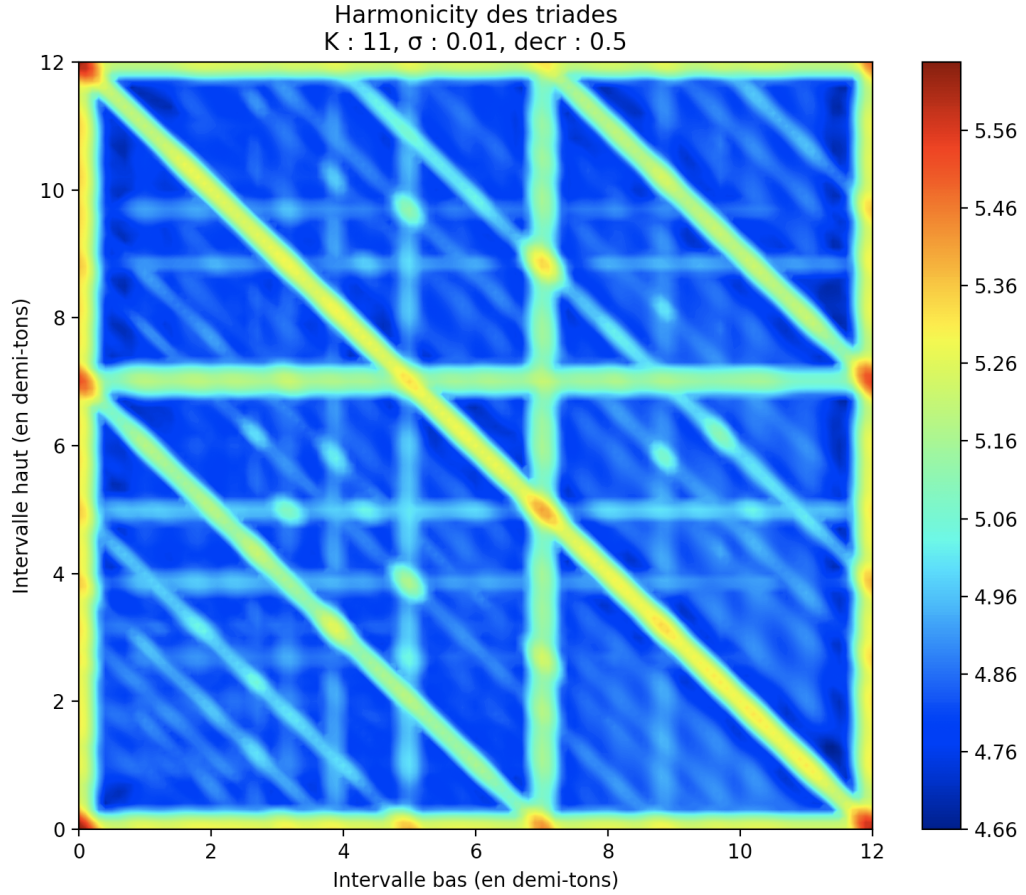


FIGURE 8.23 – Surface d'harmonicit  des triades, avec $K = 11$ (nombre de partiels), $\sigma = 0.01$ ( paisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude  gale   $1/\sqrt{k}$). Pour une meilleure visibilit  on a repr sent  le logarithme de l'harmonicit .

La surface d'harmonicit  a de fortes ressemblances avec la surface de concordance¹, mais contrairement   celle-ci elle n'admet pas de sym trie diagonale. L'harmonicit  ne peut pas en effet  tre d duite du contenu intervallique.

1. Voir figure 8.15

L'harmonicité met en avant les triades présentes dans la série des partiels. La figure 8.24 recense toutes les triades présentes dans la série des 7 premiers harmoniques dont les intervalles haut et bas ne dépassent pas l'octave, et indique leur position sur la surface d'harmonicité. Leur position coïncide avec des maximums locaux d'harmonicité.

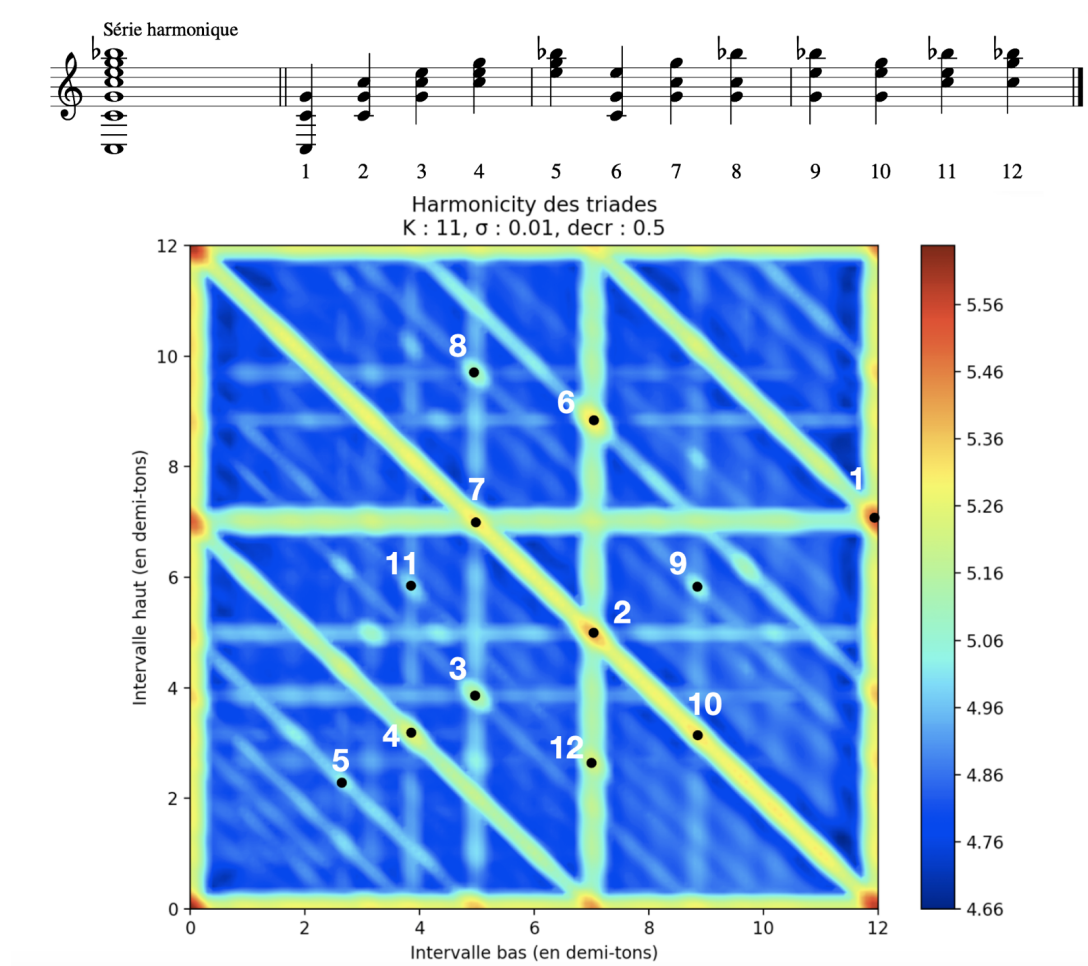


FIGURE 8.24 – Positionnement sur la surface d'harmonicité des triades présentes dans la série des 7 premiers partiels harmoniques. La surface étant invariante par transposition, tous les accords peuvent être représentés sur la même figure bien que leur note inférieure diffère. Paramètres de spectre ; $K = 11$ (nombre de partiels), $\sigma = 0.01$ (épaisseur des partiels) et $\text{decr} = 0.5$ (le k^{e} partiel a une amplitude égale à $1/\sqrt{k}$).

Concordance d'ordre 3 La figure 8.25 montre la surface de concordance d'ordre 3 des triades, calculée avec 11 partiels.

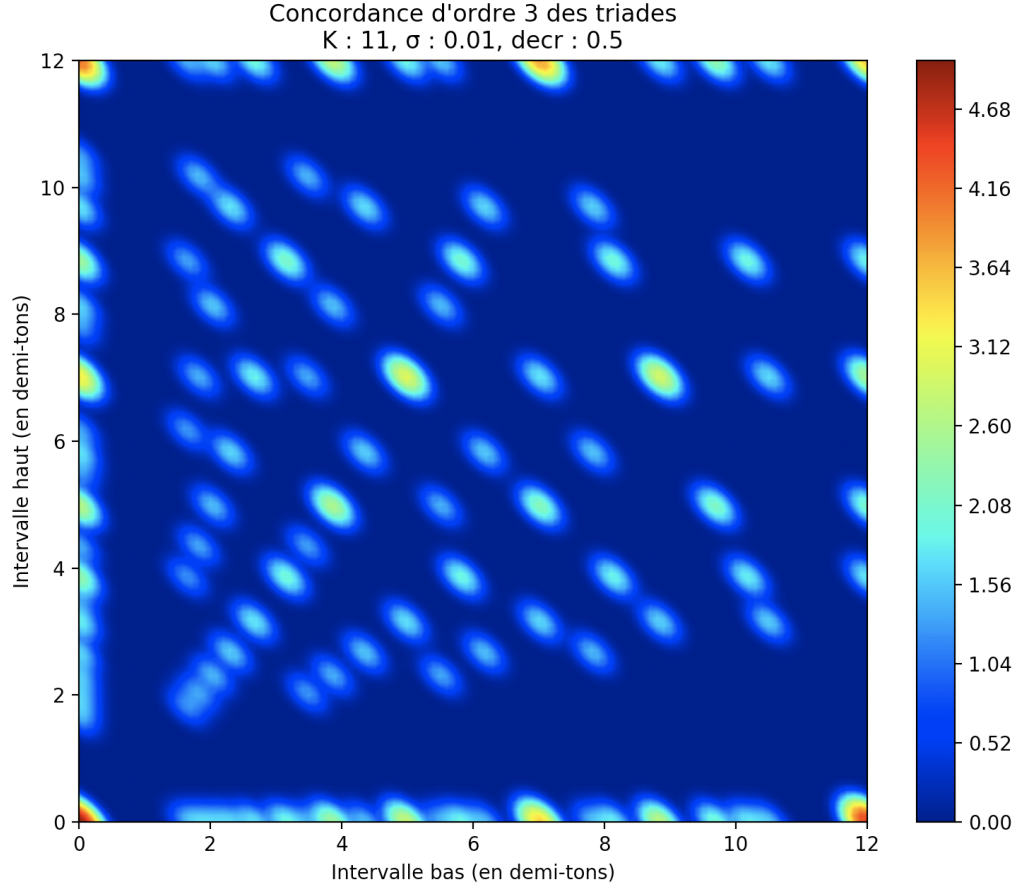


FIGURE 8.25 – Surface de concordance d'ordre 3 des triades, avec $K = 11$ (nombre de partiels), $\sigma = 0.01$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

La géométrie de la surface de concordance d'ordre 3 semble moins régulière que pour les autres descripteurs statiques, du fait des lignes courbes dont nous détaillerons plus loin l'origine. La concordance d'ordre 3 met en avant les triades qui apparaissent dans la série des harmoniques inférieures¹.

1. Voir page 218 et partie 6.1.3.

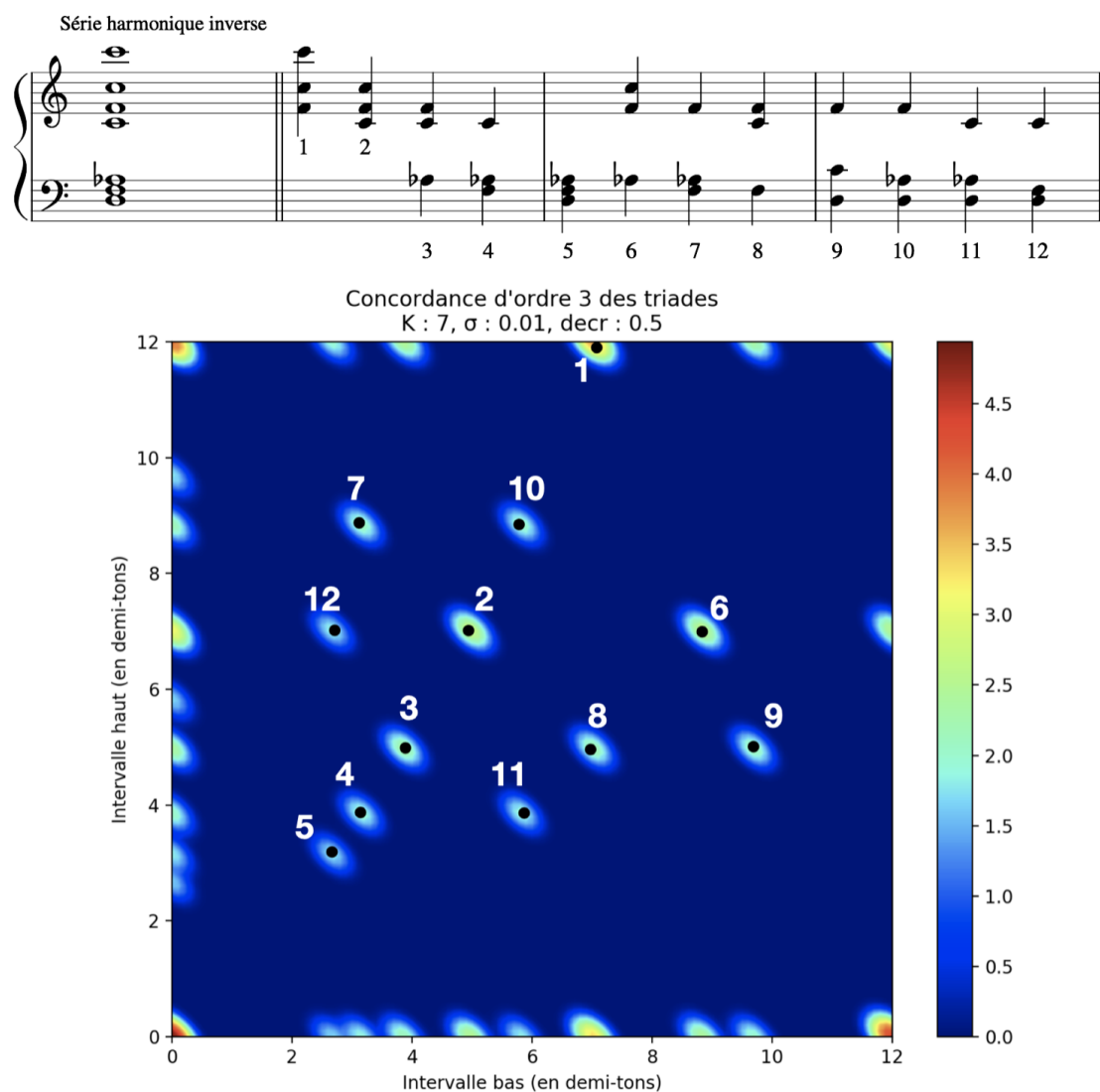


FIGURE 8.26 – Positionnement des triades présentes dans la suite des 7 premiers partiels harmoniques inférieurs sur la surface de concordance d'ordre 3 obtenue avec 7 partiels. Les harmoniques inférieurs sont représentés à partir de *do* 6. La surface étant invariante par transposition, tous les accords peuvent être représentés sur la même figure bien que leur note inférieure diffère. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.01$ (épaisseur des partiels) et $\text{decr} = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

La figure 8.26 montre le positionnement des triades présentes dans la suite des 7 premiers partiels harmoniques inférieurs sur la surface de concordance d'ordre 3 à 7 partiels. L'ensemble de ces triades est obtenu à partir de l'ensemble des triades issues de la série harmonique¹ par échange des intervalles supérieur et inférieur de chaque triade, ce qui se traduit graphiquement par la symétrie diagonale d'axe passant par l'origine. Ces accords correspondent exactement aux maximums locaux de concordance d'ordre 3.

La figure précédente nous permet de mieux comprendre les lignes courbes se dégageant des surfaces de concordance d'ordre 3, qui sont d'autant plus visibles que le nombre de partiels est grand. Considérons-en trois :

1. Celle qui contient les accords n°1, 2, 3, 4, 5 et 10 (dits du 1^{er} groupe) ;
2. Celle qui contient les accords n°6, 8 et 11 (du 2^e groupe) ;
3. Celle qui contient les accords n°7 et 12 (du 3^e groupe).

Pour toute triade incluse dans la série harmonique inférieure, notons i, j, k les numéros des partiels de la série harmonique inférieure desquels ils sont constitués. Pour l'accord n°1, on a ainsi $i = 1, j = 2$ et $k = 3$.

Il est aisé de vérifier que les accords au sein de chaque groupe vérifient des égalités communes :

$$\begin{cases} j - i = k - j & \text{pour les accords du 1^{er} groupe} \\ k - j = 2(j - i) & \text{pour les accords du 2^e groupe} \\ j - i = 2(k - j) & \text{pour les accords du 3^e groupe} \end{cases} \quad (8.1)$$

Il est alors possible de paramétriser ces courbes. Considérons une triade incluse dans la série harmonique inférieure, et notons X et Y les intervalles respectivement inférieurs et supérieurs, en octaves (et donc situés entre 0 et 1). Si f_i, f_j et f_k désignent les fréquences de la triade rangées de la plus aiguë à la plus grave, alors :

$$X = \log_2 \left(\frac{f_j}{f_k} \right) \text{ et } Y = \log_2 \left(\frac{f_i}{f_j} \right).$$

Or, comme les fréquences f_i, f_j et f_k sont en rapport harmonique, on a :

1. Voir figure 8.24.

$$\frac{f_j}{f_k} = \frac{k}{j} \text{ et } \frac{f_i}{f_j} = \frac{j}{i}.$$

Ainsi

$$2^X = \frac{k}{j} \text{ et } 2^Y = \frac{j}{i}. \quad (8.2)$$

En substituant i dans le second membre de l'équation (8.2) par son expression donnée dans chacune des égalités (8.1), on obtient les équations de courbes suivantes :

$$\begin{cases} Y = -\log_2(2 - 2^X) & \text{pour les accords du 1}^{er} \text{ groupe} \\ Y = 1 - \log_2(3 - 2^X) & \text{pour les accords du 2}^e \text{ groupe} \\ Y = -\log_2(3 - 2^{X+1}) & \text{pour les accords du 3}^e \text{ groupe.} \end{cases} \quad (8.3)$$

Ces courbes sont représentées sur la figure 8.27.

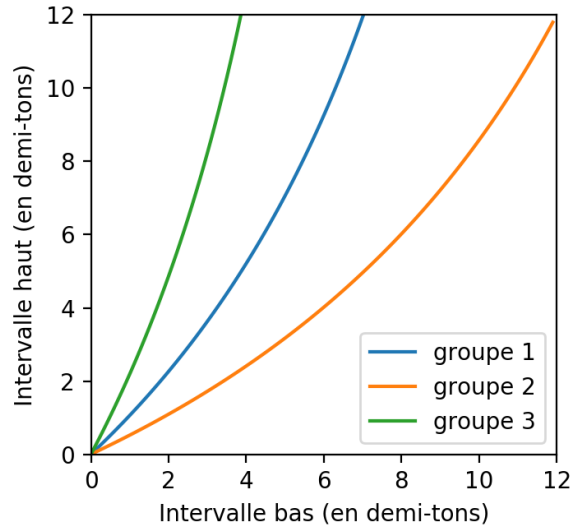


FIGURE 8.27 – Courbes de localisation des triades au sein de la série harmonique inférieure, visibles sur les figures 8.25 et 8.26.

De telles courbes deviennent de plus en plus nombreuses avec un nombre de partiels élevé, étant donné que les configurations combinatoires des triades au sein de la série harmonique inférieure sont de plus en plus nombreuses.

8.3 Aspect dynamique : ajout d'une troisième note à une dyade

8.3.1 Représentation des espaces de transition

L'ensemble des dyades, si l'on fixe la note inférieure, est de dimension 1, représentable par une droite, tandis que l'espace des triades, toujours en fixant la note inférieure, est de dimension 2, représentable par un plan. Cela rend aisée la visualisation des descripteurs harmoniques statiques sur ces ensembles, comme nous l'avons vu précédemment ¹. Par contre, dans le cas des descripteurs harmoniques dynamiques, la visualisation n'est pas si évidente. En effet, l'ensemble des enchaînements de dyades est de dimension 4, qui se réduit à 3 si l'on fixe la note inférieure du premier accord, tandis que l'ensemble des enchaînements de triades, la note inférieure du premier accord fixée, est de dimension 5, ce qui pose des problèmes de visualisation.

Ajout d'une troisième note à une dyade Nous considérons dans cette partie les transitions harmoniques qui consistent en l'ajout d'une troisième note à une dyade. Les normalisations des descripteurs par l'énergie de l'accord et par le nombre de notes ² permettent de calculer sur ces transitions les descripteurs dynamiques ainsi que l'évolution des descripteurs statiques.

L'ensemble de ces transitions, si l'on fixe la note inférieure de la dyade, est de dimension 2. La figure 8.28 montre une visualisation plane de cet espace, dans laquelle l'axe des abscisses représente la dyade initiale, tandis que l'axe des ordonnées représente l'intervalle ajouté au-dessus de la note supérieure de la dyade initiale, compté algébriquement. Ainsi, la partie supérieure de cet espace, correspondant à un intervalle ajouté positif, est analogue à la représentation de l'espace des triades ³ utilisé jusqu'ici, si l'on se restreint au second accord de la transition. Si l'intervalle ajouté est négatif (vers le grave), deux cas se présentent : soit l'intervalle ajouté est plus petit en valeur absolue que l'intervalle de la dyade initiale, auquel cas la troisième note se situe entre les deux notes de la dyade, ce qui correspond à l'espace au-dessus de la diagonale en bleue, soit il est plus grand, auquel cas la note ajoutée devient la note inférieure de la triade.

1. Voir parties 8.1 et 8.2.

2. Voir parties 7.2 et 7.3.

3. Voir figure 8.13.

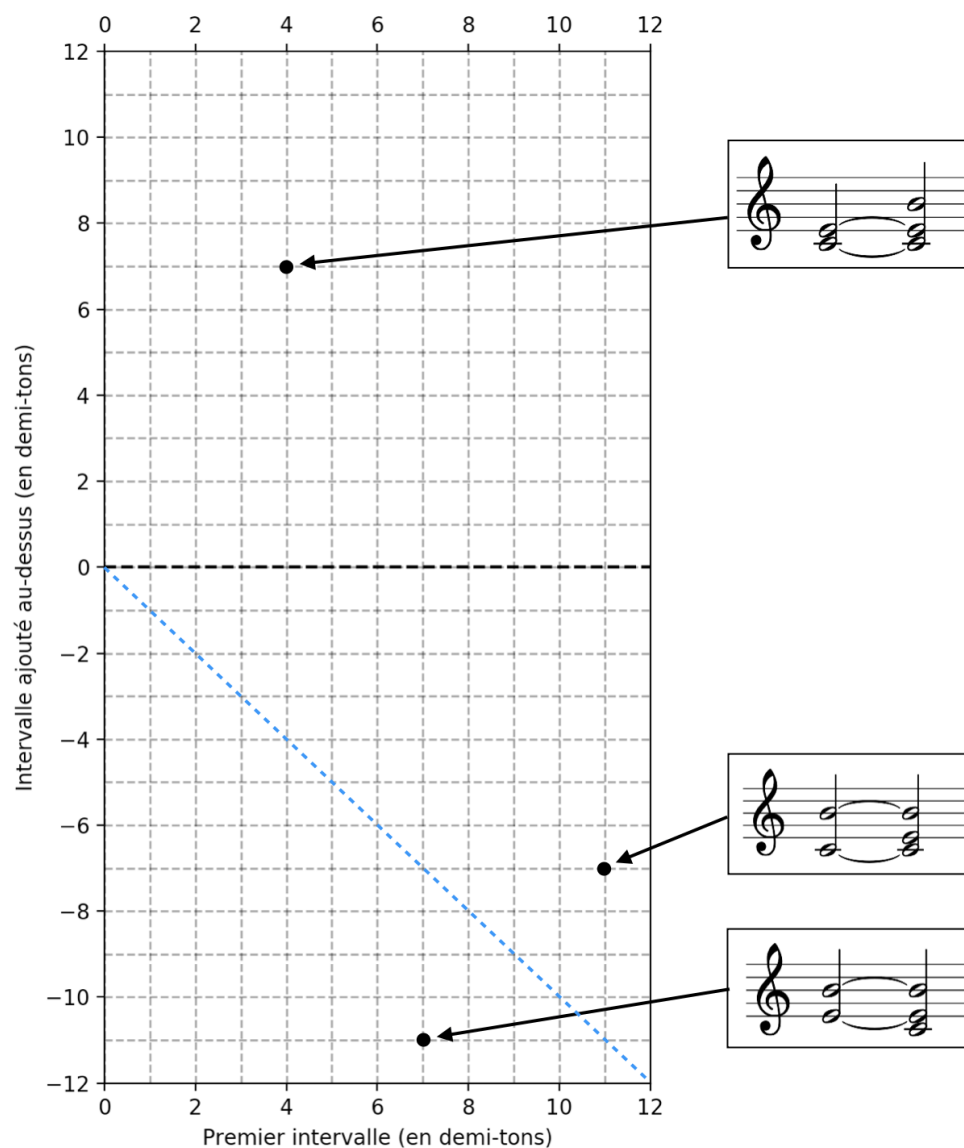


FIGURE 8.28 – Représentation de l'espace des transitions : ajout d'une troisième note à une dyade. L'axe des abscisses représente la dyade initiale, tandis que l'axe des ordonnées représente l'intervalle ajouté au-dessus de la note supérieure de la dyade initiale, compté algébriquement. Quelques transitions sont indiquées dans la gamme tempérée.

8.3.2 Cartes de gain des descripteurs statiques

Les transitions peuvent être caractérisées de deux manières : par les descripteurs dynamiques non triviaux, tels le changement harmonique, la concordance ou la rugosité différentielle, ainsi que par les descripteurs dynamiques triviaux, qui correspondent à l'évolution des descripteurs statiques. Nous nous penchons dans cette partie sur les descripteurs triviaux. De la dyade à la triade obtenue par ajout d'une note, la valeur d'un descripteur statique peut augmenter ou diminuer, ce qui définit deux régions distinctes sur la carte, respectivement celles du gain et de la perte.

Perte de rugosité La figure 8.29 montre la carte de perte de rugosité. La région bleue correspond à une perte de rugosité, tandis que la région rouge correspond à un gain.

En raison des conditions sur les nuances et sur les unissons¹, l'équation qui régit la rugosité² dépend à la fois du nombre de notes et de l'énergie de l'accord. C'est pour cette raison que l'ajout de la troisième note ne fait pas nécessairement augmenter la rugosité.

On peut séparer le lieu de perte de rugosité en trois régions :

- Les diagonales et la ligne horizontale centrale. La diagonale partant du coin gauche est le lieu où la 3^e note crée une octave avec la basse. La diagonale partant de (0,0) et la ligne horizontale d'ordonnée nulle sont le lieu des transitions où la 3^e note crée un unisson. La diagonale qui part de (7,12) est quant à elle le lieu des transitions où la 3^e note crée une octave plus une quinte avec la note inférieure.
- Les transitions dont la dyade initiale est située autour de l'intervalle de rugosité maximale, c'est-à-dire le demi-ton³, avec *do* 4 comme note de référence.
- Les transitions qui n'entrent pas dans les deux premières catégories, notamment des transitions avec une tierce mineure ou une septième majeure comme intervalle initial.

1. Voir chapitre 7

2. Voir équation (7.5).

3. Voir figure 8.4

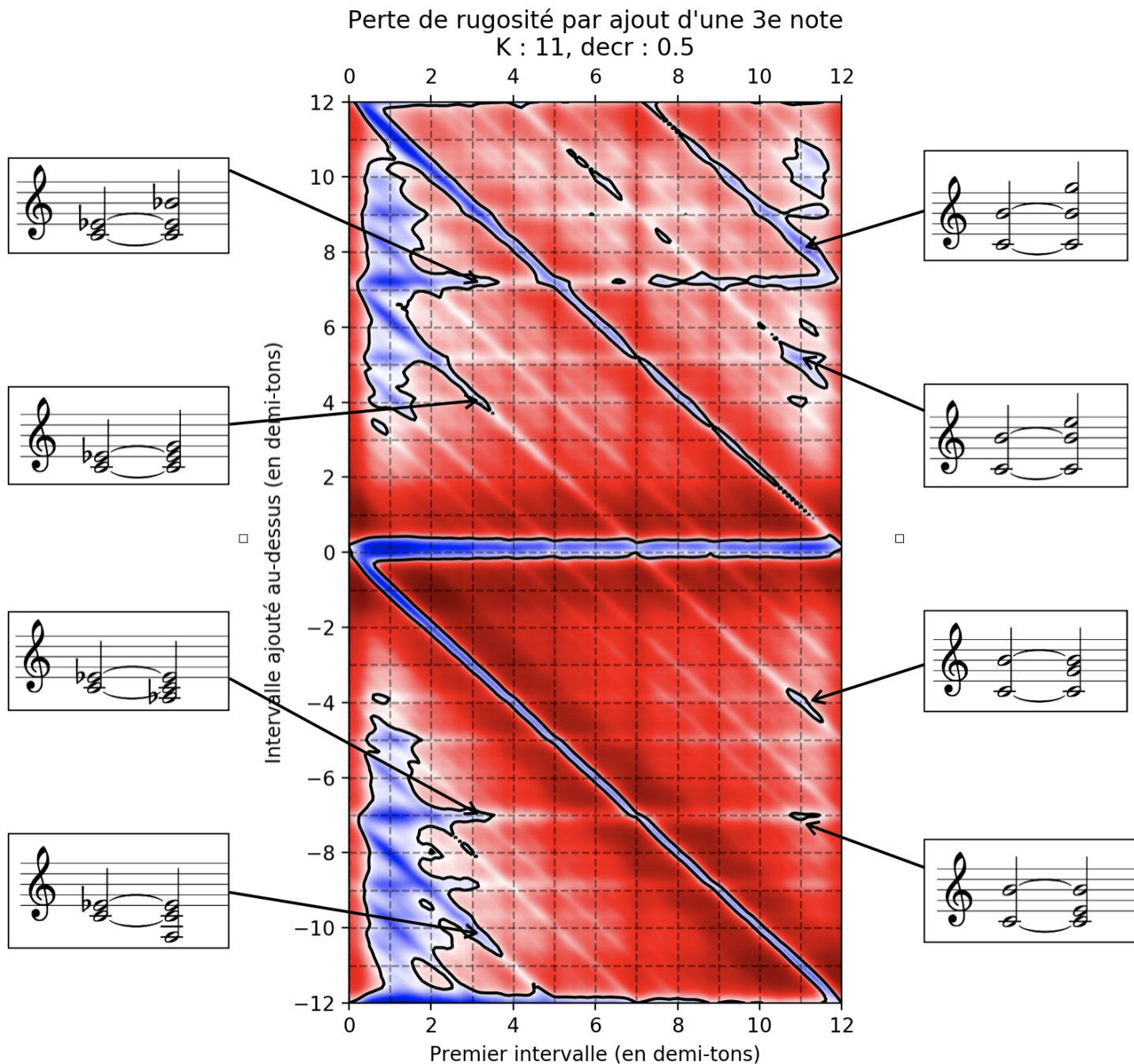


FIGURE 8.29 – Carte de perte de rugosité. La région rouge correspond à un gain, la bleue à une perte. Les transitions indiquées correspondent à une intonation juste. En noir, les transitions qui laissent la rugosité invariante. La note de référence est ici *do* 4. Paramètres de spectre utilisés : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Sur les deux premières régions, la perte de rugosité était prévisible. En effet, l'unisson, l'octave et l'octave plus la quinte sont les trois intervalles avec la rugosité la plus faible. Pour ces deux régions, la 3^e note ajoute des intervalles qui sont moins rugueux que la dyade initiale, d'où l'idée qu'« en moyenne¹ » la rugosité diminue.

La troisième région nous intéresse particulièrement. Quelques transitions ont été indiquées. Là encore, les transitions indiquées sont en intonation juste, d'où le décalage avec le quadrillage correspondant aux intervalles tempérés. Notamment, à partir de la septième majeure *do-si*, l'ajout de la tierce majeure *mi* et de la quinte *sol*, à la fois à la même octave et à l'octave supérieure, fait diminuer la rugosité. Ceci répond, du moins partiellement, au problème soulevé dans le prologue de ce travail², qui était celui de comprendre pourquoi l'accord *do-mi-si* « sonne mieux³ » que l'intervalle de septième majeure seul.

Rappelons ici les quatre explications plausibles données par Parncutt :

- **Explication 1** : Les battements de l'intervalle *do-si* sont partiellement masqués quand le *mi* est rajouté.
- **Explication 2** : La rugosité de l'accord *do-mi-si* est distribuée sur les bandes critiques de manière plus homogène que pour l'intervalle *do-si*.
- **Explication 3** : L'accord *do-mi-si* est un accord de septième majeure sans la quinte, on l'identifie en tant que tel et il est plus familier que l'intervalle *do-si*.
- **Explication 4** : L'accord à 3 notes aurait une propriété de fusion ou d'euphonie que n'a pas l'intervalle *do-si*.

Notre approche ne touche aucune de ces explications, si ce n'est éventuellement la dernière, puisque le calcul de la rugosité tient compte du nombre de notes.

1. Ce n'est pas là la seule considération, puisque le calcul de la rugosité tient également compte de l'énergie, voir l'équation (7.5).

2. Voir page 29.

3. Cette expression « sonner mieux », est à prendre avec toutes les précautions nécessaires. Elle garde un sens volontairement imprécis ici, l'étude perceptive n'étant pas le sujet direct de ce travail.

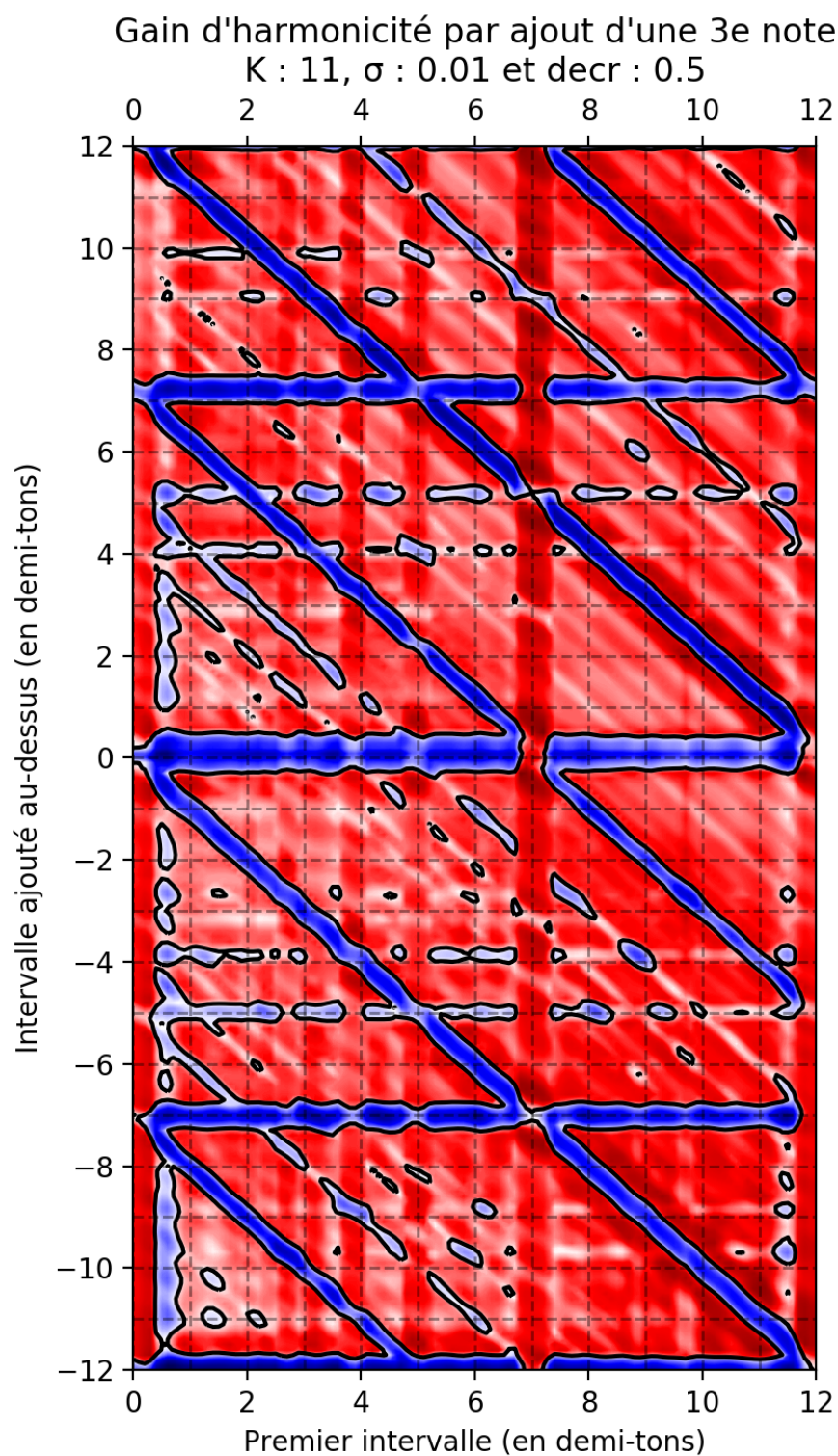


FIGURE 8.30 – Carte de gain d’harmonicit . La r gion bleue correspond   un gain d’harmonicit , la rouge   une perte. En noir, les transitions laissant invariante l’harmonicit . Param tres de spectre utilis s : $K = 11$ (nombre de partiels), $\sigma = 0.01$ ( paisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude  gale   $1/\sqrt{k}$).

Gain d’harmonicité La figure 8.30 montre la carte de gain d’harmonicité. La région bleue correspond à un gain d’harmonicité, la rouge à une perte.

Les régions de gain d’harmonicité sont constituées essentiellement en diagonales — celles qui partent des points (7,12), (0,12), (0,7), (0,0) et (0,-7), et de lignes horizontales — d’ordonnées 12, 7, 0, -7, -12. Celles-ci correspondent aux transitions vers des triades contenant des unissons, quintes, octaves ou octaves plus quintes. D’autres lignes et diagonales de gain d’harmonicité apparaissent, le gain étant moins important que les précédentes. Ces transitions contiennent notamment des quartes — diagonale partant de (0,5) ou ligne horizontale d’ordonnée 5, ou des tierces majeures — diagonale partant de (0,4) ou ligne horizontale d’ordonnée 4.

Gain de concordance totale La figure 8.31 montre la carte de gain de concordance totale. La région bleue correspond à un gain de concordance totale, la rouge à une perte.

Les régions de gain de concordance totale se présentent comme des îlots isolés. Une comparaison de la partie supérieure avec la surface de concordance d’ordre 3 des triades¹ montre que ces régions correspondent toutes à des transitions vers des maximums locaux de la concordance d’ordre 3, mis à part certaines transitions à partir de la seconde mineure, dont le gain de concordance totale est quasiment nul. Il en va de même de la partie basse de la carte de gain. Cependant, toutes les transitions vers des maximums locaux de la concordance d’ordre 3 ne correspondent pas à un gain de concordance totale. Les transitions vers de maximums locaux de la concordance d’ordre 3 à partir des intervalles les plus concordants, à savoir l’unisson, l’octave ou la quinte, s’accompagnent d’une perte de concordance totale.

1. Voir page 304.

Gain de concordance totale par ajout d'une 3e note

K : 11, σ : 0.01 et decr : 0.5

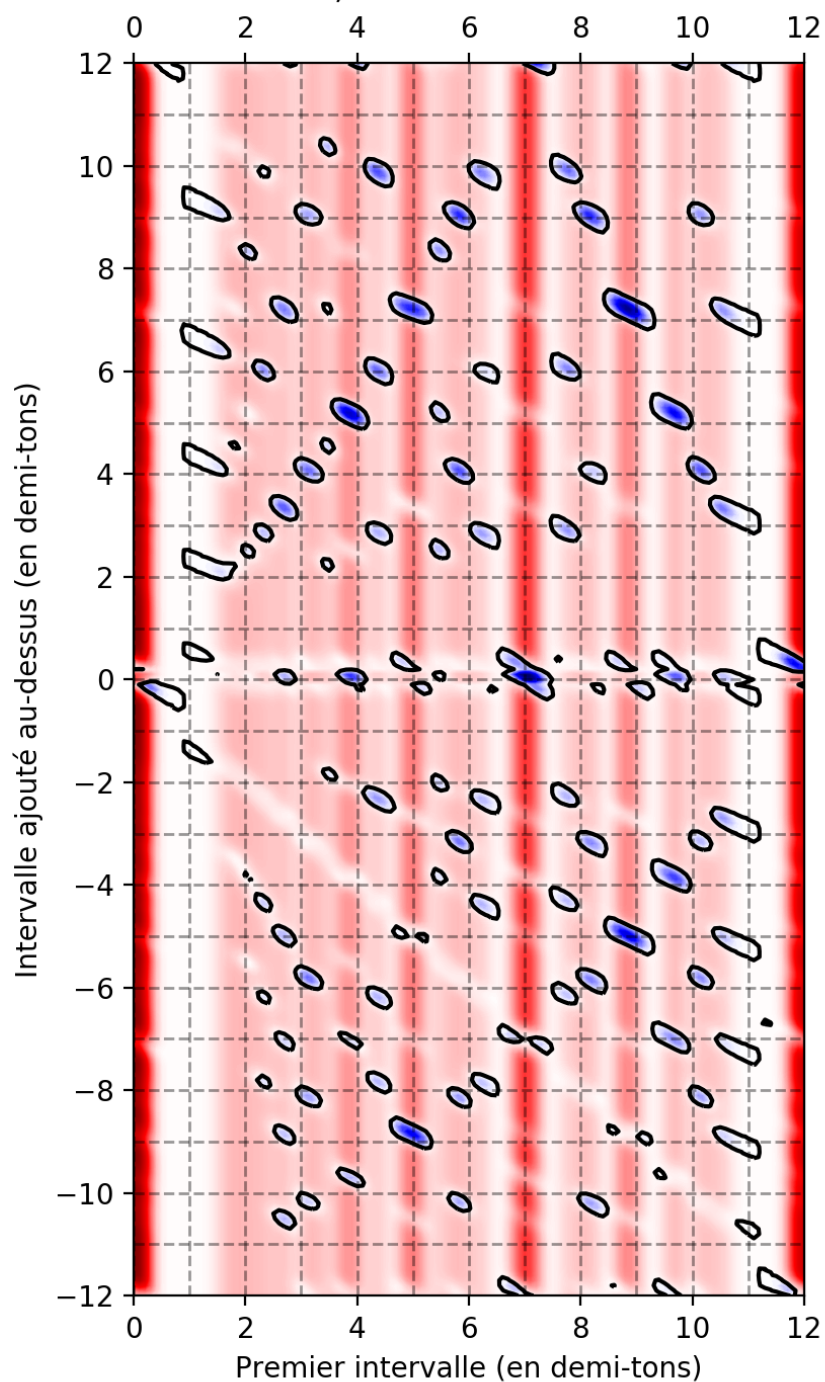


FIGURE 8.31 – Carte de gain de concordance totale. La région bleue correspond à un gain de concordance totale, la rouge à une perte. En noir, les transitions laissant invariante la concordance totale. Paramètres de spectre utilisés : $K = 11$ (nombre de partiels), $\sigma = 0.01$ (épaisseur des partiels) et $\text{decr} = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

8.3.3 Surfaces de descripteurs dynamiques

Penchons-nous maintenant sur la topographie induite par les descripteurs dynamiques sur l'espace des transitions par ajout d'une troisième note à une dyade. Ces descripteurs ne sont plus une différence entre deux valeurs successives de descripteurs statiques, mais décrivent la transition elle-même. Cependant, étant donnée la configuration particulière des transitions considérées, et le statut particulier de la troisième note par rapport aux deux autres, les surfaces obtenues ont de fortes similarités avec les surfaces de descripteurs statiques des triades. Contrairement au calcul de la rugosité, de la concordance ou de la concordance totale, le calcul des descripteurs dynamiques non triviaux ne fait pas intervenir le nombre de notes des accords, traitant de la même manière la dyade et la triade, avec toutefois une normalisation par les énergies ¹.

Changement harmonique et concordance différentielle La figure 8.32 montre les surfaces de concordance différentielle et de changement harmonique. La partie supérieure des deux surfaces présente des motifs qui entretiennent de fortes similarités avec ceux présents sur les surfaces de concordance et d'harmonicité des triades ². Contrairement à la surface de concordance des triades, la grande diagonale passant par (0,0) et (12,12) n'est plus un axe de symétrie (la ligne verticale et la ligne horizontale passant par l'origine diffèrent). le statut particulier de la 3^e note brise en effet la symétrie des descripteurs entre la dyade initiale et l'intervalle formé par la 3^e note avec la note supérieure de la dyade.

Les surfaces de concordance différentielle et de changement harmonique ont des variations opposées, ce qui montre leur forte corrélation sur ce type particulier de transitions.

Rugosité différentielle La figure 8.33 montre la surface de rugosité différentielle. Là-aussi, la partie supérieure entretient de fortes similarités avec la surface de rugosité des triades ³.

1. Voir page 255.

2. Voir figures 8.15 et 8.23.

3. Voir figure 8.19.

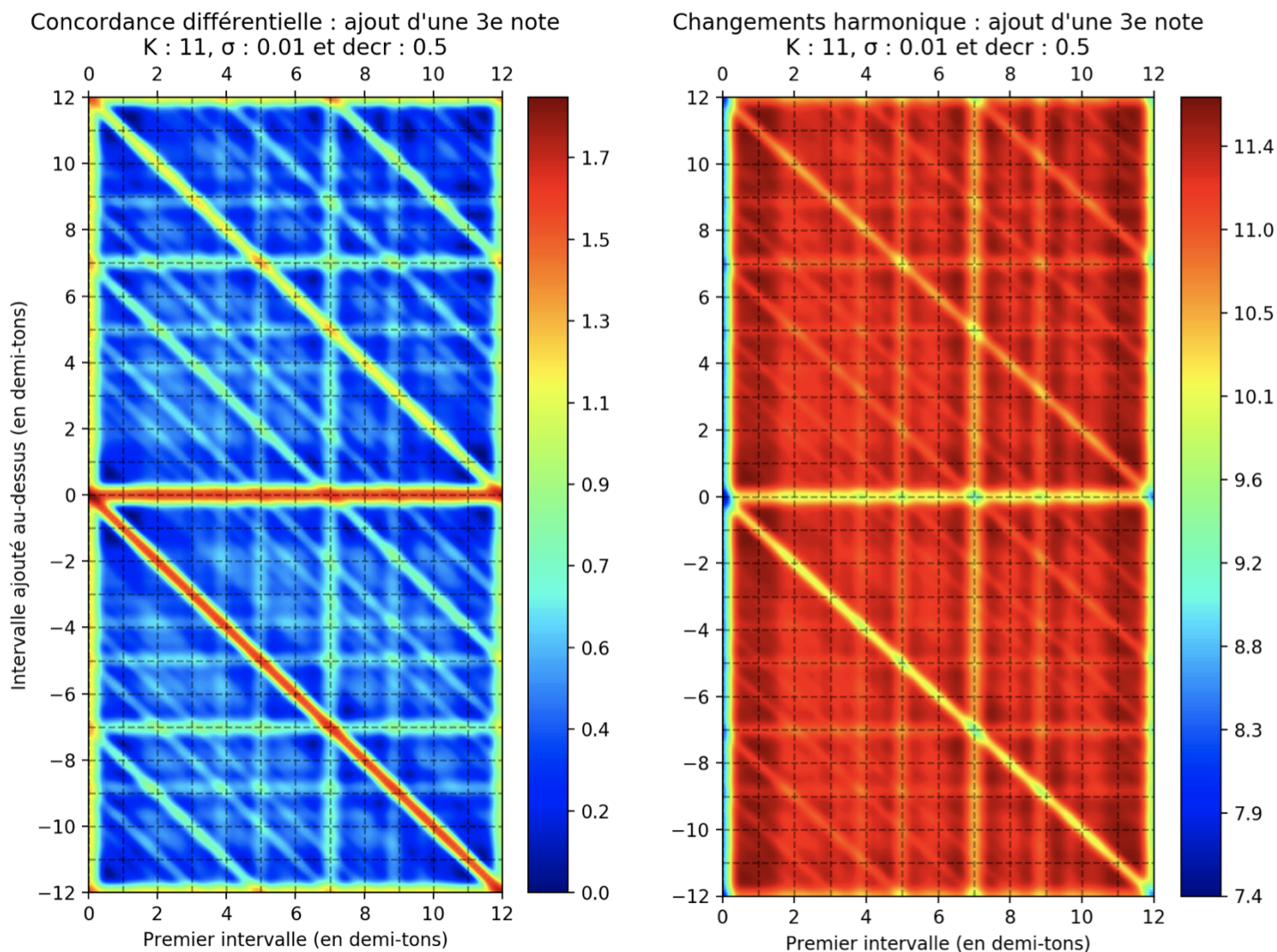


FIGURE 8.32 – Ajout d’une note : surfaces de concordance différentielle et changement harmonique. Pour une meilleure visualisation, on a représenté le logarithme de la concordance différentielle. le changement harmonique reste inchangé. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.01$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

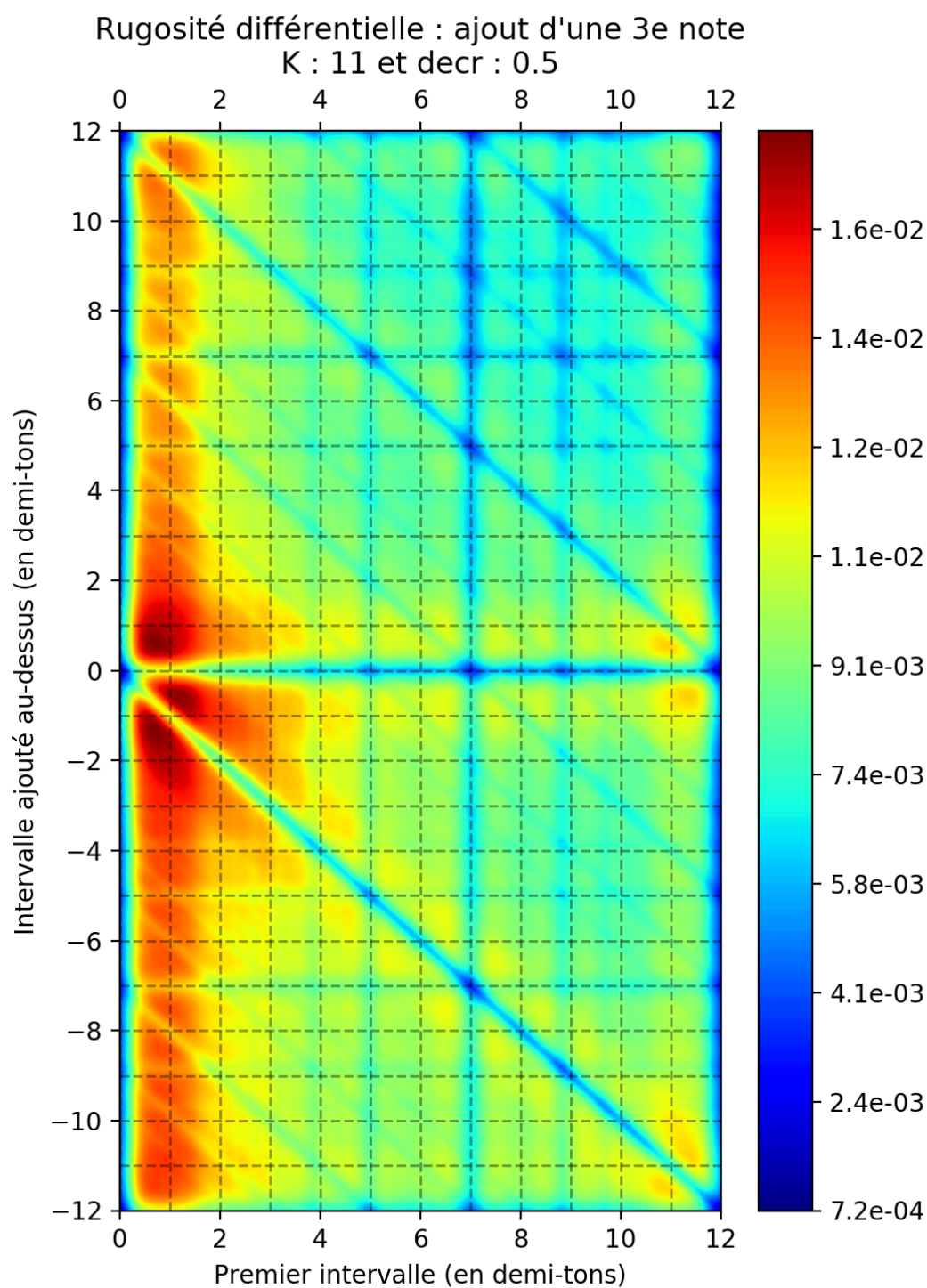


FIGURE 8.33 – Ajout d’une note - surface de rugosité différentielle. Paramètres de spectre : Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

8.3.4 Descripteurs dynamiques triviaux et non triviaux : synthèse des apports

Nous avons présenté deux approches pour l'étude de transitions particulières, qui consistent en l'ajout d'une 3^e note à une dyade : la première s'attache à mesurer l'augmentation ou la diminution des valeurs des descripteurs statiques occasionnées par l'ajout de la 3^e note, et donne lieu à une carte de gain, tandis que la seconde traite la transition en temps que telle, par le calcul et la représentation des descripteurs dynamiques non triviaux.

La première approche mène à des représentations plus facilement interprétables que la seconde, et semble plus appropriée à l'étude de ces transitions, étant donnée qu'elle met en valeur des comportements non triviaux. C'est notamment le cas de l'évolution de la rugosité, qui fournit un élément de réponse au problème présenté dans le prologue, qui est celui de comprendre les différences perceptives entre la triade *do-mi-si* et l'intervalle de septième *do-si*. C'est le cas également de l'évolution de la concordance totale. Le calcul de ces deux descripteurs statiques dépend du nombre de notes, contrairement aux descripteurs dynamiques non triviaux. De là vient leur pertinence dans l'étude des transitions avec changement de la cardinalité, en particulier l'ajout d'une note.

Chapitre 9

Classes et classifications d'accords

Dans ce chapitre nous présentons d'abord le calcul des descripteurs harmoniques sur les classes d'accord, avant de l'appliquer à deux classifications et représentations d'accords : le potentiel combinatoire d'Estrada et le Tonnetz.

9.1 Calcul des descripteurs harmoniques sur les classes d'accords

9.1.1 Affranchissement de la dépendance à la disposition

Lequel, de l'accord mineur et de l'accord majeur, a la plus grande rugosité ? Ou bien, entre l'accord de septième diminuée et l'accord de septième de dominante, la plus grande tension ? Ou encore, entre la transition de la sous-dominante vers la dominante et la résolution de la dominante vers la tonique, quelle transition admet la plus grande concordance différentielle ?

De telles questions paraissent naturelles, mais y répondre en toute généralité exige de résoudre des problèmes de taille. Ces problèmes peuvent être pressentis dans les quelques réponses que l'on est déjà en mesure de donner, après l'étude des espaces de triades continus :

— Le demi-ton a une plus grande rugosité que le quart-de-ton sur *do* 3,

mais plus petite sur *do* 5¹.

- Les triades majeures et mineures ont globalement (c'est-à-dire considérées dans leur ensemble) la même rugosité et la même concordance, du fait de la symétrie entre les intervalles supérieurs et inférieurs². Cependant, les valeurs de rugosité et de concordance dépendent de la disposition de l'accord³.
- La concordance d'ordre 3 met en avant les triades mineures, et non les triades majeures. Cependant, avec un timbre à 7 partiels, seuls l'état fondamental et le premier renversement est concerné⁴, le second renversement n'étant mis en valeur qu'à partir du 8^e partiel.
- Un accord majeur à 5 notes avec un unisson à 3 notes a une concordance d'ordre 3 plus grande qu'une triade mineure.

Ces affirmations sont vraies avec les spectres utilisés dans le chapitre 8, mais ne le sont plus nécessairement avec d'autres spectres.

Outre la dépendance des descripteurs harmoniques au timbre, les exemples précédents mettent en avant un autre obstacle dans la réponse à ces questions. Il s'agit de la dépendance des descripteurs harmoniques :

- Au registre,
- Au renversement,
- Aux doublures,
- Plus généralement, à la disposition des accords.

En effet, les notions d'accord majeur, mineur, diminué, de septième de dominante ne se réfèrent pas à des accords, mais à des classes d'accord, en l'occurrence des classes normales⁵.

Le calcul des descripteurs harmoniques sur les classes d'accords nécessite donc de pouvoir se départir de la dépendance à la disposition des accords. Une classe d'accord, qu'elle soit un ensemble de hauteurs, une classe normale, première ou une identité, est une abstraction, qui admet une grande multiplicité de réalisations. On peut tirer parti de la variation des descripteurs sur ces réalisations pour étudier les différentes dispositions. Cependant, dans

1. Voir figure 8.4.
2. Voir figure 8.15.
3. Voir figure 8.16.
4. Voir figure 8.26.
5. Voir page 65.

le cas qui nous intéresse, nous souhaitons rendre les descripteurs invariants sur ces réalisations. Cela ne peut se faire impunément, car ce qu'on gagne en abstraction, on le perd nécessairement dans le lien avec la réalité sonore et perceptive.

État de l'art La littérature offre relativement peu d'exemples de calcul de descripteurs harmoniques sur les classes d'accords. La contribution majeure est celle de Huron ([Huron, 1994](#)), présentée page [260](#). Huron associe une valeur de consonance à toutes les classes normales obtenues à partir de l'échelle chromatique, afin de tester deux hypothèses, qui sont les suivantes :

- *Hyp 1* : Les échelles utilisées en musique (« échelles » pris au sens large, incluant les gammes, les modes, les accords) sont-elles celles qui maximisent la consonance ?
- *Hyp 2* : Les échelles utilisées sont-elles celles dans lesquelles le nombre d'occurrences des intervalles est proportionnel à leur consonance ?

Afin de se départir de la multitude de réalisations des classes normales, la méthode de Huron est la suivante : il associe à chaque classe d'intervalle une valeur de consonance, en additionnant les consonances des deux réalisations de cette classe d'intervalle ne dépassant pas l'octave (par exemple tierce mineure et sixte majeure), puis multiplie chaque valeur de consonance d'une classe d'intervalle par le nombre de ses occurrences dans la classe d'accord. En sommant sur les six classes d'intervalles, il obtient une valeur de consonance de classe, l'"*aggregate dyadic consonance*" (ADC). La consonance des intervalles, quand à elle, est calculée en amalgamant les données expérimentales de trois études ([Malmberg, 1918](#); [Hutchinson et Knopoff, 1979](#); [Kameoka et Kuriyagawa, 1969b](#)).

La table des échelles les plus consonantes par cette méthode est reproduite à la page [262](#).

Comme le montrent ses hypothèses, Huron ne s'occupe pas ici de la dimension perceptive de la consonance de classe. Celle-ci est d'ailleurs problématique a priori, puisque si la consonance traduit un aspect de la perception, sa variation sur les réalisations d'une classe d'accords traduit à son tour une variation perceptive des différentes réalisations. Dès lors, une consonance de classe, qui ne différencie pas les réalisations, serait éloignée de la réalité per-

ceptive.

Cependant, Kuusi ([Kuusi, 2010](#)), dans une étude visant à déterminer les rôles respectifs des classes d'accords et de leurs réalisations dans la perception des accords non traditionnels de 4 et 5 notes, montre que certaines propriétés de classe influencent la perception. Parmi elles figure la consonance de classe de Huron, à côté de la cardinalité (nombre de notes) et de l'« implication tonale » (présence de sous-accords majeurs ou mineurs).

9.1.2 Spectres de Shepard

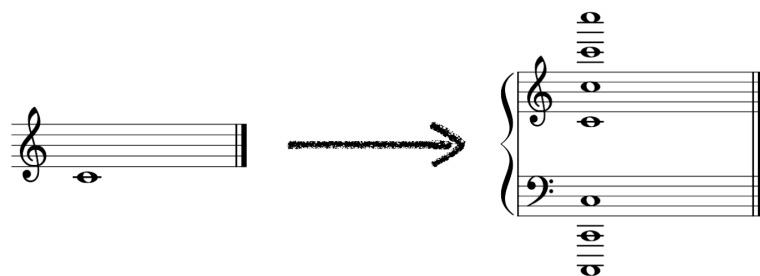
L'approche de Huron ne saurait nous satisfaire, cela pour deux raisons :

- La consonance d'une classe d'intervalle est obtenue en sommant les consonances de ses deux réalisations ne dépassant pas l'octave. Ce choix semble arbitraire, et ne tient pas compte des autres réalisations de la classe d'intervalle.
- Les descripteurs présentés dans le chapitre 6 sont des descripteurs spectraux, prenant en entrée des spectres en amplitude. La généralisation des descripteurs aux classes d'accords nécessite donc d'avoir un équivalent du spectre des notes, mais pour les hauteurs¹, ce que ne fournit pas l'approche de Huron.

Aussi adoptons-nous ici une autre approche. Nous cherchons à associer un spectre à une hauteur, et non plus à une note, autrement dit à avoir un spectre associé à *do*, à *ré*, ou à *mi*, sans avoir à préciser l'octave. Une solution consiste à démultiplier les partiels d'une note *do* sur toutes les octaves de l'ambitus disponible, ce qui est illustré par la figure ci-dessous. Cela a du sens seulement lorsqu'on admet l'équivalence des octaves. On se placera pour se faire dans le cadre des sons harmoniques. Dans le cas de sons non harmoniques non octavians, la méthode pourra être adaptée en remplaçant l'octave par un intervalle idoine², à la fois dans la définition des hauteurs, qui deviendront des fréquences modulo cet intervalle, et dans la construction des spectres de Shepard.

1. Voir page 60

2. Voir ([Sethares, 1999](#), chap.1) : « *The octave is dead* » (« l'octave est morte »).



En pratique, nous utilisons les spectres de Shepard généralisés aux sons complexes, présentés dans la partie 5.2.2. Sur ces spectres, tous les partiels sont démultipliés sur toutes les octaves, avec une enveloppe en gaussienne pour atténuer les fréquences extrêmes¹. Ce faisant, un spectre devient associé à une fréquence modulo l'octave, autrement dit une hauteur, et non plus à une fréquence fondamentale.

Sur les accords, du moment qu'est imposée la condition d'absence de doublure, l'usage des spectres associés aux hauteurs permettent de se départir de la dépendance à la disposition. En effet, les notions de renversement ou de disposition n'ont plus alors de sens. Il devient alors possible d'associer un spectre artificiel aux ensembles de hauteurs *Do M* ou *Do m*. De plus, en fixant arbitrairement une hauteur de référence à partir de laquelle est construit l'accord, on peut s'abstraire des hauteurs absolues et associer un spectre à une classe normale, par exemple aux accords de septième de dominante.

Harmonie avec des sons de Shepard Les sons de Shepard peuvent être synthétisés à partir des spectres de Shepard. Plusieurs aspects perceptifs de l'harmonie créée avec de tels sons ont été étudié par Parncutt ([Parncutt, 1993](#)) : la multiplicité, c'est-à-dire le nombre de sons perçus dans un accord, l'affinité de deux accords successifs et la similarité de deux accords, proche de notre notion de concordance. Parncutt crée ainsi un modèle de perception de l'harmonie avec les sons de Shepard, dans la continuité du modèle développé quelques années auparavant dans ([Parncutt, 1989](#)).

1. Voir figure 5.8.

9.1.3 Descripteurs de classes normales, descripteurs de classes premières et descripteurs dynamiques entre ensembles de hauteurs.

Descripteurs de classes normales et descripteurs de classes premières Tous les descripteurs harmoniques présentés dans la partie 6 sont généralisables aux ensembles de hauteurs, ainsi qu’aux transitions entre ensembles de hauteurs, en remplaçant les spectres associés à une note par les spectres de Shepard associés à une hauteur.

De plus, l’homogénéité de l’étalement du spectre sur toutes les hauteurs assure que le choix arbitraire de la hauteur de référence n’a pas d’influence sur le calcul des descripteurs harmoniques statiques, ce qui permet d’abstraire le calcul de ces descripteurs aux classes normales.

La rugosité et la concordance décrivent les interactions spectrales entre les notes d’un accord prises deux à deux, elles ne dépendent que des intervalles présents dans l’accord. De ce fait, la rugosité et la concordance sur les ensembles de hauteurs ne dépendent que du contenu intervallique¹. On peut donc les calculer sur les classes homométriques², et, a fortiori, sur les classes premières³.

Ce n’est pas le cas des autres descripteurs considérés, qui ne sont pas constants sur les classes premières. On a vu, par exemple⁴, que la concordance d’ordre 3 met en avant les accords mineurs, et non pas les accords majeurs, qui appartiennent pourtant à la même classe première.

Descripteurs dynamiques entre ensembles de hauteurs Le calcul des descripteurs dynamiques, contrairement à celui des descripteurs statiques, ne peut pas se faire au niveau des classes normales, mais seulement au niveau des enchaînements entre ensembles de hauteurs. En effet, les interaction spectrales entre deux hauteurs consécutives dépendent de la classe d’intervalle entre ces deux hauteurs. Il n’y a donc pas de sens à les rendre équivalentes en quotientant chacune par la relation de transposition.

1. Le contenu intervallique est défini page 63.

2. Une classe homométrique regroupe les ensembles de hauteurs ayant le même contenu intervallique, voir page 64.

3. La partition en classes premières affine la partition en classes homométriques.

4. Voir figure 8.25.

9.1.4 Espaces continus de classes de dyades et de triades

Nous avons défini dans l'introduction les ensembles de hauteurs, les classes normales, les classes premières et les identités. Nous mettons de côté ici les classes homométriques. Les ensembles de hauteurs forment une sous-partition des classes normales, lesquelles forment une sous-partition des classes premières, elles-mêmes sous-partition des identités. Ces notions ont du sens sur les échelles musicales octaviantes ou sur un *continuum* des fréquences.

Classes d'intervalles Dans le cas des intervalles, les identités ne se distinguent pas des classes premières, qui elles-mêmes ne se distinguent pas des classes normales. On parle de classe d'intervalles.

Le passage d'un intervalle à une classe d'intervalle consiste en deux étapes :

- Réduction de l'intervalle à un intervalle inférieur à l'octave par changement d'octave,
- Réduction du nouvel intervalle à sa distance absolue à la quarte augmentée (dans l'échelle tempérée) qui divise l'octave en deux intervalles égaux.

Les intervalles complémentaires se retrouvent alors dans la même classe d'intervalles. Avec l'échelle chromatique, on obtient ainsi 7 classes d'intervalle. Dans le cas d'un continuum fréquentiel, les classes d'intervalles peuvent se représenter par le segment entre 0 et 6 inclus, 0 correspondant aux unissons ou aux octaves, et 6 aux quartes augmentées modulo l'octave, comptés en demi-tons.

Présentation réduite d'une triade On peut ramener toute triade à une triade ne dépassant pas l'octave, en abaissant si besoin l'octave des deux notes supérieures, et à cette nouvelle triade associer sa présentation réduite¹. Celle-ci est directement visualisable sur le plan des triades, comme le montre la figure 9.1 où deux accords ont été localisés : l'accord *do-sol[#]-la*, de présentation réduite [8,1,3], et l'accord majeur *do-mi-sol*, de présentation réduite [4,3,5].

Les deux premières composantes X et Y de la présentation réduite [X,Y,Z] correspondent aux coordonnées sur le plan des triades. La troisième composante Z se lit comme la distance horizontale (ou verticale) entre le point de

1. Voir page 1.5.2

localisation de l'accord et la grande diagonale passant par (0,12) et (12,0), dite dans la suite « diagonale de l'octave » . On a alors $X + Y + Z = 12$. L'ensemble des présentations réduites des triades est formé des triplets $[X,Y,Z]$ vérifiant cette égalité.

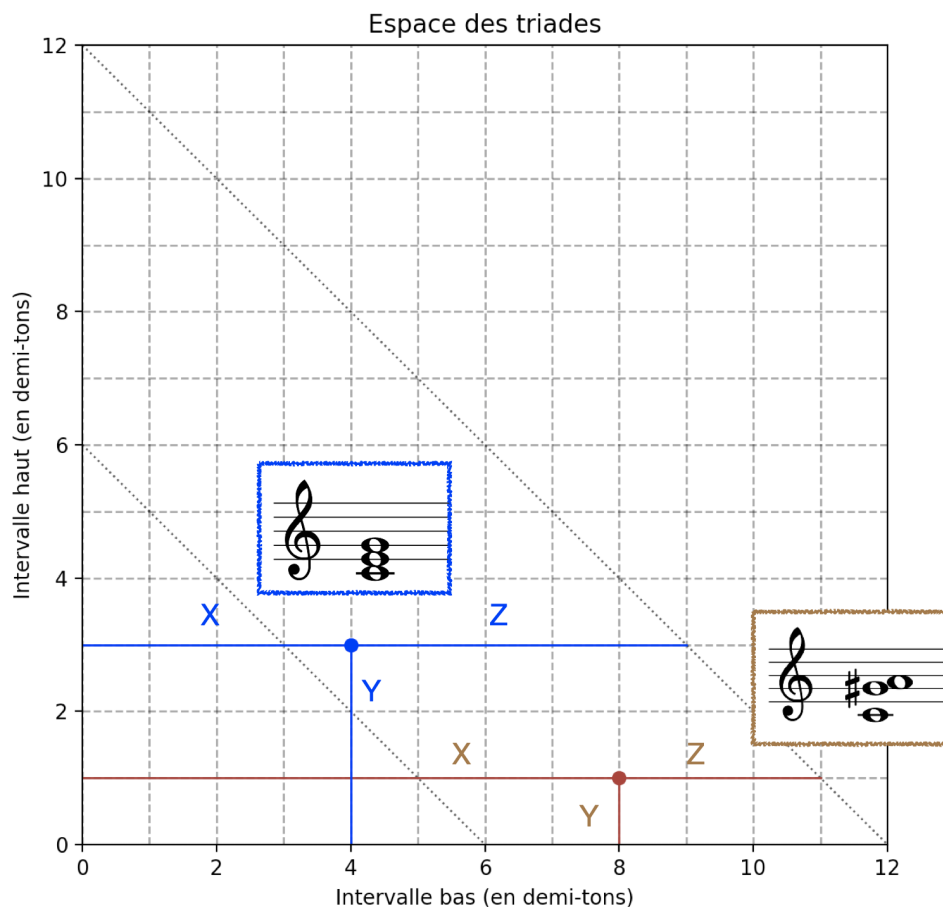


FIGURE 9.1 – Présentation réduite - visualisation sur l'espace des triades. Deux accords sont localisés, de présentation réduite $[X,Y,Z]$. Chacune des composantes de la présentation réduite est visualisable.

Classes normales de triades Nous avons défini dans l'introduction¹ les classes normales d'accords, caractérisées par leur forme normale. Pour les

1. Voir page 65.

triades, sont regroupés dans la même classe normale les accords dont les présentations réduites sont des permutations cycliques l'une de l'autre. La forme normale correspond à la présentation réduite dont la dernière composante majore les deux premières; et si plusieurs présentations réduites vérifient cette propriété, la forme normale est celle d'entre elles dont la première composante minore les autres. Ainsi, parmi les deux accords représentés sur la figure précédente, l'accord majeur est bien une forme normale, mais ce n'est pas le cas de l'autre accord, de présentation réduite $[8,1,3]$, dont la forme normale est $[1,3,8]$.

On identifiera désormais une classe normale avec la présentation réduite de sa forme normale.

Localisons maintenant sur le plan des triades le lieu des formes normales. Pour qu'une présentation réduite $[X,Y,Z]$ soit une forme normale, il est nécessaire que $X, Y \leq Z$. Les présentations réduites vérifiant $X + Y + Z = 12$, la condition de majoration des deux premières composantes par la dernière est équivalente à :

$$\begin{cases} Y \leq 12 - 2X \\ Y \leq 6 - \frac{X}{2} \end{cases}$$

De plus, si $[X,Y,Z]$ est une forme normale, $X = Z$ implique $X = Y = Z$, ce qui exclue la droite $Y = 6 - \frac{X}{2}$ à l'exception de l'accord $[4,4,4]$.

Ainsi, une présentation réduite $[X,Y,Z]$ avec $X, Y, Z > 0$, est une forme normale si et seulement si :

$$\begin{cases} Y \leq 12 - 2X \\ Y \leq 6 - \frac{X}{2} \text{ ou } (X, Y, Z) = (4, 4, 4) \end{cases} \quad (9.1)$$

La figure 9.2 montre le lieu des formes normales dans le plan des triades. On a représenté les réalisations au sein de l'octave de plusieurs classes normales.

Mis à part la classe de forme normale *do-mi-sol* $\sharp$ (en rouge) qui n'est réduite qu'à un représentant sur l'ambitus d'une octave, toutes les autres classes ont trois représentants dans l'espace des triades ne dépassant pas l'octave, qui correspondent à la position fondamentale et aux deux renversements de

la forme normale. En bleu foncé et en marron sont indiquées les classes des accords de la figure 9.1. La classe en vert, de forme normale *do-ré-sol*, est symétrique par rapport à la grande diagonale passant par l'origine, ce qui montre la nécessité d'exclure le bord d'équation $Y = 6 - \frac{X}{2}$.

Classes premières et identités de triades Dans le cas particulier des triades, les identités sont confondues avec les classes premières, ce qui n'est plus vrai à partir de 4 notes. La forme première d'une triade¹ s'obtient en mettant dans l'ordre croissant les composantes de sa présentation réduite. Ainsi, l'accord majeur en position fondamentale, de présentation réduite $[4,3,5]$, a pour forme première $[3,4,5]$, qui correspond à l'accord mineur en position fondamentale. Soit la forme première s'identifie à la forme normale, soit elle est obtenue à partir de cette dernière en échangeant les deux premières composantes, ce qui revient à une inversion².

Ainsi, une présentation réduite $[X,Y,Z]$, avec $X, Y, Z > 0$, est une forme première si et seulement si

$$X \leq Y \leq 12 - 2X. \quad (9.2)$$

La figure 9.3 montre le lieu des formes premières dans le plan des triades. On a représenté les réalisations dans l'ambitus de l'octave de plusieurs classes premières, pour illustrer différents comportements.

Les classes possèdent toutes une symétrie par rapport à la diagonale passant par l'origine. En effet, l'échange des intervalles inférieur et supérieur laisse invariante la classe première. Les classes en bleu et en marron correspondent aux classes des accords représentés figure 9.1. Elles sont toutes deux de cardinal 6. En vert et en noir, des classes à trois représentants, dont la forme première est située sur le bord de l'ensemble des formes premières. Enfin, l'accord *do-mi-sol*[♯] est le seul représentant de sa classe première.

Les exemples précédents nous permettent d'énoncer les propriétés suivantes, qui se vérifient sans difficulté particulière : les classes premières dont la forme première est située sur la frontière de l'ensemble des formes premières ne sont constituées que d'une classe normale. Par contre, celles dont la forme première est située à l'intérieur de l'ensemble de formes premières sont constituées de deux classes normales.

1. Voir page 65
2. Voir page 1.5.2

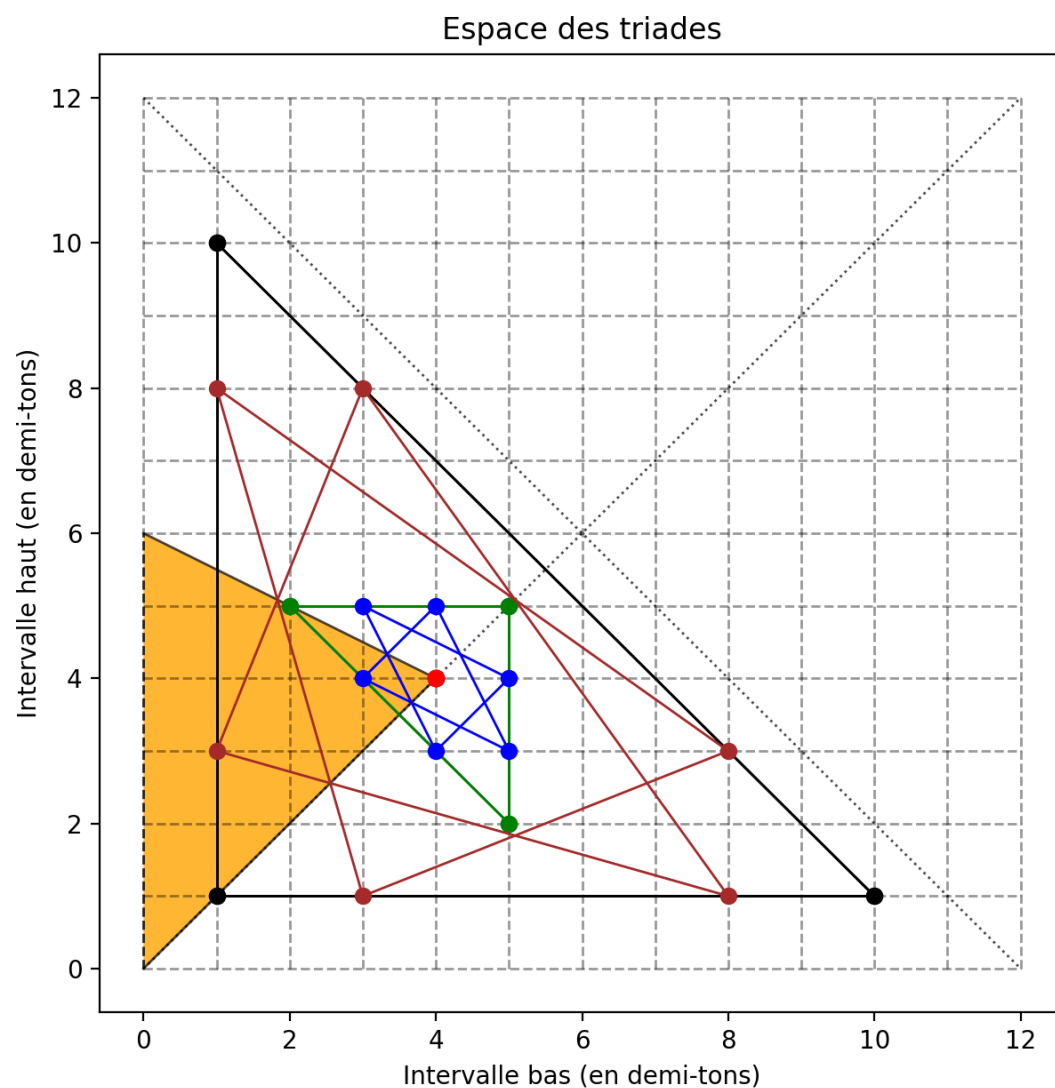


FIGURE 9.3 – Localisation des formes premières dans le plan des triades (en jaune). Cet espace est assimilé à l'espace des classes premières. Les traits pleins sur le bord indique l'appartenance, et le trait en pointillés l'exclusion. Sont représentées les réalisations dans l'ambitus de l'octave de plusieurs classes premières.

9.1.5 Structuration des classes d'intervalles

Shepard vs. non Shepard : comparaison des courbes de concordance

Avant de calculer les descripteurs harmoniques statiques sur les classes d'intervalles, comparons les courbes de concordance sur une octave obtenues avec des spectres harmoniques et avec des spectres de Shepard. La figure 9.4 met en parallèle les deux courbes.

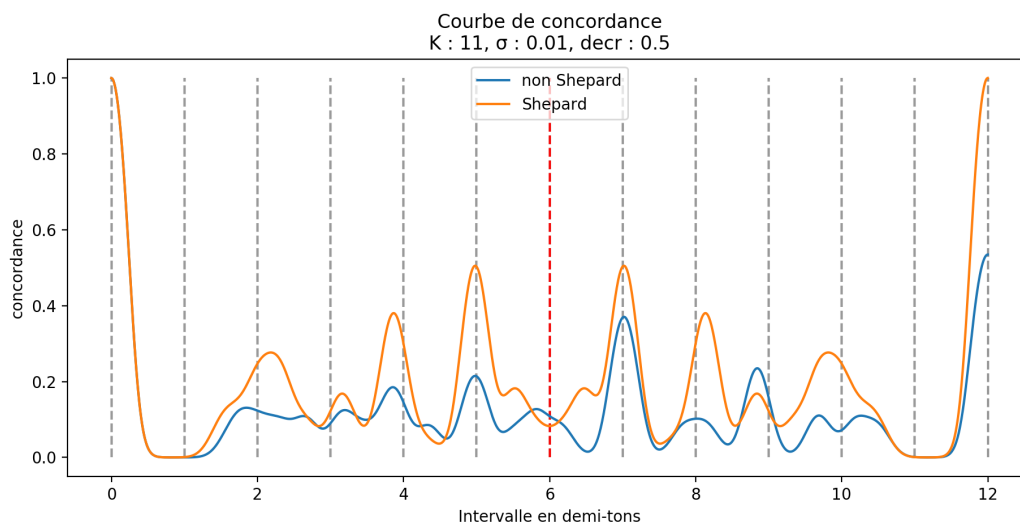


FIGURE 9.4 – Shepard vs. non Shepard : comparaison des courbes de concordance, normalisées à 1 sur l'unisson. En bleu : courbe obtenue avec des spectres modélisant les spectres réels (voir figure 8.5). En orange : courbe obtenue avec des spectre de Shepard. Paramètres de spectres : $K = 11$, $\sigma = 0.01$ et $decr = 0.5$.

La courbe obtenue avec les spectres de Shepard présente un axe de symétrie autour de l'intervalle de 6 demi-tons, ce qui n'est pas le cas de l'autre courbe. Cela vient du fait qu'au sein de l'octave les classes d'intervalle sont déterminées par la distance à cet intervalle. Cette disparité crée de grandes différences qualitatives : avec les spectres de Shepard, la concordance de la sixte mineure, qui était plus petite que celle de la tierce majeure, lui devient égale, ou encore la seconde majeure devient plus concordante que la tierce mineure.

Nous représentons désormais les courbes de descripteurs sur les classes

d'intervalles en se restreignant à la représentation des classes d'intervalles de 0 à 6 demi-tons.

Courbes de descripteurs sur les classes d'intervalles La figure 9.5 montre les courbes de concordance, d'harmonicité et de rugosité sur les classes d'intervalles. Nous avons pris $\sigma = 0.005$ de manière à ce que les pics proches de concordance et d'harmonicité soient distinguables.

En dehors de la classe de l'unisson/octave, les classes d'intervalle de quarte/quinte et de tierce M/sixte m (en tempérament juste) sont celles qui maximisent la concordance et l'harmonicité, en même temps qu'elles minimisent la rugosité.

Le maximum de rugosité est situé entre le demi-ton et le ton, sur une classe proche de la classe du maximum de rugosité obtenue avec des spectres classiques et *do* 3 comme note basse¹. Notons que l'emplacement de maximum de rugosité dépend de la fréquence centrale de l'enveloppe du spectre de Shepard, fixée ici à *do* 4². La rugosité est le seul descripteur statique, parmi ceux envisagés, qui soit sensible à cette fréquence.

1. Voir figure 8.4.

2. Voir l'implémentation des spectres partie 5.2.2.

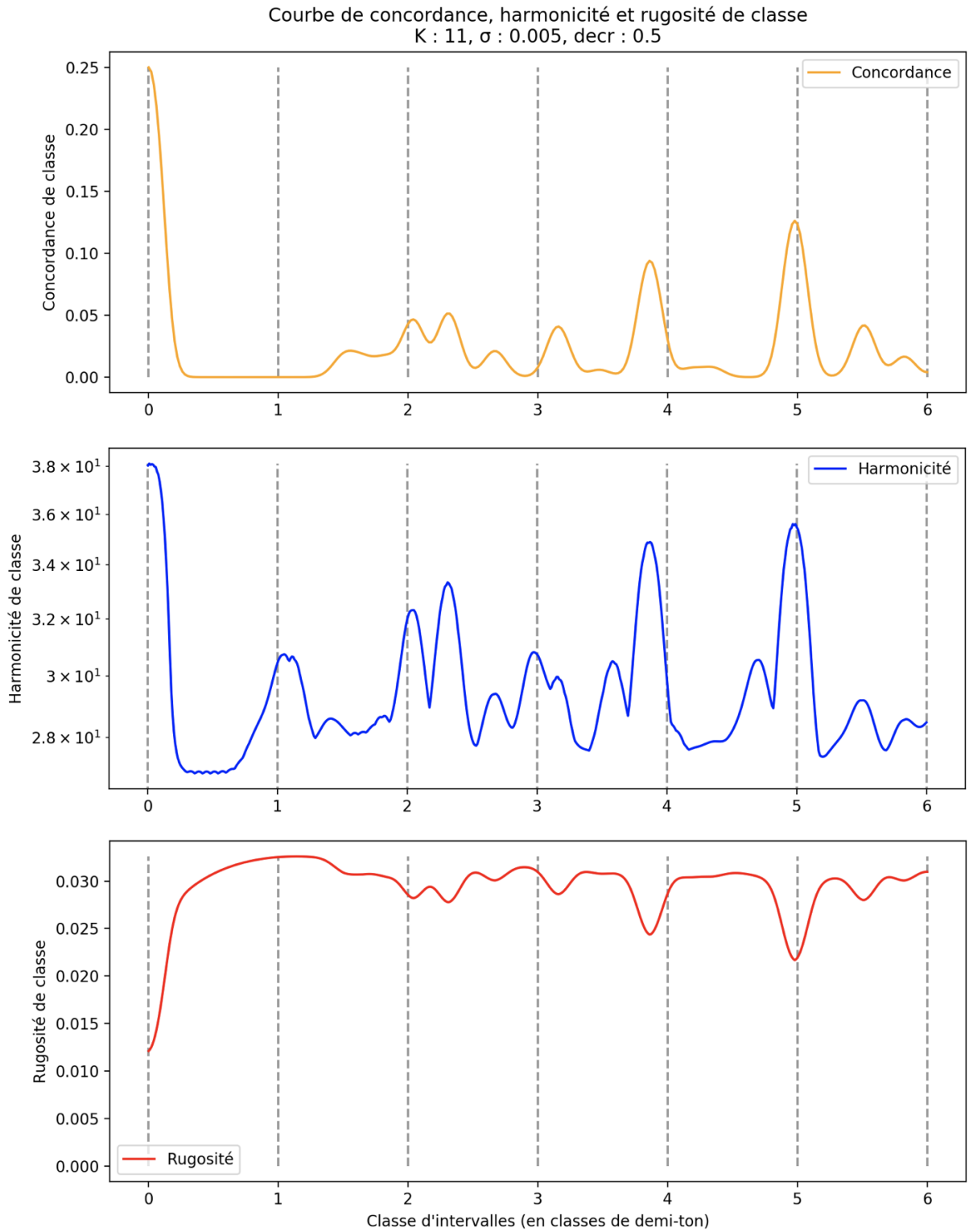


FIGURE 9.5 – Courbes de concordance (en orange), d’harmonicit  (en bleu) et de rugosit  (en rouge) de classe. Param tres de spectres : $K = 11$ (nombre de partiels), $\sigma = 0.005$ ( paisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude  gale   $1/\sqrt{k}$).

9.1.6 Topographie des classes premières de triades

Pour les intervalles, on ne distingue pas les classes normales des classes premières ou des identités, toutes trois étant confondues en classes d'intervalles. Pour les triades, les classes normales sont par contre des sous-classes des classes premières, lesquelles sont confondues avec les identités.

Les descripteurs statiques calculables sur les classes premières sont ceux qui ne dépendent, une fois les paramètres de spectre fixés, que du contenu intervallique. C'est le cas de la concordance et de la rugosité.

La figure 9.6 montre la surface de concordance des classes premières de triades. Nous avons pointés les classes premières de plus grande concordance, en indiquant les classes normales qui les constituent. La classe de plus grande concordance est représentée par l'accord *do-ré-sol*, vient ensuite la classe des accords majeurs et mineurs. En effet, l'accord *do-ré-sol*, de présentation réduite [2,5,5] possède deux occurrences de la classe d'intervalle la plus consonnante, en dehors de l'unisson/octave : la classe de la quarte/quinte.

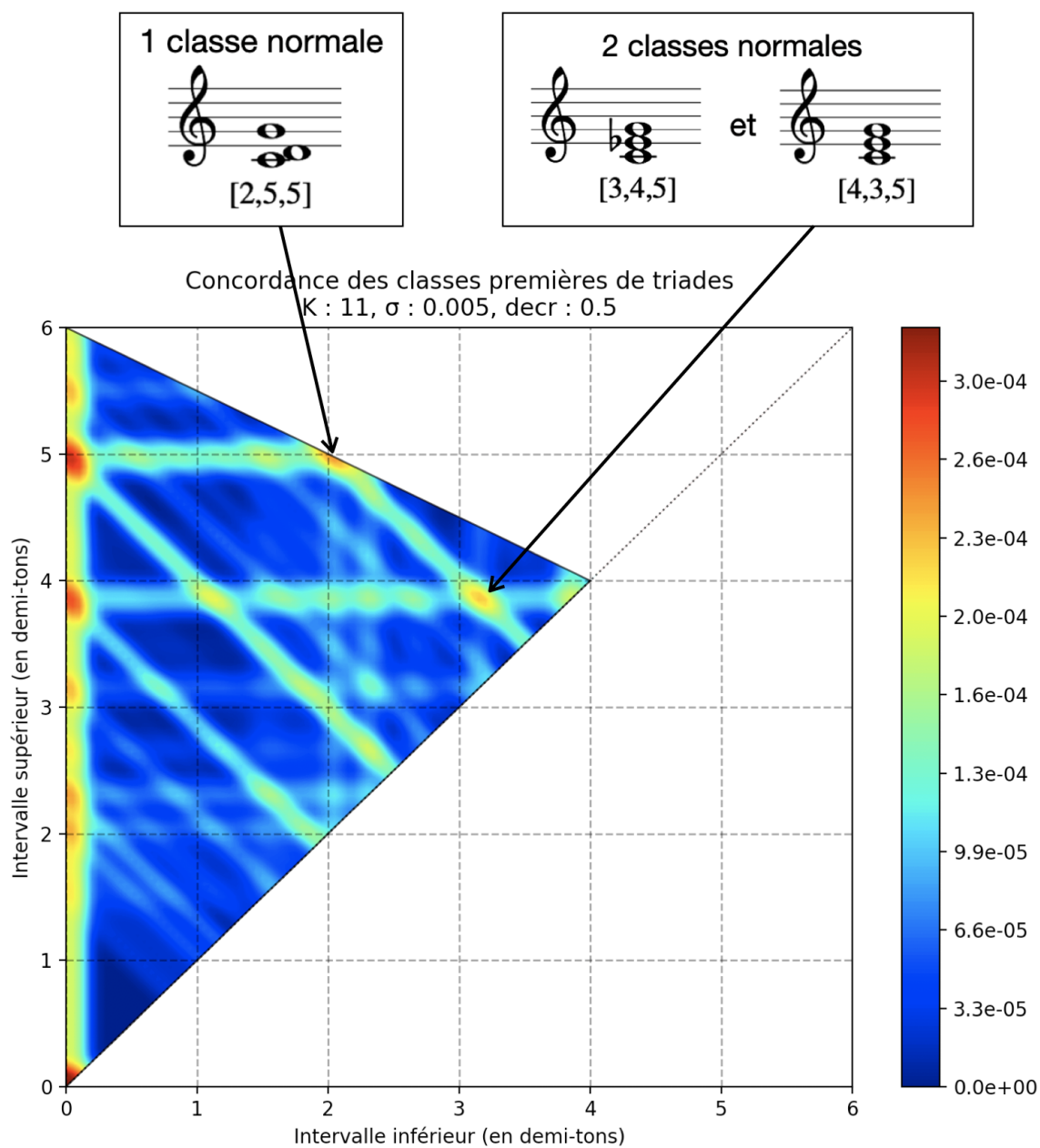


FIGURE 9.6 – Surface de concordance des classes premières de triades. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $\text{decr} = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

La figure 9.7 montre la surface de rugosité des classes premières de triades.

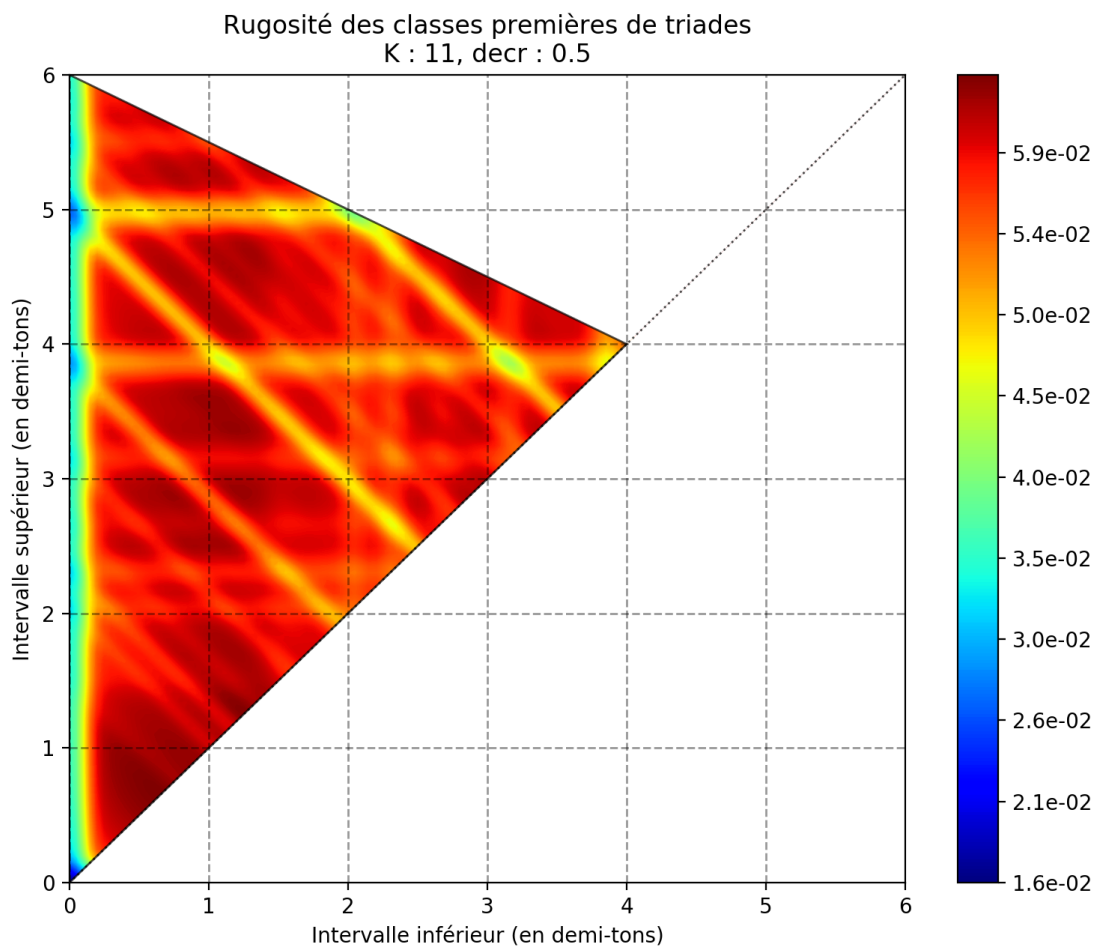


FIGURE 9.7 – Surface de rugosité des classes premières de triades. Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

9.1.7 Topographie des classes normales de triades

La concordance d'ordre 3, l'harmonicit  et la tension sont calculables sur les classes normales de triades.

La figure 9.8 montre la surface de concordance d'ordre 3 des classes normales de triades.

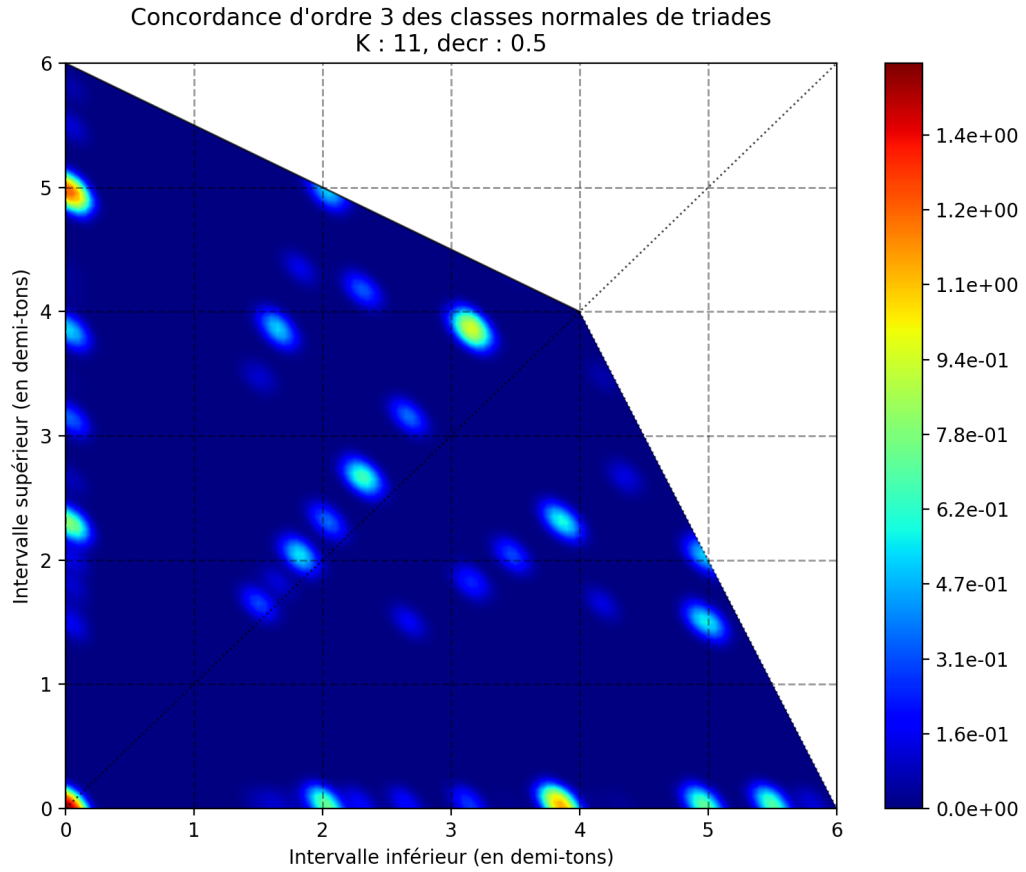


FIGURE 9.8 – Surface de concordance d'ordre 3 des classes normales de triades. Param tres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ ( paisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude  gale   $1/\sqrt{k}$). Pour une meilleure visualisation on a repr sent  $\log(1 + 100 * conc3)$.

La classe de plus grande concordance d'ordre 3 qui n'est pas situ e sur

les bords de l'espace des classes premières est la classe de l'accord mineur, représenté par l'accord *do-mib-sol*. De plus, on sait ¹ que la diagonale passant par l'origine est un axe de symétrie du bord de la surface, ce qui se voit par l'égalité de la concordance d'ordre 3 des accords *do-ré-sol* et *do-fa-sol*, qui appartiennent à la même classe normale.

La figure 9.9 montre la surface d'harmonicité des classes normales de triades.

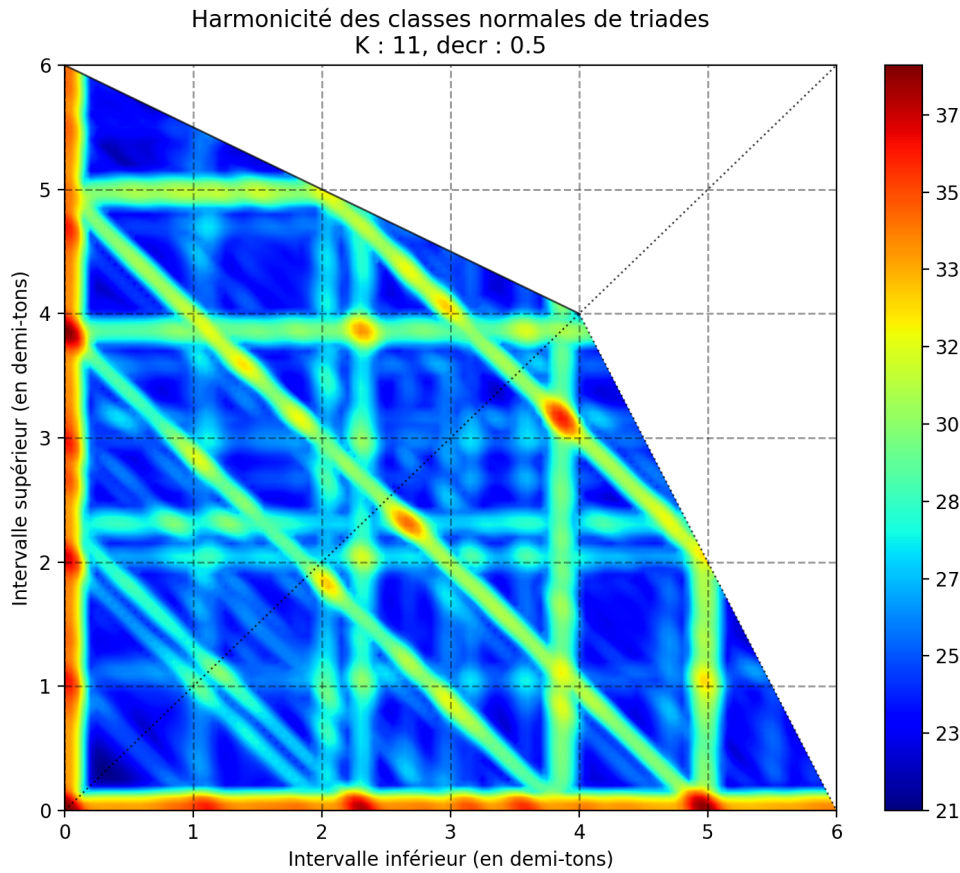


FIGURE 9.9 – Surface d'harmonicité des classes normales de triades. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

1. Voir figure 9.2

9.2 Continuum et discontinuum des classes : application au potentiel combinatoire des intervalles

9.2.1 Continuum et discontinuum

Dans la partie [9.1](#) nous avons considéré les classes normales, les classes premières et les identités obtenues à partir d'un continuum des fréquences. Les espaces de ces classes sont eux-même continus, au sens géométrique du terme. Ils constituent un espace des possibles, qu'une musique, du moins si elle est constituée de notes, ne pourra jamais épuiser. En effet, l'ensemble des identités présentes dans une œuvre donnée, ou encore associées à un langage ou à une échelle donnés, constitue un sous-espace dénombrable de l'espace continu des classes.

Dénombrement des classes : approche géométrique Les subdivisions de l'octave par les échelles tempérées créent un quadrillage de l'espace des classes. La figure [9.10](#) montre la localisation des identités de triades (confondues avec les classes premières) avec une subdivision de l'octave en demi-tons puis en quarts de ton.

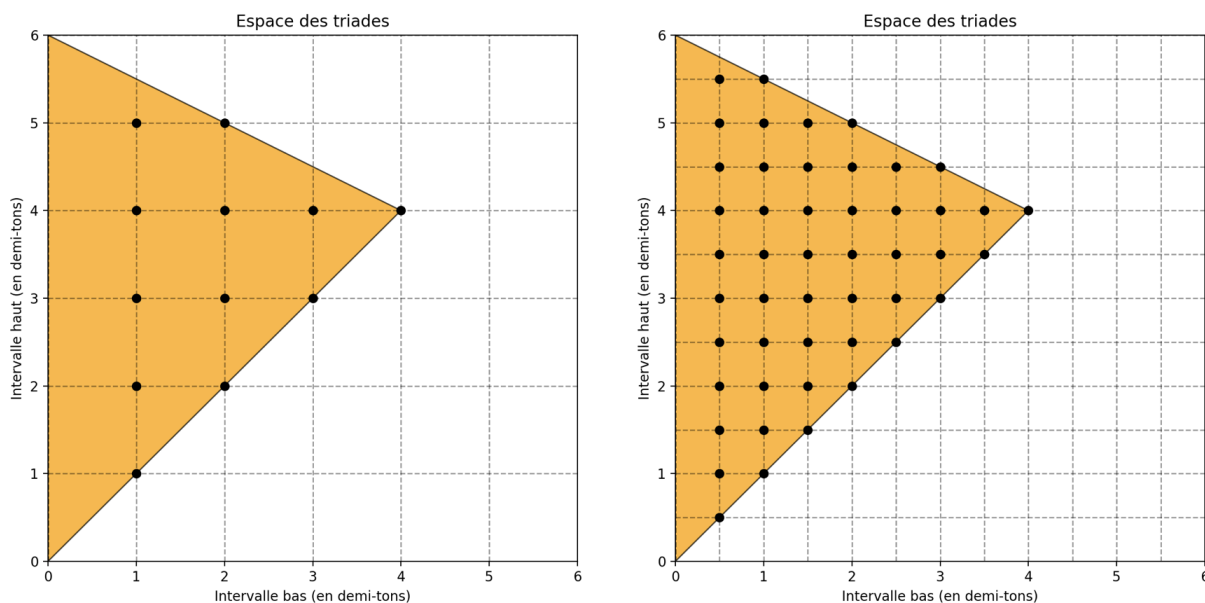


FIGURE 9.10 – Localisation des identités de triades avec une subdivision de l’octave en demi-tons (à gauche) et en quarts de ton (à droite). Avec les demi-tons, il y a 12 identités, contre $12 \times 2^2 = 48$ avec les quarts de ton. En jaune, l’espace des identités de triades.

L’espace des identités de triades étant une surface d’aire égale à 12 demi-tons au carré, et une « case » du quadrillage dans l’échelle des demi-tons correspondant à un demi-ton au carré, il y a avec cette échelle 12 identités de triades, contre $12 \times 2^2 = 48$ avec une subdivision en quarts de tons, $12 \times 3^2 = 108$ en sixièmes de tons, et ainsi de suite. L’aire de l’espace des classes normales étant deux fois plus grande¹, le nombre de classes normales tend, lorsque la subdivision devient de plus en plus petite, vers le double de celui des classes premières. L’inégalité est due aux effets de bords.

Le tableau 9.1 synthétise le nombre d’identités pour les accords de 2, 3, 4 et 5 notes dans plusieurs subdivision de l’octave : la subdivision en demi-tons, en quarts de ton et en sixièmes de ton.

1. Voir figure 9.2.

	Demi-tons	$\frac{1}{4}$ de ton	$\frac{1}{6^e}$ de ton
N = 2	6	12	18
N = 3	12	48	108
N = 4	15	108	351
N = 5	13	164	748

TABLE 9.1 – Table du nombre d’identités pour les accords de 2, 3, 4 et 5 notes en fonction de la subdivision de l’octave, obtenue par génération et dénombrement.

Pour un nombre de notes donné, le nombre de classes d’accords augmente avec la subdivision. Un raisonnement géométrique permet de comprendre la progression géométrique de cette augmentation. Nous avons déjà traité les cas des dyades et des triades. De manière générale, avec N notes, il est possible, en fixant la note inférieure, de représenter l’ensemble des accords à N notes dans un espace à $N - 1$ dimensions¹, où chaque axe correspond à un intervalle entre deux notes consécutives de l’accord. Les formes identitaires étant déterminées par N inégalités linéaires, leur espace est délimité par N hyperplans², c’est donc un $(N-1)$ -simplexe de $\mathbb{R}^{N-1}$, où $\mathbb{R}$ désigne les nombres réels.

Toute subdivision de l’octave en S intervalles égaux crée un quadrillage de l’espace des accords à N notes où chaque accord est associé à un volume élémentaire $\nu(N, S) = \left(\frac{12}{S}\right)^{N-1}$, compté en demi-tons à la puissance N .

Pour un nombre N de notes fixé, on a : $\nu(N, S) \xrightarrow{S \rightarrow \infty} 0$. Il en résulte que le volume « occupé » par les formes identitaires tend vers le volume de l’espace continu des formes identitaires quand la subdivision devient plus fine. Ainsi, on a :

$$\mathcal{P}(N, S) \nu(N, S) \xrightarrow{S \rightarrow \infty} \mathcal{V}(N),$$

où $\mathcal{P}(N, S)$ désigne le nombre d’identités des accords à N notes dans la subdivision S , et $\mathcal{V}(N)$ le volume de l’espace continu des formes identitaires

1. Voir figure 8.13

2. Un hyperplan dans un espace de dimension N est un plan de dimension $N-1$.

à N notes.

Au premier ordre en S , cela fournit une approximation du nombre de classes premières :

$$\mathcal{P}(N, S) \approx \left(\frac{S}{12}\right)^{N-1} \mathcal{V}(N) \quad (9.3)$$

Pour $N = 2$, $\mathcal{V}(2) = 6$ demi-tons, et l'approximation devient une égalité. Pour $N = 3$, $\mathcal{V}(3) = 12$, et les exemples du tableau 9.1 montre qu'on a aussi une égalité, puisque $\mathcal{P}(3, 12) = 12$ pour les demi-tons, $\mathcal{P}(3, 2 \times 12) = 48$ pour les quarts de ton et $\mathcal{P}(3, 3 \times 12) = 108$ pour les sixièmes de ton.

9.2.2 Le cas des tetrachordes

Intéressons-nous maintenant aux formes normales et aux identités des ensembles de 4 hauteurs. Un problème supplémentaire se pose par rapport aux triades : celui de la représentation. En effet, l'impossibilité de visualiser spatialement un espace à 4 dimensions rend complexe voir impossible la représentation des descripteurs harmoniques dans le volume continu des tetrachordes. La discrétisation de ce volume apporte la possibilité d'une visualisation.

Formes identitaires Commençons par déterminer le lieu des formes identitaires des tetrachordes. En généralisant la représentation plane des triades, on peut représenter les accords de 4 notes, la première fixée, dans un espace à 3 dimensions, où les trois axes correspondent respectivement aux intervalles entre les 1^{re} et 2^e, 2^e et 3^e, 3^e et 4^e notes.

Une présentation réduite $[X, Y, Z, W]$ est une forme identitaire si et seulement si :

$$\begin{cases} X + Y + Z + W = 12 \\ X \leq Y \leq Z \leq W. \end{cases}$$

Les formes identitaires sont donc déterminées par les inégalités suivantes :

$$\begin{cases} 0 < X \leq Y \leq Z \\ X + Y + 2Z \leq 12. \end{cases} \quad (9.4)$$

Le lieu des formes identitaires, délimité par les quatre plans provenant des équations précédentes, est représenté figure 9.11.

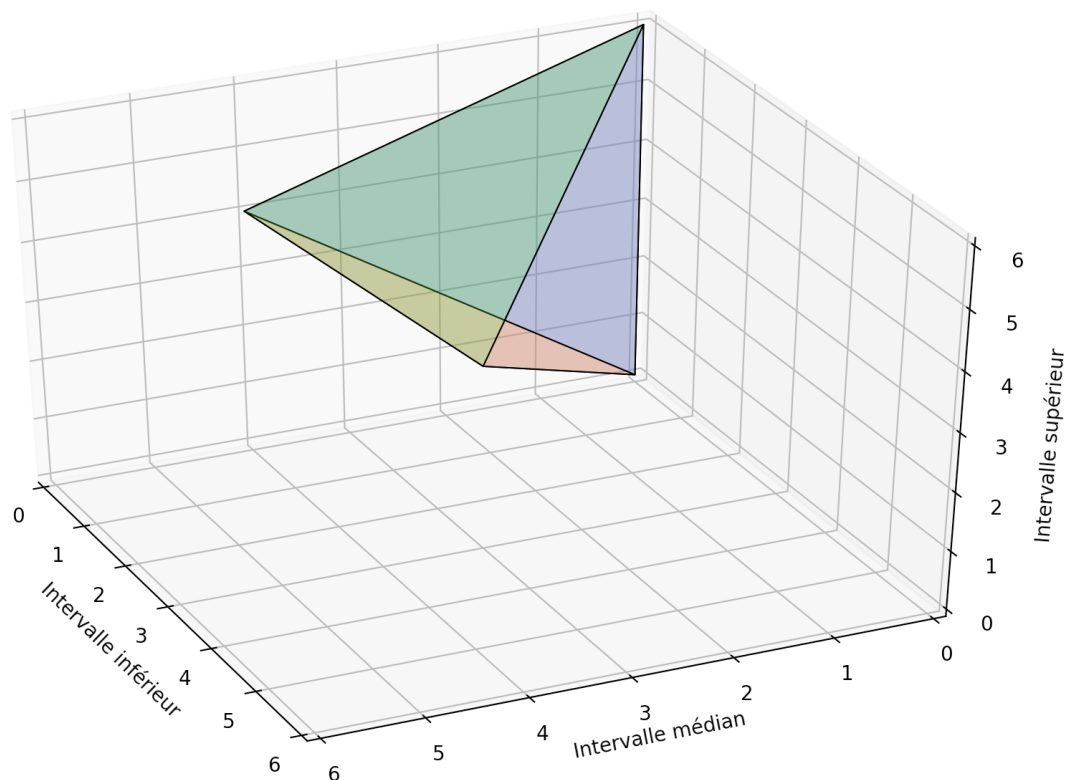


FIGURE 9.11 – Tetrachordes : localisation des formes identitaires

Comme dans le cas des triades, les formes identitaires dans une subdivision donnée de l'octave correspondent à l'intersection du quadrillage des accords issus de cette subdivision avec l'espace continu des formes identitaires. En se restreignant à un ensemble fini d'identités, il est possible de visualiser les valeurs des descripteurs statiques sur ces identités par un code couleur, de la même manière que dans le cas d'une caractérisation de l'espace continu de triades¹.

La figure 9.12 montre la localisation dans l'espace continu des formes identitaires des 15 formes identitaires issues d'une subdivision des fréquences

1. Voir partie 8.2.2.

en demi-tons. Leur rugosité est représentée par un dégradé de couleurs, où les minimums et maximums de rugosité sont associés respectivement aux deux limites du dégradé. Étant donné qu'une identité peut contenir plusieurs classes premières, ces couleurs sont associées aux formes identitaires ou à leur classe première, mais pas aux identités.

Les formes identitaires sont listées à droite dans l'ordre alphabétique de leurs composantes spatiales : elles sont d'abord classées dans l'ordre croissant de l'intervalle inférieur, puis de l'intervalle médian, puis de l'intervalle supérieur. La numérotation tient compte des identités des accords d'une note (une seule), de deux notes (au nombre de 6) et de trois notes (12).

On observe la tendance globale que la rugosité de la forme identitaire diminue avec le numéro de l'identité. En effet, les numéros bas correspondent à des formes premières resserrées, qui favorisent la rugosité, tandis que les dernières correspondent aux formes identitaires les moins resserrées.

Aucune relation claire de ce type n'est observable avec la concordance, représentée sur la figure 9.13. La forme identitaire la plus concordante est celle de l'accord *do-ré-mi-sol*.

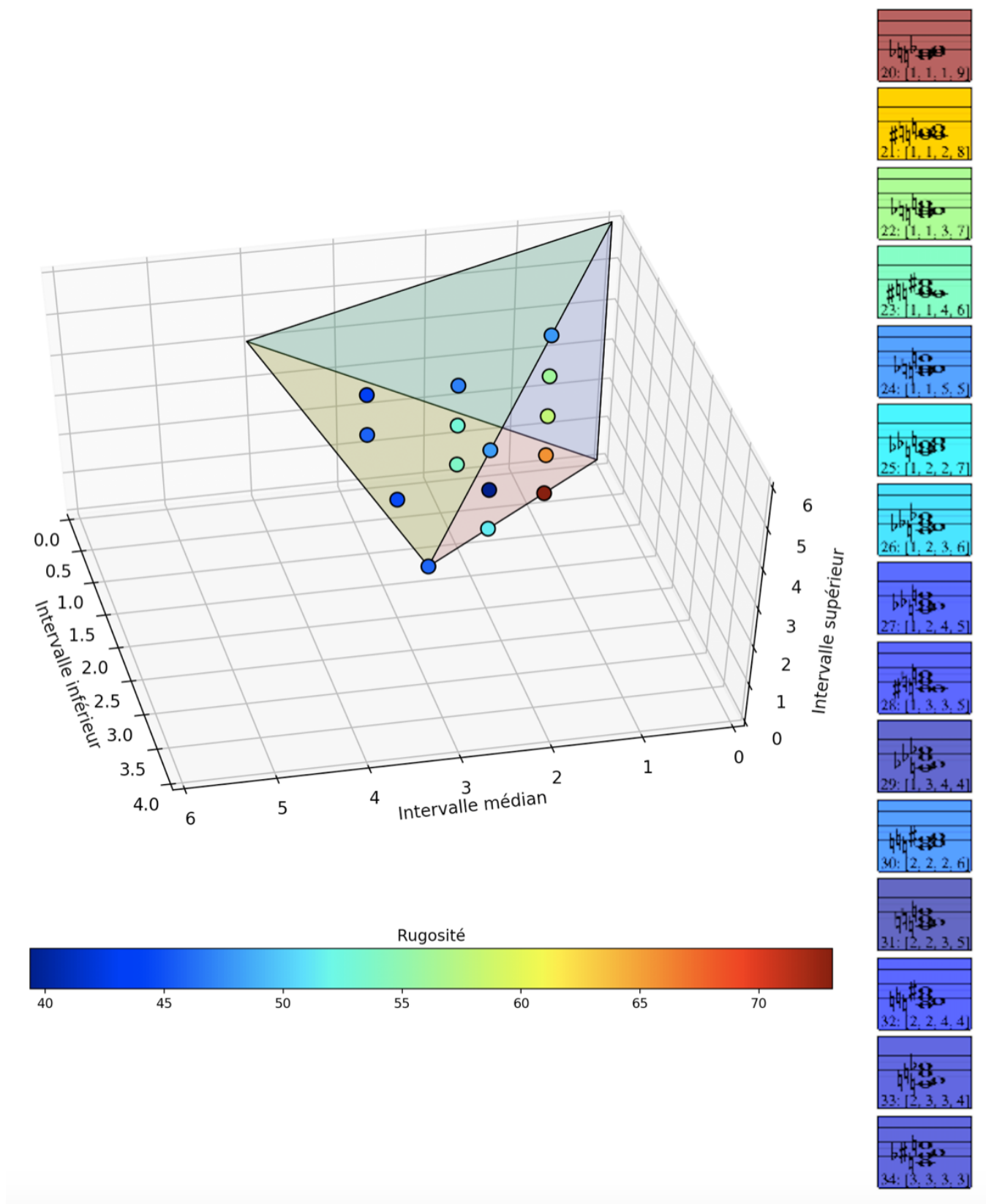


FIGURE 9.12 – Localisation dans l’espace continu des formes identitaires des 15 formes identitaires issues d’une subdivision en demi-tons, et représentation de leur rugosité sur une échelle de couleurs. Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

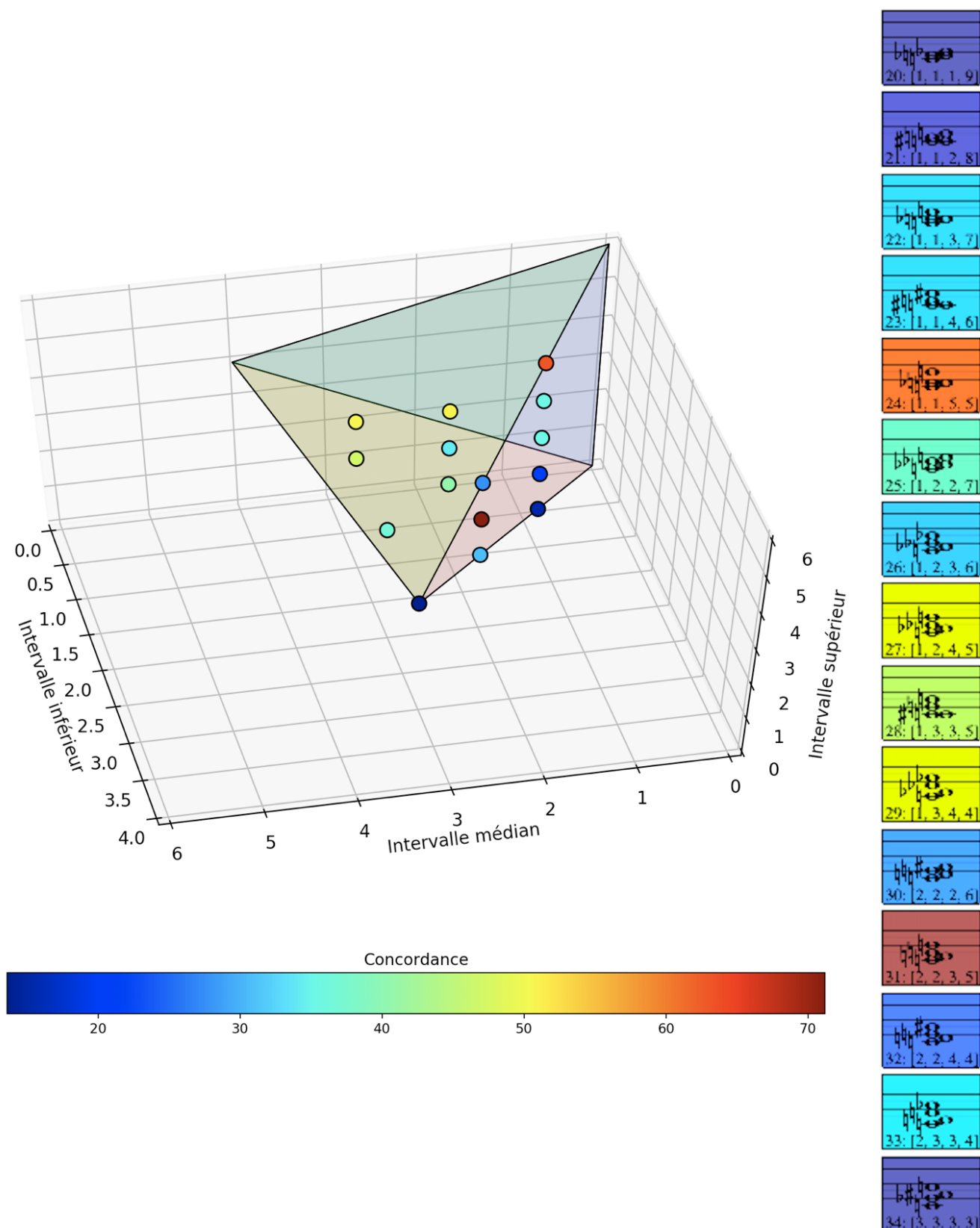


FIGURE 9.13 – Localisation dans l’espace continu des formes identitaires des 15 formes identitaires issues d’une subdivision en demi-tons, et représentation de leur concordance sur une échelle de couleurs. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Formes premières À partir d'une présentation réduite $[X, Y, Z, W]$, on obtient une présentation réduite de son inverse en échangeant les première et troisième coordonnées : $[Z, Y, X, W]$.

De là, étant donnée une forme identitaire « non-dégénérée », c'est-à-dire de la forme $[X, Y, Z, W]$ avec $X < Y < Z < W$, son identité contient trois classes premières, de formes premières $[X, Y, Z, W]$, $[Y, X, Z, W]$ et $[X, Z, Y, W]$. Le volume des formes premières de triades est ainsi trois fois plus grand que celui des formes identitaires.

Une forme réduite $[X, Y, Z, W]$ est une forme première si et seulement si

$$\begin{cases} X \leq Z \\ Y, Z \leq W \end{cases}$$

Ainsi, les formes premières sont déterminées par les inégalités suivantes :

$$\begin{cases} 0 \leq X, Y \\ X \leq Z \\ X + Y + 2Z \leq 12 \\ X + 2Y + Z \leq 12 \end{cases} \quad (9.5)$$

Le lieu des formes premières est représenté figure [9.14](#).

L'espace des formes premières constitue l'espace optimal pour calculer les descripteurs de classes premières que sont la rugosité et la concordance de classe, au sens où c'est le plus petit espace sur lequel la connaissance de ces descripteurs permet d'inférer leur valeur sur tous les ensembles de hauteurs.

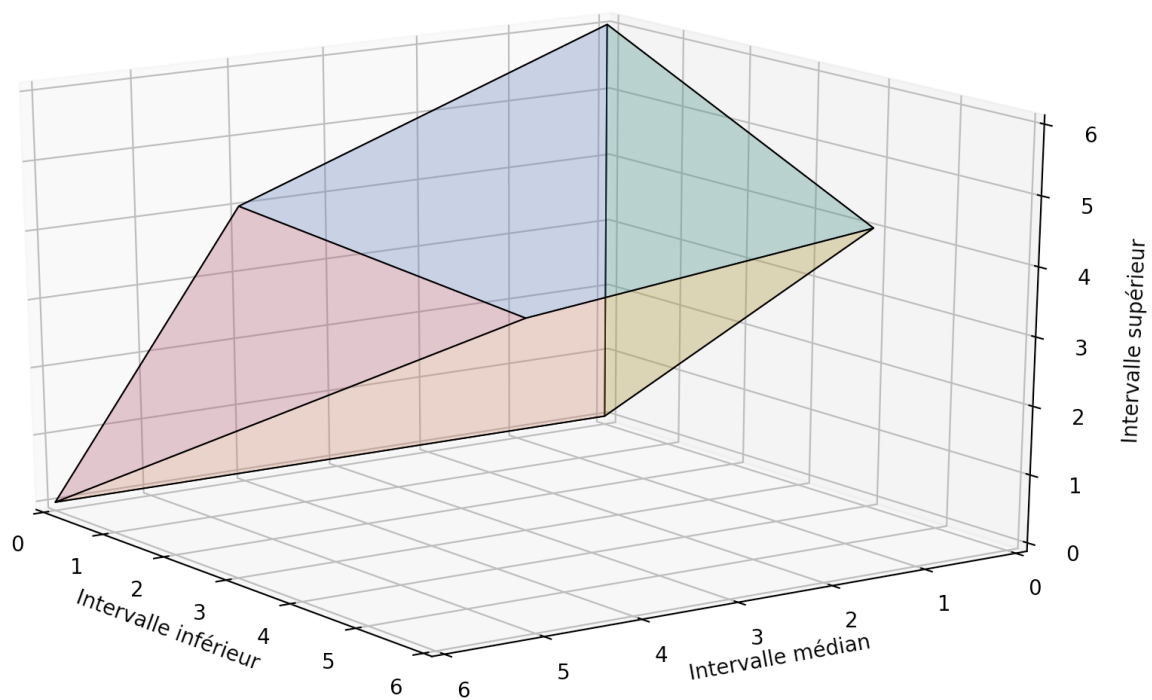


FIGURE 9.14 – Tetrachordes : localisation des formes premières

Formes normales Les formes normales, quant à elles, sont déterminées par les inégalités suivantes :

$$\left\{ \begin{array}{l} 0 \leq X, Y, Z \\ X + Y + 2Z \leq 12 \\ 2X + Y + Z \leq 12 \text{ avec } (2X + Y + Z = 12 \Leftrightarrow X = Y = Z = 3) \\ X + 2Y + Z \leq 12 \text{ avec } (X + 2Y + Z = 12 \Rightarrow X \leq Z) \end{array} \right. \quad (9.6)$$

Le lieu des formes normales, délimité par les six plans provenant des équations précédentes, est représenté figure 9.15.

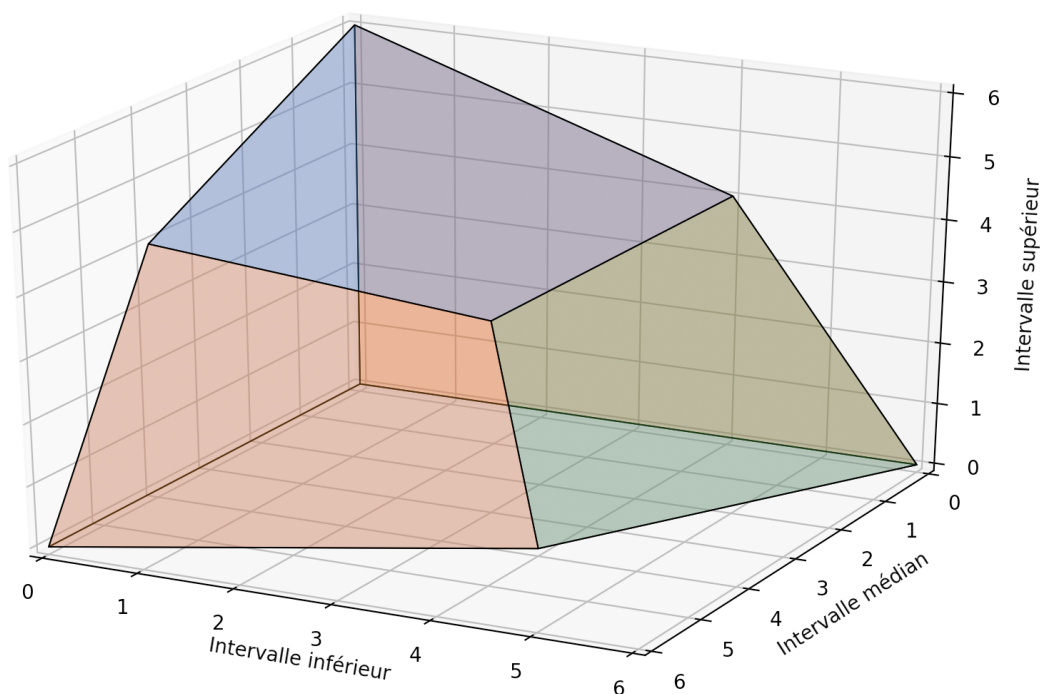


FIGURE 9.15 – Tetrachordes : localisation des classes normales

La frontière constituée par la face verte est exclue de l'espace des formes normales, mis à part l'accord diminué *do-mib-solb-la*, ainsi qu'une partie de de la face orange.

Orbites identitaires L'ensemble continu des classes normales a un volume six fois supérieur au volumes des classes identitaires, et deux fois supérieur à celui des classes premières. En effet, chaque classe identitaire n'étant pas située sur la frontière de l'ensemble continu des classes identitaires¹ est constituée de six classes normales, comme le montre la figure 9.16, qui représente, à l'intérieur de l'espace des formes normales, l'identité associée à l'accord *do-réb-mib-sol* de présentation réduite [1,2,4,5].

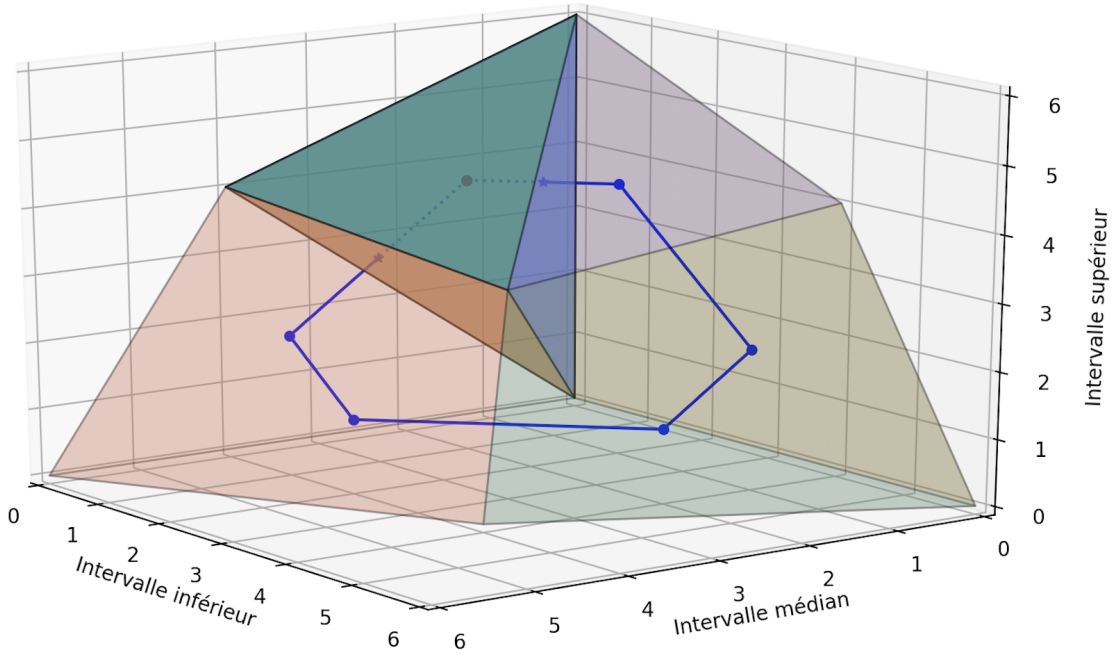


FIGURE 9.16 – Représentation, à l'intérieur de l'espace des formes normales, de l'identité associée à l'accord *do-réb-mib-sol*

Le chiffre 6 correspond au nombre de permutations des triplets. De manière générale, pour les accords à N notes, une identité « non dégénérée », c'est-à-dire ayant un nombre maximal de classe normales, est constituée de $|\mathbf{S}_{N-1}| = (N-1)!$ classes normales, où $\mathbf{S}_{N-1}$ désigne le groupe symétrique à $N-1$ éléments. Pour $N = 2$, on retrouve bien la coïncidence des classes

1. Auquel cas le cardinal de l'identité peut être inférieur à six, comme c'est le cas de la classe première associée à l'accord de 7^e diminuée, qui n'a qu'une classe normale.

normales et des identités.

La figure 9.17 montre les valeurs de concordance des classes normales de l'identité associée à *do-réb-mib-solb*, de forme identitaire [1,2,3,6]. La normalisation des couleurs est cette fois réalisée sur cette orbite, et non plus sur l'ensemble des formes premières comme c'était le cas sur la figure 9.13. Les couleurs se correspondent deux-à-deux pour former trois groupes de formes normales, qui sont en fait les trois classes premières de l'identité associée à *do-réb-mib-solb*.

Les classes normales d'une même classe première ne peuvent être distinguées que par des descripteurs non constants sur les classes premières. C'est le cas de la concordance d'ordre 3, la concordance totale, la tension ou l'harmonicité. La figure 9.18 représente la tension sur l'orbite identitaire de l'ensemble de hauteurs *do-réb-mib-sol*. Chaque classe normale de l'orbite a des valeurs différentes de tension.

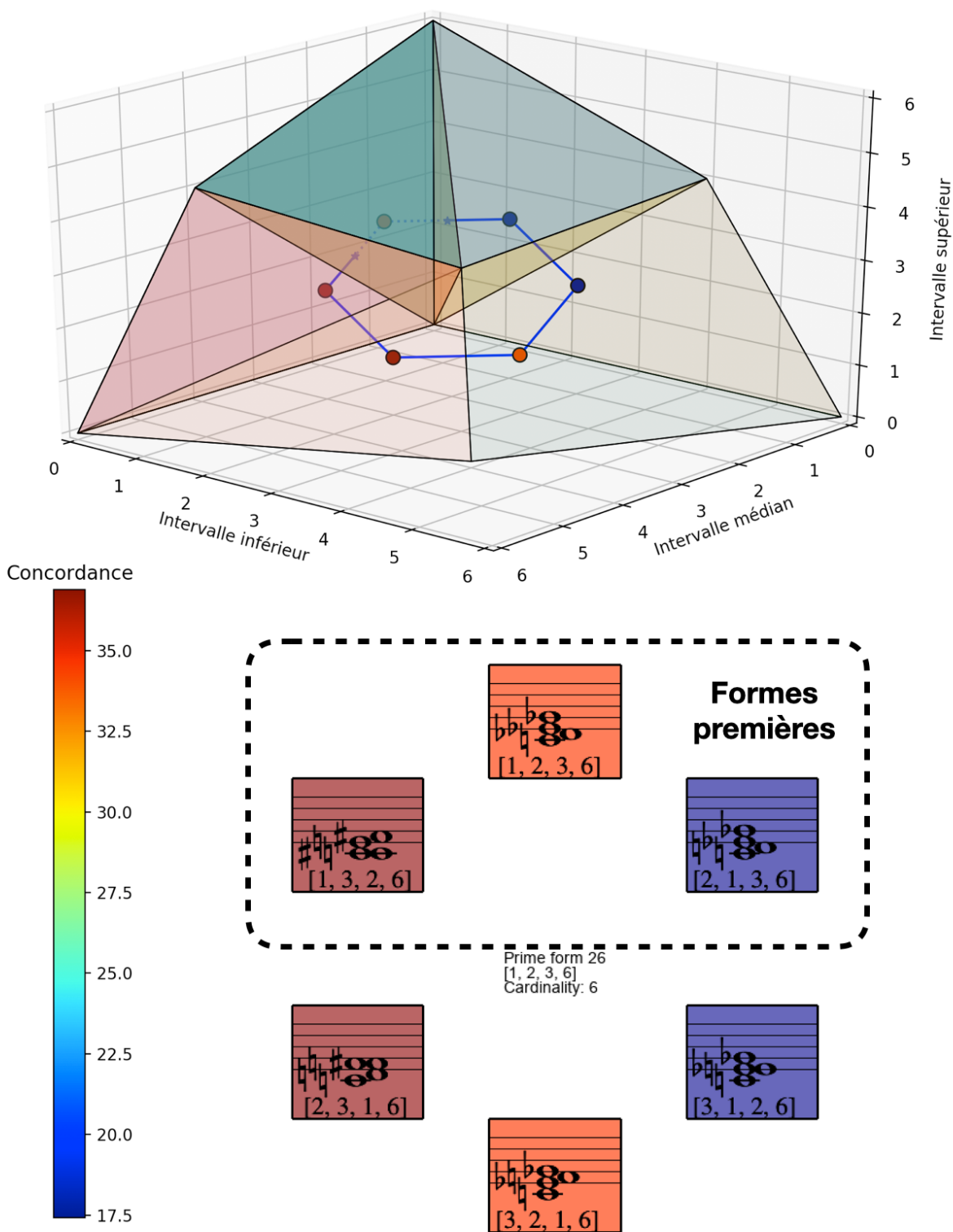


FIGURE 9.17 – Représentation de la concordance de classe sur l’orbite identitaire de l’accord *do-réb-mib-solb*. L’espace des formes identitaires est foncé, l’espace des formes premières d’un foncé intermédiaire, et l’espace des formes normales d’un foncé atténué. L’identité est composé de trois formes premières. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

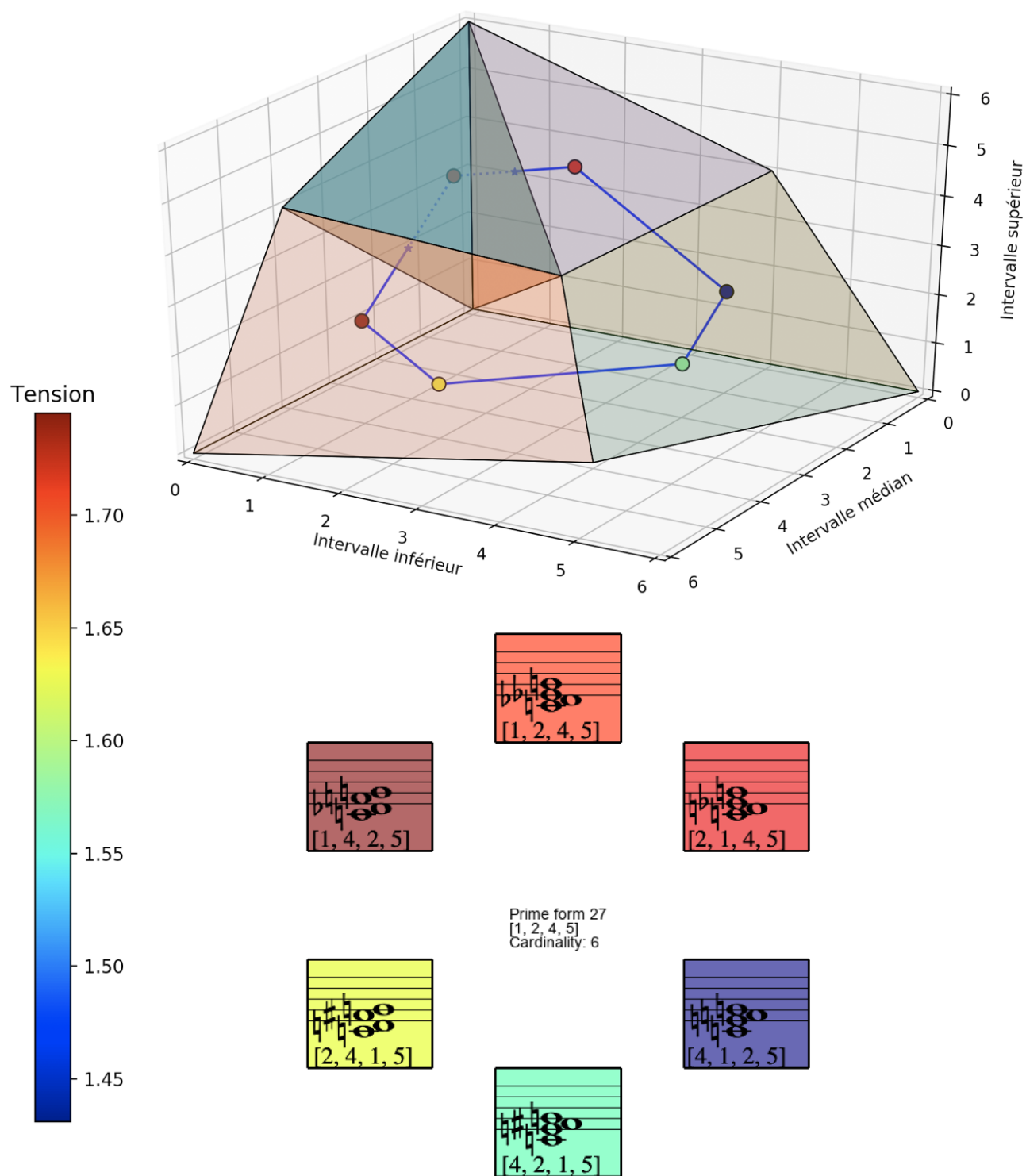


FIGURE 9.18 – Représentation de la tension de classe sur l'orbite identitaire de l'accord *do-réb-mib-sol*. L'espace des formes identitaires est foncé. Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Chaque classe normale, à son tour, est constituée de plusieurs accords. La figure 9.19 montre l'orbite de la classe normale associée à l'ensemble de hauteurs *do-réb-mib-sol*, dans l'espace des présentations réduites des tetrachordes. Une classe normale est composée de la forme normale et de ses renversements, qui sont au nombre de 3 si la classe est non dégénérée¹.

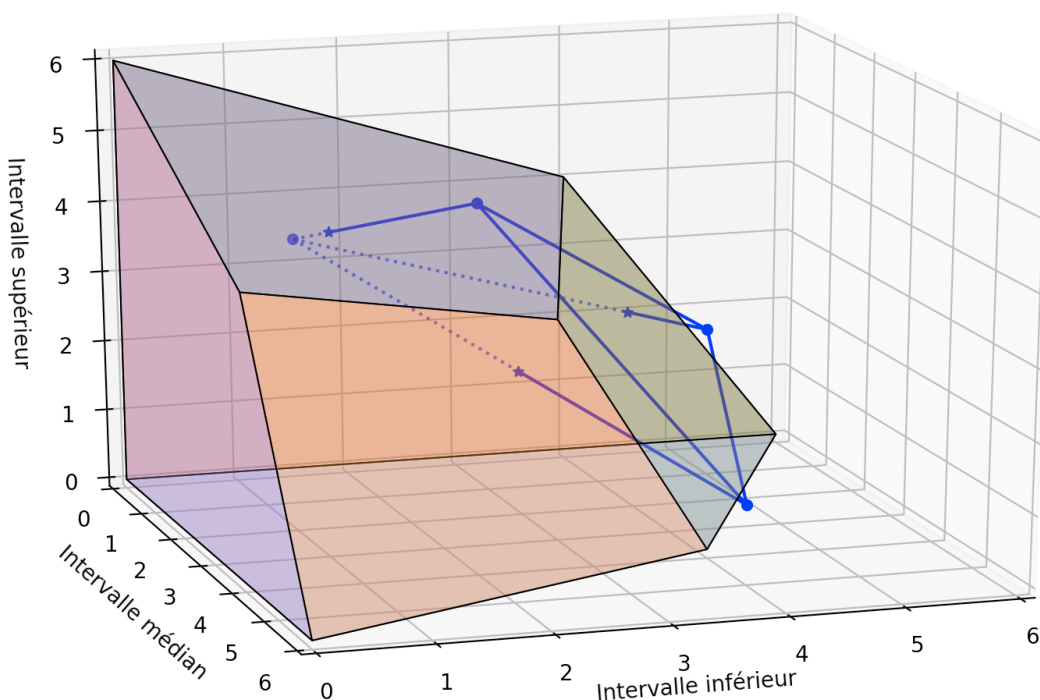


FIGURE 9.19 – Orbite de la classe normale associée à l'accord *do-réb-mib-sol*, dans l'espace des tetrachordes dont les intervalles entre notes successives ne dépassent pas l'octave.

En se restreignant aux spectres de Shepard, les valeurs des descripteurs sont constants sur tous les accords d'une classe normale. Là réside la raison de l'introduction de ce type de spectre. En revanche, avec des spectres réels, les valeurs des descripteurs varient sur les accords d'une classe normale, ce qui ne permet pas d'abstraire les descripteurs aux classes normales.

1. Une classe normale est dite non dégénérée lorsqu'il n'y a pas de répétition dans sa forme normale.

Continuum et discontinuum Les considérations précédentes amènent à distinguer les espaces continus des espaces discontinus.

Sans discrétisation préalable des fréquences, pour chaque entier $N \geq 2$, l'espace des identités, des classes premières et des classes normales des accords à N notes, confondus avec leurs espace de formes, est infini et continu. Chaque identité contient en revanche un nombre fini de classes premières, qui sont composées à leur tour d'un nombre fini de classes normales, lesquelles contiennent par contre un nombre infini d'ensemble de hauteurs, et chaque ensemble de hauteurs admet une infinité de réalisations.

Une subdivision égale des fréquences entraîne une majoration du nombre de notes des classes par le nombre de subdivisions dans l'octave. À cette limitation s'ajoute une discrétisation des classes normales et premières, et des identités, qui sont alors en nombre fini. La subdivision n'a par contre aucun effet sur le contenu des classes.

9.2.3 La « carotte » ou potentiel combinatoire des intervalles

Présentation Le potentiel combinatoire, aussi appelé « *carotte* ¹ » du fait de sa forme, est une structure générale synthétisant l'ensemble des combinaisons possibles des divisions d'une unité dans une subdivision donnée de cette unité. Elle a été introduite par le compositeur et musicologue Julio Estrada dans sa thèse ([Estrada, 1994](#)), pour rendre compte des combinaisons de hauteurs dans une échelle de hauteurs donnée, ou des combinaisons rythmiques au sein d'une unité temporelle, à partir d'une subdivision fixe de cette unité. Estrada, en représentant les combinaisons harmoniques et rythmiques dans le même espace, entend mettre en avant la continuité entre rythme, timbre et harmonie.

L'approche harmonique donne à voir l'ensemble de l'univers harmonique (au sens des identités et des classes normales) lié à une échelle de hauteurs. Elle nécessite d'abord de fixer un « intervalle de duplication de fréquence », qui rend équivalentes les notes séparées de cet intervalle. On parle alors de hauteur, et non plus de note ². Dans la suite, nous prenons l'octave comme intervalle de duplication de fréquence, mais il est tout à fait envisageable de prendre un autre intervalle, ce qui peut être pertinent physiquement et perceptivement dans le cas de spectres inharmoniques ³. Dans de tels cas, la notion de hauteur définie dans l'introduction nécessitera d'être adaptée, de même que la construction des spectres de Shepard, dont les partiels ne devront plus être répliqués à l'octave, mais à l'intervalle de duplication.

L'intervalle de duplication de fréquence fixé, il y a autant de « carottes » que de subdivisions égales de cet intervalle. Dans la suite, nous nous restreignons à la subdivision de l'octave en demi-tons, soit l'échelle chromatique. Il convient cependant de garder à l'esprit que la carotte constitue un outil particulièrement adapté à la musique microtonale, puisqu'elle permet de représenter de manière compacte les accords formés avec des quarts de tons ou toute autre subdivision régulière.

La carotte contient deux niveaux. Le premier niveau est constitué de l'ensemble des identités, regroupées en colonnes selon leur cardinalité. Chaque identité est représentée par sa forme identitaire. Au sein de chaque colonne,

1. Le nom vient de Julio Estrada, voir ([Estrada, 1994](#)).

2. Voir page 60.

3. Voir ([Sethares, 1999](#)), en particulier le chapitre 1 « *The octave is dead* » (« l'octave est morte »).

les identités sont rangées selon l'ordre alphabétique de la présentation réduite de leur forme identitaire, comme sur la figure 9.12.

La figure 9.20, tirée de (Estrada, 1994), montre le premier niveau de la carotte associée à l'échelle chromatique. La représentation des 77 identités a une forme globale de double cône caractéristique d'une « carotte », d'où son nom. À partir de la seule identité des accords à une note, le nombre d'identités dans chaque colonne augmente jusqu'à $N = 4$, puis diminue progressivement jusqu'à ne contenir qu'une identité pour les accords de 11 et 12 notes.

La forme en double cône du potentiel combinatoire est générale, et le nombre d'identités augmente avec la subdivision.

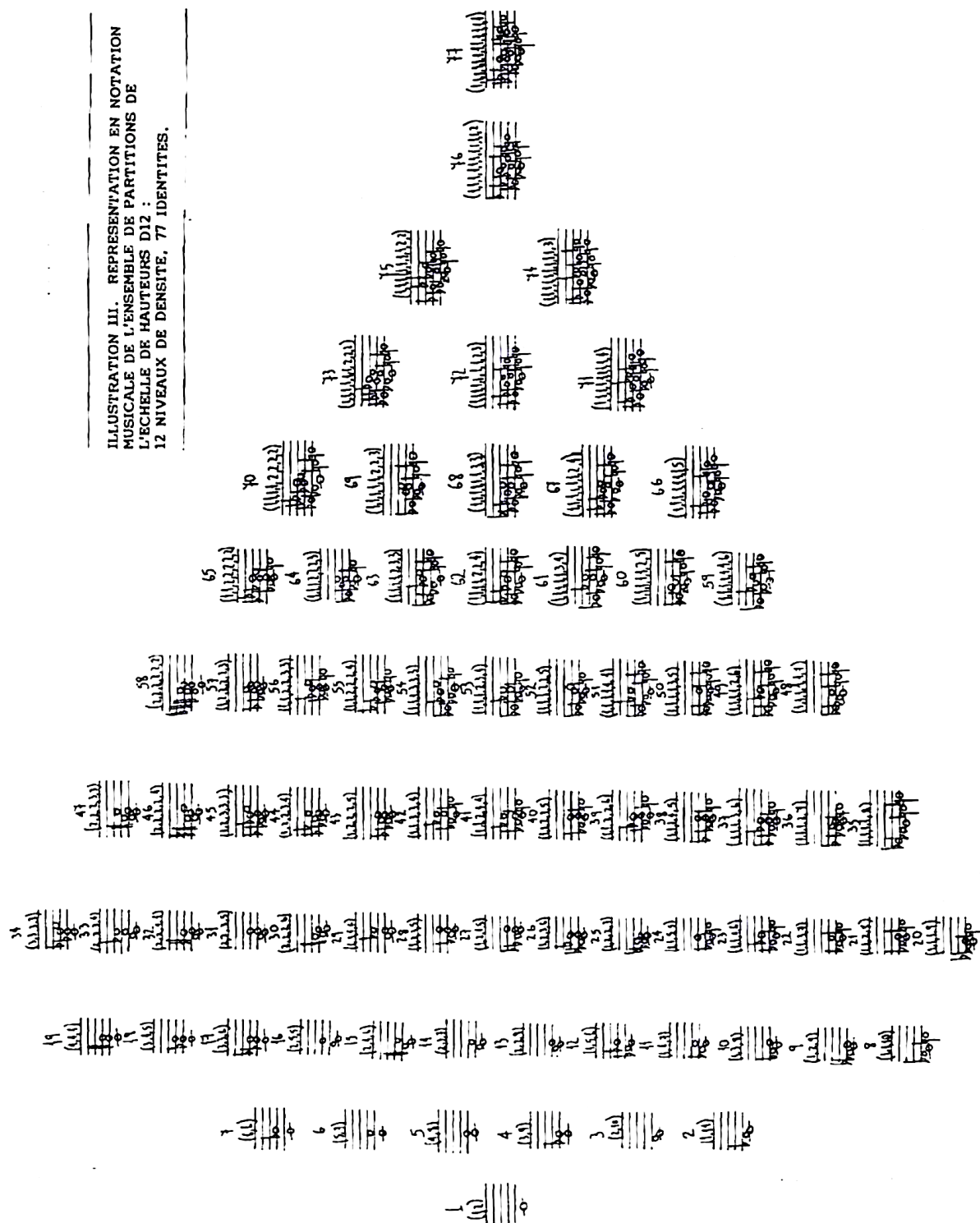


FIGURE 9.20 – Potentiel combinatoire de l'échelle chromatique. Tiré de (Estrada, 1994, p.176). Toutes les identités sont représentées, regroupées en colonnes selon leur cardinalité.

Le second niveau de la carotte correspond à celui des classes normales, représentées par les formes normales. Celles-ci sont regroupées par identités au sein des orbites identitaires. La figure 9.21 représente l'ensemble des formes normales de l'identité de forme identitaire *do-réb-mib-sol*, de cardinalité 6, déjà rencontrée sur la figure 9.18.

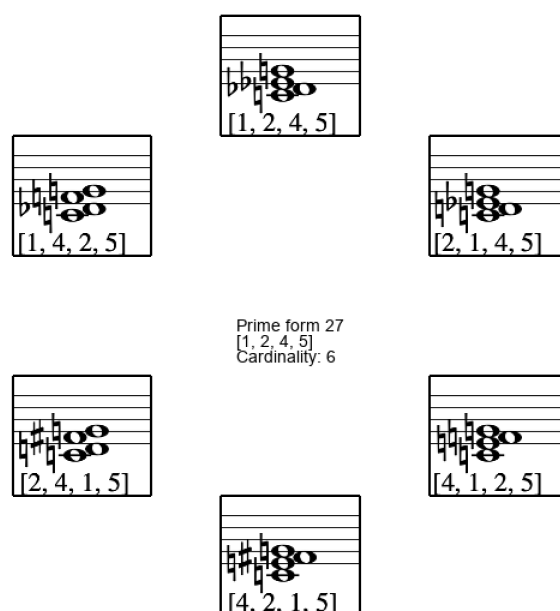


FIGURE 9.21 – Orbite identitaire associée à la forme identitaire *do-réb-mib-sol*

Les orbites identitaires du second niveau ont une structure qui généralise celle connue en théorie des graphes sous le nom de « *permutahedron* ». Alfonsin et Romero (Alfonsin et Romero, 2002) introduisent le terme de « *combinohedron* » et généralisent des résultats connus sur le *permutahedron* au *combinohedron*.

Usage en analyse musicale Estrada, en introduisant la carotte (Estrada, 1994), a développé une théorie dite « théorie *d1*, » qui met en place une distance entre les identités. Essentiellement deux raisons motivent cet effort :

- Cette théorie intègre le rythme et l’harmonie (en tant que combinaison d’intervalles) dans un même espace. Le rythme et l’harmonie sont vus comme deux aspects d’un même continuum de fréquences.
- Cette approche fournit des outils pour étudier la musique qui se veulent généraux et objectifs, sans présupposé esthétique et indépendants du système musical¹. Ils permettent en outre d’aborder la musique microtonale.

La théorie *d1* permet l’étude d’une œuvre dans ses aspects à la fois statiques et dynamiques.

Tout d’abord, l’approche statique consiste à projeter toutes les verticalités sur la carotte, ce qui amène à circonscrire l’espace harmonique des identités utilisées par le compositeur dans cette œuvre. En comparant des œuvres d’écritures ou d’époques différentes, l’approche diachronique permet d’identifier les nouveautés, propres à une œuvre, un compositeur, un langage, et de se rendre compte du degré d’exploration de l’univers harmonique. Dans (Estrada, 2011), Estrada analyse de cette manière, sur le premier niveau de la carotte, les *6 Kleine Klavierstücke* d’Arnold Schoenberg, montrant comment le compositeur explore, dans chaque pièce, l’espace harmonique atonal.

L’approche dynamique nécessite ce qu’Estrada appelle la « distance *d1* », qui est une distance à valeurs entières entre les identités. Celle-ci fait appel à un ensemble de transitions élémentaires entre identités de distance minimale. La distance entre deux identités correspond au nombre minimal de transitions élémentaires pour passer de l’une à l’autre. De là, il est possible de représenter les transitions dans une œuvre sous forme de graphe dont les sommets sont les identités et les flèches indiquent les transitions ainsi que la distance entre les classes. Cette approche permet à Estrada, dans l’étude des *6 Kleine Klavierstücke*, de mettre en évidence des pôles harmoniques.

Généralisation aux autres subdivisions Le premier niveau de la carotte fournit une représentation compacte de l’ensemble de l’« univers harmonique potentiel » associé à une subdivision de l’intervalle de duplication, au sens précis des identités. Plus la subdivision de l’intervalle de duplication est fine, plus l’univers harmonique est riche, ce qui se traduit par une augmentation

1. Avec la restriction toutefois à une musique de notes avec des échelles fréquentielles ou rythmiques qui peuvent être envisagées comme des subdivision d’un certain intervalle ou d’une certaine durée fixes.

du nombre d'identités.

La question du nombre d'identités dans une échelle donnée n'est pas immédiate. Considérons sans perte de généralité que l'intervalle de duplication est l'octave, et que l'échelle divise l'octave en n intervalles égaux. Les présentations réduites des formes identitaires fournissent une correspondance entre les identités et les manières d'écrire n comme somme d'entiers strictement positifs, à l'ordre des termes près. Le nombre d'identités se ramène au problème connu du nombre de partitions d'un entier.

Soit $n \geq 1$ le nombre de subdivisions de l'octave, égal à 12 si l'échelle considérée est l'échelle chromatique, et k un entier positif. Désignons par $\mathcal{P}(n, k)$ le nombre d'identités à k hauteurs, égal à 0 si $k = 0$ ou $k > n$. Si $1 \leq k \leq n$, $\mathcal{P}(n, k)$ correspond à la taille de la k^e colonne du premier niveau de la carotte.

Pour dénombrer les identités à k hauteurs, fixons une identité et raisonnons sur la première composante de la présentation réduite de la forme identitaire. Le raisonnement n'est ici pas musical, mais purement mathématique. Deux cas se présentent :

- Soit la première composante est égale à 1, auquel cas la présentation réduite est obtenue à partir d'une présentation réduite d'une identité à $k - 1$ hauteurs dans une subdivision de l'octave en $n - 1$ intervalles égaux, par adjonction à cette première composante.
- Soit toutes les composantes valent au-moins deux, auquel cas la présentation réduite est obtenue à partir de la présentation réduite d'une identité à k hauteurs dans une subdivision de l'octave en $n - 1$ intervalles égaux, par incrémentation de chaque composante.

Ce raisonnement fournit la relation de récurrence suivante :

$$\mathcal{P}(n, k) = \mathcal{P}(n - 1, k - 1) + \mathcal{P}(n - k, k), \quad (9.7)$$

avec les conditions initiales :

$$\begin{cases} \mathcal{P}(n, k) = 0 \text{ si } k > n \\ \mathcal{P}(n, 1) = \mathcal{P}(n, n) = 1 \end{cases} \quad (9.8)$$

Le nombre $\mathcal{P}(n)$ d'identités dans une subdivision de l'octave en n inter-

valles égaux se calcule alors par une sommation sur les colonnes :

$$\mathcal{P}(n) = \sum_{k=1}^n \mathcal{P}(n, k). \quad (9.9)$$

La relation de récurrence (9.7) assure au calcul du nombre d'identités une complexité quadratique.

Hardy et Ramanujan ([Hardy et Ramanujan, 1918](#)) ont démontré la croissance exponentielle de $\mathcal{P}$, donnant l'équivalent suivant :

$$\mathcal{P}(n) \sim \frac{1}{4n\sqrt{3}} \exp \left(\pi \sqrt{\frac{2n}{3}} \right) \quad (9.10)$$

La figure 9.22 dresse un tableau du nombre de partitions des entiers, qui correspond au nombre d'identités dans les subdivisions égales de l'octave en n intervalles. L'échelle chromatique, nous l'avons vu, donne lieu à 77 identités. Lorsque la subdivision devient plus fine, ce nombre croît rapidement : 1575 identités pour l'échelle en quart-de-ton, contre près de 18000 pour l'échelle en sixième-de-ton.

Cette croissance exponentielle rend difficile la représentation au format papier des carottes associées à de telles échelles. Pour cette raison nous nous restreignons, dans cette partie, à l'échelle chromatique, mais la méthode de coloration de la carotte présentée dans la suite est généralisable à toutes les échelles égales.

1 ... 1	39 ... 31185	77 ... 10619863	115 ... 1064144451
2 ... 2	40 ... 37338	78 ... 12132164	116 ... 1188908248
3 ... 3	41 ... 44583	79 ... 13848650	117 ... 1327710076
4 ... 5	42 ... 53174	80 ... 15796476	118 ... 1482074143
5 ... 7	43 ... 63261	81 ... 18004327	119 ... 1653668665
6 ... 11	44 ... 75175	82 ... 20506255	120 ... 1844349560
7 ... 15	45 ... 89134	83 ... 23338469	121 ... 2056148051
8 ... 22	46 ... 105558	84 ... 26543660	122 ... 2291320912
9 ... 30	47 ... 124754	85 ... 30167357	123 ... 2552338241
10 ... 42	48 ... 147273	86 ... 34262962	124 ... 2841940500
11 ... 56	49 ... 173525	87 ... 38887673	125 ... 3163127352
12 ... 77	50 ... 204226	88 ... 44108109	126 ... 3519222692
13 ... 101	51 ... 239943	89 ... 49995925	127 ... 3913864295
14 ... 135	52 ... 281589	90 ... 56634173	128 ... 4351078600
15 ... 176	53 ... 329931	91 ... 64112359	129 ... 4835271870
16 ... 231	54 ... 386155	92 ... 72533807	130 ... 5371315400
17 ... 297	55 ... 451276	93 ... 82010177	131 ... 5964539504
18 ... 385	56 ... 526823	94 ... 92669720	132 ... 6620830889
19 ... 490	57 ... 614154	95 ... 104651419	133 ... 7346629512
20 ... 627	58 ... 715220	96 ... 118114304	134 ... 8149040695
21 ... 792	59 ... 831820	97 ... 133230930	135 ... 9035836076
22 ... 1002	60 ... 966467	98 ... 150198136	136 ... 10015581680
23 ... 1255	61 ... 1121505	99 ... 169229875	137 ... 11097645016
24 ... 1575	62 ... 1300156	100 ... 190569292	138 ... 12292341831
25 ... 1958	63 ... 1505499	101 ... 214481126	139 ... 13610949895
26 ... 2436	64 ... 1741630	102 ... 241265379	140 ... 15065878135
27 ... 3010	65 ... 2012558	103 ... 271248950	141 ... 16670689208
28 ... 3718	66 ... 2323520	104 ... 304801365	142 ... 18440293320
29 ... 4565	67 ... 2679689	105 ... 342325709	143 ... 20390962757
30 ... 5604	68 ... 3087735	106 ... 384276336	144 ... 22540654445
31 ... 6842	69 ... 3554345	107 ... 431149389	145 ... 24908858009
32 ... 8349	70 ... 4087968	108 ... 483502844	146 ... 27517052599
33 ... 10143	71 ... 4697205	109 ... 541946240	147 ... 30388671978
34 ... 12310	72 ... 5392783	110 ... 607163746	148 ... 33549419497
35 ... 14833	73 ... 6185689	111 ... 679903203	149 ... 37027355200
36 ... 17977	74 ... 7089500	112 ... 761002156	150 ... 40853235313
37 ... 21637	75 ... 8118264	113 ... 851376628	151 ... 45060624582
38 ... 26015	76 ... 9289091	114 ... 952050665	152 ... 49686288421

FIGURE 9.22 – Dénombrement des identités pour différentes subdivision, tiré de (Hardy et Ramanujan, 1918, p.114). Nous avons souligné en rouge les subdivisions correspondant aux échelles en demi-tons, en quart-de-tons et en sixième-de-tons.

9.2.4 Coloration de la « carotte » par les descripteurs harmoniques

Les analyses musicales citées précédemment s'appuyant sur le potentiel intervallique sont fondées sur des considérations purement combinatoires. Dans l'idée d'ajouter des considérations d'ordre acoustique, deux approches sont a priori possibles qui sont une application directe des descripteurs harmoniques : l'approche statique et l'approche dynamique. Or, il se trouve qu'une transition entre identités ou classes normales ne peut être décrite par les descripteurs dynamiques non triviaux¹. En effet, quand bien même les spectres considérés seraient des spectres de Shepard, les interactions entre spectres de deux accords successifs dépendent des hauteurs relatives des deux accords, hauteurs qui ne sont pas préservées lorsque l'on passe aux identités ou aux classes normales. Ainsi, les descripteurs dynamiques ne sont pas directement exploitables dans l'étude des transitions dans la carotte.

En revanche, les descripteurs harmoniques sont tout à fait exploitables dans une étude statique. Les figures 9.12, 9.13 et 9.18 ont déjà montré des colorations de parties de la carotte correspondant à une subdivision en demi-tons (respectivement 4^e colonne du premier niveau et orbite identitaire du second niveau) en fonction des valeurs d'un descripteur. De manière générale, il est possible de colorier toute ou partie de la carotte par les descripteurs harmoniques statiques.

Implémentation Nous avons développé en Python un logiciel proposant une coloration de la carotte associée à une subdivision en demi-tons par les descripteurs harmoniques statiques. Son principe est expliqué ci-dessous.

L'affichage par défaut est l'affichage du premier niveau de la carotte, mais on peut aussi se restreindre à une colonne ou un intervalle de colonnes. Un clic sur une identité a pour effet de faire apparaître une seconde fenêtre avec son orbite identitaire, c'est-à-dire l'ensemble des classes normales qu'elle contient, en même temps qu'il affiche sur le terminal les valeurs des descripteurs courants de sa forme identitaire. Les descripteurs courants sont les descripteurs impliqués dans la coloration, que l'on peut choisir dans l'interface. De même, un clic sur une classe normale du second niveau fait apparaître sur le terminal les valeurs des descripteurs courants de cette classe normale.

1. Voir partie 9.1.3

Une interface, présentée figure 9.23, permet de régler les paramètres de spectre ainsi que les paramètres de coloration des classes.



FIGURE 9.23 – Interface du logiciel de coloration de la carotte par les descripteurs harmoniques statiques

Les valeurs des descripteurs, normalisées de 0 à 100, sont pré-rentrées sur plusieurs spectres de Shepard, qui résultent d’une combinaison de différentes valeurs du nombre de partiels, de leur décroissance et du facteur de largeur des partiels σ .

L’assignation des couleurs est la même pour les deux niveaux. Deux modes de coloration sont proposés. Le premier, le mode « arc-en-ciel », consiste à assigner la couleur des classes aux valeurs d’un descripteur selon une échelle allant du bleu foncé pour les valeurs les plus basses au rouge foncé pour les

valeurs les plus hautes¹. Il n'y a alors qu'un seul descripteur courant. Le second mode de coloration consiste à assigner un descripteur à chacune des 3 composantes du codage RVB (rouge - vert - bleu).

Pour la rugosité et la concordance, c'est la valeur du descripteur qui détermine directement la coloration. Par contre, pour l'harmonicité, la tension, la concordance d'ordre 3 et la concordance totale, une assignation directe des valeurs aux échelles de couleur donne lieu à une répartition des couleurs concentrée vers les bleus (dans le mode « arc-en-ciel ») ou vers les faibles valeurs du codage RVB. Pour avoir une répartition des couleurs qui permette de discriminer davantage les formes identitaires, on fait agir au préalable un logarithme sur les valeurs de ces descripteurs, afin d'observer plus de variations sur les faibles valeurs des descripteurs.

Contrairement aux identités, les classes normales sont trop nombreuses pour pouvoir être visualisées en une fois², c'est pourquoi on a fait le choix de restreindre leur visualisation à une seule orbite identitaire à la fois. Deux choix sont possibles pour la coloration d'une orbite identitaire : on peut normaliser les couleurs sur l'ensemble des 351 classes normales (mode « valeurs absolues ») — c'est l'affichage par défaut — ou bien sur les seules classes normales de l'orbite (mode « valeurs relatives »). Le premier choix apporte une cohérence de couleurs entre les niveaux des identités et celui des classes normales. Il permet en outre de comparer les classes normales appartenant à différentes identités, mais avec le risque d'aboutir à des orbites monochromes quand les valeurs des descripteurs sur les classes normales d'une même identité sont trop proches. Le second choix discrimine les valeurs des descripteurs au sein d'une identité, mais cette discrimination peut être artificielle si les valeurs sont très proches.

Dispersion des couleurs sur les orbites identitaires Les identités sont coloriées selon les valeurs des descripteurs statiques sur les formes identitaires. Il faut garder à l'esprit que cette couleur ne décrit pas l'identité dans son ensemble, mais seulement une de ses formes normales.

On peut cependant considérer en première approximation certains descripteurs comme des descripteurs sur les identités, ce qui permet d'interpréter les couleurs qui en résultent comme décrivant les identités, et non plus uni-

1. Comme sur les figures 9.12 et 9.13 déjà rencontrées.

2. La subdivision en demi-tons donne lieu à 351 classes normales, contre 77 identités.

quement les formes identitaires. L'erreur de cette approximation est donnée par la dispersion, qui mesure les variations relatives de couleurs sur les orbites identitaires. La dispersion de couleur sur une identité vis-à-vis d'un descripteur est définie comme le rapport de la variation de couleur sur l'orbite identitaire par la variation totale de couleur sur les classes normales de toutes les orbites, lorsque la couleur est assignée à ce descripteur (mode arc-en-ciel).

Prenons l'exemple de la rugosité. La figure 9.26 montre la rugosité des formes identitaires sur le premier niveau de la carotte, avec un spectre constitué de 11 partiels, et la figure 9.27 la rugosité sur l'orbite identitaire de l'identité 40 à 5 hauteurs, de forme identitaire [1,1,2,3,5], composée de 12 classes normales. Les couleurs, bien que non constantes, sont relativement homogènes au sein de l'orbite identitaire. La dispersion de couleur sur l'identité 40 vis-à-vis de la rugosité est égale à 8.4%. Autrement dit, on fait une erreur de 8.4% en considérant la rugosité constante sur cette identité. L'erreur sur l'ensemble de la carotte est indiqué par le maximum de dispersion vis-à-vis de la rugosité sur toutes les orbites identitaires, ici égal à 15.5%.

Le tableau 9.2 donne les dispersions de couleurs maximales et moyennes sur toutes les identités vis-à-vis des descripteurs. L'harmonicité et la tension ont une dispersion maximale inférieure à 10%, ce qui permet de considérer avec une erreur de 10% que leur variation au sein de toute orbite identitaire est négligeable par rapport à leur variation sur l'ensemble du potentiel combinatoire.

	Dispersion maximale	Dispersion moyenne
Rugosité	15.5 %	6.8 %
Concordance	37.7 %	14.6 %
Harmonicité	7.6 %	2.6 %
Tension	2.9 %	0.7 %
Concordance ordre 3	38 %	5.9 %
Concordance totale	32.7 %	0.6 %

TABLE 9.2 – Table de la dispersion des couleurs sur les orbites identitaires, avec les paramètres de spectre suivants : $K = 11$, $\sigma = 0.005$ et $decr = 0.5$. La dispersion de couleur vis-à-vis d’un descripteur est définie comme le rapport de la variation de la couleur sur l’orbite identitaire par la variation totale sur les classes normales de toutes les orbites, lorsque la couleur est assignée à ce descripteur (mode arc-en-ciel).

Coloration du premier niveau

Rugosité La figure 9.26 montre la rugosité des formes identitaires sur le premier niveau de la carotte, avec un spectre constitué de 11 partiels. Les formes bleues ont une faible rugosité, les rouges une forte rugosité.

On observe une augmentation globale de la rugosité avec la cardinalité, et cela malgré la normalisation par le nombre de notes résultant de la condition d’égalité des descripteurs sur les unissons¹. De plus, au sein d’une colonne, la rugosité a tendance à diminuer avec le numéro d’identité, ce qui généralise à toutes les cardinalités le résultat déjà observé pour $N = 4$ ².

Concordance La figure 9.28 montre la concordance des formes identitaires sur le premier niveau de la carotte. Les formes bleues correspondent aux classes de faible concordance, les rouges à celles de concordance forte, avec les mêmes paramètres spectraux que sur la figure précédente.

1. Voir partie 7.3.

2. Voir figure 9.12.

On observe là aussi une augmentation globale de la concordance avec le nombre de notes, quoique moins homogène que dans le cas de la rugosité. Les identités de concordance maximale pour chaque cardinalité sont facilement repérables : la classe de la quarte, puis l'identité de l'accord *do-ré-sol* pour les triades, *do-ré-mi-sol* pour les tetrachordes, *do-ré-mi-sol-si* pour les accords à 5 notes.

Harmonicité La figure 9.29 montre l'harmonicité des formes identitaires sur le premier niveau de la carotte. En bleu, les formes de faible harmonicité, en rouge, celles de forte harmonicité.

La classe à une hauteur ayant un spectre harmonique, son harmonicité est maximale. Globalement, l'harmonicité diminue quand le nombre de hauteurs augmente. la figure 9.24, qui restreint l'affichage aux colonnes 3 à 5, permet de mieux apprécier les variations d'harmonicité sur les cardinalités intermédiaires.

Concernant les triades, la forme *do-ré-sol* a une harmonicité plus grande que la classe normale majeure *do-mi-sol*. Cela vient du tempérament égal. En effet, la figure 9.9 montre qu'avec des intervalles justes la classe normale majeure a une plus grande harmonicité que la classe normale de *do-ré-sol*. Une manière de s'approcher du tempérament juste sur l'accord mineur serait de considérer une subdivision plus fine de l'octave, par exemple une subdivision en 12^e de tons.

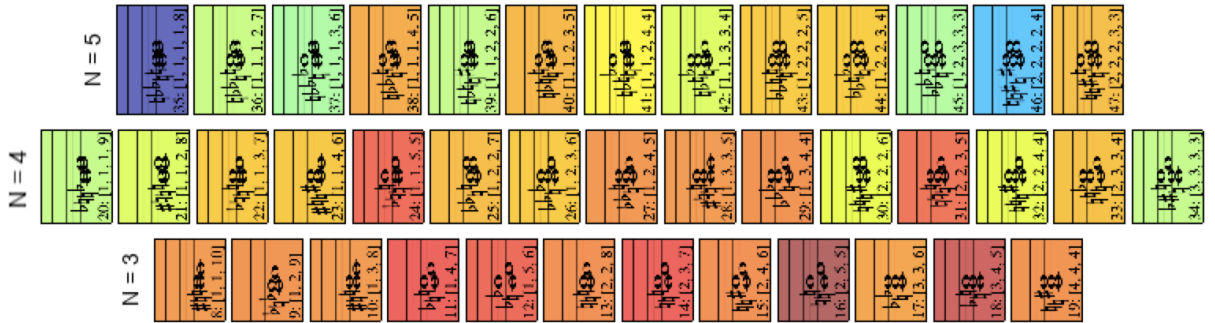


FIGURE 9.24 – Harmonicity des formes identitaires entre les 3^e et 5^e colonnes de la carotte. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Tension La figure 9.30 montre la tension des formes identitaires sur le premier niveau de la carotte. En bleu, les formes de faible tension, en rouge, celles de forte tension.

La tension augmente avec le nombre de hauteurs, et semble ne dépendre que ce paramètre, puisque la variation de couleur au sein de chaque colonne est négligeable par rapport à la variation entre les colonnes.

Pour se rendre compte de la variation de tension au sein d’une colonne, il faut réduire l’affichage à cette colonne, comme c’est le cas sur la figure 9.25, qui montre la tension des formes identitaires à trois hauteurs. La forme de tension maximale est *do-mi-sol*_#, de présentation réduite [4,4,4], dont l’identité ne contient qu’une classe première et qu’une classe normale. Les hauteurs divisent l’octave en trois intervalles égaux, d’où une tension maximale.

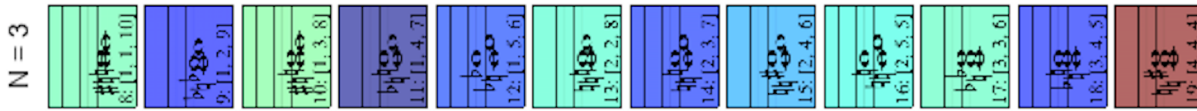


FIGURE 9.25 – Tension des formes identitaires sur la 3^e colonne de la carotte. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

Concordance Totale La coloration de la carotte par la concordance totale et la concordance d’ordre 3 dépend fortement des paramètres de spectre, plus qu’avec les autres descripteurs. De manière générale, la concordance totale est nulle sur une grosse partie de la carotte, mis à part quelques accords qui ressortent pour les faibles cardinalités.

La figure 9.31 montre la concordance totale des formes identitaires sur le premier niveau de la carotte, avec 11 partiels et trois valeurs successives du facteur de largeur des partiels : $\sigma = 0.001, 0.005$ et 0.01 . En bleu, les formes de faible concordance totale, en rouge, celles de forte concordance totale. Avec $\sigma = 0.001$, seule la forme identitaire *do-ré-sol* se distingue des autres. En augmentant les valeurs de σ , de plus en plus de formes sont mises en valeur : *do-mib-sol* et *do-ré-mi* pour $\sigma = 0.005$ et des formes de cardinalité 4 et 5 pour $\sigma = 0.01$.

Quand σ diminue, les pics de concordance totale dans les espace continus de présentations réduites se resserrent de plus en plus autour des maximums, dont les intervalles constitutifs sont des intervalles justes, c’est à dire pré-

sents dans la suite des harmoniques. Ainsi, pour les petites valeurs de σ , la concordance totale ne tolère pas l'écart du tempérament égal au tempérament juste.

Sur la forme *do-mib-sol*, la concordance totale est due à l'interaction entre les harmonique n° 3, 5 et 1 des trois hauteurs respectives *do*, *mib* et *sol*¹; et l'harmonique 5 a 14 cents de décalage par rapport au tempérament égal. Aussi $\sigma = 0.001$ correspond-t-il à une largeur des partiels de 1.2 cents, ce qui est trop peu pour mettre en avant la forme *do-mib-sol*. Par contre, avec $\sigma = 0.005$, soit une largeur des partiels de 6 cents, la forme *do-mib-sol* est détectée. Sur la forme *do-ré-sol*, la concordance totale est due à l'interaction entre les harmonique n° 9, 1 et 3 des trois hauteurs respectives *do*, *ré* et *sol*; et l'harmonique 3 a 3 cents de décalage par rapport au tempérament égal, contre 4 cents pour l'harmonique 9. La proximité sur cette forme du tempérament égal avec les intervalles justes explique qu'elle ait une plus grande concordance totale que la forme *do-mib-sol* pour $\sigma = 0.001$, à condition que le spectre contienne au moins 9 harmoniques, ce qui est le cas ici.

Coloration mixte La figure 9.32 donne un exemple de coloration mixte du premier niveau de la carotte. L'assignation des trois composantes rouge-vert-bleu du codage RVB à des descripteurs introduit une « couleur harmonique » propre à chaque forme normale. Cette couleur indique la prévalence de certains descripteurs par rapport à d'autres. Chaque descripteur voit ses valeurs normalisés de 0 à 100, linéairement pour la concordance et la rugosité, logarithmiquement pour les autres², et ces valeurs normalisées font office de coefficients dans le codage RVB. La « couleur harmonique » dépend donc fortement du type de normalisation sur chaque descripteur.

Sur la figure 9.32, la rugosité est assignée à la composante rouge, l'harmonicité au vert et la concordance au bleu. La dominante verte (identité 1) indique la prépondérance de l'harmonicité sur les deux autres descripteurs, la dominante cyan (identité 16) une prédominance de l'harmonicité et de la concordance sur la rugosité, et la composante magenta (identité 77) une prédominance de la rugosité et de la concordance sur l'harmonicité. Les tons gris indiquent que les valeurs des trois descripteurs sont proches l'une de l'autre.

1. Voir figure 8.26.

2. Voir page 369.

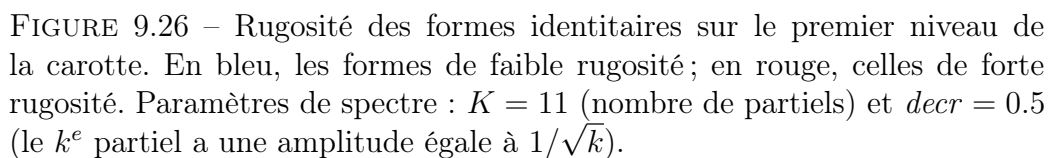


FIGURE 9.26 – Rugosité des formes identitaires sur le premier niveau de la carotte. En bleu, les formes de faible rugosité ; en rouge, celles de forte rugosité. Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

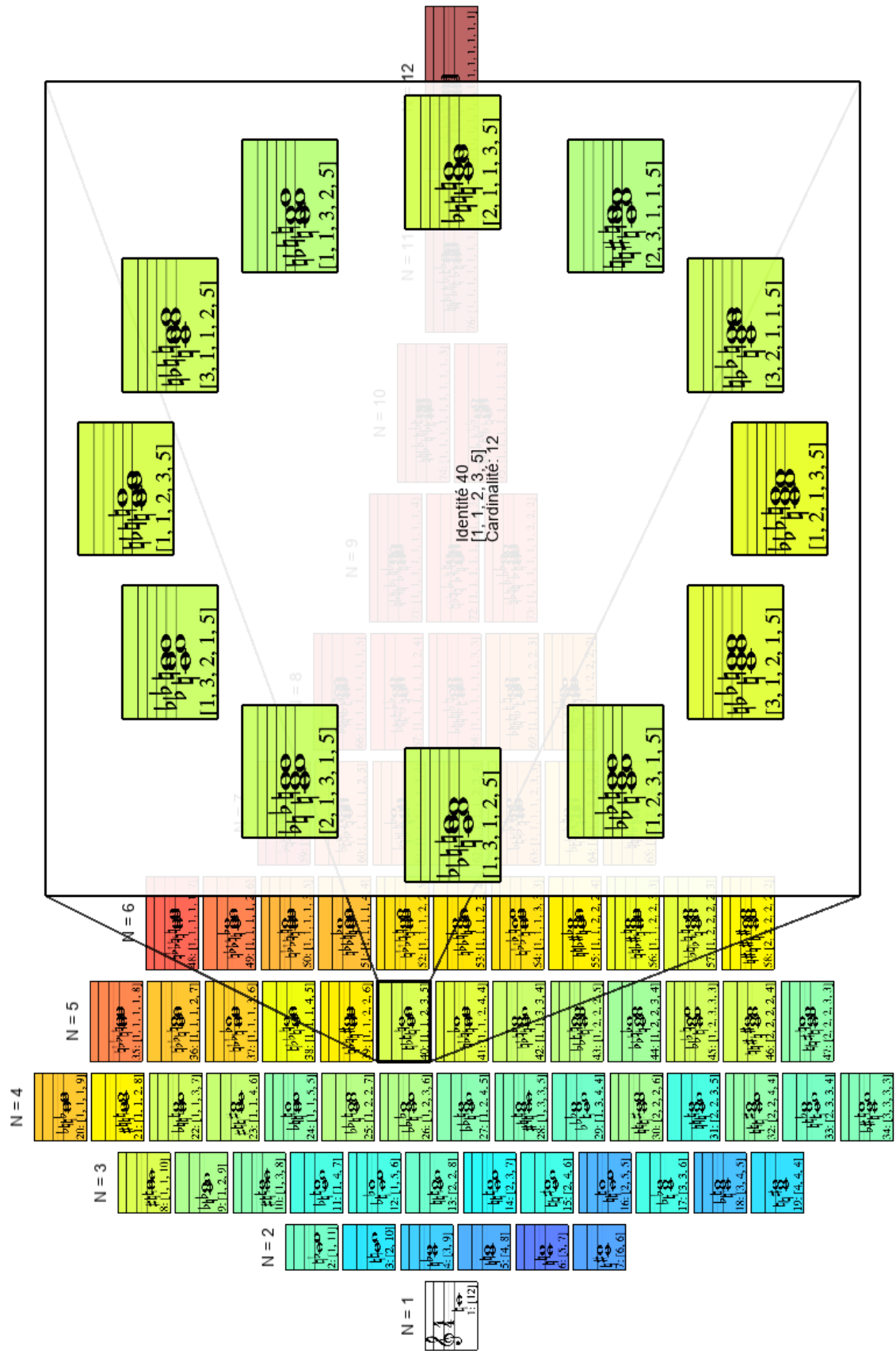


FIGURE 9.27 – Rugosité des formes identitaires sur le premier niveau de la carotte. En bleu, les formes de faible rugosité ; en rouge, celles de forte rugosité. Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

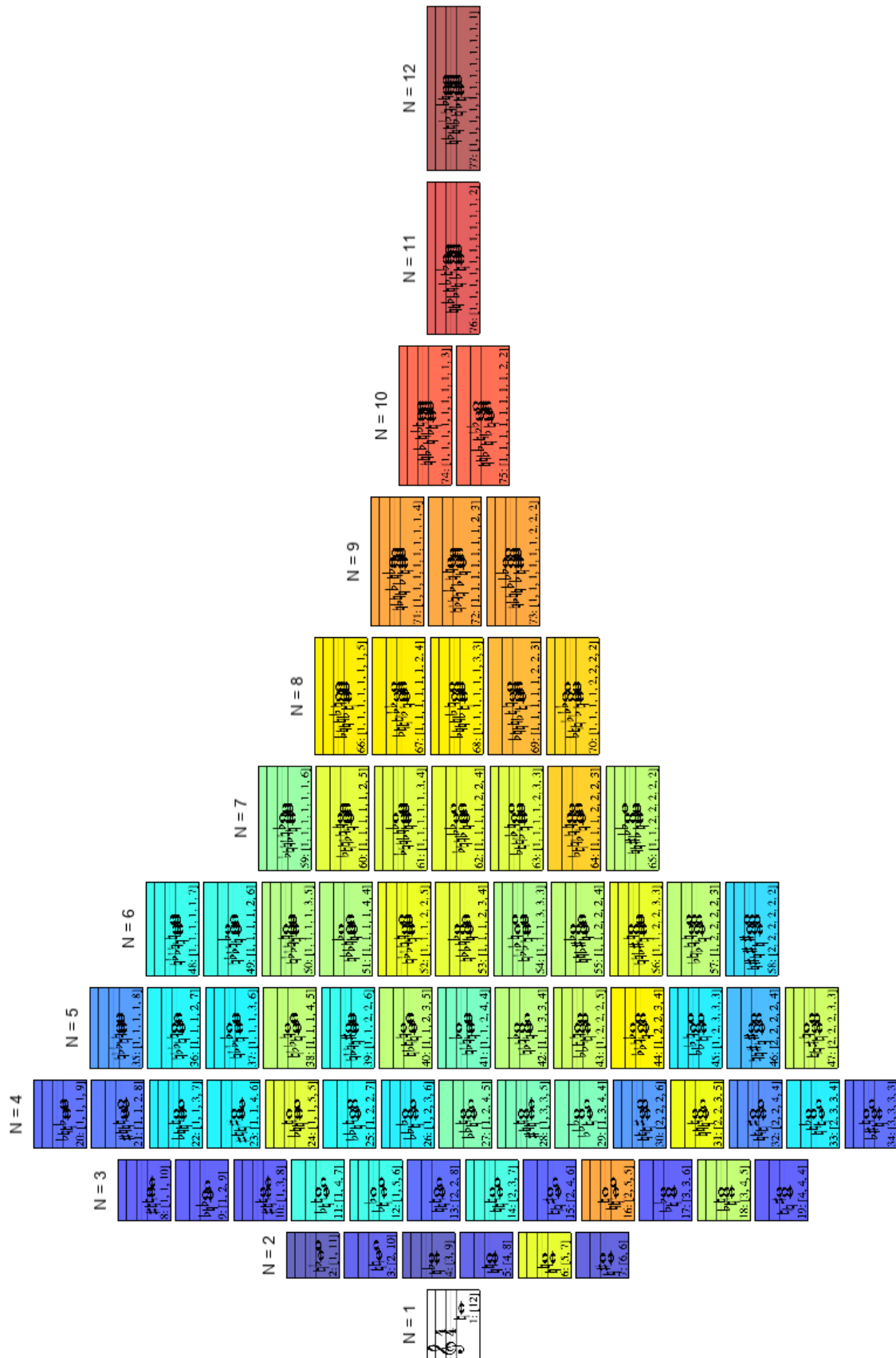


FIGURE 9.28 – Concordance des formes identitaires sur le premier niveau de la carotte En bleu, les formes de faible concordance ; en rouge, celles de forte concordance. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

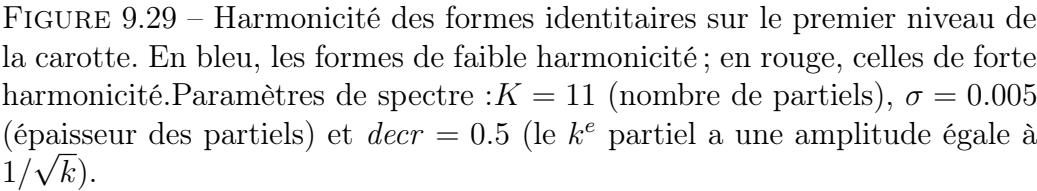


FIGURE 9.29 – Harmonicité des formes identitaires sur le premier niveau de la carotte. En bleu, les formes de faible harmonicité ; en rouge, celles de forte harmonicité. Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

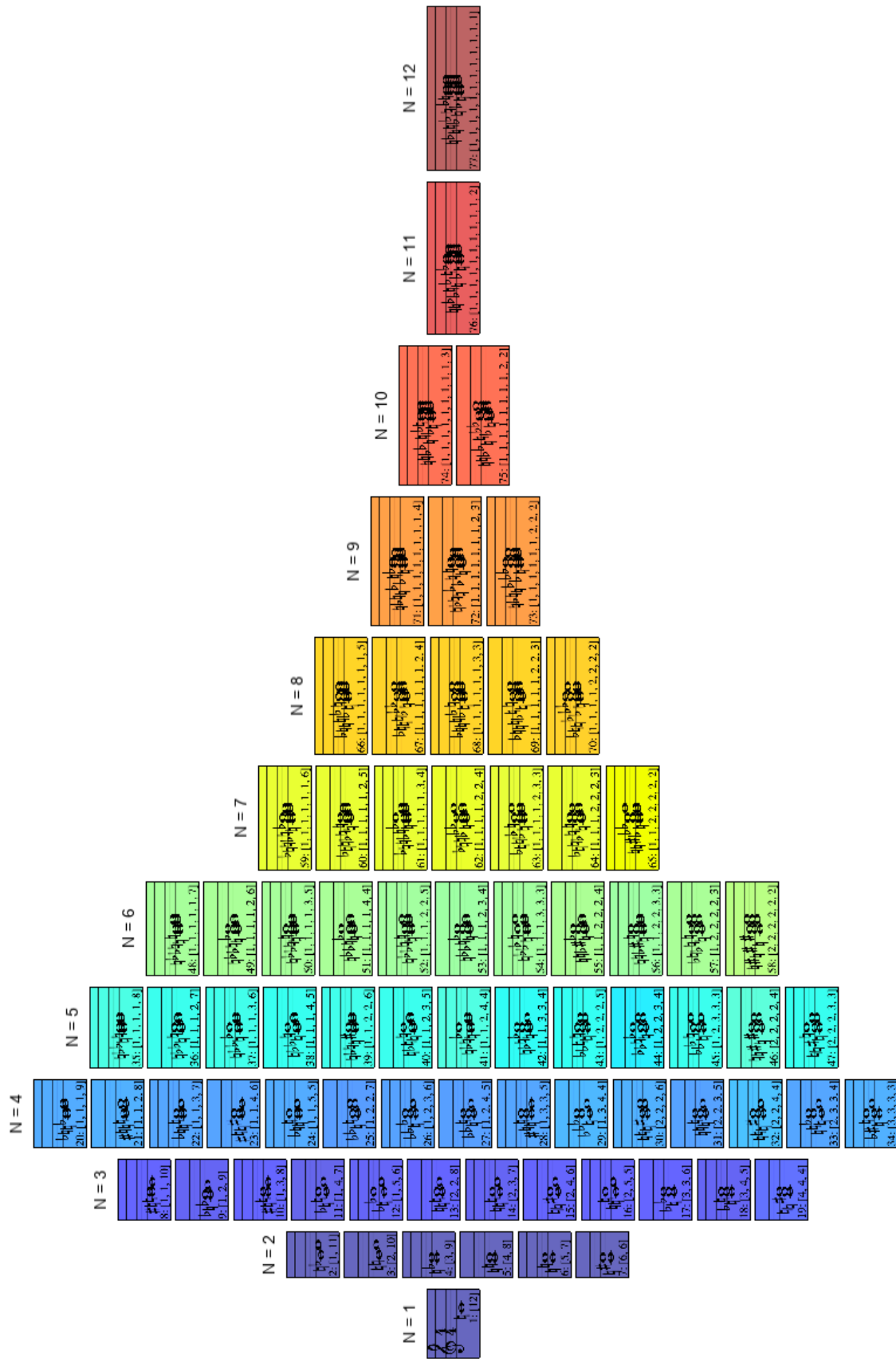


FIGURE 9.30 – Tension des formes identitaires sur le premier niveau de la carotte. En bleu, les formes de faible tension ; en rouge, celles de forte tension. Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

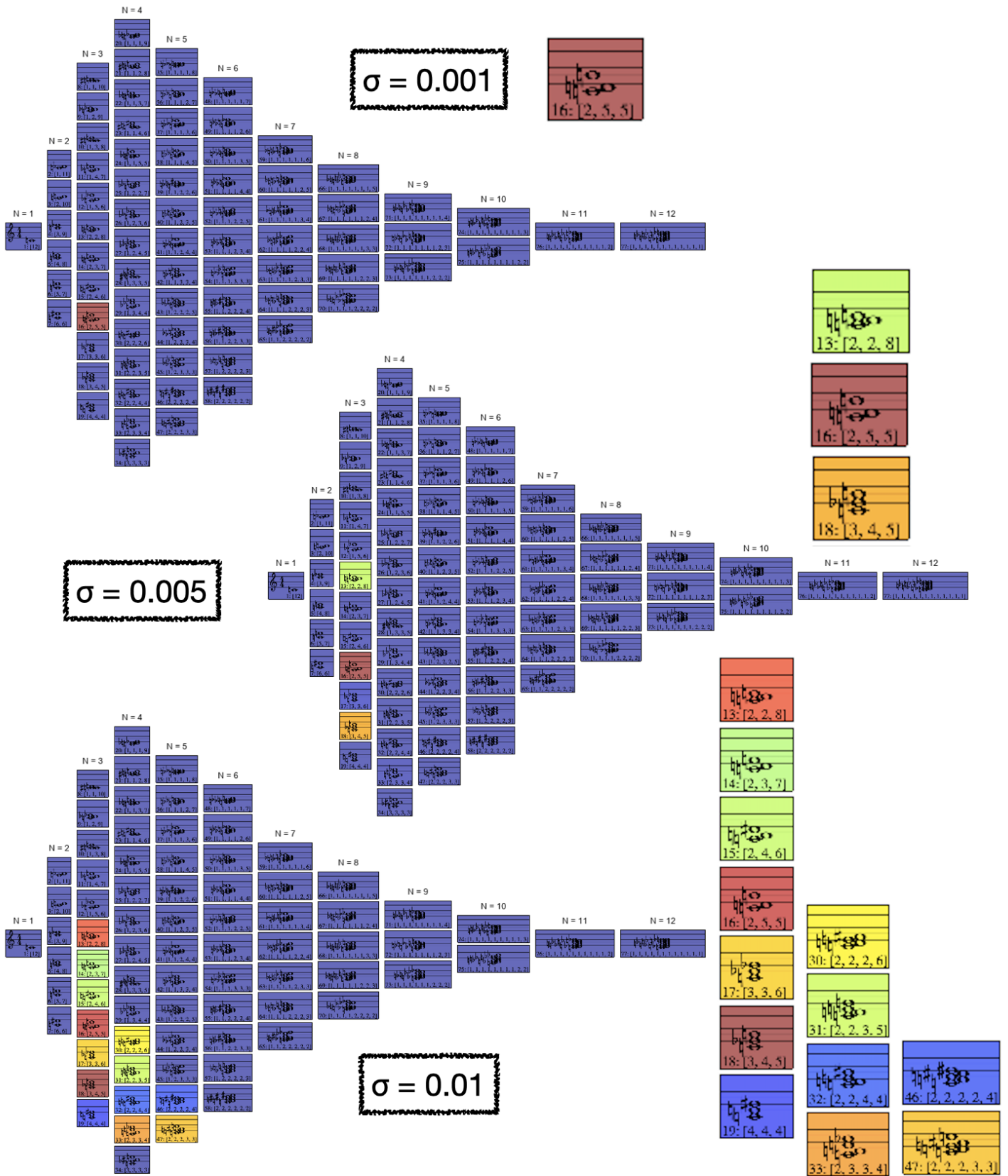


FIGURE 9.31 – Concordance totale des formes identitaires sur le premier niveau de la carotte, avec $\sigma = 0.001, 0.005$ et 0.01 (épaisseur des partiels). En bleu, les formes de faible concordance totale; en rouge, celles de forte concordance totale. Paramètres de spectre : $K = 11$ (nombre de partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

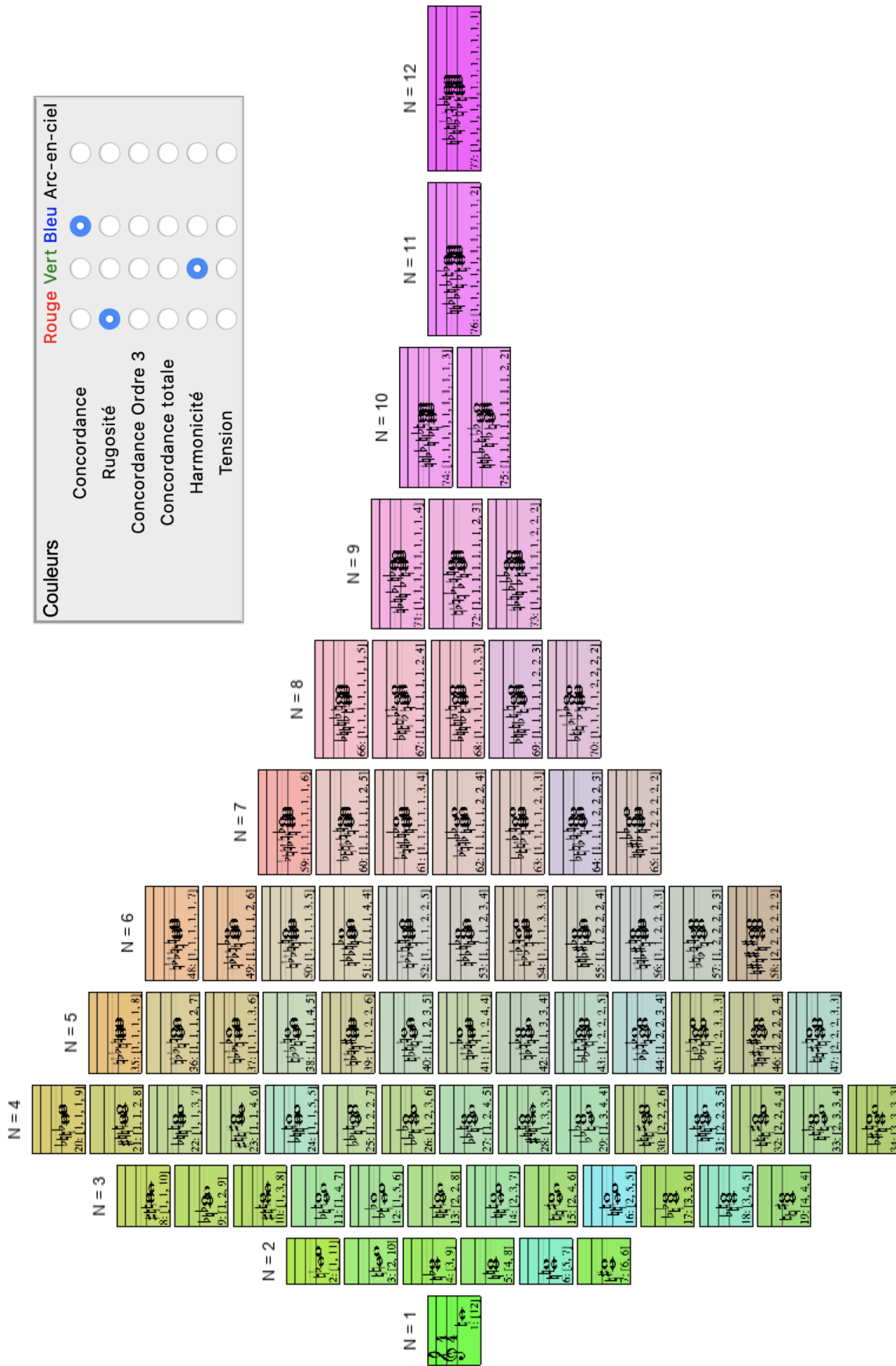


FIGURE 9.32 – Coloration mixte du premier niveau de la carotte par la rugosité (rouge), l’harmonicité (vert) et la concordance (bleu). Paramètres de spectre : $K = 11$ (nombre de partiels), $\sigma = 0.005$ (épaisseur des partiels) et $decr = 0.5$ (le k^e partiel a une amplitude égale à $1/\sqrt{k}$).

9.2.5 Conclusions et perspectives

Le monde harmonique des accords, dans ce qu'il a de plus général, est inépuisable. En effet, sans hypothèse sur les fréquences disponibles ni sur le nombre de notes, les combinaisons de notes sont infinies, de même que les ensembles de hauteurs et les classes d'accords, et des accords existent de toutes les cardinalités.

Pour un entier N fixé, l'ensemble des classes d'accords, que celles-ci soient normales, premières, ou des identités, est assimilable à un volume borné de dimension $N-1$, plus précisément un $(N-1)$ -simplexe pour les identités.

Par contre, en se restreignant à une subdivision de l'octave, l'univers harmonique est limité à un nombre fini d'ensembles de hauteurs. Le potentiel combinatoire d'Estrada, appelé aussi « carotte », est une représentation compacte de tout l'univers harmonique, dont la taille dépend de la subdivision. Le premier niveau organise les identités par colonne selon leur cardinalité, tandis que le second niveau organise les classes normales par orbites identitaires.

Les descripteurs harmoniques statiques sont calculables sur les classes normales, à condition d'utiliser des spectres de type Shepard. Le passage d'un volume continu à un dénombrement fini permet en outre de visualiser les descripteurs sur les formes identitaires et sur les orbites identitaires, offrant ainsi des cartographies colorées de l'espace harmonique.

L'avantage de la carotte est double, il réside à la fois dans sa compacité et dans son adaptation à toutes les échelles divisant régulièrement un intervalle de référence. Aussi offre-t-elle un outil particulièrement pertinent d'investigation des univers micro-tonaux, puisqu'elle permet de synthétiser en une représentation compacte un univers harmonique complexe et souvent peu exploré. Les descripteurs harmoniques statiques apportent une description acoustique des éléments de cette représentation.

La coloration de la carotte peut avoir des applications en analyse musicale, où elle prolongerait le type d'analyse déjà mené sur la carotte, ainsi qu'en création musicale, puisqu'elle peut guider les compositeurs dans les univers micro-tonaux.

9.3 Enchaînements harmoniques dans le Tonnetz

Les parties 9.1 et 9.2 traitent le cas des descripteurs harmoniques statiques calculés sur les classes premières et normales, respectivement dans des espaces de classes continus, avec un axe des fréquences lui-même continu, et discrets, liés à une subdivision de l’octave. Qu’en est-il des descripteurs harmoniques dynamiques ?

Comme nous l’avons vu¹, il n’y a pas de sens à vouloir caractériser les enchaînements entre classes normales ou premières par les descripteurs dynamiques non triviaux. En effet, les interactions entre spectres de deux accords successifs dépendent des hauteurs relatives des deux accords, informations qui ne sont pas préservées lorsque l’on passe aux classes normales.

Par contre, le calcul des descripteurs dynamiques devient possible dès lors que l’on manipule des ensembles de hauteurs, sans considérer comme équivalents les ensembles transposés l’un de l’autre. Les structures du type Tonnetz donnent à voir, sous l’allure d’un graphe, les ensembles de hauteurs de cardinal 3.

9.3.1 Le Tonnetz : présentation

Le Tonnetz est un réseau de hauteurs permettant une visualisation de certains types d’ensembles de hauteurs. Il a été introduit par Euler en 1739 dans son *Tentamen novae theoriae musicae* (Euler, 1739)², et repris trente-cinq ans plus tard dans *De harmoniae veris principiis* (Euler, 1774).

D’abord utilisé pour illustrer les questions liées à l’intonation juste, il est maintenant principalement employé avec le tempérament égal, notamment dans la théorie néo-riemannienne³.

1. Voir page 367.

2. Voir figure 9.33.

3. Voir (Cohn, 1998).

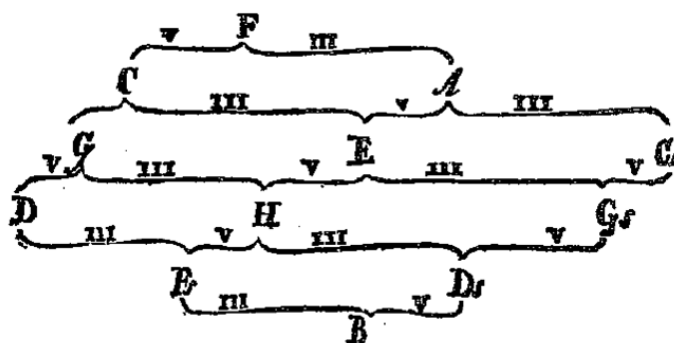


FIGURE 9.33 – Première apparition du Tonnetz, tirée de (Euler, 1739, p.147). L'axe diagonal met en rapport les quintes, et l'axe horizontal les tierces mineures.

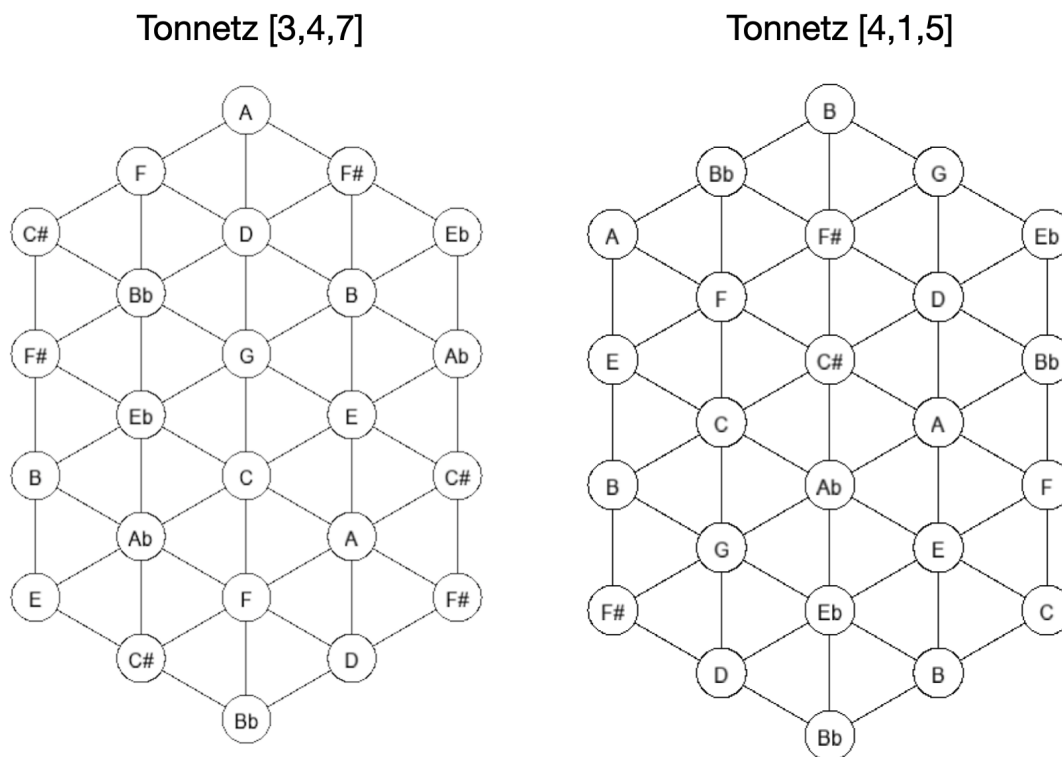


FIGURE 9.34 – Deux géométries de Tonnetz. À gauche, le Tonnetz standard [3,4,7], à droite le Tonnetz [4,1,5].

Le Tonnetz standard est un réseau plan dont les sommets sont les hauteurs. Celles-ci sont organisées selon trois axes : l'axe vertical est celui des quintes, l'axe diagonal allant vers le haut et la droite est celui des tierces majeures, et l'axe allant vers le haut et la gauche celui des tierces mineures. Les cellules fondamentales figurent des espaces de trois hauteurs. Ils sont de deux types : ceux tournés vers la droite sont les triades majeures, ceux tournés vers la gauche les triades mineures. La périodicité des hauteurs en tempérament égal confère au Tonnetz une structure toroïdale, et on lui préfère parfois la représentation toroïdale à la représentation plane, nécessairement redondante.

À côté de cette géométrie standard, il existe d'autres géométries qui diffèrent par les intervalles (ou les classes d'intervalles) portés par les axes. La figure 9.34 montre le tonnetz standard correspondant au réseau [3,4,7] (en demi-tons), ainsi que le tonnetz [4,1,5], dans lequel les cellules fondamentales sont du type *do-do♯-fa* et *do-mi-fa*. Nous restreignons ici la représentation à une partie du Tonnetz plan permettant la visualiser chacun des cellules fondamentales.

Les différentes géométries des réseaux de notes permettent d'explorer a priori toutes sortes d'univers harmoniques, que ceux-ci relèvent ou non d'une logique tonale. Bigo (Bigo, 2013) a introduit une notion de compacité dans le Tonnetz qui permet, pour un univers harmonique donné¹, de lui associer la géométrie optimale du Tonnetz, au sens de celle qui rend sa représentation la plus compacte. Le logiciel Hexacorde (Bigo et al., 2015)² donne à visualiser, à partir de morceaux midi, les trajectoires harmoniques sur les différentes géométries de Tonnetz.

Le Tonnetz est particulièrement adapté à l'étude des transformations harmoniques. Notamment, les trois enchaînements élémentaires entre accords parfaits de la théorie néo-riemannienne correspondent à des mouvements élémentaires sur le Tonnetz. Ces transformations, représentées sur la figure 9.35, sont les suivantes³ :

- R (*Relative*), met en relation les triades partageant une tierce majeure, par exemple *Fa M* et *Ré m* ;
- L (*Leading-tone*), met en relation les triades partageant une tierce mineure, par exemple *Fa M* et *La m* ;

1. Lequel peut être constitué de l'ensemble des accords présents dans une œuvre donnée.

2. Téléchargeable à l'adresse <https://louisbigo.com/hexachord>.

3. Tiré de (Cohn, 1998).

- P (*Parallel*), met en relation les triades partageant une quinte, par exemple *Fa* M et *Fa* m.

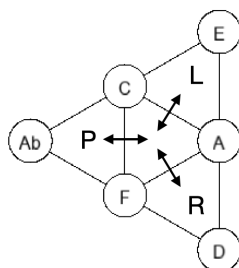


FIGURE 9.35 – Les trois transformations élémentaires de la théorie néo-riemannienne.

Extensions du Tonnetz Il existe de nombreuses variantes et extensions du Tonnetz, qui se proposent d'étendre sa portée. Citons notamment les Tonnetz en 3 dimensions, dans lesquels les volumes élémentaires sont des tetrachordes ([Gollin, 1998](#)), le *Spiral Array*, dans laquelle l'organisation spirale du Tonnetz permet de se passer de l'hypothèse du tempérament égal ([Chew, 2000](#)), ou encore un plongement du Tonnetz en 4 dimensions, qui à résoudre les incohérences résultant du plongement en 3 dimensions de la représentation toroïdale ([Baroin, 2011](#)).

Exploitation des arêtes Dans l'optique d'une coloration du Tonnetz par les descripteurs statiques, nous proposons ici une version étendue à peu de frais du Tonnetz, qui permet de représenter un plus grand nombre d'ensemble de hauteurs. L'idée est d'épaissir les arêtes de manière à ce qu'elles figurent des triades.

À partir de Tonnetz standard, cette version augmentée, présentée figure [9.36](#), donne accès aux ensembles de hauteurs des accords diminués, des accords augmentés et des accords de quarts superposés.

- Les accords diminués sont représentés par les segments partant de chaque hauteur dans la direction des tierces mineures.
- Les ensembles d'accords augmentés, qui sont au nombre de 4 dans l'échelle chromatique, sont représentés par les grandes diagonales sur l'axe des tierces majeures.

- Les accords de quarts superposées sont représentés par les segments partant de chaque hauteur dans la direction verticale des quintes.

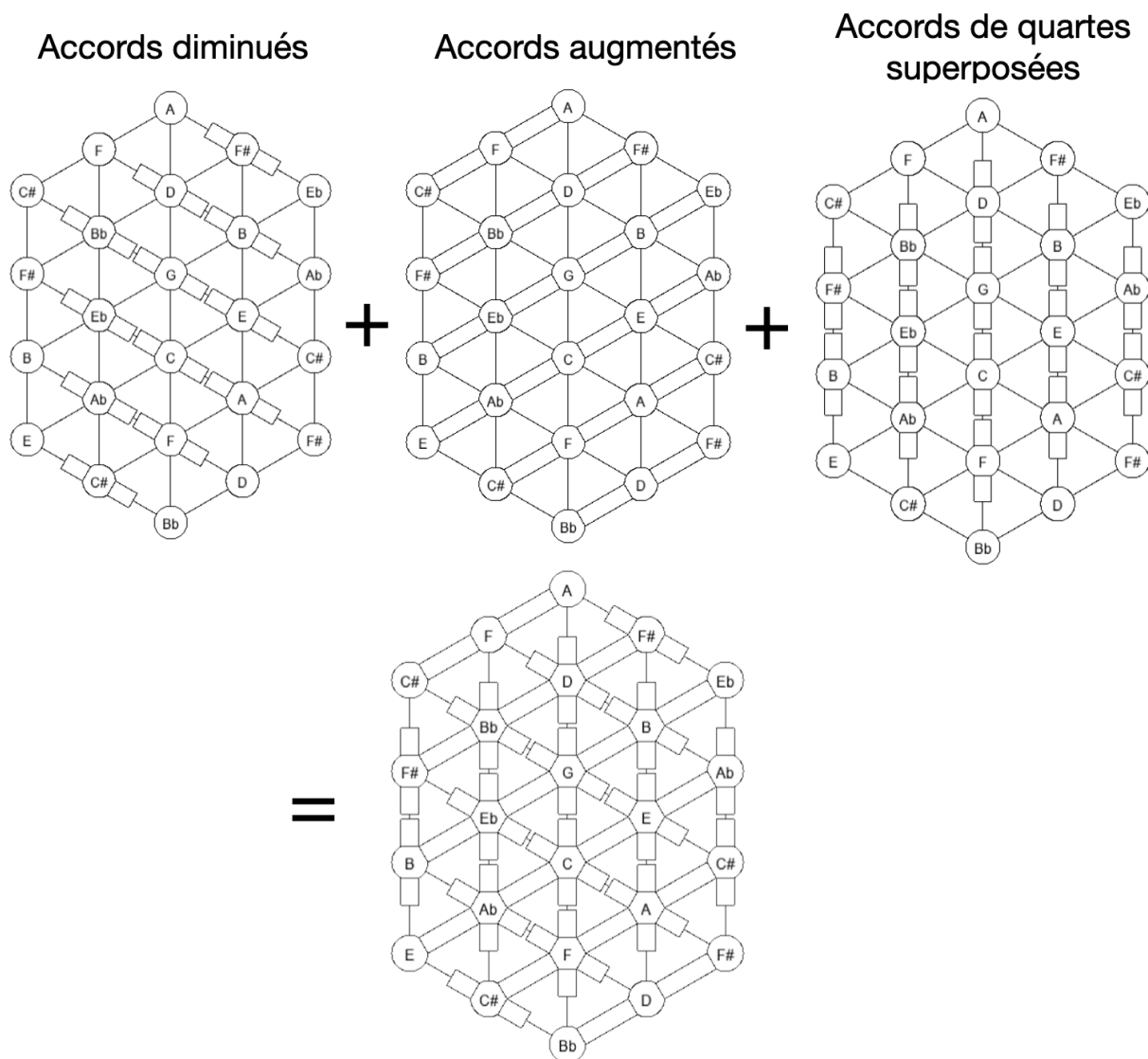


FIGURE 9.36 – Tonnetz augmenté : exploitation des arêtes du Tonnetz standard pour figurer les accords diminués, les accords augmentés et les accords de quartes superposées.

9.3.2 Coloration du Tonnetz par les descripteurs harmoniques dynamiques

Le Tonnetz augmenté, dans la géométrie standard, constitue un espace de représentation de 5 types d'ensembles de hauteurs parmi les triades : les accords majeurs, mineurs, diminués, augmentés et de quartes superposées, dans l'échelle chromatique en tempérament égal. Aussi, contrairement à la carotte constituée d'identités et de classes normales¹, le Tonnetz offre-t-il un espace de visualisation des descripteurs dynamiques sur les enchaînements harmoniques².

Implémentation Nous avons développé en Python un logiciel de visualisation des descripteurs dynamiques sur les enchaînements harmoniques par coloration du Tonnetz augmenté, avec une prise en compte du contexte harmonique par un modèle de mémoire à court terme³.

La géométrie par défaut est la géométrie standard [3,4,7], mais le logiciel donne la possibilité de changer de géométrie, auquel cas les types d'accords représentés diffèrent de ceux cités précédemment. Les exemples sont donnés ici dans la géométrie standard.

Le principe du logiciel est le suivant : on commence par cliquer sur un des ensembles de hauteurs (qu'on appelle ici « accords » par abus de langage) du Tonnetz, qui sera le premier accord de l'enchaînement. Une portée s'affiche alors au-dessus du Tonnetz, qui sert à garder la mémoire des accords joués. Dans le même temps, tous les accords se colorent, comme sur la figure 9.39. La couleur correspond, via une assignation préalable des descripteurs dynamiques à un code couleur, aux valeurs des descripteurs sur la transition potentielle entre les accords présents sur la partition, en l'occurrence le premier accord, et l'accord colorié. Les valeurs de chaque descripteur sont normalisées linéairement de 0 à 100 sur l'ensemble des transitions potentielles. Elles ne sont pas pré-enregistrées, comme c'était le cas avec la carotte, mais sont calculées en direct. Lorsqu'on clique sur un deuxième accord, la portée s'actualise, et les couleurs décrivent alors la transition entre les deux premiers accords et le prochain accord potentiel. Les accords de l'enchaînement sont ainsi rentrés un à un. Sans prise en compte de la mémoire (ce qui correspond

1. Voir partie 9.2.3.

2. Voir page 326.

3. Voir page 164.

à une longueur de mémoire de 1), la coloration est déterminée seulement par le dernier accord de la portée. En revanche, avec prise en compte de la mémoire, la coloration est déterminée également par les accords précédents. Un passage de la souris sur les accords affiche dans le terminal les valeurs numériques des descripteurs.

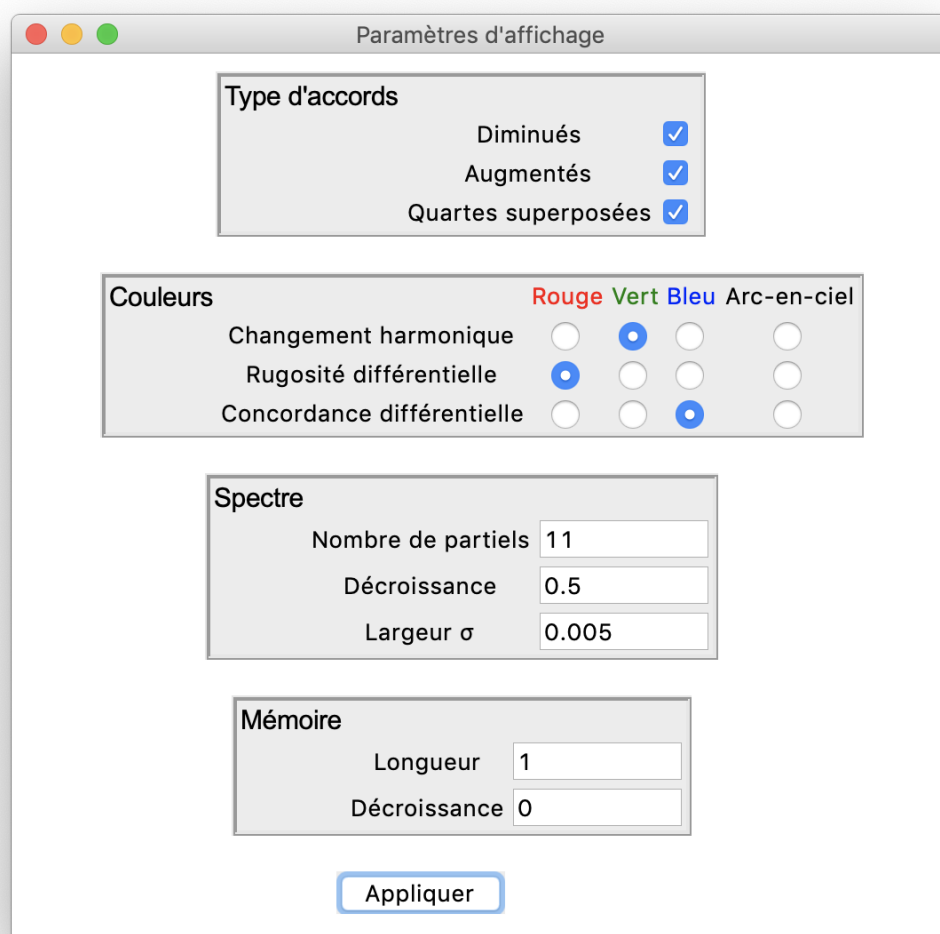


FIGURE 9.37 – Interface du logiciel de coloration du Tonnetz par les descripteurs dynamiques.

Une interface, présentée figure 9.37, permet de contrôler les types d'accords affichés, les paramètres de spectre, de mémoire, ainsi que l'assignation des descripteurs aux couleurs.

Les descripteurs dynamiques introduits dans ce travail sont au nombre de trois¹ : le changement harmonique, la concordance différentielle et la rugosité différentielle. L'assignation des couleurs aux différents descripteurs suit les mêmes principes que dans le cas de la carotte², avec le mode RVB et le mode « arc-en-ciel ». Les spectres utilisés sont là-aussi des spectres de type Shepard, de sorte que les descripteurs dynamiques peuvent être calculés sur les transitions entre ensembles de hauteurs. Une mémoire de longueur 1 revient à caractériser les transitions uniquement par le dernier accord de l'enchaînement, une mémoire de longueur 2 correspond à une prise en compte des deux derniers accords de l'enchaînement³.

Descripteurs dynamiques sans mémoire La figure 9.38 montre la concordance différentielle, le changement harmonique et la rugosité différentielle des transitions à partir de l'accord majeur *do-mi-sol*. Il convient ici de garder à l'esprit l'interprétation donnée page 236 de ces descripteurs en termes respectifs d'identité, de distinction et d'opposition.

Sur cet exemple, le changement harmonique et la concordance différentielle sont redondants car ils évoluent en sens inverse - les grandes valeurs de l'un correspondent aux petites valeurs de l'autre, et inversement.

On peut séparer le Tonnetz en quatre zones, selon le nombre de hauteurs communes dans l'enchaînement. Malgré la continuité de l'échelle des couleurs, ces zones se distinguent entre elles par le comportement des descripteurs sur les transitions :

1. La 1^{re} zone correspond à la répétition du même accord. Elle se caractérise comme prévu par un changement harmonique et une rugosité différentielle nulle ainsi qu'une concordance différentielle maximale.
2. La 2^{de} zone correspond aux transitions laissant deux hauteurs invariantes. Parmi elles figurent les transformations L, P et R de la théorie néo-riemannienne⁴, plus cinq autres transitions. Elle est caractérisée par un changement harmonique et une rugosité différentielle très

1. Voir partie 6.2.

2. Voir page 368.

3. Voir figure 4.9.

4. Voir figure 9.35.

faible, ainsi qu’une très forte concordance différentielle.

3. La 3^e zone est composée des transition laissant une seule hauteur invariante. Visuellement, elle correspond à la « couronne » autour de la seconde zone, contenant 17 transitions. Elle est caractérisée par un changement harmonique et une rugosité différentielle faible, ainsi qu’une forte concordance différentielle.
4. La 4^e zone correspond aux transitions ne partageant aucune note avec l’accord précédent. Elle est caractérisée par un changement harmonique et une rugosité différentielle importantes, ainsi qu’une faible concordance différentielle.

La distinction des zones à partir des colorations par les descripteurs dynamiques, cela malgré le continuum de l’échelle des couleurs, montre que les représentations sont au moins cohérentes avec les données acoustiques de base.

La figure 9.39 synthétise les trois colorations de la figure 9.38 dans une coloration mixte, où la rugosité différentielle est associée à la composante rouge, le changement harmonique à la composante verte et la concordance différentielle à la composante bleue.

La transition identité qu’est la répétition maximise la concordance différentielle en minimisant les deux autres descripteurs. L’accord de *Fa*♯ M, très éloigné tonalement de *Do* M, maximise le descripteur de distinction, tandis que l’accord diminué *ré-fa-lab*, interprété dans la tonalité de *do* majeur comme le deuxième degré de l’emprunt à la tonalité homonyme *do* mineur, maximise le descripteur d’opposition, à savoir la rugosité différentielle.

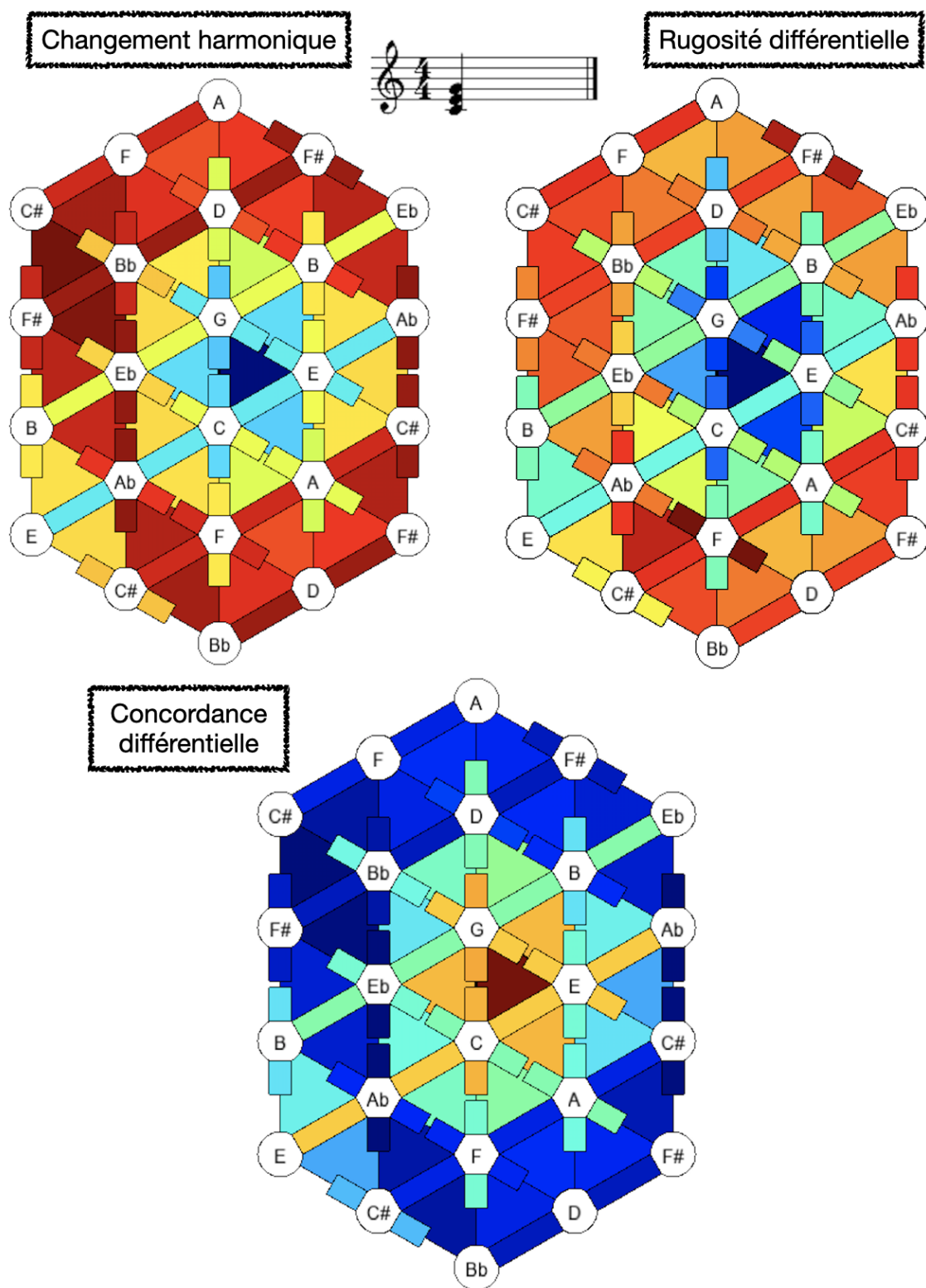


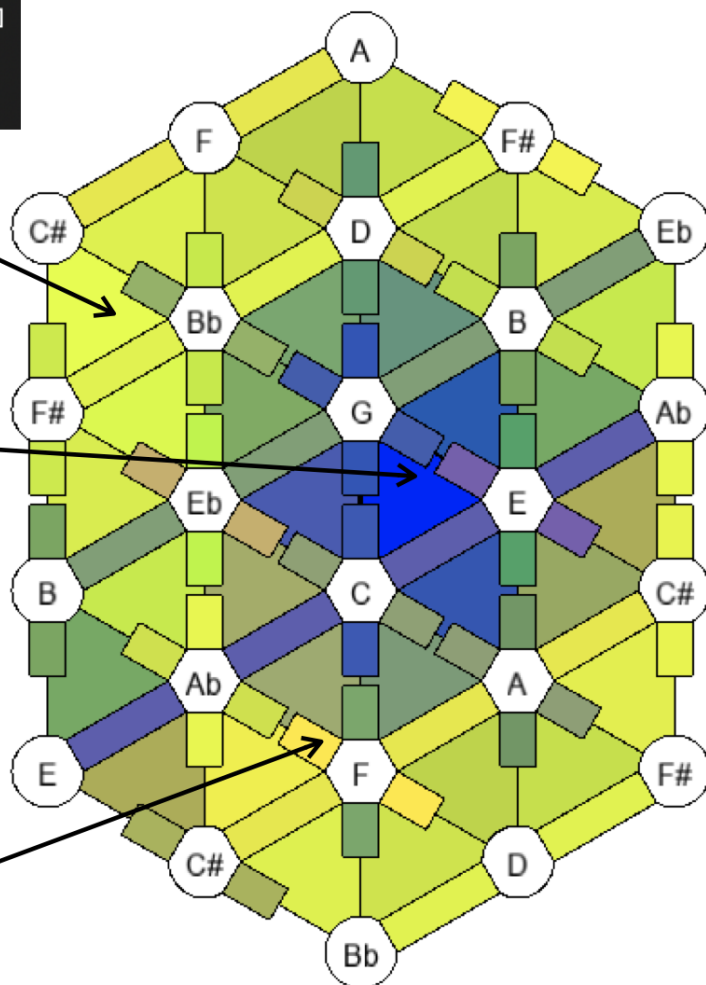
FIGURE 9.38 – Coloration du Tonnetz caractérisant les transitions à partir de l'accord majeur *do-mi-sol*. Caractérisation par le changement harmonique, la rugosité différentielle et la concordance différentielle. Rouge : valeurs élevées de descripteurs ; bleu : valeurs faibles. Paramètres de spectre : $K = 11$, $\sigma = 0.005$ et $decr = 0.5$.



Accord major [F# - Bb - C#]
 HarmonicChange: 100
 DiffConcordance: 0
 DiffRoughness: 87

Accord major [C - E - G]
 HarmonicChange: 0
 DiffConcordance: 100
 DiffRoughness: 0

Accord dim [D - F - Ab]
 HarmonicChange: 91
 DiffConcordance: 10
 DiffRoughness: 100



Couleurs	Rouge	Vert	Bleu	Arc-en-ciel
Changement harmonique	☐	☒	☐	☐
Rugosité différentielle	☒	☐	☐	☐
Concordance différentielle	☐	☐	☒	☐

FIGURE 9.39 – Coloration mixte des enchaînements avec l'accord majeur *do-mi-sol* par la rugosité différentielle (rouge), le changement harmonique (vert) et la concordance différentielle (bleu). Paramètres de spectre : $K = 11$, $\sigma = 0.005$ et $decr = 0.5$.

9.3.3 Description acoustique des résolutions cadentielles

Le modèle d'intégration de la mémoire à court terme¹ permet de traiter des enchaînements de plus de deux accords. Nous nous plaçons ici dans le cadre du langage tonal, en étudiant les résolutions de cadences.

Positionnement par rapport aux théories de l'harmonie tonale Les théories de l'harmonie tonale — théories des fonctions, théories génératives, analyse schenkérienne — sont des approches cognitives de l'harmonie décrivant les structures porteuses de signification musicale dans le langage tonal (Meeùs, 1992). À cette description « de l'extérieur », on peut opposer une description « de l'intérieur » de l'harmonie tonale, s'attachant à ce qui se situe en amont du traitement cognitif. Celle-ci peut être, comme nous l'avons vu, de nature arithmétique (théorie des coups au XVII^e siècle², degré de suavité d'Euler³), combinatoire (théorie néo-riemannienne (Cohn, 1998), *Set Theory* (Forte, 1973), représentations de trajectoires sur le Tonnetz (Bigo, 2013) ou sur la « carotte » d'Estrada), acoustique (représentations temps-fréquence, descripteurs audio) ou sensorielle (modèles de perception, courbes de rugosité). En contrepartie de leur aspect fondamental et de leur polyvalence, leur champ d'action n'étant en général pas cantonné au langage tonal, ces approches dites « de l'intérieur », en troquant le qualitatif contre le quantitatif, ne donnent pas accès aussi directement que les approches cognitives à ce qui relève de la signification musicale⁴. Leur pertinence en musicologie tient à leur capacité de structurer le matériau musical, et ainsi éclairer, sous les angles tantôt arithmétiques, combinatoires, acoustiques ou sensoriels, des éléments de langage ou d'écriture.

Nous utilisons ici les descripteurs harmoniques dynamiques pour caractériser les résolutions tonales « régulières », via une visualisation de l'ensemble des résolutions potentielles sur le Tonnetz. Nous appelons « résolutions régulières » les résolutions canoniques, autrement dit attendues, dans le langage tonal. Cette approche est motivée par trois points communs aux théories tonales, à savoir l'aspect dynamique, l'importance du contexte et la considérations des ensembles de hauteurs. Précisons chacun d'eux.

-
1. Voir page 164.
 2. Voir partie 2.1.3.
 3. Voir partie 2.1.4.
 4. Voir partie 1.1.4.

Les notions de conduite des voix, de résolution ou de modulation sont des notions dynamiques, centrales dans la musique tonale. Aussi, plusieurs théories tonales sont-elles fondées sur les transitions : la théorie néo-riemannienne (Cohn, 1998), les théories transformationnelles d’Arnold Schoenberg (Schoenberg et Stein, 1969) et Yizhak Sadaï (Sadai, 1980) ou la théorie des vecteurs harmoniques de Nicolas Meeùs (Meeùs, 2003). Aussi adoptons-nous, pour caractériser acoustiquement des éléments du langage tonal, une approche dynamique.

Le deuxième point concerne la prise en compte du contexte harmonique. Dans la théorie tonale, la fonction tonale d’un accord est déterminée par sa relation aux accords environnants, c’est-à-dire à son contexte (Meeùs, 1992). Certaines notions, comme celle de tonalité, sont globales, et ne peuvent pas être déterminées à partir d’un seul accord. En outre, dans la perception des accords, le contexte agit, nous l’avons vu¹, par la mémoire à court terme. Pour ces raisons nous utilisons ici un modèle de mémoire à court terme² pour tenir compte du contexte harmonique.

Même si elle ne le fait pas systématiquement, la théorie tonale traite parfois les ensembles de hauteurs, sans considérations relatives au renversement, aux doublures, au registre, à la disposition. C’est le cas lorsqu’il s’agit de déterminer les degrés tonaux. Nous calculons ici les descripteurs dynamiques sur les transitions entre ensembles de hauteurs, à la fois pour se rapprocher de ce versant de la théorie tonale, ainsi que pour proposer une visualisation sur le Tonnetz.

Cadence mineure Intéressons-nous à un enchaînement harmonique dans la tonalité de *do* mineur, commençant par le premier degré et allant à la dominante en passant par la sous-dominante, comme sur la portée de la figure 9.40. Nous considérons ici des ensembles de hauteurs, pour lesquels la notion de renversement n’est pas pertinente.

De la même manière que pour les transitions à partir de l’accord de *Do* M figure 9.39, on peut caractériser les transitions à partir de cet enchaînement par les descripteurs dynamiques, ce qui est fait figure 9.40. Le premier membre de la transition n’est plus un seul accord, mais une moyenne pondérée des trois accords, avec un poids d’autant plus important que l’accord est

1. Voir partie 5.1.3.

2. Ce modèle est détaillée page 167.

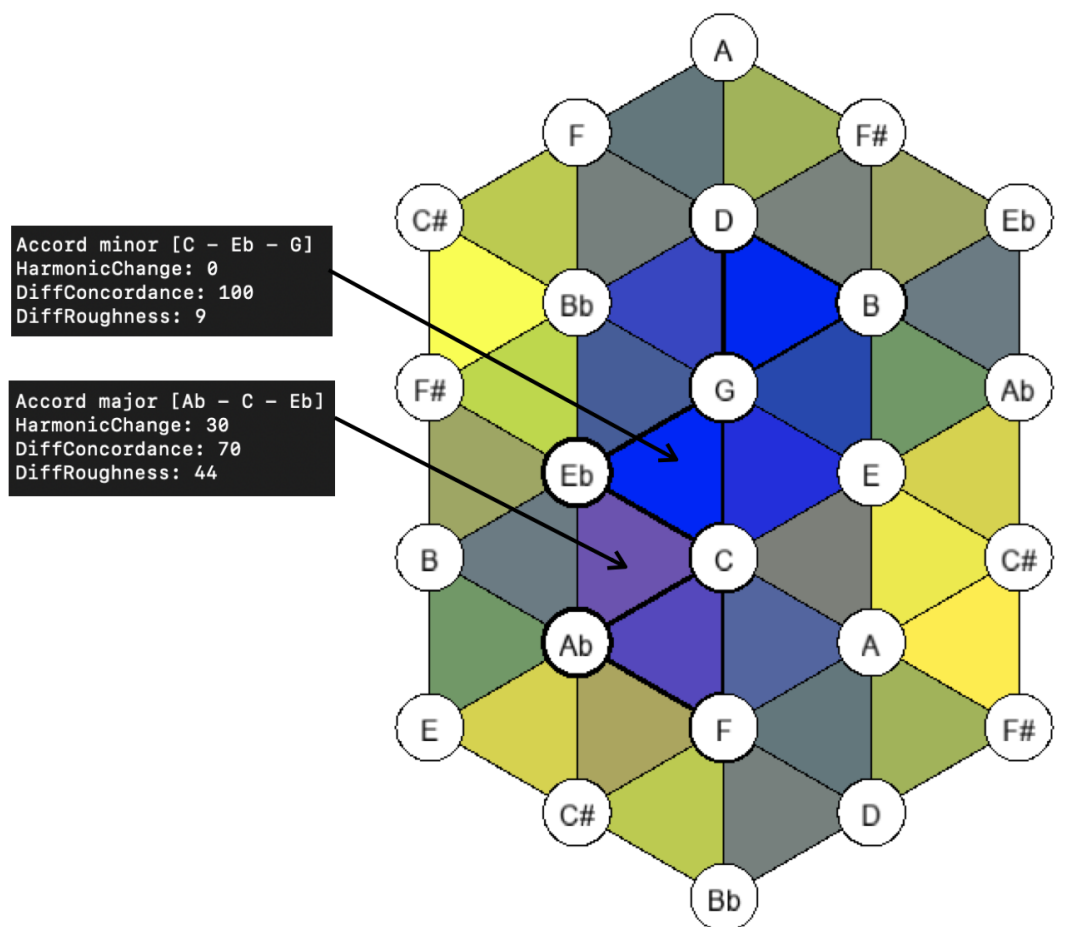
récent³. Avec une mémoire de longueur 3 et une décroissance de 0.5, l'accord de *Sol* M a un poids de 1, *Fa* m un poids de $\frac{1}{\sqrt{2}}$ et *Do* m un poids de $\frac{1}{\sqrt{3}}$. Le code couleur est le même que celui de la figure 9.39.

La figure 9.40 peut être interprétée comme une carte de résolution de la cadence mineure. On a restreint l'affichage, par soucis de clarté, aux accords mineurs et majeurs. La résolution par le premier degré *Do* m se démarque des autres résolutions possibles en ce qu'elle maximise la concordance différentielle, en plus d'avoir une très faible rugosité différentielle (la plus faible après la répétition de *Sol* M). De plus, la cadence rompue avec résolution sur le VI^e degré *Lab* M se distingue des autres résolutions par sa couleur violette, caractéristique d'un certain degré de rugosité différentielle, en plus d'une concordance différentielle élevée. Ces observations fournissent les premières pistes d'une interprétation perceptive des descripteurs dynamiques, dans le langage tonal et avec prise en compte de la mémoire : la concordance différentielle serait liée à la notion d'affinité, c'est à dire de bon enchaînement des accords, tandis que la rugosité différentielle serait liée à la surprise. Le premier point est déjà connu, et c'est justement en tant que modèle d'affinité que la concordance harmonique a d'abord été introduite¹.

On peut se demander dans quelle mesure les résultats sur la résolution de la cadence mineure étaient prévisibles à partir de la seule considération des hauteurs fondamentales, lisibles sur la partition. La figure 9.41 vient y répondre : la considération des hauteurs fondamentales ne suffit pas. Cette fois, les spectres utilisés n'ont qu'un seul partiel. Les descripteurs peuvent alors être approchés sans peine par des calculs à la main directement sur la partition, puisque les seuls partiels présents sur le spectre correspondent aux hauteurs notées. En réduisant le spectre à un unique partiel, l'allure de la « carte de résolution » de la cadence diffère. La transition qui maximise la concordance différentielle n'est plus la résolution vers le premier degré, mais la répétition de la dominante, qui minimise d'ailleurs la rugosité différentielle. La résolution vers le premier degré a une forte concordance différentielle, mais aussi une rugosité différentielle relativement importante, ce qui n'était pas le cas avec un spectre plus riche. De plus, la résolution vers le VI^e degré ne se démarque plus des résolutions qui l'entourent.

3. Voir partie 5.1.3.

1. Voir partie 6.2.2.



Tonnetz [3,4,7], timbre (11,0.5,0.005), mémoire (4 acc, decr: 0.5)

Couleurs	Rouge	Vert	Bleu	Arc-en-ciel
Changement harmonique	☐	☒	☐	☐
Rugosité différentielle	☒	☐	☐	☐
Concordance différentielle	☐	☐	☒	☐

FIGURE 9.40 – Résolution d’une cadence mineure - coloration mixte du Tonnetz par la rugosité différentielle (rouge), le changement harmonique (vert) et la concordance différentielle (bleu), avec prise en compte de la mémoire (décroissance 0.5) Paramètres de spectre : $K = 11$, $\sigma = 0.005$ et $decr = 0.5$.

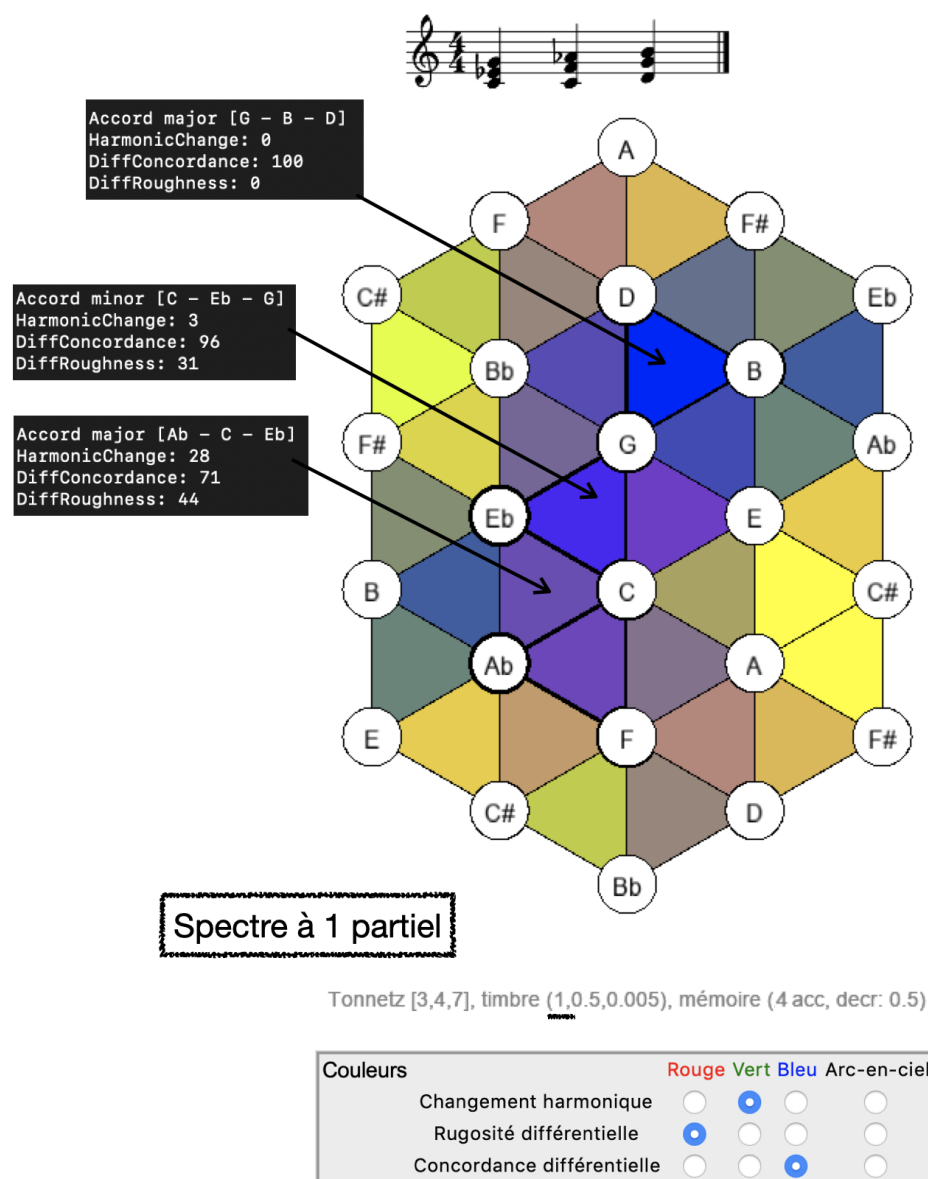


FIGURE 9.41 – Résolution d’une cadence mineure - coloration mixte du Tonnetz par la rugosité différentielle (rouge), le changement harmonique (vert) et la concordance différentielle (bleu), avec prise en compte de la mémoire (décroissance 0.5) Paramètres de spectre : $K = 1$, $\sigma = 0.005$ et $decr = 0.5$.

Cadence majeure avec suspension de la tierce Considérons maintenant une cadence majeure qui se résout sur le premier degré, mais avec suspension de la tierce, comme sur la portée de la figure 9.42. La figure pré-

sente le profil des transitions à partir de cet enchaînement. Le coefficient de décroissance de la mémoire est le même que précédemment, de sorte que l'accord de quartes superposées *do-fa-sol* a un poids de 1, *Sol* M un poids de $\frac{1}{\sqrt{2}}$, *Fa* M un poids de $\frac{1}{\sqrt{3}}$ et *Do* M un poids de $\frac{1}{\sqrt{4}}$. La transition attendue est alors la résolution de la tierce par mouvement conjoint du *fa* vers le *mi*.

Parmi toutes les transitions possibles vers les accords majeurs ou mineurs, la transition régulière vers l'accord de *Do* M maximise les descripteurs de distinction et d'opposition (changement harmonique et rugosité différentielle).

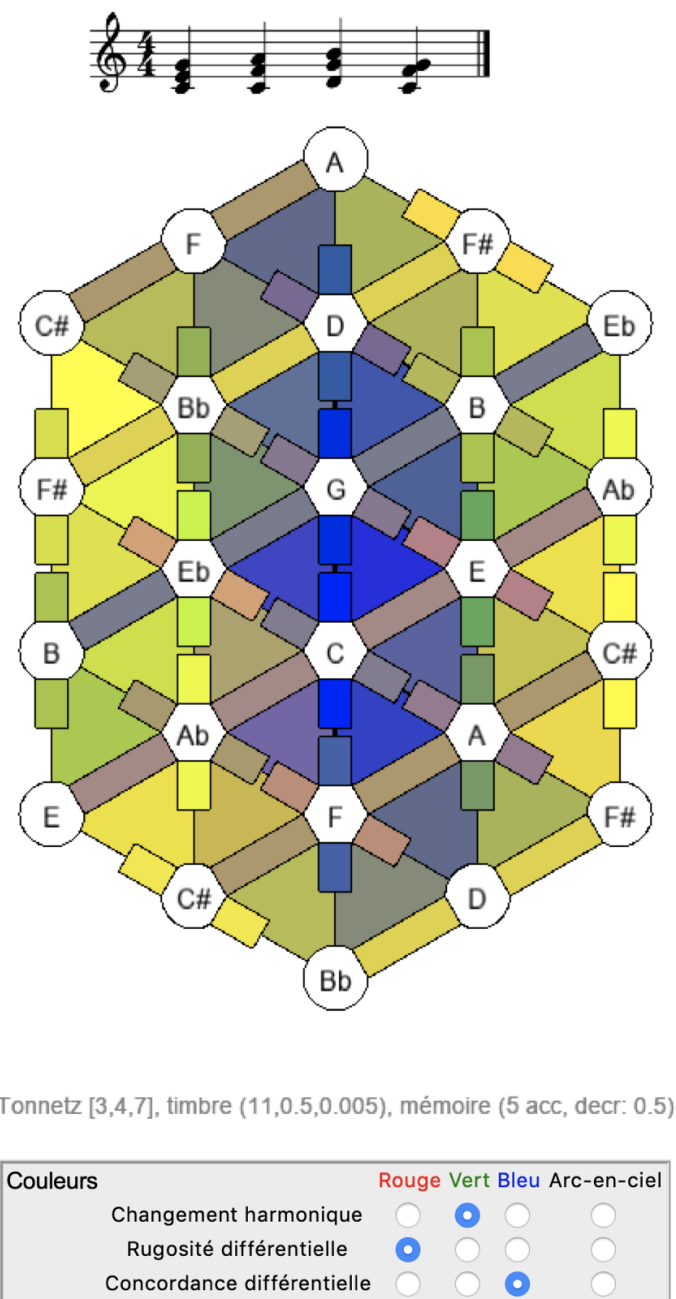


FIGURE 9.42 – Résolution du retard de la tierce dans une cadence majeure - coloration mixte du Tonnetz par la rugosité différentielle (rouge), le changement harmonique (vert) et la concordance différentielle (bleu), avec prise en compte de la mémoire (décroissance 0.5) Paramètres de spectre : $K = 11$, $\sigma = 0.005$ et $decr = 0.5$.

9.3.4 Conclusions et perspectives

Synthèse des résultats Sur les deux exemples de cadences traités, nous avons caractérisé acoustiquement la résolution régulière vers le premier degré, par le biais des interactions spectrales successives entre l'accord d'arrivée et le contenu de la mémoire. Parmi les résolutions vers les accords majeurs et mineurs, la résolution vers le premier degré maximise la concordance différentielle et minimise le changement harmonique ainsi que la rugosité différentielle. L'interprétation de ces trois descripteurs en termes d'identité, de distinction et d'opposition permet d'envisager la résolution régulière vers le premier degré comme la transition vers ce qui est le plus identique et le plus en accord avec la mémoire. La cadence rompue avec résolution sur VI^e degré, dans le premier exemple, se distingue par sa rugosité différentielle et sa concordance différentielle simultanément élevées. La cadence rompue, laquelle crée de la surprise sans pour autant moduler, peut être interprétée, vis-à-vis de la mémoire, comme la transition vers un accord qui lui est relativement identique tout en créant avec elle une opposition.

Cette approche montre les liens entre les descripteurs dynamiques et certains enchaînements « canoniques » constitutifs d'un langage donné, en l'occurrence le langage tonal. Elle ne prétend pas justifier ni légitimer ces éléments de langage, ce qu'elle ne saurait faire, mais enrichir leur compréhension par une description « de l'intérieur » du phénomène harmonique à partir de modèles acoustiques.

Perspectives

Suivi des transitions Une coloration du Tonnetz comme celles données dans les figures précédentes constitue une carte des transitions potentielles à un instant donné d'une séquence d'accords. À l'instant d'après, un choix est réalisé parmi toutes les transitions potentielles, qui peut ou non privilégier tel descripteur dynamique. Ce choix réactualise la carte des transitions potentielles. Ce processus est équivalent à une chaîne de Markov homogène d'ordre égal à $L - 1$, où L désigne la longueur de la mémoire, et dont les probabilités de transition correspondent aux valeurs des descripteurs dynamiques.

Dans la même approche consistant à mettre en avant les liens entre les descripteurs dynamiques et certains enchaînements d'un langage donné, il peut être intéressant de suivre l'évolution des descripteurs dynamiques sur

toutes les transitions d'une séquence. Cette vision diachronique devrait en outre permettre de relier avec plus de finesse les descripteurs dynamiques à l'évolution du matériau musical. En étendant l'échelle temporelle des séquences d'accords, elle peut notamment avoir des applications en analyse musicale.

Intégration des aspects rythmiques et métriques La mémoire à court terme agit ici par intégration temporelle à un instant donné des éléments précédemment entendus. D'un point de vue cognitif, le modèle utilisé est rudimentaire, étant donné que la mémoire n'y est limitée que par le nombre d'accords, sans considérations sur leur durée respective. Or, il paraît raisonnable d'intégrer les aspects temporels que sont le rythme et la métrique dans le modèle. En effet, il est souhaitable qu'un accord joué précédemment prenne d'autant plus de place dans la mémoire qu'il dure longtemps, et que son attaque soit située sur un temps fort. De cette manière, l'impact cognitif d'un accord dans la mémoire n'est fonction plus seulement du nombre d'événements s'étant déroulés depuis, mais aussi de sa durée et sa position métrique. Un tel modèle permettrait en outre de se rapprocher de la réalité perceptive.

Le modèle de Jensen¹(Jensen et Hjortkjær, 2012) intègre l'aspect rythmique, mais pas la métrique. Une autre possibilité, à moindres frais, pour intégrer le rythme, serait d'échantillonner les accords avec un pas qui serait un dénominateur commun de toutes les durées considérées. Un accord long serait alors équivalent à un accord court répété un certain nombre de fois. Une telle implémentation est envisageable uniquement pour les sons stables, sans attaque.

Intégration des descripteurs statiques Seuls les descripteurs dynamiques entrent en compte dans la coloration du Tonnetz. On peut cependant envisager une approche hybride, qui intègre les descripteurs statiques. En effet, les descripteurs statiques, calculables sur les classes normales, le sont *a fortiori* sur les ensembles de hauteurs. Tous les ensembles de hauteurs du même type, c'est à dire transposés l'un l'autre, ont les mêmes valeurs de descripteurs, mais les valeurs diffèrent entre chaque type² (majeur, mineur, diminué, augmenté, de quartes superposées).

1. Présenté page 166 et figure 5.1.

2. Voir partie 9.2.

Lors d'une transition, l'intégration des descripteurs statiques permettrait d'avoir une approche mixte combinant deux types d'interactions spectrales : les interactions dynamiques entre le contenu de la mémoire et le nouvel accord, ainsi que les interactions statiques entre les différentes hauteurs du nouvel accord. Les modalités d'une approche mixte mériteraient d'être développées, ce qui représente un travail théorique conséquent. Notamment des études expérimentale visant à comprendre comment s'articulent les dimensions statiques et dynamiques de l'harmonie dans la perception seraient les bienvenues.

Quatrième partie

Exemples musicaux

Dans la partie [III](#) nous avons essentiellement abordé le volet descriptif des descripteurs harmoniques, en étudiant la manière dont ils structurent les espaces harmoniques, à la fois les espaces continus et les espaces discrets d'accords ou de classes d'accords. Ce volet a un intérêt avant tout théorique, il permet d'aller plus en avant dans la compréhension des briques de base du phénomène harmonique.

L'intérêt des descripteurs harmoniques ne se borne pas, cependant, à l'aspect théorique, mais des applications sont envisageables dans les domaines de l'analyse et de la création musicale. Nous n'abordons pas dans ce travail leur potentiel comme guide à la composition, mais nous efforcerons de montrer par des exemples leur apport en analyse. Cet apport vient de leur capacité à mettre en avant des phénomènes acoustiques à l'œuvre dans une musique, qui ne peuvent pas nécessairement être déduits directement de l'écriture, c'est-à-dire de la partition, ni même de la perception. En cela réside leur valeur heuristique, qui met en lumière certains éléments d'écriture ou d'interprétation liés à des logiques d'interactions entre partiels.

Nous proposons ici de traiter un corpus de trois œuvres écrites à quelques années d'intervalle : *Fratres* d'Arvo Pärt (1977), la *Pièce pour orgue n°2* d'Alfred Schnittke (1980) et les *Cartas Celestes I* d'Almeida Prado (1974). Malgré leur proximité historique, ces œuvres font appel à des langages radicalement différents, intimement liés à la personnalité de leur auteur : le style *tintinnabuli* de Pärt, l'écriture du contraste et de l'ambiguïté harmoniques de Schnittke et le *transtonalisme* d'Almeida Prado. Ces pièces ont été choisies pour l'intérêt de leur écriture harmonique.

Les deux premières analyses sont très locales, éclairant de courts extraits des œuvres de Pärt et Schnittke. Celle des *Cartas Celestes I* est plus globale, car traite une section entière, en plus d'un catalogue d'« accords-alphabet ». Dans cette analyse, le rapport du compositeur à son écriture est questionné par le biais des descripteurs harmoniques.

Chapitre 10

Arvo Pärt - *Fratres*

Fratres est une œuvre emblématique du style « *tintinnabuli*² » du compositeur estonien Arvo Pärt. Composée en 1977 à l'origine pour ensemble mixte de 10 à 25 instruments, elle a été arrangée à six reprises pour des ensembles différents. Nous nous intéressons ici à la version pour violon et piano de 1980, parfois jouée à l'orgue à la place du piano, dont la partition est jointe en annexe 1.

Structure générale L'œuvre suit la structure suivante : après une introduction virtuose au violon³ est exposé au clavier un thème de 6 mesures⁴, ensuite répété 7 fois⁵, en étant transposé à chaque fois d'une tierce descendante tantôt majeure tantôt mineure.

Sur chaque occurrence du thème la répartition des voix est similaire entre pédale, voix mélodiques et voix *tintinnabulantes* :

- Le piano tient une pédale de quinte *la-mi*, parfois réduite à une note pour des raisons de réalisation au clavier.
- Sur cette pédale viennent se greffer 3 voix homorythmiques (note contre note), dont les deux extrêmes sont mélodiques. Ces deux voix mélodiques, évoluant en sixtes parallèles, sont conjointes et gravitent

2. Voir ([Amblard, 2019](#)).

3. Chiffre 1 sur la partition.

4. Chiffre 2 sur la partition.

5. Chiffres 3 à 9 sur la partition.

- autour d'un centre mélodique, comme souvent chez le compositeur ¹.
- La voix médium du piano, insérée entre les deux voix mélodiques, est la voix dite « *tintinnabulante* », non conjointe, jouant un rôle harmonique. La voix du violon, absente dans la version originale, reprend la voix *tintinnabulante* du piano en l'étoffant harmoniquement et rythmiquement dans chaque variation.

Le thème est formé de deux parties, chacune constituée de trois mesures 7/4, 9/4 et 11/4. La première mesure est composée de quatre notes formant deux mouvements conjoints, et les mesures suivantes procèdent par ajout de deux notes au centre qui viennent prolonger d'un cran chacun des deux mouvements conjoints.

La seconde partie du thème est obtenue par imitation rétrograde de chaque mesure de la première partie. La voix *tintinnabulante* est seule à ne pas suivre ce procédé, ce qui a pour effet de renouveler l'harmonisation entre les deux parties du thème.

10.1 Rôle harmonique et perceptif de la voix tintinnabulante

Nous nous concentrons ici sur le rôle de la voix *tintinnabulante* dans la première exposition du thème. La figure 10.1 met en évidence les changements de cette voix entre les parties 1 et 2, à la rétrogradation près de chaque mesure.

Les accords qui subissant une modification de la voix *tintinnabulante* sont ceux dont la voix *tintinnabulante* est en rapport de septième ou de neuvième avec l'une des deux voix mélodiques. La modification a pour effet de transformer ces intervalles en intervalles de quintes, diminuées ou justes, ou de quartes. Pour l'accord *sib-la-ré* (hors pédale), l'abaissement du *la* au *mi* élimine l'intervalle de septième majeure, mais fait apparaître un intervalle de septième mineure avec l'autre voix mélodique.

1. C'est le cas de *Für Alina* pour piano, composé un an avant *Fratres*.

Partie 1 du thème : version droite



Partie 2 du thème : version rétrogradée



FIGURE 10.1 – Partie de piano à la première exposition du thème (chiffre 2). Sur la première partie du thème sont entourés en rouge les intervalles de septième ou de neuvième formés par la voix *tintinnabulante* avec les voix mélodiques. Sur la version rétrogradée du thème sont indiquées en bleu les notes de la voix *tintinnabulante* qui diffèrent de la rétrogradation mesure par mesure de la voix *tintinnabulante* de la partie 1.

On peut faire l'hypothèse que cette modification de la voix *tintinnabulante* entre les deux parties du thème a pour effet, en éliminant des intervalles typiquement dissonants, d'abaisser la dissonance.

Pour s'en convaincre, procédons à un calcul des descripteurs harmoniques statiques sur une synthèse avec des sons d'orgue de la partie pour piano. Cela est cohérent musicalement puisque la partie de piano est parfois jouée à l'orgue¹.

La figure 10.2 met en regard la concordance, la rugosité et l'harmonicité des deux parties du thème. La partie 2 a été de nouveau rétrogradée mesure par mesure, de manière à faire correspondre les lignes mélodiques. Lorsque la voix *tintinnabulante* diffère entre les deux parties, elle est notée en rouge pour la partie 1, en bleu pour la partie 2.

1. Voir par exemple le disque (Nevolovitch et Van de Velde, 2012).

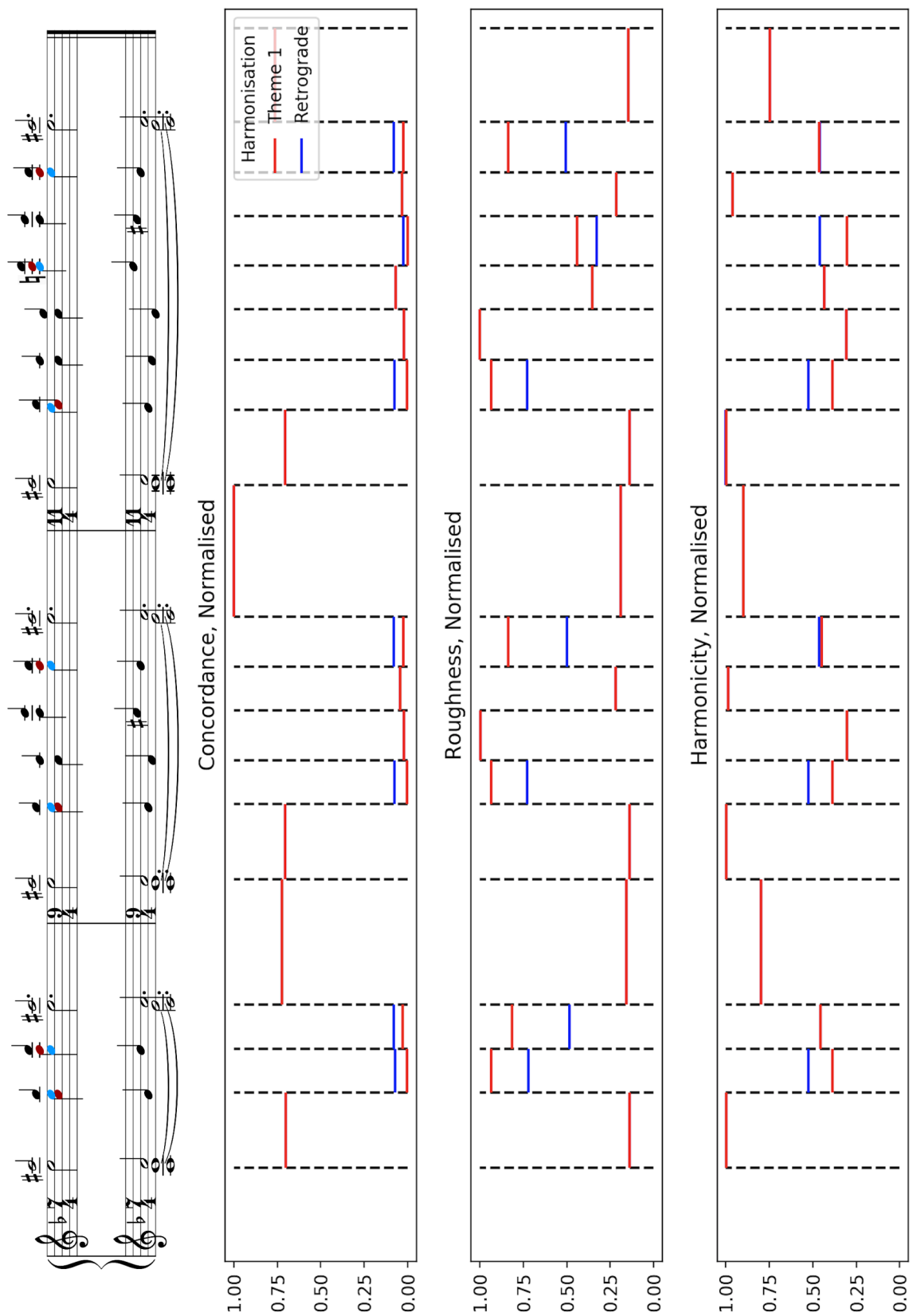


FIGURE 10.2 – *Fratres* de Pärt — comparaison de la concordance, de la rugosité et de l’harmonicité des harmonisations du thème droit et de sa version rétrogradée. Lorsque la voix *tintinnabulante* diffère entre les deux parties, elle est notée en rouge pour la partie 1, et en bleu pour la partie 2. Sons synthétisés avec les jeux *bourdon + Flûte 4*.

On observe que les modifications de la voix *tintinnabulante* entre les parties 1 et 2 ont pour effet, sur les accords concernés par le changement, de faire augmenter la concordance et l’harmonicité et de faire diminuer la rugosité, ce qui confirme notre hypothèse.

L’harmonicité et la faible rugosité étant considérées comme les deux principales causes psychoacoustiques de la consonance¹, la ré-harmonisation du thème dans sa présentation rétrogradée a pour effet de faire augmenter la consonance sur les accords concernés. Cela fournit une interprétation perceptive de cette ré-harmonisation. On peut raisonnablement faire l’hypothèse que celle-ci constitue une intention musicale de la part d’Arvo Pärt.

10.2 Ajout progressif d’accords

Revenons maintenant sur l’ajout progressif des accords entre les mesures 7/4 et 9/4, et 9/4 et 11/4, qui font prolonger d’un cran chacun des deux mouvements mélodiques conjoints du thème. Intéressons-nous à leur incidence sur les transitions harmoniques, par le biais des descripteurs dynamiques.

Il est possible de représenter les « trajectoires harmoniques » d’un extrait musical dans les espaces de descripteurs harmoniques à 2 ou 3 dimensions, en suivant les différentes positions prises par les valeurs des descripteurs. L’information temporelle est absente dans la considération des seules trajectoires, mais une numérotation permet de suivre l’ordre des valeurs.

La figure 10.3 montre les trajectoires harmoniques des premières trois mesures du thème dans l’espace des descripteurs dynamiques constitué de la concordance différentielle et de la rugosité différentielle. Là encore, les descripteurs sont calculés sur une synthèse avec des sons d’orgue (jeux *Bourdon* + *Flûte4*), et seule la partie de clavier est prise en compte.

Les transitions au sein de chaque mesure sont représentées par des points dans l’espace des descripteurs dynamiques, et les ensembles de points forment des trajectoires harmoniques.

Les trajectoires des trois mesures ont les mêmes points de départ et d’arrivée, étant donné qu’elles commencent et finissent par les mêmes transitions. De même, les transitions 2 et 4 de la mesure 9/4 sont les mêmes que les transitions 2 et 6 de la mesure 11/4. L’ajout de deux accords sur les mesures

1. Se référer au chapitre 3.

9/4 et 11/4 enrichit les trajectoires de 3 nouveaux points. Ainsi, la trajectoire de la 2^e mesure (orange) est plus étendue que la trajectoire de la 1^{re} mesure (en bleu), notamment par la transition n°2 caractérisée par une faible concordance différentielle et une forte rugosité différentielle, et la transition n°4 caractérisée par une forte concordance différentielle. De même, la trajectoire de la 3^e mesure enrichit l'espace exploré dans le plan des descripteurs, avec la transition n°3 caractérisée par sa faible concordance différentielle et les transitions n°4 et 5 caractérisées par leur faible rugosité différentielle.

Cette analyse nous montre qu'une régularité dans la conduite des voix n'est pas synonyme d'une régularité dans les interactions spectrales dynamiques. En effet, les points les plus éloignés l'un de l'autre sur les trajectoires, qui correspondent aux interactions spectrales dynamiques les plus divergentes, sont les transitions n°2 et 4 de la 2^e mesure (ou 2 et 6 de la 3^e mesure). Or, ces transitions, avec la transition n°3 de la 3^e mesure, suivent le même schéma en ce qui concerne la conduite des voix : la voix *tintinnabulante* ne bouge pas, tandis que les deux voix mélodiques procèdent par mouvement conjoint descendant. Cet écart entre la similarité de l'écriture et les disparités des interactions spectrales ne peut qu'être dû à la pédale tenu *la-mi*.



Comparaison des mesures
Fratres (DiffConcordance, DiffRoughness)
Normalisation on all the curves

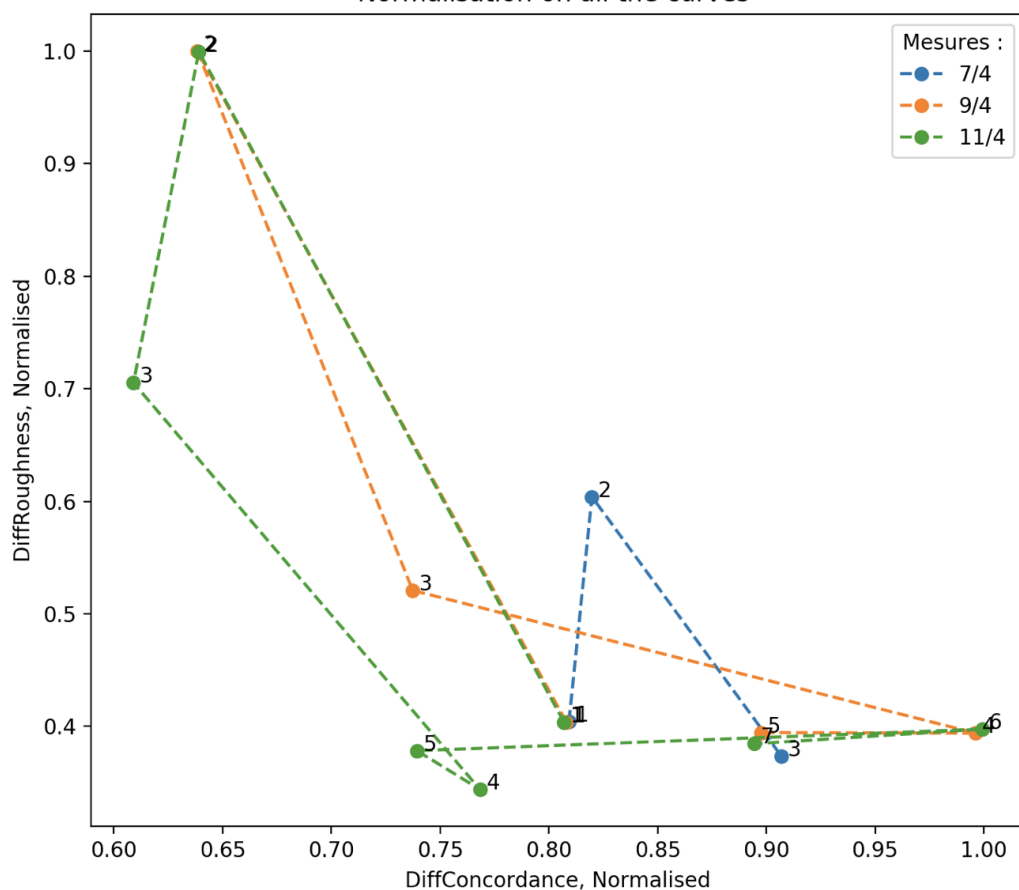


FIGURE 10.3 – *Fratres* de Pärt — trajectoires harmoniques des trois premières mesures du thème dans l’espace des descripteurs dynamiques constitué de la concordance différentielle et de la rugosité différentielle. Les descripteurs sont calculés sur une synthèse avec des sons d’orgue (jeux *Bourdon* + *Flûte4*), et seule la partie de clavier est prise en compte.

Chapitre 11

Alfred Schnittke - *Pièce pour orgue* op.149 n°2

Les deux petites *Pièces pour orgue* d'Alfred Schnittke, d'une durée approximative de 8 minutes, constituent la seule contribution du compositeur russe à l'orgue solo. Écrites en 1980, elles sont emblématiques des œuvres de la fin des années 70 - début des années 80, comme le *Concerto pour violon n°3* (1978), la *Première sonate pour violoncelle et piano* (1978), le *Concerto pour piano et orchestre à cordes* (1979) ou encore le *Quatuor n°2* (1980). Comme elles, les deux petites pièces pour orgue se caractérisent par le rôle des chromatismes, la recherche de dissonances, une écriture procédant par ajout de strates sonores, et une harmonie ambiguë qui oscille entre tonalité et atonalité, avec l'emploi d'amas faisant se superposer accords majeurs et mineurs². La partition des deux *Pièces pour orgue* est donnée en annexe 2.

Nous nous proposons d'étudier ici le début de la seconde pièce comme application de la coloration de la carotte³, en comparant les résultats obtenus par cette visualisation à trois interprétations différentes de l'extrait.

2. Voir ([Gaulhiac, 2019](#)), consacré à la sonate n°1 pour violoncelle et piano d'Alfred Schnittke.

3. Voir partie [9.2](#)

II

Sw: Gambe 8'
Hw: 16', 8', 4'. Sesquialter.
Trompète 8'
Ped: 16', 8', Fagott 16'.
Trompète 8'. Hw/Ped

FIGURE 11.1 – *Pièce pour orgue n°2* d’Alfred Schnittke — premières mesures.

Écriture de l’accumulation Les mesures 1-2 et 2-4 de la pièce n°2 suivent le même schéma. Les deux mains jouent en mouvements parallèles sur le même clavier, avec une octave d’écart. Les notes jouées sont tenues, créant une accumulation de strates sonores. Les deux mouvements sont entourés de silences.

Le *Tutti* de *Do* M mesure 5 est superposé mesure 6 à l’accord de *Mib* m, créant une ambiguïté modale.

Exploration des identités La figure 11.2 montre la trace des accords des 6 premières mesure de la pièce sur les identités de la carotte.

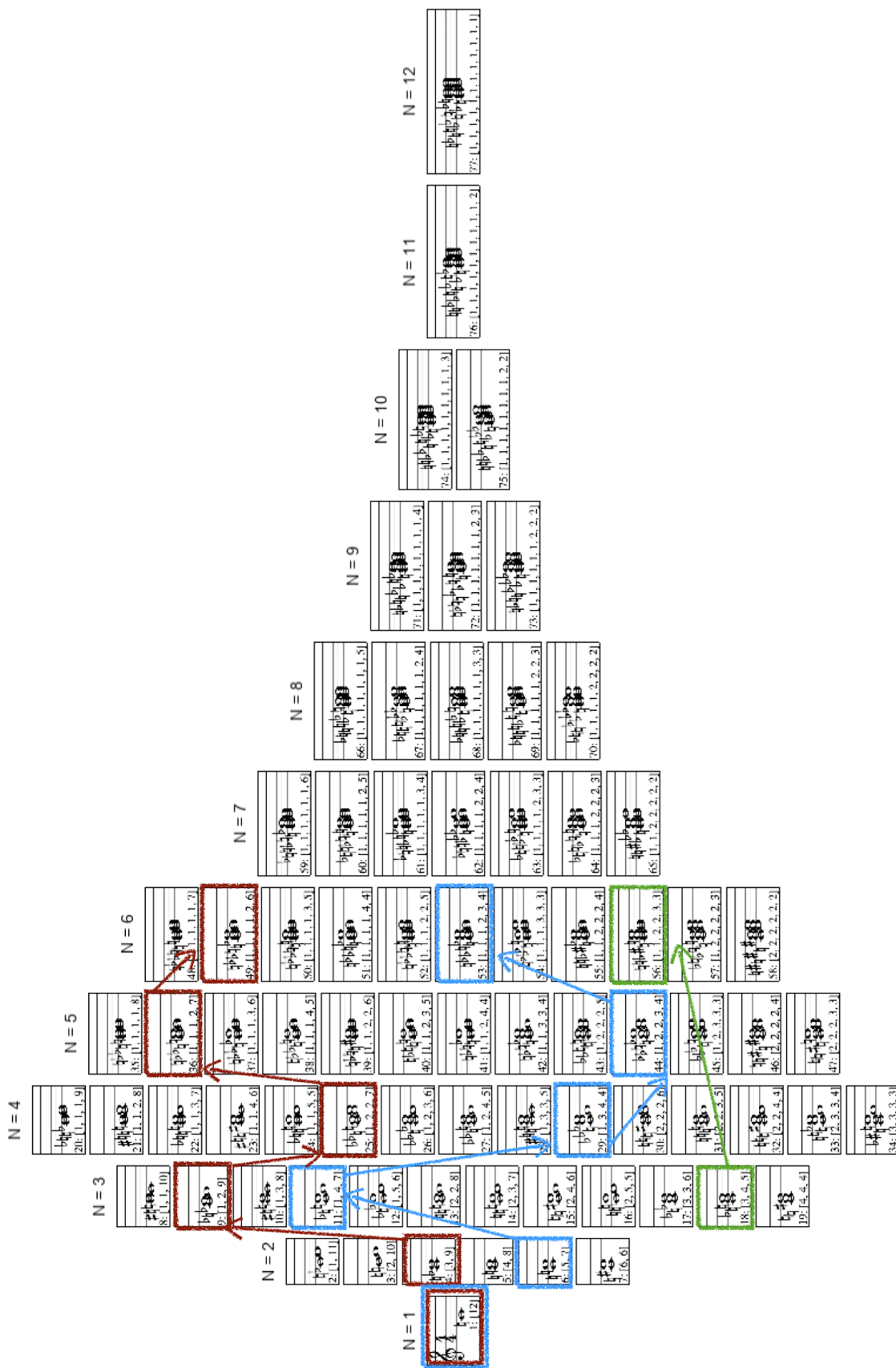


FIGURE 11.2 – Six premières mesures de la *Pièce pour orgue n°2* d'Alfred Schnittke - exploration des identités sur la carotte. En rouge : mesures 1-2, en bleu : mesures 3-4, en vert : mesures 5-6.

Les deux premiers cycles d'accumulation se traduisent par une trajectoire vers la droite, puisqu'ils procèdent par ajout d'une note. Mis à part l'unisson, chacune des trajectoires explore différentes identités, la trajectoire du deuxième cycle restant géométriquement en-dessous de celle du premier cycle. Les accords des mesures 5 et 6 sont également en-dessous des identités les précédant.

Ces observations vont dans le sens d'une volonté d'exploration du matériel harmonique, qui met de côté la réutilisation des mêmes identités.

Analyse par les descripteurs statiques de classes Il semble que les questions de consonance/dissonance et d'harmonicité/inharmonicité sont des éléments moteurs de l'écriture de Schnittke. En effet, chacun des trois groupes de deux mesures commence par des accords considérés comme consonants et harmoniques (unisson et accord majeur), pour aller vers des amas de 6 notes, considérés comme inharmoniques et dissonants.

Cette hypothèse invite à mettre en parallèle les trajectoires avec les colorations du Tonnetz présentées dans la partie 9.2.4. Cependant, une difficulté apparaît ici : ces colorations sont obtenues avec des spectres de Shepard, alors que l'œuvre est pour orgue. Des précautions s'imposent quant à l'interprétation que l'on pourra tirer de l'analyse. Néanmoins, il ne paraît pas déraisonnable, en première approximation, de procéder à une analyse avec des spectres de Shepard, du fait de la ressemblance des sons d'orgue avec les sons de Shepard. En effet, ces derniers, synthétisés¹ à partir des spectres employés pour la coloration de la carotte, s'approchent relativement bien perceptivement de sons d'orgue. De plus, la ressemblance est favorisée par les jeux employés et les doublures à l'octave sur le clavier.

La figure 11.3 montre un spectrogramme des deux premiers accords, dans l'interprétation de l'organiste Robert Pecksmith². Sur le spectre du premier accord, le *mib* est présent sur 7 octaves différentes, de *mib* 2 à *mib* 8. Cette duplication sur les différentes octaves rapproche ce son d'orgue des sons de Shepard.

1. Les sons de Shepard peuvent être synthétisés par synthèse additive, avec des logiciels comme *CSound*.

2. Interprétation disponible sur youtube à l'adresse <https://www.youtube.com/watch?v=GNzgpi1H4K0>.

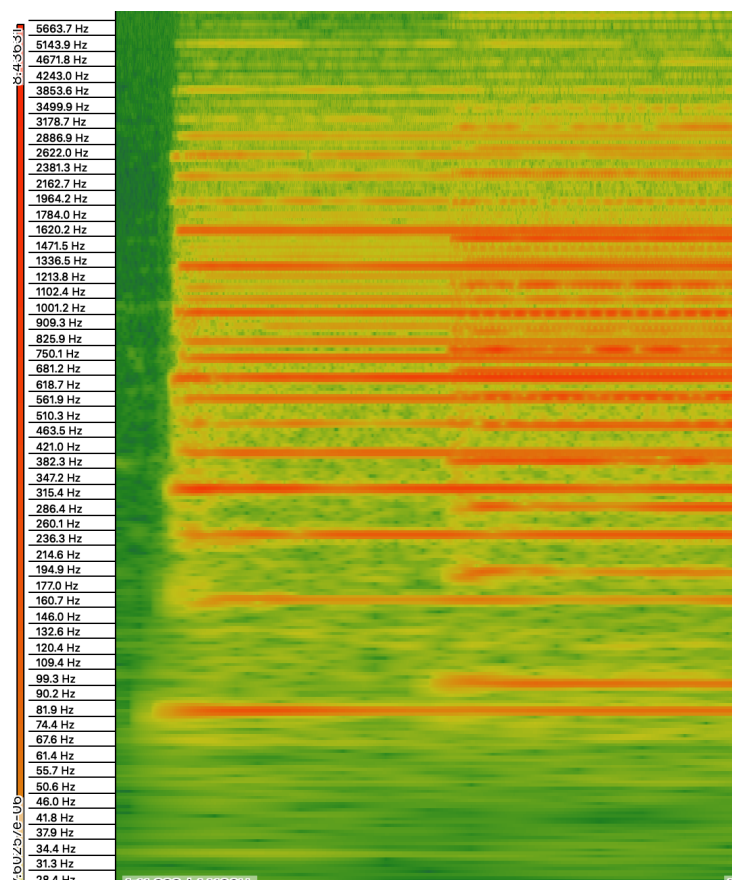


FIGURE 11.3 – Spectrogramme des deux premiers accords de la *Pièce pour orgue n°2* d'Alfred Schnittke, dans l'interprétation de Robert Pecksmith. TQC avec une précision de $1/36^e$ d'octave, obtenu avec *Sonic Visualiser*.

Les figures 11.4, 11.5 et 11.6 montrent respectivement les colorations des trajectoires sur la carotte par la rugosité, l'harmonicité et la tension. Sur chacune des trajectoires les mêmes variations de couleur sont observées : passage des couleurs froides aux couleurs chaudes pour la rugosité et la tension, des couleurs chaudes aux couleurs froides pour l'harmonicité. Ainsi, chacun des cycles d'accumulation, correspondant aux mesures 1-2, 3-4 et 5-6, est caractérisé par une augmentation de la rugosité et de la tension, en même temps qu'une diminution de l'harmonicité.

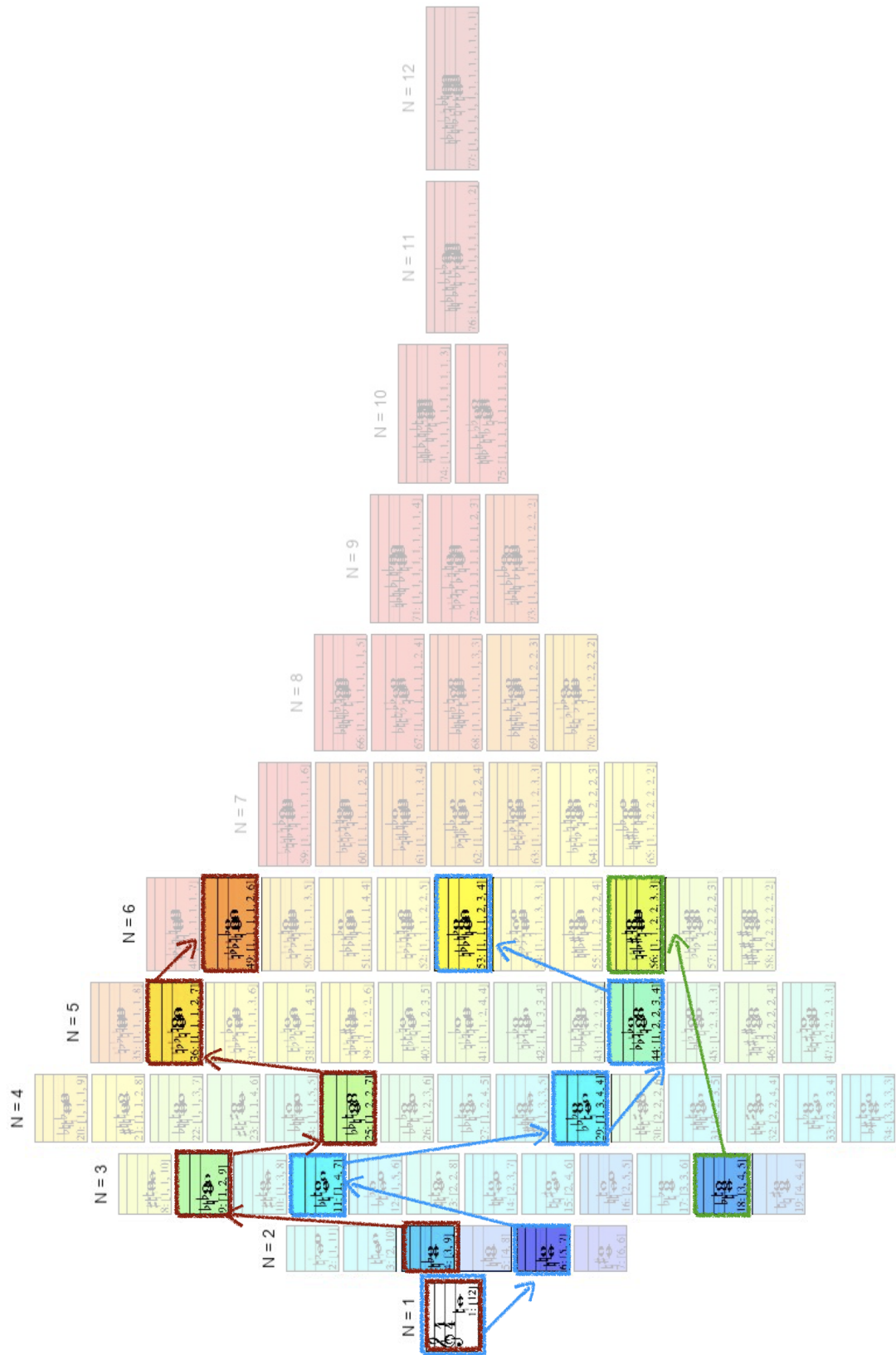
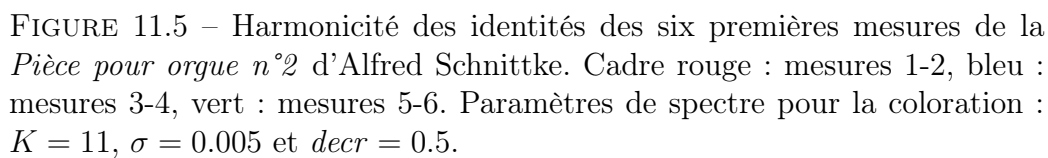


FIGURE 11.4 – Rugosité des identités des six premières mesures de la *Pièce pour orgue n°2* d'Alfred Schnittke. Cadre rouge : mesures 1-2, bleu : mesures 3-4, vert : mesures 5-6. Paramètres de spectre pour la coloration : $K = 11$, $\sigma = 0.005$ et $decr = 0.5$.



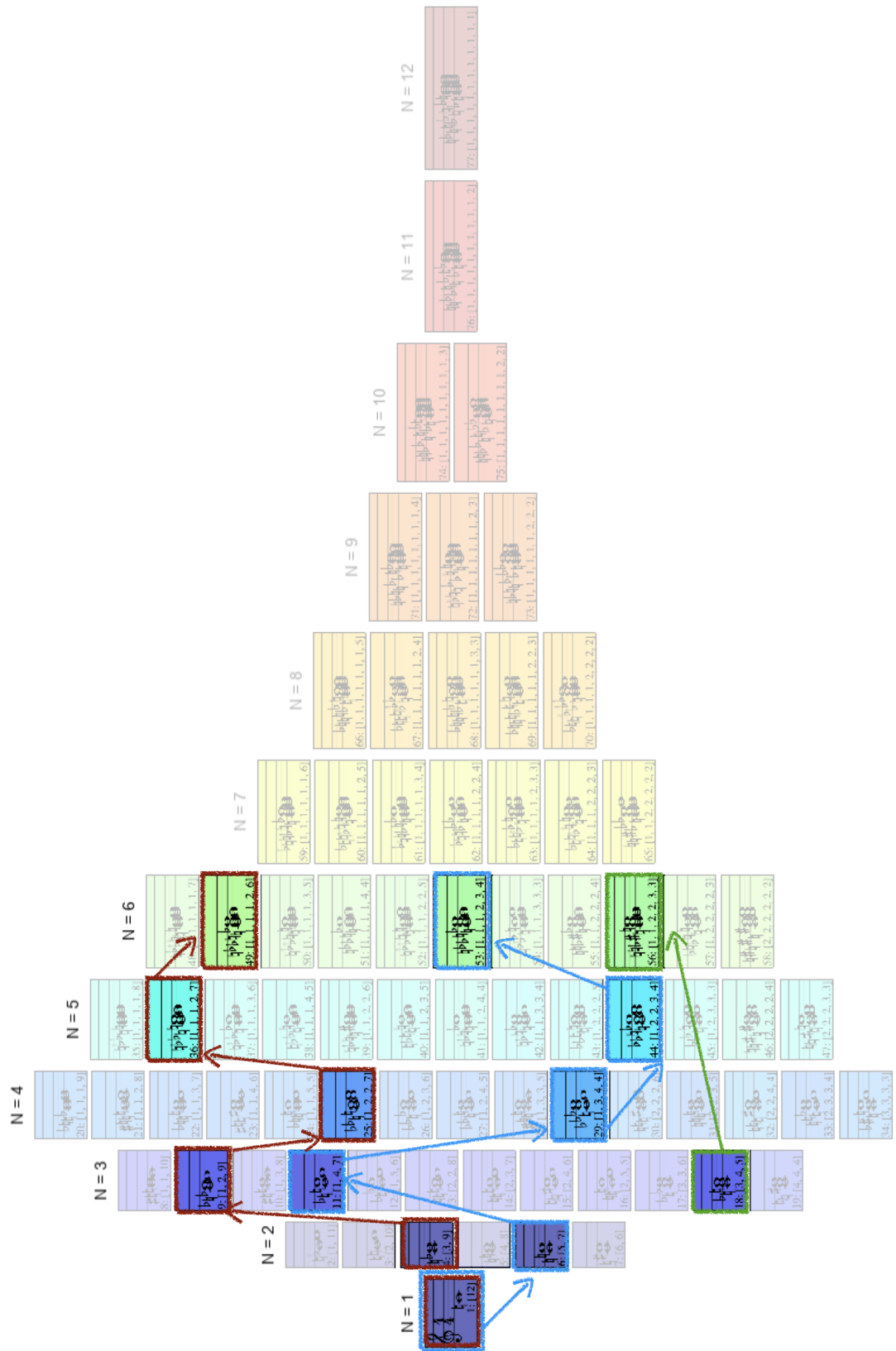


FIGURE 11.6 – Tension des identités des six premières mesures de la *Pièce pour orgue n°2* d'Alfred Schnittke. Cadre rouge : mesures 1-2, bleu : mesures 3-4, vert : mesures 5-6. Paramètres de spectre pour la coloration : $K = 11$, $\sigma = 0.005$ et $decr = 0.5$.

Comparaison d'interprétations Pour valider les résultats précédents, calculons les descripteurs harmoniques statiques sur trois versions différentes de l'œuvre : l'interprétation de Robert Pecksmith¹, celle de Daniel Bruun² et celle d'Oleg Yanchenko³. La figure 11.7 superpose les profils de rugosité, d'harmonicité et de tension des trois interprétations. La concordance, la concordance d'ordre 3 et la concordance totale nécessitent quant à eux d'avoir le spectre des notes séparées, ils ne sont donc pas calculables dans la configuration « enregistrement⁴ », c'est-à-dire lorsqu'on a à disposition seulement le signal global. Comme nous considérons ici de vrais enregistrements et non plus des synthèses à partir de sons échantillonnés, nous faisons le choix, dans un souci de réalisme, de ne pas séparer la partie harmonique de la partie percussive et du bruit⁵, mais de synchroniser le spectre entier.

Nous observons des différences qualitatives entre les versions, dues aux disparités d'instruments, d'acoustique, de prise de son et d'interprétation. Malgré tout, les évolutions sont globalement les mêmes, et viennent corroborer les résultats obtenus à partir des colorations des trajectoires sur la carotte. Sur les deux premiers cycles d'accumulation - jusqu'à la coupure de notes au clavier, la rugosité et la tension augmentent, tandis que l'harmonicité diminue.

La variation d'harmonicité n'est cependant pas uniforme sur le premier cycle dans la version de Yanchenko. On peut expliquer l'augmentation importante d'harmonicité lors de l'ajout de la pédale *ré* par la saillance de cette note dans l'enregistrement - elle ressort plus que dans les deux autres versions. Une grande part de l'énergie de l'accord est alors contenue dans le spectre harmonique de la pédale. L'affichage des hauteurs fondamentales virtuelles⁶ indique d'ailleurs *ré* 1 sur cette verticalité pour les trois versions, ce qui montre que la partie harmonique vient essentiellement de la pédale. À la fin du cycle 2, n'a pédale n'a pas le même effet, en raison du registre plus aigu et du trille. On a cette fois une incohérence des hauteurs fondamentales virtuelles : *si* 2 pour la version de Smith, *mi* 0 pour celle de Bruun, et *la* 0 pour celle de Yanchenko. Cette disparité, qui montre une relative indétermi-

1. Voir la note 2 page 420.

2. Tirée du disque (Windekilde, 2002).

3. Tirée du disque (Yanchenko, 2012).

4. Voir page 163.

5. Voir page 185.

6. Voir partie 6.1.4.

nation de la hauteur fondamentale virtuelle, est symptomatique des faibles valeurs d'harmonicité.

L'ajout de l'accord de *Mib* m sur l'accord de *Do* M a également pour effet de faire augmenter la rugosité et la tension et de faire diminuer l'harmonicité.

Conclusion La comparaison de plusieurs interprétations de l'œuvre à travers les descripteurs harmoniques statiques met en évidence le rôle structurant dans le début de l'œuvre de la rugosité, l'harmonicité et la tension, qui suivent les mêmes schémas sur les cycles d'accumulation.

De plus, la coloration des trajectoires entre identités sur la carotte constitue une bonne approximation des analyses signal, et permet déjà de mettre en valeur le rôle de ces descripteurs.

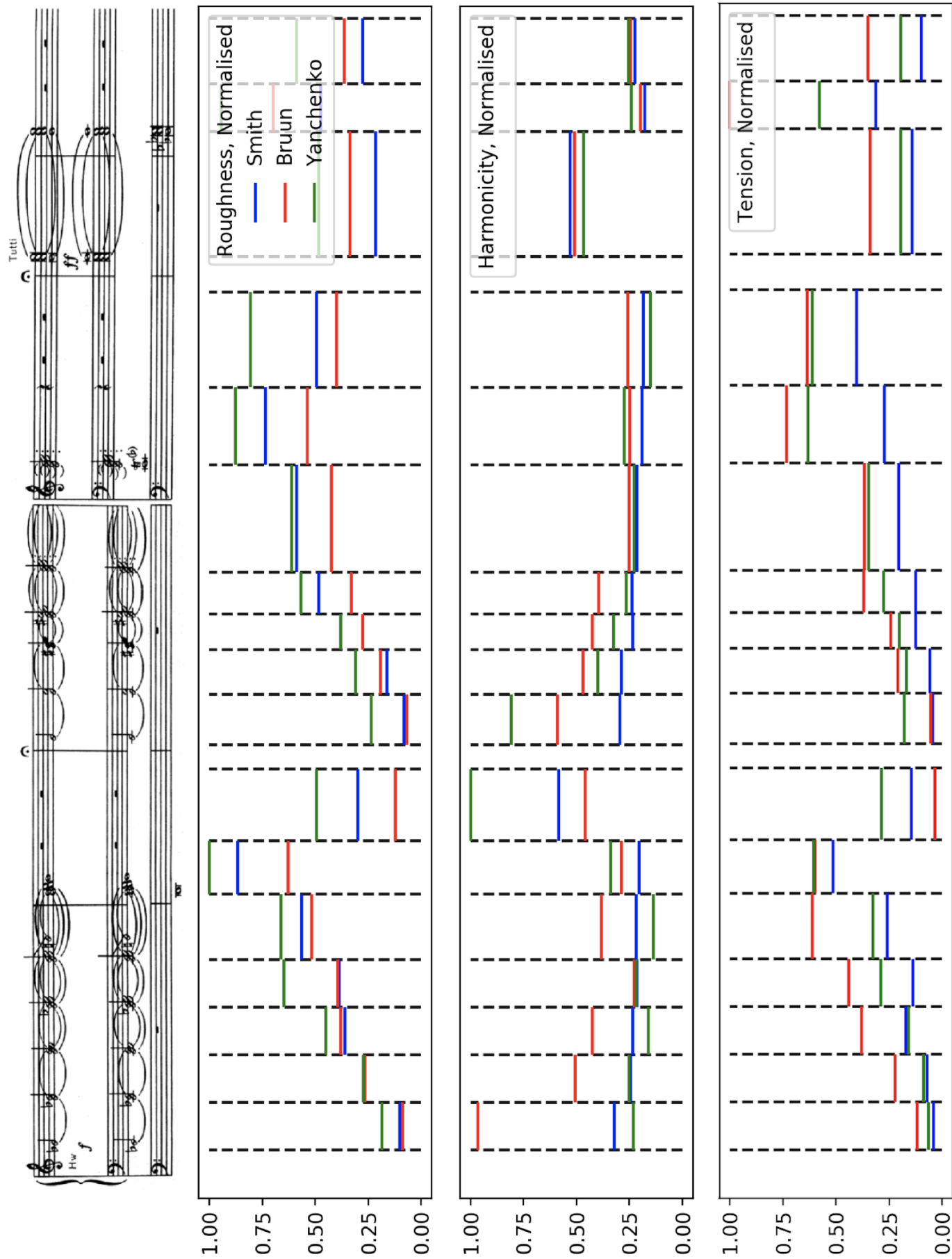


FIGURE 11.7 – Rugosité, harmonicité et tension calculées sur 3 interprétations de la *Pièce pour orgue n°2* de Schnittke. En bleu, la version de Smith, en rouge celle de Bruun et en vert celle de Yanchenko.

Chapitre 12

Almeida Prado - *Cartas Celestes*

I

José Antonio Rezende de Almeida Prado (1943-2010) est un compositeur brésilien, dont l'œuvre prolifique est marquée par le cycle des *Cartas Celestes* constitué de 18 pièces, la plupart pour piano. Leur composition s'étale de 1974 jusqu'à l'année de sa mort, sur une période de 36 ans. Rompant tôt avec le courant musical nationaliste marqué par le folklore brésilien et porté par son professeur le compositeur Mozart Camargo-Guarnieri, Almeida Prado suit les cours de Ligeti et Stockhausen à Darmstadt et part étudier à Paris sous la houlette de Nadia Boulanger et Olivier Messiaen.

La pièce *Cartas Celestes I* est une commande destinée à accompagner des spectacles au Planétarium de São Paulo. Écrite en 1974, soit la même année que *Des Canyons aux Étoiles* d'Olivier Messiaen, elle s'inscrit dans la thématique cosmique, à côté de *Sternklang* (1971) de Stockhausen ou *Le noir de l'étoile* (1990) de Grisey. Le compositeur s'inspire du ciel brésilien des mois d'août et septembre et divise l'œuvre en 21 sections faisant chacune le portrait d'une étoile, d'une constellation ou d'une nébuleuse. La partition de l'œuvre est jointe en annexe 3.

Transtonalisme Les *Cartas Celestes I* occupent une place particulière dans l'œuvre d'Almeida Prado puisqu'elles marquent le début d'un nouveau système de composition, que le musicologue Yulo Brandao qualifiera en 1975 de « *transtonalismo* ». Le compositeur théoriserait après coup ce système dans sa thèse de doctorat dédiée au cycle des *Cartas Celestes* ([Almeida Prado, 1985](#)).

Le transtonalisme est fondé sur la résonance, en particulier la résonance des séries harmoniques supérieures et inférieures¹, qui structurent les aspects tant harmoniques et timbraux que rythmiques. Ce dualisme dans la résonance peut faire songer à Rameau² ou à Riemann³. Almeida Prado, ne se réfère pas à ces théoriciens, son approche étant avant tout fondée sur l'écoute. D'ailleurs, ces résonances l'amènent bien au-delà des accord parfaits et des accords de septième.

Cette primauté de la résonance peut se traduire par l'usage de vocabulaire tonal tel un accord majeur, mais sorti des hiérarchies tonales et de l'aspect fonctionnel du système tonal. Le terme « transtonalisme » doit d'ailleurs être compris en ce sens, en tant que « *présence tonale transfigurée, car libérée des réminiscences des cadences tonales* »⁴, pour reprendre les mots du compositeur. Contrairement à ce que fait Alfred Schnittke dans les mêmes années⁵, le rapport entre tonalité et atonalité ne relève pas d'une situation conflictuelle d'opposition et de frottement stylistique⁶, mais plutôt d'une fusion des écritures.

Pour Guigue (Guigue, 2009), ce système est « *apte à absorber et recycler des constructions non seulement atonales, mais aussi tonales, sérielles, minimalistes ou spectrales* », et « *possède dans la tolérance un de ses principaux atouts* »⁷. Almeida Prado définit lui-même le transtonalisme comme :

l'observation d'harmoniques supérieurs et inférieurs résultant d'une note donnée et l'incorporation de tout ce qui provient des tech-

1. Le terme « série inférieure » employé par Almeida Prado vient de Rameau. Nous avons déjà mis en avant le lien entre la série des harmoniques inférieurs avec la concordance totale, voir page 216 et figure 6.13. Les « séries harmoniques supérieures » sont les séries harmoniques usuelles.

2. Le dualisme chez Rameau consiste à justifier la naturalité de l'accord mineur par l'usage des séries harmoniques inférieures, calquant le raisonnement sur l'accord majeur justifié par la série harmonique supérieure. Rameau justifie la série harmonique inférieure, qui n'a pas la réalité physique ni sonore de la série harmonique supérieure, par un principe de symétrie qu'il considère prescrit dans la nature. Se référer à la partie 2.2.1 pour plus de détails.

3. Riemann parle déjà de résonances « supérieures » et « inférieures » en se référant aux séries harmoniques supérieures et inférieures (Riemann, 1887).

4. Tiré de (Mariz, 1994), cité dans (Ferraz, 2009, p.44).« *Transfigured tonal presence, but free from the reminiscences of tonal cadences* ».

5. Voir l'exemple de la *Pièce pour orgue n°2* traité dans le chapitre précédent. Dans cette pièce, l'accord Do M de la mesure 5 est perturbé par un accord hautement dissonant.

6. Voir (Gaulhiac, 2019).

7. Tiré de (Guigue, 2009).

*niques contemporaines d'écriture, comme le sérialisme et le minimalisme, dans l'utilisation d'amas sonores, plus la richesse rythmique de Messiaen et Villa-Lobos*¹.

L'écriture d'Almeida Prado a de nombreux points communs avec l'écriture de Messiaen, son professeur dans ses années parisiennes². Outre Messiaen et Villa-Lobos, Almeida Prado cite l'école spectrale française, allant jusqu'à qualifier son écriture de « spectrale ». Le transtonalisme partage avec le spectralisme de Grisey une démarche fondée sur la décomposition spectrale du son, une exploration des propriétés acoustiques et perceptives du son ainsi qu'une écriture où le son structure organiquement l'œuvre musicale. Les deux approches, qui étaient concomitantes mais s'ignoraient mutuellement³, diffèrent en ce que le transtonalisme est moins rigoureux que le spectralisme. Guigue (Guigue, 2009) écrit à ce propos :

Cette approche [...] dénote des préoccupations qui côtoient celles de l'école « spectrale », s'en démarque cependant clairement dans la mesure où le spectre n'est pas ausculté de manière exhaustive par des moyens « scientifiques » pour constituer la matrice d'où serait dérivée la totalité des aspects de la structure, en particulier son élaboration dans le temps.

Comme le fait remarquer Ferraz (Ferraz, 2009, p.45), Almeida Prado, contrairement à Grisey, n'utilise pas d'ordinateur pour composer, que ce soit pour visualiser des spectres ou pour utiliser des techniques issues de la musique électroacoustique. Son approche de la résonance est avant tout sensible et poétique, liée à sa connaissance intime du piano.

En écrivant pour cet instrument, Almeida Prado approxime les harmoniques supérieurs ou inférieurs d'une note par les demi-tons de l'échelle chro-

1. Tiré de (Mariz, 1994), cité dans (Ferraz, 2009, p.44). « *the observation of ascending and descending harmonics resulting from a given note and the incorporation of everything that could be obtained from contemporary techniques, such as serialism and minimalism, into the utilization of sonorous blurs (clusters), plus all the rhythmical richness of Messiaen and Villa-Lobos.* » Traduction de l'auteur.

2. Voir (Morales Taffarello, 2010), où les techniques de compositions des deux compositeurs sont mises en parallèle.

3. Voir (Ferraz, 2009, p.45).

matique¹. Malgré l'inharmonicité des cordes due à leur raideur² et la discrétisation des fréquences sur le clavier, le choix du piano pour l'écriture des six premières *Cartas Celestes* est motivé par les propriétés résonantes de l'instrument. En effet, les cordes du piano vibrent par sympathie. Ce phénomène invite à appréhender la résonance sous l'angle des séries harmoniques supérieures ou inférieures. De plus, les propriétés rayonnantes de la table d'harmonie et de la structure permet d'atteindre une forte puissance sonore. La pédale *forte*, qui soulève les étouffoirs, joue un rôle important dans les *Cartas Celestes I*, où elle est notée et souvent tenue pendant des sections entières, comme c'est le cas dans *Scorpio*. Elle permet le mélange des sonorités pour créer des nuages de résonance, appelés « *spectres de résonance* ».

Système de résonances organisées Almeida Prado, dans sa thèse (Almeida Prado, 1985), consolide d'un point de vue théorique son approche de la résonance en distinguant 4 zones, qui correspondent à différents degrés de « *densité d'espaces sonores* », qu'il nomme également « *zones de perception de résonance* » (Almeida Prado, 1985, p.559). Ces zones décrivent non pas des accords isolés, mais des situations de jeu. Il écrit à leur propos :

*J'ai essayé de prouver qu'il était possible, à travers l'usage rationnel et organisé des résonances, de transmettre à l'auditeur une émotion d'intense vibration, le mettant face-à-face avec un Espace Sonore d'un nouveau type, non plus structuré en mélodies et en rythmes, mais matérialisé par des zones denses ou transparentes de masses sonores*³.

Ces quatre zones sont les suivantes :

1. Zone de résonance explicite : c'est lorsqu'une série harmonique supérieure ou inférieure, dont la note fondamentale est présente, structure l'écriture du passage.

1. Se référer aux séries harmoniques supérieures et inférieures qu'il présente dans sa thèse aux pages 559 et 560. Celles-ci sont constituées de 16 partiels et approximées au demi-ton, ce qui peut paraître aberrant du point de vue acoustique, étant donné que le partiel 11 est situé presque exactement entre deux demi-tons (à 1 *cent* du quart-de-ton).

2. Se référer à (Young, 1952) pour plus de détails.

3. « *Tentei provar ser possível, através do uso racional e organizado das ressonâncias, passar ao ouvinte uma emoção de intensa vibração, colocando o face a uma nova proposta de Espaço Sonoro, não mais comprometido com melodias ou ritmos, mas materializado por zonas espessas ou transparentes de massas sonoras.* » (Almeida Prado, 1985, p.29), traduction de l'auteur.

2. Zone de résonance implicite : elle est caractérisée par la présence d'une série harmonique supérieure ou inférieure , éventuellement incomplète, dans un contexte atonal. Les notes de la série ne sont plus au premier plan mais ont un rôle d'« *indices* » et d'« *insinuations* ».
3. Zone de résonance multiple : elle est caractérisée par une accumulation de résonances créant un nuage de résonance, appelé « *spectre de résonances* ». Le compositeur la décrit ainsi :

*Lorsque l'emploi d'accords simultanés ou successifs, constitués de résonances mélangées, crée un tourbillon de résonances, les rendant presque impossibles à distinguer à l'oreille. Ce procédé d'accumulation de notes dégage une puissance sonore incroyable due à un battement désordonné des vibrations simultanées*¹.

4. Zone de non-résonance : elles proviennent d'une intention délibérée du compositeur de minimiser les résonances, au sens des séries harmoniques supérieures ou inférieures. Elles créent « *une zone opaque, neutre, un élément vital de contraste* »².

Le compositeur utilise des extraits des *Cartas Celestes I* pour illustrer ces quatre zones. On se référera à sa thèse (Almeida Prado, 1985, p.559-568) pour plus de détail.

Motivation du choix de l'œuvre et de l'analyse L'écriture de la résonance d'Almeida Prado explore les interactions simultanées ou successives entre les spectres. Pour cette raison, elle invite assez naturellement à être auscultée sous l'angle des descripteurs harmoniques. Comme le note Ferraz (Ferraz, 2009) :

Les différents harmoniques sont amplifiés ou atténués par plusieurs résonances à l'intérieur du piano, de sorte que certains harmoniques peuvent être mis en valeur tandis que d'autres peuvent

1. « *Quando o uso de acordes simultâneos ou sequenciais, constituídos de ressonâncias misturadas, criam um turbilhão de ressonâncias, - tornando quase impossível a distinção pelo ouvido. Este processo acumulativo de notas é de incrível poder sonoro devido ao batimento desordenado das vibrações simultâneas.* » (Almeida Prado, 1985, p.565), traduction de l'auteur.

2. « *zona de opacidade, neutralidade, elemento também vital de contraste com os outros.* » (Almeida Prado, 1985, p.566), traduction de l'auteur.

*être supprimés, selon la manière dont le compositeur les combine*¹.

La combinaison des harmoniques peut être décrite par les descripteurs comme la concordance ou la rugosité, statiques et dynamiques. De plus, l'harmonicité paraît particulièrement pertinente pour mettre en avant le rôle structurant des séries harmoniques supérieures.

Outre ces raisons, le choix de se pencher sur les *Cartas Celestes I* dans le présent travail est motivé par la présence d'un catalogue d'accords faisant office de vocabulaire harmonique de l'œuvre.

Le but de la présente analyse est d'éclairer de manière rigoureuse une écriture qui, si elle puise sa substance à la source de considérations acoustiques et perceptives, se dit explicitement non scientifique car avant tout poétique et fondée sur l'écoute et l'imagination. Nous nous pencherons dans un premier temps sur le catalogue d'accord, avant d'analyser un exemple canonique de zone à résonance multiple, à savoir la section *Scorpio*. Les résultats seront mis en parallèle avec les écrits du compositeur.

L'analyse se fera à partir d'enregistrements du pianiste César Birschner, réalisés pour l'occasion. Ce choix, visant à rendre compte au maximum des phénomènes résonants ayant inspiré les accords du catalogue ainsi que l'écriture de la pièce, est motivé par le rôle crucial que joue le piano dans le transtonalisme d'Almeida Prado. En employant des sons synthétisés d'orgue comme dans le chapitre 10, nous serions passés à côté de la décroissance des partiels, des effet de pédale, ainsi que des attaques, lesquelles font partie intégrante, comme nous allons le voir, du phénomène de résonance. En utilisant des sons échantillonnés de piano, nous serions passés à côté des phénomène acoustique de couplage comme la résonance des cordes par sympathie.

1. « *The different harmonics are amplified or attenuated by various resonances inside the piano in such a way that certain harmonics may be greatly emphasized while others may be deeply suppressed according to how the composer combines them* ». (Ferraz, 2009, p.41), traduction de l'auteur.

12.1 Analyse statique du catalogue d'accord

12.1.1 Un alphabet d'accords

Le catalogue d'accords présenté au début de la partition des *Cartas Celestes I* est reproduit à la page suivante. Il est constitué de 24 accords, chacun représenté par une lettre grecque. Les accords sont formés d'un nombre variable de notes, allant de 4 (accords α et ω) à 18 (accord ξ), et couvrent toute l'étendue du clavier. Cet « *alphabet d'accords* », pour reprendre les mots du compositeur, fait office de « *matériel sonore fixe* » qui sert d'« *élément unificateur de l'œuvre* ». Unificateur, il l'est aussi du cycle des 6 premières *Cartas Celestes*. Dans la première, chacune des constellations, lesquelles constituent le cœur de l'œuvre, est construite uniquement sur ces accords-alphabet.

α



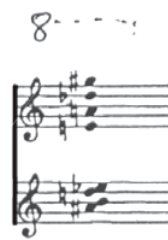
β



γ



δ



ε



ζ



η



θ



ι



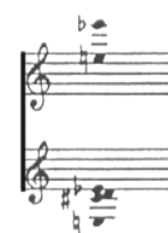
κ



λ



μ



ν



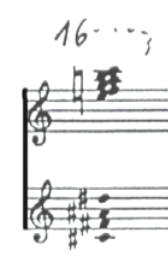
ξ



ο



π



ρ



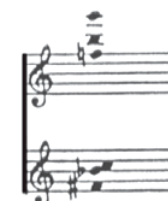
σ



τ



ι2



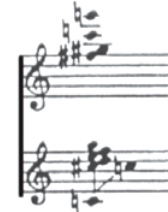
φ



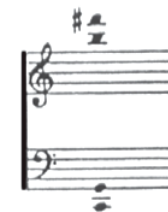
χ



ψ



ω



Almeida Prado dit avoir choisi les accords pour leurs propriétés résonnantes, ainsi que par « *la fantaisie et l'intuition*¹ ». Dans sa thèse², il présente chaque accord en matière de structure, de propriétés acoustiques, du rôle joué dans les *Cartas Celestes I*, ainsi qu'en termes métaphoriques, qui concernent avant tout le cosmos et la luminosité. De l'accord ϵ , il écrit par exemple, en ajoutant la figure ci-dessous :

*Cet accord est construit sur des résonances d'octaves, qui produisent un battement rapide et intense du demi-ton si-do, créant un timbre d'une intense luminosité, comme l'étincelle bleu-violette d'une étoile. De par cette qualité, il est utilisé pour produire une lumière intense*³.



FIGURE 12.1 – Accord ϵ , tiré de (Almeida Prado, 1985, p.13)

Dans les *Cartas Celestes I*, les accords sont présentés, à quelques exceptions près, dans une disposition inchangée. Ce sont des entités fixes, qui ont toujours le même registre, les mêmes doublures, et ne sont ni transposables ni renversables. Ils varient par contre par la durée, la nuance, le toucher et la présentation. Ils ne sont pas toujours joués d'un bloc, mais parfois arpégés. Cette fixité des positions apparaît comme un avantage de taille pour leur analyse, car il se prête alors au calcul direct des descripteurs harmoniques à partir d'un enregistrement. Cette démarche analytique se situe à l'opposé des artifices du chapitre 9 consistant à calculer les descripteurs sur les classes normales en passant par les spectres de Shepard. Les propriétés résonnantes de chaque accord sont ici intimement liées à sa disposition, qui ne peut être

1. (Almeida Prado, 1985, p.7)

2. Voir (Almeida Prado, 1985, p.7-25).

3. « ϵ e feito de ressonâncias de oitavas, produzindo uma vibração rápida e intensa do semitom si-do, criando um timbre de intensa luminosidade, como o brilho azul-violeta de uma estrela. Por essa qualidade, ele é usado sempre com muita intensão de brilho e luz ». (Almeida Prado, 1985, p.13), traduction de l'auteur.

modifiée sous peine de modifier l'accord. Cette approche a ceci de commun avec le spectralisme qu'elle considère les sons « *non comme des objets morts, facilement et arbitrairement permutable, mais comme des objets vivants avec une naissance, une durée et une mort*¹. »

Aussi, plusieurs accords distincts du catalogue partagent-ils la même classe normale, comme le montre la figure 12.2.

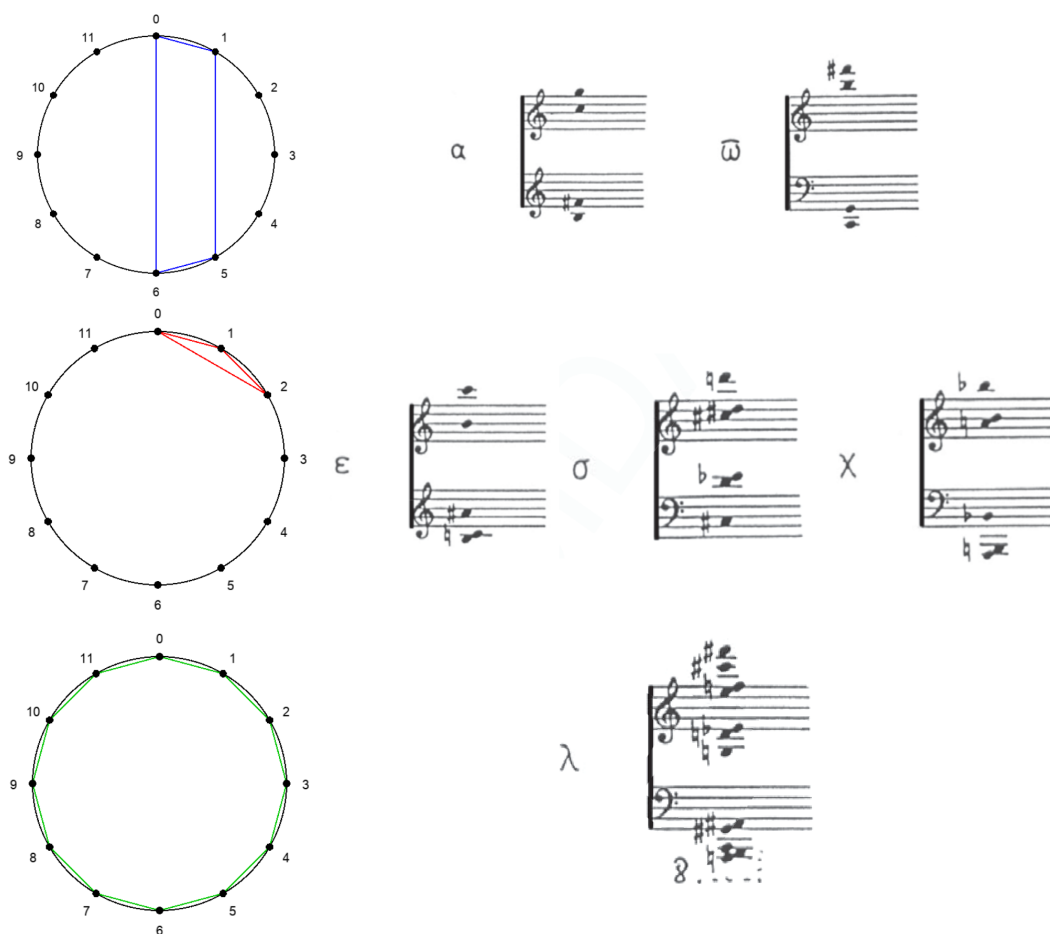


FIGURE 12.2 – Quelques accords du catalogue, regroupés par classes normales.

1. « *not as dead objects that you can easily and arbitrarily permute in all directions, but as being like living objects with a birth, lifetime and death* ». Tiré d'une interview par David Bundler (18 juin 1996), disponible à l'adresse www.angelfire.com/music2/davidbundler/grisey.html, traduction de l'auteur.

12.1.2 Cartographies des accords

En raison de la fixité des accords du catalogue, ceux-ci se prêtent volontiers à une étude comparative par les descripteurs harmoniques statiques. En outre, la normalisation par l'énergie et par le nombre de notes¹ permettent de comparer de manière non triviale des accords de cardinalités différentes, et minimisent les écarts liés aux différences de nuances. On peut faire l'hypothèse, du fait du rôle crucial des séries harmoniques dans l'écriture d'Almeida Prado, que l'harmonicité fournit un outil particulièrement pertinent pour comprendre le choix des accords et leur utilisation dans la pièce.

Méthodologie Les accords du catalogue, ainsi que l'extrait dont il sera question plus tard, ont été enregistrés par le pianiste brésilien César Birschner sur un piano Yamaha C2X. Un appareil *Tascam DR-09*, monté sur pied, et placé à deux mètres derrière le piano, a été utilisé pour la captation. Le pianiste a joué chaque accord pendant 6 secondes, avec des pauses entre chaque. La consigne lui avait été donnée de jouer tous les accords avec un toucher et une nuance les plus homogènes possibles, malgré les disparités de registres et de cardinalité. Les accords qui ne peuvent être joués d'un bloc - comme l'accord λ , ont été arpégés en deux temps.

La fenêtre de synchronisation² a été fixée à une longueur donnée pour tous les accords, à savoir 5 secondes, à partir de l'attaque comprise. Étant donné que l'attaque, pour Almeida Prado, fait partie intégrante du phénomène de résonance, comme l'atteste l'emploi fréquent de trémolos et d'accords répétés *forte*³, nous avons fait le choix de ne pas supprimer la partie percussive du signal⁴. L'aspect percussif de l'instrument, la décroissance de l'enveloppe et le fait que les partiels aigus disparaissent plus vite que les partiels graves induisent une instabilité temporelle du contenu spectral sur les verticalités, qui fait sortir du cadre établi dans la partie II⁵. Les résultats doivent ainsi être considérés avec précaution ; nous verrons cependant qu'ils sont suffisamment cohérents pour pouvoir être exploités.

1. Voir parties 7.2 et 7.3.

2. Voir partie 5.3.2.

3. Par exemple les sections *Hércules* ou *Scorpio*.

4. L'algorithme de décomposition en partie harmonique et partie percussive est décrit page 185.

5. Voir page 5.3.2 et partie 7.1.2.

Rugosité - harmonicité Étant donné que nous sommes ici dans une configuration d'enregistrement ¹, les seuls descripteurs statiques calculables parmi ceux présentés dans ce travail sont l'harmonicité, la rugosité et la tension. Nous mettons de côté la tension, étant donné qu'aucune considération en rapport ne semble apparaître dans les écrits du compositeur. La figure 12.3 montre la cartographie des accords du catalogue dans l'espace (rugosité, harmonicité). Les valeurs des descripteurs ont été divisées par leurs valeurs maximales. Pour le calcul de l'harmonicité, le spectre-type a été pris avec $\kappa = 15$ partiels ². Ce choix est motivé par la série harmonique supérieure représentée par Almeida Prado dans sa thèse ³, qui couvre quatre octaves. Il serait pertinent de tester la stabilité des résultats avec d'autres valeurs de κ , mais nous n'entreprendrons pas ce travail ici.

Tous les accords du catalogue occupent des positions distinctes, et certains se singularisent par leur position excentrée. L'accord maximisant l'harmonicité est ω , formé de deux quintes dans deux registres différents du clavier. Sa hauteur fondamentale est *do* 2, ce qui montre que l'harmonicité est principalement due à la quinte *do-sol* dans le grave. L'accord α , renversement de ω possédant également deux quintes, est un des accords avec la plus basse harmonicité, ce qui montre la grande sensibilité à la disposition. L'accord ν a également une harmonicité importante, en raison essentiellement de la quarte à la basse, puisque sa note fondamentale virtuelle est *fa*♯ 1. La rugosité est quant à elle maximisée sur l'accord λ , accord du total chromatique, autrement dit constitué de toutes les hauteurs de l'échelle chromatique ⁴. S'étendant sur près de cinq octaves et demie, il est décrit par Almeida Prado comme « *somptueux* », « *hautement dramatique* », « *comme un grand spectre ultra-violet* » capable de créer « *une infinité cosmique* ⁵ ». La richesse des qualificatifs est à l'image de sa singularité vis-à-vis de la rugosité.

-
1. Les différentes configurations sont présentées dans la partie 5.1.2.
 2. Pour les rappels techniques, voir page 219.
 3. Voir (Almeida Prado, 1985, p.560), lorsque le compositeur présente les zones de résonance explicites.
 4. Voir figure 12.2.
 5. (Almeida Prado, 1985, p.16).

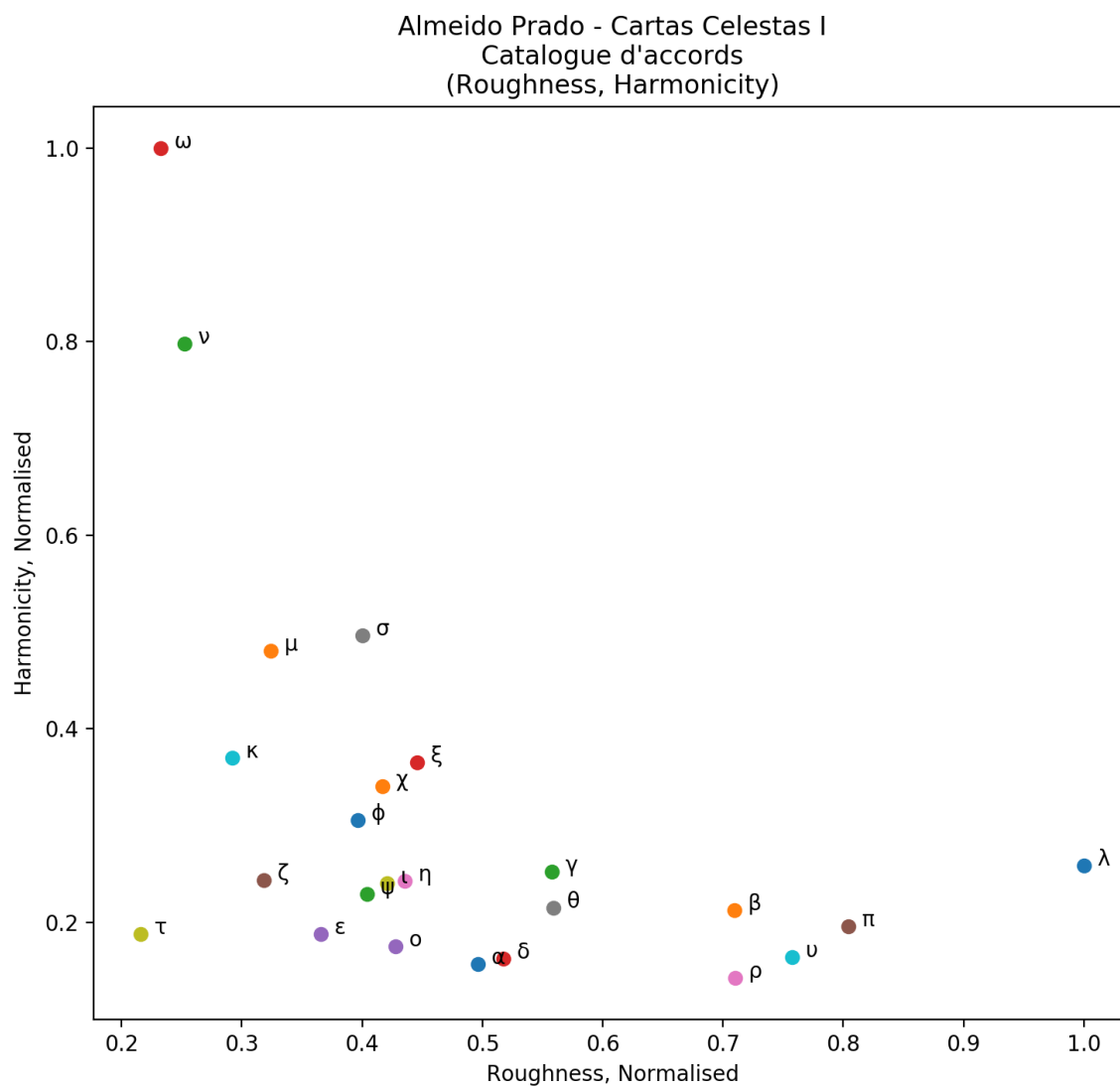


FIGURE 12.3 – *Cartas Celestes I* - Cartographie du catalogue d'accord dans l'espace (rugosité, harmonicité).

Série harmonique inférieure et harmonicité inverse Le « *système de résonances organisées*¹ » d’Almeida Prado est construit sur les séries harmoniques supérieures et inférieures, toutes deux placées sur le même plan et jouant le même rôle dans la perception de la résonance, selon les propos du compositeur. Qu’en est-il réellement ? Il serait instructif d’avoir des outils à disposition pour quantifier leur importance relative, ne serait-ce que dans la constitution du catalogue d’accords.

L’harmonicité statique détecte le degré de ressemblance d’un spectre donné à un spectre harmonique supérieur. Il serait souhaitable d’avoir un descripteur équivalent pour les spectres harmoniques inférieurs.

Nous avons vu que la concordance totale met en avant les accords dont les notes appartiennent simultanément à une même série harmonique inférieure². Or, Almeida Prado, fidèle à son désir de « *concilier les expériences atonales avec l’usage rationnel des harmoniques inférieurs et supérieurs*³ », n’utilise pas, dans son catalogue, d’accord dont les notes forment une série harmonique — supérieure ou inférieure — « pure », c’est-à-dire dont toutes les notes sont incluses dans la série. Les séries harmoniques présentes dans les accords du catalogue sont entourées de notes étrangères à la série. Dans la description des accords, Prado évoque notamment les résonances d’octaves — σ ou ξ — pour des accords contenant plusieurs octaves, et donc des fragments de séries harmoniques distinctes. En contexte de jeu, les zones de résonance implicite⁴ font entendre des fragments de séries harmoniques dans un contexte atonal, donc superposées à des notes qui n’en font pas partie. Dans l’extrait de la section *Via-Lactea*, dont le compositeur se sert pour illustrer les zones de résonance explicite⁵, les notes sont toutes issues de la série harmonique supérieure de fondamentale *do* 2, jusqu’au cluster dans le grave, qui libère d’autres résonances en faisant sortir de la résonance de cette série harmonique. De ce fait, la concordance totale, si elle peut permettre localement de mettre en avant l’appartenance à une même série harmonique inférieure, semble être un point de vue peu fécond pour différencier les accords du catalogue.

1. Voir page 432.

2. Voir figures 6.12 et 6.13 page 217.

3. « *colocar juntos as experiências atonais com o uso racional dos Harmônicos Superiores e Inferiores* », (Almeida Prado, 1985, p.559), traduction de l’auteur.

4. Voir page 432.

5. Voir (Almeida Prado, 1985, p.560).

La concordance d'ordre 3, quant à elle, permettrait de mettre en avant les accords contenant des triplets de notes issues d'une même série harmonique inférieure. Cela paraît déjà plus pertinent que la concordance totale pour étudier les accords du catalogue, mais l'arbitraire du chiffre 3, lequel ne ressort d'aucune façon dans les écrits du compositeur, nous dissuade de l'employer. De plus, le calcul de la concordance totale ou de la concordance d'ordre 3 nécessiterait d'enregistrer chaque note séparément, ce qui éliminerait les effets de couplage dans l'instrument.

Plutôt que de se ramener à un descripteur déjà présenté, construisons le pendant de l'harmonicité pour les séries harmoniques inférieures, qui mesure le degré de ressemblance d'un spectre donné avec un spectre harmonique inférieur. Ce descripteur, appelé « harmonicité inverse ¹ », est défini et implémenté de la même manière que l'harmonicité ², en remplaçant le spectre-patron harmonique par un spectre-patron harmonique inférieur. Là-aussi, nous faisons le choix d'un spectre-patron à 15 partiels, couvrant 4 octaves, pour être cohérent avec la série harmonique inverse présentée par Almeida Prado dans sa thèse ³. Comme l'harmonicité, l'harmonicité inverse fournit également une fréquence fondamentale virtuelle, cette fois appelée « fréquence fondamentale virtuelle inverse », qui correspond à la fréquence de laquelle est issue la série harmonique inverse la plus ressemblante avec le spectre donné.

La figure 12.5 montre la cartographie des accords du catalogue dans l'espace (harmonicité, harmonicité inverse). Les valeurs de chacun des descripteurs ont été normalisées linéairement de 0 à 1. Elle est à lire en parallèle de la figure 12.4, qui expose les fréquences fondamentales virtuelles et fréquences fondamentales virtuelles inverses des accords maximisant l'harmonicité et l'harmonicité inverse.

Deux observations s'imposent. Premièrement, il apparaît que l'harmonicité et l'harmonicité inverse sont fortement corrélées sur les accords du catalogue, avec un coefficient de corrélation positif égal à $r = 0.84$. Comprendre si cette corrélation est propre à ce catalogue ou est une propriété des descripteurs, voir du timbre, nécessiterait d'étudier un nombre important d'accords

1. La série harmonique inférieure est obtenue par inversion d'une série harmonique supérieure, au sens défini page 62, d'où le nom d'« harmonicité inverse ».

2. Pour l'implémentation de l'harmonicité, se référer à la partie 6.1.4.

3. (Almeida Prado, 1985, p.561).

généérés aléatoirement, ce que nous n’aborderons pas ici. Ensuite, une forte symétrie est visible entre les deux descripteurs, matérialisée sur la figure par l’axe en pointillés, et indépendante de leur corrélation. Cette symétrie n’est absolument pas triviale, n’étant pas une propriété des descripteurs, mais bien une propriété des accords constituant le catalogue. Elle permet de répondre par l’affirmative à la problématique ayant motivé l’introduction de l’harmonicité inverse : les séries supérieures et inférieures jouent des rôles comparables dans la constitution du catalogue d’accords.

Observons maintenant de plus près la disposition des accords. L’accord maximisant l’harmonicité inverse est ν , qui a la propriété de contenir l’accord mineur $do\sharp-fa\sharp-la$ ¹. La hauteur fondamentale virtuelle inverse de ω est le *sol* situé sous la quinte $si-fa\sharp$, ce qui montre que sa forte harmonicité inverse est due à la quinte *do-sol* dans le grave. Notons que pour les trois accords maximisant l’harmonicité et l’harmonicité inverse, ω , ν et σ , la hauteur fondamentale virtuelle inverse est située à une quinte ascendante plus quelques octaves (respectivement 3, 4 et 4 octaves) de la hauteur fondamentale virtuelle. Notons également que les accords ω et ν partagent un même rôle discursif dans le *Cartas Celestes I*, celui d’accords de transition entre parties ou sections. Le compositeur l’écrit explicitement à propos de ω ² ; quant à ν , il « *punctue*³ » chaque cycle de *Scorpio*.

1. Nous avons vu dans la partie 6.1.3 que les triades mineures en position resserrée font partie d’une même série harmonique inverse.

2. (Almeida Prado, 1985, p.24).

3. Tiré de (Ferraz, 2009, p.68)

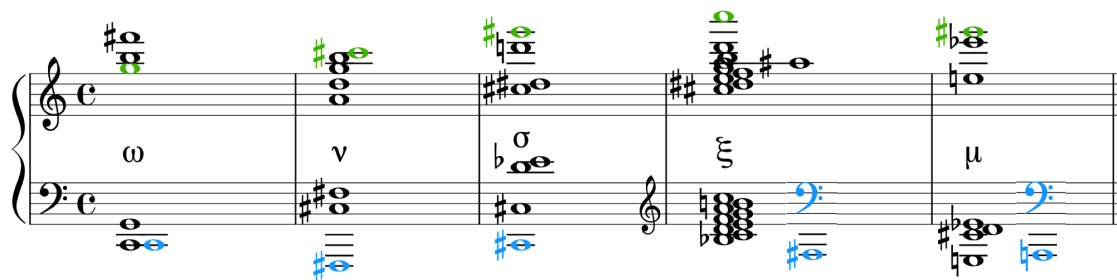


FIGURE 12.4 – *Cartas Celestes I* - Hauteurs fondamentales virtuelles (en bleu) et hauteurs fondamentales virtuelles inverses (en vert) des accords du catalogue maximisant l'harmonicité et l'harmonicité inverse. Les spectrepatron de l'harmonicité et de l'harmonicité inverse ont tout deux $\kappa = 15$ partiels.

Almeido Prado - Cartas Celestas I
Catalogue d'accords
(Harmonicity, HarmonicityInv)

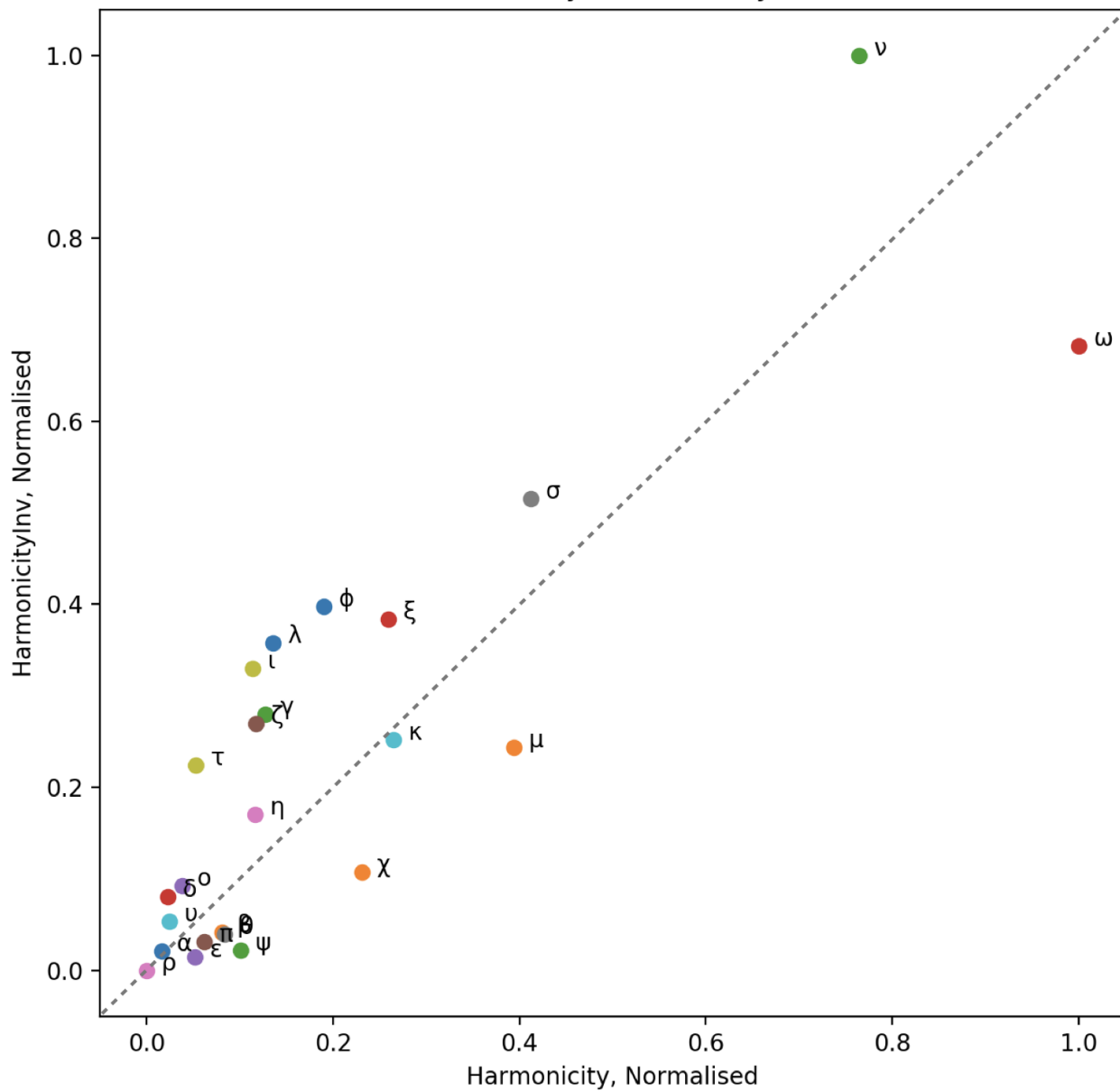


FIGURE 12.5 – *Cartas Celestes I* - Cartographie du catalogue d'accord dans l'espace (harmonicité, harmonicité inverse). En pointillés, la ligne de pseudo-symétrie qui montre l'égalité des rôles joués par les séries harmoniques supérieures et inférieures dans la constitution du catalogue d'accords.

Les images 12.3 et 12.5 peuvent être condensées en une cartographie dans un espace de descripteurs à trois dimensions, comme sur la figure 12.6. Celle-ci est montrée à titre de curiosité, étant donné que de telles représentations sont peu exploitables sur du papier.

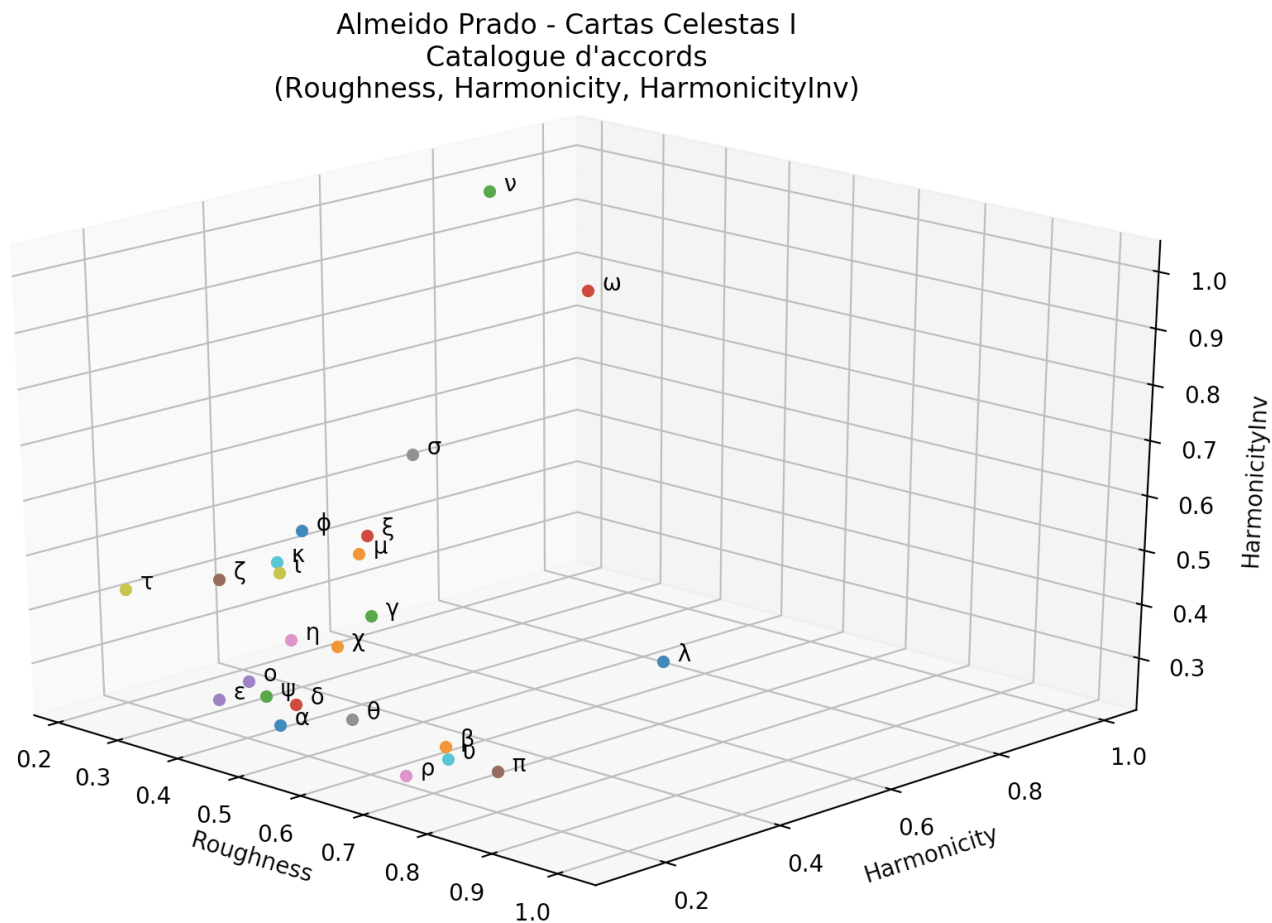


FIGURE 12.6 – *Cartas Celestes I* - Cartographie du catalogue d'accord dans l'espace (rugosité, harmonicité, harmonicité inverse).

12.2 Analyse de *Scorpio*

Penchons-nous maintenant sur l'utilisation du « vocabulaire-alphabet » en situation de jeu, avec la section *Scorpio*, dernière des constellations des *Cartas Celestes I*. Cette section, véritable explosion d'accords martelés *fortissimo* et répétés à grande vitesse, est éprouvante pour l'interprète. Indiqué *lumineux* et *fulgurant*, *Scorpio* se veut une « louange à la virtuosité de la lumière », qui prend la forme d'une « danse féroce des étoiles avec une agressivité fantastique¹ ». La pédale tenue participe du caractère dense et massif, et mélange les sonorités pour créer un spectre de résonances complexe. À ce titre, *Scorpio* est un exemple type de *zone de résonance multiple*².

La section enchaîne 4 cycles de 15 accords tirés du catalogue, dont l'ordre est présenté figure 12.7. Les accords notés avec un *prime* subissent des modifications par rapport à leur version originelle de l'alphabet : π voit sa main gauche transposée d'une octave vers le haut ; le *mi* de σ , de b , devient $\natural$, tandis qu'un *do* $\natural$ aigu est rajouté à la main droite ; le *la* $\natural$ de ζ devient *la* $\sharp$; η perd son *si* ; et χ est éclaté sur le registres extrême du clavier, avec sa main gauche transposée d'une octave vers le grave, et sa main droite de deux octaves dans l'aigu. De telles modifications des accord-alphabet restent l'exception : outre *Scorpio*, on en trouve dans deux autres constellations, *Hércules* et *Lyra*. Les raisons acoustiques de ces modifications mériteraient d'être creusées, notamment par le biais des descripteurs harmoniques, mais nous nous attacherons ici à d'autres aspects de l'analyse.

Le nombre de répétitions varie pour chaque accord, allant sur le premier cycle de 2 (pour ι) à 39 (pour λ). De cycle en cycle ce nombre diminue pour chaque accord, créant par-là une compression rythmique. La pédale étant tenue sur chaque cycle, cette compression rythmique a une incidence directe sur l'harmonie puisqu'elle densifie le spectre de résonance.

Analyse Nous proposons une analyse en deux temps de *Scorpio*. Tout d'abord, en tirant parti des cartographies des *accord-alphabet*³, nous montrons comment persistent des résidus cadentiels dans une écriture qui se veut pourtant « libérée des réminiscences des cadences tonales ». Nous effectuons

1. « *elogio à alta velocidade, à virtuosidade da luminosidade [...] dança feroz das estrelas, com uma agressividade fantástica* », (Almeida Prado, 1985, p.41)

2. Voir page 432.

3. Voir partie 12.1.2.

ensuite une analyse dynamique de *Scorpio* permettant de comprendre l'incidence de la compression rythmique sur la résonance.

Scorpio CONSTELAÇÃO III

π' δ β

simili luminoso, fulgurante simili

σ' α ι ϵ μ ζ'

η' θ ι χ'

λ ν 16

* ped.

The image shows a musical score for a piece titled "Scorpio CONSTELAÇÃO III". The score is written for piano, with a tempo of 110 and a 16-measure phrase. The music is in a key with two sharps (F# and C#) and a 16-measure phrase. The score is divided into four systems, each with a treble and bass staff. Above the notes, Greek letters are placed in red boxes: π' , δ , β in the first system; σ' , α , ι , ϵ , μ , ζ' in the second system; η' , θ , ι , χ' in the third system; and λ , ν in the fourth system. The score includes dynamic markings like "simili", "luminoso, fulgurante", and "ped." (pedal). The tempo is marked as 110, and the phrase length is 16 measures.

FIGURE 12.7 – Premier cycle de *Scorpio*

12.2.1 Résidus cadentiels

De par la densité de son l'écriture, la dernière constellation des *Cartas Celestes I* constitue un *climax* de la pièce, tant des points de vue de la structure globale, de l'expression que de la virtuosité pour l'interprète. L'insistance à ce moment-là sur le même enchaînement de quinze accords se répétant quatre fois donne de bonnes raisons de penser que cet enchaînement est loin d'être anodin, qu'il joue un rôle particulier dans l'œuvre. Cette hypothèse est renforcée par les écarts par rapport au catalogue¹, symptomatiques d'un choix raisonné et intentionné, car exceptionnels dans la pièce. Si le compositeur tient tant à cet enchaînement d'accords, il ne s'en est pas expliqué dans ses écrits, indiquant seulement qu'il oblige le pianiste à « *changer continument de registres*² ».

Pour tenter de comprendre le choix des accords de *Scorpio*, regardons du côté des cartographies des accords du catalogue³. Nous avons vu que les accords ω , λ et ν se distinguent en maximisant respectivement l'harmonicité, la rugosité et l'harmonicité inverse. ω est absent de *Scorpio*, mais λ et ν forment les deux derniers accords du cycle. La transition de λ vers ν , qui constitue la résolution du cycle d'accords, est indiquée par une flèche rouge sur la cartographie des accords du catalogue dans l'espace (rugosité, harmonicité), sur la figure 12.8. Elle correspond à une diminution de la rugosité en même temps qu'une augmentation de l'harmonicité, et fait passer de l'accord le plus rugueux à l'un des accords les plus harmoniques et les moins rugueux du catalogue.

Cette évolution de l'harmonicité et de la rugosité est typique des résolutions des cadences parfaites tonales. Pour s'en convaincre, calculons ces descripteurs sur une cadence parfaite en *Do* majeur, synthétisée avec des sons de piano du logiciel *Musescore*⁴. La figure 12.8 montre l'évolution des descripteurs sous la partition (en haut) ainsi que la trajectoire induite dans l'espace (rugosité, harmonicité) (en bas à droite). L'accord de septième de dominante (n°4) est le plus rugueux des accords de la cadence, et l'accord de tonique de la résolution (n°5) l'un des plus harmoniques et des moins

1. Ces écarts correspondent aux accords notés avec un *prime* sur la figure 12.7.

2. « *uma contínua mudança de registros* », (Almeida Prado, 1985, p.41), traduction de l'auteur.

3. Voir figures 12.3, 12.5 et 12.6.

4. Timbre *Grand Piano* de la bibliothèque de sons par défaut *Musescore_General.sf3*.

rugueux. Ce qui est observé ici n'est pas un cas particulier, mais constitue une tendance générale des résolutions cadentielles régulières faisant intervenir une septième de dominante, comme le montrent des tests réalisés sur différentes réalisations de cadences parfaites majeures et mineures. Malgré la disparité des trajectoires harmoniques, ces résolutions sont caractérisées par une augmentation de l'harmonicité et une diminution de la rugosité.

En conséquence, la similarité de la résolution du cycle de *Scorpio* avec la résolution de la cadence parfaite est frappante. L'accord λ joue le rôle de la septième de dominante tandis que ν joue le rôle de la tonique.

Nous sommes ici en présence d'un exemple paradigmatique d'écriture *transtonale*, qui « *absorbe et recycle des constructions non seulement atonales, mais aussi tonales*¹ ». Il convient, pour ne pas faire de contre-sens, de préciser les propos du compositeur quand il considère son écriture « *libérée des réminiscences des cadences tonales*² ». Comme le remarque Guigue (Guigue, 2009) à propos d'exemples similaires de pseudo-cadences transtonales tirées d'œuvres d'Almeida Prado, le sens cadentiel, dans le transtonalisme, n'est pas fourni par un hypothétique rapport tonal fonctionnel entre deux accords, mais par un rapport dialectique non fonctionnel. C'est dans cette absence de rapport fonctionnel qu'il convient de comprendre l'affirmation du compositeur.

Comme dans les résolutions tonales, le caractère cadentiel des résolutions transtonales est donné par un rapport dialectique qui concerne plusieurs dimensions de la musique. Dans la résolution du cycle de *Scorpio*, ce rapport intervient par les aspects suivants, communs aux cadences tonales :

- Aspect rythmique : la « dominante transtonale » qu'est λ domine la « tonique transtonale » en matière de durée. En effet, dans le premier cycle, elle est l'accord le plus répété avec 39 répétitions, contre seulement 3 répétitions pour ν .
- Relation de fermeture : λ , s'étendant sur près de six octaves, couvre tous les registres du clavier, tandis que ν est plus resserré, couvrant seulement trois octaves dans le registre médium.
- Richesse harmonique et structure des accords : λ , qui comporte les 12 hauteurs de l'échelle harmonique, est plus riche harmoniquement que

1. Voir note 7 page 430.

2. Voir note 4 page 430.

- ν , comportant seulement 6 hauteurs. Il est intéressant de noter que ν superpose un accord mineur ($Fa\sharp$ m) et un accord majeur (Sol M).
- Propriétés acoustiques : c'est ce que nous avons mis en avant dans cette section. La résolution, comme pour les cadences tonales, est caractérisée par une augmentation de l'harmonicité et une diminution de la rugosité.

Le dernier point montre que l'affinité qu'entretient la résolution de *Scorpio* avec les résolutions cadentielles tonales est profonde, résultat non pas uniquement des effets rythmiques, des effets de registre et de complexité harmonique, mais également de la structure intime des accords choisis. Les descripteurs harmoniques nous permettent de révéler cette affinité profonde, cet héritage tonal.

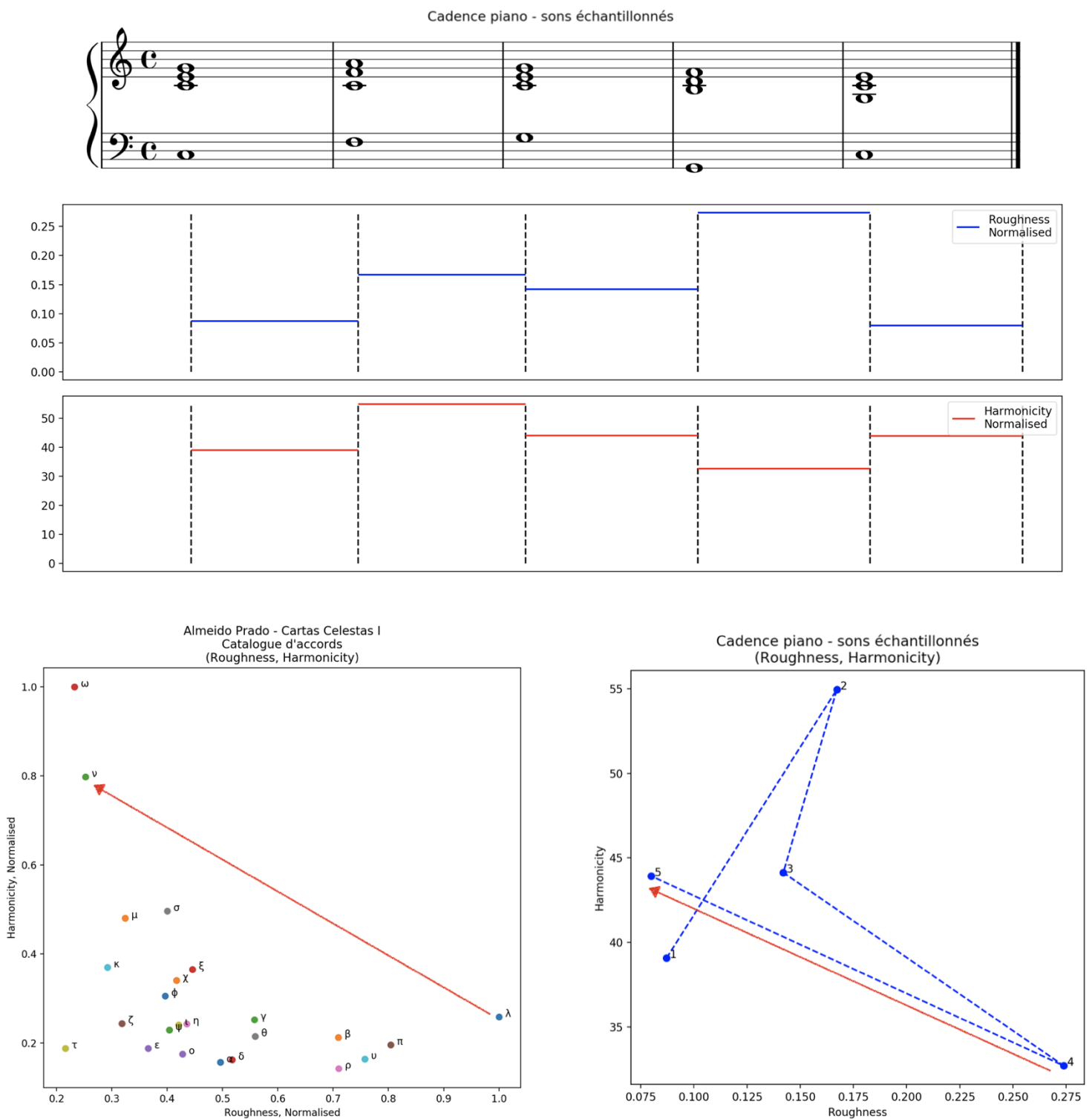


FIGURE 12.8 – Aspect cadentiel de *Scorpio* : comparaison avec une cadence majeure. En haut : évolution de la rugosité et de l’harmonicité sur la cadence, synthétisés avec des sons échantillonnés de piano. En bas à droite : représentation de la trajectoire harmonique de cette cadence dans l’espace (rugosité, harmonicité), la résolution est indiquée par une flèche rouge. En bas à gauche : cartographie du catalogue d’accords dans le même espace, et représentation du dernier enchaînement de *Scorpio* par une flèche rouge.

12.2.2 Incidence de la compression rythmique sur la résonance

L'analyse de la séquence d'accords du cycle de *Scorpio* à partir de la cartographie du catalogue d'accords est une analyse *hors-texte*, qui isole les accords en les sortant de leur contexte de jeu. Or ce contexte joue un rôle perceptif primordial, comme en témoigne la tenue de la pédale sur chaque cycle. Cette pédale, associée aux répétitions rapides des accords, mélange les sonorités pour créer un spectre de résonance. Aussi, à la dimension linéaire, temporelle, dont nous avons montré les ressorts tonaux, s'ajoute la dimension verticale, harmonique, à laquelle elle semble subordonnée. Outre les propos du compositeur qui présente sa musique comme une « *musique de la résonance* », l'écriture-même de *Scorpio* nous permet d'affirmer que la première est au service de la seconde. En effet, à partir d'une matière fixe constituée de l'enchaînement des 15 accords du cycle¹, tous les aspects de l'écriture d'Almeida Prado semblent aller dans le sens d'une recherche de résonance : pédale tenue sur chaque cycle, répétition des accords et surtout compression rythmique au fil des cycles. Il s'agit de créer par des moyens musicaux un espace cosmique dans lequel temps et espace sont intimement liés :

*L'immersion dans un espace résonant, finement construit par une série d'accords choisis selon leurs propriétés sonores, nous mène dans un lieu où temps et espace se côtoient et se bouleversent*².

Analyse Nous nous proposons d'entrer dans le détail de l'agencement du matériel harmonique initial pour comprendre comment Almeida Prado crée des *espace-temps* résonants, en mettant en lumière l'interdépendance entre la dimension temporelle et la dimension spatiale dans cette résonance. L'analyse se fera en deux temps, d'abord en déconstruisant l'écriture, puis en comparant les cycles entre eux.

1. À un iota près. Il convient de signaler une modification de ι , qui apparaît deux fois dans chaque cycle : le $la\sharp$ de la main droite disparaît au cycle 4 ainsi qu'à la première occurrence du cycle 3, et la main droite est transposée d'une octave ascendante dans la première occurrence du cycle 4.

2. Tiré de (Birschner Lira, 2021, p.29)

La résonance comme phénomène mouvant La résonance, pour Almeida Prado, n'est pas un phénomène statique, figé, mais une entité mouvante, en devenir. Aussi l'objectif des masses sonores est-il de « *remplir un espace dans le temps, de manière continue*¹ ». D'ailleurs, les répétitions rapides d'accords sont un moyen pour le compositeur d'entretenir le *continuum* sonore, de même que la pédale, qui, en mélangeant les sons, brouille les attaques. À propos de *Scorpio*, Almeida Prado lui-même ne parle-t-il pas de « *changement continu*² » ?

Cette dimension mouvante de la résonance nous incite à adopter une approche dynamique plutôt que statique. La concordance différentielle, qui quantifie les fréquences communes aux spectres de deux verticalités successives³, semble particulièrement pertinente pour étudier cette écriture, puisque :

*Il [le compositeur] intensifie avec des degrés différents une couleur ou une autre en utilisant une séquence d'accords dont les spectres partagent des harmoniques.*⁴

Nous utiliserons dans l'analyse la concordance différentielle ainsi que le changement harmonique.

Méthodologie Afin de pouvoir traiter différentes phases de l'articulation du matériel harmonique et de la création du spectre de résonance, nous avons enregistré le pianiste César Birschner dans trois configurations :

- Accords de *Scorpio* joués sans répétition et sans pédale. Chaque accord, joué *fortissimo*, est tenu pendant 5 secondes.
- Accords de *Scorpio* joués sans répétition et avec pédale. Chaque accord est là-aussi joué *fortissimo* et tenu pendant 5 secondes. La même pédale est commune à tous les accords.
- *Scorpio* dans son intégralité, en situation de jeu.

1. « *For the composer, the objective of sounds and sonorous masses is to fill a space in time, continuously, as the sound takes over the musical form or structure* ». (Ferraz, 2009, p.47), traduction de l'auteur.

2. Voir note 7 page 430.

3. Voir partie 6.2.2.

4. « *He also intensifies to varying degrees one color or another by using a sequence of chords whose spectrums share some harmonics* ». (Ferraz, 2009, p.41), traduction de l'auteur.

Cette déconstruction de l'écriture vise à isoler les effets de la pédale et des répétitions sur les descripteurs harmoniques employés.

Pour l'analyse de *Scorpio*, nous procédons à une segmentation de l'audio correspondant à la segmentation en mesures. Ce faisant, nous élargissons la notion de verticalité définie dans l'introduction¹, qui englobe maintenant toutes les répétitions du même accord. En synchronisant les spectres sur les verticalités², nous associons un spectre moyen à chaque mesure. Étant donné le rôle évident des répétitions et des attaques dans la section, nous ne procédons pas à une séparation de la partie harmonique et de la partie percussive.

L'intérêt d'une telle méthode réside dans le fait que pour un accord donné du cycle de *Scorpio*, son spectre associé, en situation de jeu, dépend du nombre de répétitions ainsi que de la résonance portée par la pédale, elle-même fonction de ce qui a été entendu précédemment. Cela va nous permettre de discriminer les différents cycles de *Scorpio* en mettant en avant l'effet de la compression rythmique sur les transitions.

Effet de la pédale et des répétitions La figure 12.9 montre l'effet de la pédale sur les transitions, en comparant les accords sans répétition et sans pédale (en bleu) avec les accords sans répétition et avec pédale (en orange). Le cycle de *Scorpio* étant composé de 15 accords, les trajectoires dans l'espace (concordance différentielle, changement harmonique) comportent 14 transitions.

On observe une différence notable entre le calcul des descripteurs dynamiques sur les classes d'accords³ ou sur des sons réels. Sur les classes d'accords, la concordance différentielle et le changement harmonique étaient redondants car très fortement corrélés négativement. Sur les sons réels, par contre, cette corrélation disparaît, comme en témoigne le coefficient de corrélation positif $r = 0.36$ de la trajectoire de transitions pour les accords sans répétition et sans pédale (en bleu).

Nous ne nous intéressons pas spécifiquement au détail de chaque trajectoire mais à la comparaison des trajectoires entre elles.

1. Voir page 58.

2. Voir partie 5.3.2.

3. Voir partie 9.3 et figure 9.38.

Effet de la pédale La pédale a pour effet global d'augmenter la concordance différentielle en diminuant le changement harmonique. En effet, sur une transition, sa tenue prolonge la première verticalité sur la seconde, d'où un plus grand partage de partiels sur les spectres synchronisés, et une réduction des différences entre ces spectres.

Certaines transitions ne sont que peu affectées, en matière de position dans l'espace de descripteurs dynamiques, par la tenue de la pédale. C'est le cas des transitions 5, 12 et 13, ce qui indique que pour ces transitions le spectre du premier accord est déjà en grande proportion contenu dans le spectre du nouvel accord. Penchons-nous sur la transition 5, entre α et ι . Le *si* et le *fa*♯ de α sont compris dans le spectre du *si* de ι (harmoniques n°2 et 3), tandis que le *do* et le *sol* sont compris dans le spectre du *fa* grave de ι (harmoniques n°6 et 9).

Effet de la répétition La figure 12.10 montre l'effet des répétitions du premier cycle sur les transitions, en comparant les deux trajectoires précédentes avec le premier cycle de *Scorpio* (en vert). Dans ce dernier, le nombre de répétitions, et donc la durée de la fenêtre de synchronisation, varie en fonction des accords.

Par rapport à la trajectoire correspondant aux accords joués sans répétition, avec une durée constante et avec pédale (en orange), les répétitions dans le premier cycle de *Scorpio* ont pour effet global de diminuer la concordance différentielle en augmentant le changement harmonique. En effet, la répétition d'un accord modifie dans le spectre associé à la verticalité le rapport entre sa contribution et la résonance des accords précédents dans la pédale, au détriment de cette dernière. Cela est dû autant à l'entretien de l'accord qu'à la décroissance de la pédale. Ainsi, la trajectoire se rapproche globalement de la trajectoire obtenue sans pédale (en bleu). Elle est évidemment sensible au nombre de répétitions de chaque accord. Comme c'est ce dernier qui différencie les 4 cycles de *Scorpio*, les trajectoires des cycles diffèrent, comme nous allons le voir, mais sont toujours situées, dans leur globalité, à l'intermédiaire entre la configuration sans répétition et sans pédale (en bleu) et la configuration sans répétition avec pédale (en orange).

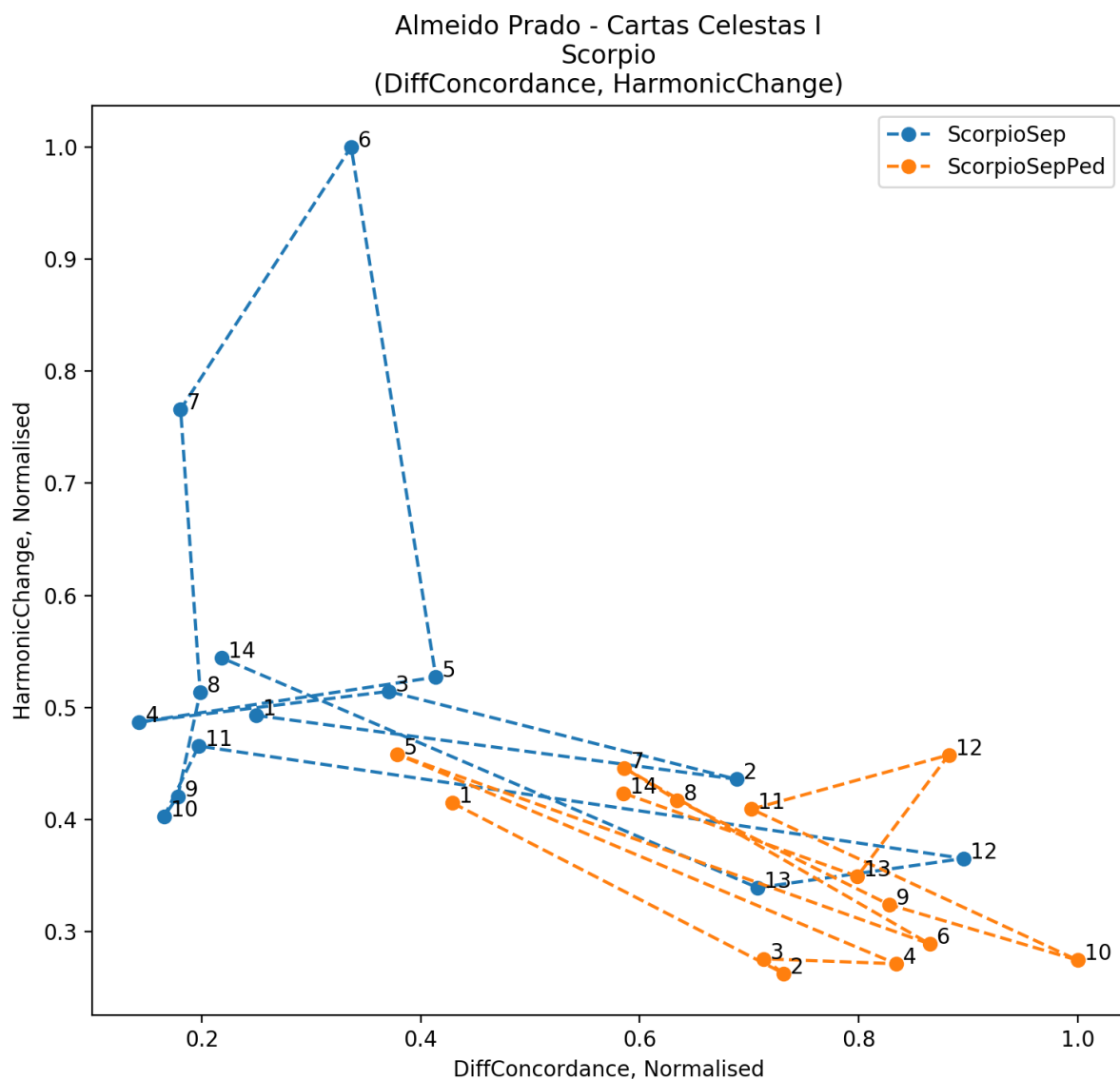


FIGURE 12.9 – *Scorpio* - effet de la pédale sur les transitions. Comparaison des accords sans répétition et sans pédale (en bleu) avec les accords sans répétition et avec pédale (en orange), sur les trajectoires des transitions dans l'espace (Concordance différentielle, Changement harmonique).

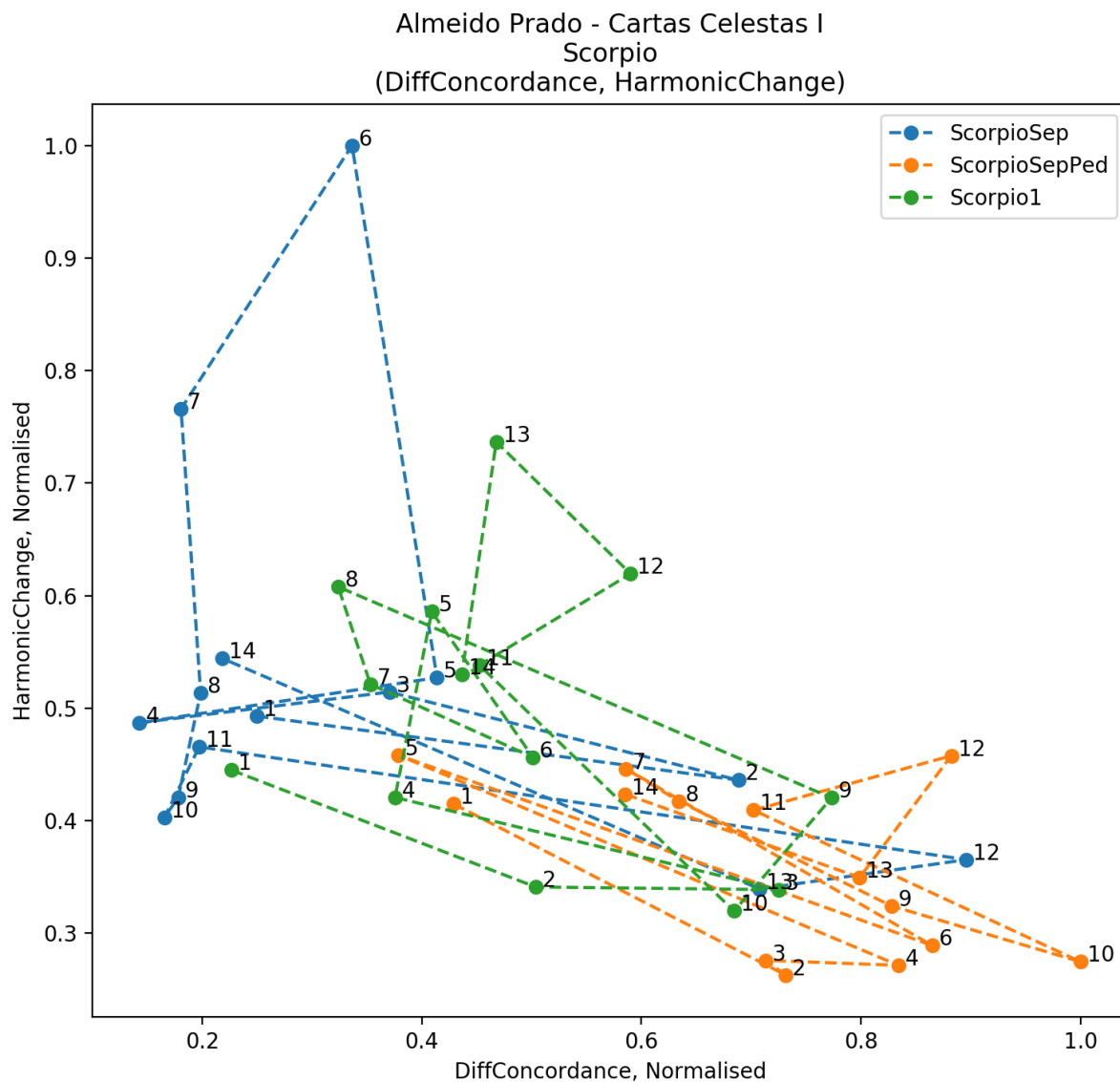


FIGURE 12.10 – *Scorpio* - effet des répétitions sur les transitions. Comparaison des accords sans répétition et sans pédale (en bleu), des accords sans répétition et avec pédale (en orange) et du premier cycle de *Scorpio* en situation de jeu (en vert), sur les trajectoires des transitions dans l'espace (Concordance différentielle, Changement harmonique).

Effet de la compression rythmique La succession des quatre cycles de *Scorpio* est caractérisée par une diminution des répétitions de cycle en cycle pour chaque accord, créant ainsi une compression rythmique. Celle-ci a pour effet de donner plus d'importance à la résonance de la pédale, c'est-à-dire d'augmenter la concordance différentielle en diminuant le changement harmonique. La figure 12.11 l'atteste, qui compare les trajectoires de transitions des cycles successifs de *Scorpio*. Ces effets vont dans le sens d'un déroulement de plus en plus continu de la résonance, en même temps qu'un enrichissement maximal du spectre de résonance.

Conclusion Les descripteurs harmoniques dynamiques nous permettent d'explicitier la manière dont l'écriture temporelle de *Scorpio* contribue à créer un *espace-temps* résonnant en mouvement, par l'effet de répétition. Notamment, ils quantifient les rôles respectifs de la pédale, de la répétition et de la contraction rythmique.

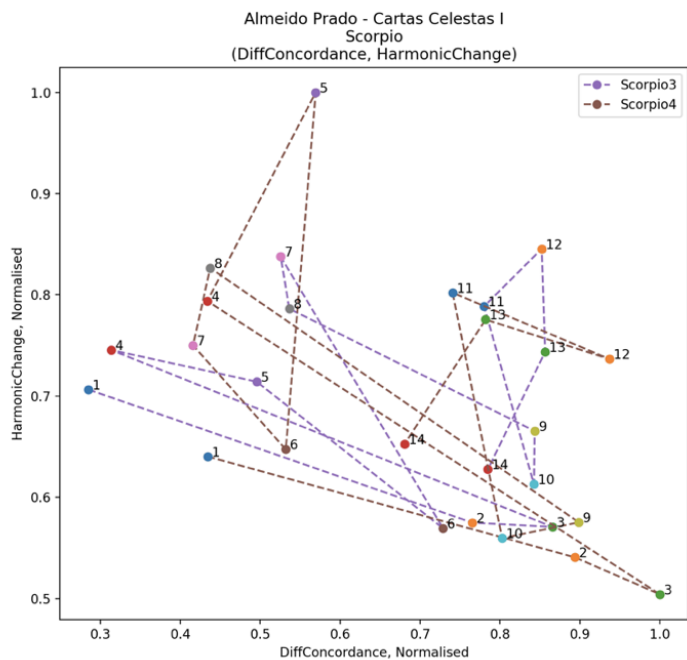
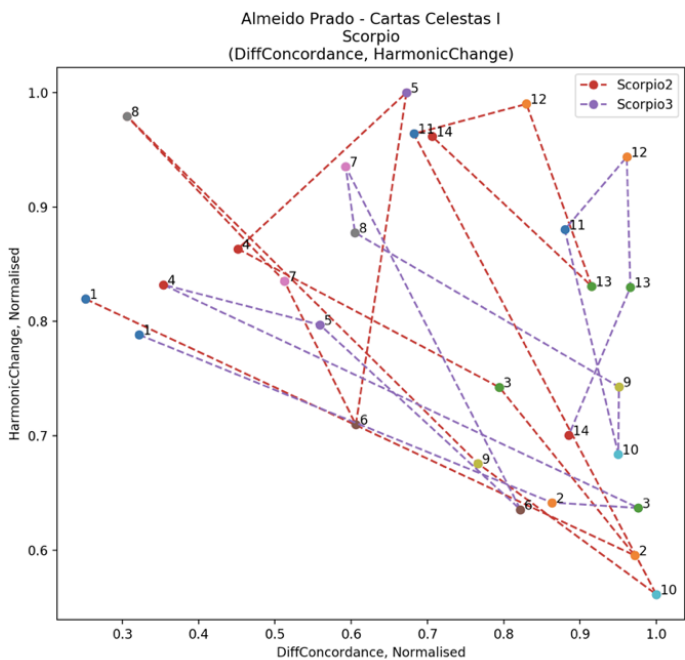
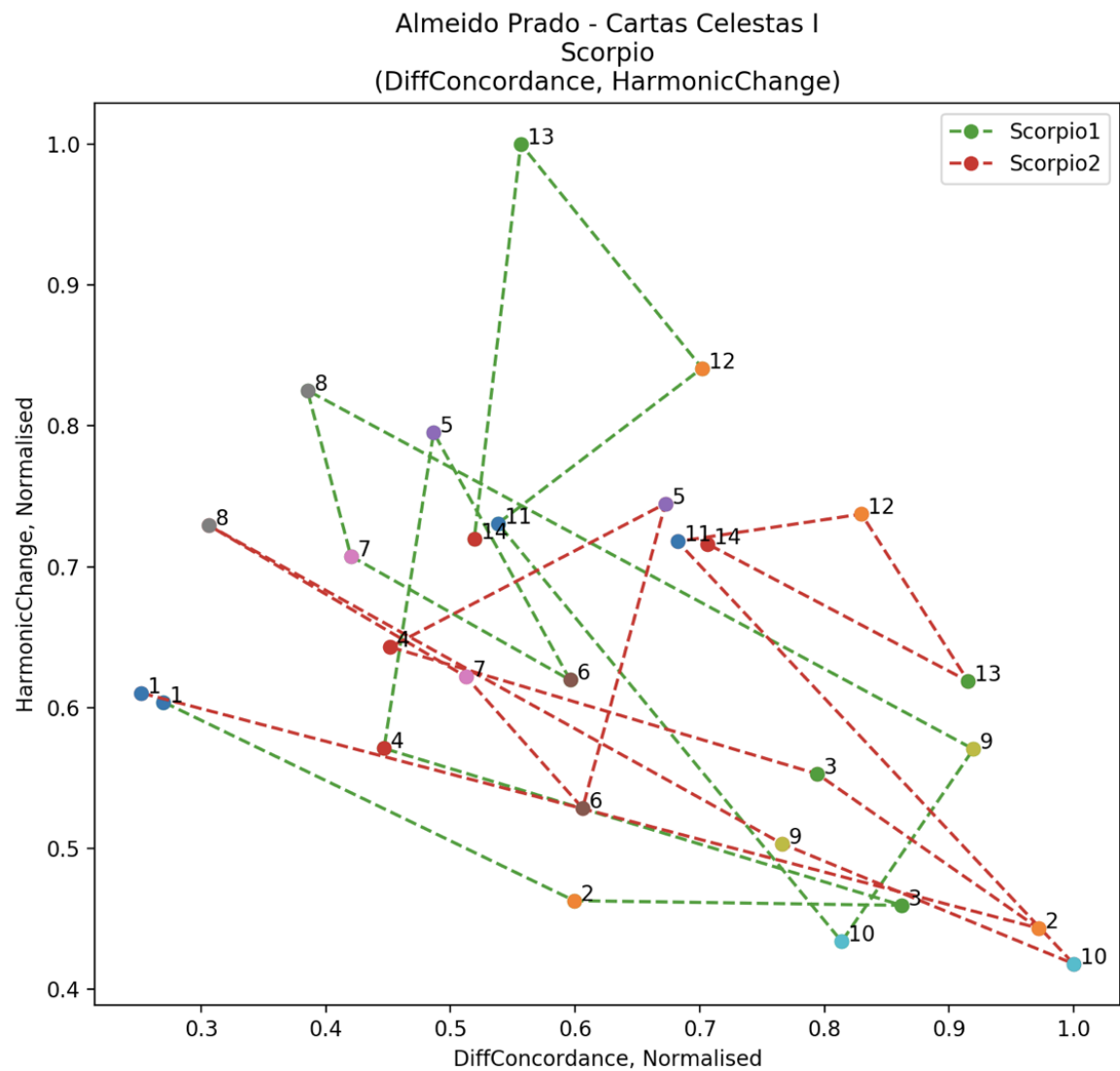


FIGURE 12.11 – *Scorpio* - effet de la compression rythmique sur les transitions. Comparaison des cycles 1 (en vert) et 2 (en rouge) en haut, 2 et 3 (en violet) en bas à gauche, et 3 et 4 (en marron) en bas à droite, sur les trajectoires des transitions dans l'espace (Concordance différentielle, Changement harmonique).

Conclusion

Résumé des travaux

Au point de départ de ce travail se trouve une constatation : celle de l'isolement relatif de trois champs de recherche, à savoir le domaine de la perception musicale, le monde des descripteurs audio, et enfin l'analyse musicale. Et pour cause, ces champs divergent à la fois dans leurs objectifs et leurs méthodes. Les recherches perceptives visent à comprendre les mécanismes de la perception et à la modéliser, via des expériences perceptives. Les descripteurs audio sont des méthodes de traitement du signal décrivant certains aspects du signal, dans des buts variés : segmentation automatique, détection d'accords, reconnaissance de style... L'analyse musicale, quant à elle, vise à comprendre les œuvres musicales sous le prisme de méthodes et de théories, dont beaucoup partent de la partition.

La question de l'harmonie cristallise cette constatation en frustration, car elle est dans chaque domaine très poussée. On peut citer plusieurs dizaines, si ce n'est plus, de modèles de la consonance et de la dissonance, un des aspects fondamentaux de la perception de l'harmonie. Les théories musicales traitant le matériau harmonique sont nombreuses : chiffage harmonique, analyse schenkérienne, théorie générative de la musique tonale, vecteurs harmoniques, analyse néo-riemannienne, *set theory*... Les descripteurs liés de près ou de loin à l'harmonie, eux aussi, sont nombreux : inharmonicité, centroïde spectral, rapport des harmoniques pairs sur les harmoniques impairs... Les modèles de perception sont rarement employés en analyse musicale, alors qu'ils auraient toute la légitimité de l'être, de même que les descripteurs audio. L'application à l'analyse d'une œuvre des modèles de perception est souvent trop rudimentaire, peu adaptée, et n'intègre pas nécessairement les techniques de traitement de signal développées dans le domaine des descripteurs audio.

C'est dans la perspective d'une porosité de ces disciplines que nous avons introduit la notion de descripteur harmonique, comme type particulier de modèle acoustique adapté à l'usage musicologique et rendant compte du phénomène harmonique. Les descripteurs harmoniques peuvent avoir une origine perceptive, mais pas nécessairement. Ils utilisent des méthodes de traitement du signal, et leur but est d'ordre musicologique : ils participent d'une compréhension fine de l'intérieur du phénomène harmonique, sous l'angle acoustique. Appliqués au contexte musical, ils viennent éclairer des éléments de langage ou de choix d'écriture.

Cette visée spécifique nous a amené à développer une implémentation

hybride, à l'interface entre le signal et le symbolique, en accord avec notre approche de l'harmonie comme l'ensemble des phénomènes sonores ou relations formelles impliquant l'interaction simultanée ou successive de plusieurs notes. Les descripteurs harmoniques prennent en entrée soit un signal audio, soit plusieurs signaux audio, soit encore une partition associée à un modèle de spectre. Afin que les représentations issues des modèles soient adaptées à l'usage musicologique, nous avons mis en place une méthodologie qui diffère de celle des descripteurs audio classiques, essentiellement par deux points. Le premier vient de l'exigence de lisibilité et d'interprétabilité des représentations, plus spécifiquement de la possibilité de la mise en parallèle avec une partition. À cette fin, les descripteurs sont calculés non pas sur les fenêtres de l'analyse spectrale, mais sur l'intervalle temporel séparant deux notes successives. Cela passe par un moyennage des spectres qui réduit par segmentation le *continuum* temporel en un *discontinuum*, et limite les applications à des sons spectralement stables. Le second point à différer des approches standard résulte de notre définition de l'harmonie mettant en jeu les interactions simultanées ou successives de plusieurs notes. La quantification des interactions simultanées des différentes notes nécessite, pour une partie des descripteurs, de pouvoir distinguer les contributions spectrales de chaque note, ce qui requiert d'avoir en entrée l'audio des notes séparées ou des voix, en plus du signal global. De plus, pour rendre les représentations cohérentes avec l'usage musical, des contraintes déterminant des normalisations sont imposées sur les descripteurs : l'indépendance aux nuances globales et l'égalité des descripteurs statiques sur l'unisson indépendamment du nombre de notes. Ces normalisations motivées par des considérations musicales représentent une nouveauté par rapport aux implémentations classiques des modèles de perception.

Nous avons introduit dans ce travail plusieurs descripteurs harmoniques, qui ont en commun de décrire les interactions de nature spectrale entre les notes simultanées ou successives. Une partie d'entre eux sont des modèles de perception, liés à la question de la consonance et de la dissonance (rugosité, harmonicité, tension) ou de l'affinité de sons successifs (concordance différentielle). La rugosité différentielle, sans être elle-même un modèle perceptif, généralise à l'approche dynamique un modèle de perception statique. Les autres descripteurs harmoniques ont des motivations essentiellement acoustiques (concordance, concordance totale, concordance d'ordre supérieur). Enfin, tandis que les autres descripteurs n'ont pas été pensés pour un contexte historique ou esthétique particulier, l'harmonicité inverse est introduite pour

analyser un langage harmonique spécifique, le « transtonalisme ». Elle est motivée par les propos du compositeur ayant utilisé ce langage.

Nous avons tiré deux usages distincts des descripteurs harmoniques : ils nous ont permis de révéler les structures acoustiques du matériau harmonique abstrait, ainsi que d'éclairer des éléments de langage et des choix d'écriture dans les œuvres étudiées.

L'approche symbolique nous a donné les moyens de caractériser des espaces d'accords et de transitions continus fréquemment. À partir d'une représentation linéaire des intervalles et plane des triades, les descripteurs harmoniques statiques induisent de véritables cartographies acoustiques de ces espaces continus, dépendantes du timbre et du registre, qui invitent à interpréter les accords en matière de propriétés physiques. La concordance d'ordre 3 nous a en outre permis de reformuler l'interprétation naturelle de l'accord mineur de Rameau. Les structures acoustiques des espaces de transitions, notamment la carte de gain de rugosité, fournissent quant à elle des éléments de réponse possibles à des questions perceptives encore irrésolues.

Nous avons utilisé des spectres de Shepard pour généraliser le calcul des descripteurs harmoniques aux classes d'accords, conciliant ainsi approche acoustique et approche formelle. Là encore, via des représentations géométriques, les descripteurs statiques révèlent les structures acoustiques des espaces continus de classes d'accords. Une application particulièrement riche concerne la classification de Julio Estrada, qui fournit une vue d'ensemble compacte de tous les accords possibles dans une échelle égale donnée, chromatique ou microtonale. Les descripteurs statiques apportent une « coloration acoustique » à cette vue d'ensemble. La seconde application concerne le réseau de notes connu sous le nom de *Tonnetz*. En se restreignant à un nombre limité de types d'accords, nous avons associé les descripteurs dynamiques à un modèle de mémoire à court terme pour caractériser les transitions dans les enchaînements harmoniques sur le *Tonnetz*. Par cette approche nous avons ébauché une interprétation des résolutions régulières de cadences tonales en matière d'optimisation de la concordance différentielle et du changement harmonique.

Le corpus d'étude nous a permis de mettre en avant le potentiel des descripteurs harmoniques en analyse musicale. Dans l'extrait de *Fratres* de Pärt, les irrégularités dans l'harmonisation du thème rétrogradé sont comprises comme un moyen d'augmenter la concordance et l'harmonicité tout en diminuant la dissonance. Sur une comparaison de plusieurs interprétations du début de la *Pièce pour orgue n°2* de Schnittke, les descripteurs

harmoniques confirment les intuitions perceptives. L'abstraction aux classes d'accords et la coloration des trajectoire sur la classification de Julio Estrada aboutissent aux mêmes profils de descripteurs, ce qui montre l'intérêt de la représentation. Avec les *Cartas Celestes I* d'Almeida Prado, les descripteurs harmoniques ont été utilisés pour confronter les propos du compositeur à la réalité acoustique de l'œuvre, dans le cadre du système harmonique « trans-tonal ». Les *Cartas Celestes* ont la particularité de contenir un catalogue d'accords constituant le vocabulaire harmonique de l'œuvre. L'analyse du catalogue par les descripteurs statiques donne lieu à une cartographie des accords qui confère un statut spécial à deux accords particuliers, dont l'un joue un rôle précis dans la section *Scorpio*. De plus, l'introduction d'un descripteur *ad hoc*, l'harmonicité inverse, nous a permis de confirmer les propos du compositeur quant à l'égalité des rôles respectifs des séries harmoniques supérieure et inférieure dans le phénomène de résonance ayant guidé le choix des accords du catalogue. En comparant la trajectoire harmonique dans un espace de descripteurs statiques avec la trajectoire harmonique d'une cadence tonale, nous avons mis en évidence la présence d'un héritage tonal de résidus cadentiels.

Apports

Les apports de ce travail sont d'ordre à la fois théorique, méthodologique et analytique.

Du point de vue théorique, la contribution principale est la notion de descripteur harmonique, développée pour motiver une étude des structures acoustiques du phénomène harmonique. Nous avons introduit dans le champ de la musicologie de nouveaux modèles statiques : la concordance totale et la concordance d'ordre fixe, ainsi que dynamiques : changement harmonique et rugosité différentielle. La normalisation des descripteurs sous des contraintes musicales liées aux nuances et au nombre de notes constitue une nouveauté par rapport à ce qui se fait dans les domaines de la perception et du traitement du signal pour l'analyse musicale. Les descripteurs ont fourni des interprétations nouvelles de questions traitées dans la littérature : le caractère « naturel » de l'accord mineur tel que le présente Rameau, ou encore le fait que la triade *do-mi-si* « sonne mieux » que l'intervalle de septième *do-si*. Nous avons décrit géométriquement les espaces continus de classes d'accords pour les triades et les tétrachordes. Enfin, nous avons concilié les approches acoustiques et formelles en ajoutant une « couleur harmonique » aux espaces d'accords abstraits. Sur la classification de Julio Estrada, cette couleur caractérise les propriétés acoustiques des classes d'accords, tandis que sur le *Tonnetz* elle caractérise les transitions entre classes d'accords.

D'un point de vue méthodologique, nous avons mis en place une approche hybride à l'interface entre le signal et le symbolique, traduisant la volonté d'une description acoustique adaptée à l'objet musical. L'approche diffère de celle des descripteurs audio classiques par le fait que les modèles sont calculés sur l'intervalle temporel séparant deux notes successives, et certains modèles nécessitent les pistes audio des voix séparées. Elle diffère également des méthodes d'analyse harmoniques partant de la partition en ce qu'elle exige des données spectrales, et s'attache par là à l'aspect sonore de l'objet étudié. Une prise en compte de la mémoire à court terme est proposée. Plusieurs représentations nouvelles ont été introduites : les trajectoires harmoniques dans les espaces de descripteurs statiques ou dynamiques, ainsi que les cartes de gain dans les espaces de transition par ajout d'une note. De plus, les spectres de Shepard ont été utilisés pour caractériser acoustiquement les classes d'accords.

L'apport du travail dans le domaine de l'analyse musicale consiste en l'introduction d'une méthodologie spécifique permettant d'adapter des modèles

acoustiques à l'usage musicologique. Cette méthodologie nous a permis de révéler à la fois les structures acoustiques des espaces harmoniques, et, dans le cas d'extraits musicaux, des logiques acoustiques à l'œuvre dans l'écriture. Nous avons notamment apporté de nouveaux éléments de compréhension de l'écriture « transtonale » d'Almeida Prado, en mettant en évidence la présence de résidus cadentiels révélateurs d'un héritage tonal.

Perspectives

Ce travail relève d'une approche pluridisciplinaire qui a de quoi déconcerter, et pour cause.

Les psychologues pourront s'étonner d'une démarche qui utilise des modèles perceptifs sans réaliser d'expériences perceptives ni même se préoccuper de leur pertinence perceptive dans le contexte où ils sont utilisés. Bien qu'en grande partie motivé par des questions perceptives, il est vrai que ce travail ne réalise pas d'apports dans ce domaine. La concordance, la concordance totale et la rugosité différentielle mériteraient cependant de faire l'objet d'études expérimentales.

Les physiciens et acousticiens pointeront du doigt l'aspect rudimentaire des modèles, notamment une vision simplifiée voire erronée de l'objet sonore. En effet, la réduction du son à un spectre en amplitude, ajoutée à son moyennage sur les verticalités, exclue de l'analyse tout un aspect du son, à savoir la phase et l'évolution temporelle, plus la dimension spatiale (spatialisation, dispositif d'écoute). Les modèles pourraient aussi tenir compte de l'acoustique des salles. Leurs remarques seraient tout à fait justes et légitimes. En outre, l'étape suivante de ce travail serait de proposer des modèles harmoniques non directement spectraux, rendant compte de la forme d'onde, d'effets de phase et d'un suivi temporel fin. L'implémentation nécessiterait alors d'être adaptée.

Les chercheurs en sciences cognitives objecteront que la modélisation de la mémoire à court terme est là-aussi très rudimentaire, et que le travail ne dit rien de la manière dont les aspects statiques et dynamiques de l'harmonie interagissent. Concernant la simplicité du modèle de mémoire à court terme, nous l'avons seulement appliqué à la caractérisation des transitions dans le *Tonnetz*, donc à un niveau déjà abstrait. L'usage d'un modèle de mémoire à court terme dans le contexte d'une œuvre ne peut être pertinent, à notre avis, qu'à condition qu'il tienne compte des aspects rythmiques et métriques. Le modèle de mémoire proposé a plus une valeur d'ébauche, et mériterait d'être affiné. La question de l'interaction des aspects statiques et dynamiques de l'harmonie constitue une question théorique de première importance, sur laquelle peu de choses sont réellement connues. Trop peu, en tout cas, pour pouvoir proposer un modèle qui ne serait pas spéculatif.

Les mathématiciens pourront être frustrés par le fait que la dimension de la représentation limite les espaces continus de transitions considérés au rajout d'une troisième note à un intervalle. Cette limitation n'est en effet pas

satisfaisante au vu de la réalité musicale, même si elle aboutit déjà à des résultats. Cependant, il n'est pas exclu qu'une représentation plus adaptée permette d'élargir le spectre des transitions. De plus, dans notre implémentation du second niveau de la représentation du potentiel intervallique d'Estrada, les formes normales des orbites identitaires sont listées, et non pas structurées géométriquement par les permutations de proche en proche, comme invitait à le faire leur représentation dans le volume des formes normales pour le cas des tetrachordes. C'est par souci de simplicité que nous avons fait le choix de les lister. Les structures des *combinohedron* se déploient dans plusieurs dimensions, et leur représentation en toute généralité n'est pas sans poser des difficultés. Enfin, la représentation des espaces continus de classes d'accords pour les triades et les tetrachordes pourrait être généralisée aux dimensions supérieures. Espérons que ce travail incite à approfondir ces questions de représentations.

Les informaticiens pourront exiger, et à raison, une mise à disposition des outils développés dans ce travail. Nous envisageons de les rendre disponibles sur Github sous forme de programme Python, ou d'intégrer l'analyse par suivi temporel dans un *plugin Vamp* pour *Audacity* ou *Sonic visualiser*. Du point de vue informatique, une des perspectives envisageables est le calcul et l'affichage des descripteurs harmoniques en temps réel, par exemple pour projeter une « couleur harmonique » calculée en direct. Il est possible pour ce faire d'imaginer un dispositif prenant en entrée la sortie midi d'un synthétiseur.

Les musicologues pourront avoir l'impression que les outils présentés aident à comprendre le son, mais pas la musique. La musique étant un objet culturel, une approche limitée à l'acoustique serait inapte à saisir le fait musical. Si elle est inapte, en effet, à saisir la totalité des facettes d'un fait musical, comme l'est d'ailleurs toute autre approche, elle l'éclaire sous un point de vue précis, et par là participe de sa compréhension, sans prétention à le justifier. En outre, la compréhension qu'elle offre du phénomène harmonique est une compréhension fine, de l'intérieur. De plus, les musicologues pourront être frustrés que nous ne soyons allés plus loin dans l'interprétation des résolutions régulières de cadences tonales sur le *Tonnetz* en matière d'optimisation des descripteurs dynamiques. En effet, ces résultats mériteraient d'être creusés et d'être confrontés aux théories tonales comme la théorie des fonctions ou l'analyse schenkérienne. On pourrait notamment imaginer un suivi de l'évolution des descripteurs dynamiques sur toutes les transitions d'un enchaînement, et non uniquement sur la résolution. Ces résultats ouvrent de

nombreuses perspectives de recherche. Le nombre d'applications potentielles des descripteurs harmoniques nous a obligé à faire des choix, et à ne pas tout aborder, faute de temps. Citons quelques applications qui auraient eu un intérêt musicologique certain. Tout d'abord, la dépendance des descripteurs harmoniques au timbre fournit des outils pour étudier les questions d'orchestration et éclairer certains usages. L'étude historique des tempéraments est également une application envisageable. Ensuite, la configuration mélodie-accompagnement, que nous avons présentée sans en donner d'application, paraît particulièrement adaptée à une démarche de comparaison de différentes harmonisations d'un même motif. De plus, le calcul des descripteurs sur une version altérée d'une cadence du *Via Crucis* de Liszt¹ invite à aller plus loin dans la description acoustique des « fautes d'harmonie » par rapport aux « réalisations justes », l'optique étant d'éclairer un usage canonique par des considérations acoustiques. Espérons que les descripteurs harmoniques motivent des recherches dans ces directions.

Dans les structures acoustiques des espaces harmoniques révélées par les descripteurs harmoniques, les compositeurs pourront voir des guides ou prétextes à la composition. Ces structures invitent en effet à être explorées non pas uniquement sous l'angle théorique ou analytique, mais aussi sous l'angle de la création. Les espaces harmoniques deviendraient alors à proprement parler des espaces de potentialité devant la feuille blanche. Ces structures paraissent d'autant plus pertinentes dans le cas d'univers musicaux peu explorés par la théorie harmonique, comme celui des échelles microtonales ou des spectres inharmoniques. Même si le corpus traité dans ce travail utilise la gamme chromatique, les descripteurs harmoniques peuvent sans restriction être appliqués à d'autres échelles. Avec des sons inharmoniques, la construction des spectres de Shepard doit cependant être modifiée, l'intervalle de réplique pouvant différer de l'octave. De même, quelques ajustements s'imposent dans la construction du potentiel harmonique d'Estrada, concernant à la fois l'intervalle de duplication de fréquence et le nombre d'identités, celle-ci dépendant du nombre de divisions de l'intervalle de duplication de fréquence. De plus, la « couleur harmonique », introduite dans un but analytique sur cette classification ainsi que sur le *Tonnetz*, pourrait acquérir une dimension visuelle et poétique dans une démarche de création. L'emploi des descripteurs harmoniques par les compositeurs serait facilitée par leur implémentation et leur mise à disposition sur *Max* et *Open Music*.

1. Voir figure 6.23 page 234.

Bibliographie

- JL Ramirez ALFONSIN et David ROMERO : Embeddability of the combino-hedron. *Discrete mathematics, Elsevier*, 254(1-3):473–483, 2002.
- Daniel ALGOM et Lawrence E MARKS : Range and regression, loudness scales, and loudness processing : Toward a context-bound psychophysics. *Journal of Experimental Psychology : Human Perception and Performance, American Psychological Association*, 16(4):706, 1990.
- José Antonio Rezende de ALMEIDA PRADO : *Cartas celestes : uma uranografia sonora geradora de novos processos composicionais*. Thèse de doctorat, Instituto de Artes, Universidade Estadual de Campinas, 1985.
- Jacques AMBLARD : Arvo Pärt - Éléments de style. *Brahms (encyclopédie en ligne de l'IRCAM)*, 2019.
- Rytis AMBRAZEVIČIUS : Dissonance/roughness and tonality perception in lithuanian traditional schwebungsdiaphonie. *Journal of interdisciplinary music studies*, 8:39–53, 2017.
- John R ANDERSON et Christian J LEBIERE : *The atomic components of thought*. Psychology Press, New-York, 2014.
- ARISTOXÈNE : *Éléments harmoniques*. P. Haffner, 1870. traduction Charles-Émile Ruelle.
- Yuko ARTHURS : *The creation of consonance : how musical context influences chord perception*. Thèse de doctorat, University of Sheffield, 2015.
- Yuko ARTHURS et Renee TIMMERS : On the fluidity of consonance and dissonance : The influence of musical context. *Psychomusicology : Music, Mind, and Brain, Educational Publishing Foundation*, 26(1):1, 2016.

- Mitchell G ASH : *Gestalt psychology in German culture, 1890-1967 : Holism and the quest for objectivity*. Cambridge University Press, 1998.
- Patrice BAILHACHE : *Leibniz et la théorie de la musique*. Domaine musico-logique. Klincksieck, Paris, 1992.
- Patrice BAILHACHE, Céline VAUTRIN et Antonia SOULEZ : *Helmholtz, du son à la musique*. Vrin, Paris, 2011.
- Llorenç BALSACH : *Los fundamentos de las tensiones harmónicas*. Coedición : Editorial Boileau y Edicions La mà de Guido, Barcelone, 2016.
- Gilles BAROIN : *Applications de la théorie des graphes à des objets musicaux : modélisations, visualisations en hyperspace*. Thèse de doctorat, Université de Toulouse, 2011.
- Pierre BARROUILLET, Sophie BERNARDIN et Valérie CAMOS : Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology : General, American Psychological Association*, 133(1):83, 2004.
- Arthur H BENADE : *Fundamentals of musical acoustics*. Courier Corporation, 1990.
- Jamshed BHARUCHA et Carol L KRUMHANSL : The representation of harmonic structure in music : Hierarchies of stability as a function of context. *Cognition*, 13(1):63–102, 1983.
- Gavin M BIDELMAN et Michael G HEINZ : Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing. *The Journal of the Acoustical Society of America*, 130(3):1488–1502, 2011.
- Gavin M BIDELMAN et Ananthanarayan KRISHNAN : Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience, Soc Neuroscience*, 29(42):13165–13171, 2009.
- Louis BIGO : *Représentations symboliques musicales et calcul spatial*. Thèse de doctorat, Université Paris Est, 2013.

- Louis BIGO, Daniele GHISI, Antoine SPICHER et Moreno ANDREATTA : Representation of musical structures and processes in simplicial chord spaces. *Computer Music Journal, MITP*, 39(3):9–24, 2015.
- César BIRSCHNER LIRA : Cartas celestes I d’Almeida Prado - transtonalité et espace sonore cosmique. Mémoire de Master 2, Sorbonne Université, 2021.
- Marcel BITSCH : *Précis d’harmonie tonale*. Alphonse Leduc, Paris, 1957.
- Dmitry BOGDANOV, Nicolas WACK, Emilia GÓMEZ GUTIÉRREZ, Sankalp GULATI, Herrera BOYER, Oscar MAYOR, Gerard ROMA TREPAT, Justin SALAMON, José Ricardo ZAPATA GONZÁLEZ, Xavier SERRA *et al.* : Essentia : An audio analysis library for music information retrieval. In *Britto A, Gouyon F, Dixon S, editors. 14th Conference of the International Society for Music Information Retrieval (ISMIR) ; 2013 Nov 4-8 ; Curitiba, Brazil.[place unknown] : ISMIR ; 2013. p. 493-8*. International Society for Music Information Retrieval (ISMIR), 2013.
- Oliver BONES, Kathryn HOPKINS, Ananthanarayan KRISHNAN et Christopher J PLACK : Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia*, 58:23–32, 2014.
- Georges BONNET : Considérations sur la représentation et l’analyse harmonique des signaux déterministes ou aléatoires. In *Annales Des Télécommunications*, volume 23, pages 62–86, 1968.
- Daniel L BOWLING : Harmonicity and roughness in the biology of tonal aesthetics. *Music Perception : An Interdisciplinary Journal, University of California Press*, 38(3):331–334, 2021.
- Daniel L. BOWLING, Marisa HOESCHELE, Kamraan Z. GILL et W. Tecumseh FITCH : The nature and nurture of musical consonance. *Music Perception : An Interdisciplinary Journal*, 35(1):118–121, 2017.
- Daniel L BOWLING et Dale PURVES : A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences*, 112(36):11155–11160, 2015.

- Daniel L BOWLING, Dale PURVES et Kamraan Z GILL : Vocal similarity predicts the relative attraction of musical chords. *Proceedings of the National Academy of Sciences*, 115(1):216–221, 2018.
- Austin M BRUES : The fusion of non-musical intervals. *The American Journal of Psychology*, 38(4):624–638, 1927.
- Samuele CARCAGNO, Saday LAKHANI et Christopher J. PLACK : Consonance perception beyond the traditional existence region of pitch. *The Journal of the Acoustical Society of America*, 146(4):2279–2290, 2019.
- Norman CAZDEN : The definition of consonance and dissonance. *International Review of the Aesthetics and Sociology of Music*, pages 123–168, 1980.
- Antoine CHAIGNE et Jean KERGOMARD : *Acoustique des instruments de musique*. Belin, Paris, 2008.
- Paul Yaozhu CHAN, Minghui DONG, Haizhou LI *et al.* : The science of harmony : A psychophysical basis for perceptual tensions and resolutions in music. *Research*, 2019.
- André CHARRAK : *Raison et perception - fonder l'harmonie au XVIIIe siècle*. Mathesis. Vrin, Paris, 2001.
- Elaine CHEW : *Towards a mathematical model of tonality*. Thèse de doctorat, MIT, Cambridge, États-Unis, 2000.
- Cinzia CHIANDETTI et Giorgio VALLORTIGARA : Chicks like consonant music. *Psychological Science*, 22(10):1270–1273, 2011.
- Jean-Marc CHOUVEL : *Esquisses pour une pensée musicale. Les métamorphoses d'Orphée*. L'Harmattan, Paris, 1998.
- Jean-Marc CHOUVEL, Jean BRESSON et Carlos AGON : L'analyse musicale différentielle : principes, représentation et application à l'analyse de l'interprétation. In *EMS conference*, 2007. URL <http://www.ems-network.org/IMG/JMChouve1EMS07/index.html>.
- Richard COHN : Introduction to neo-riemannian theory : a survey and a historical perspective. *Journal of Music Theory*, pages 167–180, 1998.

- Tom COLLINS, Barbara TILLMANN, Frederick S BARRETT, Charles DELBÉ et Petr JANATA : A combined model of sensory and cognitive representations underlying tonal expectations in music : From audio signals to behavior. *Psychological review*, 121(1):33, 2014.
- Ning CONG : *A Multi-Dimensional Analytical Model for Musical Harmony Perception*. Thèse de doctorat, University of York, 2016.
- N COOK : Towards the compleat musicologist ? invited talk. *In Sixth International Conference on Music Information Retrieval*, 2005.
- Norman D. COOK : Harmony Perception : Harmoniousness is More Than the Sum of Interval Consonance. *Music Perception*, 27(1):25–42, 2009.
- Pierre COUPRIE : Quelques propos sur les outils et les méthodes audionumériques en musicologie. l’interdisciplinarité comme rupture épistémologique. *Revue musicale OICRM*, 6(2):25–44, 2020.
- Marion COUSINEAU, Josh H MCDERMOTT et Isabelle PERETZ : The basis of musical consonance as revealed by congenital amusia. *Proceedings of the National Academy of Sciences*, 109(48):19858–19863, 2012.
- Nelson COWAN : The magical number 4 in short-term memory : A reconsideration of mental storage capacity. *Behavioral and brain sciences, Cambridge University Press*, 24(1):87–114, 2001.
- Gregory COX : On the relationship between entropy and meaning in music : An exploration with recurrent neural networks. *In Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 32, 2010.
- Luciano da FONTOURA COSTA : Modeling consonance and its relationships with temperament, harmony, and electronic amplification. *arXiv*, 2019.
- Carl DAHLHAUS : *La tonalité harmonique : étude des origines [Untersuchungen über die Entstehung der harmonischen Tonalität]*. Editions Mardaga, Liège, 1967. Traduction de l’allemand par Anne-Emmanuelle Ceulemans, Ed.Bärenreiter Verlag, 1993.
- Jean le Rond D’ALEMBERT : *Éléments de Musique Théorique et Pratique suivant les principes de M. Rameau*. éd. David, le Breton, Durand, Paris, 1752. Seconde édition de 1779 (Bruyset), enrichie d’un *Discours préliminaire*.

- Alain DANÉLOU : *Sémantique musicale, essai de psychologie auditive*. Hermann, Paris, 1967.
- Gregory DANNER : The use of acoustic measures of dissonance to characterize pitch-class sets. *Music Perception : An Interdisciplinary Journal*, 3(1):103–122, 1985. Publisher : University of California Press Journals.
- Christopher J DARWIN : Pitch and auditory grouping. *In Pitch*, pages 278–305. Springer, 2005.
- Prosdocimo de BELDEMANDI : *Tractacus de Contrapunctu*. Coussemaker, 1412.
- René DESCARTES : *Compendium Musicae*. Paris, 1618.
- Diana DEUTSCH, Kevin DOOLEY et Trevor HENTHORN : Pitch circularity from tones comprising full harmonic series. *The Journal of the Acoustical Society of America*, 124(1):589–597, 2008.
- Lucinda A. DEWITT et Robert G. CROWDER : Tonal fusion of consonant musical intervals : The oomph in Stumpf. *Perception & Psychophysics*, 41(1):73–84, 1987.
- Simon DIXON : Automatic extraction of tempo and beat from expressive performances. *Journal of New Music Research, Taylor & Francis*, 30(1):39–58, 2001.
- Jonathan DRIEDGER, Meinard MÜLLER et Sascha DISCH : Extending harmonic-percussive separation of audio signals. *In ISMIR*, pages 611–616, 2014.
- Joseph Guichard DUVERNEY : *Traité de l’organe de l’ouïe ; contenant la structure, les usages & les maladies de toutes les parties de l’oreille*. Estienne Michallet, Paris, 1683.
- Chris DUXBURY, Juan Pablo BELLO, Mike DAVIES, Mark SANDLER *et al.* : Complex domain onset detection for musical signals. *In Proc. Digital Audio Effects Workshop (DAFx)*, volume 1, pages 6–9. Queen Mary University London, 2003.

- Martin EBELING : Neuronal periodicity detection as a basis for the perception of consonance : A mathematical model of tonal fusion. *The Journal of the Acoustical Society of America*, 124(4):2320–2329, 2008.
- R EBERLEIN : Die entstehung der tonalen klangsyntax [the emergence of the syntax of tonal harmony]. *Frankfurt : Lang*, 1994.
- Tuomas EEROLA et Imre LAHDELMA : The anatomy of consonance/dissonance : Evaluating acoustic and cultural predictors across multiple datasets with chords. *OSF Preprints*, 2020.
- Pierre ESTÈVE : *Nouvelle découverte du principe de l’harmonie avec un examen de ce que M. Rameau a publié sous le titre de Démonstration de ce principe*. Jorry, Paris, 1752.
- Julio ESTRADA : *Théorie de la composition : discontinuum-continuum. (une théorie du potentiel intervallique dans les échelles tempérées et une méthodologie d’inscription graphique et de transcription de paramètres rythmico-sonores. Quelques applications à l’analyse et à la composition musicale*. Thèse de doctorat, Université de Strasbourg II, 1994.
- Julio ESTRADA : La teoría d1, música-win y algunas aplicaciones al análisis musical : Seis piezas para piano, de arnold schoenberg. *In Memoirs of the Fourth International Seminar on Mathematical Music Theory*, volume 4, pages 113–145, 2011.
- Leonhard EULER : *Tentamen novae theoriae musicae ex certissimis harmoniae principiis dilucide expositae*. Academiae scientiarum Petropolitanae, 1739.
- Leonhard EULER : Du véritable caractère de la musique moderne. *Mémoires de l’académie des sciences de Berlin*, 20:174–199, 1766.
- Leonhard EULER : De harmoniae veris principiis per speculum musicum repraesentatis. *Novi commentarii academiae scientiarum Petropolitanae*, pages 330–353, 1774.
- Anton FAIST : Versuche uber tonverschmelzung. *Zsch Psychol Physio Sinnesorg*, 15:102–131, 1897.

- Morwaread M FARBOOD : A parametric, temporal model of musical tension. *Music Perception, University of California Press USA*, 29(4):387–428, 2012.
- Morwaread M FARBOOD et Khen C PRICE : The contribution of timbre attributes to musical tension. *The Journal of the Acoustical Society of America*, 141(1):419–427, 2017.
- Maria Helena Peixoto FERRAZ : *An overview of Almeida Prado’s “Transtonal” system in his Cartas Celestes, first six volumes, for piano solo*. Thèse de doctorat, The University of Arizona, 2009.
- Yonatan I FISHMAN, Igor O VOLKOV, M Daniel NOH, P Charles GARELL, Hans BAKKEN, Joseph C AREZZO, Matthew A HOWARD et Mitchell STEINSCHNEIDER : Consonance and dissonance of musical chords : neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology*, 2001.
- Derry FITZGERALD : Harmonic/percussive separation using median filtering. *In International Conference on Digital Audio Effects (DAFx)*, volume 13, 2010.
- Allan FORTE : *The structure of atonal music*. Yale university press, New Haven, 1973.
- Joseph FOURIER : *Théorie analytique de la chaleur*. Firmin Didot, Père et fils, Paris, 1822.
- Robert FRANCÈS : *La perception de la musique*, volume 14. Vrin, Paris, 1984.
- Ronald S FRIEDMAN, Douglas A KOWALEWSKI, Dominique T VUVAN et W Trammell NEILL : Consonance preferences within an unconventional tuning system. *Music Perception : An Interdisciplinary Journal, University of California Press*, 38(3):313–330, 2021a.
- Ronald S FRIEDMAN, Douglas A KOWALEWSKI, Dominique T VUVAN et W Trammell NEILL : Response to invited commentaries on “consonance preferences within an unconventional tuning system”. *Music Perception : An Interdisciplinary Journal, University of California Press*, 38(3):340–341, 2021b.

- Franchinus GAFURIUS : *De Harmonia*. Milan, 1518.
- Nicolas Alexandrovitch GARBOUZOV : *Музыкальная акустика [Acoustique musicale]*. Musgis, Moscou, 1954.
- Jean de GARLANDE : *De mensurabili musicae*. 1250.
- Manuel GAULHIAC : La synthèse des possibles dans la musique d'Alfred Schnittke : le cas de la sonate pour violoncelle et piano n°1. Mémoire de Master, Sorbonne Université, 2019.
- John M GEARY : Consonance and dissonance of pairs of inharmonic sounds. *The Journal of the Acoustical Society of America*, 67(5):1785–1789, 1980.
- Bruno GINGRAS, Marcus T PEARCE, Meghan GOODCHILD, Roger T DEAN, Geraint WIGGINS et Stephen MCADAMS : Linking melodic expectation to expressive performance timing and perceived musical tension. *Journal of Experimental Psychology : Human Perception and Performance, American Psychological Association*, 42(4):594, 2016.
- Edward GOLLIN : Some aspects of three-dimensional "tonnetze". *Journal of Music Theory*, pages 195–206, 1998.
- Nathalie GOSSELIN, Sébastien PAQUETTE et Isabelle PERETZ : Sensitivity to musical emotions in congenital amusia. *Cortex*, 71:171–182, 2015.
- Roni Y GRANOT et Zohar EITAN : Musical tension and the interaction of dynamic auditory parameters. *Music Perception, University of California Press USA*, 28(3):219–246, 2011.
- Martha GUERNSEY : The role of consonance and dissonance in music. *The American Journal of Psychology*, pages 173–204, 1928.
- Didier GUIGUE : Les compositions pour piano d'Almeida Prado. Quelques éléments techniques. In *"Le piano brésilien", actes de la journée d'étude du 30 janvier 2008*, volume 40, pages 93–101. Université Paris-Sorbonne, Observatoire Musical Français, 2009.
- Christophe GUILLOT-NOETHMANN : Les signes musicaux et leur étude par l'informatique. le statut épistémologique du numérique dans l'appréhension du sens et de la signification en musique. *Revue musicale OICRM*, 6(2):45–72, 2020.

- ER GUTHRIE et H MORRILL : The fusion of non-musical intervals. *The American Journal of Psychology*, 1928.
- Stephen HAINSWORTH, Malcolm D MACLEOD *et al.* : Onset detection in musical audio signals. *In ICMC*, 2003.
- Godfrey H HARDY et Srinivasa RAMANUJAN : Asymptotic formulæ in combinatory analysis. *Proceedings of the London Mathematical Society*, 2 (1):75–115, 1918.
- Peter HARRISON et Marcus T. PEARCE : Simultaneous consonance in music perception and composition. *Psychological Review*, 127(2):216, 2020.
- Peter MC HARRISON : Three questions concerning consonance perception. *Music Perception : An Interdisciplinary Journal*, University of California Press, 38(3):337–339, 2021.
- Christopher HARTE : *Towards automatic extraction of harmony information from music signals*. Thèse de doctorat, Queen Mary University of London, Department of Electronic Engineering, 2010.
- Christopher HARTE, Mark SANDLER et Martin GASSER : Detecting harmonic change in musical audio. *In Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, pages 21–26. ACM, 2006.
- Hermann von HELMHOLTZ : *Théorie physiologique de la musique [Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik]*. J. Gabay, Paris, 1863. Réimpression en 1990 par les éditions Gabay de la traduction originale en français de M.G. Guérout de 1868 aux éditions Masson. Éditeur original allemand : Braunschweig.
- Hermann von HELMHOLTZ : *Optique physiologique*. Masson, Paris, 1867.
- Dorien HERREMANS et Elaine CHEW : Tension ribbons : Quantifying and visualising tonal tension. *In Second International Conference on Technologies for Music Notation and Representation (TENOR)*, 2016.
- Jens HJORTKJÆR : *Toward a cognitive theory of musical tension*. PhD Thesis, University of Copenhagen, Faculty of Humanities, 2011.
- Ludger J HOFMANN-ENGL : Virtual pitch and the classification of chords in minor and major keys. 2008.

- David HUME : *Essai sur la règle du goût [On the Standard of Taste]*. A. Millar in the Strand., London, 1757. In 'Four dissertations'.
- David HURON : Tonal consonance versus tonal fusion in polyphonic sonorities. *Music Perception, University of California Press*, 9(2):135–154, 1991.
- David HURON : Interval-class content in equally tempered pitch-class sets : Common scales exhibit optimum tonal consonance. *Music Perception, University of California Press*, 11(3):289–305, 1994.
- David HURON : Tone and voice : A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19(1):1–64, 2001.
- David Brian HURON : *Sweet anticipation : Music and the psychology of expectation*. MIT press, Cambridge, États-Unis, 2006.
- William HUTCHINSON et Leon KNOPOFF : The significance of the acoustic component of consonance in Western triads. *Journal of Musicological Research*, 3(1-2):5–22, 1979.
- Kristoffer JENSEN et Jens HJORTKJÆR : An Improved dissonance measure based on auditory memory. *Journal of the Audio Engineering Society*, 60(5):350–354, 2012.
- Phil N. JOHNSON-LAIRD, Olivia E. KANG et Yuan Chang LEONG : On musical dissonance. *Music Perception : An Interdisciplinary Journal*, 30(1):19–35, 2012.
- Akio KAMEOKA et Mamoru KURIYAGAWA : Consonance theory part I : Consonance of dyads. *The Journal of the Acoustical Society of America*, 45(6):1451–1459, 1969a.
- Akio KAMEOKA et Mamoru KURIYAGAWA : Consonance theory part II : Consonance of complex tones and its calculation method. *The Journal of the Acoustical Society of America*, 45(6):1460–1469, 1969b.
- Hans KAYSER : *Lehrbuch der Harmonics [Le livre de l'Harmonique]*. Verlag, Berlin, 1950.
- Johannes KEPLER : *Harmonices Mundi*. Linz, 1619.

- Anssi P K LAPURI : Multiple fundamental frequency estimation based on harmonicity and spectral smoothness. *IEEE Transactions on Speech and Audio Processing*, 11(6):804–816, 2003.
- Felix KRUEGER : Differenzttöne und Konsonanz. *Archiv f. d. gesamte Psychologie 1 (2 u. 3)*, pages 205–275, 1903.
- Carol L KRUMHANSL et Edward J KESSLER : Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological review, American Psychological Association*, 89(4):334, 1982.
- Tuire KUUSI : Comparing nontraditional tetrachords and pentachords : Both set-class and chord voicing guide evaluations. *Journal of New Music Research, Taylor & Francis*, 39(3):215–225, 2010.
- Imre LAHDELMA : *At the interface between sensation and emotion : perceived qualities of single chords*. Thèse de doctorat, University of Jyväskylä, Finlande, 2017.
- Imre LAHDELMA, James ARMITAGE et Tuomas EEROLA : Affective priming with musical chords is influenced by pitch numerosity. *Musicae Scientiae, SAGE Publications Sage UK : London, England*, 2020.
- Imre LAHDELMA et Tuomas EEROLA : Mild dissonance preferred over consonance in single chord perception. *i-Perception, Sage Publications Sage UK : London, England*, 7(3), 2016.
- Imre LAHDELMA et Tuomas EEROLA : cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Scientific Reports, Nature Publishing Group*, 10(1):1–11, 2020.
- Philippe LALITTE : The theories of Helmholtz in the work of Varese. *Contemporary Music Review*, 30(5):327–342, 2011.
- Philippe LALITTE : *Analyser l'interprétation de la musique du XXe siècle. Une analyse de l'interprétation des Dix pièces pour quintette à vent de György Ligeti*. Hermann, Strasbourg, gream édition, 2015.
- Charles LALO : *Esquisse d'une esthétique musicale scientifique*. Félix Alcan, Paris, 1908.

- Olivier LARTILLOT, Petri TOIVIAINEN et Tuomas EEROLA : A matlab toolbox for music information retrieval. *In Data analysis, machine learning and applications*, pages 261–268. Springer, 2008.
- Fred LERDAHL et Carol L KRUMHANSL : Modeling tonal tension. *Music perception, University of California Press USA*, 24(4):329–366, 2007.
- Scott LEVINE et Julius SMITH : A sines + transients + noise audio representation for data compression and time/pitch scale modifications. *In Audio Engineering Society Convention 105*. Audio Engineering Society, 1998.
- Theodor LIPPS : *Grundtatsachen des seelenlebens [Faits fondamentaux de la vie psychique]*. Cohen, Bonn, 1880.
- Jacobus de LIÈGE : *Speculum musicae*. 1325.
- Robert W LUNDIN : Toward a cultural theory of consonance. *The Journal of Psychology*, 23(1):45–49, 1947.
- Ernst MACH : *Introduction à la théorie de la musique d’Helmholtz, présentation populaire pour les musiciens [Einleitung in die Helmholtz’sche Musiktheorie, populär für musiker dargestellt]*. Leuchner und Lubensky, Graz, 1866.
- Constantine Frithiof MALMBERG : The perception of consonance and dissonance. *Psychological Monographs, Psychological Review Company*, 25 (2):93, 1918.
- Nikita MAMEDOV : *The Interval Dissonance Rate in Chopin’s Études Op. 10, Nos. 1-4 : Dissecting Arpeggiation, Chromaticism, and Linear Progressions*. Thèse de doctorat, State University of Louisiana, 2018.
- Manuela M MARIN, William Forde THOMPSON, Bruno GINGRAS et Lauren STEWART : Affective evaluation of simultaneous tone combinations in congenital amusia. *Neuropsychologia*, 78:207–220, 2015.
- Vasco MARIZ : *História da Música no Brasil*, volume 4. Ed. Civilização Brasileira, Rio de Janeiro, 1994.
- Nobuo MASATAKA : Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental science*, 9(1):46–50, 2006.

- Keith MASHINTER : Calculating sensory dissonance : Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. *Empirical Musicology Review*, 2006.
- Matthias MAUCH : *Automatic chord transcription from audio using computational models of musical context*. Thèse de doctorat, Queen Mary University of London, School of Electronic Engineering and Computer Science, 2010.
- Matthias MAUCH et Simon DIXON : Approximate note transcription for the improved identification of difficult chords. *In ISMIR*, pages 135–140, 2010.
- Stephen MCADAMS : Spectral fusion and the creation of auditory images. *In Music, mind, and brain*, pages 279–298. Springer, 1982.
- Josh H. McDERMOTT, Andriana J. LEHR et Andrew J. OXENHAM : Individual Differences Reveal the Basis of Consonance. *Current Biology*, 20 (11):1035–1041, juin 2010.
- Josh H McDERMOTT, Alan F SCHULTZ, Eduardo A UNDURRAGA et Ricardo A GODOY : Indifference to dissonance in native amazonians reveals cultural variation in music perception. *Nature, Nature Publishing Group*, 535(7613):547–550, 2016.
- James MCGOWAN : “consonance” in tonal jazz : A critical survey of its semantic history. *Jazz Perspectives*, 2(1):69–102, 2008.
- Cory MCKAY, Ichiro FUJINAGA et Philippe DEPALLE : jaudio : A feature extraction library. *In Proceedings of the International Conference on Music Information Retrieval*, pages 600–3, 2005.
- Martin MCKINNEY et Jeroen BREEBAART : Features for audio and music classification. *Johns Hopkins University*, 2003.
- Neil McLACHLAN, David MARCO, Maria LIGHT et Sarah WILSON : Consonance and pitch. *Journal of Experimental Psychology : General*, 142 (4):1142, 2013.
- Nicolas MEEÛS : Transitivité, rection, fonctions tonales. une approche cognitive de la tonalité. *Analyse musicale*, 26:26–29, 1992.

- Nicolas MEEÛS : Vecteurs harmoniques. *Musurgia*, pages 7–34, 2003.
- Alexius MEINONG et Stephan WITASEK : *Zur experimentellen Bestimmung der Tonverschmelzungsgrade*. Verlag von Johann Ambrosius Barth, 1897.
- Marin MERSENNE : *Harmonie universelle, contenant le théorie et la pratique de la musique*. éditeurs : Sébastien Cramoisy, Pierre I Ballard et Richard Charlemagne, Paris, 1636.
- Max Friedrich MEYER : *Contributions to a psychological theory of music*, volume 1. University of Missouri, 1901.
- George A MILLER : The magical number seven, plus or minus two : Some limits on our capacity for processing information. *Psychological review, American Psychological Association*, 63(2):81, 1956.
- Andrew J. MILNE, Robin LANEY et David B. SHARP : Testing a spectral model of tonal affinity with microtonal melodies and inharmonic spectra. *Musicae Scientiae*, 20(4):465–494, 2016.
- Andrew J MILNE, William A SETHARES, Robin LANEY et David B SHARP : Modelling the similarity of pitch collections with expectation tensors. *Journal of Mathematics and Music, Taylor & Francis*, 5(1):1–20, 2011.
- Ludovico MINATI, Cristina ROSAZZA, Ludovico D’INCERTI, Emanuela PIETROCINI, Laura VALENTINI, Vidmer SCAIOLI, Catherine LOVEDAY et Maria Grazia BRUZZONE : Functional mri/event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *Neuroreport*, 20(1):87–92, 2009.
- M. MIRON, Carabias-Orti J. J., Bosch J., Gómez E. et Janer J. : PHENICX-Anechoic : note annotations for Aalto anechoic orchestral database, novembre 2016. URL <https://doi.org/10.5281/zenodo.1289821>.
- Dalibor MITROVIĆ, Matthias ZEPPELZAUER et Christian BREITENEDER : Features for content-based audio retrieval. In *Advances in computers*, volume 78, pages 71–150. Elsevier, 2010.
- Brian CJ MOORE : Frequency difference limens for short-duration tones. *The Journal of the Acoustical Society of America*, 54(3):610–619, 1973.

- Henry Thomas MOORE : The genetic aspect of consonance and dissonance. *The psychological monographs, psychological review publications*, 1914.
- Tadeu MORALES TAFFARELLO : *O percurso da interseccao Olivier Messiaen – Almeida Prado : Momentos, La Fauvette des Jardins e Cartas Celestes*. Thèse de doctorat, Universidade Estadual de Campinas, Brésil, 2010.
- Meinard MÜLLER : *Fundamentals of music processing : Audio, analysis, algorithms, applications*. Springer, Heidelberg, 2015.
- WA MUNSON : The growth of auditory sensation. *The Journal of the Acoustical Society of America*, 19(4):584–591, 1947.
- Nadja NEVOLOVITSCH et Peter Van de VELDE : Encounters. Aeolus, 2012. Enregistrement. Dans ce disque, version orgue-violon de Frates.
- Frede NIELSEN : Musical tension and related concepts. *The semiotic web*, 86:491–513, 1987.
- Hella OELMANN et Bruno LAENG : The emotional meaning of harmonic intervals. *Cognitive Processing*, 10(2):113–131, 2009.
- François PACHET et Pierre ROY : Analytical features : a knowledge-based approach to audio feature generation. *EURASIP Journal on Audio, Speech, and Music Processing*, 2009(1):153017, 2009.
- Giambattista PARASCANDOLO, Heikki HUTTUNEN et Tuomas VIRTANEN : Recurrent neural networks for polyphonic sound event detection in real life recordings. In *2016 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 6440–6444. IEEE, 2016.
- Richard PARNCUTT : *Harmony : A Psychoacoustical Approach*. Springer Science & Business Media, Berlin, Heidelberg, 1989.
- Richard PARNCUTT : Pitch properties of chords of octave-spaced tones. *Contemporary Music Review, Taylor & Francis*, 9(1-2):35–50, 1993.
- Richard PARNCUTT : Commentary on keith Mashinter's "calculating sensory dissonance : Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopff". *Empirical Musicology Review*, 2006.

- Richard PARNCUTT, Daniel REISINGER, Andreas FUCHS et Fabio KAISER : Consonance and prevalence of sonorities in Western polyphony : Roughness, harmonicity, familiarity, evenness, diatonicity. *Journal of New Music Research*, 48(1):1–20, 2019.
- Roy D PATTERSON : Spiral detection of periodicity and the spiral form of musical scales. *Psychology of Music*, 14(1):44–61, 1986.
- TH PEAR : The experimental examination of some differences between the major and the minor chord. *British Journal of Psychology, Cambridge University Press*, 4(1):56, 1911.
- Geoffroy PEETERS : A large set of audio features for sound description (similarity and classification) in the cuidado project. *CUIDADO IST Project Report*, 54(0):1–25, 2004.
- Geoffroy PEETERS, Bruno L. GIORDANO, Patrick SUSINI, Nicolas MISDARIIS et Stephen MCADAMS : The timbre toolbox : Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, 130(5):2902–2916, 2011.
- John R PIERCE : Attaining consonance in arbitrary scales. *The Journal of the Acoustical Society of America*, 40(1):249–249, 1966.
- Christopher J PLACK et Andrew J OXENHAM : The psychophysics of pitch. *In Pitch*, pages 7–55. Springer, 2005.
- Judy PLANTINGA et Sandra E TREHUB : Revisiting the innate preference for consonance. *Journal of Experimental Psychology : Human Perception and Performance*, 40(1):40, 2014.
- R. PLOMP et W. J. M. LEVELT : Tonal Consonance and Critical Bandwidth. *The Journal of the Acoustical Society of America*, 38(4):548–560, 1965.
- Daniel PRESSNITZER et Stephen MCADAMS : Two phase effects in roughness perception. *The Journal of the Acoustical Society of America*, 105(5):2773–2782, 1999.
- Daniel PRESSNITZER, Stephen MCADAMS, Suzanne WINSBERG et Joshua FINEBERG : Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Perception & psychophysics, Springer*, 62(1):66–80, 2000.

- Claude PTOLÉMÉE : *Les Harmoniques [Die Harmonielehre des Klaudios Ptolemaios]*. Elanders, Göteborg, 1930. Texte en grec édité par Ingemar Düring.
- Gabriel A RADVANSKY : *Human memory*. Routledge, London and New-York, 2015.
- Jean-Philippe RAMEAU : *Traité de l'harmonie reduite à ses principes naturels*. Ballard, Paris, 1722.
- Jean-Philippe RAMEAU : *Nouveau système de musique théorique, où l'on découvre le Principe de toutes les règles néceaires à la pratique, pour servir d'introduction au Traité de l'harmonie*. Ballard, Paris, 1726.
- Jean-Philippe RAMEAU : *Génération harmonique, ou traité de musique théorique et pratique*. Prault, Paris, 1737.
- Jean-Philippe RAMEAU : *Démonstration de Principe de l'harmonie, servant de base à tout l'art musical et théorique et pratique*. Durant et Pissot, Paris, 1750.
- François RASTIER : Sur l'immanentisme en sémantique. *Cahiers de linguistique française*, 15:325–336, 1994.
- François RASTIER : De la signification au sens - pour une sémiotique sans ontologie. *Cahiers de linguistique française*, 1999.
- Bruno H REPP et Jacqueline M THOMPSON : Context sensitivity and invariance in perception of octave-ambiguous tones. *Psychological Research, Springer*, 74(5):437–456, 2010.
- Hugo RIEMANN : *Handbuch der harmonielehre*. Breitkopf & Härtel, 1887.
- Jean-Claude RISSET : Pitch control and pitch paradoxes demonstrated with computer-synthesized sounds. *The Journal of the Acoustical Society of America*, 46(1A):88–88, 1969.
- LA ROBERTS : Consonance judgements of musical chords by musicians and untrained listeners. *Acta Acustica United with Acustica, S. Hirzel Verlag*, 62(2):163–171, 1986.

- Susan E. ROGERS : *The influence of sensory and cognitive consonance/dissonance on musical signal processing*. Thèse de doctorat, McGill University, Montreal, Canada, 2010.
- Mario ROSSI : *Audio*. Presses polytechniques et universitaires romandes, Lausanne, 2007.
- Jean-Jacques ROUSSEAU : *Dictionnaire de musique*. Duchesne, Paris, 1768.
- Frank A RUSSO et William Forde THOMPSON : An interval size illusion : The influence of timbre on the perceived size of melodic intervals. *Perception & psychophysics, Springer*, 67(4):559–568, 2005.
- Yizhak SADAI : *Harmony in its systemic and phenomenological aspects*. Yanetz, Jerusalem, 1980.
- Ferdinand de SAUSSURE : *Cours de linguistique générale*. Payot, Paris, 1916.
- Joseph SAUVEUR : Sur la détermination d’un son fixe. *Mémoires de l’Académie Royale des Sciences*, pages 134–144, 1700.
- Joseph SAUVEUR : *Principes d’acoustique et de musique, ou Système général des intervalles des sons*. Paris, 1701.
- Patrick E SAVAGE, Psyche LOUI, Bronwyn TARR, Adena SCHACHNER, Luke GLOWACKI, Steven MITHEN et W Tecumseh FITCH : Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, pages 1–36, 2020.
- Albrecht SCHNEIDER : “Verschmelzung”, tonal fusion, and consonance : Carl Stumpf revisited. *In Music, Gestalt, and Computing*, Lecture Notes in Computer Science, pages 115–143. Springer, Berlin, Heidelberg, 1996.
- Arnold SCHOENBERG et Leonard STEIN : *Structural functions of harmony*. Numéro 478. WW Norton & Company, Londres, 1969.
- Christian SCHÖRKHUBER et Anssi KLAPURI : Constant-Q transform toolbox for music processing. *In 7th Sound and Music Computing Conference*, pages 3–64, Barcelone, Espagne, 2010.

- Nahoko SEKIMOTO : Jean-adam serre (1704-1788) était-il dualiste ? la justification de l'origine du mode mineur dans ses essais sur les principes de l'harmonie (1753). *Musurgia*, pages 7–34, 2007.
- Xavier SERRA, Michela MAGAS, Emmanouil BENETOS, Magdalena CHUDY, Simon DIXON, Arthur FLEXER, Emilia GÓMEZ GUTIÉRREZ, Fabien GOUYON, Herrera BOYER, Sergi JORDÀ PUIG *et al.* : *Roadmap for music information research*. The MIRES Consortium, 2013.
- Jean-Adam SERRE : *Essais sur les principes de l'harmonie*. Prault fils, 1753.
- William A SETHARES : *Tuning, timbre, spectrum, scale*. Springer, Heidelberg, 1999.
- Inbal SHAPIRA LOTS et Lewi STONE : Perception of musical consonance and dissonance : an outcome of neural synchronization. *Journal of The Royal Society Interface*, 5(29):1429–1434, 2008. Publisher : The Royal Society London.
- Roger N SHEPARD : Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, 36(12):2346–2353, 1964.
- Kai SIEDENBURG : Timbral shepard-illusion reveals ambiguity and context sensitivity of brightness perception. *The Journal of the Acoustical Society of America*, 143(2):EL93–EL98, 2018.
- Eline A SMIT et Andrew J MILNE : The need for composite models of music perception : Consonance in tuning systems (familiar or unfamiliar) cannot be explained by a single predictor. *Music Perception : An Interdisciplinary Journal*, University of California Press, 38(3):335–336, 2021.
- Eline A SMIT, Andrew J MILNE, Roger T DEAN et Gabrielle WEIDEMANN : Making the unfamiliar familiar : The effect of exposure on ratings of unfamiliar musical chords. *Musicae Scientiae*, SAGE Publications UK, 2020.
- Eline Adrienne SMIT, Andrew J MILNE, Roger T DEAN et Gabrielle WEIDEMANN : Perception of affect in unfamiliar musical chords. *Plos one*, 14(6), 2019.
- Robert SMITH : *Harmonics, or the philosophy of musical sounds*. T. and J. Merrill, Cambridge, London, 1759.

- Bob SNYDER et Robert SNYDER : *Music and memory : An introduction*. MIT press, Cambridge, États-Unis, 2000.
- Stanley S STEVENS : On the psychophysical law. *Psychological review, American Psychological Association*, 64(3):153, 1957.
- Stanley Smith STEVENS, John VOLKMANN et Edwin Broomell NEWMAN : A scale for the measurement of the psychological magnitude pitch. *The journal of the acoustical society of america*, 8(3):185–190, 1937.
- Frieder STOLZENBURG : Harmony perception by periodicity detection. *Journal of Mathematics and Music, Taylor & Francis*, 9(3):215–238, 2015.
- Gregory J. STROHMAN : *Psychoacoustic entropy theory and its implications for performance practice*. Thèse de doctorat, Temple University, Philadelphie, 2014.
- Carl STUMPF : *Tonpsychologie*, volume I et II. S. Hirzel, Leipzig, 1883, 1890.
- Tasuku SUGIMOTO, Hiromi KOBAYASHI, Noritomo NOBUYOSHI, Yasushi KIRIYAMA, Hideko TAKESHITA, Tomoyasu NAKAMURA et Kazuhide HASHIYA : Preference for consonant music over dissonant music by an infant chimpanzee. *Primates*, 51(1):7–12, 2010.
- Marie TAHON, Jean-Marc CHOUVEL, Zaher BELGHITH et Pierre MICHEL : Segmentation automatique de catégories musicales : étude exploratoire sur le free jazz. In *Journées d’informatique musicale*, Bayonne, France, 2019.
- Giuseppe TARTINI : *Trattato di Musica seconda la vera Scienza dell’Armonia*. Giovanni Manfrè, Padoue, 1754.
- David TEMPERLEY *et al.* : *Music and probability*. Mit Press, Cambridge, Massachusetts, 2007.
- James TENNEY : *A history of consonance and dissonance*. Excelsior Music Publishing Company, New-York, 1988.
- Ernst TERHARDT : On the perception of periodic sound fluctuations (roughness). *Acta Acustica united with Acustica, S. Hirzel Verlag*, 30(4):201–213, 1974.

- Ernst TERHARDT : The Concept of Musical Consonance : A Link between Music and Psychoacoustics. *Music Perception : An Interdisciplinary Journal*, 1(3):276–295, 1984.
- Ernst TERHARDT, Gerhard STOLL et Manfred SEEWANN : Algorithm for extraction of pitch and pitch salience from complex tonal signals. *The Journal of the Acoustical Society of America*, 71(3):679–688, 1982.
- Sten TERNSTRÖM : Choir acoustics : an overview of scientific research published to date. *International Journal of Research in Choral Singing, American Choir Directors Association*, 1(1):3–12, 2003.
- William Forde THOMPSON et Richard PARNCUTT : Perceptual judgments of triads and dyads : Assessment of a psychoacoustic model. *Music Perception*, 14(3):263–280, 1997.
- Juan M TORO et Paola CRESPO-BOJORQUE : Consonance processing in the absence of relevant experience : Evidence from nonhuman animals. *Comparative Cognition & Behavior Reviews*, 12, 2017.
- Mark Jude TRAMO, Peter A CARIANI, Bertrand DELGUTTE et Louis D BRAIDA : Neurobiological foundations for the theory of harmony in western tonal music. *Annals of the New York Academy of Sciences*, 930(1):92–116, 2001.
- Lluís L TRULLA, Nicola DI STEFANO et Alessandro GIULIANI : Computational approach to musical consonance and dissonance. *Frontiers in psychology*, 9:381, 2018.
- Dmitri TYMOCZKO : *A geometry of music : Harmony and counterpoint in the extended common practice*. Oxford University Press, Oxford, 2010.
- Panteleimon Nestor VASSILAKIS : *Perceptual and physical properties of amplitude fluctuation and their musical significance*. Thèse de doctorat, University of California, Los Angeles, 2001.
- Pantelis N VASSILAKIS : Auditory roughness as a means of musical expression. *Selected Reports in Ethnomusicology*, 12(119-144):122, 2005.
- Flemming WINDEKILDE : Alfred schnittke - requiem, concerto for mixed chorus. Label Scandinavian Classic, Chamber Choir Hymnia, 2002. Enregistrement.

- István WINKLER, Oleg KORZYUKOV, Valentina GUMENYUK, Nelson COWAN, Klaus LINKENKAER-HANSEN, Risto J ILMONIEMI, Kimmo ALHO et Risto NÄÄTÄNEN : Temporary and longer term retention of acoustic information. *Psychophysiology, Wiley Online Library*, 39(4):530–534, 2002.
- Yiming WU, Tristan CARSAULT et Kazuyoshi YOSHII : Automatic chord estimation based on a frame-wise convolutional recurrent neural network with non-aligned annotations. *In 2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5. IEEE, 2019.
- Wihelm WUNDT : *Grundzüge der Physiologischen Psychologie*. Engelmann, Leipzig, 1902.
- Oleg YANCHENKO : Alfred schnittke - concerto for choir, two small pieces for organ. Label Melodia, 2012. Enregistrement.
- Chunghsin YEH, Axel ROEBEL et Xavier RODET : Multiple fundamental frequency estimation and polyphony inference of polyphonic music signals. *IEEE transactions on audio, speech, and language processing*, 18(6):1116–1126, 2009.
- Robert W YOUNG : Inharmonicity of plain wire piano strings. *The Journal of the Acoustical Society of America*, 24(3):267–273, 1952.
- Gioseffo ZARLINO : *Istituzioni harmoniche*. Venice, 1558.
- Eberhard ZWICKER, Georg FLOTTORP et Stanley Smith STEVENS : Critical band width in loudness summation. *The Journal of the Acoustical Society of America*, 29(5):548–557, 1957.

Annexe 1 : Partition de *Fratres* d'Arvo Pärt

fratres (1980)

für violine und klavier

arvo pärt

*1935

1♩ = 63
legato

Viol.

ppp poco a poco crescendo sino al fff

The image displays the violin part of the musical score for 'fratres' by Arvo Pärt. The score is written on eight staves. The first staff includes a tempo marking of ♩ = 63 and the instruction 'legato'. The dynamics range from 'ppp' (pianissimo) to 'fff' (fortissimo), with the instruction 'poco a poco crescendo sino al fff'. The music features a complex, continuous melodic line with many triplets, indicated by the number '3' below the notes. The key signature has one sharp (F#) and the time signature is 7/8. The notation is dense, with many beamed notes and slurs.

This musical score is a single melodic line written on a grand staff (treble and bass clefs). It consists of ten staves of music. The key signature is one sharp (F#), and the time signature is 2/4. The melody is highly rhythmic, primarily consisting of eighth notes and quarter notes, often grouped in triplets. The notation includes various slurs, ties, and dynamic markings. The first staff begins with a treble clef and a key signature of one sharp. The second staff through the fourth staff continue the melody. The fifth staff introduces a change in the bass line, with the treble clef staff playing a more active role. The sixth staff through the eighth staff continue the melody. The ninth staff ends with a double bar line and a repeat sign. The tenth staff concludes the piece with a final double bar line and a repeat sign. The dynamic marking 'ff' (fortissimo) appears at the end of the piece.

2

pizz.
f

arco
ppp

mp

8.

8.

3

This musical score is for a piece that begins with a piano introduction and then transitions into a main melody. The score is written for a single melodic line and a piano accompaniment.

Introduction (Measures 1-4): The piano introduction is marked *pizz.* (pizzicato) and *p* (piano). It features a series of chords in the left hand and a single note in the right hand. The tempo is marked *mp* (moderato). The introduction ends with a measure marked *arco* (arco) and *mp*.

Main Melody (Measures 5-11): The main melody is marked *legato* (legato) and *mp*. It begins with a series of eighth notes in the right hand, followed by a series of chords in the left hand. The melody is marked *mp* and *legato*. The piano accompaniment consists of chords in the left hand and a single note in the right hand. The piano introduction is marked *pizz.* and *p*. The main melody is marked *arco* and *mp*. The piano accompaniment is marked *p* and *mp*. The tempo is marked *mp* and *legato*. The score is written for a single melodic line and a piano accompaniment.

Measures 12-14: The melody continues with a series of eighth notes in the right hand, followed by a series of chords in the left hand. The piano accompaniment consists of chords in the left hand and a single note in the right hand. The piano introduction is marked *pizz.* and *p*. The main melody is marked *arco* and *mp*. The piano accompaniment is marked *p* and *mp*. The tempo is marked *mp* and *legato*. The score is written for a single melodic line and a piano accompaniment.

Measures 15-17: The melody continues with a series of eighth notes in the right hand, followed by a series of chords in the left hand. The piano accompaniment consists of chords in the left hand and a single note in the right hand. The piano introduction is marked *pizz.* and *p*. The main melody is marked *arco* and *mp*. The piano accompaniment is marked *p* and *mp*. The tempo is marked *mp* and *legato*. The score is written for a single melodic line and a piano accompaniment.

Measures 18-20: The melody continues with a series of eighth notes in the right hand, followed by a series of chords in the left hand. The piano accompaniment consists of chords in the left hand and a single note in the right hand. The piano introduction is marked *pizz.* and *p*. The main melody is marked *arco* and *mp*. The piano accompaniment is marked *p* and *mp*. The tempo is marked *mp* and *legato*. The score is written for a single melodic line and a piano accompaniment.



First system of musical notation. It consists of three staves. The top staff is in treble clef with a key signature of one flat and a 4/4 time signature. It contains a continuous eighth-note melody. The middle staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords. The bottom staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords, with a double bar line at the end.



Second system of musical notation. It consists of three staves. The top staff is in treble clef with a key signature of one flat and a 4/4 time signature. It contains a continuous eighth-note melody. The middle staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords. The bottom staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords, with a double bar line at the end.



Third system of musical notation. It consists of three staves. The top staff is in treble clef with a key signature of one flat and a 4/4 time signature. It contains a series of chords, with a double bar line at the end. The middle staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords. The bottom staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords, with a double bar line at the end. The system is marked with a box containing the number 4. Dynamic markings include *pizz.*, *p*, *mf*, *arco*, *f non legato*, *p*, *mp*, and *mf*. There are also markings for *8* and *8* with dashed lines.



Fourth system of musical notation. It consists of three staves. The top staff is in treble clef with a key signature of one flat and a 4/4 time signature. It contains a continuous eighth-note melody. The middle staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords. The bottom staff is in treble clef with a key signature of one flat and a 9/4 time signature. It contains a series of chords, with a double bar line at the end.



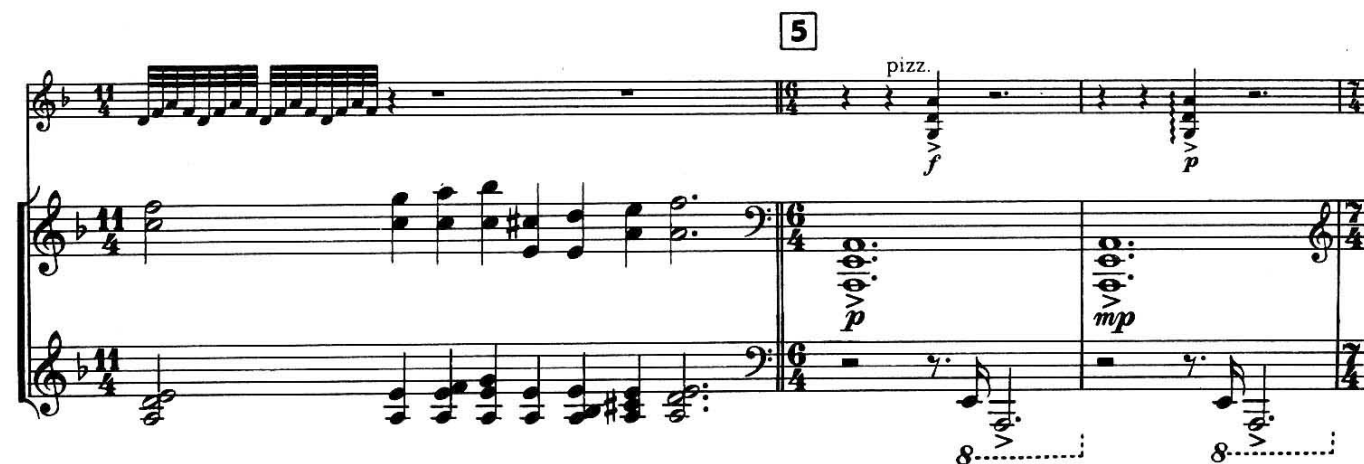
First system of musical notation. It consists of three staves. The top staff is a single melodic line with a complex, fast-moving pattern. The middle and bottom staves are piano accompaniment, featuring chords and single notes in a more rhythmic, steady pattern.



Second system of musical notation. Similar to the first, it has three staves. The top staff continues the complex melodic line. The piano accompaniment in the middle and bottom staves shows some changes in chord structure and rhythm.



Third system of musical notation. The top staff continues the melodic line. The piano accompaniment in the middle and bottom staves maintains a consistent rhythmic pattern with some harmonic shifts.



Fourth system of musical notation, starting with a box containing the number 5. The top staff has a rest followed by a section marked "pizz." (pizzicato) with a forte (f) dynamic. The middle and bottom staves show a change in the piano accompaniment, with some notes marked with a piano (p) dynamic and a mezzo-piano (mp) dynamic. The system concludes with a double bar line and a key signature change to one sharp.

rubato
arco
p (legato)

mp

mp

mp

mp

6

pizz.

p

p

p

arco

ff sub.

ff sub.

The musical score is written for a string quartet, featuring four staves. The first system shows a violin part with a pizzicato (pizz.) instruction and a piano (p) dynamic, and a cello/bass part with a piano (p) dynamic. The second system shows a violin part with an arco instruction and a fortissimo (ff) dynamic, and a cello/bass part with a fortissimo (ff) dynamic. The third system shows a violin part with an arco instruction and a fortissimo (ff) dynamic, and a cello/bass part with a fortissimo (ff) dynamic. The fourth system shows a violin part with an arco instruction and a fortissimo (ff) dynamic, and a cello/bass part with a fortissimo (ff) dynamic. The score includes various musical notations such as notes, rests, and dynamic markings.

7

Measures 7 and 8 of the musical score. The top staff is marked *pizz.* (pizzicato) and features a forte (*f*) chord in measure 7 and a mezzo-forte (*mf*) chord in measure 8. The middle staff shows a forte (*f*) chord in measure 7 and a mezzo-forte (*mf*) chord in measure 8. The bottom staff has a whole rest in measure 7 and a half note in measure 8, with an 8-measure rest indicated by a dotted line.

Measures 9, 10, and 11 of the musical score. The top staff is marked *arco* (arco) and features a mezzo-forte (*mf*) chord in measure 9. The middle staff shows a mezzo-forte (*mf*) chord in measure 9. The bottom staff has a whole rest in measure 9 and a half note in measure 10, with an 8-measure rest indicated by a dotted line.

Measures 12, 13, and 14 of the musical score. The top staff features a mezzo-forte (*mf*) chord in measure 12. The middle staff shows a mezzo-forte (*mf*) chord in measure 12. The bottom staff has a whole rest in measure 12 and a half note in measure 13, with an 8-measure rest indicated by a dotted line.

Measures 15, 16, and 17 of the musical score. The top staff features a mezzo-forte (*mf*) chord in measure 15. The middle staff shows a mezzo-forte (*mf*) chord in measure 15. The bottom staff has a whole rest in measure 15 and a half note in measure 16, with an 8-measure rest indicated by a dotted line.

8

This musical score is for the 'The Swan' movement from the Suite for Piano and Violin by Camille Saint-Saëns. The score is written for a piano and a violin. The piano part is in 6/8 time, and the violin part is in 7/4 time. The key signature is one flat (B-flat). The score is divided into two systems. The first system contains measures 1 through 8. The second system contains measures 9 through 14. The piano part features a variety of dynamics, including *pizz* (pizzicato), *f* (forte), *mf* (mezzo-forte), *mp* (mezzo-piano), and *p* (piano). The violin part includes *arco* (arco) markings and features complex rhythmic patterns, including triplets and sixteenth notes. The score is written in a standard musical notation style, with a treble clef for the violin and a bass clef for the piano. The piano part is written in a 6/8 time signature, while the violin part is written in a 7/4 time signature. The key signature is one flat (B-flat). The score is divided into two systems. The first system contains measures 1 through 8. The second system contains measures 9 through 14. The piano part features a variety of dynamics, including *pizz* (pizzicato), *f* (forte), *mf* (mezzo-forte), *mp* (mezzo-piano), and *p* (piano). The violin part includes *arco* (arco) markings and features complex rhythmic patterns, including triplets and sixteenth notes. The score is written in a standard musical notation style, with a treble clef for the violin and a bass clef for the piano.

9

(pizz.) arco

pizz. mp

pp

pp

pp

col legno

pp

ppp

8.

8.

Annexe 2 : Partition de la *Pièce*
pour orgue n°2 d'Alfred
Schnittke

Две пьесы

А. Шнитке

Sw: 4', 2', Quinte 1 $\frac{1}{3}$ ', Scharf
Pos: Flöte 8'
Hw: Pleno
Ped: 4', 8', Sw / Ped

I

$\text{♩} = 56$

The first system of the musical score consists of three staves. The top staff is a treble clef with a wavy line above it, indicating a sustained or tremolo effect. The middle staff is a bass clef with a wavy line below it, also indicating a sustained or tremolo effect. The bottom staff is a bass clef. The music is in 4/4 time, with a tempo of 56 beats per minute. The key signature has one flat (B-flat). The first staff has a dynamic marking of *f* (forte) and a wavy line above it. The second staff has a dynamic marking of *p* (piano) and a wavy line below it. The third staff has a dynamic marking of *f* (forte) and a wavy line above it. The system ends with a double bar line.

The second system of the musical score consists of three staves. The top staff is a treble clef with a wavy line above it, indicating a sustained or tremolo effect. The middle staff is a bass clef with a wavy line below it, also indicating a sustained or tremolo effect. The bottom staff is a bass clef. The music is in 4/4 time, with a tempo of 56 beats per minute. The key signature has one flat (B-flat). The first staff has a dynamic marking of *f* (forte) and a wavy line above it. The second staff has a dynamic marking of *p* (piano) and a wavy line below it. The third staff has a dynamic marking of *f* (forte) and a wavy line above it. The system ends with a double bar line.

Sw: +8' =

p

Sw

f Hw

Tutti

Hw *fff*

Hw: + Trompete 8'

Sw

Pos

7:3

5:2

5:2

4:3

Ped: — Sw / Ped' + Pos / Ped

Pos — Ped Pos: Flöte Pos Sw: Spitzgambe 8' solo

p *pp*

(dim.)

II

Sw: Gambe 8'
Hw: 16', 8', 4', Sesqualter.
Trompette 8'
Ped: 16', 8', Fagott 16'.
Trompette 8', Hw / Ped

f

This musical score is for a section titled "Tutti". It features a large, complex musical structure with multiple staves. The score is written in a grand staff format, with a treble clef on the left and a bass clef on the right. The music is characterized by a large, sweeping melodic line that spans across the staves, with a dynamic marking of *ff* (fortissimo) indicating a very loud section. The score includes various musical notations, including notes, rests, and dynamic markings, and is presented in a clear, professional layout.

Musical score for "The Rose Tree" in G major, 2/4 time. The score is written for voice and piano. The piano part features a complex texture with multiple staves, including a grand staff (treble and bass clef) and a lower staff. The score includes various musical notations such as notes, rests, and dynamic markings like *pp* (pianissimo) and *p* (piano). A section of the score is marked "Zarter 16' 8' solo" and "ped" (pedal). The score concludes with a double bar line and a repeat sign.

Hw: + Flöte 2'
 + Sesquialter, Trompete 8'

The score consists of two systems of staves. The first system has two staves: the top staff is for Flöte 2' and the bottom staff is for Sesquialter, Trompete 8'. The second system has two staves: the top staff is for Flöte 2' and the bottom staff is for Sesquialter, Trompete 8'. The Flöte 2' part features a melodic line with various ornaments and a final flourish. The Sesquialter, Trompete 8' part features a series of chords and a final flourish. The score is marked with dynamics such as *mf* and *f*, and includes a pedal marking (*ped*) with a double bar line.

Musical score for organ, featuring multiple systems of staves (treble and bass clefs) and various musical notations. The score includes performance instructions and dynamic markings.

System 1:

- Top staff: Treble clef, key signature of two flats (B-flat, E-flat), time signature of 4/4. Includes a melodic line with a slur and a fermata.
- Middle staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Bottom staff: Bass clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Annotations: *Sw pp* (Swell, pianissimo) above the top staff.

System 2:

- Top staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Middle staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Bottom staff: Bass clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Annotations: *Ped: +4', 2', Mixtur, Shalmei 4'* above the top staff. *Ped: 16', 8'* above the bottom staff. *5* (finger number) above the middle staff.

System 3:

- Top staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Middle staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Bottom staff: Bass clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Annotations: *Sw: + Nachthorn 4'* above the top staff. *f* (forte) above the middle staff.

System 4:

- Top staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Middle staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Bottom staff: Bass clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Annotations: *mf* (mezzo-forte) above the top staff. *Ped: +4'* above the bottom staff.

System 5:

- Top staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Middle staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Bottom staff: Bass clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Annotations: *Ped: Zarter 16', 8' solo* above the top staff. *Hw: 8' solo* above the middle staff. *Sw: Gedackt 8' solo* above the bottom staff.

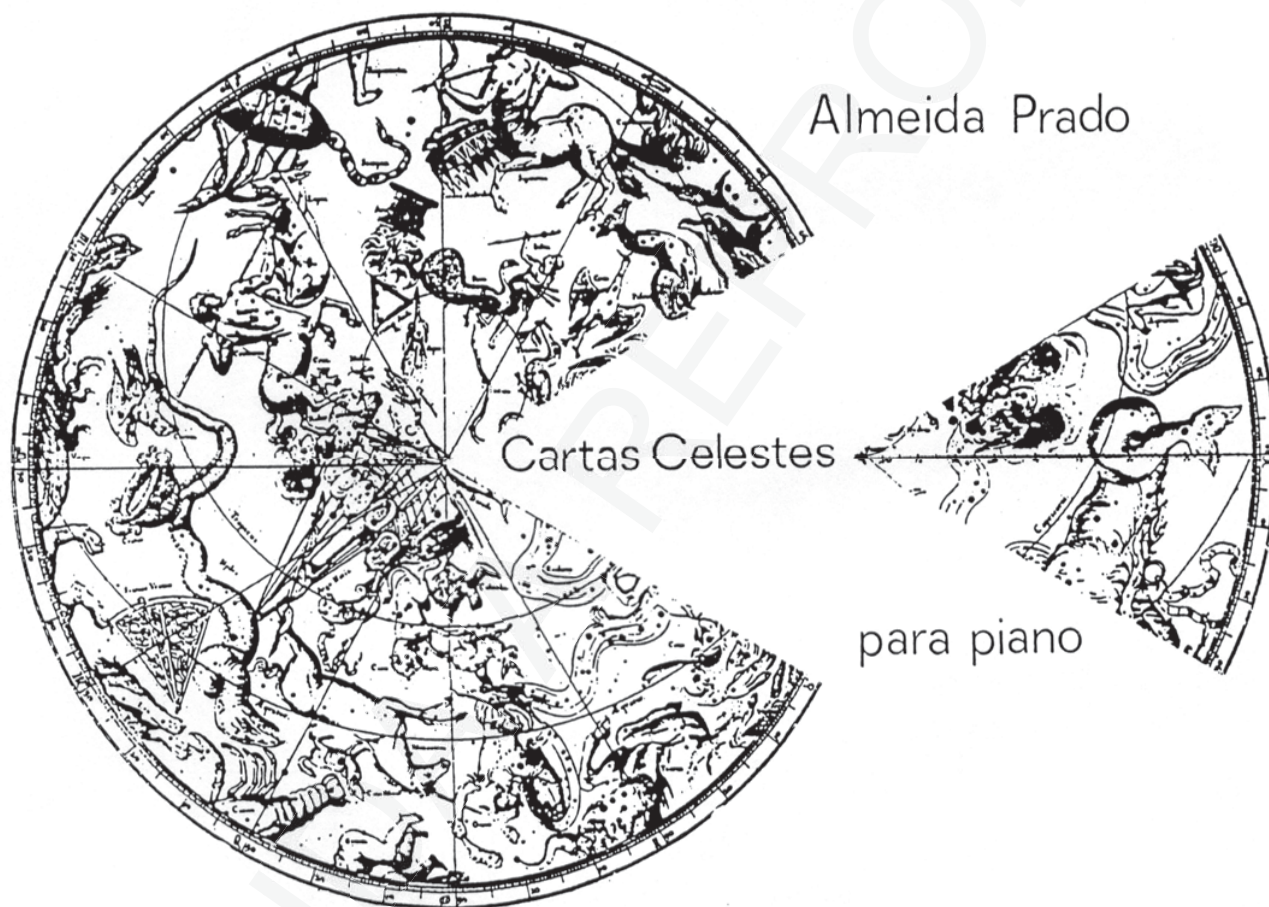
System 6:

- Top staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Middle staff: Treble clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Bottom staff: Bass clef, key signature of two flats. Includes a melodic line with a slur and a fermata.
- Annotations: *Sw tremolo* above the top staff.

Annexe 3 : Partition des *Cartas Celestes I* d'Almeida Prado

Almeida Prado

Cartas Celestes
para piano



TONOS international DARMSTADT

PREFACIO

Quando José Luis Paes Nunes me fez a encomenda para o Planetário Municipal, me veio logo a mente a ideia para uma baseada nas cartas celestes, por isso o próprio título da obra ficou sendo *Cartas Celestes*.

Galáxias, a via láctea, constelações, nebulosas, meteoros, tudo reunido formando um mural sonoro, as vinte e quatro letras do alfabeto grego, dando origem a vinte e quatro acordes diferentes, os quais utilizo para representar sonoramente cada estrela nas constelações.

O livro *"Atlas Celeste"* de Ronaldo Rogério de Freitas Mourão muito me ajudou na difícil tarefa de percorrer o céu. ...

Preferi compor para piano, por ser este o instrumento mais adequado para as passagens vertiginosamente rápidas, as passagens percussivas e a imensa possibilidade de uma grandiosa ressonância.

Ousar representar o Universo infinito na música! Uma pretenciosa ambição! Mas a música oferece Também um Universo infinito e mágico, logo, a imaginação ousa o que a razão titubeia ...

Almeida Prado

São Paulo, 4 de Agosto de 1974

PREFACE

When Jose Luis Paes Nunes gave me the order to write a composition for the presentations of the Planetarium of Sao Paulo, I had immediately the imagination of the musical structure and also of the title: *"Cartas Celestes"* - "Map of Firmament".

Galaxies, Constellation of Stars, Nebulas, Shower of Meteors: for all this I created a sound pattern. I coordinated the 24 letters of the Greek alphabet, for each star of the constellations another chord.

The *"Atlas Celeste"* - "Atlas of Firmament" - of Ronaldo Rogério de Freitas Mourão was much help for me to get a feeling for the width of the firmament.

I intentionally choose the piano as medium for this composition: it has a large spectrum of overtones, is capable of quick figurations, its percussive possibilities and its enormous resonance met my intentions.

Eternity reproduced by music: a proud presumption! But does n't offer us music also a magical and eternal universe? Therefore, imagination may dare, what reason hesitates to do ...

Almeida Prado

VORWORT

Als José Luis Paes Nunes mir den Auftrag erteilte, eine Komposition für die Vorführungen des Planetariums von Sao Paulo zu schreiben, hatte ich sogleich eine musikalische Idee für das Werk - auch für dessen Titel: *"Cartas Celestes"* (*"Himmels-Karten"*).

Galaxien, Milchstraße, Sternbilder, Nebelfelder, Meteore: All dies baute ich in eine Klangkulisse. Jedem der 24 Buchstaben des griechischen Alphabets habe ich einen anderen Akkord zugeordnet. Mit ihnen stellte ich gewissermaßen ein klangliches Abbild der Sterne in den Sternbildern dar.

"Atlas Celeste" (*"Himmels-Atlas"*), ein Buch von Ronaldo Rogério de Freitas Mourão, half mir sehr, etwas von der Weite des Himmels zu erahnen.

Bewußt habe ich als Medium für diese Komposition das Klavier gewählt: Es hat ein großes Obertöne-Spektrum, ist geeignet für schnelle Figuration, auch kann seine Schlagwerk-Möglichkeit und seine großartige Resonanz meiner Vorstellung entgegen.

Unendliches in der Musik wiederzugeben? Eine stolze Anmaßung! Doch kann uns die Musik nicht auch ein magisches und unendliches Universum anbieten? Möge die Vorstellung wagen, was die Vernunft zu tun zögert ...

Almeida Prado

PRÉFACE

Quand Jose Luis Paes Nunes m'invita à composer pour le Planétaire Municipal, tout de suite, m'ai venu à l'idée de m'inspirer dans les lettres célestes: pour cela le propre titre de l'ouvrage est resté *"Lettres Célestes"*.

Galaxies, Voie lactée, Constellations, Nébuluses, Méteores, tout réuni, formant un sonore mural, les vingt quatre lettres de l'alphabet grec, donnant origine à vingt quatre accords différents, lesquels j'emploie pour représenter, sonoreusement, chaque étoile dans les Constellations.

Le bel *"Atlas Céleste"* de Ronaldo Rogério de Freitas Mourão m'aïda beaucoup dans la difficile tâche de parcourir le Ciel.

J'ai préféré composer pour le piano parce que, celui-ci, est l'instrument le plus adequat pour les passages vertigineusement rapides, les passages percussifs et l'immense possibilité d'une grandieuse resonance.

Oser représenter l'Univers infini dans la musique! Prétençieuse ambition! Mais la Musique offre aussi un Univers infini et magique, par conséquence l'imagination essaye ce que la raison hésite.

Almeida Prado

dedicada "in Memoriam"
de Hugh Robertson

Disse Também Deus:
sejam feitos luzeiros
no firmamento do céu,
e separem o dia da noite,
e sirvam para sinais,
e para distinguir
os tempos, os dias e os anos;
e resplandeçam
no firmamento do céu,
e alumiem a terra.
E assim se fez.

Gênesis

Und Gott sprach:
Es werden Lichter
an der Feste des Himmels
die da scheiden Tag und Nacht
und werden geben Zeichen
Zeiten, Tage, Jahre.
Es seien Lichter
an der Feste des Himmels,
daß sie scheinen mögen
auf die Erde.
Und also geschah es.

And God said:
Let there be lights
on the firmament
to divide day and night
to give signs, times,
days and years.
Let there be lights
on the firmament
so that they
light the earth.
And it was so.

Dieu a aussi dit:
soient faits Astres
dans le Firmament du Ciel
et séparent le jour de la nuit
et ils servent comme signaux
pour distinguer
les temps, les jours et les annés
et brillent
dans le Firmament du Ciel
et illuminent la terre.
Et ainsi a été fait.

CARTAS CELESTES

Esta obra foi uma encomenda do Movimento Mário de Andrade, da Prefeitura de São Paulo, especialmente para o Planetário Municipal de São Paulo - 1974.

A obra segue a seguinte ordem:

I.	Pórtico do Crepúsculo	6
	Noite (Vesper — Venus)	7
II.	Via-láctea	8
	Galáxia NGC 224 = M 31 (Nebulosa de Andromeda)	12
	Meteoros	14
	CONSTELAÇÃO I (Hercules)	14
	Aglomerado globular Messier 13	15
	Meteoros	16
	Aglomerado globular Messier 13	16
	CONSTELAÇÃO II (Lyra)	16
	Nebulosa NGC 696095	17
	CONSTELAÇÃO III (Scorpio)	18
	Aglomerado globular Messier 13	19
	Nebulosa NGC 696095	20
	Meteoros	20
	Alpha Piscium	20
	Meteoros	20
	Via-láctea	21
	Venus	21
III.	Pórtico da Aurora	22
	Manhã	24

α		β		γ		δ	
ε		ζ		η		θ	
ι		κ		λ		μ	
ν		ξ		ο		π	
ρ		σ		τ		υ	
φ		χ		ψ		ω	

CARTAS CELESTES

para piano

PÓRTICO DO CREPÚSCULO

ALMEIDA PRADO

O mais rápido possível

1

fff ca. 17''

2

ff ca. 13''

*Ped. até o fim deste movimento**(Ped. bis Schluß des Bewegungsteiles liegen lassen)*

3

f ca. 9''

4

ff ca. 8''

5

mf ca. 7''

6

ff ca. 6''

7

p ca. 5''

8

ff ca. 5''

9

ff ca. 5''

10

pp ca. 5''

11

pp ca. 5''

12

ca. 5''

13

ca. 4''

14

ca. 4''

15

ppp ca. 4''

16

ca. 4''

17

ca. 4''

18

pppp ca. 3''

*

NOITE

VESPER (Venus)

ppp

8

Ped. até o sinal * (Ped. liegen lassen bis Zeichen *)

Ped.

VIA-LACTEA

$\text{♩} = 92$

ppp

ppp

Ped. até o fim (Pedal liegen lassen)

Ped.

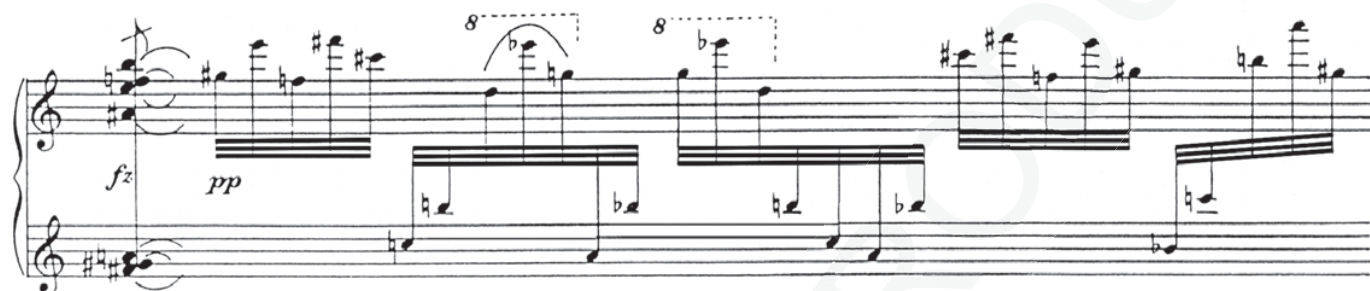
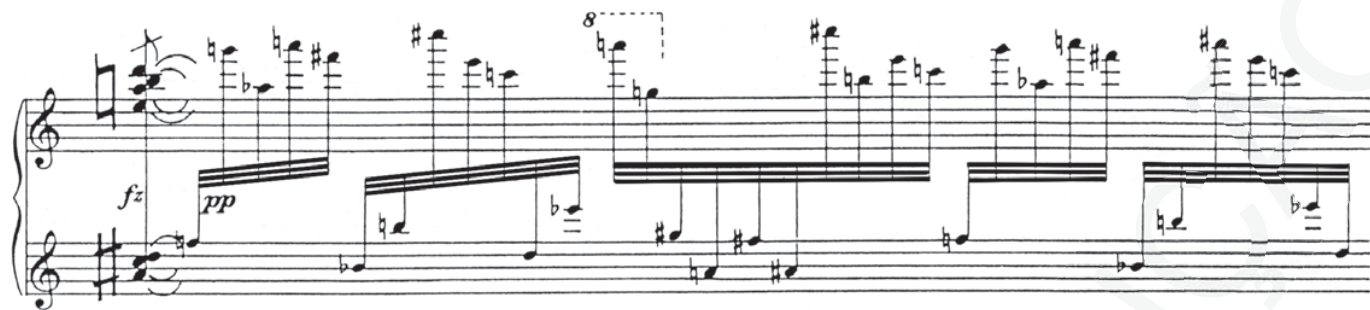
Musical score for piano, featuring multiple systems of staves with complex notation, including dynamics (*pp*, *p*, *cresc.*, *ff*), articulation (accents, slurs), and fingerings (7, 8, 16). The score includes a tempo marking of 60 beats per minute and a final section with a 6-measure repeat.

The musical score consists of several systems of piano music. The first system shows a treble clef with a complex chromatic passage marked with a 5-measure bracket and a 7-measure bracket. The bass clef has a trill. The second system continues the chromatic passage in the treble, marked with a 7-measure bracket and a *cresc.* marking. The third system shows the chromatic passage continuing, with a *cresc.* marking and a *até ff* marking. The fourth system shows a bass clef with a chromatic passage marked *sub. ppp* and a *repetir varias vezes* instruction. The fifth system shows a bass clef with a chromatic passage marked with a 5-measure bracket and a *repetir varias vezes crescendo violentamente até fff* instruction. The sixth system shows a treble clef with a chromatic passage marked *ppp*. The seventh system shows a treble clef with a chromatic passage marked *cresc.*

Proibida a reprodução total ou parcial © Academia Brasileira de Música

desaparecendo - - *

+ $\text{♩} = 120, 144$



(METEOROS)

First system of the musical score for 'METEOROS'. It features a piano introduction with a glissando in the right hand and a steady eighth-note bass line in the left hand. The tempo is marked 'And.' (Andante).

Second system of the musical score for 'METEOROS'. It continues the piano introduction with glissandi and sustained chords in both hands.

Third system of the musical score for 'METEOROS'. The piano introduction continues with complex glissandi and harmonic textures.

(Hércules) CONSTELAÇÃO I

♩ = 80

First system of the musical score for '(Hércules) CONSTELAÇÃO I'. It begins with a forte (ff) piano introduction, followed by a section marked 'p' (piano). The tempo is marked 'And.' (Andante). The system includes fingering numbers 5, 6, and 7.

Second system of the musical score for '(Hércules) CONSTELAÇÃO I'. It continues the piano introduction with various dynamics including ff, p, and pp, and includes a 'rall.' (rallentando) marking. The system includes fingering numbers 5 and 6.

AGLOMERADO GLOBULAR Messier 13

(METEOROS)

First system of the musical score for (METEOROS). It consists of two staves. The upper staff features a series of ascending sixteenth-note runs, while the lower staff has a more rhythmic accompaniment with eighth and sixteenth notes. A dynamic marking of *f* (forte) is present at the beginning.

Second system of the musical score for (METEOROS). It continues the melodic and harmonic development from the first system, with similar ascending runs in the upper staff and accompaniment in the lower staff.

Third system of the musical score for (METEOROS). The upper staff shows a continuation of the melodic lines, and the lower staff includes some rests and sustained notes. Dynamics include *ff* (fortissimo) and *pp* (pianissimo).

AGLOMERADO GLOBULAR MESSIER 13

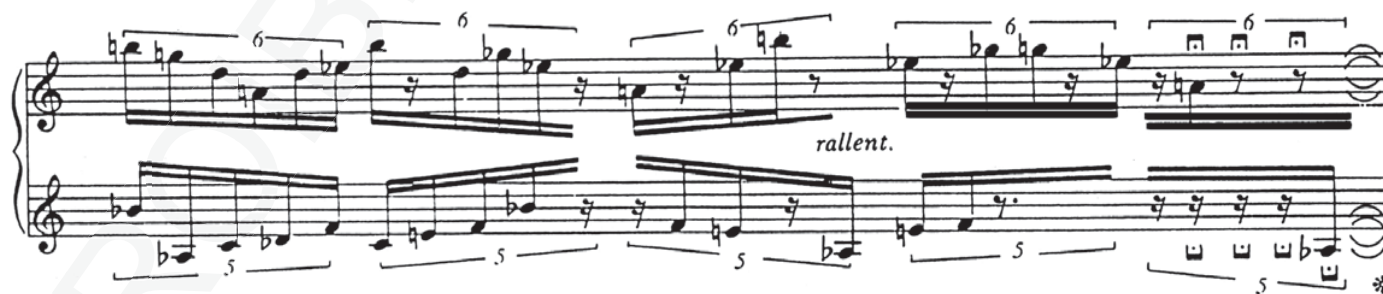
Fourth system of the musical score, titled "Lyra CONSTELAÇÃO II". It begins with a boxed Greek letter alpha (α) and a dynamic marking of *fff* (fortississimo). The upper staff has a melodic line with some grace notes, and the lower staff provides a harmonic base. A specific note is labeled "Vega". Dynamics include *ppp* (pianississimo) and *p* (piano).

Fifth system of the musical score for Lyra CONSTELAÇÃO II. It starts with a boxed Greek letter gamma (γ) and a dynamic marking of *ppp*. The upper staff continues the melodic exploration, and the lower staff has a more active accompaniment. Dynamics include *pp* (pianissimo).



NEBULOSA NGC 696095

♩ = 50



Scorpio CONSTELAÇÃO III

Musical score for "Scorpio CONSTELAÇÃO III". The score is written for piano (piano) and includes various musical notations such as treble and bass staves, chords, and melodic lines. The tempo is marked as 110, and the time signature is 16. The score is divided into measures, with some measures marked with "8" and "16". The score includes dynamic markings such as "simili", "luminoso, fulgurante", and "x ped.". The score is written in a key signature of one sharp (F#) and one flat (Bb).

This page of musical notation is for a piano piece, likely a transcription of a work by Messiaen. It features complex chordal textures and dynamic markings. The notation is arranged in four systems, each with two staves (treble and bass clef). The first system includes a 'ped.' marking. The second system includes a '16' marking and a 'ped.' marking. The third system includes a '16' marking, a 'ff' (fortissimo) marking, and a 'pp' (pianissimo) marking. The fourth system includes a 'pp' marking, a 'ped.' marking, and an 'accel.' (accelerando) marking. The music is written in a key with one sharp (F#) and a 4/4 time signature. The notation is dense, with many chords and rapid changes in dynamics.

♩ = 50 NEBULOSA NGC 696096

♩ = 120

ppp

repetir 3 a 4 vezes

ppp

(METEOROS)

f

tr

f

tr

f

tr

f

tr

First system of the musical score. It features a treble and bass staff. The treble staff has a melodic line with sixteenth and eighth notes, marked with '16' and '8' above it. The bass staff has a more complex texture with sixteenth notes and rests, marked with '8', '16', and 'ff'. Dynamics include 'pp' and 'ff'. A fermata is present over a measure in the bass staff. A small asterisk and 'Ad.' are at the bottom.

Second system of the musical score. The treble staff continues the melodic line with '16' and '8' markings. The bass staff has a more active line with 'ff' and 'pp' dynamics. A fermata is present over a measure in the bass staff.

Third system of the musical score. The treble staff has a melodic line with 'ppp' dynamics. The bass staff has a more active line with 'ppp' dynamics. A fermata is present over a measure in the bass staff.

Fourth system of the musical score. The treble staff has a melodic line with '8' markings. The bass staff has a more active line with 'ppp' dynamics. A fermata is present over a measure in the bass staff. The word 'Venus' is written above the treble staff.

Fifth system of the musical score. The treble staff has a melodic line with '8' markings. The bass staff has a more active line with '8' markings. A fermata is present over a measure in the bass staff.

PORTICO DA AURORA

1



2



3



Red. —————>

4



5



6



7



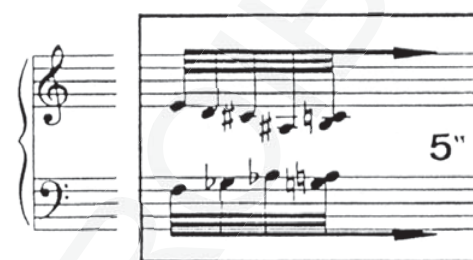
8



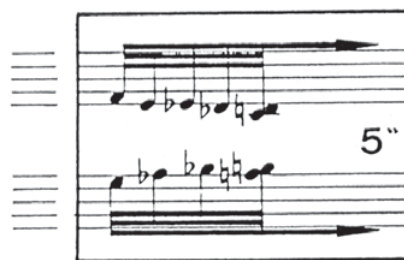
9



10



11



12



13

6"

14

7"

15

8"

16

9"

17

13"

18

17"

19

11"

MANHÃ

O mais rápido possível

Solar!

varias vezes 9"

16

8

varias vezes 7"

16

8

15

8

fff

Cubatão, 2 de Agosto de 1974

Les descripteurs harmoniques : étude théorique et applications musicologiques

Résumé

Cette thèse propose des modèles acoustiques adaptés à l'usage musicologique capables de rendre compte du phénomène harmonique, décrit comme l'ensemble des phénomènes sonores et des relations formelles impliquant l'interaction simultanée ou successive de plusieurs notes. Nous introduisons dans ce travail la notion de descripteur harmonique, comme type particulier de descripteur audio décrivant ce type d'interaction. Les modèles développés s'attachent aux aspects à la fois statique (accords isolés) et dynamique (enchaînements d'accords) de l'harmonie. Les descripteurs proposés — rugosité, harmonicité, concordance, concordance totale, tension, changement harmonique, concordance différentielle et rugosité différentielle — sont définis à travers une implémentation hybride, à l'interface entre les approches symbolique et signal.

Les descripteurs harmoniques fournissent des représentations que nous mettons à profit pour étudier les ensembles d'accords et de transitions. Nous proposons des cartographies des espaces harmoniques continus mettant en avant des structures liées à différentes propriétés spectrales. Le recours à des spectres de Shepard permet de généraliser les modèles aux classes d'accords. Nous appliquons cette généralisation à deux classifications : le potentiel combinatoire d'Estrada et le Tonnetz, en associant une «~couleur harmonique~» aux identités et aux transitions entre classes d'accords.

Les descripteurs harmoniques offrent des outils d'analyse à même de mettre en évidence des logiques acoustiques à l'œuvre dans la musique. Nous en tirons parti pour analyser un corpus de trois œuvres de la même époque qui font appel à des langages radicalement différents : *Fratres* d'Arvo Pärt (1977), la *Pièce pour orgue n°2* d'Alfred Schnittke (1980) et les *Cartas Celestes I* d'Almeida Prado (1974).

Mots-clés : Harmonie ; descripteurs harmoniques ; analyse musicale ; modèles de perception ; consonance et dissonance ; espaces d'accords et de transitions entre accords ; Tonnetz ; potentiel combinatoire des intervalles.

Harmonic descriptors: theoretical study and musicological applications

Summary

This thesis proposes acoustic models designed for musicological use that describe harmonic phenomena, defined as the set of sound phenomena or the formal relationships between several simultaneous or successive notes. We introduce the notion of harmonic descriptors, which characterize the interactions between notes, and model both static (isolated chords) and dynamic (chord sequences) aspects of harmony. Proposed descriptors — roughness, harmonicity, concordance, total concordance, tension, harmonic change, differential concordance and differential roughness — are defined through a hybrid implementation at the intersection of symbolic and signal-based approaches.

Harmonic descriptors provide representations that are then used to study sets of chords and transitions. We present maps of continuous harmonic spaces, highlighting structures linked to different spectral properties. The use of Shepard spectra makes it possible to generalize the models to chord classes. This generalization is applied to two classifications: the combinatorial potential of Estrada and the Tonnetz, by associating a “harmonic color” to identities and transitions between chord classes.

Finally, we demonstrate the analytical potential of harmonic descriptors in highlighting the acoustic logics at work in a corpus of three works from the same period that use radically different languages: *Fratres* by Arvo Pärt (1977), the *Piece for organ n°2* by Alfred Schnittke (1980) and Almeida Prado's *Cartas Celestes I* (1974).

Keywords: Harmony ; harmonic descriptors ; musical analysis ; perception models ; consonance and dissonance ; chord and transition spaces ; Tonnetz ; combinatorial potential of intervals.

UNIVERSITÉ SORBONNE UNIVERSITÉ

ÉCOLE DOCTORALE :

ED V – Concepts et langages

Maison de la Recherche, 28 rue Serpente, 75006 Paris, FRANCE

DISCIPLINE : Musique et musicologie